BOOK OF ABSTRACTS

University of Beira Interior
Covilhã - PORTUGAL
November 9–10, 2018

Committees

Executive Committee

Christos Kitsos, University of West Attica, Athens, Greece
M. Ivette Gomes, Faculdade de Ciências, Universidade de Lisboa, Portugal
Stanislaw Mejza, Poznan University of Life Sciences, Poland
Teresa A. Oliveira, Universidade Aberta and CEAUL, Portugal

Scientific Committee

Ana Caballero, University of Extremadura, Spain
Alex Karagrigoriou, University of the Aegean, Greece
Amílcar Oliveira, Universidade Aberta and CEAUL, Portugal
António Araújo, Universidade Aberta, Portugal
Arminda Manuela Gonçalves, University of Minho, Portugal
Carla Santos, Politechnic Institute of Beja, Portugal
Carlos Agra Coelho, Nova University of Lisbon, Portugal
Carlos Braumann, Universidade de Évora, Portugal
Célia Nunes, University of Beira Interior, Portugal
Christos Kitsos, University of West Attika, Greece
Christos Skiadas, Technological University of Crete, Greece
Cristina Dias, Politechnic Institute of Portalegre, Portugal
Dário Ferreira, University of Beira Interior, Portugal
Dora Prata Gomes, Nova University of Lisbon, Portugal
Elisa Henning, University of Estado de Santa Catarina, UDESC, Brazil
Fernanda Figueiredo, University of Porto, Portugal
Fernando Carapau, University of Évora, Portugal
Filomena Teodoro, Naval School, Portugal
Francisco Carvalho, Instituto Politécnico de Tomar, Portugal
Frederico Caeiro, Nova University of Lisbon, Portugal
Iwona Mejza, Poznan University of Life Sciences, Poland
Izabela Pruchnicka-Grabias, College of Economic and Social, Banking Institute, Warsaw, Poland
James R. Bozeman, American University of Malta, Malta
João Tiago Mexia, Nova University of Lisbon, Portugal
José António Macias-Seijas, University of Corunha, Spain
José António Pereira, University of Porto, Portugal
José Luís Cardoso, University of Trás-os-Montes e Alto Douro, Portugal
Karl-Moder, University of Natural Resources and Life Sciences, Austria
Lidia Filus, Northeastern Illinois University, USA

Luís Grilo, Politechnic Institute of Tomar, Portugal
Manuela Neves, University of Lisbon , Portugal
Manuel L. Esquível, Nova University of Lisbon, Portugal
Maria das Neves Rebocho, Universidade da Beira Interior, Portugal
Maria do Rosário Ramos, Universidade Aberta, Portugal
Marina A. P. Andrade, ISCTE-IUL/ISTAR, Portugal
M. Ivette Gomes, University of Lisbon, Portugal
Nuno Brites, Instituto Superior de Economia e Gestão, Universidade de Lisboa, Portugal
Robert G. Aykroyd, University of Leeds, UK
Sandra Ferreira, University of Beira Interior, Portugal
Sandra Nunes, Politechnic Institute of Setúbal, Portugal
Stanislaw Mejza, Poznan University of Life Sciences, Poland
Teresa A. Oliveira, Universidade Aberta and CEAUL, Portugal

Organizing Committee

Amílcar Oliveira, Open University and CEAUL, Portugal
Célia Nunes, University of Beira Interior, Portugal
Dário Ferreira, University of Beira Interior, Portugal
Luís Grilo, Politechnic Institute of Tomar, Portugal
Sandra Ferreira, University of Beira Interior, Portugal - Local Chair
Teresa A. Oliveira, Open University and CEAUL, Portugal

Sponsored by

Preface

Dear Friends and Colleagues,

WELCOME to the 12th Workshop on Statistics, Mathematics and Computation!

We celebrate this meeting in Honour of Professor Carlos Braumann, for his brilliant career and outstanding contribution to the field of Statistics in Portugal and abroad, and to whom we are deeply grateful for all the support, kind and precious collaboration in our meetings!

We are delighted to have this celebration fostering strong interaction between national and international researchers, leading to a successful commitment and enthusiasm on promoting research in and between the broad areas of Statistics, Mathematics and Computation. It is a great pleasure to receive all our guests and contributors from 9 to 10 th November in UBI-Universidade da Beira Interior, expressing our huge gratitude to the University Coordinators for kindly accepted to embrace this challenge of receiving us again! The WSMC was successfully organized in several places along these last 12 years and we believe that our meetings have been creating very nice opportunities for showcasing the growth and development of the focused main areas, at a time when so many new technologies are available and huge challenges are emerging. The idea exchanges between participants are always helpful for generating positive impact on propelling the advancement of science and technology and some of these results have been published in high standard Journals, special issues and Springer Series Books. Also in this 12th WSMC edition the participants will have several journal opportunities for papers submission. Selected papers, after review, will appear in Journal of Applied Statistics, Biometrical Letters, Biometrics & Biostatistics International Journal (BBIJ) and in a new Springer volume of the Series Contributions to Statistics.

We are hightly grateful to all the participants, Invited Speakers, Session Organizers and Authors who submitted abstracts, for their valuable contribution and for the enthusiastic way how they assume their participation. We also acknowledge all the sponsors and contributors who made this meeting a reality. Furthermore, we acknowledge the Rector of Universidade da Beira Interior, the Rector of Universidade Aberta, the Coordinator of the Centro de Estatística e Aplicações da Universidade de Lisboa and the President of the Committee on Risk Analysis of International Statistical Institute, for

their support. We are most grateful to all the members of the Organizing Committee and of the Scientific Committee for their crucial help and suggestions. We address a deeply thanks to the Local Chair, Sandra Ferreira, and also to Amílcar Oliveira, Célia Nunes, Dário Ferreira and Luís Grilo for their invaluable contribution on organizing the Webpage, the final programme, the Book of Abstracts and so many details in a such incredible short time!

Finally, the venue adds an important attraction to the meeting. The destination city, Covilhã, is a fantastic place close to beautiful and highest mountain in Portugal, Serra da Estrela. This region is very famous not only for awesome landscapes and sports opportunities but also for the famous delights like the wool cheese, ham, licors and also for the traditional artesanat. Covilhã offers a full calendar of traditional festivals and holidays celebrated throughout the year adds to its cultural appeal.

We trust that you will find the conference and your visit to the very beautiful city of Covilhã in November 2018 both valuable and enjoyable, we believe that this meeting will be rewarding to all of us and that it will draw us to the next edition -WSMC13, which is planned for 2019 in Lisbon.

By the Executive Committee,

Teresa A. Oliveira

Covilhã, 9-10th November 2018

Programme

NOVEMBER 9th, 2018

08:30 – 09:00 Registration Desk

09:00 – 09:30 Opening Ceremony [Room 6.3]

09:30 – 10:00

Plenary Session 1   **Room 6.3**

Population growth in random environments
Carlos Braumann

10:00 – 10:30 Coffee Break

10:30 – 11:00

Plenary Session 2   Room 6.3

Confidence Ellipsoids for Additive Models

João T. Mexia

11:00 – 12:30 Parallel Sessions

Contributed Session C2   Room 6.3

Chair: Dário Ferreira

Estimation of foreseeable and unforeseeable risks in motor insurance

Alfredo Reis

Monte Carlo solution of boundary value problem for semilinear isotropic diffusion equation

A.S. Rasulov

Optimal control problem for a non-autonomous SIR model for the Ebola virus

Paulo Rebelo

Contributed Session C1 Room 6.2

Chair: Célia Nunes

Adaptive estimation for light-tailed models
Lígia Henriques-Rodrigues

Efficient approximate sampling from k-DPP for large graphs
Diala Wehbe

Mixed effects ANOVA with random sample sizes
Anacleto Mário

12:30 – 14:00 Lunch

14:00 – 14:30

Plenary Session 3 Room 6.3

Testing interaction in unrepeated designs
Karl Moder

14:30 − 16:00 Parallel Sessions

Organized Session S3   **Room 6.3**

Statistical Challenges and its applications

Organizer: Dora Prata Gomes

Chair: Dora Prata Gomes

Pareto index estimation through a modification of the probability weighted moment method

Ayana Mateus

Introducing nonparametric predictive inference methods for reproducibility of likelihood ratio tests

Filipe Marques

Comparison of classes of generalized Hill estimators

Frederico Caeiro

Challenges in extremal index estimation

Dora Prata Gomes

Contributed Session C3   **Room 6.2**

Chair: Amílcar Oliveira

Analysis of the Skewness and Kurtosis for the Product of Two Normally Distributed Variables

José António Seijas-Macias

Structural Equation Modelling with missing data - application to PISA 12 data

Conceição Leal

Multinomial regression models to estimate the risk of diabetes and hypertension in Pima Indian Women

José António Pereira

Forecast time series in presence of structural breaks: a proposal

M. Rosário Ramos

16:00 – 16:30 Coffee Break and Poster Session

16:30 – 17:00

Plenary Session 4   Room 6.3

Measuring the Shape of Boundary Districts Using the Convexity Ratio

James Boseman

19:30   Conference dinner at the Puralã Hotel

NOVEMBER 10th, 2018

08:30 – 09:00 Registration Desk

09:00 – 09:30

Plenary Session 5   **Room 6.1**

    Development of tools for numerical inversion of the characteristic functions

    and the exact distributions of multivariate test statistics

    Viktor Witkovský

09:30 – 10:00 Coffee break and Poster Session

10:00 – 11:30 Parallel Sessions

Organized Session S1   **Room 6.1**

Statistical Applications in Environmental Science

Organizer: Manuela Neves

Chair: Manuela Neves

Extreme value theory in the analysis of environmental events
M. Ivette Gomes

Application of Statistics of Extremes to the Identification of Biodiversity Patterns
Luís Borda de Água

Statistical Ecology: a few examples
Russell Alpizar Jara

Organized Session S2 Room 6.2

Stochastic Differential Equations: Applications in Finance, Management and Medicine
Organizer: Nuno Brites
Chair: Nuno Brites

The application of clustering algorithms to the handŠs biometrics for personal recognition for personal recognition
Lúcia Sousa

From an Ordinary Differential Equation Model to an Open Population
Manuel Esquível

Markov Chain Model, via Stochastic Differential Equations; models for HIV infection in individuals and populations

Valuation of lookback options and turbo warrants on defaultable stocks
José Carlos Dias

Stochastic differential equations driven by mixed fractional

Brownian motion and na application in finance

João Guerra

Stochastic differential equations in fisherie management with stepwise effort management with stepwise effort

Nuno Brites

Organized Session S6 Room 6.3

Mathematics Education and Society – PART I

Organizers: Carla Santos, Cristina Dias, Ana Caballero

Chair: Carla Santos

Assessing higher education studentsŠ ability to read and interpret statistical tables and graphs

Carla Santos

On the pedagogical practices in teaching and learning Mathematics presented in CNaPPES

Cristina Dias

How is the employment situation of mathematics teachers?

Ana Caballero

Reading and understanding of statistical graphs in secondary students of Badajoz

José Ruiz

11:30 − 12:00

Plenary Session 6 Room 6.1

Adopting Mathematical Aspects and Computational Techniques

to Optimal Experimental Design Theory

Christos Kitsos

12:00 – 13:15 Lunch

13:15 – 15:00 Parallel Sessions

Organized Session S5 Room 6.1

Special functions and applications

Organizers: José Luís Cardoso, Maria das Neves Rebocho

Chair: Maria das Neves Rebocho

Höder and Minkowski's inequalities related with q-integrals

José Luís Cardoso

Orthogonal polynomials via polynomial mappings and some applications

Márcio N. de Jesus

Explicit relations between cumulants and moments

Patrícia Antunes

Product formulas, generalized convolutions and integral transforms

Rúben Sousa

On Laguerre–Hahn orthogonal polynomials on the real line

Maria das Neves Rebocho

Organized Session S4   **Room 6.2**

Computational Mathematics and Statistics and its Applications

Organizers: M. Filomena Teodoro, Marina A.P. Andrade

Chair: M. Filomena Teodoro

On the distribution of linear combinations of chi-square random variables

Carlos A. Coelho

Inferences on multivariate normal covariance matrices based on synthetic data

Ricardo Moura

Water meters inaccuracies registrations: A first approach of a portuguese case study

Marina Andrade

Three-dimensional velocity ?eld for blood ?ow using the power-law viscosity function

Fernando Carapau

Students performance and individual profile, a generalized linear models approach

M. Filomena Teodoro

Organized Session S6   **Room 6.3**

Mathematics Education and Society – PART II

Organizers: Carla Santos, Cristina Dias, Ana Caballero

Chair: Carla Santos

Interpretation and communication of statistical data in primary education

Luis Maya Jaramillo

Application of SLP methodology in a factory of aluminium Doors and medium density fiber plates aided by multicriteria decision analysis

Maria Varadinov

Income, consumption and saving of Portuguese households in numbers

João Romacho

The use of computers in Mathematics classes of secondary education

Adelaide Proença

Teaching mathematical concepts using a videogame prototype

Ana Caballero

15:00-15:15   **Closing Ceremony / (Prize SILABO/Book for the two student's with the best presentations) [Room 6.1]**

15:15   **City Tour (Discovering Urban Art)**

Contents

Keynote Speakers

Organized Sessions

Contributed Talks

Contributed Poster

Keynote Speakers

Population growth in random environments

<u>Carlos A. Braumann</u>[1,2]

Centro de Investigação em Matemática e Aplicações, Instituto de Investigação e Formação Avançada, Universidade de Évora[1]
Departamento de Matemática, Escola de Ciências e Tecnologia, Universidade de Évora[2]

E-mail address: *braumann@uevora.pt*

Abstract

The growth of populations was traditionally modelled by ordinary differential equations (ODE), but there are random fluctuations in the environmental conditions that affect the growth rate. Particularly in wildlife populations, such fluctuations may cause considerable deviations from the mean dynamics and may also produce new qualitative features that can be studied using stochastic differential equation (SDE) models. The other advantage of using SDE models is that the statistical issues (estimation, model choice, prediction) become quite natural and, unlike ODE models, do not have to be artificially imposed as an outer layer (like a regression structure, which actually is inadequate to deal with randomness affecting the very dynamics).

The main results concerning extinction, extinction time, conditions for existence of a stochastic equilibrium and how Allee effects affect the dynamics will be reviewed for general models, so as to obtain qualitative results that are model robust. The particular case of fished populations and the issues of optimizing the harvesting yield or the profit will also be addressed. In what concerns quantitative predictions in applications, one needs to use a specific model, usually a reasonable approximation of the unknown correct model, and so we will also study the resulting error in such predictions.

This brief account reviews publications (that cannot be referenced in the space available here) by several authors, particularly by this author and co-authors Patrícia A. Filipe, Clara Carlos and Nuno M. Brites.

The issue of the stochastic calculi of Itô and Stratonovich and its incidence on modelling will also be briefly addressed.

Keywords: stochastic differential equations, randomly varying environment, population growth, extinction, harvesting.

——⊙——

Acknowledgements

Adopting Mathematical Aspects and Computational Techniques to Optimal Experimental Design Theory

Christos P. Kitsos

Department of Informatics, University of West Attica[1]
([1]ex Technological Educational Institute of Athens)

Abstract

The target of this paper is to investigate, analyze and discuss the Optimal Experimental Design Theory from the point of view of a Mathematical extension and the adoption of Computational techniques in practice. From a Mathematical point of view the invariance either as a group theory or through affine Geometry is essential, as well as the Group theory or Projective Geometry. The optimality criteria, through the Mathematical Analysis, play also an important role on the development of the Design Theory . When the evaluation of optimal Design points is requested, Computational, Techniques are adopted especially when the sequential principle of Design is adopted. This excellent merge of various Mathematical Aspects with the computational support created the Optimal Experimental Design Theory - the back bone of Statistics, as some researchers believe.

Confidence Ellipsoids for Additive Models

João T. Mexia[1] **Sandra S. Ferreira**[2,3]**, Célia Nunes**[2,3] **and Dário Ferreira**[2,3]

Center of Mathematics and its Applications, Faculty of Science and Technology, New University of Lisbon, Monte da Caparica, Portugal[1]
Department of Mathematics of University Beira Interior[2]
Center of Mathematics and Applications, University of Beira Interior, Covilhã, Portugal[3]

E-mail address: *jtm@fct.unl.pt; sandraf@ubi.pt; celian@ubi.pt;dario@ubi.pt*

Abstract

We use a classical result on cumulant generation functions to obtain Least Square Estimators for the components assumed to be independent and identical distributed of the vectors $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_m$ of an additive model

$$\boldsymbol{Y} = \boldsymbol{X}_0\boldsymbol{\beta}_0 + \sum_{i=1}^{m} \boldsymbol{X}_i\boldsymbol{Z}_i + \boldsymbol{e}, \tag{1}$$

where the $\boldsymbol{Z}_1, ..., \boldsymbol{Z}_m$ and $\boldsymbol{e}$ are assumed to be independent and $\boldsymbol{e}$ is normal and homocedastic.

We then obtain the cumulants of order 2,3 and 4 of the $V_p = \boldsymbol{\alpha}_l^\top(\boldsymbol{Y} - \boldsymbol{\mu})$, where $\boldsymbol{\mu} = X_0\beta_0$ and $\boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_n$ constituting an orthonormal basis for $\mathbb{R}^n$. Applying the Edgeworth expressions we obtain (approximate) quantiles $u_{p,l}$ for U_h, $h = 1, ..., n$. Lastly we adjust confidence ellipsoids to the quantiles, reducing the quadratic sum to an inner product. Thus

$$\boldsymbol{v}^\top \boldsymbol{M}\boldsymbol{v} = \boldsymbol{v}^{<2>t}\boldsymbol{s}(\boldsymbol{M}), \tag{2}$$

where, with $\boldsymbol{v} = (v_1, ..., v_m)$ and $\boldsymbol{M} = [m_{l,b}]$, we have

$$\begin{cases} \boldsymbol{v}^{<2>} = (v_1^2, ..., v_1 v_m, ..., v_m^2) \\ \boldsymbol{s}(M) = (m_{1,1}, ..., m_{1,w}, ..., m_{w,w}) \end{cases}. \tag{3}$$

The adjusted ellipsoid will be the (approximate) p level ellipsoid.

Keywords: Additive Models, Cumulants, Cumulant Generating Function, Quantiles, Moments.
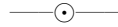
———⊙———

Acknowledgements

References

[1]Akhiezer, N. I. (1965), The Classical Moment Problem, *N. Kemme, Transl.*, Hafner, New York.

[2]Balakrishnan, N., Johnson,N. L., and Kotz, S. (1998), A note on relationships between moments, central moments and cumulants from multivariate distributions, *Statistics & probability letters*, **39**: 49–54, Elsevier.

[3]Daniels, H. E. (1987). Tail probability approximations. *Review of the International Statistical Institute*, **55**: 37–46.

[4]Kendall, M. and Stuart, A. (1958). *The Advanced Theory of Statistics*. Charles Griffin & Company.

[5]Landau, H. J. (1987), Moments in Mathematics, Proceedings of Symposia in Applied Mathematics, **37**, *Amer. Math. Soc.*, Providence, RI.

[6]Schultze, J. and Steinebach, J. (1996). On least squares estimates of an exponential tail coecient. Statistics & Decisions **14**, 353–372.

Testing interaction in unrepeated designs

Karl Moder

Department of Landscape, Spatial and Infrastructure Sciences University of Natural resources and Life Sciences in Vienna, Austria

**E-mail address:** *karl.moder@boku.ac.at*

Abstract

Unrepeated experimental designs like Randomized Complete Block Designs, Incomplete Block Designs. split-plot designs are probably the most widely used experimental designs. Despite many advantages they suffer from one serious drawback: In a common linear model there is no test on interaction effects in ANOVA, as there is only one observation for each combination of a block and factor effects.

Several people tried to overcome this problem by using some additional restrictions to the model. None of these methods are used in practice, especially as most of them are non-linear. A review on such tests is given by Karabatos [2005] and Alin & Kurt [2006].

Here a new method is introduced which permits a test of interactions in non-repeated designs. The underlying models are linear and identical to those of common factorial designs. A big addvantage of the propoesed model is, that one can use any common statistical program packages like SAS, SPSS, R ... for analysis by doing a few additional calculations.

Keywords: unrepeated designs, block designs, interaction, power.

————⊙————
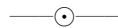
References

[1]Karabatos G. 2005. *Additivity Test.* Encyclopedia of Statistics in Behavioral Science. Wiley, New York. 25–29.

[2]Alin A., Kurt S. 2006. *Testing non-additivity (interaction) in two-way anova tables with no replication.* Statistics in Medicine. 15: 63–85.

Measuring the Shape of Boundary Districts Using the Convexity Ratio

J R Bozeman[1] **and J Mori**[2]

American University of Malta[1] ; *University of Vermont*[2]

E-mail address: *james.bozeman@aum.edu.mt; jillian.mori@uvm.edu*

Abstract

When measuring the shape of voting districts in redistricting plans, for example when trying to identify potential gerrymanders, some districts receive poor scores due to the fact that other objects beyond the redistricters control protrude into the district. This happens when, for example, another State intrudes on a district, or the district is bounded by a body of water, etc. Hence the shape of a state can occasionally lead to *poorly shaped* districts, by definition, but which cannot be considered as such. This occurs regardless of the scoring method used ('compactness' or convexity, e.g.) since these boundaries are then unchangeable. Dealing with this issue is the subject of much current research. We solve this problem assuming the *convexity ratio* is used when measuring the shape of a district.

In this talk we exhibit different methods for measuring the shape of districts with immoveable boundaries. The flaws of some of these ideas are shown. We then solve the unchangeable boundary district problem in the context of the convexity ratio by introducing convex hulls at the immoveable portions. This technique is readily programmable, one of our goals. The result is then applied to artifiical and real boundary districts so that the geometry can be visualized. We conclude by discussing future work.

Keywords: boundary districts, convexity ratio, convex hull, gerrymandering, redistricting.

——⊙——
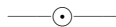
References

[1]Bozeman JR, Davey M, Hutchins S, et al. (2018) Redistricting without gerrymandering, utilizing the convexity ratio, and other applications to business and industry. *Appl Stochastic Models Bus Ind.*:1–17. https://doi.org/10.1002/asmb.2396

Development of tools for numerical inversion of the characteristic functions and the exact distributions of multivariate test statistics

Viktor Witkovský

Institute of Measurement Science,
Slovak Academy of Sciences, Bratislava, Slovakia

E-mail address: *witkovsky@savba.sk*

Abstract

The exact statistical inference frequently leads to a non-standard probability distributions of the considered test statistics, which can be fully specified by their moments or their characteristic functions. Typically, distribution of many estimators and test statistics can be structurally expressed as a linear combination or product of independent random variables with known distributions, see e.g. [1–3], and in particular, it is true for many standard multivariate test criteria, see e.g. [4].

However, analytical inversion from the moments (by using the inverse Mellin transform) or from the characteristic function (by using the inverse Fourier transform) frequently leads to complicated expressions of the distribution functions (PDF/CDF/QF).

Alternative approaches are frequently based on the approximation methods (small sample or asymptotic), or the simulation methods. In particular, the null-distribution of the most common likelihood ratio based test statistics can be approximated by the standard asymptotic chi-square approximation. The small sample approximations based on first moments/cumulants include the Edgeworth or Gram-Charlier expansions about the known distribution. In specific situations the more sophisticated approximations can be used, such as the near-exact distributions suggested by Carlos A. Coelho and his co-authors, see e.g. [5-7]. For more details see
`https://sites.google.com/site/nearexactdistributions/home`.

For most applications, the method based on numerical inversion of the characteristic functions is fully sufficient. Applicability of the approach is illustrated by computing the exact null and non-null distributions of selected test statistics used in multivariate hypothesis testing. The suggested methods

and algorithms have been implemented as *CharFunTool: The Characteristic Functions Toolbox* in MATLAB and R programming environment for statistical computing. The work is still under progress and cooperation with interested researchers on its further development is very welcome, see [8, 9].

Keywords: characteristic function, numerical inversion, multivariate test statistics, exact distribution.

Acknowledgements

References

[1]Witkovský, V. (2001). On the exact computation of the density and of the quantiles of linear combinations of t and F random variables. *Journal of Statistical Planning and Inference*, **94**, 1–13.

[2]Witkovský, V. (2001). Computing the distribution of a linear combination of inverted gamma variables. *Kybernetika*, **37**, 79–90.

[3]Witkovský, V. (2002). Exact distribution of positive linear combinations of inverted chi-square random variables with odd degrees of freedom. *Statistics & Probability Letters*, **56**, 45–50.

[4]Mathai, A.M. (1973). A review of the different techniques used for deriving the exact distributions of multivariate test criteria. *Sankhyā: The Indian Journal of Statistics, Series A*, 39–60.

[5]Coelho, C.A. (2004). The generalized near-integer gamma distribution: A basis for a near-exact approximations to the distribution of statistics which are the product of an odd number of independent beta random variables, *Journal of Multivariate Analysis*, **89**, 191–218.

[6]Marques, F.J., Coelho, C.A. and Arnold, B.C. (2011). A general near-exact distribution theory for the most common likelihood ratio test statistics used in multivariate analysis. *Test*, **20**, 180–203.

[7]Marques, F.J., Coelho, C.A. and Rodrigues, P.C. (2017). Testing the equality of several linear regression models. *Computational Statistics*, **32**, 1453–1480.

[8]Witkovský, V. (2018) CharFunTool: The Characteristic Functions Toolbox (MATLAB). `https://github.com/witkovsky/CharFunTool/`.

[9]Gajdoš, A. (2018). CharFunToolR: The Characteristic Functions Toolbox (R). `https://github.com/gajdosandrej/CharFunToolR`.

Organized Sessions

Organized Session 1

Statistical Applications in Environmental Science

Organizer: Manuela Neves (Portugal)

Extreme value theory in the analysis of environmental events

**M. Ivette Gomes**[1,2] and **Fernanda Figueiredo**[3,2]

Faculdade de Ciências, Universidade de Lisboa[1]
Centro de Estatística e Aplicações da Universidade de Lisboa[2]
Faculdade de Economia, Universidade do Porto[3]

E-mail address: *ivette.gomes@fc.ul.pt; otilia@fep.up.pt*

Abstract

Climatic changes have contributed for the occurrence of severe winters and summers, with incredible low or high temperatures, flows, hurricanes, huge areas of forest destroyed by fires, and so on. To deal and try preventing this type of catastrophic events, it is important an adequate modeling of the tails, or at least, a reliable estimation of the risk of exceedance of an adequate safety level. Extreme value theory plays then a crucial role in modeling those tails, through an initial estimation of the *extreme value index* (EVI), the primary parameter of rare events. In this study we consider datasets from the areas of ecology and hydrology, and after a preliminary analysis of the data, we proceed with a semi-parametric estimation of the EVI and of related parameters of extreme events, like the return period of a high level.

Keywords: data analysis, semi-parametric estimation, statistics of extremes.

——⊙——
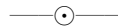
Modeling rare large events is one of the most important issues in the field of *extreme value theory* (EVT). On the basis of an assumed non-degenerate asymptotic behavior for linearly normalized maxima, it is common to consider the *general extreme value* (GEV) *cumulative distribution function* (CDF), GEV_ξ, to model such events. The standard GEV_ξ CDF is defined by

$$\text{GEV}_\xi(x) := \begin{cases} \exp(-(1+\xi x)^{-1/\xi}), & 1+\xi x > 0, \text{ if } \quad \xi \neq 0, \\ \exp(-\exp(-x)), & x \in \mathbb{R}, \qquad \text{if} \quad \xi = 0, \end{cases}$$

where ξ denotes the EVI, a parameter associated with the frequency of such extreme events. Regarding EVI-estimation, several studies have been carried out, to achieve efficient estimators for ξ. For details see, among others, the

books by Beirlant *et al.* (2004), de Haan and Ferreira (2006), Gomes *et al.* (2013) and the recent overview by Gomes and Guillou (2015).

After an adequate transformation of the available data, let us assume that we have access to $\{X_i,\ 1 \leq i \leq n\}$, a random sample of size n from a model F, and let us denote by $\{X_{i:n},\ 1 \leq i \leq n\}$ the sample of associated ascending order statistics (OSs). In statistical EVT the estimation is often based on the largest k upper OSs, with k an intermediate sequence of integers. For heavy tailed models ($\xi > 0$), the most common semi-parametric EVI-estimators are the Hill estimators (Hill, 1975), denoted by H(k), the average of the log-excesses $V_{ik} := \ln X_{n-i+:n} - \ln X_{n-k:n},\ 1 \leq i \leq k$. Since these estimators have the disadvantage of having a reasonably high bias for low thresholds, i.e., for large k, several reduced-bias EVI-estimators have been introduced in the literature, associated not only with the H EVI-estimators, but also with generalized means' EVI-estimators (see, Caeiro *et al.*, 2016, where some of those estimators are discussed), dependent on an extra tuning parameter, say p. For the choice of these tuning parameters, (k,p), bootstrap methodologies are crucial, enabling to obtain reliable semi-parametric estimates of any parameter of rare events, like a high quantile, the return period of a high level or the two primary parameters of extreme events, the EVI and the *extremal index* (EI), related to the degree of local dependence in the extremes of a stationary sequence (see, Gomes, 1993, for the EI-estimation, together with an application to environmental data). After a brief discussion of these estimators, and extending the discussion in Figueiredo and Gomes (2007), we proceed with the analysis of data sets in the field of environment, related to burned areas in forests and river flows, trying to motivate the use of EVT to solve problems in environmental sciences.

Acknowledgements

References

[1] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*, Wiley, England.

[2] Caeiro, F., Gomes, M.I., Beirlant, J. and de Wet, T. (2016). Mean-of-order-p reduced-bias extreme value index estimation under a third-order framework. *Extremes* **19**(4): 561–589.

[3] Figueiredo, F. and Gomes, M.I. (2007). Analysis of some environmental extremal events. In Fraga Alves *et al.* (eds.), *Statistical Extremes and Environmental Risk*, CEAUL Editions, Portugal, 39–42.

[4] Gomes, M.I. (1993). On the estimation of parameters of rare events in environmental time series. In V. Barnett and K.F. Turkman (eds.). *Statistics for the Environment*, Wiley, New York, 225–241.

[5] Gomes, M.I. and Guillou, A. (2015). Extreme value theory and statistics of univariate extremes: a review. *International Statistical Review* **83**(2): 263–292.

[6] Gomes, M.I., Fraga Alves, M.I. e Neves, C. (2013). *Análise de Valores Extremos: Uma Introdução*, Edições SPE, Lisboa, Portugal.

[7] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: an Introduction*, Springer Science+Business Media, LLC, New York.

[8] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics* **3**: 1163–1174.

# Statistical Ecology: a few examples

Russell Alpizar-Jara[1,2]

University of Évora, Departament of Mathematics, ECT[1]
Research Center of Mathematics and Applications, CIMA-UE, IIFA[2]

E-mail address: *alpizar@uevora.pt*

Abstract

Last few decades have been characterized by a rapid increase in computer power, new technological advances for data collection, data accessibility, and by the need to develop adequate statistical methodology to address challenges on data and model complexities. Statistical Ecology has not fell behind. Recent developments provide tools to answer complex ecological problems. There is a very large set of statistical applications in Ecology. For instance, an intensive research area is the estimation of animal abundance and related demographic parameters. I will illustrate how a simple idea for a single population analysis can be extended to answer more general community-level questions. I will also refer to recent applications and extensions of classical methods to account for detection or capture probabilities in elusive populations. Finally, I will highlight the importance of collaborative and interdisciplinary work to contribute to the advancement of the field.

Keywords: capture-recapture, community dynamics, integrated modelling, non-invasive genetics analyses.
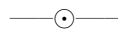
———⊙———

References

[1]Hilborn, R. and Mangel, M. (1997) *The Ecological detective: Confronting Models with Data*, Princeton, Princeton University Press.
[2]Williams, B., Nichols, J.D., and Conroy, M. (2002). *Analysis and Management of Animal Populations.* Academic Press.

Application of Statistics of Extremes to the Identification of Biodiversity Patterns

Luís Borda de Água[1,2]

University of Porto[1]
CIBIO/INBIO[2]

E-mail address: *lbagua@cibio.up.pt*

Abstract

In a landmark paper, He and Hubbell [1] pointed out that to determine the number of species as area size increases - the so-called species accumulation curve - what matters is to find the first individual of each species. They also pointed out that to find out how many species go extinct as a function of the size of the area cleared - the so-called endemic species accumulation curve - what matters is to find the last individual. From a practical point of view, one can think of a sampling scheme consisting of concentric circles where we record the distance from the central point to the first individual of each species (the minimum distance), and the distance to the furthest individual (the maximum distance). Therefore, given an area with S species, this sampling scheme leads to S points for the minimum distances and S for the maximum distances. The analysis of these maxima and minima (extremes) falls within the realm of the statistics of extremes. Here we show preliminary results of the application of statistics of extremes to species richness to data on tropical trees collected in a 50 ha plot in Barro Colorado Island, Panama, and show the resulting "species accumulation curves" (based on the minima) and the "endemic accumulation curves" (based on the maxima). More importantly, we suggest that the asymptotic theorems of statistics of extremes provide a theoretical framework to the study the species accumulation curves and other biodiversity scaling patterns.

Keywords: Biodiversity, Endemic Accumulation curves, Species Accumulation Curves, Statistics of Extremes.
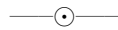
———⊙———

Acknowledgements

References

[1]He, F. and Hubbell, S.P., 2011. Species-area relationships always overestimate extinction rates from habitat loss. Nature, 473: 368-371.

Organized Session 2

Stochastic Differential Equations: Applications in Finance, Management and Medicine

Organizer: Nuno Brites (Portugal)

From an Ordinary Differential Equation Model to an Open Population Markov Chain Model, via Stochastic Differential Equations; models for HIV infection in individuals and populations.

Manuel L. Esquível[1,2] and Paula Patrício[1,2] and Gracinda R. Guerreiro[1,2]

Departamento de Matemática, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, Campus de Caparica, 2829-516, Caparica, Portugal[1]
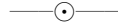*Centro de Matemática e Aplicações FCT NOVA UNL[2]*

E-mail address: *mle@fct.unl.pt; pcpr@fct.unl.pt; grg@fct.unl.pt*

Abstract

We present an initial exploration of a method for the association of a open population Markov chain model – with a finite number of states – to some phenomena that may be, by force of its intrinsic characteristics, best modeled by ODE, at least in some average sense. The ODE model here presented is formulated as a dynamic change between two regimes; one regime is of mean reverting type and the other is of inverse logistic type. For the general purpose of defining a open Markov chain model for an human population, we associate an Ito process to the OD equations, by means of the addition of Gaussian noise terms which may be thought to model non essential characteristics of the phenomena with small and undifferentiated influences. The next step consists on discretizing the Ito processes and using the sequence of values obtained to define, by simulation, trajectories that, in turn, may define transitions of a finite valued Markov chain, if the state space of the Ito process is partitioned according to some rule. We detail the application of these ideas to the study of the evolution of a Portuguese population of newly diagnosed with HIV. For that purpose the state space of the Ito process referred is partitioned in six infection classes. We detail the evolution of the population in these classes under two different projections for the evolution of the newly diagnosed. The method here presented connects the model for the evolution of the HIV viral load and the CD4 leucocytes

count to a Markov chain open model for the Portuguese population of HIV diagnosed.

Keywords: Ordinary Differential Equations, Stochastic Differential Equations, Markov Chains, Infection Modeling, Population Dynamics.

—⊙—

Acknowledgements

References

[1]Esquível M. L., Fernandes J. M., Guerreiro G.R. (2014) On the evolution and asymptotic analysis of open Markov populations: application to consumption credit. *Stoch. Models* **30(3)**:365–389.

[2]Esquível M. L., Guerreiro G.R., Fernandes J. M. 2017) Open Markov chain scheme models. *REVSTAT* **15(2)**: 277–297.

[3]Esquível M. L., Patrício P., Guerreiro G.R. (2018) From ODE to MC, via SDE. Associated models for HIV infection in individuals and populations *(submitted)* .

Valuation of lookback options and turbo warrants on defaultable stocks

José Carlos Dias[1,2] **and João Pedro Vidal Nunes**[1,2]

Instituto Universitário de Lisboa (ISCTE-IUL), Lisbon, Portugal[1]
*Business Research Unit (BRU-IUL), Lisbon, Portugal*[2]

E-mail address: *jose.carlos.dias@iscte-iul.pt; joao.nunes@iscte-iul.pt*

Abstract

A turbo call (resp., put) warrant is a contract whose payoff is similar to the one of a standard call (resp., put) option if a pre-specified barrier has not been hit by the underlying asset price before maturity. If the underlying asset price hits the barrier, a rebate is paid. For turbo call warrants the rebate is the difference between the lowest recorded stock price during a pre-specified period after the barrier is hit and the strike price, and for turbo put warrants the rebate is calculated as the difference between the strike price and the largest recorded stock price during a pre-specified period after the barrier is hit. Hence, the price of turbo warrants can be decomposed into two components: a barrier option and a lookback option.

This paper offers three concrete contributions to the financial engineering literature. First, we extend the first passage time approach of [3] to price fixed-strike and floating-strike lookback options on defaultable stocks assuming the underlying stock price follows the jump to default extended constant elasticity of variance (JDCEV) model proposed by [1]. Therefore, and as our second contribution, we are able to efficiently price turbo warrants on defaultable equity since the barrier option component of the contract is effortlessly evaluated through the efficient pricing solutions recently offered by [3]. Finally, our novel pricing methodology is compared with the Laplace transform scheme developed in [4] for the CEV model of [2], because this model is known to be nested as a special case of the more general JDCEV modeling setup.

Keywords: lookback options, turbo warrants, default, CEV model, JDCEV model.
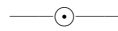
———⊙———

Acknowledgements

References

[1] P. Carr, V. Linetsky (2006). A jump to default extended CEV model: An application of Bessel processes. *Finance and Stochastics* 10, 303–330.

[2] J. C. Cox (1975). Notes on option pricing I: Constant elasticity of variance diffusions, Working Paper, Stanford University. Reprinted in *Journal of Portfolio Management*, 23 (1996), 15-17.

[3] J. C. Dias, J. P. Nunes, J. P. Ruas (2015). Pricing and static hedging of European-style double barrier options under the jump to default extended CEV model. *Quantitative Finance* 15, 1995–2010.

[4] H. Y. Wong, C. M. Chan (2008). Turbo warrants under stochastic volatility. *Quantitative Finance* 8, 739–751.

Stochastic differential equations driven by mixed fractional Brownian motion and an application in finance

João Guerra[1,2]

ISEG - Universidade de Lisboa[1]
CEMAPRE[2]

E-mail address: *jguerra@iseg.ulisboa.pt*

Abstract

We discuss the definition of a stochastic integral with respect to fractional Brownian motion and also the existence and uniqueness of solutions for stochastic differential equations driven simultaneously by a fractional Brownian motion with Hurst parameter $H > 1/2$ and a standard Brownian motion (mixed fractional Brownian motion). A financial application of these type of stochastic differential equations is the so-called mixed Brownian-fractional Brownian pricing model, which will also be discussed.

Keywords: Fractional Brownian motion, Stochastic Differential equations, mixed fractional Brownian motion, arbitrage.

The application of clustering algorithms to the hand's biometrics for personal recognition

Lúcia Paiva Sousa

ESTGV - IPV/CI&DETS

E-mail address: *lsousa@estv.ipv.pt*

Abstract

In data analysis, the hierarchical clustering algorithms are powerful tools allowing to identify natural clusters, often without any priori information of the data structure, and are quite often used because provide a graphical representation of the resulting partitions, a hierarchy or dendrogram, revealing more information than non-hierarchical algorithms that returns a unique partition. Moreover, it is not necessary specify the number of clusters à priori. This work focuses on the problem of choosing the best partition in hierarchical clustering. The procedure to search for the best partition is made in the nested set of partitions, defined by the hierarchy. It is applyed an improved method, SEP/COP, to obtain the best partition, based on a wide set of partitions. The study is considering a set of experiments using two-dimensional synthetic data sets and a real-world data set, based on the biometrics of the hands. This database is provided from Bosphorus Hand Database, in the context of recognition of the identity of a person by using the features of her hand/biometrics. Regarding to real data, the results of the experiments demonstrated that SEP/COP hierarchical clustering algorithm can contribute to identification systems based on the biometrics of the hands.

Keywords: Hand biometrics, Recognition, Hierarchical clustering, Partition, Validation.

Acknowledgements

References

[1]Gurrutxaga, I., Albisua, I., Arbelaitz, O., Martin, J., Muguerza, J., Perez, J. and Perona, I. (2010) SEP/COP: An efficient method to find the best partition in the hierarchical clustering based on a new cluster validity index. Pattern Recognition, 43: 3364-3373.

[2]Hubert, L. and Arabie, P. (1985) Comparing Partitions. Journal of Classification, 2: 193-218.

[3]Konukoglu, E., Yoruk, E., Darbon, J. and Sankur, B. (2006) Shape-based hand recognition. IEEE Trans. on Image Processing, 15 (7): 1803-1815.

Stochastic differential equations in fisheries management with stepwise effort

<u>Nuno M. Brites</u>[1,2] and Carlos A. Braumann[2,3]

Instituto Superior de Economia e Gestão, Universidade de Lisboa[1]
*Centro de Investigação em Matemática e Aplicações, Instituto de
Investigação e Formação Avançada,
Universidade de Évora*[2]
*Departamento de Matemática, Escola de Ciências e Tecnologia,
Universidade de Évora*[3]

E-mail address: *nbrites@iseg.ulisboa.pt; braumann@uevora.pt*

Abstract

Stochastic optimal control methods have been applied to derive optimal harvesting policies in a randomly varying environment. With these policies, the fishing effort, $E(t)$, must be adjusted at every instant and will have sudden frequent transitions between maximum or high harvesting efforts and low or null harvesting efforts. These transitions in effort are not compatible with the logistics of fisheries. Besides, the period of low or no harvesting poses social and economical undesirable implications (intermittent unemployment is just one of them). In addition to such shortcomings, these optimal policies require the knowledge of the population size at every instant, to define the appropriate level of effort. Therefore, these policies should be considered unacceptable and inapplicable.

In [1,2], a constant fishing effort, $E(t) \equiv E$, was assumed. For a large class of models it was found that, taking a constant fishing effort, there is, under mild conditions, a stochastic sustainable behaviour. Namely, the probability distribution of the population size at time t will converge, as $t \to +\infty$, to an equilibrium probability distribution (the so-called stationary or steady-state distribution) having a probability density function (the so-called stationary density). In particular, for the Gompertz model, the stationary density was found, and the effort E that optimizes the steady-state yield was determined. The issue of profit optimization, however, was not addressed.

In previous works (see, for instance, [3,4]), we have developed as an alternative to the variable effort policies, a sustainable constant effort policy based on profit optimization, which is extremely easy to implement and leads

to a stochastic steady-state. We determine the constant effort that maximizes the expected profit per unit time at the steady-state in the general case and for the specific cases of the logistic and the Gompertz models. One might think that a constant effort policy would result in a substantial profit reduction compared with the optimal variable effort policy, but we show, using data on the particular stock of the Bangladesh shrimp (*Penaeus monodon*), that this is not the case. This new policy, rather than switching between large and small or null fishing effort, keeps a constant effort and is therefore compatible with the logistics of fisheries. Furthermore, this alternative policy does not require knowledge of the population size.

Since the optimal variable effort policy is not applicable, we present here, for the Gompertz model, sub-optimal policies, named stepwise policies, where the harvesting effort under the optimal variable effort policy is determined at the beginning of each year (or of each biennium) and kept constant during that year (biennium). These policies are not optimal and still pose some social problems, but have the advantage of being applicable, since the changes on effort are less frequent and compatible with the fishing activity. Furthermore, although we still need to keep estimating the fish stock size, we do not need to do it so often. Replacing the optimal variable effort policy by these stepwise policies has the advantage of applicability but, at best, considerably reduces the already small profit advantage the optimal variable effort policy has over the optimal constant effort policy. In some cases, the optimal sustainable policy even outperforms these stepwise policies in terms of profit.

Keywords: stochastic differential equations, fisheries management, profit optimization, stepwise effort.
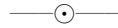
—⊙—

Acknowledgements

References

[1] C. A. Braumann (1981). Pescar num mundo aleatório: um modelo usando equacões diferenciais estocásticas. In *Proceedings of the XII Congresso Luso-Espanhol de Matemática, Coimbra*, 301–308.

[2] C. A. Braumann (1985). Stochastic differential equation models of fisheries in an uncertain world: extinction probabilities, optimal fishing effort, and parameter estimation.

In *Mathematics in Biology and Medicine*, eds.: V. Capasso, E. Grosso, S. L. Paveri-Fontana, Springer, Berlin, 201–206.

[3] N. M. Brites, C. A. Braumann (2017). Fisheries management in random environments: Comparison of harvesting policies for the logistic model. In *Fisheries Research*, 238–246.

[4] N. M. Brites (2017). Stochastic differential equation harvesting models: sustainable policies and profit optimization, Ph.D. Thesis, Universidade de Évora, Portugal.

Organized Session 3

Statistical Challenges and its applications

Organizer: Dora Prata Gomes (Portugal)

Comparison of classes of generalized Hill estimators

Frederico Caeiro[1,2], **Ivanilda Cabral**[3] **and M. Ivette Gomes**[4,5]

[1] *Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal*
[2] *Centro de Matemática e Aplicações (CMA), Universidade Nova de Lisboa*
[3] *Departamento de Ciência e Tecnologia, Universidade de Cabo Verde, Cabo Verde*
[4] *DEIO, Faculdade de Ciências, Universidade de Lisboa, Portugal*
[5] *Centro de Estatística e Aplicações da Universidade de Lisboa (CEAUL)*

E-mail address: *fac@fct.unl.pt; ivanilda.cabral@docente.unicv.edu.cv; ivette.gomes@fc.ul.pt*

Abstract

In this work, we are interested in the estimation of the extreme value index, the primary parameter of extreme events. For heavy tails, classical extreme value index estimators, such as the Hill [7] estimator, have usually a strong bias. Consequently those estimators are quite sensitive to the number of top order statistics used in the estimation. To improve the mean squared error of the aforementioned estimators, many alternative estimators have appeared in the literature ([1,4]). We analyse several generalizations of the Hill estimator, like the ones in [2,3,5,6,8]. Our aim is to study their non degenerate asymptotic behaviour and to compare them altogether.

Keywords: Extreme value index, heavy tails, semi-parametric estimation, Statistic of extremes.

Acknowledgements

References

[1] Beirlant, J., Caeiro, F. and Gomes, M.I. (2012) An overview and open research topics in statistics of univariate extremes. *Revstat* **10**:1 1–31.

[2] Brilhante, M.F., Gomes, M.I. and Pestana, D. (2013) A simple generalization of the Hill estimator. *Comput. Statist. and Data Analysis* **57**:1 518–535.

[3] Caeiro, F., Gomes, M.I., Beirlant, J. and de Wet, T. (2016) Mean-of-order p reduced-bias extreme value index estimation under a third-order framework. Extremes **19**:4, 561–589.

[4] Gomes, M.I. and Guillou, A. (2015). Extreme value theory and statistics of univariate extremes: a review. *International Statistical Review* **83**:2 263–292.

[5] Gomes, M. I. and Martins, M.J. (2001) Generalizations of the Hill estimator – asymptotic versus finite sample behaviour. *Journal of Statistical Planning and Inference* **93** 161–180.

[6] Gomes, M.I. and Martins, M.J. (2004) Bias reduction and explicit semi-parametric estimation of the tail index, *Journal of Statistical Planning and Inference* **124** 361–378.

[7] Hill, B. M. (1975) A Simple General Approach to Inference About the Tail of a Distribution. *Ann. Statist.* **3**:5 1163–1174.

[8] Penalva, H., Gomes, M.I., Caeiro, F. and Neves, M.M. (2018) A couple of non Reduced Bias Generalized Means in Extreme Value Theory: an Asymptotic Comparison. Accepted for publication in Revstat–Statistical Journal.

Introducing nonparametric predictive inference methods for reproducibility of likelihood ratio tests

Filipe J. Marques[1,2]**, Frank P. A. Coolen and Tahani Coolen-Maturi**[3]

Departamento de Matemática, FCT, UNL[1]
Centro de Matemática e Aplicações (CMA), FCT, UNL[2]
Durham University, Durham, UK[3]

E-mail address: *fjm@fct.unl.pt; frank.coolen@durham.ac.uk; tahani.maturi@durham.ac.uk*

Abstract

This work introduces the nonparametric predictive inference approach for reproducibility of likelihood ratio tests. The general idea of this approach is outlined for tests between two simple hypotheses, followed by an investigation of reproducibility for tests between two Beta distributions. The work reports on the first steps of a wider research programme towards tests involving composite hypotheses and substantial computational challenges.

Keywords: Beta distribution, lower and upper probabilities, nonparametric predictive inference, likelihood ratio test, reproducibility probability.
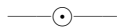
———⊙———

Pareto index estimation through a modification of the probability weighted moment method

Ayana Mateus[1,2] and **Frederico Caeiro**[1,2]

Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal[1]
Centro de Matemática e Aplicações (CMA)[2]

E-mail address: *amf@fct.unl.pt; fac@fct.unl.pt*

Abstract

The Pareto distribution result from the work of the italian Vilfredo Pareto in 1897 who noticed that the number of taxpayers with income greater than x could be approximated by $b\,x^{-a}$, for some positive parameters a and b. Nowadays it has been extensively used for modelling events in fields such as bibliometrics, demography, insurance, finance, among others. In this work, we propose a consistent estimator for any positive shape parameter the Pareto distribution. This objective is achieved through a modification of the probability weighted moments method. The asymptotic limit distribution for this estimator is also derived. Finally, we compare, through a Monte Carlo simulation study, the finite sample performance of the proposed estimator, in terms of the mean value and root mean square error, with the most usual estimators from the literature.

Keywords: Pareto distribution, Monte Carlo method, Probability weighted moments, Moment estimator, Maximum likelihood estimator.

Acknowledgements

References

[1]B. C. Arnold, N. Balakrishnan and H. N. Nagaraja. *A first course in order statistics*, New York, Wiley, 1992.

[2] F. Caeiro and M. I. Gomes. Semi-parametric tail inference through probability-weighted moments. *Journal of Statistical Planning and Inference*, 141, 937–950, 2011.

[3] F. Caeiro, A.P. Martins and I.J. Sequeira. Finite sample behaviour of Classical and Quantile Regression Estimators for the Pareto distribution. In T.E. Simos and C. Tsitouras (eds.), AIP Conf. Proc. 1648, 540007, 2015.

[4] F. Caeiro and A. Mateus. Log Probability Weighted Moments Method for Pareto distribution. In Skiadas, C.H. (ed.) 17th ASMDA Conference Proceedings, 211–218 (2017)

[5] J.A. Greenwood, J.M. Landwehr, N.C. Matalas and J.R. Wallis. Probability Weighted Moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, 15, 1049–1054, 1979.

[6] R. E. Quandt. Old and new Methods of Estimation and the Pareto Distribution. *Metrika*, 10, 55–82, 1966.

[7] V. Pareto. Cours d'Economie Politique, Vol. 2, Lausanne, 1897.

Challenges in extremal index estimation

**<u>Dora Prata Gomes</u>**[1,2] **and M. Manuela Neves**[3,4]

Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa[1]
Centro de Matemática e Aplicações da Universidade Nova de Lisboa[2]
Instituto Superior de Agronomia, Universidade de Lisboa[3]
Centro de Estatística e Aplicações da Universidade de Lisboa[4]

E-mail address: *dsrp@fct.unl.pt; manela@isa.ulisboa.pt*

Abstract

The extremal index $\theta \in [0, 1]$ is a primary measure of extremal serial dependence, characterizing the tendency that extremal observations tend to occur in clusters. It has important applications in areas such as hydrology, telecommunications, insurance, finance, and environmental studies, among others. An adequate characterization of the extremal behaviour of a time series in those areas is important because the human and economic impact that those clusters can produce. Several estimators of θ have appeared in the literature, among which we will refer to those based on the blocks method and in the runs method. However those estimators depend on two tuning parameters to be chosen by the statistician: a threshold sequence and a cluster identification scheme parameter (such as a block length). Under certain additional conditions we will discuss an estimator that only depends on the block length. Those conditions and an heuristic procedure for obtaining the estimate will be discussed.

Keywords: blocks method, clusters of extremes, extremal index, stationary time series.

———⊙———
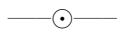
Let $X_1, X_2, \ldots$ be a strictly stationary process with marginal distribution function F and $M_n = max\{X_1, ..., X_n\}$. Let be $\widehat{X_1}, \widehat{X_2}, \ldots$ the independent sequence of variables with the same marginal distribution, F, and $\widehat{M_n} = max\{\widehat{X_1}, \ldots, \widehat{X_n}\}$. Under appropriate mixing conditions, if there exists a nondegenerate limiting distribution G for the variable $(\widehat{M_n} - b_n)/a_n$ with some sequences of constants $\{a_n > 0\}$ and $\{b_n\}$, then the normalized maximum $(M_n - b_n)/a_n$ of the dependent series also has a nondegenerate

limit distribution G^*, and they are related to each other by $G^* = G^\theta$, where θ is the extremal index ([6]). The problem of estimating the extremal index based on a finite stretch from a time series is well-known in literature, some references are [1], [3], [4], [7] and [9], among others. The most common methods of estimation include the blocks methods, the runs method, and the inter-exceedance method. We focus on blocks method in order to demonstrate the effect of the threshold u_n and the block size choice r_n much smaller than n in the block method estimates. The blocks estimator is defined as the reciprocal of average number of exceedances of the threshold u_n per block, among blocks with at least one exceedance. Assuming $r_n[1 - F(u_n)] \to 0$ but $n[1 - F(u_n)] \to \infty$ as $n \to \infty$, and under certain mixing conditions, this estimator is has been shown to be consistent and asymptotically normal, see [5] and [8].

Under some additional conditions on the behaviour of the extremes of the time series we will study an estimator, introduced in [2] who suggested to define the threshold inside each block such as a previous condition is verified. Now the estimates depend essentially on the block size r_n.

Calculations in certain models and a finite-sample Monte Carlo simulation study show that the procedure proposed performs quite well. An application in the hydrological domain is shown.

Acknowledgements

References

[1]Berghaus, B. and Bucher, A. (2018) Weak convergence of a pseudo maximum likelihood estimator for the extremal index. *The Annals of Statistics* **46** (5): 2307–2335.

[2]Canto e Castro, L. (1992) Estudo de um método de estimação do indice extremal. I Congresso Ibero-Americano de Estadística e Investigación Operativa.

[3]Ferro, A. T. and Segers, J. (2003) Inference for Clusters of Extreme Values. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **65** (2): 545–556.

[4]Hsing, T. (1993) Extremal index estimation for a weakly dependent stationary sequence. *The Annals of Statistics* **21** (4) :2043–2071.

[5]Hsing, T. (1991) Estimating the parameters of rare events. *Stochastic Processes and their Applications* **37** (1) :117–139.

[6]Leadbetter, M. R.(1983). *Extremes and local dependence in stationary sequences.* Z. Wahrscheinlichkeitstheor. Verw. Geb. **65** (2) :291–306.

[7]Northrop, P. N. (2005) An efficient semiparametric maxima estimator of the extremal index. *Extremes* **18** (4): 585–603.

[8]Novak, S. and Weissman, I. (1998) On Blocks and runs estimators of the extremal index, *J. Statist. Plann. Inference* **66** (2): 281–288.

[9]Smith, R. L. and Weissman, I. (1994) Estimating the extremal index. *Journal of the Royal Statistical Society. Series B (Methodological)* **56** (3): 515–528.

Organized Session 4

Computational Mathematics and Statistics and its Applications

Organizer: Filomena Teodoro (Portugal)

On the distribution of linear combinations of chi-square random variables

Carlos A. Coelho[1,2]

Departamento de Matemática, Faculdade de Ciências e Tecnologia,
Universidade Nova de Lisboa, Portugal.[1]
CMA-FCT/UNL–Centro de Matemática e Aplicações[2]

E-mail address: *cmac@fct.unl.pt*

Abstract

The distribution of linear combinations of independent chi-square random variables is intimately related with the distribution of quadratic forms in normal random variables [1,6–11,13,14] and thus it also appears as the limit distribution of quadratic forms in non-normal random variables. As such, this distribution has been studied by many authors [2,4–15]. However, there is still much room left for improvement, since while some simpler approximations do not yield sufficiently good results, other approximations which show a better performance are sometimes too complicated to be implemented in practical terms.

In this paper the exact distribution of linear combinations of independent chi-square random variables is obtained, for some particular cases, in a closed finite highly manageable form, while for the general case a near-exact approximation [3] is obtained, which is able to yield very manageable and well-performing approximations.

Keywords: Characteristic function, Gamma distribution, Generalized integer gamma distribution, Generalized near-integer gamma distribution, Mixtures.

—⊙—

Acknowledgements

References

[1]Baksalary, J. K., Hauke, J., Styan, G. P. H. (1994). On some distributional properties of quadratic forms in normal variables and on some associated matrix partial orderings. In *Multivariate Analysis and its Applications* (T. W. Anderson, K. T. Fang & I. Olkin, eds.), *IMS Lecture Notes-Monograph Series*, Institute of Mathematical Statistics, Hayward, California, vol. 24, 111-121.

[2]Castaño-Martínez, A., López-Blázquez, F. (2005). Distribution of a sum of weighted noncentral chi-square variables. *Test*, 14, 397-415.

[3]Coelho, C.A. (2004). The generalized near-integer Gamma distribution: a basis for 'near-exact' approximations to the distribution of statistics which are the product of an odd number of independent Beta random variables. *Journal of Multivariate Analysis 89*, 191–218.

[4]Davis, A. W. (1977). A differential equation approach to linear combinations of independent chi-squares. *J. Amer. Statist. Assoc.*, 72, 212-214.

[5]Gabler, S., Wolff, C. (1987). A quick and easy approximation to the distribution of a sum of weighted chi-square variables. *Statist. Papers*, 28, 317-325.

[6]Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48, 419-426.

[7]Jensen, D. R., Solomon, H. (1972). A Gaussian approximation to the distribution of a definite quadratic form. *J. Amer. Statist. Assoc.*, 67, 898-902.

[8]Johnson, N. L., Kotz, S. (1968). Tables of distributions of positive definite quadratic forms in central normal variables. *Sankhya*, Ser. B, 30, 303-314.

[9]Kotz, S., Johnson, N. L. (1967). Series representations of distributions of quadratic forms in normal variables. I. central case. *Ann. Math. Statist.*, 38, 823-837.

[10]Kotz, S., Johnson, N. L. (1967). Series representations of distributions of quadratic forms in normal variables. II. non-central case. *Ann. Math. Statist.*, 38, 838-848.

[11]Lu, Z.-H. (2006). The numerical evaluation of the probability density function of a quadratic form in normal variables. *Comput. Statist. Data Anal.*, 51, 1986-1996.

[12]Moschopoulos, P. G., Canada, W. B. (1984). The distribution function of a linear combination of chi-squares. *Comput.Math. Appl.*, 10, 383-386.

[13]Robbins, H. (1948). The distribution of a definite quadratic form. *Ann. Math. Statist.*, 19, 266-270.

[14]Robbins, H., Pitman, E. J. G. (1949). Application of the method of mixtures to quadratic forms in normal variates. *Ann. Math. Statist.*, 20, 552-560.

[15]Solomon, H., Stephens, M. A. (1977). Distribution of a sum of weighted chi-square variables. [*J. Amer. Statist. Assoc.*, 72, 881-885.

Three-dimensional velocity field for blood flow using the power-law viscosity function

Fernando Carapau[1,2] **Paulo Correia**[1,2] **and Ricardo Conceição**[3,4]

University of Évora, Mathematics Departament[1]
CIMA-UÉ[2]
*University of Évora*[3]
Renewable Energies Chair and IES[4]

E-mail address: *flc@uevora.pt; pcorreia@uevora.pt; rfc@uevora.pt*

Abstract

The three-dimensional model associated with blood flow in vessels with viscosity depending on shear-rate, e.g., power-law type, is a complex model to obtaining computational implementation, which in many important situations reveals impracticable. In order to simplify the three-dimensional model, and as an alternative to classic one-dimensional models, we will use the Cosserat theory related fluid dynamics to approximate the three-dimensional velocity field, and thus obtain a one-dimensional system. Therefore, this system consists on an ordinary or partial differential equation depending only on time and on a single spatial variable, the flow axis. From this reduce system, we obtain the unsteady equation for the mean pressure gradient depending on the volume flow rate, Womersley number and the flow index over a finite section of the tube geometry. Attention is focused on some numerical simulations for constant and non-constant mean pressure gradient using a Runge-Kutta method and on the analysis of perturbed flows. In particular, given a specific data we can get information about the volume flow rate and consequently we can illustrate the three-dimensional velocity field on the constant circular cross-section of the tube. Moreover, we compare the three-dimensional exact solution for steady volume flow rate with the corresponding one-dimensional solution obtained by the Cosserat theory.

Keywords: Cosserat Theory, Blood Flow, Shear-Thinning Fluid, One-Dimensional Model, Power-Law Model.
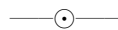
——⊙——

Acknowledgements

References

[1] Caulk, D.A. and Naghdi, P.M. (1987) Axisymmetric motion of a viscous fluid inside a slender surface of revolution, *Journal of Applied Mechanics* **54(1)**: 190–196.

[2] Carapau, F. and Sequeira, A. (2006) 1D Models for Blood Flow in Small Vessels Using the Cosserat Theory, *WSEAS Transactions on Mathematics* **5(1)**, 54–62.

[3] Bird, B.R., Armstrong, R.C. and Hassager, O. (1987) *Dynamics of Polymeric Liquids*, John Wiley & Sons Edition, vol. 1, no. 2.

Water meters inaccuracies registrations: A first approach of a portuguese case study

<u>Marina A. P. Andrade</u>[1], M. Filomena Teodoro[2,3], Sérgio Fernandes[4] and Nelson Carriço[5]

ISCTE-IUL/ISTAR, Lisboa, Portugal[1]
CINAV, Naval Academy, Portuguese Navy, Almada, Portugal[2]
CEMAT, Instituto Superior Técnico, Lisboa, Portugal[3]
Setúbal School of Technology, Polytechnic Inst. Setúbal, Setúbal, Portugal[4]
Barreiro School of Technology, Polytechnic Inst. Setúbal, Barreiro, Portugal[5]

E-mail address: *marina.andrade@iscte.pt; maria.alves.teodoro@marinha.pt*

Abstract

This manuscript deals with the case study – Evaluating Water Meters Performance – as proposed by a Portuguese water utility for the 140th European Study Group with Industry (ESGI140), which took place in Barreiro School of Technology, Polytechnic Institute of Setúbal (ESTBarreiro/IPS).

The objective is to evaluate water meters performance using historical data, knowing that, like any other mechanical device, the water meters suffer a deterioration with time and use, losing accuracy throughout its cycle of use. We intend to approach a problem capable of identifying anomalies on water consumption pattern. The insight gained with this study should help the water utility to identify malfunctioning water meters.

Keywords: arima, water comsumption modeling, water meter performance.

——⊙——
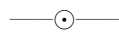
This work starts with a description of the problem, the data provided and some possible statistical approaches, from elementary to intermediate level solutions. The data collected were registered in hours and in months. Being intended it is possible to perform different analysis according with the pretended data registration. Apart from that, for different consumers, the company has different registered data sets, depending on the type of water meters installed and on the period of registered time.

Therefore, first we have analyzed graphically the comparison of the registered values of two distinct water meters (id's 1440 and 2180) in an hourly period of time data comparison, and subsequently, by aggregation of data, the measurements were compared in a daily registration, see Figs. 1 and 2. Two peaks are possible to identify in distinct time periods. These irregularities are discussed and since the water meters are in a serial disposition the measurement difference of data is used in the analysis.
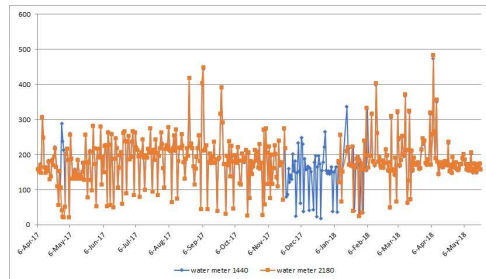
Fig. 1. Water meters series for both 1440 and 2180.
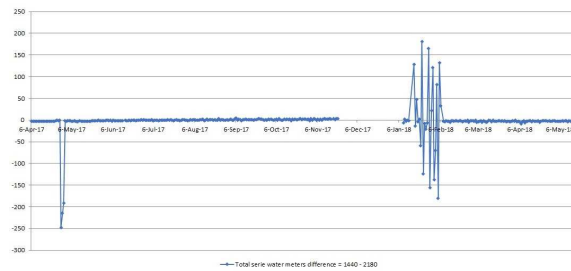
Fig. 2. Water meters difference series.

Another analysis here presented is modeling of a monthly recorded data set provided from the company named water meter 4773. The consumption has a monthly registration, it seems important to consider external influence factors as seasons (winter/summer) or weather variables that may influence the customers needs. In a first approach to model our data, we intend to apply the usual techniques in time series domain such as ARIMA approach [1] or econometric methods [2].

Acknowledgements

References

[1]Gonçalves, E., Lopes, N. M., (2008) Séries temporais – Modelações lineares e não lineares, 2nd ed., Lisboa, Sociedade Portuguesa de Estatística.

[2]Hill, R. C. and et al., (2018) Principles od Econometrics, 5th ed., New Jersey, Wiley.

Inferences on multivariate normal covariance matrices based on synthetic data

Ricardo Moura

CINAV, Marinha Portuguesa, Portugal
CMA, Faculty of Sciences and Technology, Nova University of Lisbon, Portugal

E-mail address: *pinto.moura@marinha.pt; rp.moura@fct.unl.pt*

Abstract

Methods of creating synthetic datasets are being used more and more often by statistical agencies to protect the confidentiality of respondents, while guarantying that its information is being released and that it can be analyzed. Some of these agencies are also demanding that only one single imputation version of the original data may be released due to high risk of disclosure. In the literature are already available exact procedures to draw inference for the mean of a multivariate normal model and the regression coefficients matrix of a multivariate linear regression model, when the single imputed data is generated via Plug-in Sampling and Posterior Predictive Sampling. The objective of this work is to also make available inferential procedures to test different aspects of the structure of the covariance matrix in the referred models, such as the generalized variance and the sphericity test. These inferential procedures will rely on deriving the distribution of test statistics which are somewhat similar to the usual test statistics used when one analyzes the original data. By providing procedures for the multivariate normal covariance matrix, one is completing the task of allowing the general public to analyze the released data thoroughly, when in the past was impossible, when only one synthetic version was released. Simulation studies are carried out to show that the derived procedures perform as expected for each one of the tests.

Keywords: Statistical Disclosure Control, Covariance Matrix, Multivariate Normal Distribution, Multivariate Linear Regression, Single Imputation.

References

[1] Anderson, T. W., Anderson, T. W., Anderson, T. W., Anderson, T. W., & Mathématicien, E. U. (1958) An introduction to multivariate statistical analysis (Vol. 2, pp. 5–3) New York, Wiley.

[2] Klein, M. and B. Sinha (2015) Inference for singly imputed synthetic data based on posterior predictive sampling under multivariate normal and multiple linear regression models. Sankhya B 77 (2), 293-311.

[3] Klein, M., & Sinha, B. (2015) Likelihood based finite sample inference for singly imputed synthetic data under the multivariate normal and multiple linear regression models. Journal of Privacy and Confidentiality 7(1).

[4] Moura, R., M. Klein, C. Coelho, and B. Sinha (2016) Inference for multivariate regression model based on synthetic data generated using plug-in sampling. Technical Report - Centro de Matemática e Aplicações (CMA) 2/2016 (submitted for publication in Journal of the American Statistical Association).

[5] Moura, R., M. Klein, C. Coelho, and B. Sinha (2017) Inference for Multivariate Regression Model based on synthetic data generated under Fixed-Posterior Predictive Sampling: comparison with Plug-in Sampling. Revstat 15 (2) 155-186.

Students performance and individual profile, a generalized linear models approach

M. Filomena Teodoro[1,2], Alcindo Gomes[1] and José Maia Martins[1]

CINAV, Naval Academy, Portuguese Navy, Base Naval de Lisboa, Alfeite, 1900-001 Almada, Portugal[1]
CEMAT, Instituto Superior Técnico, Lisbon University, Av. Rovisco Pais, 1,1046-01 Lisboa, Portugal[2]

E-mail address: *maria.alves.teodoro@marinha.pt, mteodoro64@gmail.com*

Abstract

The Portuguese Naval Academy has a significant failure rate, losing about half of the admitted individuals during the course. This work aims to understand the causes associated to low income, through the analysis of the data of the individuals from the application to the end of the course, in order to identify the characteristics of individuals with greater and less likely to succeed in the Naval Academy. This work was started in [1]. In an initial stage, some techniques of descriptive statistics of data analysis were used [4]. The first step of this analysis, the candidates are analyzed independently, and the admitted students and the finalists are analyzed together comparing the variables at the beginning and at the end of the courses. Simple statistical inference techniques were used [4], namely confidence intervals, parametric tests, contingency tables. Here we extend such analysis using some inference techniques, namely generalized linear models [2,3]. The study evidences greater success for individuals entering in Naval Academy with better grades and for individuals taking notice of the application competition over the internet.

Keywords: Indicators of academic success, academic performance, optimization of Education, generalized linear models.
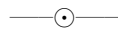
——⊙——

References

[1]Gomes, A. D., (2016) Rentabilização de Capital Humano, Master Thesis, Almada, Portuguese Naval Academy.

[2]McCullagh, P. and Nelder, J. A., (1989) Generalized Linear Models, Londres, Chapman & Hall.

[3]Tamhane, A. C. and Dunlop, D. D. (2000) Statistics and Data Analysis: from Elementary to Intermediate, New Jersey, Prentice Hall.

[4]Turkman, M. A. Silva, G., (2000) Modelos Lineares Generalizados da teoria a prática, Lisboa, Portuguese Society of Statistics.

Organized Session 5

Special functions and applications

Organizers: José Luís Cardoso and Maria das Neves Rebocho (Portugal)

Hölder and Minkowski's inequalities related with q-integrals

José Luis Cardoso[1,2]

University of Trás-os-Montes e Alto Douro[1]
CMAT - Polo da UTAD[2]

E-mail address: jluis@utad.pt

Abstract

The Jackson q-integral and the Jackson-Thomae-Nörlund q-integral are the inverse of the Jackson's $q-$operator D_q and the Hahn's $D_{q,\omega}$ operator, respectively. In [2] is presented a generalization of these operators and the corresponding inverse operator. We show how the Hölder and Minkowski's inequalities related with these operators hold.

Keywords: Jackson $q-$integral, Jackson-Nörlund $(q,\omega)-$integral, q-analogue of the Lebesgue spaces, $q-$analogues.

$$\underline{\qquad}\odot\underline{\qquad}$$
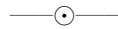
Let $0 < q < 1$ and $\omega \geq 0$. The Jackson's $q-$operator, D_q, is defined by

$$D_q[f](x) := \frac{f(qx) - f(x)}{(q-1)x} , \quad x \neq 0 ,$$

and $D_q[f](0) := f'(0)$, while the Hahn's $D_{q,\omega}$ operator is defined by

$$D_{q,\omega}[f](x) := \frac{f(qx + \omega) - f(x)}{(q-1)x + \omega} , \quad x \neq 0 ,$$

and $D_{q,\omega}[f](\omega_0) := f'(\omega_0)$, where $\omega_0 := \omega/(1-q)$.
Under certain restrictions on the function $\beta : I \to I$, (I=interval), the operator

$$D_\beta[f](x) := \frac{f(\beta(x)) - f(x)}{\beta(x) - x} , \quad x \neq s_0 ,$$

and $D_\beta[f](s_0) := f'(s_0)$ where s_0 is a fixed point (unique) of the function β in the interval I, generalizes the previous operators.

References

[1] Cardoso J. L. and Petronilho, J. (2016) Variations around Jackson's quantum operator. *Methods Appl. Anal.* **22** no. 4: 343–358.

[2] Hamza, A. E., Sarhan, A.-S. E., Shehata, E. M. and Aldwoah, K. A. (2015) A general quantum difference calculus. *Advances in Difference Equations* **2015:182**: 1–19.

Product formulas, generalized convolutions and integral transforms

Rúben Sousa[1]

CMUP, Departamento de Matemática, Faculdade de Ciências,
Universidade do Porto[1]

E-mail address: *rubensousa@fc.up.pt*

Abstract

It is well-known that the ordinary convolution is closely related with the Fourier transform. It is therefore natural to ask: for other important integral transforms, can we define generalized convolution operators having analogous properties? Actually, the answer depends on the existence of a product formula for the kernel of the integral transform. In this talk, I will describe the general connection between product formulas, generalized convolutions and integral transforms. I will explain the role of such generalized convolutions in answering to a natural question in probability theory: if it is possible to define families of Lévy-like processes which are generated by diffusion process other than the Brownian motion. Emphasis will be given to convolution structures related with families of orthogonal polynomials (such as the Jacobi and the Bessel polynomials). I will report on recent progress in constructing the product formula and convolution associated with the index Whittaker transform.

Keywords: Generalized convolutions, Integral transforms, Lévy processes, Product formula, Whittaker function.

References

[1] Sousa R., Guerra M. and Yakubovich S.(2018) On the product formula and convolution associated with the index Whittaker transform, arXiv preprint arXiv:1802.06657

[2] Sousa R., Guerra M. and Yakubovich S. (2018) Lévy processes with respect to the index Whittaker convolution, arXiv preprint arXiv:1805.03051

[3] Berezansky Y. M., Kalyuzhnyi A. A., Harmonic Analysis in Hypercomplex Sys- tems, Kluwer Academic Publishers, Dordrecht, 1998.

[4] Rentzsch C. and Voit M., Lévy Processes on Commutative Hypergroups, in Proba- bility on Algebraic Structures: AMS Special Session on Probability on Algebraic Structures, American Mathematical Society, Providence RI, pp. 83–105 (2000).

Orthogonal polynomials via polynomial mappings and some applications

<u>M. N. de Jesus</u>[1]

CI & DETS, Área Científica de Matemática, Instituto Politécnico de Viseu[1]

E-mail address: *mnasce@estv.ipv.pt*

Abstract

We consider orthogonal polynomials (OP) via polynomial mappings in the framework of the semiclassical class. In particular, we analyze in detail cubic transformations for semiclassical OP of class at most 2, giving new examples of semiclassical OP of class 2 and providing integral representations for the regular functionals with respect to which these new semiclassical families are orthogonal. Furthermore, we derive in an unified way old and new properties concerning the sieved ultraspherical polynomials of the first and of the second kind, introduced by W. Al-Salam, W. R. Allaway, and R. Askey. This allow us to derive infinitely many examples of semiclassical functionals such that the pair of polynomials appearing in the corresponding canonical Pearson-type distributional differential equation is non-admissible. Finally, we present an application involving an electrostatic model. The results presented are joint work with K. Castillo and J. Petronilho.

Keywords: orthogonal polynomials, semiclassical class, electrostatic model.

References

[1] Castillo K., de Jesus M.N. and Petronilho J. (2018) On semiclassical orthogonal polynomials via polynomial mappings II: Sieved ultraspherical polynomials revisited. (Submitted.)

[2] Castillo K., de Jesus M.N. and Petronilho J. (2017) On semiclassical orthogonal polynomials via polynomial mappings. *J. Math. Anal. Appl.* 455: 1801–1821.

[3] de Jesus M.N. and Petronilho J. (2010) On orthogonal polynomials obtained via polynomial mappings. *J. Approx. Theory* 162: 2243–2277.

On Laguerre-Hahn orthogonal polynomials on the real line

Maria das Neves Rebocho[12]

Departamento de Matemática, Univ. da Beira Interior[1]
CMUC, Univ. Coimbra[2]

E-mail address: *mneves@ubi.pt*

Abstract

Laguerre-Hahn orthogonal polynomials on the real line were introduced in [2], within the framework of theory of distributions and moment functionals, and, since then, studied by several authors within a vast list of problems, for instance, related to Hermite-âĂŞPadé Approximation [4], measure modifications, and perturbations of orthogonal polynomials, [1,5,7].

A fundamental property of Laguerre Hahn orthogonal polynomials is the Riccati type differential equation for the corresponding Stieltjes function,

$$AS' = BS^2 + CS + D\,, \quad A \neq 0\,, \tag{4}$$

where A, B, C, D are polynomials. From (4), the Laguerre-Hahn orthogonal polynomials are generalizations of semi-classical orthogonal polynomials, as the later ones appear upon the specification $B \equiv 0$. Furthermore, upon the bounds $\deg(A) \leq 2, \deg(C) = 1$, we recover the classical orthogonal polynomials - the Hermite, Laguerre and Jacobi polynomials.

In this talk I will survey some recent results on Laguerre-Hahn orthogonal polynomials, focusing on the difference-differential equations for the systems of polynomials and (recent) connections with discrete Painlevé equations [3,6].

Keywords: orthogonal polynomials, Stieltjes function, discrete Painlevé equations.

References

[1]Dehesa J.S., Marcellán F. and Ronveaux A. (1990) On orthogonal polynomials with perturbed recurrence relations. *J Comput Appl Math.* 30: 203âĂŞ-212.

[2]Dini J., Sur les formes linéaires et les polynômes orthogonaux de LaguerreâĂŞHahn [doctoral dissertation]. Paris (France): University Pierre et Marie Curie; 1988.

[3]Filipuk G. and Rebocho M.N. (2016) Discrete Painlevé Equations for Recurrence Coefficients of Laguerre-Hahn Orthogonal Polynomials of Class One *Integral Transforms Spec. Funct.* 27: 548âĂŞ-565.

[4]Magnus A.P., Riccati acceleration of the Jacobi continued fractions and Laguerre-Hahn polynomials. In: Werner H, Bunger HT, editors. Padé Approximation and its Applications. Lect. Notes in Math. 1071, Berlin: Springer Verlag; 1984. pp. 213-230.

[5]Ronveaux A. and Van Assche W. (1996) Upward extension of the Jacobi matrix for orthogonal polynomials, *J. Approx Theory* 86: 335-âĂŞ357.

[6]Van Assche W, Discrete Painlevé equations for recurrence coefficients of orthogonal polynomials. In: Elaydi S, Cushing J, Lasser R, Ruffing A, Papageorgiou V, Van Assche W, editors. Difference equations, special functions and orthogonal polynomials. Hackensack, NJ: World Scientific; 2007. pp. 687-725.

[7]Zhedanov A. (1997) Rational spectral transformations and orthogonal polynomials. *J. Comput. Appl. Math.* 85: 67-âĂŞ86.

Explicit relations between cumulants and moments

<u>Patrícia Antunes</u>[1], **Sandra S. Ferreira**[1,2], **Célia Nunes**[1,2],
Dário Ferreira[1,2] **and João T. Mexia**[3]

*Center of Mathematics and Applications, University of Beira Interior,
Covilhã, Portugal*[1]
Department of Mathematics of University Beira Interior[2]
*Center of Mathematics and its Applications, Faculty of Science and
Technology, New University of Lisbon, Monte da Caparica, Portugal*[3]

E-mail address: *antunes.patricia03@gmail.com; sandraf@ubi.pt;
celian@ubi.pt;dario@ubi.pt;jtm@fct.unl.pt*

Abstract

This communication is dedicated to the study of higher order probability distribution cumulants, which corresponds to sets of quantities that provide an alternative to the distribution moments. Well known higher order cumulants are variance (second order), skewness (third order), measuring asymmetry, and kurtosis (fourth order), measuring 'peakiness' of the probability distribution. Of particular interest to us, is the fourth cumulant, for which we will use two models to estimate the skewness and kurtosis coefficients. Explicit relations between cumulants and moments are given, as well as a numerical example.

Keywords: Central Moments, Cumulants, Cumulant Generating Function, Kurtosis, Moments.
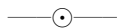
—⊙—

Acknowledgements

References

[1]Akhiezer, N. I. (1965), The Classical Moment Problem, *N. Kemme, Transl.*, Hafner, New York.

[2]Balakrishnan, N., Johnson,N. L., and Kotz, S. (1998), A note on relationships between moments, central moments and cumulants from multivariate distributions, *Statistics & probability letters*, **39**: 49–54, Elsevier.

[3]Dodge Y. and Valentin R. (1999), The Complications of the Fourth Central Moment, *The American Statistician*, **53** (3): 267–269.

[4]Harvey, J. R., (1972), On Expressing Moments in Terms of Cumulants and Vice Versa. *The American Statistician*, **26** (4): 38–39.

[5]Kendall, M. and Stuart, A. (1958). *The Advanced Theory of Statistics.* Charles Griffin & Company.

[6]Landau, H. J. (1987), Moments in Mathematics, Proceedings of Symposia in Applied Mathematics, **37**, *Amer. Math. Soc.*, Providence, RI.

[7]Ruppert, D. (1987) What is Kurtosis? An Influence Function Approach, *The American Statistician*, **41**: 1–5.

[8]Smith, P. J. (1995), A Recursive Formulation of the Old Problem of Obtaining Moments from Cumulants and Vice Versa. *The American Statistician*, **49** (2): 217–218.

Organized Session 6

Mathematics Education and Society

Organizers: Carla Santos, Cristina Dias and Ana Caballero (Portugal)

On the pedagogical practices in teaching and learning Mathematics presented in CNaPPES

Dias, C.[1,2], **Santos, C.**[1,3], **Varadinov, M.**[3] **and Romacho, J.**[3]

CMA -Center of Mathematics and its Applications, New University of Lisbon, Portugal[1]
College of Technology and Management, Polytechnical Institute of Portalegre, Portugal[2]
Department of Mathematics and Physical Sciences, Polytechnical Institute of Beja, Portugal[3]

E-mail address: *cpsilvadias@gmail.com; carla.santos@ipbeja.pt; mjvaradinov@ipportalegre.pt; jromacho@ipportalegre.pt*

Abstract

The National Congress of Pedagogical Practices in Higher Education (CNaPPES), which is held annually since 2014, is one of the rare opportunities in Portugal for exchanging experiences, discussing pedagogical practices and learning with peers among polytechnic higher education teachers and university. Currently in its 6th edition, CNAPPES has registered an increasing number of participants, as well as higher education institutions represented.

Considering the need to adapt the teaching-learning processes to the challenges of 21st century society, many of the papers, presented in the different editions of the CNaPPES, reflect a growing concern with the diversification and innovation of pedagogical practices in higher education.

In this article we identify the pedagogical practices proposed in CNaPPES, for teaching and learning Mathematics, which gave rise to works published in the conference proceeding books, classifying these pedagogical practices by area, provenance and type, and reflecting on the success achieved by them and their transferability to other contexts.

Keywords: CNaPPES, Higher education, Mathematics, Pedagogical practices.

Acknowledgements

References

[1]Beetham, H, McGill, L and Littlejohn, A (2009) 'Thriving in the 21st century: Learn-
    ing Literacies for the Digital Age' in Attwell, G and Hughes, J (2010) Pedagogic
    approaches to using technology for learning: Literature review, Wales: Ponty-
    dysgu: Lifelong Learning UK (LLUK), `http://webarchive.nationalarchives.`
    `gov.uk/20110414152025/http:/www.lluk.org/wp-content/uploads/2011/01/`
    `Pedagogical-appraches-for-using-technology-literature-reviewjanuary-11-`
    `FINAL.pdf`

[2]Cook-Sather, A. (2011). Lessons in higher education: Five pedagogical practices that
    promote active learning for faculty and students.Journal of Faculty Development, 25
    (3), 33-39.

[3]Gordon, N (2014) Flexible Pedagogies: Technology-Enhanced Learning, York, England:
    The Higher Education Academy (HEA), `https://www.heacademy.ac.uk/system/`
    `files/resources/tel_report_0`.

[4]Maloney, E (2007) 'What Web 2.0 can teach us about learning', The Chronicle of Higher
    Education, 53(18):B26.

How is the employment situation of mathematics teachers?

Rosa Gómez, <u>Ana Caballero</u> and Eloísa Guerrero

University of Extremadura
Faculty of Education

E-mail address: *rosagomez@unex.es; acabcar@unex.es; eloisa@unex.es*

Abstract

This paper presents a part of the results obtained in a broader study on sources of stress and emotions in high school math teachers. It carried out an investigation with 253 teachers of mathematics, who worked in public centers, private and concerted in the province of Badajoz (Extremadura, Spain). The instrument used was a self-elaboration questionnaire, which underwent a validation process (Cronbach coefficient =. 95; coefficient of Sperman Brown. 89), from which we obtained the 9 factors that make up our questionnaire. This instrument is composed of 66 items and measures, through a scale type Likert, the level of stress that gives teachers the various situations that are proposed, but only in the event that they actually manifest in their day to day work. On the other hand, teachers also had to indicate whether these situations produced joy, sadness, anger or Fear (Ekman, 1992) or, if on the contrary did not produce any emotion. In this work we will show the results that describe the real situation of the teachers of secondary mathematics. Some of the results obtained are that 97.6% of teachers say that there is a low social recognition of teaching; 99.63% of students ' assessments are low; A 99.26% is faced with a lack of interest and motivation of students; 94.38% say that their students have difficulty understanding mathematical content and 79.7% cannot transfer mathematical contents to real life. It will also show what are the main sources of stress of the teachers of the sample and what emotions are those that manifest to these. Based on the results obtained in this study, we believe it is necessary to develop specific programmes for the prevention and control of the stress of teachers, as well as to adjust the initial and permanent teacher training, in order to prepare it for the Context in which it will develop and optimize the teaching and learning process of mathematics.

Keywords: stress, emotions, teachers, mathematics, teaching.

Acknowledgements

References

[1]Ekman, P. (1992). An argument for basic emotion. Cognition and emotion, 6, 169–200.

Interpretation and communication of statistical data in primary education

Luis Maya and <u>Ana Caballero</u>

University of Extremadura
Faculty of Education

E-mail address: *malloso@hotmail.com; acabcar@unex.es*

Abstract

The mathematical research about statistics has grown up in the last years. Besides, the majority of the current researches related to statistics are being assigned to the training of teachers, also to the students that are in Secondary Education. The curriculum in Primary Education, in the contents block of statistics and probability, gives more importance to the comprehensive and communication aspect of statistic data about the students' surroundings. This article analysis the situation about how we are using the communicational and interpretative variable in the classes in Primary.

The situation about statistical communication in Primary Education is having real problems because of the great differences between the curriculum design, the teaching activity, and what books offer; this tenacious idea in Spain is also noticed in many other countries. Besides, the international proofs support that the statical development is getting increasingly more necessary and due to this fact, they settle the guidelines of the future statical work.

Keywords: basic statistic, statistical communication, statistical basic language.

References

[1] Ainley, J., Monteiro, C. (2008). Comparing curricular approaches for statistics in primary school in England and Brazil: A focus on graphing. In C.Batanero, G. Burril, C.Reading, and A. Rossman (Eds). *Joint ICMI/IASE Study: Teaching Statistic in School Mathematics. Challenges for Reaching and Teacher Education.*

[2] Mullis, I., Martin, M. (Eds) (2018). *TIMSS 2019. Marcos de la Evaluación* (S. Montoya y J. Rodríguez, Trans). Instituto Nacional de Evaluación Educativa, Ministerio de Educación y Formación Profesional. NIPO línea 030-18-150-06

The use of computers in Mathematics classes of secondary education

Adelaide Proença[1,2], Luis Manuel Casas García[3,4] and Ana Paula Canavarro[5,6]

School of Education and Social Sciences, Polytechnic Institute of Portalegre, Portugal[1]
C3i - Polytechnic Institute of Portalegre, Portugal[2]
Faculty of Education, University of Extremadura, Spain[3]
Department of Didactics of Experimental Sciences and Mathematics[4]
School of Social Sciences, University of Évora, Portugal[5]
Education Research and Development Unit, Institute of Education of University of Lisbon[5]

E-mail address: *adelaideproenca@ipportalegre.pt; luisma@unex.es; apc@uevora.pt*

Abstract

The research question "How do secondary education Mathematics teachers integrate technology, in particular the computer, into their teaching practices, in Portugal?" points out the problem studied in this communication: The interest in this issue focuses on the role that the potentials of technology, especially the computer, can have in the development of innovative and effective teaching practices that allow the student to actively acquire skills, learning how to question, make and test conjectures, discover, argue and reason about different abstract concepts, relating them to the physical and social reality. (Miranda, 2007; Ricoy & Couto, 2012). Generally, this study intends, on the one hand, to characterise the situation regarding the integration of the computer in Mathematics secondary education classrooms (Mathematics A, Mathematics B and Mathematics Applied to Social Sciences), on the other hand, to identify factors that can influence the teaching practices of the teachers using this technological tool (Barrantes, Casas & Luengo, 2011; Ross, Morrison & Lowther, 2010).

In methodological terms, the study followed a mixed approach, with a strong quantitative perspective, using a non-experimental or descriptive cross-sectional plan. The data collection process was carried out through a questionnaire, sent online, to all School Clusters/Schools in the country

where there were teachers who taught Mathematics' subjects in secondary education, in the academic year 2015/2016. The questionnaire consists of five parts, comprising 31 questions (closed, open and semi-open questions).

Keywords: Teaching practices, Mathematics, Computer.

Some results

119 teachers voluntarily answered the questionnaire. The first results indicate that the computer is not yet fully integrated in Mathematics' classes, in secondary education. Its use by teachers stands out more in the accomplishment of certain professional tasks such as lesson planning, construction of teaching materials, preparation of worksheets and evaluation tests than specifically in the accomplishment of teaching activities with students.

Of all Mathematics' themes taught in secondary education, the study of functions and geometry are those in which teachers most frequently use the computer. In Statistics, the themes with the highest levels of computer use are linear correlation and regression and data organization. The aspects that teachers consider most important in relation to the role that the computer can have in Mathematics' subjects are related to the possibility of accessing more information, working with large data sets, working with real data and simulating phenomena and experiences.

When asked about the conditions they consider relevant for an effective use of the computer in the classroom, a significant percentage of teachers identify aspects related to the equipment available in the school where they teach and also curriculum aspects related to the subjects' programs, namely the fact that programs are too long and the contents are not suitable for an effective use of the computer.

References

[1]Barrantes, G.; Casas, L. M., Luengo, R. (2011) Obstáculos percibidos para la integración de las TIC por los professores de Infantil e Primaria en Extremadura. Pixel-Bit. Revista de Medios y Educación, 39: 83-94.

[2]Miranda, G. L. (2007) Limites e possibilidades das TIC na educação. Sísifo. Revista de Ciências da Educação, 3, 41-50.

[3]Ricoy, M. C. and Couto, M. J. V. S. (2012). Os recursos educativos e a utilização das TIC no Ensino Secundário na Matemática. Revista Portuguesa de Educação, 25 (2): 241-262.

[4]Ross, S. M., Morrison, G. R., and Lowther, D. L. (2010) Educational Technology Research Past and Present: Balancing Rigor and Relevance to Impact School Learning. Contemporary Educational Technology, 1 (1): 17-35.

Income, consumption and saving of Portuguese households in numbers

<u>João Romacho</u>[1,2], Carla Santos[3], Cristina Dias[1,4] and Maria José Varadinov[1,2]

Polytechnic Institute of Portalegre, Portugal[1]
Interdisciplinary Coordination for Research and Innovation (C3i), Portalegre, Portugal[2]
Polytechnic Institute of Beja, Portugal[3]
CMA- Center for Mathematics and Applications of the Nova University of Lisbon[4]

E-mail address: *jromacho@ipportalegre.pt; carla.santos@ipbeja.pt; cpsd@ipportalegre.pt; mjvaradinov@ipportalegre.pt*

Abstract

In this work the income, consumption and saving of Portuguese households during the last two decades (1995 to 2015) are analyzed. Through the use of PORDATA portal data different variables are used for each of the three aggregates.

The results show that the majority of households live with a low income (below e 10,000 per year), which justifies the great inequality in the distribution of income in Portuguese households (in 2015, the richest households earn six times more than the poor) and the high risk of poverty (almost 50%). This latter indicator is only attenuated to 20% due to the important role of Social Security and other social support institutions.

It is also verified that more than 50% of households consumption is concentrated in three sectors: Food, beverages and tobacco; Housing, water, electricity, gas and fuels; and, Transport and communications

Finally, the results show that the saving rate has been decreasing in the last two decades, standing at 5% of disposable income in 2015. Although this decrease in saving in relative terms, it shows an increase in absolute terms of the amount invested in term deposits from 2006. This is a reflection of a concern of the households to safeguard the future in a period of economic difficulties that the country faced.

Keywords: Portuguese households, Income, Consumption, Saving.

References

[1] Attanasio, O. and G. Weber (2010). Consumption and saving: models of intertemporal allocation and their implications for public policy. Journal of Economic Literature, 48: 693-751.

[2] Bover, O., Schürz, M., Slacalek, J. and Teppa, F. (2016). Eurosystem household finance and consumption survey: Main results on assets, debt, and saving. International Journal of Central Banking, 12(2): 1-13.

[3] Cardoso, F., Farinha, L. and Lameira, R. (2008), "Household wealth in Portugal: revised series", Occasional Paper 1, Banco de Portugal.

[4] Carlos, G., Ribeiro, H., Alves, S., Veloso, C. and Pereira, J. (2017). The determinants of the evolution of family savings in the context of a high leveraged society. Proceedings of 23rd International Scientific Conference on Economic and Social Development, 78-99.

[5] Guerello, C. (2018). Conventional and unconventional monetary policy vs. households income distribution: An empirical analysis for the Euro Area. Journal of International Money and Finance, 12(2): 1-13.

Databases

www.apb.pt

www.bportugal.pt

www.cmvm.pt

http://ec.europa.eu/eurostat

www.ine.pt

www.pordata.pt

Reading and understanding of statistical graphs in secondary students of Badajoz

Ruiz, J.

University of Extremadura

E-mail address: *ruizruizjoseisrael@gmail.com*

Abstract

The ability to interpret the information represented in graphs is a key task within Statistical Education. This study presents different levels of reading and comprehension of statistical graphs obtained by 52 students of a private school of Compulsory Secondary Education, who answered a questionnaire. In addition, it was analyzed whether there were significant differences with respect to the level of reading and comprehension achieved by the students according to the variables sex and the course. Also, if there was any correlation between the level of reading and comprehension and the variables age and academic achievement in Mathematics.

Keywords: Statistical Education, Reading and Understanding of Graphs, Secondary Students, Reading and Understanding Levels.

References

[1] Aoyama, K. (2007). Investigating a hierarchy of studentsťs interpretations of graphs. International Electronic Journal of Mathematics Education, 2(3), 298-318. Retrieved from `http://www.iejme.com/makale_indir/337`

[2] Contreras, J. M., Molina Portillo, E., Batanero, M. D., & Díaz, J. (2017). Construction of a questionnarie to evaluate the critical interpretation of statistical graphs by future teachers. In J.M. Muñoz-Escolano, A. Arnal-Bailera, P.Beltrán-Pellicer, M.L. Callejo & J. Carrillo (Eds.), XXI Symposium of SEIEM (pp. 207-216). Zaragoza: SEIEM. Retrieved from `https:\\dialnet.unirioja.es/servlet/articulo?codigo=6262028`

[3] Grandío, A. (2014). The 2012 PIAAC results in Mathematics and the interpretation of graphs (Masterťs thesis). University of Cantabria, Spain. Retrieved from `https://repositorio.unican.es/xmlui/bitstream/handle/10902/4896/Grand$\%$C3$\%$ADoBarrialAlberto.pdf?sequence=1`

[4] Monteiro, C., & Ainley, J. (2007). Investigating the interpretation of media graphs among student teachers. International Electronic Journal of Mathematics Education, 2(3), 187-207. Retrieved from `http:\\citeseerx.ist.psu.edu/viewdoc/download;jsessionid=AA5CDE080CD04D124999E823CF17CBDC?doi=10.1.1.152.9454&rep=rep1&type=pdf`

[5]Vigo, J. M. (2016). Understanding of statistical graphs by students of Basic Vocational Training (Master's thesis). Retrieved from `www.ugr.es/~batanero/documentos/TFMVigo.pdf`

[6]Wu, Y. (2004). Singapore Secondary School Students' Understanding of Statistical Graphs. 10th Internarional Congress on Mathematics Education (ICME-10). Retrieved from `https:\\iase-web.org/documents/papers/icme10/Yingkang.pdf`

Assessing higher education students' ability to read and interpret statistical tables and graphs

Santos, C.[1,2], **Dias, C.**[1,3], **Varadinov, M.**[3] and **Romacho, J.**[3]

CMA -Center of Mathematics and its Applications, New University of Lisbon, Portugal[1]
Department of Mathematics and Physical Sciences, Polytechnical Institute of Beja, Portugal[2]
College of Technology and Management, Polytechnical Institute of Portalegre, Portugal[3]

E-mail address: *carla.santos@ipbeja.pt ; cpsilvadias@gmail.com; mjvaradinov@ipportalegre.pt; jromacho@ipportalegre.pt*

Abstract

From the very beginning of human history Mathematics has played a crucial role in responding to the needs of humanity, contributing decisively to the amazing evolution that has led to today's society.

In the 21st century society, science and technology evolution has made statistical data an important asset in decision-making and a key source for knowledge. In an increasingly data-driven world new skills became fundamental to all citizens, leading to an evolution of the concept of literacy that underline the skills needed to access, process and communicate information, highlighting the relevance of quantitative and statistical skills.

Nowadays, much of the information that we deal with appears as data expressed in tables and graphs, reason why it is mandatory to have the capacity of critical evaluation, decoding, understanding and use this information.

In Portugal, Statistics education begins during the early years of schooling, however, there are many students who do not acquire the basic skills of Statistics, reaching higher education with difficulties in key aspects, such as reading and interpreting statistical information in tables and graphs.

Based on an activity developed in the classroom, we reflect on the difficulties that a sample of students, from polytechnic higher education, presents in the interpretation of statistical graphs and tables similar to those used in technical-scientific reports or those that can be found every day in the media. The activity was structured considering the aspects that, in the literature, are identified as influencing the understanding of tables and graphs with statistical information.

Keywords: Graphs, Higher education, Statistical data presentation, Statistical literacy, tables.

References

[1]Arteaga, P., Batanero, C., Cañadas, G. & Contreras, J.M.(2011) Las tablas y gráficos estadísticos como objetos culturales, Números 76:55-67.

[2]Batanero, C., Godino, J. D., Vallecillos, A., Green, D. R. & Holmes, P. (1994). Errors and difficulties in understanding statistical concepts. International Journal of Mathematics Education in Science and Technology. 25(4): 527-547

[3]Bertin, J. (1973). Semiologie Graphique. The Hague: Mouton-Gautier. English translation by William Berg & Howard Wainer (1983) and published as Semiology of graphics, Madison, WI: University of Wisconsin Press.

[4]Curcio, F. R. (1989). Developing graph comprehension. Reston, N.C.T.M.

[5]Gal, I. (2000). The numeracy challenge. In: I. Gal (ed.). Adult numeracy development: theory, research, practice. 9-31. Hampton Press.

[6]Garfield, J. (2003). Assessing statistical reasoning. Statistics Education Research Journal. 2: 22-38

[7]Huff, D. (1954) How to Lie with Statistics. Norton, New York. W. W. Norton and Company.

[8]Kosslyn, S.M. (1985). Graphics and Human Information Processing: A review of five books. Journal of the American Statistical Association. 80(391): 499-512.

[9]Shah, P., Freedman, E. G. & Vekiri, I. (2005). The comprehension of quantitative information in graphical displays. In: P. Shah. & A. Miyake (eds.). The Cambridge handbook of visuospatial thinking. 426-476. New York, NY: Cambridge University Press.

[10]Wainer, H. (1992) Understanding graphs and tables. Educational Researcher. 21: 14-23.

[11]Watson, J. (1997). Assessing Statistical Thinking Using the Media. In: I. Gal & J. B. Garfield (eds.). The Assessment Challenge in Statistics Education. 107-121. Amsterdam: IOS Press and The International Statistical Institute.

Teaching mathematical concepts using a videogame prototype

Luis Manuel Soto, Ricardo Luengo, José Luis Torres and <u>Ana Caballero</u>

University of Extremadura
Faculty of Education
Grupo de Investigación Ciberdidact

E-mail address: *luismanuel@unex.es; rluengo@unex.es; jltc@unex.es; acabcar@unex.es*

Abstract

This study presents an educational intervention initiative based on the importance that videogames can have as a teaching method, used responsibly. For this study, we are going to focus on the use of a videogame created by us, to teach mathematical contents in a playful and motivating way.

We used a specific software to develop the videogame, where the characters, scenarios and dialogues were designed, to allow the student to enter into a fantastic world in which they will learn and have fun.

The purpose of the videogame is to teach certain mathematical concepts such as area, length, surfaces, etc., taking advantages of the benefits that these games give to the children. This constructivist methodology helps the student learning mathematics by himself, using videogames. Another of our objectives is checking if when the students use the videogame, they get better results than those who do not use it.

We got a sample composed of 50 students of primary education. The sample was divided into two similar groups: a control group, in which the concepts was taught in a traditional way, and an experimental group, in which the concepts mathematics was taught with our videogame. These two groups will allow us to evaluate and compare the results obtained by children using a traditional methodology and the innovative methodology based of the use of the videogames.

In relation to the results that will be obtained in the study, they will allow teachers new teaching alternatives based on new technologies, updating their way of teaching and adapting it to the requirements of students and the new technological and social demands.

Keywords: Videogames, learning, teaching mathematics, learning with videogames, mathematics.

References

[1] Gisbert, M. (2000). El siglo XXI, hacia la sociedad del conocimiento. En Cabero Almenara, J.; Martínez Sánchez, F. and Salinas Ibáñez, J. Medios audiovisuales y NNTT para la formación en el S. XXI (pp. 277-288). Murcia, España: edutec.

[2] Graetz, J. M. (1981). The origin of spacewar. Creative Computing, 7(8), 56-67.

[3] Green, S. and Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. Journal of Experimental Psychology: Perception and Performance, 6(32), 1465-1478.

[4] Marcano, B. (2008). Juegos serios y entrenamiento en la sociedad digital. Revista Electrónica Teoría de la Educación. Educación y Cultura en la Sociedad de la Información, 9(3): 93-107.

[5] Martín, A., López, E., and González, J. (2013). Reflexiones sobre la Sociedad de la Información y las Tecnologías de la Información y la Comunicación. II Seminario científico Internacional sobre Formación Didáctica con Tecnologías Web, 2, 1-7.

[6] Moncada, J. and Chacón, Y. (2012). El efecto de los videojuegos en variables sociales, psicológicas y fisiológicas en niños y adolescentes. Retos. Nuevas tendencias en Educación física, deporte y recreación, (21), 43-49.

[7] Soto, L (2015). Estudio sobre la aplicación de los videojuegos en la enseñanza de la matemática: elaboración de un prototipo de juego (Trabajo de fin de Máster). Universidad de Extremadura, Badajoz.

Databases

www.apb.pt

www.bportugal.pt

www.cmvm.pt

http://ec.europa.eu/eurostat

www.ine.pt

www.pordata.pt

Application of SLP methodology in a factory of aluminium Doors and medium density fiber plates aided by multicriteria decision analysis

Varadinov, M.[1,2]**, Bertuzzi, C.**[3]**, Eleutério, D.**[4]**, Dias, C.**[5,6]**, Santos, C.**[7,8] **and Romacho, J.**[9,10]

College of Technology and Management, Polytechnical Institute of Portalegre, Portugal[1,4,5,9]
Coordenação Interdisciplinar para a Investigação e a Inovação (C3i) do Instituto Politécnico de Portalegre, Portugal[2,10]
Universidade Tecnológica Federal do Paraná, Brasil[3]
CMA -Center of Mathematics and its Applications, New University of Lisbon, Portugal[6,8]
Department of Mathematics and Physical Sciences, Polytechnical Institute of Beja, Portugal[7]

E-mail address: *mjvaradinov@ipportalegre.pt ; carolinebertuzzi@gmail.com; daniloeleuterio1902@gmail.com; cpsd@ipportalegre.pt; carla.santos@ipbeja.pt; jromacho@ipportalegre.pt*

Abstract

The use of the correct physical arrangement for a factory helps to develop the flow of material and information, creating several benefits for the company. This article proposes the best structural design based on the method of Systematic Layout Planning (SLP) in an aluminium and MDF door factory and proposes the use of multicriteria decision support tools in systematic layout planning, refining the selection phase of the SLP method. From the results obtained with the research, it was necessary to eliminate accumulated solid waste, reallocation of machinery and stock.

Keywords: Layout, SLP, Optimization, Production, Aluminium.

Some results

The presented layout may not be the best or the most suitable, however, it shows significant improvements compared to the current one, especially when compared to the implementation cost.

References

[1]Wilde, E. (1996). Functional Planning.Facilities, v. 14, July-August, 35-39.

[2]Yang, T.; Su, C.; Hsu, Y. (2000). Systematic Layout Planning: a study on semiconductor wafer fabrication facilities.International Journal of Operations Production Management, v. 20, p. 1359-1371.

[3]Tompkins, J. A. et al. (1973). Facilities planning. 4.ed. New York: John Wiley & Sons, 2010

[4]Muther, R. Systematic Layout Planning. 2.ed. Boston: Cahners Books.

[5]Lee, Q. (1998). Projeto de instalações e do local de trabalho. São Paulo: IMAM.

[6]Chase, R. B.; Jacobs, R. F.; Aquilano, N. J. (2006) Operation management for competitive advantage with global cases. 11. ed. Editora New York McGraw - Hill.

Databases

www.apb.pt

www.bportugal.pt

www.cmvm.pt

http://ec.europa.eu/eurostat

www.ine.pt

www.pordata.pt

Contributed Talks

Monte Carlo solution of boundary value problem for semilinear isotropic diffusion equation

<u>A.S. Rasulov</u> and M.T. Bakoev

University of World Economy and Diplomacy

E-mail address: *arasulov@uwed.uz; mbakoev@uwed.uz*

Abstract

In the present work a mixed boundary value problem (BVR) for nonlinear isotropic diffusion equation is considered. Assuming the existence of solution BVR obtained a probabilistic representation of the solution as a mathematical expectation of some random variable. In accordance with a probabilistic representation on the trajectories of the branching random process were constructed unbiased estimator of the solution. An unbiased estimator of the solution has finite variance, based on the trajectories of a branching process with a finite average number of branching and easily simulated.

Keywords: Isotropic diffusion equation, statistical modeling, unbiased estimators, mean value theorem,branching Markov process.

——⊙——
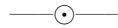
The mean value theorem for the Laplace equation plays an important role in the theory of harmonic functions. It is well known, the popular algorithm in Monte Carlo is the "random walk on spheres" for solving the Dirichlet problem for an elliptic equation with given conditions on the boundary. The basis of the "random walk on spheres" method is a mean value theorem. This type of theorem is less known for parabolic type operators. For parabolic equations in general, the work of Pini [1] was the first to publish such a theorem.The Kernel in such a parabolic mean value theorem is normally singular in time. From possible mean value theorems, which are known for solving some classes of parabolic equations, we will give one of them. The application of this theorem will be given in the work. With the purpose of further application, we present mean value theorem with a positive and bounded kernel in the domain of integration.We will not study the properties of the above mean value formula. The only important fact is that

the constructed integral equation allows us to construct Monte Carlo algorithms for the solution of BVP , with the integral operators being connected with jump Markov processes.

The analog of the "random walk on the spheres" for the heat equation was considered in the work of Haji-Sheik and Sparrow [2]. Assuming that a solution of considered BVP exists using known methods [3] of statistical simulation we propose a new method to construct unbiased estimators for solving boundary value problem at an arbitrary point $(x^0, t^0) \in$ $considered domain Q$. Furthermore, in each specific case we can derive a corresponding mean value formula. The methods of simulating the branching Markov processes are given in [4] . On trajectories of these processes we constructed unbiased estimators for solving some nonlinear BVP problems.The obtained results can be used for further studying and constructing Monte Carlo algorithms for the solution more general diffusion BVP with polynomial nonlinearities. In the future we want to apply these results for the solution to the practical physic problems. The results obtained in this work were used in solving some model problems.

References

[1]Pini V., (1954) Sulla solutzone generalizzata di Wiener peril primo problema di valori al contorno nel caso parabolico. Rend. Del. Sem. Math. Padova.**23**: p.422-434.

[2]Haji-Sheikh A., Sparrow E.M., (1967) Monte-Carlo method for solving the thermal equations. Proc. Amer. Assoc. Engeen. and mechan. **2**: p.1-18

[3]Ermakov S.M., Nekrutkin V.V., Sipin A.S. (1984),Random processes for solutions of classical equations of mathematical physics. Moskov.: Nauka, p.206.

[4]Rasulov A.S., Mascagni M., Raimova G.M. (2006) Monte Carlo Methods for the Solution of Linear and Nonlinear Boundary Value Problems. (Monograph) Printing house of UWED, Tashkent, p. 346.

Estimation of foreseeable and unforeseeable risks in motor insurance

Alfredo D. Egídio dos Reis[1,2], **Corina Constantinescu**[3,4], **Véronique Maume-Deschamps**[5,6] and **Weihong Ni**[7,8]

ISEG, Universidade de Lisboa[1]
CEMAPRE[2]
University of Liverpool, UK[3]
Institute for Financial and Actuarial Mathematics[4]
Université Claude Bernard Lyon 1, Fr[4]
Institut Camille Jordan, Institut de Science Financiére et d'Assurances[5]
Universidade de Lisboa[6]
CEMAPRE, ISEG[7]

E-mail address: *alfredo@iseg.ulisboa.pt; c.constantinescu@liverpool.ac.uk; veronique.maume@univ-lyon1.fr; wni.grace@gmail.com*

Abstract

This project works with the risk model introduced by [3] and quests modeling, estimating and pricing insurance for risks brought in by innovative technologies, or some other emerging risks. The model considers together two diferent risks streams that arise together, however not clearly separated or observed. Specifically, we consider a risk surplus process where premiums are adjusted according to past claim frequencies, like in a Bonus-Malus (BM) system in motor insurance, where we consider a *classical* or *historical risk* stream and an *unforeseeable risk* one. Particularly, this latter stream represent unknown risks which can be of high uncertainty that, when pricing insurance (ratemaking and experience rating), suggest a sensitive premium adjustment strategy. When a claim arrives, it is not clear for the Actuary to observe which claim comes from one or the other stream, when modelling such risks it is of utmost importance to figure out the nature of both the probability of the occurrence of such claims and its amount. Subsequently, premium calculation must be estimated and fairly reflect these two risk streams. This is not an easy task. We propose here an estimation procedure for the distributions of the claim counts and the corresponding severities for both streams of risks, as well as the premium to be charged.

We assume a Bayesian approach as used in credibility theory for premium calculation/estimation.

The model starts assuming a sum of two mixed counting processes, one representing the *historical risk* stream and the other the *unforeseeable risk* one. Concerning the first process the randomized intensity parameter, say $\Lambda^{(1)}$, has a classical behaviour, it is positive, whereas in the second process, with corresponding parameter $\Lambda^{(2)}$, we set a positive probability of $\Lambda^{(2)}$ being zero.

Keywords: bonus-malus; credibility theory; foreseeable and unforeseeable risks; mixed Poisson processes; ratemaking and experience rating.

——⊙——
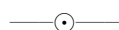
Modelling two different streams of risks for the same portfolio can be done exclusively either on the claim count process like [1] and [3] tried, or in the claim severity or on both. Here, we must consider a dependence between the claim counts process, $\{N(t) = N(t)^{(1)} + N(t)^{(2)}, t \geq 0\}$, and sequence of individual severities, say $\{Y_j\}_{j=1}^{\infty}$, separated into two different sequences. The introduction of the severity consideration means that the different streams may also *bring* different severity behaviour. This is not dealt by previous authors. We start dealing with the claim frequency component only, after that we deal with the claim severity behaviours, considering that the severity may bring some information on the stream origin.

We calculate Bayesian and credibility premia. Parameter estimation is not straightforward, we use the Expectation-Maximisation (EM) algorithm.

Acknowledgements

References

[1]Dubey, A. (1977) Probabilité de ruine lorsque le paramètre de Poisson est ajusté a posteriori. *Mitteilungen der Vereinigung Schweiz Versicherungsmathematiker* **2**: 130-141.

[2]Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2012) *Loss models: From data to decisions*, volume 715, John Wiley & Sons.

[3]Li, B., Ni, W., and Constantinescu, C. (2015) Risk models with premiums adjusted to claims number. *Insurance: Mathematics and Economics* **65** (2015): 94-102.

[4]Lukacs, E. (1955) A characterization of the gamma distribution. *The Annals of Mathematical Statistics* **26** (2): 319-324.

[5]Waimo (2016). Waimo team. `https://waymo.com/`

Mixed effects ANOVA with random sample sizes

Anacleto Mário[1]**, Célia Nunes**[1,2]**, Dário Ferreira**[1,2]**, Sandra S. Ferreira**[1,2] **and João T. Mexia**[3]

Center of Mathematics and Applications, University of Beira Interior, Portugal[1]
Department of Mathematics, University of Beira Interior, Portugal[2]
Center of Mathematics and Applications, Faculty of Sciences and Technology, New University of Lisbon, Portugal[3]

E-mail address: *anitomario@yahoo.com.br; celian@ubi.pt; dario@ubi.pt; sandraf@ubi.pt; jtm@fct.unl.pt*

Abstract

In this work we aim to present a new approach considering orthogonal mixed models, under situations of stability, when the samples dimensions are not known in advance. We will assume that the occurrences of observations correspond to counting processes, which lead us to consider the sample sizes as realizations of independent random variables with Poisson distribution. The applicability of the proposed approach is illustrated through an application considering the incidence of unemployed persons in the European Union. The interest of this approach lies in avoiding false rejections obtained when using the classical ANOVA.

Keywords: mixed models, random sample sizes, stability situations, counting processes, unemployment in European Union.

Acknowledgements

References

[1] Mexia, J.T., Nunes, C., Ferreira, D., Ferreira, S.S. and Moreira, E. (2011) Orthogonal fixed effects ANOVA with random sample sizes, *Proceedings of the 5th International Conference on Applied Mathematics, Simulation, Modelling (ASM'11)*, Corfu, Greece, pp. 84–90.

[2]Nunes, C., Ferreira, D., Ferreira S.S. and Mexia, J.T. (2012) F-tests with a rare pathology, *Journal of Applied Statistics*, **39**(3): 551–561.

[3]Nunes, C., Ferreira, D., Ferreira, S.S. and Mexia, J.T. (2014) Fixed effects ANOVA: an extension to samples with random size, *Journal of Statistical Computation and Simulation* **84**(11): 2316–2328.

[4]Nunes, C., Capistrano, G., Ferreira, D., Ferreira, S. S. and Mexia, J. T. (2018) Exact critical values for one-way fixed effect models with random sample sizes. *Journal of Computational and Applied Mathematics.* doi: 10.1016/j.cam.2018.05.057.

Structural Equation Modelling with missing data - application to PISA 12 data

Conceição Leal[1,2]**, Teresa Oliveira**[1,2] **and Amílcar Oliveira**[1,2]

Universidade Aberta[1]
CEAUL[2]

E-mail address: *conceicao.leal2010mail.com; teresa.Oliveira@uab.pt, amilcar.oliveira@uab.pt*

Abstract

Structural Equation Modelling (SEM) is a collection of multivariate statistical techniques, which, in most cases, are used to formulate, adjust, and test a wide variety of models.

Solving complex problems requires the ability to examine multiple influences and multiple responses simultaneously. Given the flexibility and comprehensiveness of SEM applications, this methodology provides a means to develop and evaluate ideas about complex multivariate relationships which makes it capable of responding to problems and challenges of both Social and Human Sciences and Natural Sciences.

A common problem in SEM concerns the frequent existence of missing data, a problem that is transversal to all types of data analysis. Omitting data can reduce the statistical power of a study and produce biased estimates, leading to invalid conclusions. SEM is not immune to these problems either.

In this work, we will present the theoretical framework of SEM with missing data, namely the mechanisms that lead to this type of data and the procedures to handle with them in the context of the methodology. Special attention will be given to the resources available in the R software to handle with missing data, specifically for SEM analysis. For this purpose, we will present an application to data of PISA 2012, concerning the literacy in Mathematics of Portuguese young people.

Keywords: Structural Equation Modelling, Missing data, SEM package of R, PISA 2012.

References

[1]Allison, P. D. (2003). Missing data techniques for structural equation modeling. Journal of abnormal psychology, 112(4), 545.

[2]Beaujean, A. A., & Beaujean, M. A. A. (2018). Package 'BaylorEdPsych'. Version 0.5

[3]Buuren, Stef van, Groothuis-Oudshoorn, Karin (2011). mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1-67. URL `https://www.jstatsoft.org/v45/i03/`.

[4]Epskamp, Sacha (2017). semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output. An R package version 1.1

[5]Fox, J., Nie, Z., Byrnes, J., Culbertson, M., DebRoy, S., Friendly, M., ... & Fox, M. J. (2017). Package 'sem'.

[6]Gelman, A., Hill, J., Su, Y. S., Yajima, M., Pittau, M., Goodrich, B., ... & Goodrich, M. B. (2015). Package 'mi'.

[7]Graham, J. W., & Coffman, D. L. (2012). Structural equation modeling with missing data. In R. H. Hoyle (Ed.), Handbook of structural equation modeling (pp. 277-295). New York, NY, US: Guilford Press.

[8]Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. Journal of statistical software, 45(7), 1-47.

[9]Jamshidian, M., Jalal, S. J., & Jansen, C. (2014). MissMech: an R package for testing homoscedasticity, multivariate normality, and missing completely at random (MCAR).

[10]Jamshidian, M., Jalal, S. J., & Jansen, C. (2014). MissMech: an R package for testing homoscedasticity, multivariate normality, and missing completely at random (MCAR).

[11]Oberski, D. (2014). lavaan. survey: An R package for complex survey analysis of structural equation models. Journal of Statistical Software, 57(1), 1-27. URL `http://www.jstatsoft.org/v57/i01/`.

[12]Peugh, J. L., & Enders, C. K. (2004). Missing Data in Educational Research: A Review of Reporting Practices and Suggestions for Improvement. Review of Educational Research, 74(4), 525-556.

[13]Rosseel, Y. (2018). The lavaan tutorial. Department of Data Analysis: Ghent University.

Efficient approximate sampling from k-DPP for large graphs

Diala Wehbe[1,2] **and Nicolas Wicker**[1]

University of Lille[1] */ Paul Painlevé Laboratory*
Lebanese University/Azm Center for Research in Biotechnology and its Applications[2]

E-mail address: *diala.wehbe@net.usj.edu.lb*

Abstract

Recently, the amount of graph-structured data available has been increasing rapidly. While the graph might contain similar nodes, our goal is to find a diverse subset of nodes which authorizes getting an outline of different types of information related to the ground set. The aim of this work is to sample a set of nodes from a large graph with accordance to fixed-size determinantal point processes. A polynomial bound on the mixing time for Markov chain sampling from a determinantal point process with fixed cardinality is given under certain conditions on the eigenvalues of the Laplacian matrix.

Keywords: determinantal point process, Laplacian kernel, Markov chain, Metropolis-Hastings, mixing time.

——⊙——
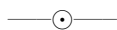
A k-Determinantal Point Processes denoted k-DPPs which conditions a standard DPP on the event that the modeled set is of fixed size is defined by a kernel L, basically a symmetric and positive semidefinite matrix [2]. In this work, a clarification of how k-DPP offer a powerful approach to modeling diversity is stated by finding interesting nodes in connected graphs where the goal is to select a diverse set of relevant nodes according to the users query.

The most important step in this sampling method is constructing a Markov chain whose stationary distribution is the k-DPP. The aim behind this approach is to suggest a new configuration by choosing two elements: one to remove from the set of size k and another to add. Based on the main theorem of Anari et al. [1], choosing the Moore-Penrose pseudo-inverse of the

normalized Laplacian matrix as the kernel will be the right tool utilized to generate an approximate sample from k-DPP.

The following results provide a fast mixing sampler for k-DPP, for which polynomial bounds on the mixing time are presented:

For any $\epsilon > 0$, if $\lambda_2(\mathcal{L}_G) - \frac{1}{\sqrt{d_{\min}}} > \frac{d_{\max}}{m}$ then the Markov chain described above generates ϵ-approximate sample of k-DPP in time

$$\tau_\epsilon \leq 2k^2 n \cdot \log\left[\frac{2nm}{d_{\min}\epsilon^{\frac{1}{k}}}\right],$$

where $\lambda_2(\mathcal{L}_G)$ is the second smallest eigenvalue of the normalized Laplacian, $d_{\min} = \min_i d_i$, $d_{\max} = \max_i d_i$, m is the number of edges of the graph G and n the number of nodes.

It is shown that the results provide a fast mixing sampler for k-DPP, for which polynomial bounds on the mixing time are presented if for all $n \geq 0$,

$$\frac{B_n}{C_n} \longrightarrow 0 \text{ and } \lambda_2(\mathcal{L}_G) - \frac{1}{\sqrt{d_{\min}}} > \frac{d_{\max}}{m},$$

with

$$B_n = 2\sqrt{\frac{d_n}{2m}\left(1 + \frac{1}{d_{\min}}\right)} \text{ and } C_n = \lambda_2(\mathcal{L}_G) - \frac{1}{\sqrt{d_{\min}}} + 2\frac{d_n}{m}.$$

Intuitively, these conditions are satisfied if $\lambda_2(\mathcal{L}_G)$ is large compared to $\frac{d_{\max}}{m}$. Finally, the resulting mixing time depends on whether the graph bottleneck is large or narrow. The larger the better for the convergence.

Acknowledgements

References

[1] Anari, N. , Gharan, S. O. and Rezaei A. (2016) Monte Carlo Markov chain algorithms for sampling strongly Rayleigh distributions and determinantal point processes. *In COLT.*

[2] A. Kulesza and B. Taskar.(2012) Determinantal point processes for machine learning. *Foundations and Trends in Machine Learning* **5**: 123–286.

Analysis of the Skewness and Kurtosis for the Product of Two Normally Distributed Variables

J. Antonio Seijas-Macias[1,2], Amílcar Oliveira and Teresa Oliveira[3,2]

University of A Corunha (Spain)[1]
CEAUL - Universidade de Lisboa[2]
Universidade Aberta (Portugal)[3]

**E-mail address:** *antonio.smacias@udc.gal*

Abstract

The analysis of the product of two normally distributed variables does not seem to follow any known distribution that can be applied in a generalised manner to the different situations that may occur. From the first approach to the problem [1], until the most recent studies ([2],[3]), the shape of the product is an open problem. Some advances are reported during these years: first difference occurs between whether or not both variables are correlated; the second one would be if both normal distributions are the same or different. Depending on these circumstances we will have different situations that can be modelled using different strategies. Under some circumstances, we can represent the product using a function (Bessel Function, Pearson Type III function, ...), other cases are modelling using a normal distribution, or a skew-normal distribution.

Although the PDF of the product can not be calculated could be represented as a difference between two integrals, but solution of these integrals requires of numerical algorithms [6]. Fortunately, the moment-generating function is available and could be used to calculate the statistics of the distribution of the product: mean, variance, skewness and kurtosis. The study of these stastistics was considered in [4] where the evolution of mean and variance is considered a function of the values of the mean and variance of the two normal distributions considered in the product. More recently ([5]) the influence of these parameters into skewness of the product is analysed.

In this work, we considered the role that the parameters of the two normal distributions (mean and variance) have into the evolution of the skewness and kurtosis of the product. The paper consider a first approach when there is no

correlation between the two elements of the product. Next approach consider the existence of correlation and its influence on skewness and kurtosis. Range for kurtosis and skewness are defined. Determination of the evolution of skewness and kurtosis of the product could be used to establish the normality of the product and the best approach to it.

Keywords: Skewness, Kurtosis, Product of Normal Variables, Skewnormal Distribution, Normal Approach.
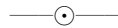
—⊙—

Acknowledgements

References

[1]Craig, C. (1936) On the Frequency Function of xy. *Annals of Mathematical Society* **7**:1–15

[2]Cui, G, Yu, X., Iommelli, S. and Kong, L. (2016) Exact Distribution for the Product of Two Correlated Gaussian Random Variables. *IEEE Signal Processing Letters* **11**:1662–1666

[3]Nadarajah, S. and Pogány, T. K. (2016) On the distribution of the product of correlated normal random variables. *Comptes Research de la Academie Sciences Paris, Serie I* **354**: 201–204

[4]Seijas-Macias, A. and Oliveira, A. (2012) An approach to distribution of the product of two normal variables. *Discussiones Mathematicae. Probability and Statistics* **32:1-2**: 87–99

[5]Oliveira, A., Oliveira, T. and Seijas-Macias, A. (2016) Skewness into the product of two normally distributed variables and the risk consequences. *REVSTAT-Statistical Journal* **14:2**: 119–138

[6]Ware, R. and Lad, F. (2003) *Approximating the Distribution for Sums of Product of Normal Variables*, Research-Paper 2003-15. Department of Mathematics and Statistics (University of Canterbury - New Zealand)

Adaptive estimation for light-tailed models

<u>Lígia Henriques-Rodrigues</u>[1,3] and M. Ivette Gomes[2,3]

Universidade de São Paulo, IME, São Paulo, Brasil[1]
*Universidade de Lisboa, FCUL, Portugal*[2]
Centro de Estatística e Aplicações, CEAUL[3]

E-mail address: *ligiahr@ime.usp.br; ivette.gomes@fc.ul.pt*

Abstract

In this work we propose a new class of consistent semi-parametric tail index estimators for light-tailed models, i.e., models with a negative extreme value index. Light-tailed models are very common in practice, in areas such as environment and hydrology, among others. The extreme value index is the primary parameter in statistics of extremes and in a semi- parametric context the classical estimators depend upon an integer tuning parameter k, related to the number of upper order statistics used in the estimation. The new class of estimators is also dependent on another real tuning parameter, θ, that properly chosen leads us to an estimator with null second-order component of asymptotic bias for a large variety of models. In this work we address the adaptive choice of the tuning parameters, (k, θ), under play providing an algorithm based on a combination of a bootstrap and a data-driven heuristic criterion. Applications to simulated data sets are also provided.

Keywords: extreme value index; semi-parametric estimation; generalized Hill estimator; bootstrap.
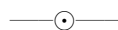
—⊙—

References

[1]Beirlant, J., Vynckier, P. and Teugels, J.L. (1996). Tail index estimation, Pareto quantile plots and regression diagnostics, *J. Amer. Statist. Assoc.*, **91**: 1659–1667.

[2]Beirlant, J., Dierckx, G. and Guillou, A. (2005). Estimation of the extreme-value index and generalized quantile plots, *Bernoulli*, **11** (6): 949–970.

[3]Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*, Wiley series in Probability and Statistics.

[4]Caeiro, F. and Gomes, M.I. (2010). An asymptotically unbiased moment estimator of a negative extreme value index, *Discussiones Mathematicae – Probability and Statistics* **30**: 5–19.

[5]Dekkers, A., Einmahl, J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution, *Annals of Statistics*, **17**: 1833–1855.

[6]Fraga Alves, M.I., Gomes M.I., Haan, L. de and Neves, C. (2009). Mixed moment estimator and location invariant alternatives, *Extremes* **12** (2): 149–185.

[7]Gomes, M.I. and Neves, C. (2008). Asymptotic comparison of the mixed moment and classical extreme value index estimators, *Statistics and Probability Letters* **78** (6): 643–653.

[8]Gomes, M.I., Henriques-Rodrigues, L. and Caeiro, F. (2013). Refined estimation of a light tail: an application to environmental data. In Nicola Torelli *et al.* (eds.) *Advances in Theoretical and Applied Statistics*, 143–153, Springer Berlim Heildelberg.

[9]Haan, L. de and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*, Springer-Verlag.

[10]Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *Ann. Statist.* **3**: 1163–1174.

[11]Reiss, R.-D. and Thomas, M. (2007). *Statistical Analysis of Extreme Values, with Application to Insurance, Finance, Hydrology and Other Fields*, 3rd edition, Birkhäuser Verlag.

Forecast time series in presence of structural breaks: a proposal

<u>M. Rosário Ramos</u>[1] and Clara Cordeiro[2,3]

Universidade Aberta[1]
*Faculty of Sciences and Technology, Universidade do Algarve*[2]
Centre of Statistics and Applications[3]

E-mail address: *MariaR.Ramos@uab.pt; ccordei@ualg.pt*

Abstract

For a time series $\{y_1, \cdots, y_n\}$, is assumed that a structural change exists at a unknown time if $\{y_1, \cdots, y_t\}$ differ from $\{y_{t+1}, \cdots, y_n\}$, in its trend behavior. In order to develop an improved methodology to forecast in this context, a combined procedure is proposed, and a first study is presented. The starting phase is the estimation and removing of the seasonality using the Seasonal-Trend decomposition by Loess (STL) [1]. The selection of the best STL fit was performed by the algorithm stl.fit() developed by [2], which runs all the possible combinations of the smoothing parameters and in the end, finds the optimal combination of parameters which minimized an accuracy measure. Moreover, the detetion of a structural break in the seasonally adjusted time series is performed by the package strucchange [3] in R [4]. If the trend after the breakpoint shows a persistence of magnitude significantly different from the previous period, the subset until the first breaking point is partially or totaly discarded and forecasting is performed over the remaining observations of the time series. The performance of the approach is evaluated by comparing the forecasts from the original and the fitted time series and using several real time series.

Keywords: time series, Seasonal-Trend decomposition by Loess, breaking points, forecast.
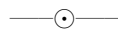
——⊙——

References

[1] Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I.(1990) Stl: A seasonal-trend decomposition procedure based on loess. J. Off. Stat. **6**(1), 3–73

[2] Cristina, S., Cordeiro, C., Lavender, S., Goela, P.C., Icely, J. and Newton, A. (2016) MERIS Phytoplankton Time Series Products from the SW Iberian Peninsula (Sagres) Using Seasonal-Trend Decomposition Based on Loess. *Remote Sensing* **8** (6): 449, doi:10.3390/rs8060449, http://www.mdpi.com/2072-4292/8/6/449.

[3] Achim Zeileis, Friedrich Leisch, Kurt Hornik and Christian Kleiber (2002). strucchange: An R Package for Testing for Structural Change in Linear Regression Models. *Journal of Statistical Software*, **7**(2), 1–38. URL http://www.jstatsoft.org/v07/i02/

[4] R Develpment core team,*R: A Language and Environment for Statistical Computing;* software available at http://www.R-project.org.

Optimal control problem for a non-autonomous SIR model for the Ebola virus

<u>Paulo Rebelo</u>[1,2], Silvério Rosa[1,3] and César M. Silva[1,2]

Universidade da Beira Interior (UBI)[1]
Centro de Matemática e Aplicações, UBI[2]
Instituto de Telecomunicações, Polo da Covilhã, UBI[3]

E-mail address: *rebelo@ubi.pt; rosa@ubi.pt; csilva@ubi.pt*

Abstract

The aim of this work is to present and study an optimal control problem for a non-autonomous SIR (susceptible-infected-recovered) epidemic model for the propagation of the Ebola virus. The controls of the epidemic model mean vaccination, information campaigns of the susceptible individuals and the treatment of infectious elements of the population. Numerical simulation is provided to validate the model.

Keywords: Mathematical Biology, Optimal Control, Numerical Simulation, Epidemic model, Ebola outbreak.

References

[1] Yoshida, N., Hara, T. (2007), Global stability of a delayed SIR epidemic model with density dependent birth and death rates, *J. Comput. Appl. Math.* 201*, 339 − 347.*

[2] Zaman, Gul, Han Kang, Yong and Hyo Jung, Il, (2009), Optimal treatment of an SIR epidemic model with time delay, *BioSystems* 98*, 43 − 50.*

[3] Rosa, S., Rebelo, P., Silva, C.M., Alves, H. and Carvalho, P.G., (2018) Optimal control of the customer dynamics based on marketing policy, *Applied Mathematics and Computation*, v. 330, pp. 42 − 55.

[4] Mateus, J. P., Rebelo, P., Rosa, S., Silva, C.M. and Torres, D.F.M., (2018) Optimal control of non-autonomous SEIRS models with vaccination and treatment, Discrete and Continuous Dynamical Systems - Series S, Volume 11, Number 6, December.

Multinomial regression models to estimate the risk of diabetes and hypertension in Pima Indian Women

<u>Pereira J.A.</u>[1], Pereira A.L.[2] and Oliveira T.[2]

University of Porto[1]
University Aberta[2]

E-mail address: *jpereira@fmd.up.pt*

Abstract

The obesity defined as an excess of body fat is usually assessed, by approximation, with the body mass index with values over 30Kg/m2. Obesity is strongly associated with serious health issues that are potentially life threatening, including type II diabetes mellitus and hypertension and these relationships are well established [1].

The diabetes mellitus is a metabolic disorder characterized by persistent hyperglycemia and is the most common form of diabetes, affecting $85 - 90\%$ of all people with diabetes. This condition results from insulin resistance, which results from a combination of factors, among the most relevant are family history, obesity, advanced age, ethnicity [2].

The hypertension, defined by persistently elevated blood pressure in the arteries, is usually asymptomatic. However is a major risk factor for several severe heart conditions, vision loss, chronic kidney disease, and dementia. The some of the main risk factors are age, ethnicity, and Being overweight or obese [2].

The population of Pima Indians from Arizona currently have the highest prevalence of diabetes and obesity in the United States, with more than 50% of adults over 35 being type 2 diabetic, therefore one of the most studied and with abundant data available [3].

The objective of this work is to evaluate the relationships between obesity and age with hypertension, and obesity and age with type 2 diabetes.

The data used is the Pima Indians Diabetes Database [4] of women available online on data.world.

To accomplish our aims multinomial logistic regression and binomial models were fitted to the data in R environment, the p-values for the regression coefficients were computed using Wald tests. The packages used were "foreign", "nnet", "ggplot2", "aod", and "reshape2".

The results showed that, the probability of prehypertension increase with age for the three considered ponderal conditions (normoponderal, overweight and obesity) being higher in obese individuals and evolving similarly in the normoponderal and overweight individuals. Regarding the evolution of the probability of diabetes considering the age, those of overweight and obese behaved differently from normoponderal being the probabilities of the obese larger.

We conclude that in the Pima women the risk of hypertension and diabetes increase with age and differ among ponderal states.

Keywords: Ponderal condition, Blood pressure, Diabetes, Multinomial regression, Pima Indians.

References

[1] Prospective Studies Collaboration. (2009) Body-mass index and cause-specific mortality in 900000 adults: collaborative analyses of 57 prospective studies. The Lancet 373 (9669) : 1083-1096.

[2] Finnish Diabetes Prevention Study Group (2001) Prevention of Type 2 Diabetes Mellitus by Changes in Lifestyle among Subjects with Impaired Glucose Tolerance. The New England Journal of Medicine 344 : 1343-1350.

[3] Knowler, W. C., Pettitt D. J., Saad M. F., Bennett P. H. (1990) Diabetes mellitus in the pima indians: Incidence, risk factors and pathogenesis. Diabetes/Metabolism/Reviews, 6 (1) : 1-27.

[4] `https://www.kaggle.com/uciml/pima-indians-diabetes-database`

[5] R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org/`.

Contributed Poster

Statistical Modelling: a Study on Customer Retention and Customer Lifecycle Management

A. Manuela Gonçalves[1,2], Guadalupe Costa[3] and Alexandre Freitas[4]

University of Minho[1]
Center of Mathematics, Portugal[2]
University of Minho, Department of Mathematics and Applications, Portugal[3]
*University of Porto, Faculty of Economy, Portugal[4]*

E-mail address: *mneves@math.uminho.pt; guadalupefilipa95@hotmail, alexandremrfreitas@gmail.com*

Abstract

With the increasing competitiveness of the health fitness industry, implementing strategies and tactics to prevent the loss of customers is highly important, particularly because the acquisition of new customers entails high costs for companies. Their retention and loyalty may be vital in the medium- to long-term financial health of a company. It is important for marketing planners to develop initial strategies to attract customers and engage them to patronize the products in the long run. The customer can only be motivated to enter into such a contract if attractive incentives are available.

In this work, in the context of a Portuguese health fitness company, we intend to analyze customers' behaviour-customer retention and customer lifecycle management-namely through the understanding of customer retention, which is the most important key performance indicator (KPI) in this industry; learning from customer cancellations; discovering how to prevent customer cancellations by improving customer touch points; exploring the importance of the first 30 days of customer retention; and learning about customer retention as a team effort, but especially as a fitness team effort. We develop statistical methodologies within the context of generalized linear models and survival analysis to predict and forecast customer retention-the customer is continuing (or active) or non-continuing (historic)-in a health fitness company comprising nine clubs, in the registration period from January 2013 to December 2015. The data were collected on February 1, 2018.

Keywords: Health Fitness Industry, Customer Lifecycle, Retention, Linear Models, Generalized Linear Models, Survival Analysis.
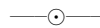
———⊙———

Acknowledgements

References

[1]Glady, N., Baesens, B. and Croux, C. (2009) Modeling customer lifetime value. *European Journal of operational Research* 197: 402-411.

[2]R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

[3]Reinartz, W. and Kumar, V. (2003) The impact of customer relationship characteristics on profile lifetime duration. *Journal of Marketing* 67: 77-99.

[4]Wood, S.N. (2006) Generalized Additive Models. An Introduction with R. CRC/Chapman-Hall.

Accuracy Assessment of Time Series Forecasting Models

A. Manuela Gonçalves[1,2], Susana Lima[3] and Marco Costa[4,5]

University of Minho[1]
CMAT-Center of Mathematics, Portugal[2]
University of Minho, Department of Mathematics and Applications, Portugal[3]
University of Aveiro[4]
CIDMA-Center for Research and Development in Mathematics and Applications, Portugal[5]

E-mail address: *mneves@math.uminho.pt; susanamrlima@gmail.com, marco@ua.pt*

Abstract

This study deals with forecasting economic time series that have strong trend and seasonal patterns. How to best model and forecast these patterns has been a long-standing issue of time series analysis. In this work, we propose different approaches to time series forecasting by combining different models in order to increase the chance of capturing different patterns in the data and thus improve forecasting performance. Therefore, the main propose of this study is to compare the accuracy of the SARIMA models and their extensions, the classical decomposition time series associated with multiple regression models, and Holt-Winters models. These methods are chosen because of their ability to model trend and seasonal fluctuations present in economic data.

Models are fitted to time series of retail sales (retail time series in the Eurostat retail databases). Finally, a comparison is made and discussed.

Keywords: Forecast Accuracy, Time Series Models, SARIMA Models, Multiple Regression Models, Holt-Winters Models, Retail Sales.
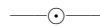
——⊙——

References

[1]Box, G., Jenkins, G., Reinsel, G. and Ljung, G. (2016) Time series analysis: forecasting and control. New Jersey: John Wiley and Sons.

[2]Gooijer, J.G.D. and Hyndman, R.J. (2006) 25 years of time series forecasting. International Journal of Forecasting 22: 443-473.

[3]R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Models with commutative orthogonal block structure: study of models joining

Santos, C.[1,2], Dias, C.[1,3] Nunes, C.[4] and Mexia, J. T.[1,5]

CMA -Center of Mathematics and its Applications, New University of Lisbon, Portugal[1]
Department of Mathematics and Physical Sciences, Polytechnical Institute of Beja, Portugal[2]
College of Technology and Management, Polytechnical Institute of Portalegre, Portugal[3]
Department of Mathematics and Center of Mathematics and Applications, University of Beira Interior, Portugal[4]
*Department of Mathematics, New University of Lisbon, Portugal[5]*

E-mail address: *carla.santos@ipbeja.pt ; cpsilvadias@gmail.com; celian@ubi.pt*

Abstract

Imposing a commutativity condition on models with orthogonal block structure we obtain a special class of linear mixed models, named COBS (models with orthogonal block structure) where the least squares estimators, LSE, for estimable vectors are uniformly best linear unbiased estimators, UBLUE.

Considering n initial COBS with segregation and matching, resorting to the algebraic structure of the models and the Cartesian product, through joining we get models, consisting in the overlap of the observations of the initial models, that retain these characteristics.

Keywords: Commutativity, Matching, Models joining, Mixed models, Segregation.
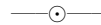
——⊙——

References

[1]Nelder, J. A. (1965a) The analysis of randomized experiments with orthogonal block structure I. Block structure and the null analysis of variance, *Proceedings of the Royal Society*, Series A 283, 147-162.

[2]Nelder, J. A. (1965b) The analysis of randomized experiments with orthogonal block structure II. Treatment structure and the general analysis of variance, *Proceedings of the Royal Society*, Series A283, 163-178.

[3]Santos, C., Nunes, C., Dias, C., & Mexia, J. T. (2017). Joining models with commutative orthogonal block structure. *Linear Algebra and its Applications*, 517, 235-245.

[4]Seely, J. (1970) Quadratic subspaces and completeness. *Ann. Math. Statist.*,42, 710-721.

[5]Zmy?lony, R. (1978) A characterization of best linear unbiased estimators in the general linear model, Mathematical Statistics and Probability Theory 2, 365-373.

Rank one mean matrix: application to cross product matrices

## <u>Cristina Dias</u>[1,2], Carla Santos[2,3] and João Tiago Mexia[2,4]

Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Portalegre[1]
CMA - Centro de Matemática e Aplicações da Universidade Nova de Lisboa[2]
Departamento de Matemática e Ciências Físicas do Instituto Politécnico de Beja[3]
Departamento de Matemática da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa[4]

E-mail address: *cpsd@ipportalegre.pt; carla.santos@ipbeja.pt, jtm@fct.unl.pt*

Abstract

Degree-one models can be applied to cross products matrices and Hilbert-Schmidt scalar products matrices. The latter have an important role in the first stage (inter-structure) of STATIS methodology, while the former matrices (in particular the $\mathbf{AA}^t$ and $\mathbf{A}^t\mathbf{A}$ cross products, which have the same non-null eigenvalues) have an important role in inference. The case of rank one is interesting since the first eigenvector of matrix $\mathbf{XX}^t$ may be used to describe the behavior of the variables corresponding to the columns of $\mathbf{X}$. We now consider the estimators of the pair (λ, α) and testing that the mean matrix as rank one. We apply our results to cross product matrices $\mathbf{AA}^t$, given an numerical example.

Keywords: Eigenvalues, Mean matrix, Inference; Acceptance regions, Simulation.
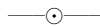
———⊙———

References

[1] Dias, (C. 2013) Models and Families of Models for Symmetric Stochastic Matrices. Ph.D. Thesis, University of Évora, Évora.

[2] Dias, C., Oliveira, M. M. & Mexia, J. T. (2012) Models of Symmetric Stochastic Matrices. Proceedings of the International Conference of Numerical Analysis and Applied Mathematics. Vol. 1479 (pp. 1666-1669).

[3] Mexia, J. T. (1990) Best Linear Unbiased Estimates, Duality of F Tests and the Scheffé Multiple Comparison Method in Presence of Controlled Heterocedasticity. Comp. Stat & Data Analysis, 10(3), 271-281.

[4] Mexia, J. T. (1995) Introdução à Inferência Estatística Linear. Centro de Estudos de Matemática Aplicada. Edições Universitárias Lusófonas: Lisboa.

[5] Ito, P. K. (1980) Robustness of Anova and Macanova Test Procedures, P. R. Krishnaiah (ed), Handbook of Statistics1, Amsterdam: North Holland, pp. 199-236.

Taguchi's Method vs Fractional Factorial Plans

Dina Salvador, Sandra Monteiro and Sandra Nunes[1,2]

Polytechnic Institute of Setúbal[1]
Centre for Mathematics and Applications of Faculty of Sciences and Technology of NOVA University of Lisbon[2]

E-mail address: *dina.salvador@estsetubal.ips.pt;*
sandra.ines.monteiro@esce.ips.pt;sandra.nunes@esce.ips.pt

Abstract

The Taguchi method was originally developed for applications in the field of quality engineering, having acquired many enthusiasts due to the improvements introduced in this area [3]. This method is also popular in design of experiments in many other areas, as we can see in [2], [4] and [5]. Many experimenters use orthogonal matrices developed by Taguchi in design of experiments. The orthogonal matrices defined by Taguchi allow to save time and costs because they considerably reduce the number of treatments used in the planning of experiments.

In this work, we intend to expose the Taguchi method in a simplified way, presenting its main guidelines, with an approach to fractional factorial plans.

Keywords: Taguchi Method's, Orthogonal Matrix, Design of Experiments, Fractional Factorial Plans.

———⊙———
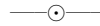
References

[1] Montgomery, D. (1997) *Design and analysis of experiments*, New York: Wiley.

[2] Nalbant, M., Gokkaya, H. and G. Sur, G. (2007). Application of Taguchi method in the optimization of cutting parameters for surface roughness in turning. *Mater. Des.*, **28**, (4): 1379-1385.

[3] Ross, P. J. (1996) *Taguchi Techniques for Quality Engineering*, 2nd ed, McGraw-Hill.

[4] Roy, R. K. (2001) *Design of experiments using the Taguchi approach*, John Wiley Sons Inc., New York.

[5] Taner, T. and Antony, J. (2000). The assessment of quality in medical diagnostic tests: a comparison of ROC/Youden and Taguchi methods. *Int. J. Health Care Qual. Assur.*, **13**, (7): 300-307.

[6] Wu, Y. and Wu, A. (2000) *Taguchi methods for robust design*, New York: ASME Press.

SEM to evaluate the impact of psychosocial risks on workers' health and wellbeing

Luís M. Grilo

Polytechnic Institute of Tomar, Department of Mathematics and CMA, FCT, New University of Lisbon, Portugal

E-mail address: *lgrilo@ipt.pt*

Abstract

The number of workers under psychosocial risks has been increasing worldwide, which has negative consequences on workers' health and wellbeing, in their organizations and in economies in general. For the workers, the psychosocial risks can be translated into emotional, behavioural and/or physical symptoms, such as stress, sleep disorders, alcohol and drug abuse, migraine, depression, conflicts, burnout, aggression, harassment, suicide, etc. To evaluate these impacts in workers from a Portuguese services institution, a survey was conducted applying an adapted version of the international validated Copenhagen Psychosocial Questionnaire (COPSOQ), with variables expressed in a Likert-type scale. Although the sociodemographic variables that characterize the workers (such as gender, age and educational stage) were not available, a theoretical Structural Equation Model (SEM) was proposed for the burnout syndrome. The estimated model was obtained using the Partial Least Squares (PLS) approach and a nonparametric bootstrap procedure was also used to assess the statistical significance of the estimated path coefficients with PLS-SEM. The latent construct 'quantitative demands' has a direct effect on both 'stress' and 'work-family conflict' and an indirect effect on 'burnout', through these constructs, which also have a direct effect on 'burnout'. These results were considered of great importance for the occupational health specialists, who champion the protection of workers' health, avoiding bad work atmosphere, poor social relationships and absenteeism, since these entail huge costs, in terms of human suffering as well as a decline in economic performance.

Keywords: Burnout, Stress, SmartPLS, Survey, Work-family conflict.

References

[1] Directorate-General for the Humanisation of Labour (2013). Guide to the prevention of psychosocial risks at work. Universit de Namur (Flohimont, V., Lambert, C., Berrewaerts, J., Zaghdane, S. and Fzfa, A.). www.employment.belgium.be.

[2] Grilo, L. M. and Vairinhos, V. (2018). PLS-PM to evaluate workers health perception. V Workshop on Computational Data Analysis and Numerical Methods, Instituto Politécnico do Porto, Felgueiras, Portugal, May 11-12 (p. 61, Book of Abstracts).

[3] Grilo, L. M., Grilo, H. L., Gonçalves, S. P. and Junça, A. (2017). Multinomial Logistic Regression in Workers Health. ICCMSE 2017, AIP Conf. Proc. 1906, 110001.

[4]Hair, J. F., Hult, G. T. M., Ringle, C. M. and Sarstedt, M. (2017). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) (2nd Ed., Thousand Oakes, CA: Sage).

[5]Kristensen, T. S., Borritz, M., Villadsen, E. and Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. Work and Stress 19: 192 207.

[6]Maslach, C. and Leiter, M.P. (2016). Understanding the burnout experience: recent research and its implications for psychiatry. World Psychiatry 152: 103-111.

[7]Maslach, C. and Leiter, M.P. (2008). Early predictors of job burnout and engagement. Journal of Applied Psychology 93: 498-512.

[8]Montero-Marin J., Zubiaga, F., Cereceda, M., Piva Demarzo, M.M., Trenc, P. and Garcia Campayo, J. (2016). Burnout Subtypes and Absence of Self-Compassion in Primary Healthcare Professionals: A Cross-Sectional Study. PLoS ONE 11(6): e0157499.

[9]Ringle, C. M., Wende, S. and Becker, J.-M. (2015). SmartPLS 3. Bnningstedt: SmartPLS. Retrieved from http://www.smartpls.com.

[10]Silva, C. F. (2016). Copenhagen Psychosocial Questionnaire (COPSOQ). Portuguese version. FCT MEC, Universidade de Aveiro, Análise Exacta.

[11]Stehlík, M., Helpersdorfer, Ch., Hermann, P., Supina, J., Grilo, L. M., Maidana, J. P., Fuders, F. and Stehlíková, S. (2017). Financial and risk modelling with semicontinuous covariances. Information Sciences 394-395, 246-272.