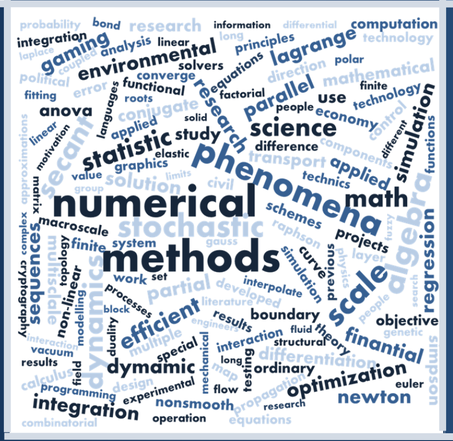
May,

11-12/2018



Instituto Politécnico do Porto, Portugal
Escola Superior de Tecnologia e Gestão, Felgueiras

V Workshop
on
COMPUTATIONAL DATA ANALYSIS AND NUMERICAL METHODS

BOOK OF ABSTRACTS

Joint organization:

P.PORTO

ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO

ipt Instituto Politécnico de Tomar

UNIVERSIDADE DE EVORA

BOOK OF ABSTRACTS

INSTITUTO POLITÉCNICO DO PORTO
ESTG–ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO
11^{st}–12^{nd} May, 2018
Felgueiras – PORTUGAL

Participating Institutions

Participating Institutions

ESCOLA
SUPERIOR
DE TECNOLOGIA
E GESTÃO

Scientific Research Centers

Scientific Research Centers

Projecto UID/MAT/04674/2013

Scientific Sponsor

Scientific Sponsor

List of Participants

List of Participants

- A. Manuela Gonçalves (mnves@math.uminho.pt)
- Alberto Simões (asimoes@ubi.pt)
- Aldina Correia (aic@estg.ipp.pt)
- Ana Nogueira (anogueira@civiparts.com)
- Ana Rodrigues (amr@inesctec.pt)
- Ana Isabel Borges (aib@estg.ipp.pt)
- Ana Nata (anata@ipt.pt)
- Ana Teixeira (8130097@estg.ipp.pt)
- Andrew Prokhorenkov (andrew.prokhorenkov@gmail.com)
- Carla Francisco (carla.francisco.1@ulaval.ca)
- Carla Santos (carla.santos@ipbeja.pt)
- Carlos Ramos (ccr@uevora.pt)
- Célia Nunes (celian@ubi.pt)
- Cheraitia Hassen (hassen.stat@hotmail.fr)
- Cristina Dias (cpsd@ipportalegre.pt)
- Dário Ferreira (dario@ubi.pt)
- Dinis Pestana (ddpestana@fc.ul.pt)
- Domingos Silva (domingosjlsilva@gmail.com)
- Dora Prata Gomes (dsrp@fct.unl.pt)
- Eliana Costa e Silva (eos@estg.ipp.pt)

- Fátima De Almeida (mff@estg.ipp.pt)
- Fernanda A. Ferreira (faf@esht.ipp.pt)
- Fernanda Figueiredo (otilia@fep.up.pt)
- Fernando Carapau (flc@uevora.pt)
- Filipe Marques (fjm@fct.unl.pt)
- Filomena Teodoro (maria.alves.teodoro@marinha.pt)
- Flora Ferreira (flora.ferreira@gmail.com)
- Graça Soares (gsoares@utad.pt)
- Gustavo Soutinho (gustavosoutinho@sapo.pt)
- Helena Penalva (helena.penalva@esce.ips.pt)
- Ilda Inácio (ilda@ubi.pt)
- Isabel Cristina Lopes (cristinalopes@iscap.ipp.pt)
- Ivana Malá (malai@vse.cz)
- José Lobo Pereira (lobopereiramail@gmail.com)
- José G. Dias (jose.dias@iscte-iul.pt)
- Luís Bandeira (lmzb@uevora.pt)
- Luís Miguel Bandeira (ee12106@fe.up.pt)
- Luis Miguel Grilo (lgrilo@ipt.pt)
- Luísa Novais (luisa_novais92@hotmail.com)
- M. Ivette Gomes (ivette.gomes@fc.ul.pt)
- Manuel Cruz (mbc@isep.ipp.pt)
- Manuela Neves (manela@isa.ulisboa.pt)

- Maria Clara Grácio (mgracio@uevora.pt)
- Maria Varadinov (mjvaradinov@ipportalegre)
- Maria de Fátima Pilar (8120109@estg.ipp.pt)
- Oliva Martins (oliva.martins@ipt.pt)
- Oussama Rida (oussamar@gmail.com)
- Patrícia Antunes (antunes.patricia03@gmail.com)
- Paulo Rebelo (rebelo@ubi.pt)
- Raquel Oliveira (rmro_17@hotmail.com)
- Sandra Ferreira (sandraf@ubi.pt)
- Sergey Frenkel (fsergei51@gmail.com)
- Sérgio Nunes (spnunes@ipt.pt)
- Stella Abreu (sau@isep.ipp.pt)
- Susana Faria (sfaria@math.uminho.pt)
- Telmo Matos (tsm@estg.ipp.pt)
- Weronika Wojtak (w.wojtak@gmail.com)

Committees

Committees

Organizing Committee

Aldina Correia, Instituto Politécnico do Porto, Portugal (Local Chair)
Eliana Costa e Silva, Instituto Politécnico do Porto, Portugal
Fátima De Almeida, Instituto Politécnico do Porto, Portugal
Fernando Carapau, Universidade de Évora, Portugal
Luís M. Grilo, Instituto Politécnico de Tomar, Portugal (Chair)

Executive Committee

Anuj Mubayi, Arizona State University, USA
Fernando Carapau, Universidade de Évora, Portugal
Luís M. Grilo, Instituto Politécnico de Tomar, Portugal
Milan StehlíK, Johannes Kepler University, Áustria

Scientific Committee

Adélia Sequeira, IST, Portugal
Ana Borges, Instituto Politécnico do Porto, Portugal
Ana Cristina Nata, Instituto Politécnico de Tomar, Portugal
Ashwin Vaidya, Montclair State University, USA
Anuj Mubayi, Arizona State University, USA
Arminda Manuela Gonçalves, Universidade do Minho, Portugal
Ana Silvestre, Instituto Superior Técnico, Portugal
Amílcar Oliveira, Universidade Aberta, Portugal
Alexandra Moura, Instituto Superior de Economia e Gestão, Portugal
Alberto Simões, Universidade da Beira Interior, Portugal
Adelaide Figueiredo, Universidade do Porto, Portugal
Aldina Correia, Instituto Politécnico do Porto, Portugal
Bong Jae Chung, George Mason University, USA
Carla Santos, Instituto Politécnico de Beja, Portugal
Carlos Ramos, Universidade de Évora, Portugal
Carlos Agra Coelho, Universidade Nova de Lisboa, Portugal
Célia Nunes, Universidade da Beira Interior, Portugal
Cristina Dias, Instituto Politécnico de Portalegre, Portugal
Dário Ferreira, Universidade da Beira Interior, Portugal
Delfim F. M. Torres, Universidade de Aveiro, Portugal

Dora Gomes, Universidade Nova de Lisboa, Portugal
Eliana Costa e Silva, Instituto Politécnico do Porto, Portugal
Fátima De Almeida Ferreira, Instituto Politécnico do Porto, Portugal
Fernanda Figueiredo, Universidade do Porto, Portugal
Filipe Marques, Universidade Nova de Lisboa, Portugal
Fernando Carapau, Universidade de Évora, Portugal
Filomena Teodoro, Escola Naval, Portugal
Gabriel Pires, Instituto Politécnico de Tomar, Portugal
Henrique Pinho, Instituto Politécnico de Tomar, Portugal
Ilda Inácio, Universidade da Beira Interior, Portugal
Isabel Silva Magalhães, Universidade do Porto, Portugal
Irene Rodrigues, Universidade de Évora, Portugal
João Janela, Instituto Superior de Economia e Gestão, Portugal
João Romacho, Instituto Politécnico de Portalegre, Portugal
João Luís Miranda, Instituto Politécnico de Portalegre, Portugal
José Gonçalves Dias, Instituto Universitário de Lisboa, Portugal
José António Pereira, Universidade do Porto, Portugal
José Augusto Ferreira, Universidade de Coimbra, Portugal
Lígia Ferreira, Universidade de Évora, Portugal
Luís Bandeira, Universidade de Évora, Portugal
Luís M. Grilo, Instituto Politécnico de Tomar, Portugal
Maria Luísa Morgado, Uni. de Trás-os-Montes e Alto Douro, Portugal
Maria Clara Grácio, Universidade de Évora, Portugal
Manuel Branco, Universidade de Évora, Portugal
Milan Stehlík, Johannes Kepler University, Áustria
Mouhaydine Tlemcani, Universidade de Évora, Portugal
Óscar Oliveira, Instituto Politécnico do Porto, Portugal
Paulo Correia, Universidade de Évora, Portugal
Pedro Lima, Instituto Superior Técnico, Portugal
Pedro Oliveira, Universidade do Porto, Portugal
Russell Alpizar-Jara, Universidade de Évora, Portugal
Rafael Santos, Universidade do Algarve, Portugal
Sara Fernandes, Universidade de Évora, Portugal
Sandra Ferreira, Universidade de Beira Interior, Portugal
Sérgio Nunes, Instituto Politécnico de Tomar, Portugal
Susana Faria, Universidade do Minho, Portugal
Tanuja Srivastava, Indian Institute of Technology, India
Teófilo Melo, Instituto Politécnico do Porto, Portugal
Teresa Oliveira, Universidade Aberta, Portugal
Valter Vairinhos, Centro de Investigação Naval, Portugal

Technical Specifications

Technical Specifications

Title:

V Workshop on Computational Data Analysis and Numerical Methods: Book of Abstracts

Web page:

http://www.wcdanm-ipporto18.uevora.pt/

Editor:

Instituto Politécnico do Porto
Escola Superior de Tecnologia e Gestão
Centro de Inovação e Investigação em Ciências Empresariais e Sistemas de Informação
Rua do Curral – Margaride
4610–156 Felgueiras, Portugal

Workshop place:

ESTG–Escola Superior de Tecnologia e Gestão, Instituto Politécnico do Porto

Authors:

Aldina Correia, Eliana Costa e Silva, Fátima De Almeida, Luís M. Grilo and Fernando Carapau

Published and printed by:

INE - Instituto Nacional de Estatística

ISBN: 978-989-98447-6-6 (print version)

Preface

Preface

Dear Participants, Colleagues and Friends,

It is a great honour and a privilege to give you all a warmest welcome to the Fifth Annual Workshop of Computational Data Analysis and Numerical Methods (V WCDANM).

This Workshop is being held at the beautiful Campus 3 of Instituto Politécnico do Porto, located in the city of Felgueiras, Portugal. The host institution, as well as Instituto Politécnico de Tomar and Universidade de Évora have been fully committed on this challenge from the beginning, hoping that the final result could exceed expectations for participants, sponsors and organizers. The contributions from Plenary Speakers, the high scientific level of oral and poster presentations and an active audience, will certainly contribute to the success of the meeting. A special thanks to all of them, since this event could not be possible without any of these essential and complementary parts.

A thank is also due to the Members of the Executive, Scientific and Organizing Committees, specially to Aldina Correia (Local Chair), Eliana Costa e Silva, Fátima De Almeida (hosts in Instituto Politécnico do Porto) and Fernando Carapau (Universidade de Évora), who have been continually working hard together to yield a balanced, wide-scoped and interesting programme, having accomplished an outstanding result.

The number of participants in the CDANM Workshop, coming from Portugal and other countries, have been increasing every year, with particular interest in applications to specific research fields, namely in Health and Social Sciences, Environmental and Engineering.

For the second consecutive year authors have the opportunity to publish full versions of the abstracts were presented in a special issue of the International

Journal of Applied Mathematics and Statistics (IJAMAS), after refereeing process and according to the conditions of the Journal.

It is a pleasure joining you in Felgueiras, hoping that the Workshop could be intellectually stimulating and an opportunity for the researchers to work together, providing all unforgettable moments!

Felgueiras, 11^{st}-12^{nd} May, 2018

Chairman of the Executive Committee of VWCDANM

Luís Miguel Grilo
Instituto Politécnico de Tomar

Programme

Programme

V WCDANM, ESTG – Instituto Politécnico do Porto
May 10^{st}–12^{nd}, 2018, Felgueiras, Portugal

Event Place

School of Management and Technology, Polytechnic Institute of Porto (ESTG–P. Porto)
Felgueiras, Portugal

Friday (May 11^{st})

09:30 | Registration

10:00 | Open Ceremony of the V WCDANM

10:15 | Plenary Session: Milan Stehlík, Johannes Kepler University, Áustria

11:00 | Coffee Break

11:15 | Paralell Sessions

12:45 | Lunch

14:00 | Social Program

20:00 | Dinner

Saturday (May 12^{nd})

09:30 | Registration

10:15 | Plenary Session: Anuj Mubayi, Arizona State University, USA

11:00 | Coffee Break

11:15 | Paralell Sessions

12:45 | Lunch

14:00 | Plenary Session: Pedro Oliveira, Biomedical Sciences Institute of Abel Salazar, Porto University, Portugal

14:45 | Paralell Sessions

16:15 | Coffee Break and Poster Session

17:00 | Paralell Sessions

18:30 | Closing Ceremony: Local and Executive Committees

Contents

Contents

Contributed Posters

Invited Speakers

The hidden outbreaks: the role of biodiversity, scarcity of information, and health disparity on mathematical modeling and analysis of a neglected vector borne disease in a complex landscape of global health

Anuj Mubayi[1]

Collaborators:
Mugdha Thakur (Arizona State University-Tempe, USA), Xin Liu (Clemson University-Clemson, USA), Edwin Michael (The University of Notre Dame-South Bend, USA), Malay Banerjee (Indian Institute of Technology-Kanpur, India), Shyam Sundar (Banaras Hindu University-Varanasi, India), Niyamat Ali-Siddiqui (Rajendra Memorial Research Inst. of Medical Sciences-Patna, India), Pradeep Das (Rajendra Memorial Research Inst. of Medical Sciences-Patna, India), Carlos Castillo-Chavez (Arizona State University-Tempe, USA)

[1]Simon A. Levin Mathematical Computational and Modeling Science Center, School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA

E-mail: amubayi@asu.edu

Abstract

Neglected vector-borne diseases (NVDs) pose some of the greatest challenges in public health, especially in tropical and sub-tropical regions of the world. NVDs have caused repeated outbreaks in poorest of poor communities in spite of effective and lasting intervention tools available in richer communities. NVDs are a result of a medical-social-political problem, where the economic cost of the disease burden, the human suffering from diseases, and the transformation of the communities due to disease can only be tackled by strong will to address challenges using medical, scientific, and mathematical expertise. Efforts to control or eliminate NVD have been underpinned by a theoretical and mathematical framework developed some 100 years ago by Ross and Macdonald, including models, metrics for measuring transmission, and theory of control that identifies key vulnerabilities in the transmission cycle. The analysis of their modeling framework resulted in tipping point

theory for epidemics via the computation of disease reproduction number, entomological transmission metrics and vectorial capacity, which are now used extensively to study dynamics and design of interventions for many NVDs. Since their pioneer work on describing and modeling of malaria life cycle at the start of last century, the vast majority of modeling literature on NVDs attempted to use some variation of the Ross-Macdonald modeling framework that involve of homogeneous transmission Neglected vector-borne diseases (NVDs) pose some of the greatest challenges in public health, especially in tropical and sub-tropical regions of the world. NVDs have caused repeated outbreaks in poorest of poor communities in spite of effective and lasting intervention tools available in richer communities. NVDs are a result of a medical-social-political problem, where the economic cost of the disease burden, the human suffering from diseases, and the transformation of the communities due to disease can only be tackled by strong will to address challenges using medical, scientific, and mathematical expertise. Efforts to control or eliminate NVD have been underpinned by a theoretical and mathematical framework developed some 100 years ago by Ross and Macdonald, including models, metrics for measuring transmission, and theory of control that identifies key vulnerabilities in the transmission cycle. The analysis of their modeling framework resulted in tipping point theory for epidemics via the computation of disease reproduction number, entomological transmission metrics and vectorial capacity, which are now used extensively to study dynamics and design of interventions for many NVDs. Since their pioneer work on describing and modeling of malaria life cycle at the start of last century, the vast majority of modeling literature on NVDs attempted to use some variation of the Ross-Macdonald modeling framework that involve of homogeneous transmission assumption in a well-mixed population. However, studies evaluating questions on modeling of changing heterogeneity in transmission processes to effective use of available scarce NVD data and the capacity to model to analyze such heterogeneity are the most important but relatively unexplored component in current literature. Local heterogeneity due to pathogen transmission pathways, biodiversity, and disparity in population may cause infection dynamics to be highly nonlinear, and poses problems for mathematical modeling, epidemiology and measurement of quantities. Novel mathematical approaches from nonlinear dynamics, sensitivity and uncertainty analysis and high performance computing used for other widely studied diseases such as HIV, and Malaria, show how heterogeneity arises from the biology and the landscape on which the processes of vector biting and feeding preferences, host competence, access to interventions and pathogen transmission unfold. In this talk, I will provide various examples using NVDs as a case study that uses emerging quantitative theory and focuses attention on the diversity of ecological and social context for vector and hosts for modeling dynamics and control. The example will include questions

on insecticide resistance, vector preference, host competence, multiple hosts, movement of individuals, health disparity, and community ecology. I will describe the established global stability and bifurcation analysis results for the considered stochastic and delayed dynamical systems. The presentation will include derivation of simple expressions to estimate the basic reproduction number of disease using the initial exponential growth rate of an outbreak for directly and indirectly transmitted NVDs and time series epidemic data. The talk will also include challenges, progress, and future of NVDs and its control.

Multi-criteria aggregation for cancer risk assessment

Milan Stehlík[1,2,3]

[1]Department of Applied Statistics, Johannes Kepler University in Linz, Austria
[2]Linz Institute of Technology (LIT), Johannes Kepler University in Linz, Austria
[3]Institute of Statistics, Universidad de Valparaíso, Valparaíso, Chile

E-mail: milan.stehlik@jku.at

Abstract

There exist a big need for discrimination between mammary cancer and mastopathy tissues (see [4]). Non invasive techniques generally may produce inverse problems, e.g. estimating a Hausdorff fractal dimension from boundary of examined tissue (see [3]). We will discuss these issues in the context of our recent results (see e.g. [2]). During the talk we will discuss several issues which bring light into both fractal based cancer modelling and more general stochastic geometry models and their comparisons. When we consider fractal based cancer diagnostic, many times a statistical procedure to assess the fractal dimension is needed. We shall look for some analytical tools for discrimination between cancer and healthy ranges of fractal dimensions of tissues. [1] discussed planar tissue preparations in mice which has a remarkably consistent scaling exponents (fractal dimensions) for tumor vasculature even among tumor lines that have quite different vascular densities and growth characteristics. In [10] we provide extensive study of cancer risk assessment on simulated and real data and fractal based cancer. Both non-random and random carpets have been validated for modelling of the cancer growth, and it was shown that only random carpets can be used. We constructed a statistical test, which is able to distinguish between the two groups, masthopathy and mammary cancer (see [9]). The inter-patient variability of fractal dimension for mammary cancer is high (see [11]) and therefore multifractality is a better concept (see [6]). This is a feasible and parsimous solution for a more delicate problem of multi-objective aggragation of information. This touches bases for general topological approach for aggregation [12], which links to Sugeno integral as a way to aggregation in bornological spaces. The algebraic and topologic properties of cancer growth are available via appropriate set structures, e.g. bornology (see [7,8]). Such structures can be very useful for defining fractal cancer hypothesis. Nephroblastoma (given by Wilms' tumour) is the typical tumour of the kidneys appearing in childhood, which

does not satisfy fractal cancer hypothesis. We illustrate on recent pre/post clinical study the effect on chemotherapy to Euclidean volumes of such tumors (see [5]).

Acknowledgements

Author acknowledge support of *FONDECYT Regular* N 1151441 and project LIT-2016-1-SEE-023.

References

[1] Baish, J.W. and Jain, R.K., (2000) Fractals and cancer. *Cancer Research*, 60, pp. 3683–3688.

[2] Filus J., Filus L. and Stehlík M., (2009) Pseudoexponential modelling of cancer diagnostic testing. *Biometrie und Medizinische Informatik, Greifswalder Seminarberichte*, Heft 15, pp. 41–54.

[3] Kiselak, J., Pardasani, K. R., Adlakha, N. Agrawal, M. and Stehlík, M., (2013) On some probabilistic aspects of diffusion models for tissue growth, *The Open Statistics and Probability Journal*.

[4] Hermann, P., Mrkvicka, T., Mattfeldt, T., Minarova, M., Helisova, K., Nicolis, O., Wartner, F. and Stehlík, M., (2015) Fractal and stochastic geometry inference for breast cancer: a case study with random fractal models and Quermass-interaction process, *Statistics in Medicine*,34, pp. 2636–2661.

[5] Hermann, P., Giebel, S.M., Schenk, J.-P. and Stehlík, M., (2015) Dynamic Shape Analysis - before and after chemotherapy, *Proc. International Conference on Risk Analysis (ICRA6)*, pp. 339–346.

[6] Nicolis, O., Kiselak, J. Porro, F. and Stehlík, M. Multi-fractal cancer risk assessment, Stochastic Analysis and Applications, DOI: 10.1080/07362994.2016.1238766.

[7] Paseka, J., Solovyov, S. A. and Stehlík, M., (2015) Lattice-valued bornological systems, *Fuzzy Sets and Systems*, 259, 15, pp. 68–88.

[8] Paseka, J., Solovyov, S. A. and Stehlík, M., (2016), On a topological universe of L-bornological spaces, *Soft Computing* 20, pp. 2503–2512.

[9] Stehlík M., Mrkvička T., Filus J. and Filus, L., (2012) Recent development on testing in cancer risk: a fractal and stochastic geometry, *Journal of Reliability and Statistical Studies*, Vol. 5, pp. 83–95.

[10] Stehlík,M., Giebel, S. M., Prostakova, J. and Schenk, J.P., (2014) Statistical inference on fractals for cancer risk assessment *Pakistan Journal of Statistic*, Vol. 30(4), pp. 439–454.

[11] Stehlík M., Hermann, P. and Nicolis, O., (2016) Fractal based cancer modelling, *Revstat*, 14(2), pp. 139–155.

[12] Stehlík, M., (2016) On convergence of topological aggregation functions, *Fuzzy Sets and Systems*, 287, pp. 48–56.

A multilevel approach in the study of paratuberculosis in milk production in Portuguese dairy farms

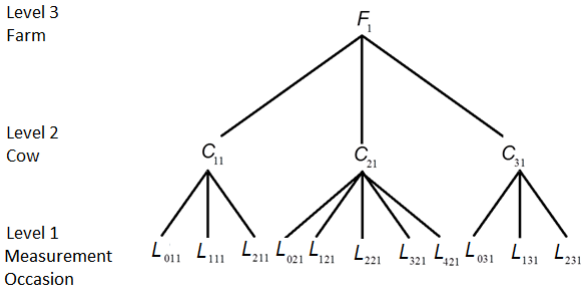
E.G. Martins[1,2,3], Pedro Oliveira[1,2], B.M. Oliveira[1], D. Mendonça[4] and J. Niza-Ribeiro[1,2]

[1]Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Portugal
[2]EPIUnit - Instituto de Saúde Pública da Universidade do Porto, Porto, Portugal
[3]Department of Veterinary Medicine, Escola Univ. Vasco da Gama, Coimbra, Portugal
[4]CINTESIS - Center for Health Technology and Services Research, Faculty of Medicine, University of Porto, Porto, Portugal

E-mail: pnoliveira@icbas.up.pt

Abstract

Mycobacterium avium subspecies paratuberculosis (MAP) is the cause of a chronic granulomatous enteric disease, also known as Paratuberculosis, or Johne's disease, affecting both ruminant and non-ruminant animals. The aim of this work was to discuss the relevance of Multilevel Mixed Models (MLM) [1, 2] in the analysis of the effect of individual cow's seropositivity to MAP on 305 days corrected milk production (D305MP) and Somatic Cell Count (SCC), based on data from the first five lactations of each cow in Portuguese dairy herds. Collected data have a natural hierarchical structure considering the "measurement occasion" (first level) nested within cows (second level), and cows nested in farms (third level) [3]. It should note that the data have



a longitudinal component which takes into account the different numbers of observations by cow, given that for some cows, just one lactation and for others, two, three or up to five lactations may be registered. A total

of 191 farms, including 14,829 cows and their respective 36,219 lactations were retained. Several multivariable multilevel models, having as dependent variable D305MP or lnCCS and as explanatory variables Lactation (L), Cow Status (CS) or Farm Status (FS) and cross-level interactions were fitted. All models have at least three variance components: a residual variance at level 1, random intercept variances at level 2 and level 3, allowing for the three-level data structure. Models were fitted using Maximum Likelihood Estimation and an unstructured random-effects variance/covariance matrix [3]. The effect of Lactation, with linear and quadratic terms, is significant in all models; these terms account for the curvature associated with D305MP and lnSCC along lactations: the production increases from first to third lactation, decreasing afterwards. The variance/covariance analysis confirms the importance of the selected structure. Each level: measurement occasion, Cow and Farm retains a significant amount of the observed variance in the data. Our study provides the first report of MAP effects in Portuguese dairy farms using a multilevel approach [3]. Strengths of this study are that it was performed using a large size data collection spanning a broad temporal scope, in a three-level structure.

Keywords: multilevel mixed model, paratuberculosis, milk Production, somatic cell count.

Acknowledgements

The authors would like to thank SEGALAB SA, EABL SNIRB, and Abigaíl Barbosa, Adelaide Pereira, António Ferreira, Salles Henriques for their contribution to this work.

References

[1] Correia-Gomes, C., Mendonça, D., Vieira-Pinto, M. and Niza-Ribeiro, J., (2013) Risk factors for Salmonella spp in Portuguese breeding pigs using a multilevel analysis, *Preventive Veterinary Medicine*, 108, pp. 159–166.

[2] Hox, J.J., (2010) Multilevel Analysis-Techniques and Applications, 2nd edition, Routledge.

[3] Martins, E.G., Oliveira, P., Oliveira, B.M., Mendonça, D. and Niza-Ribeiro, J., (2018) Association of Paratuberculosis sero-status with milk production and somatic cell counts across five lactations, using multilevel mixed models, in dairy cows, *Journal of Dairy Science* (accepted for publication).

Contributed Talks

The numerical range of a banded 3-Toeplitz operators in a Hilbert space

Ana Nata[1,2] and Natália Bebiano[2,3]

[1]Polytechnic Institute of Tomar, Portugal
[2]CMUC - Centro de Matemática da Universidade de Coimbra, Portugal
[3]Department of Mathematics, University of Coimbra, Portugal

E-mail: anata@ipt.pt

Abstract

Let M_n be the algebra of $n \times n$ complex matrices. A matrix $T_n = [t_{ij}]_{i,j=1}^n \in M_n$ is said to be a 3-Toeplitz matrix if $t_{i+3,j+3} = t_{ij}$, for $i, j = 1, 2, \ldots, n-3$. If there exists an integer $m \in \mathbb{N}$, $m < n$, such that $t_{i,j} = 0$, for $|i - j| > m$, then T_n is said to be a banded 3-Toeplitz matrix with bandwidth $2m + 1$. Let H^2 be the Hilbert space. Any infinite banded 3-Toeplitz matrix can be identified with an operator T acting on the $H^2 \times H^2$ space. In 1972, Klein prove that the numerical range of a Toeplitz operator, T_f, in a Hilbert space equals $\mathrm{conv}\,(f(\mathbb{D})$, where f is the symbol of the operator, and $\mathbb{D}$ is the complex unit circle. In this talk, we extend this result considering the characterization of the numerical range of a banded 3-Toeplitz operator defined in a Hilbert space. We prove that in this case, the numerical range coincides with the boundary of the convex hull of a family of curves. This talk is based on a joint work with Professor Natália Bebiano from the University of Coimbra, Portugal.

Keywords: numerical range, banded 3-Toeplitz operators.

References

[1] Bebiano, N. and Spitkovsky, I., (2012) Numerical ranges of Toeplitz operators with matrix symbols, *Linear Algebra Appl.*, v.436, pp. 1721–1726.

[2] Horn, R.A. and Johnson, C.R., (1991) Topics in Matrix Analysis, Cambridge University Press, New York.

[3] Klein, M., (1972) The numerical range of a Toeplitz operator, *Proc. Amer. Math. Soc.* v.35, pp. 101–103.

[4] Marcus, M. and PESCE, C., (1987) Computer generated numerical ranges and some resulting theorems, *Linear and Multilinear Algebra*, v.20 , pp. 121–157.

Signal identification: wavelet transform and structure function analysis for satellite data

Andrew Prokhorenkov[1], Kozak Liudmyla[1], Kronberg Elena[2], Grigorenko Elena[3], Petrenko Bogdan[1], Lui A.T.Y.[4] and Kundelko Iryna[1]

[1]Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
[2]Max Planck Institute for Solarsystem, Göttingen, Germany
[3]Space Research Institute of Russian Academy of Sciences, Moscow, Russia
[4]Johns Hopkins University Applied Physics Laboratory, Laurel MD, USA

E-mail: andrew.prokhorenkov@gmail.com

Abstract

Wavelet transform [1] is widely used algorithm for analysis of time series. It can be applied to any time series to decompose the signal in time-frequency domain and identify the characteristic frequencies and their amplitudes. In our study we use wavelet and Fourier transform to identify turbulence processes for in-situ magnetic field measurements. Turbulence plays an important role in plasma physics, especially for developing of magnetic field in the Earth's magnetosheath. Established turbulence model of Kolmogorov K41 [2] and more developed She-Leveque model [3] use structure function to analyse turbulences and it's characteristics power spectrum. We investigate the connection of wavelet transform and structure function of p-th order to identify turbulence processes and use both approaches to find-out their properties.

Keywords: wavelet transform, statistics, structure function, space physics.

Acknowledgements

The work is done in the frame of complex program of NAS of Ukraine on space researches for 2012–1016, the grant Az. 90 312 from the Volkswagen Foundation (VW-Stiftung) and within the framework of the educational program No.2201250 Education, Training of students, PhD students, scientific and pedagogical staff abroad launched by the Ministry of Education and Science of Ukraine.

References

[1] Torrence, C. and Compo, G.P., (1998) A Practical Guide to Wavelet Analysis, *Bull. Am. Meteorol. Soc.*, v.79, no.1, pp. 61–78.

[2] Kolmogorov, A.N., (1941) The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds Numbers, *Akademiia Nauk SSSR Doklady*, v.30, pp. 301–305.

[3] She, Z.S. and Leveque, E., (1994) Universal scaling laws in fully developed turbulence *Phys. Rev. Lett.*, v.72, no.3, pp. 336–339.

Modelling resources in selling channels for the customer service

Ana Nogueira[1], Manuel Cruz[2], Sandra Ramos[2] and Margarida Pina[1]

[1]NORS Group–Aftermarket, Portugal
[2]Mathematical Engineering Laboratory, School of Engineering, Polytechnic of Porto, Portugal

E-mail: anogueira@civiparts.com

Abstract

The customer service in a company is one of the key points to succeed as mentioned in [3] The goal as a company is to have a customer service that is not just the best but legendary. Therefore, NORS Group, particularly one of the Aftermarket businesses, wanted to improve the service level to ensure that every customer would have a great service experience. Hence, Civiparts challenge the 2Mathematical Engineering Laboratory from School of Engineering of Polytechnic of Porto (LEMA/ISEP) to develop a MSO project to model and simulate the customer's arrival rating through the phone and physical presence in the store and to optimize the number of human resources needed to fulfil the demand. NORS is a Portuguese group whose vision is to be leader in transport solutions. It sells cars, trucks, buses, machines and construction equipment, parts and service. NORS has eighty-four years of history and activity in Portugal, which started with the representation of the Volvo brand in nineteen thirty-three (1933). Civiparts, which is the business where the project was developed, is in the industry of selling parts to heavy vehicles. The company is in Portugal, represented by 8 stores, each with a warehouse to keep the parts, Spain, and Angola. There were different possible channels to communicate with the shop and make a sell, which were: going to the shop, make a phone call or send an email. The communication channels under study were the phone and the physical presence. The project began with analysing the data from the phone calls related to the year 2016 (daily data) and January 2017 (hourly data). This information was provided by an already implemented call-centre system. Furthermore, there was no data available regarding the customer's rating arrival into the stores. To solve this lack of information, the data was collected in one of the stores, in Leça da Palmeira, by measuring the number of people arriving and taking notes of the time of the different tasks that comprised the whole customer experience in the shop. The period of time was approximately 2 weeks in

different months. Analysing the data allowed to understand the behaviour of customer's demand in different months and days in the phone channel. The addition of the hourly records of January 2017 was essential to determine the hourly demand through phone call. The arrival rate of the customers that would go to the store was based on the collected data. It was developed a simulator to allocate the clients, by the rule First In First Out (FIFO), to a particular working station, in order to overcharge the minimum of employers. Thus, the simulation gave one minute of resting between each client, trying to be as close to reality as possible. The queueing theory applied was the [1] (Burnecki) Thinning algorithm using the [2] (Krzysztof Burnecki, 2010) Non-Homogeneous Poisson Process (NHPP) to simulate the arrival rate of the customers. This model allows to have different rates throughout the day, in order to increase the model adherence to reality. The simulations permitted to understand the service level by applying different number of resources available with a general customer service functions and with a specialization of only one channel of the customer service.

Keywords: queueing theory, modelling, simulation and optimization (MSO), data analysis.

Acknowledgements

The authors would like to thank and demonstrate appreciation for all the support given by the Civiparts and the rest of the Aftermarket team.

References

[1] Burnecki, K., (2004) Simulation of the NHPP - thinning, In Modelling the risk process, Hugo Steinhaus Center, Wroc Law University of Technology, pp. 17–24.

[2] Krzysztof Burnecki, Janczura J. and Weron R., (2010) Building Loss Models, Humboldt - Universität Zu Berlin, Berlin.

[3] Walton, S. (n.d.). Sam Walton Quotes. Retrieved from AZ Quotes: http://www.azquotes.com/author/15269-Sam$_W alton$.

A quantitative model for purchasing management in retail

Ana Teixeira[1], Eliana Costa e Silva[1], Isabel Cristina Lopes[2] and João Ferreira Santos

[1]CIICESI, ESTG/P.PORTO – Center for Innovation and Research in Business Sciences and Systems of Information, School of Technology and Management/Polytechnic of Porto, Portugal
[2]LEMA,CEOS.PP, ISCAP/P.PORTO – Mathematical Engineering Lab, Centre for Organisational and Social Studies of P.Porto, Accounting and Business School/Polytechnic of Porto, Portugal

E-mail: 8130097@estg.ipp.pt

Abstract

In this work we discuss a case study of Supply Chain Management in a Portuguese retail company. The increase in consumer demand caused greater uncertainty in forecasting demand and, as a result, a robust and well-designed management of the supply chain (SC) has gained increasing importance [16]. Supply Chain Management (SCM) can be seen as a complex network of interdependent organizations upstream and downstream of SC [10], where an organization's performance depends on an effective and efficient cooperation between all partners [13]. In this context, SC planning is portrayed as a critical business problem, since it implies not only an inter-organizational integration, but also four domains responsible for the different decisions: purchasing, storage, distribution and sales [11]. Of these domains, purchases are of particular importance, since they play a very important strategic role in organizations, being considered as a potential source of competitive advantage [14], mainly because it represents the entry in SC and it essentially concerns the provision of resources for the entire SC [11]. SC's planning is characterized by its complexity and strategic management can represent *imminent competitive weapons*, ensuring organizations stand out from the competition [15]. In this context, as SCM issues are developed, studies more often refer to the construction of mathematical models, in order to optimize the various problems related to the various SC domains [18]. The main objective of this study is the development of a *tailor-made* quantitative model with the aim of improving decision-making in the purchasing management of a Portuguese company in the retail sector. The proposed model has as decision variables the size of the lots and the storage mode of the products. It

considers the various associated costs, such as, picking, replacement, distribution and sales. In addition, monthly sales history information will be used as the products demand. To solve the company's challenge, a MILP model is developed and the numerical resolution will use Gurobi [17] solver with interface to AMPL [12]. Preliminary results show that the model can reduce costs by 15 %, on average. By testing the monthly demand, we found that the preference storage mode is picking by store. This means that ordering in pallets is more efficient for the company.

Keywords: supply chain management, quantitative model, decision-making, purchasing management, planning.

References

[1] Christopher, M., (2016) Logistics & supply chain management, Pearson UK.
[2] Fleischmann, B., Meyr, H. and Wagner, M., (2008) Advanced planning, *Supply chain management and advanced planning*, Springer, pp. 81–106.
[3] Fourer, R., Gay, D. and Kernighan, B., (1993) *AMPL*, volume 117, Boyd & Fraser Danvers, MA.
[4] Gold, S., Seuring, S. and Beske, P., (2010) Sustainable supply chain management and inter-organizational resources: a literature review, *Corporate social responsibility and environmental management*, 17(4), pp. 230–245.
[5] González-Benito, J., (2007) Information technology investment and operational performance in purchasing: The mediating role of supply chain management practices and strategic integration of purchasing, *Industrial Management & Data Systems*, 107(2), pp. 201–228.
[6] Ketchen Jr, D.J. and Hult, G.T.M., (2007) Bridging organization theory and supply chain management: The case of best value supply chains, *Journal of Operations Management*, 25(2), pp. 573–580.
[7] Melo, M.T., Nickel, S. and Saldanha-Da-Gama, F., (2009) Facility location and supply chain management–a review. *European journal of operational research*, 196(2), pp. 401–412.
[8] Optimization, G., (2013) Gurobi optimizer 5.0., *Gurobi: http://www. gurobi. com*.
[9] Rajeev, A., Pati, R.K., Padhi, S.S. and Govindan, K., (2017) Evolution of sustainability in supply chain management: A literature review, *Journal of Cleaner Production*, 162, pp. 299–314.
[10] Christopher, M., (2016) Logistics & supply chain management, Pearson UK.
[11] Fleischmann, B., Meyr, H. and Wagner, M., (2008) Advanced planning, *Supply chain management and advanced planning*, Springer, pp. 81–106.
[12] Fourer, R., Gay, D. and Kernighan, B., (1993) *AMPL*, volume 117, Boyd & Fraser Danvers, MA.
[13] Gold, S., Seuring, S. and Beske, P., (2010) Sustainable supply chain management and inter-organizational resources: a literature review, *Corporate social responsibility and environmental management*, 17(4), pp. 230–245.
[14] González-Benito, J., (2007) Information technology investment and operational performance in purchasing: The mediating role of supply chain management practices and strategic integration of purchasing, *Industrial Management & Data Systems*, 107(2), pp. 201–228.
[15] Ketchen Jr, D.J. and Hult, G.T.M., (2007) Bridging organization theory and supply chain management: The case of best value supply chains, *Journal of Operations Management*, 25(2), pp. 573–580.

[16] Melo, M.T., Nickel, S. and Saldanha-Da-Gama, F., (2009) Facility location and supply chain management:a review, *European journal of operational research*, 196(2), pp. 401–412.

[17] Optimization, G., (2013) Gurobi optimizer 5.0. *Gurobi: http://www.gurobi.com*.

[18] Rajeev, A., Pati, R.K., Padhi, S.S. and Govindan, K., (2017) Evolution of sustainability in supply chain management: A literature review, *Journal of Cleaner Production*, 162, pp. 299–314.

Generalized additive models - an application in health

A. L. Pereira[1], J. A. Pereira[2] and T. Oliveira[3]

[1]MBB, Universidade Aberta, Portugal
[2]Faculdade de Medicina Dentária, Porto, Portugal
[3]MBB, Universidade Aberta e Centro de Estatística e Aplicações (CEAUL), Portugal

E-mail: lobopereiramail@gmail.com

Abstract

The long QT syndrome (LQTS) is an uncommon hereditary disorder characterized by abnormal electrical activity in the heart that may lead to sudden death. LQTS can be diagnosed with an electrocardiogram (ECG). The QT interval represents the time that cardiac cells take from contraction to relaxation. An abnormally long QT leads to ventricular arrhythmias and therefore a fall in blood pressure and loss of consciousness [1]. LQTS cases may pass undiagnosed if variations of heart rate occur. In this paper we propose to model the QT interval from group of waves (complex) seen on an electrocardiogram, representing ventricular depolarization (QRS) and the interval from the peak of one QRS complex to the peak of the next as shown on an electrocardiogram (RR). It is used to assess the ventricular rate) and gender with generalized additive models (GAM) described by Hastie and Tibirani [2]. The GAM is a nonparametric extension of the linear model which replaces the linear predictor $\eta = \sum_{j=i}^{p} \beta_j X_j$ with an additive predictor of the form $\eta = \sum_{j=i}^{p} S_j(X_j)$, and is assumed that $E(Y/X_1, X_2, \ldots, X_j) = S_0 + \sum_{j=i}^{p} S_j(X_j)$, where $S(X)$ is an unspecified smooth function that can be estimated by any scatterplot smoother. Several GAMs were fitted to a sample of 825 individuals who attend routine ECG exam, using R [3]. The best GAM fit model was $QT = 372.01 + 12.53Sex + S(RR, QRS)$, where the (RR,QRS) defines the tensor product smooths that are smooths of variables RR,QRS which allow the degree of smoothing to be different with respect to different variables (RR,QRS). The dimension of the bases (k) used to represent the smooth term was 4, which is adequate, since the k-index is greater than 1. The concurvity measures provided are smaller than 1 (0.42 for the parametric component, and worst=0.187, observed=0.058 and estimate=0.019 for the smooth term), suggesting no concern about concurvity. We found $p < 0.001$ for the parametric component and for the approximate significance of smooth terms, the adjusted R^2, the deviance explained were 0.62 and 62.3%, respectively, and the AIC is 6975.35, the lower value among the fitted models. We

succeeded to model the QT interval; however we need to test the model with data from patients with LQTS.

Keywords: electrocardiogram, QT interval, generalized additive models.

References

[1] Schwartz P.J., Priori S.G., Spazzolini C., Moss A.J., Vincent G.M., Napolitano C., Denjoy I., Guicheney P., Breithardt G., Keating M.T., Towbin J.A., Beggs A.H., Brink P., Wilde A.A., Toivonen L., Zareba W., Robinson J.L., Timothy K.W., Corfield V., Wattanasirichaigoon D., Corbett C., Haverkamp W., Schulze-Bahr E., Lehmann M.H., Schwartz K., Coumel P. and Bloise R., (2001) Genotype-Phenotype Correlation in the Long-QT Syndrome Gene-Specific Triggers for Life-Threatening Arrhythmias, *Circulation*, 103, pp. 89–95.

[2] Hastie T. and Tibshirani R., (1986) Generalized Additive Models, *Journal of the Royal Statistical Society B*, 1(3), pp. 297–318.

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

σ-semi-Hyers-Ulam stability for higher order integro-differential equations

A.M. Simões[1,2,3] and L.P. Castro[3,4]

[1]CMA-UBI – Centro de Matemática e Aplicações da Universidade da Beira Interior, Universidade da Beira Interior, Portugal
[2]Departamento de Matemática, Universidade da Beira Interior, Portugal
[3]CIDMA – Centro de Investigação e Desenvolvimento em Matemática e Aplicações, Universidade de Aveiro, Portugal
[4]Departamento de Matemática, Universidade de Aveiro, Portugal

E-mail: asimoes@ubi.pt

Abstract

This talk is devoted to present the σ-semi-Hyers-Ulam stability for higher order integro-differential equations within appropriate metric spaces. We will show that σ-semi-Hyers-Ulam stability is a new kind of stability somehow between the Hyers-Ulam and the Hyers-Ulam-Rassias stabilities. Sufficient conditions are obtained in view to guarantee Hyers-Ulam, σ-semi-Hyers-Ulam and Hyers-Ulam-Rassias stabilities for such a class of higher order integro-differential equations. We will be considering finite intervals as integration domains to define the higher order integro-differential equations. Among the used techniques, we have fixed point arguments and generalizations of the Bielecki metric. Some examples of the application of the proposed theory will be included.

Keywords: Hyers-Ulam stability, σ-semi-Hyers-Ulam stability, Hyers-Ulam-Rassias stability, Banach fixed point theorem, Bielecki metric, higher order integro-differential equations, nonlinear integral equation.

Acknowledgements

This work was supported in part by FCT–*Portuguese Foundation for Science and Technology* through the *Center for Research and Development in Mathematics and Applications (CIDMA)* of University of Aveiro, within UID-MAT-04106-2013, and through the *Center of Mathematics and Applications of University of Beira Interior (CMA-UBI)*, within project UID-MAT-00212-2013.

References

[1] Castro, L.P. and Guerra, R.C., (2013) Hyers-Ulam-Rassias stability of Volterra integral equations within weighted spaces, *Lib. Math. (N.S.)*, 33(2), pp. 21–35.

[2] Castro, L.P. and Ramos, A., (2010) Hyers-Ulam and Hyers-Ulam-Rassias stability of Volterra integral equations with a delay, *Integral Methods in Science and Engineering*, 1, pp. 85–94.

[3] Castro, L.P. and Ramos, A., (2009) Hyers-Ulam-Rassias stability for a class of nonlinear Volterra integral equations, *Banach J. Math. Anal.*, 3(1), pp. 36–43.

[4] Castro, L.P. and Ramos, A., (2011) Hyers-Ulam stability for a class of Fredholm integral equations, *Mathematical Problems in Engineering Aerospace and Sciences ICNPAA 2010, Proceedings of the 8th International Conference of Mathematical Problems in Engineering, Aerospace and Science*, pp. 171–176.

[5] Castro, L.P. and Simões, A.M., (2017) A New Type of Stability: semi-Hyers-Ulam-Rassias Stability, Book of Abstracts: IV Workshop on Computational Data Analysis and Numerical Methods, pp. 42.

[6] Castro, L.P. and Simões, A.M., (2017) Different types of Hyers-Ulam-Rassias Stabilities for a class of integro-differential equations, *Filomat*, 31(17), pp. 5379–5390.

[7] Castro, L.P. and Simões, A.M., (2018) Hyers-Ulam and Hyers-Ulam-Rassias stability for a class of integro-differential equations, *Mathematical Methods in Engineering: Theoretical Aspects*, Tas, K. and Baleanu, D. and Tenreiro Machado, J.A. Eds, Springer (to appear).

[8] Castro, L.P. and Simões, A.M., (2017) Hyers-Ulam and Hyers-Ulam-Rassias stability of a class of Hammerstein integral equations, *AIP Conference Proceedings*, 1798:020036-1/10.

[9] Castro, L.P. and Simões, A.M., (2017) Hyers-Ulam and Hyers-Ulam-Rassias stability of a class of integral equations on finite intervals, *CMMSE'17: Proceedings of the 17th International Conference on Computational and Mathematical Methods in Science and Engineering*, vol. I-IV, pp. 507–515.

[10] Castro, L.P. and Simões, A.M., (2018) Hyers-Ulam-Rassias Stability of Nonlinear Integral Equations Through the Bielecki Metric, *Mathematical Methods in the Applied Sciences*, (to appear).

An efficient software for packing boxes in pallets

Ana Moura[1], Isabel Cristina Lopes[2], Stella Abreu[1] and Manuel Cruz[3]

[1]LEMA, ISEP, Politécnico do Porto, and CMUP, Portugal
[2]LEMA, CEOS.PP, ISCAP, Politécnico do Porto, Portugal
[3]LEMA, ISEP, Politécnico do Porto, Portugal

E-mail: sau@isep.ipp.pt

Abstract

In this work, we will describe a software developed to solve an industrial packing problem related to the single container loading problem and the pallet loading problem ([1]): given a set of 3D rectangular boxes of different types, the objective is to store the boxes into containers or pallets, as effectively as possible, minimizing the number of necessary containers or pallets. Our scenario considers boxes with fragile content, usually marked with the label This side up, to be packed in pallets. Some of the boxes allow rotation, others do not. The cargo is weakly heterogeneous, i.e., the assortment of boxes is small, usually less than half a dozen different references. The stability of the boxes is also considered. Given the specificities of our scenario, we proposed a tailor-made greedy heuristic, and we implemented it in Matlab. The software developed finds packing solutions, which are easy to understand and execute by a non-expert, for example a worker of some company. The graphical user interface gives a 3D visualization of the boxes that allows rotation of the 3D pallet or container, to see it from all angles. In practice, workers can usually place boxes exceeding the limits of the pallets for some centimetres. Our software compares the solutions, considering a chosen tolerance. The heuristic works best for packing into pallets, but it also works for containers of any size. We will present the results of our heuristic and will compare them with other results from the literature [2-5], using a benchmark test set.

Keywords: three dimensional packing, heuristics, container loading, pallet loading, industrial mathematics.

Acknowledgements

This work was partially funded by LEMA (Engineering Mathematical Laboratory).

References

[1] G. Wäscher, H. Haußner, H. Schumann, (2007) An improved typology of cutting and packing problems, *European Journal of Operational Research*, 183, pp. 1109–1130.

[2] E.E. Bischoff and M.S.W. Ratcliff, (1995) Issues in the development of approaches to container loading, OMEGA, *International Journal of Management Science*, 23(4), pp. 377–390.

[3] C.H. Che, W. Huang, A. Lim and W. Zhu, (2011) The multiple container loading cost minimization problem, *European Journal of Operational Research*, 214(3), pp. 501–511.

[4] N. Ivancic, K. Mathur and B.B. Mohanty, (1989) An integer-programming based heuristic approach to the three-dimensional packing problem, *Journal of Manufacturing and Operations Management*, 2, pp. 268–298.

[5] R. Morabito and M. Arenales, (1994) An AND/OR-graph Approach to the Container Loading Problem, *International Transactions in Operational Research*, 1, pp. 59-73.

Inference for structured family

Cristina Dias[1], Carla Santos[2] and João T. Mexia[3]

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

E-mail: cpsilvadias@gmail.com

Abstract

The matrices of a structured family of stochastic symmetric matrices are all of the same order k and correspond to the treatments of a base design. The most interesting case is when the matrices in the family have a dominant first eigenvalue λ_1 see [2]. We then study the action of the factors in the base design, on the components of the first structure vector $\lambda_1\alpha_1$ with α_1 the first eigenvector. When the matrices in such families correspond to the treatments of a base design we can carry out ANOVA like analysis of the action of the treatments in the model on the structured vectors see [1] and [3]. This analysis can be transversal –when we worked width homologous components and longitudinal – when we consider contrast on the components of each structure vector. In this work we consider the models for these matrices and show how to carry out inference for structured family.

Keywords: symmetric matrices, ANOVA like analysis, longitudinal and transversal analysis.

Acknowledgements

This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID-MAT-00297-2013 (Centro de Matemática e Aplicações).

References

[1] Ito, P. K., (1980) Robustness of Anova and Macanova Test Procedures, P. R. Krishnaiah (ed), Handbook of Statistics, Amsterdam, North Holland, V1, pp. 199–236.
[2] Dias, C., (2013) Modelos e Famílias de Modelos para Matrizes Estocásticas Simétricas, Ph.D. Thesis, Évora University.
[3] Scheffé, H., (1959) The Analysis of Variance, John Wiley & Sons, New York.

Woodland caribou extinction risk in boreal and open taiga canadian forests

Carla Francisco[1], Teresa Oliveira[2] and Steve Cumming[1]

[1]Laval University, Department of Wood Science and Forestry, Faculty of Forestry, Geography and Geomatics, Québec, Canada
[2]Centro de Estatística e Aplicações (CEAUL), Portugal

E-mail: carla.francisco.1@ulaval.ca

Abstract

The goal of this project is to study a new way of simulating fire regimes that accounts for inter-annual variation in fire frequency and size. The importance of this variation will be evaluated through a model of caribou population size in relation to landscape disturbance history. Existing assessments of caribou population vulnerability are based in part on simulation studies of historical fire regimes that do not account for inter-annual variation in fire weather, that effects the number and mean size of fires. Population extinction risk is sensitive to the magnitude of environmental stochasticity, such as variation in the annual area burned by fire. If this variation is underestimated, so is the risk of extinction. A population viability analysis of select caribou herds will be conducted in order to determine the sensitivity of simulated population sizes and extinction risk area to these additional sources of variation. Several models exist relating range quality to caribou demographic parameters, in Sorensen et al (2006) [2], range quality is measured by the proportional area that has burned within the last 40 years. This proportion is related to the population growth rate by a linear regression model. In simulation studies, simple landscape fire models generate time-varying landscapes which cause inter-annual variation forest age structure, and so in population growth rates. In the course of this work simulations show that the risk of extinction for most herds is very low under historical fire regimes. However, this finding may be misleading because the landscape fire models used in the simulations do not account for inter-annual variation in the fire weather. A new landscape fire models that include statistical models of the inter-annual variation in the number and mean size of fires is being developed. Specifically, annual fire ignitions will be modelled as a Poisson process whose parameter varies randomly in time. Similarly, annual fire size will be modelled as an exponential distribution with random variation in the mean. These models can be estimated from time series of fire counts and sizes. A

estimation of the herd extinction risks is being developed under these new models by simulation experiments and to quantify the increase in these risks relative to the simpler models currently in use. In this project the data is from Manitoba region, 1969-2000. The research will be carried out using the data produced by the NRCAN (Natural Resources Canada) open database (Statistics, 2017) [3]. Using these findings, allow us to evaluate the adequacy of the limits on the amount of human disturbances within caribou ranges as given the ECCC recovery plan [1].

Keywords: landscape fire model, negative binomial count model, spatial simulation, wildfire.

Acknowledgements

Research partially funded by FCT Fundação para a Ciência e a Tecnologia, Portugal, through the projects PEst-OE-MAT-UI0006-2014, UID-MAT-00006-2013.

References

[1] Environment Canada, (2012) Recovery Strategy for the Woodland Caribou (Rangifer tarandus caribou),Boreal population, in Canada. Species at Risk Act Recovery Strategy Series, Environment Canada, Ottawaxi +138 pp.

[2] Sorensen et al, (2008) Determining sustainable levels of cumulative effects for boreal Caribou. Journal of Wildlife Management, 72(4), pp. 900–905.

[3] Statistics, C., (2017) Retrieved from statistics canada, Statistics canada, NRCAN, 219, https://www.nrcan.gc.ca/.

Combining data analysis methods for forecasting liquefied petroleum gas cylinders demand

Cristina Lopes[1], Eliana Costa e Silva[2], Aldina Correia[2], Magda Monteiro[3] and Rui Borges Lopes[4]

[1]LEMA, CEOS.PP, ISCAP, Polytechnic of Porto, Portugal
[2]CIICESI, ESTG, Polytechnic of Porto, Portugal
[3]CIDMA, ESTGA, University of Aveiro, Portugal
[4]CIDMA, DEGEIT, University of Aveiro, Portugal

E-mail: cristinalopes@iscap.ipp.pt

Abstract

In this paper we address a challenge proposed by a Portuguese energy sector company at an European Study Group with Industry. The company wanted to forecast the demand and return rate of Liquefied Petroleum Gas (LPG) cylinders. To address this challenge, sales data provided by the company as well as national data were analysed. The methodology proposed in the present work consisted in the combination of several techniques used to forecast demand and sales, in order to obtain a range of forecast values and their corresponding probability, similarly to [1]. Time series techniques (moving averages and exponential smoothing), multivariate linear regression models, and artificial neural networks were used to forecast total propane gas sales and return rates of cylinders (empty bottles). Vitullo et al. in [3] suggest that the most important factors influencing gas consumption are: temperature, prices, wind, previous month's demand, humidity, precipitation, and luminosity. Using ARIMA models, Erdogdu [2] forecasts the future growth in natural gas demand in Turkey, concluding that natural gas demand elasticities are quite low, i.e., prices do not influence demand significantly. In our work, taking into account explanatory variables such as atmospheric temperatures, demand in previous periods, sales objectives, and expectation of price increase, we estimated multiple regression models for the total sales of propane and for the number of LPG bottles at a national level. Artificial neural networks were used to forecast both the total propane gas sales and the return rate of cylinders (empty bottles). Finally, these methods were combined in a single approach, by defining a probability density function for each method and using Monte Carlo simulation to draw values, which are then used in a weighted linear function (with the weights proportional to the method's accuracy).

Keywords: data analysis, linear regression, neural networks, time series forecast, propane gas.

Acknowledgements

This work was partially supported from COST Action TD1409, Mathematics for Industry Network (MI-NET), supported by COST (European Cooperation in Science and Technology). Magda Monteiro and Rui Borges Lopes were partially supported by Portuguese funds through the CIDMA - Center for Research and Development in Mathematics and Applications, and the Portuguese Foundation for Science and Technology (FCT- Fundação para a Ciência e a Tecnologia), within project UID-MAT-04106-2013. We would like to thank Ana Sapata from University of Évora, and Cláudio Henriques, Fábio Henriques e Mariana Pinto from University of Aveiro for their contributions during the European Study Group.

References

[1] Cassettari, L., Bendato, I., Mosca, M. and Mosca, R., (2017) A new stochastic multi source approach to improve the accuracy of the sales forecasts, *Foresight*, 19(1), pp. 48–64.

[2] Erdogdu, E., (2010) Natural gas demand in turkey, *Applied Energy*, 87, pp. 211–219.

[3] Vitullo, S.R., Brown, R.H., Corliss, G.F. and Marx, B.M., (2009) Mathematical models for natural gas forecasting, *Canadian Applied Mathematics Quarterly*, 17(7), pp. 807–827.

Random sample sizes in orthogonal mixed models

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

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

E-mail: celian@ubi.pt

Abstract

In this work we aim to present a new approach considering orthogonal mixed models in situations 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. An illustrative example of this is the collection of observations during a fixed time period in a study comparing, for example, several pathologies of patients arriving at a hospital, see [1], [2], [3] and [4]. The applicability of the proposed approach is illustrated through an application based on real data, considering the incidence of unemployed persons in the European Union. The dataset gathers information regarding the age of the unemployed and comes from PORDATA (Base de Dados Portugal Contemporâneo).

Keywords: random sample sizes, orthogonal mixed models, unemployment in European Union.

Acknowledgements

This work was partially supported by national founds of FCT-Foundation for Science and Technology under UID-MAT-00212-2013 and UID-MAT-00297-2013.

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), pp. 551–561.

[3] Nunes, C., Capistrano, G., Ferreira, D. and Ferreira, S.S., (2013) ANOVA with Random Sample Sizes: An Application to a Brazilian Database on Cancer Registries, *11th International Conference on Numerical Analysis and Applied Mathematics*, AIP Conf. Proc., v.1558, pp. 825–828.

[4] 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), pp. 2316–2328.

Networks and interval maps

Carlos C. Ramos [1,2]

[1]Universidade de Évora, Departamento de Matemática, Portugal
[2]Centro de Investigação em Matemática e Aplicações (CIMA), Portugal

E-mail: ccr@uevora.pt

Abstract

The relation between interval map dynamics and digraphs is well known. The digraphs are built using partitions of the interval which supports a given map. Taking this relation has a starting point we introduce a systematic correspondence between networks and a certain class of interval maps. This correspondence can be seen as a coordinate system for networks. This approach has two perspectives: On one hand, methods established for network theory can be used for the study of certain characteristics of interval map dynamics. On the other hand, associating an interval map to a given network we may obtain a characterization of the network through the topological invariants of the map.

Keywords: dynamical system, iterated maps, graphs, topological invariants.

Acknowledgements

This work was supported by FCT - Fundação para a Ciência e a Tecnologia under the project UID-MAT-04674-2013.

References

[1] Bandeira, L. and Correia Ramos, C., (2016) Transition matrices characterizing a certain totally discontinuous map of the interval, *Journal of Mathematical Analysis and Applications*, Volume 444, Issue 2, pp. 1274–1303.

[2] Bandeira, L. and Correia Ramos, C., (2015) On the spectra of certain matrices and the iteration of quadratic maps, *SeMA J.*, 67, pp. 51–69.

[3] Correia Ramos, C., Martins, N. and Pinto, P.R., (2017) Toeplitz algebras arising from escape points of interval maps, *Banach J. Math. Anal.*, Volume 11, Number 3, pp. 536–553.

[4] Correia Ramos, C., Martins, N. and Pinto, P.R., (2013) On C*-algebras from interval maps, *Complex Analysis and Operator Theory*, Vol.7, pp. 221–235.

[5] Fernandes, S., Grácio, C. and Correia Ramos, C., (2013) Systoles in discrete dynamical systems, *Journal of Geometry and Physics*, Volume 63, pp. 129–139.

Estimation of variance components in linear mixed models under some constraints

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

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

E-mail: dario@ubi.pt

Abstract

It is usual that some constraints arise when one wants to carry out a study with data application. Often researchers ignore those constraints and use standard methods, such as the analysis of variance, to process the data. This may result in a loss of power. In this talk we will show how to estimate variance components in mixed models under some constraints. Namely we will show how to estimate variance components in mixed models in which the number of random factors is not greater than the number of eigenvalues of the model variance-covariance matrix. An example will be presented to illustrate this approach.

Keywords: linear mixed models, estimation, variance components.

Acknowledgements

This work was partially supported by national founds of FCT-Foundation for Science and Technology under UID-MAT-00212-2013 and UID-MAT-00297-2013.

References

[1] Bailey, R.A., Ferreira, S.S., Ferreira D. and Nunes, C., (2017) Estimability of variance components when all model matrices commute, *Lin. Alg. Appl.*, 492, pp. 144–160.

[2] Ferreira, D., Ferreira, S., Nunes, C., Fonseca, M., Silva, A. and Mexia, J.T., (2017) Estimation and incommutativity in mixed models, *Journal of Multivariate Analysis*, 61, pp. 58–67.

[3] Nelder, J.A., (1965) The analysis of randomized experiments with orthogonal block structure, I: Block structure and the null analysis of variance, *Proc. Royal Soc. A*, 273, pp. 147–162.

[4] Nelder, J.A., (1965) The analysis of randomized experiments with orthogonal block structure, II: Treatment structure and the general analysis of variance, *Proc. Royal Soc. A*, 273, pp. 163–178.

[5] Ferreira, S.S., Ferreira, D., Nunes, C. and Mexia, J.T., (2013) Estimation of Variance Components in Linear Mixed Models with Commutative Orthogonal Block Structure, *Revista Colombiana de Estadística*, 36(2), pp. 261–271.

Extremal index blocks estimator: choosing the block size and the threshold

Dora Prata Gomes[1,3] and Manuela Neves[2,4]

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

E-mail: dsrp@fct.unl.pt

Abstract

The main objective of Statistics of Extremes is the estimation of parameters of rare events. One of these parameters is the extremal index, θ, that is a key parameter when extending the analysis of the limiting behaviour of the extreme values from independent and identically distributted sequences to sationary sequences. The extremal index measures the degree of local dependence in the extremes of a stationary process, see [1], among other authors. Clusters of extreme values are linked with incidences and durations of catastrophic phenomena, an important issue in areas like environment, finance, insurance among others. Its estimation has been considered by several authors but some dificulties still remain. Most of the semi-parametric estimators show the same type of behaviour: nice asymptotic properties, but a high variance for small k, the number of upper order statistics used in the estimation; a high bias for large k and the need for an adequate choice of k. Here we focus on the estimation of θ using blocks estimators, introduced in [2]. Blocks estimators can be constructed by using disjoint blocks or sliding blocks. The asymptotic properties for both procedures were studied and compared in [3]. These authors show that the sliding blocks estimator is more efficient than the disjoint version and has a smaller asymptotic bias. However both blocks estimators require the choice of a threshold u and a block length b. In Figure 1 the resulting estimates based on exceedances over a high threshold $u_n = X_{n-k:n}$ where $X_{k:n}$ are the ascending order statistics associated to a random sample $(X_1, X_2, \ldots, X_n)$, are shown as a function of k, for different block lengths. Some criteria have appeared for the choice of those nuisance parameters, see, for example [3,4]. In this work we will show the effect of the threshold and the block size choice to obtain reliable estimates. A large simulation study have been performed and an application to daily mean flow discharge rate in the hydrometric station of Fragas da Torre

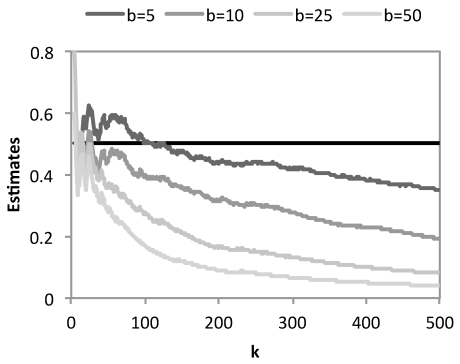


Fig. 1. Blocks-estimates as functions of k with block lengths $b = 5, 10, 20, 50$ for a *max-autoregressive* (ARMAX) process. The true extremal index value is here 0.5.

in Paiva river, data collected from 1 October, 1946 to 30 September, 2006 is shown.

Keywords: extreme value theory, stationary sequences, clusters of extreme values, daily mean flow discharge rate.

Acknowledgements

This research was partially funded by FCT–Fundação para a Ciência e a Tecnologia, Portugal, through the projects UID-MAT-00297-2013(CMA) and UID-MAT-00006-2013 (CEAUL).

References

[1] Leadbetter, M.R., Lindgren, G. and Rootzën, H., (1983) Extremes and related properties of random sequences and series, Springer-Verlag, New York.

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

[3] Robert, C.Y., Segers, J. and Ferro, C.A.T., (2009) A sliding blocks estimator for the extremal index, *Electronic Journal of Statistics*, Vol. 3, pp. 993–1020.

[4] Drees, H., (2011) Bias correction for estimators of the extremal index, *Preprint, arXiv: 1107.0935.*

Modelling of nonstationary extremes in women's hammer throw track and field competitions

Domingos Silva[1], Frederico Caeiro[2] and Manuela Oliveira[1,3]

[1]Centro de Investigação em Matemática e Aplicações, Universidade de Évora, Portugal
[2]Centro de Matemática e Aplicações, Universidade Nova de Lisboa, Portugal
[3]Departamento de Matemática, Universidade de Évora, Portugal

E-mail: domingosjlsilva@gmail.com

Abstract

Over the last century the number of women participating in sports competition has increased significantly, raising performances to levels never before imagined. This trend is an important covariable in the modelling of extreme events. The main purpose of this study is to apply extreme value models to a dataset of records from women's hammer throw in track and field competitions, collected from website https://en.wikipedia.org/wiki/Hammer_throw, and in some situations we consulted the IAAF compendium, about the progression of world athletics records [4]. Since stationarity is a basic requirement of the classical analysis of extremes, classic models may not be viable in certain sports contexts, particularly in events strongly affected by change and increasing evolution over time. To date no general theory has been established for nonstationary extremes, reason why an alternative is to consider that one or more parameters of the distribution vary over time. Therefore in this study we consider stationary and nonstationary models, for the estimation of extreme parameters in women's hammer throw. We considered the annual block maxima (BM) method for the period from 1988 to 2017 (n=30) and the generalized extreme value distribution with constant parameters and 7 different time dependent parameters. The BM method may be preferable, over other methods (e.g., POT), when the observations are not exactly independent and identically distributed [5–7,9]. The trend were verified by graphics analysis and, formally, by Mann-Kendall trend test [8,10]. The Deviance statistics, as well as the Akaike information criterion (AIC) and corrected AIC were used to select the model that best represents the available data [1–3]. The best fit model has a linear trend in location parameter, log-linear in the scale parameter and constant shape parameter. There is a strong probability of a new world record happening in the next three years.

Keywords: extreme value theory, block maxima, generalized extreme value distribution, maximum likelihood estimation, nonstationary, exceedance probability, upper right tail, return levels, hammer throw (track and field).

Acknowledgements

This research was partially funded by FCT, Fundação para a Ciência e a Tecnologia, Portugal, through the projects UID-MAT-04674-2013 (CIMA) (Centro de Investigação em Matemática e Aplicações) and project UID-MAT-00297-2013 (Centro de Matemática e Aplicações).

References

[1] Akaike, H., (1974) A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, 19(6), pp. 716–723.

[2] Bozdogan, H., (1987) Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions, *Psychometrica*, 52(3), pp. 345–370.

[3] Coles, S., (2001) An Introduction to Statistical Modeling of Extreme Values, Springer-Verlag, London., pp. 45–73 and pp. 105–123.

[4] Richard Hymans, (2015) IAAF - International Association of Athletics Federations, Progression of IAAF World Records, Edited by (ATFS) and Imre Matrahazi (IAAF), Printed by Multiprint, Monaco.

[5] Ferreira, A. and de Haan, L., (2015) On the block maxima method in extreme value theory: PWM estimators, *The Annals of Statistics*, 43(1), pp. 276–298.

[6] Katz, R.W., Parlange, M.B. and Naveau, P., (2002) Statistics of extremes in hydrology, *Advances in Water Resources*, 25, pp. 1287–1304.

[7] Katz, R.W., (2013) Statistical Methods for Nonstationary Extremes. In: Extremes in a Changing Climate: Detection, Analysis and Uncertainty, Chapter 2. AghaKouchak et al. (eds.), Springer Science + Business media Dordrecht.

[8] Kendall, M.G., (1975) Rank Correlation Methods, 4th ed., Charles Griffin, London.

[9] Madsen, H., Rasmussen, P.F. and Rosbjerg, D., (1997) Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events 1. At-site modeling. *Water Resources Research*, 33(4), pp. 747–757.

[10] Mann, H.B., (1945) Nonparametric test against trend, *Econometrica*, 13, pp. 245–259.

Numerical simulations of a third-grade fluid flow on a tube through a contraction

Fernando Carapau[1] and Paulo Correia[1]

[1]Universidade de Évora, Departamento de Matemática e CIMA, Portugal

E-mail: flc@uevora.pt

Abstract

Based on a director theory approach related to fluid dynamics we reduce the nonlinear three-dimensional equations governing the axisymmetric unsteady motion of a non-Newtonian incompressible third-grade fluid to a one-dimensional system of ordinary differential equations depending on time and on a single spatial variable. From this new system we obtain the unsteady equation for the mean pressure gradient and the wall shear stress both depending on the volume flow rate, Womersley number and viscoelastic parameters over a finite section of a straight, rigid and impermeable tube with variable circular cross-section. We present some numerical simulations of unsteady flows regimes through a tube with a contraction using a nine-directors theory.

Keywords: third-grade fluid, one-dimensional model, unsteady flow, hierarchical theory.

Acknowledgements

This work was supported by FCT - Fundação para a Ciência e a Tecnologia under the project UID-MAT-04674-2013.

References

[1] Carapau, F. and Correia, P., (2017) Numerical simulations of a third-grade fluid flow on a tube through a contraction, *European Journal of Mechanics B/Fluids*, 65, pp. 45–53.

[2] 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), pp. 190–196.

[3] Fosdick, R.L. and Rajagopal, K.R., (1980) Thermodynamics and stability of fluids of third grade, *Proc. R. Soc. Lond. A.*, 339, pp. 351–377.

Poincaré plot in gait variability analysis

Flora Ferreira[1], Miguel Gago[2,3], Estela Bicho[4], Nuno Sousa[3], João Gama[1] and Carlos Ferreira[1]

[1]Laboratório de Inteligência Artificial e Apoio à Decisão, INESC TEC, Portugal
[2]Serviço de Neurologia, Hospital da Senhora da Oliveira, Portugal
[3]ICVS, Escola de Medicina, Universidade do Minho, Portugal
[4]Centro Algoritmi, Escola de Engenharia, Universidade do Minho, Portugal

E-mail: flora.ferreira@gmail.com

Abstract

Poincaré plot is a two-dimensional geometric representation of a time series, where each point represents a pair of successive elements of the time series [1]. The resulted scatter plot allows assessment the element-to-element variability and the overall variation. A conventional Poincaré plot is analyzed quantitatively by determining two measures of point dispersion (SD1 and SD2) that there are closely related to short- and long-term variability of the time series. Poincaré plot have been widely used in the cardiovascular area to measure heart rate variability and their application on gait time series is emerging [2]. Gait stride-to-stride variability is an important marker to better understand the mechanisms of movement disorders, and in monitoring the progression of the disease under therapeutic interventions [3]. Gait variability is often reported by standard deviation and/or coefficient of variation [4], and few studies used Poincaré Plot analysis to compare gait variability of healthy subjects from patients with neurodegenerative disease, such as Parkinson's and Huntington diseases [5], and further research is needed. The aim of this study is to evaluate the stride-to-stride fluctuations through the measures derived from the Poincaré plot to help understand the neural control of locomotion in patients with neurodegenerative disease and health control subjects. For this proposed, we analyzed different gait time series that represent different gait characteristics (stride time, stance, swing, double support) from patients with Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis and healthy control subjects. All time series were obtained from the database available at https://physionet.org/physiobank/database/gaitndd/.

Keywords: time series, short-term variability, long-term variability, movement disorders.

Acknowledgements

This work was partially supported by the projects NORTE-01-0145-FEDER-000016 (NanoSTIMA) and NORTE-01-0145-FEDER-000026 (DeM), financed by the Regional Operational Program of the North (NORTE2020), under PORTUGAL2020 and FEDER.).

References

[1] Brennan M., Palaniswami M. and Kamen PW., (2001) Do existing measures of Poincaré plot geometry reflect nonlinear features of heart rate variability?, *IEEE Trans Biomed Eng*, v.48, pp. 1342–1347.

[2] Hollman, J.H., Watkins, M.K., Imhoff, A.C., Braun, C.E., Akervik, K.A. and Ness, D. K., (2016) A comparison of variability in spatiotemporal gait parameters between treadmill and over ground walking conditions, *Gait & posture*, v.43, pp. 204–209.

[3] Bryant, M.S., Rintala, D.H., Hou, J.G., Charness, A.L., Fernandez, A.L., Collins, R.L., Baker, J., Lai, E.C. and Protas, E.J., (2011) Gait variability in Parkinson's disease: influence of walking speed and dopaminergic treatment, *Neurological research*, v.33, n.9, pp. 959–964.

[4] Moon, Y., Sung, J., An, R., Hernandez, M.E. and Sosnoff, J.J., (2016) Gait variability in people with neurological disorders: a systematic review and meta-analysis, *Human movement science*, v.47, pp. 197–208.

[5] Golińska, A.K., (2013) Poincaré plots in analysis of selected biomedical signals, *Studies in logic, grammar and rhetoric*, v.35, n.1, pp. 117–127.

How to test different block diagonal structures in several covariance matrices

Filipe J. Marques[1,2] and Carlos A. Coelho[1,2]

[1]Centro de Matemática e Aplicações (CMA), FCT, UNL, Portugal
[2]Departamento de Matemática, FCT, UNL, Portugal

E-mail: fjm@fct.unl.pt

Abstract

The adequate calibration of different statistical models, such as growth curve or mixed models, requires an adequate choice of the covariance matrix structure. The complexity of some new models makes it important to be able to choose and to test elaborate patterns for covariance matrices. However, the testing procedures commonly used in this decision process are not easy to implement due to the complicated structure of the exact distributions of the test statistics involved. Therefore, the required tests are often not performed or rather are performed using approximations for the distributions of the test statistics which, in most of the cases, are unable to guarantee the necessary accuracy of the results. In this work we will show how it is possible to develop a procedure to test different block diagonal structures in several covariance matrices by splitting the null hypothesis into a set of conditionally independent hypotheses [1,3] and how does this procedure makes it easy the development of near-exact approximations [2]. The numerical studies carried out demonstrate the adequacy and accuracy of these approximations.

Keywords: covariance structures, generalized integer gamma distribution, generalized near-integer gamma distribution, mixtures.

Acknowledgements

This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID-MAT-00297-2013 (Centro de Matemática e Aplicações).

References

[1] Correia, B.R., Coelho, C.A. and Marques, F.J., (2018) Likelihood ratio test for the hyper-block matrix sphericity covariance structure – characterization of the exact distribution and development of near-exact distributions for the test statistic, *Revstat*, to appear.

[2] 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*, v. 89, pp. 191–218.

[3] Coelho, C.A. and Marques, F.J., (2009) The advantage of decomposing elaborate hypotheses on covariance matrices into conditionally independent hypotheses in building near-exact distributions for the test statistics, *Linear Algebra and its Applications*, v. 430, pp. 2592–2606.

Combining sensory and chromatographic analyses in acceptance sampling plans

Fernanda Figueiredo[1,3], Adelaide Figueiredo[2,3] and M. Ivette Gomes[1,4]

[1]CEAUL, Centro de Estatística e Aplicações, Universidade de Lisboa, Portugal
[2]LIAAD INESC TEC Porto, Universidade do Porto, Portugal
[3]Faculdade de Economia, Universidade do Porto, Portugal
[4]Departamento de Estatística e Investigação Operacional, Faculdade de Ciências, UL, Portugal

E-mail: otilia@fep.up.pt

Abstract

In certain areas of activity the quality of a product or service is not only a differentiation component that leads consumers to choose it or another competitor, but more than this, because there are standards in a developed and competitive market framework that have to be necessarily followed. For instance, some manufacturing producers as well as food and services industries, are faced with the obligation of conducting a tight quality control to their products, to detect and measure the intensity of abnormal characteristics, such as off-odors, chemical substances, off-flavors and taints, among others. The tools of statistical quality control they commonly use are the control charts for on-line inspection, and the acceptance sampling plans to decide for the acceptance or rejection of lots of raw material, and of not-finished and final products. Details about these powerful tools can be found in [1], [2] and [3]. In particular, for the detection of chemical substances, in addition to sensory evaluation, even when qualitative (presence or not) and quantitative (intensity) information is obtained, it is usually necessary to perform chromatographic analysis. In this paper we consider some acceptance sampling plans based on both sensory and chromatographic analyses, to investigate the presence (or not) of chemical substances in lots of items on the basis of the observed sample(s), and consequently, reject (or accept) such lots. The sensory evaluation, the first applied procedure to the items of the sample, is conducted by expert assessors, who must give qualitative information (Yes/No) about the identification of a chemical substance in each analyzed item. Accordingly to predefined decision rules, the lot can be accepted or rejected at this step, without subsequent analysis, or one must continue with a quantitative chromatographic evaluation. It is important to refer that most of the chromatographs in use do not register with suitable

accuracy very small or large concentrations of chemical substances, and a common practice is to truncate the results below or above a certain threshold, leading us to consider a truncated and inflated distribution to represent the chromatographic measurements. In this study such values are modeled by an inflated Pareto distribution (also considered in related works; see [4] and [5]), and some different statistics are used in the decision rule at this step, in order to decide for the acceptance or rejection of the corresponding lot of items. Some guidelines for the implementation of such plans in order to achieve predefined risk levels is provided, and the performance of the proposed sampling plans is evaluated and compared. An application of such plans to a real data set is also presented. Other details and related works on sensory testing and acceptance sampling plans can be found, for instance, in [6], [7], [8] and [9].

Keywords: acceptance sampling plans, chromatographic analysis, inflated Pareto distribution, sensory analysis.

Acknowledgements

Research partially supported by National Funds through **FCT**—Fundação para a Ciência e a Tecnologia, projects UID-MAT-0006-2013 (CEA/UL) and UID-EEA-50014-2013, and by COST Action IC1408.

References

[1] Carolino, E. and Barão, I., (2013) Robust methods in acceptance sampling, *Revstat* v.11, n.1, pp. 67–82.

[2] Gomes, M.I., (2011) Acceptance sampling, *In Lovric, M. (ed.), International Encyclopedia of Statistical Science*, Springer, Part 1, pp. 5-7.

[3] Montgomery, D.C., (2009) Introduction to Statistical Quality Control: a Modern Introduction, John Wiley & Sons, 6th ed.

[4] Figueiredo, F., Figueiredo, A. and Gomes, M.I., (2015) Acceptance sampling plans for inflated Pareto Processes, *Notas e Comunicações, CEAUL 04/15.*

[5] Figueiredo, F., Figueiredo, A. and Gomes, M.I., (2018) Acceptance sampling plans for reducing the risk associated with chemical compounds, *In Teresa A. Oliveira et al. (eds.), Recent Studies on Risk Analysis and Statistical Modeling*, Springer.

[6] Figueiredo, F., Figueiredo, A. and Gomes, M.I., (2014) Comparison of Sampling Plans by Variables using the Bootstrap and Monte Carlo Simulations, *AIP Conference Proceedings*, v.1618, pp. 535–538.

[7] Figueiredo, F., Figueiredo, A. and Gomes, M.I., (2018b) Design of Sampling Plans for Sensory Evaluation, *AIP Conference Proceedings* (in print).

[8] Meilgaard, M.C., (1991) Current progress in sensory analysis. A review, *American Society of Brewing Chemists*, v.49, n.3, pp. 101–109.

[9] Munoz, A.M., Civille, G.V. and Carr., B.T., (1992) Sensory Evaluation in Quality Control, New York, Van Nostrand Reinhold.

Some results on the determinantal range of matrix products

G. Soares[1,2], R. Lemos[3] and A. Guterman[4]

[1]Pole CMAT-UTAD, University of Trás-os-Montes and Alto Douro, Portugal
[2]Universidade de Trás-os-Montes e Alto Douro, UTAD, Escola das Ciências e Tecnologia, Portugal
[3]CIDMA, Mathematics Department, University of Aveiro, Portugal
[4]Faculty of Algebra, Department of Mathematics and Mechanics, Moscow State University, Russia

E-mail: gsoares@utad.pt

Abstract

Let matrices $A, C \in M_n$ have eigenvalues $\alpha_1, \ldots, \alpha_n$ and $\gamma_1, \ldots, \gamma_n$, respectively. The set $D_C(A) = \{\det(A - UCU^*) : U \in M_n, U^*U = I_n\}$ of complex numbers is called the $C-$*determinantal range* of A. We study different conditions under which it holds that $D_C(RS) = D_C(SR)$, for some matrix words R and S.

Keywords: numerical range, $\sigma-$points, Marcus-Oliveira conjecture.

Acknowledgements

The work of the first author was financed by Portuguese Funds through FCT- Fundação para a Ciência e Tecnologia, within the Project UID-MAT-00013-2013. The work of the second author was supported by Portuguese funds through the Center for Research and Development in Mathematics and Applications (CIDMA) and the Portuguese Foundation for Science and Technology (FCT-Fundação para a Ciência e a Tecnologia), within project UID-MAT-04106-2013. The work of the third author is partially financially supported by RSF grant 17-11-01124.

References

[1] N. Bebiano, (1986) Some analogies between the $c-$numerical range and a certain variation of this concept, *Linear Algebra Appl.*, 81, pp. 47–54.

[2] N. Bebiano, Yiu-Tung Poon and J. da Providência, (1988) On $C-$det spectral and $C-$det convex matrices, *Linear and Multilinear Algebra*, 23, pp. 343–351.

[3] N. Bebiano and J. da Providência, (1988) Some remarks on a conjecture of Oliveira, *Linear Algebra Appl.*, 102, pp. 241–246.

[4] N. Bebiano, J. K. Merikoski and J. da Providência, (1987) On a conjecture of G.N. de Oliveira on determinants, *Linear and Multilinear Algebra*, 20, pp. 167–170.
[5] N. Bebiano, (1994) New developments on the Marcus-Oliveira conjecture, *Linear Algebra Appl.*, 197, 198, pp. 791–844.
[6] N. Bebiano, A. Kovačec and J. da Providência, (1994) The validity of the Marcus-Oliveira conjecture for essentially Hermitian matrices, *Linear Algebra Appl.*, 197, 198, pp. 411–427.
[7] N. Bebiano and G. Soares, (2005) Three observations on the determinantal range, *Linear Algebra Appl.*, 401, pp. 211–220.
[8] W.S. Cheung and N.K. Tsing, (1996) The $C-$numerical range of matrices is star-shapped, *Linear and Multilinear Algebra*, 41, pp. 245–250.
[9] M.T. Chien, C.L. Ko and H. Nakazato, (2010) On the numerical ranges of matrix products, *Applied Mathematics Letters*, 23, pp. 732–737.
[10] S.W. Drury and B. Cload, (1992) On the determinantal conjecture of Marcus and de Oliveira, *Linear Algebra Appl.*, 177, pp. 105–109.
[11] M. Fiedler, (1971) Bounds for the determinant of the sum of hermitian matrices, *Proc. Amer. Math. Soc.*, 30, pp. 27–31.
[12] P. Gawron, Z. Puchala, J.A. Miszczak, L. Skowronek and K. Zyczkowski, (2010) Restricted numerical range: a versatile tool in the theory of quantum information, *J. Math. Physics*, 51, (24 p.) 02204.
[13] M. Goldberg and E.G. Straus, (1977)Elementary inclusion relations of generalized numerical ranges, *Linear Algebra Appl.*, 18, pp. 1–24.
[14] A. Guterman, R. Lemos and G. Soares, (2013) Extremal case in Marcus–Oliveira conjecture and beyond, *Linear and Multilinear Algebra*, 61(9), pp. 1206–1222.
[15] F. Hausdorff, (1919) Der Wertvorrat einer Bilinearform, *Math. Z.*, 3, pp. 314–316.
[16] R.A Horn and C. R. Johnson, (1985) Matrix analysis, Cambridge university press, Cambridge.
[17] R. A Horn and C. R. Johnson, (1991) Topics in matrix analysis, Cambridge university press, New York.
[18] A. Kovačec, (1994) On a conjecture of Marcus and de Oliveira, *Linear Algebra Appl.*, 201, pp. 91–97.
[19] C.K. Li, (1987) The $C-$ convex matrices, *Linear and Multilinear Algebra*, 21, pp. 303–312.
[20] C.K. Li, ((1994) $C-$numerical ranges and $C-$numerical radii, Special Issue: The numerical range and numerical radius, *Linear and Multilinear Algebra*, 37, no. (1)(3), pp. 51–82.
[21] C.K. Li, Y.T. Poon and N.S. Sze, (2008) Ranks and determinants of the sum of matrices from unitary orbits, *Linear and Multilinear Algebra*, 56, no. (1)(2), pp. 105–130.
[22] F.D. Murnaghan, (1932) On the field of values of a square matrix, *Proc. Nat. Acad. Sci.*, 18, pp. 246–248.
[23] M. Marcus, (1973) Derivations, Plücker relations and the numerical range, *Indiana Univ. Math. J.*, 22, pp. 1137–1149.
[24] F.D. Murnaghan, (1932) On the field of values of a square matrix, *Proc. Nat. Acad. Sci.*, 18, pp. 246–248.
[25] J.K. Merikoski and A. Virtanen, (1989) Some notes on de Oliveira's determinantal conjecture, *Linear Algebra Appl.*, 121, pp. 345–352.
[26] G.N. de Oliveira, (1982) Normal matrices (Research Problem), *Linear and Multilinear Algebra*, 12, pp. 153–154.
[27] O. Toeplitz, (1918) Das algebraische Analogon zu einem Satz von Fejer, *Math. Z.*, 2, pp. 187–197.
[28] R. Westwick, (1975) A theorem on numerical ranges, *Linear and Multilinear Algebra*, 2, pp. 311–315.

Human-likeness statistical comparison of simulated robotic arm reaching movements

Gianpaolo Gulletta[1], Eliana Costa e Silva[2], Wolfram Erlhagen[3] and Estela Bicho[1]

[1]Centre Algoritmi/Dept. of Industrial Electronics, University of Minho, Portugal
[2]CIICESI, ESTG, Polytechinic of Porto, Portugal
[3]Centre of Mathematics, Dept. of Mathematics, University of Minho, Portugal

E-mail: eos@estg.ipp.pt

Abstract

Human-like morphology and movements are often considered key features for achieving high levels of cooperation between robots and humans. In fact, robots are increasing becoming part of our daily life in several assistance scenarios. For natural human-robot interactions human-like movements are essential not only for human acceptance but also because they allow an easier interpretation of the movements of the robot as goal-directed actions. Inspired by studies in human motor control [1], we have developed a movement planning for movements of an anthropomorphic robotic arm and hand of the ARoS robotic platform [2–4]. The question that arises is on how to measure human-likeness of robotic arm movements. Several metrics have been used in literature, and the mostly used are the normalized jerk square and the -1/6 power law (see e.g.[5]). In this talk we present the statistical comparison of reaching movements using the six different planers: the Human-like Upper-limb Motion Planner (HUMP); and five popular sampling-based robotic planners, namely RRT, RRTConnect, PRM, RRT* and PRM* [6]. A total of 2400 random movements were using V-REP (Virtual Robot Experimentation Platform) on the transverse, sagittal and coronal planes, as well as 3D movements. Non-parametric Kruskal-Wallis tests were used for comparing the human-likeness of the movements generated by each planer.

Keywords: robotic arm movements, random simulated movements, V-REP, Kruskal-Wallis.

Acknowledgements

This work was partially supported by the EU Project PF7 Marie Curie NETT - Neural Engineering and Transformative Technologies and FCT PhD grant (ref. SFRH-BD-114923-2016). We would like to thank the Mobile and Anthropomorphic Robotics Laboratory at University of Minho for constant good work environment and frequent discussions.

References

[1] Rosenbaum, D.A., Meulenbroek, R.J., Vaughan, J. and Jansen, C., (2011) Posture-based motion planning: Applications to grasping, *Psychological Review*, v.108, pp. 709–734.

[2] Bicho, E., Erlhagen, W., Louro, L. and Costa e Silva, E., (2011) Neuro-cognitive mechanisms of decision making in joint action: A human–robot interaction study, *Human Movement Science*, v.30, pp. 846–868.

[3] Costa e Silva, E., Costa, F., Bicho, E. and Erlhagen, W., (2011) Nonlinear Optimization for Human-Like Movements of a High Degree of Freedom Robotics Arm-Hand System, *LNCC Computational Science and Its Applications - ICCSA*, Ed. Beniamino Murgante, Springer Berlin Heidelberg, v.6784, pp. 327–342.

[4] Costa e Silva, E., Costa, F., Erlhagen, W. and Bicho, E., (2015) Global vs. local nonlinear optimization techniques for human-like movement of an anthropomorphic robot, *AIP Conference Proceedings*, v.1648, pp. 140004.

[5] Chang, J.J., Yang, Y.S., Guo, L.Y, Wu, W.L. and Su, F.C., (2008) Differences in reaching performance between normal adults and patients post stroke a kinematic analysis, *Journal of Medical and Biological Engineering*, v.28, pp. 53–58.

[6] LaValle, S.M., (2006) Planning Algorithm, Cambridge University Press, New York.

Asymptotic and finite sample comparison of some extreme value index classes of estimators

Helena Penalva[1,3], Ivette Gomes[1,4], Frederico Caeiro[2,5] and Manuela Neves[1,6]

[1]Centro de Estatística e Aplicações, Universidade de Lisboa, Portugal
[2]Centro de Matemática e Aplicações, Universidade Nova de Lisboa, Portugal
[3]Instituto Politécnico de Setúbal, Portugal
[4]Faculdade de Ciências, Universidade de Lisboa, Portugal
[5]Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal
[6]Instituto Superior de Agronomia, Universidade de Lisboa, Portugal

E-mail: manela@isa.ulisboa.pt

Abstract

In *Extreme Value Theory* we are essentially interested in the estimation of quantities related to extreme events, and its main issue has been the estimation of the extreme value index (EVI), a parameter directly related to the tail weight of the distribution. When we are interested in large values, estimation is usually performed on the basis of the largest k order statistics in the sample or on the excesses over a high level u. In this talk we deal with the semi-parametric estimation of the EVI, for heavy tails, beginning as usual, by reviewing classical estimators. Most of those estimators show the same type of behaviour: nice asymptotic properties, but a high variance for small values of k and a high bias for large k, and therefore the need for an adequate choice of k. Some classes of EVI-estimators have appeared in the literature in order to overcome that difficulty, as those studied in [1–6], to cite a few works. The class of mean-of-order-p (MO_p) EVI-estimators [3,7], based on the *Hölder's mean-of-order-p*, revealed very nice properties, showing a mean square error smaller than that of the classical EVI-estimators, even for small values of k. Recently, a new class of EVI-estimators, the L_p EVI-estimators, based on the *Lehmer's mean-of-order p* that generalizes the arithmetic mean, was derived, see [8,9]. The study of the asymptotic behaviour of this class and the asymptotic comparison, at optimal levels to classical estimators or the members of other classes reveals that for an optimal (p, k)-choice, in the sense of minimazing the mean square error, the members of this class are able to show a very good performance, see [10]. Those estimators are also compared for finite samples, through a large simulation study.

Keywords: generalized means, heavy tails, semi-parametric estimation, statistics of extremes.

Acknowledgements

Research partially supported by National Funds through **FCT** – Fundação para a Ciência e a Tecnologia, projects UID-MAT-00006-2013 (CEA/UL), PEst-OE-MAT-UI0297-2013 (CMA/UNL), and COST Action IC1408.

References

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

[2] Caeiro, F. and Gomes, M.I., (2002) Bias reduction in the estimation of parameters of rare events, *Theory of Stochastic Processes*, 8(24), pp. 345–364.

[3] Brilhante, M.F., Gomes, M.I. and Pestana, D., (2013) A simple generalization of the Hill estimator, *Computat. Statistics and Data Analysis*, 57(1), pp. 518–535.

[4] Paulauskas, V. and Vaičiulis, M., (2013) On the improvement of Hill and some others estimators, *Lith. Math. J.*, 53, pp. 336–355.

[5] Paulauskas, V. and Vaičiulis, M., (2017) A class of new tail index estimators, *Annals of the Institute of Statistical Mathematics*, 69, pp. 661–487.

[6] Beran, J., Schell, D. and Stehlík, M., (2014) The harmonic moment tail index estimator: asymptotic distribution and robustness, *Ann. Inst. Statist. Math.*, 66, pp. 193–220.

[7] Brilhante, M.F., Gomes, M.I. and Pestana, D., (2014) The mean-of-order p extreme value index estimator revisited, *In* A. Pacheco et al. (Eds.), *New Advances in Statistical Modeling and Application*, Springer-Verlag, Berlin, Heidelberg, pp. 163–175.

[8] Penalva, H., (2017) Contributos Computacionais e Metodológicos na Estimação do Índice de Valores Extremos, Tese de Doutoramento, Instituto Superior de Agronomia, ULisboa.

[9] Penalva, H., Caeiro, F., Gomes, M.I. and Neves, M., (2016) An Efficient Naive Generalization of the Hill Estimator—Discrepancy between Asymptotic and Finite Sample Behaviour, Notas e Comunicações, CEAUL 02/2016.

[10] 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*.

A special subspace of Hessenberg-Type matrices

Henrique F. da Cruz[1,2], Ilda Inácio Rodrigues[1,2], Rogério Serôdio[1,2], A.M. Simões[1,2,3] and José Velhinho[1]

[1]Faculdade de Ciências, Universidade da Beira Interior, Covilhã, Portugal
[2]Centro de Matemática e Aplicações (CMA-UBI), Faculdade de Ciências, Universidade da Beira Interior, Covilhã, Portugal
[3]Center for Research and Development in Mathematics and Applications (CIDMA), Universidade de Aveiro, Campus Universitário de Santiago, Aveiro, Portugal

E-mail: ilda@ubi.pt

Abstract

We present and characterize subspaces of generalized Hessenberg matrices such that the determinant is convertible into the permanent by affixing $\pm$ signs. An explicit characterization of convertible Hessenberg-type matrices is described. In the end, we conclude that convertible matrices with the maximum number of nonzero entries can be reduced to a basic set.

Keywords: determinant, permanent, Hessenberg matrix, convertible matrix.

Acknowledgements

This work was supported in part by FCT–Portuguese Foundation for Science and Technology through the Center of Mathematics and Applications of University of Beira Interior, within project UID-MAT-00212-2013.

References

[1] Pólya, G., (1913) Aufgabe 424, *Arch. Math. Phys.*, n.20, pp. 271.
[2] Szegö, G. and Lösung Zu, (1913) Aufgabe 424, *Arch. Math. Phys.*, n.21, pp. 291–292.
[3] Gibson, P.M., (1971) Conversion of the permanent into the determinant, *Proc. Am. Math. Soc.*, n.27, pp. 471–476.
[4] Little, C.H.C., (1975) A characterization of convertible $(0,1)$-matrices, *J. Comb. Theory Ser. B*, n.18, pp. 187–208.
[5] Kasteleyn, P.W., (1967) Graph theory and crystal physics, *Graph Theory and Theoretical Physics, Harary, F., Ed.*, Academic Press: New York, USA, pp. 43–110.
[6] Vazirani, V.V. and Yannakakis, M., (1989) Pfaffian orientations, 0-1 permanents, and even cycles in directed graphs, *Discret. Appl. Math.*, n.25, pp. 179–190.
[7] Robertson, N., Seymour, P.D. and Thomas, R., (1999) Permanents, Pfaffian orientations, and even directed circuits, *Ann. Math.*, n.150, pp. 929–975.
[8] da Fonseca, C.M., (2011) An identity between the determinant and the permanent of Heissenberg-type matrices, *Czechoslov. Math. J.*, n.61, pp. 917–921.
[9] Gibson, P.M., (1969) An identity between permanents and determinants, *Am. Math. Mon.*, n.76, pp. 270–271.

Kalman filtering, a necessary tool to build a low cost navigation system

José Vieira Duque[1], Victor Plácido da Conceição[1] and M. Filomena Teodoro[1,2]

[1]CINAV, Center of Naval Research, Portuguese Naval Academy, Portuguese Navy, Portugal
[2]CEMAT, Center for Computational and Stochastic Mathematics, Instituto Superior Técnico, Lisbon University, Portugal

E-mail: maria.alves.teodoro@marinha.pt

Abstract

low-cost positioning and navigation systems emerged due the significant development of the technology in last twenty years, in particular with the appearance of the Micro-Electro-Mechanical Systems (MEMS)[1–3]. The main idea described in the manuscript consists in the construction of a low-cost positioning solution for small sailboats. The Kalman filtering is used to deal with data from a MEM, an inertial sensor and a GPS receiver staged on small sailing vessels. As preliminary approach, we have considered a simpler filter of Kalman where the state (signal) is defined with few variables. The validation of this low-cost equipment is done comparing the obtained results with those obtained by higher precision systems. This work is still in a preliminary stage. Due to its dimensions and costs, it is difficult to install conventional navigation systems in small boats, this work intends to give an answer to this problem using a technique widely used as a tool of excellence in signal processing, the Kalman filters (KF), see, for example [4,5]. In 1960, the engineer Rudolf Kalman published an article [4] in which he presented a new method of linear filtration. This method uses the measurements of independent variables and the noise from these measurements to filter the system signal and predicts its next state using some statistical techniques. This new method introduced by Kalman in early sixty decade came to be known as Kalman Filter (KF) and had its first use aboard the spacecraft navigation computers of the APOLLO project. In KF, the signal processing is based on stochastic models, estimation and control [6,7]. In order to create a low-cost positioning solution that is applied to small sailboats, this work, based on the different KF approaches described in [8], uses a properly adjusted KF capable of merging data from a MEMS inertial sensor and a GPS receiver. The system intends to present a solution that allows to know the position and speed of the boat and its behavior in the three main axes, yaw,

pitch and roll. At the moment, our preliminary approach, firstly introduced in [9,10], conduced to a filter that provides the position and speed, but only one of the mains axes inclination.

Keywords: systems of navigation and positioning, Kalman filters, low-cost, small vessels.

Acknowledgements

This work was supported by Portuguese funds through the *Center for Computational and Stochastic Mathematics* (CEMAT), *The Portuguese Foundation for Science and Technology* (FCT), University of Lisbon, Portugal, project UID-Multi-04621-2013, and *Center of Naval Research* (CINAV), Naval Academy, Portuguese Navy, Portugal.

References

[1] Prime Faraday Partnerships, (2002) An Introduction to MEMS (Microelectromechanical Systems), Prime Faraday Partnerships, Loughborough University, Loughborough, Accessed in 28 April 2018 from http://www.amazon.co.uk/ exec/ obidos/ASIN/1844020207.

[2] What is MEMS Technology? (n.d.). Accessed in 28 April 2018 from https:/ www.mems-exchange.org/ MEMS/ what-is.html.

[3] Ganssle, J., (2012) A Designers Guide to MEMS Sensors, Convergence Promotions LLC, Rhode Island, Accessed in 28 April 2018 from https://www.digikey.com / en /articles / techzone/ 2012/ jul/ a-designers -guide -to-mems-sensors.

[4] Kalman, R.E., (1960) A New Approach to Linear Filtering and Prediction Problems, *Journal of Basic Engineering*, 82(1), pp. 35–45.

[5] Sorenson, H.W., (1985) Kalman Filtering: Theory and Application, IEEE Press, New Jersey.

[6] Maybeck, P.S., (1979) Stochastic models, estimation and control, v.1, Academic Press, New York.

[7] Kay, S.M., (1993) Fundamentals of Statistical Signal Processing, Estimation Theory, Prentice Hall, New Jersey.

[8] Brown, R.G. and Hwang, Y.C., (2012) Introduction to random signals and applied Kalman filtering: with MATLAB exercises, 4^{th} edition, John Wiley & Sons, New Jersey.

[9] Duque, J.V., Conceição, V.P. and Teodoro, M.F., (2018) Filtragem de Kalman aplicada a sistemas de navegação de baixo custo, In Ferreira, C. and et al. (Eds.), XXV Jornadas de Classificação e Análise de Dados (JOCLAD 2018), Livro de Resumos, INE, Lisboa.

[10] Duque, J.V., Conceição, V.P. and Teodoro, M.F., (2018) Kalman Filtering Applied to Low-cost Navigation Systems: a Preliminary Approach, In Gervasi, O. et al. (Eds.), Computational Science and Its Applications ICCSA, *Lecture Notes in Computer Science* (to appear).

Country-level population structures. Insights from symbolic data clustering

José G. Dias[1]

[1]Instituto Universitário de Lisboa (ISCTE-IUL), BRU-IUL, Lisboa, Portugal

E-mail: jose.dias@iscte-iul.pt

Abstract

Symbolic data analysis (SDA) has been proposed as an extension to data analysis that handles more complex data structures [1,2]. In this general framework observations are characterized by more than one value: from two (e.g., interval-value data defined by minimum and maximum values) to multiple-valued variables (e.g., frequencies or proportions). Often these data structures result from data aggregation with the purpose of studying heterogeneity or natural clusters. Many algorithms and methodologies have been proposed to clustering data and most of them have been generalized to SDA. This research proposes a new clustering algorithm for multiple-valued data. It is based on the symmetric Kullback-Leibler distance combined with a complete-linkage rule within the hierarchical clustering framework. The algorithm is applied to the population pyramids of World countries. They define symbolic data, i.e., individual observations are aggregated at country level and summarized by population pyramids. Results show that the population structure of the 220 countries is quite diverse, but can be clustered into five groups or pyramidal shapes. These results allow a better classification and characterization of countries regarding population structure and dynamics, and can be relevant for the definition of demographic and social policies worldwide.

Keywords: symbolic data analysis, clustering algorithms, demography, population pyramids.

Acknowledgements

This research was supported by the Fundação para a Ciência e Tecnologia (Portugal), Grant UID-GES-00315-2013.

References

[1] Billard, L. and Diday, E., (2006) Symbolic Data Analysis. Conceptual Statistics and Data Mining, Chichester, UK, Wiley.

[2] Noirhomme-Fraiture, M. and Brito, P., (2011) Far beyond the classical data models: Symbolic data analysis, *Statistical Analysis and Data Mining*, v.4, n.2, pp. 157–170.

Can non-surgical periodontal treatment improve clinical and biochemical parameters of rheumatoid arthritis? A meta-analysis

J.A. Pereira[1], Ana S. Ferreira[1], Marta Resende[1] and Luzia Mendes[1]

[1]Faculdade de Medicina Dentária, Universidade do Porto, Portugal

E-mail: jpereira@fmd.up.pt

Abstract

Periodontal disease (PD) is one of the most prevalent oral disorders, affecting 10% - 60% of world population [1]. The association with rheumatoid arthritis (RA) is based on the similar humoral and cellular immune responses [2]. Moreover, these two genetic diseases share genetic and environmental risk factors [3], as well as similar pathogenic and clinical characteristics [4]. Meta-analysis combines estimates of quantities of interest, as obtained from studies addressing the same research question, to estimate the overall/mean of an outcome of interest. The research question about the effect size (ES) of non-surgical periodontal treatment (PT) on RA improvement can be answered by variations of different negative oriented, surrogated, or real endpoints. Although the overall effect tends to be a reduction of all endpoints, some discordant variations make the available studies inconclusive. In this work we approach this issue aggregating the endpoints, according to Borenstein et al. (2009) [5] procedure. Three composite ES were computed. One aggregating all ES within the study (OCE), another the subjective (symptoms) outcome measures (SCE) and the third the analytic (ACE), so that each independent study contributes only with a single EF of each type. The Edge'g effect size was computed for each endpoint. We applied the fixed-effects model by summarizing the results of $k = 4$ studies, each of which has a sample size $\eta_k, k = 1, \ldots, k$. In each study, there is a true effect β_k estimated by $\hat{\beta}_k$, with a true standard error σ_k estimated by $\hat{\sigma}_k$, or, equivalently, a true variance σ_k^2 estimated by $\hat{\sigma}_k^2$ and between-study variance τ^2. The τ^2 and β were estimated by restricted maximum-likelihood estimator (REML) and the inference was attained by first- (signed loglikelihood ratio test (rLRT)) and higher order statistics (Skovgaard statistics (rSkov)). The meta-analysis results were presented in a forest plot graph and publication bias was assessed by a funnel plot graph combined with the trim and fill method and tested by Egger's linear regression method. The two-sides confidence intervals were computed, but since the inferential interest is to evaluate the reduction of the

composite endpoints of markers of disease activity, we tested the hypothesis of composite indexes being equal to 0 (null hypothesis) against the one-sided alternative of being lesser then 0 (one-sided alternative). The results failed to estimate reliable quantitative results, but have shown a positive effect of PT on RA patients, enabling us to conclude that PT should be part of RA and PD patients' treatment.

Keywords: periodontal treatment, rheumatoid arthritis, meta-analysis, aggregate endpoints.

References

[1] Detert J., Pischon N., Burmester G.R. and Buttgereit F., (2010) The association between rheumatoid arthritis and periodontal disease, *Arthritis Res Ther.*, 12(5), pp. 218.

[2] Han J.Y. and Reynolds M.A., (2012) Effect of anti-rheumatic agents on periodontal parameters and biomarkers of inflammation: a systematic review and meta-analysis, *J Periodontal Implant Sci.*, 42(1), pp. 3–12.

[3] Monsarrat P., Vergnes J.N., Cantagrel A., Algans N., Cousty S., Kemoun P., Bertrand C., Arrivé E., Bou C., Sédarat C., Schaeverbeke T., Nabet C. and Sixou M., (2013) Effect of periodontal treatment on the clinical parameters of patients with rheumatoid arthritis: study protocol of the randomized, controlled ESPERA trial. Trials., 14, pp. 13.

[4] Ortiz P., Bissada N.F., Palomo L., Han Y.W., Al-Zahrani M.S., Panneerselvam A. and Askari A., (2009) Periodontal therapy reduces the severity of active rheumatoid arthritis in patients treated with or without tumor necrosis factor inhibitors, *J Periodontol.*, 80(4), pp. 134-137.

[5] Borenstein, (2009) Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), The handbook of research synthesis and meta-analysis (pp. 279-293), New York, Russell Sage Foundation.

Sectors and routes in transportation of non-urgent patients

Luís Miguel Bandeira[1], Ana Maria Rodrigues[2,3] and José Soeiro Ferreira[1,2]

[1]Faculdade de Engenharia, Universidade do Porto, Portugal
[2]INESC TEC, Portugal
[3]3CEOS.PP, Politécnico do Porto, Portugal

E-mail: ee12106@fe.up.pt

Abstract

Sectors appear in a large variety of contexts. Sectorization is often associated with geographic divisions such as political (re)districting, with the design of police areas, school districting, but also with home-care services, location problems and definition of sales territories. Several criteria are considered to design and assure the "quality" of the sectors. Criteria such as equilibrium, compactness or contiguity are frequently used. However, the opinion of decision makers and the type of applications may lead to other criteria, like the respect of natural boundaries, integrity or the representation of ethnic minorities (see [1]). Important surveys such as [2] and [3] can be useful for those not so familiar with the subject. Sectorization problems are usually difficult to solve, for several reasons. In this event, a real case involving sectorization and routing, applied to the daily transport of patients between their residences and a hospital centre, will be presented. The hospital centre, located in Porto, has trained staff to attend these patients and offers them daily support (Monday to Friday). Patients are from a vast area around the town and have different particularities. Not all patients require hospital visits every day. The hospital centre takes patients back to their homes, but not everyone does it at the same time. Part of the patients only returns home after dinner, while others do so early. The fleet of vehicles that transport patients is homogeneous and each vehicle has a maximum capacity of seven patients. It is intended that the transport be made as quickly as possible, (the cost of each route is given by the time of the trip) and that the occupancy rate of the vehicles is high. In the case of transportation involving patients with mental disabilities, it is important that the knowledge of drivers and caregivers can emerge in the resolution process. The method used to deal with this case resorts to SectorEl, an electrostatic based approach to sectorization, which has already been developed by the authors, see [4], including,

however, some relevant changes. The procedures are inspired on Coulomb's Law and allow to create sectors, in compliance with factors the patients have in common, and to enhance the solutions. Moreover, it is possible to take into consideration undesirable forms of a solution. By applying the method and attending specific criteria and the characteristics of each patient, it is possible to divide the entire group of patients into small groups. Within each group, the transport is done by a single vehicle. The results obtained will be shown and interpreted, especially in what concerns interesting aspects such as: the variability of patients during the week, the opportunity of rapid incorporation of new patients into the process and the likelihood that some may quit the centre, for several reasons.

Keywords: sectorization, routing, patient transport.

Acknowledgements

This work is co-financed by the ERDFEuropean Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE2020 Programme within project POCI-01-0145-FEDER-006961.

References

[1] Ricca, F., Scozzari, A. and Simeone B., (2013) Political Districting: from classical models to recent approaches, *Annals of Operations Research*, 204, pp. 271–299.

[2] Cortinhal, M.J., Mourão, M.C. and Nunes, A.C., (2016) Local search heuristics for sectoring routing in a household waste collection context, *European Journal of Operational Research*, 255, pp. 68–79.

[3] Rodrigues, A.M. and Ferreira, J.S., (2015) Measures in Sectorization Problems, *Operations Research and Big Data, Studies in Big Data*, 15, Springer, pp. 203–211.

[4] Rodrigues, A.M. and Ferreira, J.S., (2015) Sectors and Routes in Solid Waste Collection, *Operational Research, CIM Series in Mathematical Sciences*, 4, Springer, pp. 353–375.

Chains of oscillators in non-homogeneous media

Luís Bandeira[1,2] and Carlos C. Ramos[1,2]

[1]Universidade de Évora, Departamento de Matemática, Portugal
[2]Centro de Investigação em Matemática e Aplicações (CIMA), Portugal

E-mail: lmzb@uevora.pt

Abstract

We introduce a model to study the vibrational properties of non-homogeneous materials. In our work these materials, naturally idealized, are composed by one dimensional chains of harmonic oscillators represented by an alternating sequence of particles and springs. Despite the system has explicit solutions since it is linear, the formulas can be very complicated. We use homogeneous chains as building blocks for characterizing the whole system and the global dynamics. In particular, we determine the solution for a chain composed of two distinct homogeneous chains in terms of the original solutions for these two homogeneous chains, when uncoupled. This gives us a general procedure to deal easily with hon-homogeneous chains.

Keywords: dynamical system, harmonic oscillator, non-homogenous chains, recursion.

Acknowledgements

This work was supported by FCT - Fundação para a Ciência e a Tecnologia under the project UID-MAT-04674-2013.

References

[1] Bandeira, L. and Correia Ramos, C., (2018) Harmonic oscillations on non-homogeneous media, *submitted*.

[2] Bandeira, L. and Correia Ramos, C., (2016) Transition matrices characterizing a certain totally discontinuous map of the interval, *Journal of Mathematical Analysis and Applications*, 444(2), pp. 1274–1303.

[3] Bandeira, L. and Correia Ramos, C., (2015) On the spectra of certain matrices and the iteration of quadratic maps, *SeMA J.*, 67, pp. 51–69.

[4] Correia Ramos, C., Martins, N. and Pinto, P.R., (2017) Toeplitz algebras arising from escape points of interval maps, *Banach J. Math. Anal.*, 11(3), pp. 536–553.

[5] Correia Ramos, C., Martins, N. and Pinto, P.R., (2013) On C*-algebras from interval maps, *Complex Analysis and Operator Theory*, 7, pp. 221–235.

PLS-PM to evaluate worker's health perception

Luís M. Grilo[1,2,3,4,5] and Valter Vairinhos[5,6]

[1]Instituto Politécnico de Tomar, Departamento de Matemática e Física, Portugal
[2]Centro de Matemática e Aplicações (CMA), FCT/UNL, Portugal
[3]Centro de Investigação em Cidades Inteligentes (Ci2), IPT, Portugal
[4]CIICESI/ESTG–P. Porto, Portugal
[5]ICLAB–ICAA–Intellectual Capital Accreditation Association, Santarém, Portugal
[6]CINAV, Center of Naval Research, Portugal

E-mail: lgrilo@ipt.pt

Abstract

In almost every human activity sectors the number of workers under psychosocial risks has been increasing worldwide with negative impact on workers' health and wellbeing, their organizations and economies in general [1, 2, 4, 5, 6, 9, 10]. A short version (41 questions/variables measured on a 5-point scale) of the Copenhagen Psychosocial Questionnaire, which includes relevant dimensions according to several theories on psychosocial risk factors in the workplace, was applied to workers from a Portuguese financial institution [3, 8]. To describe and synthetize the sample data some multivariate methodologies were employed, such as principal components and cluster analysis. A set of seven clusters emerged as manifestations of latent variables with clear psychosocial meaning; a path model was then designed, expressing a priori empirical perceptions and authors' experience about causal relations among those latent variables, both consistent with literature. The model was estimated using Partial Least Squares Path Modeling (PLS PM), implemented through an R package [7]. A nonparametric bootstrap procedure was used to assess the statistical significance of the estimated trajectory coefficients. Although the sociodemographic variables that characterize the workers of this institution were not available (such as gender, age and educational stage), which might be seen as a limitation, the results obtained are considered easily interpretable, very useful and interesting, for the occupational health specialists.

Keywords: bootstrap, clusters, psychosocial risks, survey, R-software.

Acknowledgements

This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID-MAT-00297-2013 (Centro de Matemática e Aplicações).

References

[1] Grilo, L. M. and Vairinhos, V., (2017) Modeling health perception based on psychosocial risks, A data driven approach, *Satellite Meeting ISI Committee on Risk Analysis and XI Workshop on Statistics, Mathematics and Computation*, Universidade Aberta de Lisboa and Instituto Politécnico de Portalegre, Portugal, July 10-12 (pp. 76, Book of Abstracts).

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

[3] 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 & Stress*, 19, pp. 192–207.

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

[5] Maslach, C. and Leiter, M.P., (2008) Early predictors of job burnout and engagement, *Journal of Applied Psychology*, 93, pp. 498–512.

[6] Neto, H. V., Areosa, J. and Arezes, P., (2014) Manual sobre Riscos Psicossociais no Trabalho, Porto: RICOT.

[7] Sanchez, G., (2013) PLS path modeling with R. Trowchez Editions, Berkeley.

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

[9] Soares, J., (2018) Reload, Menos stress, Melhor performance, Porto Editora.

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

Assessing the number of components in mixtures of linear mixed models

Luísa Novais[1] and Susana Faria[1]

[1]Departamento de Matemática e Aplicações, Universidade do Minho, Portugal

E-mail: luisa_novais92@hotmail.com

Abstract

Over the last decades, linear models have been widely studied by the scientific community as an important tool of statistical modelling in a great variety of phenomena. However, in many situations the data are grouped according to one or more factors, so the introduction of random effects is required in order to consider the correlation between observations from the same individual, in which case linear mixed models are used. Additionally, it is often observed that the data arise from a heterogeneous population, which gives rise to situations where the estimation of a single linear model is not sufficient. Therefore, it is necessary to use models that incorporate this unobserved heterogeneity, as is the case of finite mixture models. As a result, in regression analysis, it has been a popular practice to model unobserved population heterogeneity through finite mixtures of regression models. Thus, finite mixtures of linear mixed models have been applied in different areas of application since they conveniently allow to account for correlations between observations from the same individual and to model unobserved heterogeneity between individuals at the same time (see [1] and [4]). Assessing the number of components in mixture models has long been considered as an important research problem which has not yet been resolved. There is wide variety of literature available on the performance of model selection statistics for assessing the number of components in mixture models (see [2] and [5]). In this work we investigate the problem of selecting the number of components in mixtures of linear mixed models, analysing the performance of different information criteria in model selection. Nonetheless, the traditional information criteria are sensitive to outliers so the presence of a single outlier may cause the estimated number of components to change [3]. In order to overcome this problem, we study a robust estimation of the number of components for mixtures of linear mixed models, based on trimmed maximum likelihood estimates, and compare the performance of the trimmed information criteria to the performance of the traditional information criteria. To evaluate the methodologies studied, we carry out a simulation study and we

illustrate these methodologies using a real data set of the Panel Study of Income Dynamics (PSID) in order to identify the most suitable mixture to model the income of a set of individuals belonging to the study. The obtained results demonstrate that the criteria HQIC, AIC4 and aBIC are the best options to assess the number of components in mixtures of linear mixed models.

Keywords: finite mixtures of linear mixed models, model selection, information criteria, trimmed information criteria, robustness, simulation study.

Acknowledgements

The collection of data used in this study was partly supported by the National Institutes of Health under grant number R01 HD069609 and R01 AG040213, and the National Science Foundation under award numbers SES 1157698 and 1623684.

References

[1] Bai, X., Chen, K. and Yao, W., (2016) Mixture of linear mixed models using multivariate t distribution, *Journal of Statistical Computation and Simulation*, v.86, pp. 771–787.

[2] Depraetere, N. and Vandebroek, M., (2014) Order selection in finite mixtures of linear regressions, *Statistical Papers*, v.55, pp. 871–911.

[3] Li, M., Xiang, S. and Yao, W., (2016) Robust estimation of the number of components for mixtures of linear regression models, *Computational Statistics*, v.31, n.4, pp. 1539–1555.

[4] McLachlan, G. and Peel, D., (2000) Finite Mixture Models, John Wiley & Sons.

[5] Young, D. and Hunter, D., (2015) Random effects regression mixtures for analyzing infant habituation, *Journal of Applied Statistics*, v.42, pp. 1421–1441.

Bad science, fake p-values and mendel variables

Maria de Fátima Brilhante[1,5], Sandra Mendonça[2,5], Dinis Pestana[3,5] and Fernando Sequeira[4,5]

[1]Departamento de Matemática e Estatística, Universidade dos Açores, Portugal
[2]Departamento de Matemática e Engenharias, Universidade da Madeira, Portugal
[3]Instituto de Investigação Científica Bento da Rocha Cabral, Portugal
[4]Departamento de Estatística e Investigação Operacional, Universidade de Lisboa, Portugal
[5]CEAUL — Centro de Estatística e Aplicações da Universidade de Lisboa, Portugal

E-mail: ddpestana@fc.ul.pt

Abstract

Assuming that the null hypothesis H_0 is true the p-value is and observation of a standard uniform random variable. In meta analysis the synthesis of the results of k independent tests H_0 vs. H_1 is often performed using a combined p-value to test the overall hypothesis H_0^*: $H_{0,i}$ is true for all $i = 1, ...k$ vs. H_1^*: there exists at least one $j \in \{1, ..., k\}$ for which $H_{1,j}$ is true. The classical theory of combined p-values assumes uniformity, cf. Pestana [1], the most popular test using Fisher's test statistic $T = -2\sum_{j=1}^{k} \ln(P_j) \frown \chi^2_{2k}$. As statistical tests are mainly used to reject the null hypothesis, assuming the validity of H_0^* is farfetched, and this led Tsui and Weerahandi [2] to introduce the concept of generalized p-values, cf. also Weerahandi [3]. However, even in the very unlikely situation of the $H_{0,j}$ being true $\forall j \in \{1, ..., k\}$ two other reasons can disrupt the classical theory of combined p-values:

1. It is unlikely that non significative p-values had been published, since those are traditionally viewed as weak results, not contributing to the advance of knowledge. This *file drawer problem* is recognized as an important source of error in meta analysis;
2. Some of the p-values are not from an uniform population, even although H_0 is true. This occurs in particular when the result of a first experiment doesn't fit the experimenter interests, and a second experimenter is performed and the "best" p-value is reported. This is in fact what we call a *fake p-value*, that being the observed value of the minimum of two standard uniform observations, comes from a $Beta(1, 2)$ population (or from a $Beta(2, 1)$, if the experimenter reports the maximum). This malpractice has been presumably done by Mendel, cf. Pires and Branco [4] explanation of the Mendel–Fisher controversy.

If there is a proportion $m/2$ of fake p-values in a sequence $\boldsymbol{p} = \{p_1, ..., p_k\}$ of reported p-values, the appropriate model is a combination of uniform (with weight $1 - m/2$) and $Beta(1, 2)$ (or $Beta(2, 1)$), with probability density function

$$f_P(x) = \left(mx + 1 - \frac{m}{2}\right) \mathbb{I}_{(0,1)}(x).$$

We denote this by $P \frown Mendel(m)$, where $m \in [-2, 2]$, the extreme cases being the $Beta(1, 2)$ and the $Beta(2, 1)$, and the central case $m = 0$ the standard uniform $U \frown Uniform(0, 1)$. On the other hand, a uniformity test H_0: $m = 0$ vs. H_1: $m \neq 0$ may allow us to deal with a combined p-value "as if" no fake p-values exist in the sequence $\boldsymbol{p}$. But in many practical situations the sample size k is very small, and the test power very tiny. Deng and George [5] presented an interesting characterization of the standard uniform: Let X and Y be independent random variables with support $\mathcal{S} = [0, 1]$. $V = \min\left(\frac{X}{Y}, \frac{1-X}{1-Y}\right)$ is uniform and independent of Y if and only if $X \frown Uniform(0, 1)$. Brilhante *et al.* [6] extended this result, showing that if $X \frown Mendel(m)$ then $V = \min\left(\frac{X}{Y}, \frac{1-X}{1-Y}\right) \frown Mendel\left([2\mathbb{E}[Y] - 1]m\right)$. This result is useful for computationally augmenting the sample $\boldsymbol{p}$, with $\boldsymbol{v} = \{v_1, ..., v_k\}$, using Y with $\mathbb{E}[Y] \approx 1$. As if $m = 0$ the sample $(\boldsymbol{p}, \boldsymbol{v})$ will be uniform with size $2k$, but if $m \neq 0$ it will be approximately $Mendel(m)$ and, as we shall show, this increasesthe power of the test (the same happening with the Fisher's χ^2_{4k} test on the combined p-value). Independence vs. Mendel auto-regressive dependence is also discussed.

Keywords: combined p-values, fake p-values, uniformity, independence.

Acknowledgements

Research financed by FCT — Fundação para a Ciência e a Tecnologia, Portugal, Project UID-MAT-00006-2013.

References

[1] Pestana, D., (2011) Combining p-values, International Encyclopaedia of Statistical Science, M. Lovric, Ed. New York, Springer Verlag, pp. 1145–1147.

[2] Tsui, K. and Weerahandi, S., (1989) Generalized p-values in significance testing of hypothesis in the presence of nuisance parameters, *The American Statistician*, 84, pp. 602–607.

[3] Weerahandi, S., (1995) Exact Statistical Methods for Data Analysis, Springer, New York.

[4] Pires, A.M. and Branco, J.A., (2010) A statistical model to explain the Mendel-Fisher controversy, *Statistical Science*, vol. 25, pp. 545–565.

[5] Deng, L.Y. and George, E.O., (1992) Some characterizations of the uniform distribution with applications to random number generation, *Annals of the Institute of Statistical Mathematics*, vol. 44, pp. 379–385.

[6] Brilhante, M.F., Mendonça, S., Pestana, D., Rocha, M.L., Sequeira, F. and Velosa, S., (2018) Fake p-Values and Mendel Variables: Testing Uniformity and Independence, *International Conference On Mathematical Applications*, Funchal.

Improving customer satisfaction in an automobile repair shop

Maria de Fátima Pilar[1], Eliana Costa e Silva[1] and Ana Borges[1]

[1]CIICESI – Center for Research and Innovation in Business Sciences and Information Systems, School of Technology and Management, Polytechnic of Porto, Portugal

E-mail: 8120109@estg.ipp.pt

Abstract

Scheduling is a classic combinatorial problem much studied by researchers in Operational Research (OR) [1], in areas such as production planning, telecommunications, logistics, and computer control [2]. Since the early 1950's it has received much attention from OR practitioners, management scientists, production and operations research workers, as well as mathematicians since the early 1950s [3]. Scheduling problems can be defined as the allocation of resources to perform a set of tasks over a period of time, to optimize some or several performance measures [4,5]. The resources and tasks in an organization may be, for example, machines in a workshop, runways at an airport, while tasks may be, for example, operations in a production process, take-offs and landings at an airport [6]. In in this paper the real case of a Portuguese automobile firm is addressed. The firm intends the improve costumer satisfaction by minimizing the waiting times of the customers in relation to the vehicles that enter their repair shop, from the moment they arrive until they are delivered to their customers. The central focus of this study is on mechanical repair. The process that each vehicle goes in the repair shop starts by the opening of a repair order form, indicating the necessary mechanical repairs to be performed. Next the manager assigns the repair taks to one of the 8 mechanics available in this repair shop. We intend to develop a mathematical model that, through the data provided by the firm, can best schedule repairs, taking into account the mechanics available, the resources available, the interventions that need to be carried out and the time these interventions take. Specifically, we develop a Mixed Integer Linear Programming (MILP). This technique has been widely used in scheduling problems, due to its rigorousness, flexibility and extensive modeling ability. Furthermore it has been employed in a wide variety of real-world problems [7]. With the provision of data by the firm, with information regarding repairs made during a year, the times of these repairs, among others. From this data set, a small subset of 10 cars was analyzed, with the purpose of elaborating the scheduling of the

same. For this, the AMPL modeling language was used to model the problem, and numerical results were obtained using Gurobi solver. This subset of data was tested for the minimization of Tardiness and starting time. From the results obtained, and taking into account the interventions performed in these vehicles (some with joint interventions), the mechanics available and the time of each intervention, it was concluded that scheduling the necessary repairs allowed a part of these vehicles to be delivered on the day of your entry. Only 8 of the 10 vehicles were repaired from immediately upon arrival, this being due to the existence of only 8 mechanics in the repair office. The sequencing of the interventions is respected, there being no overlap of interventions and mechanics, however, there were times when some vehicles were stopped, making the satisfaction of the clients and the maximization of the human resources were not satisfied.

Keywords: shceduling, mathematical programming, MILP.

Acknowledgements

This work is partially supported by CIICESI – Center for Research and Innovation in Business Sciences and Information Systems, School of Technology and Management, Polytechnic of Porto, Portugal, and it is developed within the scope of the first author's Master Thesis.

References

[1] Fukunaga, A., Rabideau, G., Chien, S. and Yan, D., (1997) Towards an application framework for automated planning and scheduling, *Aerospace Conference, Proceedings, IEEE*, 1, pp. 375–386.

[2] Unlu, Y. and Mason, S.J., (2010) Evaluation of mixed integer programming formulations for non-preemptive parallel machine scheduling problems, *Comput Ind Eng*, 58(4) pp. 785–800.

[3] Maccarthy, B.L. and Liu, J., (1993) Addressing the gap in scheduling research: a review of optimization and heuristic methods in production scheduling, *Int J Prod Res*, 31(1), pp. 59–79.

[4] Demir, Y. and İşleyen, S.K., (2013), Evaluation of mathematical models for flexible job-shop scheduling problems, *Appl Math Model*, 37(3), pp. 977–988.

[5] Leung, J.Y.T., (2004), Handbook of scheduling: algorithms, models, and performance analysis, CRC Press.

[6] Pinedo, M.L., (2008) Scheduling: Theory, Algorithms, and Systems, Springer, New York.

[7] Floudas, C.A. and Lin, X., (2005) Mixed integer linear programming in process scheduling: Modeling, algorithms, and applications, *Ann. of Oper. Res.*, 139(1), pp. 131–162.

Mean-of-order-p value-at-risk estimation: a Monte-Carlo comparison

M. Ivette Gomes[1,2], Frederico Caeiro[3,4], Fernanda Figueiredo[1,5], Lígia Henriques-Rodrigues [1,6] and Dinis Pestana[1,2]

[1]Centro de Estatística e Aplicações (CEA), Universidade de Lisboa (UL), Portugal
[2]Departamento de Estatística e Investigação Operacional, Faculdade de Ciências, UL, Portugal
[3]Centro de Matemática e Aplicações (CMA), Universidade Nova de Lisboa (UNL), Portugal
[4]Departamento de Matemática, Faculdade de Ciências e Tecnologia, UNL, Portugal
[5]Faculdade de Economia, Universidade do Porto, Portugal
[6]Universidade de São Paulo, IME, Brasil

E-mail: ivette.gomes@fc.ul.pt

Abstract

The risk of a big loss that occurs rarely is a primordial parameter of extreme events. A possible and common indicator of such a risk is the *value-at-risk* (VaR), i.e. the size of the loss that occurred with a small probability, q. For any unknown cumulative distribution function F underlying a possibly weakly dependent and stationary available sample, and denoting by $F^{\leftarrow}(y) := \inf\{x : F(x) \geq y\}$ the generalized inverse function of F, we are thus dealing with a (high) quantile, $\chi_{1-q} \equiv \mathrm{VaR}_q := F^{\leftarrow}(1-q)$. With n denoting the size of the available sample, we often have $nq \leq 1$, and this justifies theoretically the assumption that $q = q_n \to 0$, as $n \to \infty$. We thus want to extrapolate beyond the sample, being then in the area of statistical *extreme value theory* (EVT). Since in real applications in the areas of biostatistics, finance, insurance and statistical quality control, among others, one often encounters heavy right-tails, we shall assume that, for some $\xi > 0$, the *right-tail function* (RTF) satisfies the condition $\overline{F}(x) := 1 - F(x) \sim c\, x^{-1/\xi}$, as $x \to \infty$, for some positive constant c, where the notation $a(y) \sim b(y)$ means that $a(y)/b(y) \to 1$, as $y \to \infty$. The parameter ξ is a positive version of the general *extreme value index* (EVI), the primary parameter of extreme (and large) events. For heavy right-tailed or Paretian-type models, and with Q standing for quantile, Weissman ([1]) proposed the following semi-parametric VaR-estimator,

$$\mathrm{Q}^{(q)}_{\hat{\xi}}(k) := X_{n-k:n}\,(k/(nq))^{\hat{\xi}} =: X_{n-k:n}\, r_n^{\hat{\xi}}, \quad r_n \equiv r_n(k;q) = k/(nq),$$

where $X_{n-k:n}$ is the $(k+1)$-th upper order statistic and $\hat{\xi}$ is any consistent estimator for ξ. This is obviously an *asymptotic* estimator, in the sense that it provides useful estimates when the sample size n is high. For heavy RTFs, the classical EVI-estimator, usually the one which is used for a semi-parametric quantile estimation, is the Hill estimator $\hat{\xi} = \hat{\xi}(k) =: \mathrm{H}(k)$ ([2]), the average of the log-excesses, $V_{ik} := \ln(X_{n-i+1:n}/X_{n-k:n}) =: \ln U_{ik}$, $1 \le i \le k < n$. Since the Hill estimator is the logarithm of the *geometric mean* (or *mean-of-order*-0) of U_{ik}, Brilhante *et al.* ([3]) considered as basic statistics the Hölder's *mean-of-order-p* (MO_p) of U_{ik}, $1 \le i \le k$, $p \in \mathbb{R}_0^+$. More generally, Caeiro *et al.* ([4]) worked with $p \in \mathbb{R}$ and a class of MO_p EVI-estimators, $\mathrm{H}_p(k)$, which can be used for the VaR_q-estimator, through the class $\mathrm{Q}^{(q)}_{\mathrm{H}_p}(k)$. The MO_p EVI-estimators can often have a high asymptotic bias, and bias reduction has recently been a vivid topic of research in the area of statistical EVT. On the basis of partially reduced-bias ([5]) and reduced-bias ([6]) H_p EVI-estimators, respectively denoted by $\mathrm{PRB}_p(k)$ and $\mathrm{CH}_p(k)$, it is thus sensible to work with $\mathrm{Q}^{(q)}_{\mathrm{PRB}_p}(k)$ (already considered in [7]) and with the new VaR_q-estimators $\mathrm{Q}^{(q)}_{\mathrm{CH}_p}(k)$. After a brief reference to the asymptotic properties of these new VaR-estimators, we proceed to an overall comparison of VaR-estimators, through Monte-Carlo simulation techniques.

Keywords: heavy right-tails, Bias reduction, Monte-Carlo simulation, semi-parametric estimation, statistics of extremes, value-at-risk estimation.

Acknowledgements

Research partially supported by National Funds through **FCT**—Fundação para a Ciência e a Tecnologia, projects UID-MAT-0006-2013 (CEA/UL) and UID-MAT-0297-2013 (CMA/UNL), and by COST Action IC1408.

References

[1] Weissman, I., (1978) Estimation of parameters and large quantiles based on the k largest observations, *J. Amer. Statist. Assoc.*, 73, pp. 812–815.

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

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

[4] 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), pp. 561–589.

[5] Gomes, M.I., Brilhante, M.F., Caeiro, F. and Pestana, D., (2015) A new partially reduced-bias mean-of-order p class of extreme value index estimators, *Comput. Statist. and Data Analysis*, 82, pp. 223–237.

[6] Gomes, M.I., Brilhante, M.F. and Pestana, D., (2016) New reduced-bias estimators of a positive extreme value index, *Comm. Statist.—Simul. & Comput.*, 45, pp. 1–30.

[7] Gomes, M.I., Caeiro, F., Figueiredo, F., Henriques-Rodrigues, L. and Pestana, D., (2017) Corrected-Hill versus Partially Reduced-Bias Value-at-Risk Estimation, CEAUL 01/2017.

A stochastic diffusion process based on brody curve

O. Rida[1], A. Nafidi[1] and B. Achchab[1]

[1]Univ. Hassan 1, LAMSAD, École Supérieure de Technologie de Berrechid, Avenue de l'université, BP 280, Berrechid, Morocco

E-mail: oussamar@gmail.com

Abstract

The principal aim of the present work is to study a stochastic diffusion process based on the Brody curve ([1], [2]). Such a process can be considered as an extension of the nonhomogeneous lognormal diffusion process [3]. From the corresponding Ito's stochastic differential equation (SDE), firstly we establish the probabilistic characteristics of the studied process, such as the solution to the SDE, the probability transition density function and their distribution, the moments function in particular the conditional and unconditional trend functions. Secondly, we treat the parameters estimation problem by using the maximum likelihood method in basis of the discrete sampling, thus we obtain non-linear equations that can be solved by numerical methods.

Keywords: diffusion process, Brody curve, stochastic differential equation, maximum likelihood estimation.

References

[1] S. Brody, (1945) Bioenergetics and Growth, Reinhold Publishing, New York.

[2] J.E. Brown, C.J. Brown and W.T. Butts, (1972) A discussion of the genetic aspect of weight, mature weight and rate of maturing in Hereford and Angus cattle, *J. Anim. Sci.*, 34, pp. 525–537.

[3] R. Gutiérrez, P. Roman and F. Torres, (1999) Inference and first-passage time for the lognormal diffusion process with exogenous factors: application to modelling in economics, *Applied Stochastic Models in Business and Industry*, 15, pp. 325–332.

Applying structural equation model to marketing research

Oliva Maria Dourado Martins[1,2,3], Arminda Maria Finisterra do Paço[3,4], José Paulo Costa[5,6], Ana Sofia Coelho[6,7] and Ricardo Gouveia Rodrigues[3,4]

[1]Instituto Politécnico de Tomar, Portugal
[2]Centro de Investigação Aplicada em Economia e Gestão do Território (CIAEGT-IPT), Portugal
[3]Núcleo de Estudos em Ciências Empresariais (NECE–UBI), Portugal
[4]Management and Economy Departament, University of Beira Interior, Covilhã, Portugal
[5]Instituto Politécnico de Bragança, Portugal
[6]Instituto de Estudos Superiores de Fafe, Portugal
[7]Instituto Superior de Contabilidade e Administração da Universidade de Aveiro, Portugal

E-mail: oliva.martins@ipt.pt

Abstract

Constructs, also called latent variables, are unobserved variables. For this reason, it is difficult to measure directly. So, those latent variables can be measured indirectly, and need to be estimated through other observed variables, called indicators. To obtain more information about how to measure these unobserved variables, this work becomes an example on developing a theoretical and empirical validation process. The criteria for the validation process is unique to each individual model and research context. After the theoretical definition of each individual construct and the relationship between those, in terms of structural equation modeling (SEM) practical application, especially in the field of Marketing, becomes necessary to develop a process of empirical validation. SEM is a technique used in the statistical analysis field. There are two different kinds of SEM techniques: i) based on covariance, called Covariance-Based SEM (CB-SEM), or ii) performed through partial least squares path analysis, called Partial Least Square - Structural Equation Model (PLS - SEM) [4]. These two techniques are applied in different contexts and with different requirements. Furthermore, these techniques also have already been considered as complementary techniques. The Covariance-Based SEM technique (CB-SEM) is usually applied to reflexive constructs because it requires a high covariation. In turn, PLS is a non-parametric technique that aims to predict and develop the theory [2] despite the difficulty of controlling predictive efficacy. As an example, the structural

and measurement model were presented too. As a formative construct, the indicators were called items of the construct. They can be in different conceptual dimensions and, together, form the theoretical concept. They were also associated with the causes of the construct. For the reflexive construct, the indicators were considered scales of the construct, and reflected on the conceptual dimension. It was associated with an effect of the construct [1]. In a practical application, the aim of the work is to present a statistical validation process of SEM, namely a Partial Least Square - Structural Equation Model (PLS-SEM) applied on the Breastfeeding Behavioral Intention Model (BBIM) [5]. The data analysis was supported by SPSS software, version 19.0 [4] for Windows, and by SmartPLS 2.0 software [5]. All statistic results were validated also. And the results showed a path to help mothers decide about breastfeeding.

Keywords: PLS-SEM, CB-SEM, breastfeeding behavioural intention model, SmartPLS 2.0, SPSS.

Acknowledgements

The authors would like to thank all respondent or anyone who helped this research. The authors would also like to thank the Polytechnical Institute of Tomar, the research center CIAEGT of Polytechnical Institute of Tomar and the Núcleo de Estudos em Ciências Empresariais of UBI.

References

[1] Coltman, T., Devinney, T.M., Midgley, D.F. and Venaik, S., (2008) Formative versus reflective measurement models Two applications of formative measurement, *Journal of Business Research*, 61, pp. 1250–1262.

[2] Hair, J.F., Ringle, C.M. and Sarstedt, M., (2011) PLS-SEM: Indeed a Silver Bullet, *The Journal of Marketing Theory and Practice*, 19(2), pp. 139–152.

[3] Martins, O.M.D., Paço, A.M.F. and Rodrigues, R.J. de A.G., (2012) Influenciadores da intenção do comportamento do aleitamento materno – um estudo exploratório no âmbito do marketing social, *Revista Innovar*, 22(46), pp. 99–109.

[4] Oliveira, T. A., Kitsos, C. P., Oliveira, A. and Grilo, L., (2018) Recent Studies on Risk Analysis and Statistical Modeling, Contributions to Statistics Series, Springer International Publishing, pp. 1-333.

[5] Ringle, C. M., Wende, S. and Will, A., (2005) SmartPLS 2.0 (beta), SmartPLS, Hamburg.

An optimal control problem for a non autonomous SIR model for the Ebola virus

Paulo Rebelo[1], Silvério Rosa[2] and C.M. Silva[1]

[1]Departamento de Matemática, Universidade da Beira Interior, Portugal
[2]Departamento de Matemática and Instituto de Telecomunicações, Universidade da Beira Interior, Portugal

E-mail: rebelo@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. We consider a controlled SIR epidemic model where the control means the vaccination of the susceptible individuals and the treatment of infectious elements of the population. Numerical simulations are provided for both non controlled and controlled models. Also a graphical description of the optimal control is provided.

Keywords: biomathematics, optimal control, numerical simulation, epidemic model, Ebola outbreak.

Acknowledgements

This work was supported in part by CMA-UBI, Universidade da Beira Interior, Portugal.

References

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

[2] Zaman G., Kang Y.H. and Jung I.H., (2009) Optimal treatment of an SIR epidemic model with time delay, *BioSystems*, 98, pp. 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*, to appear.

Students index of satisfaction in engineering courses in Portugal years 2016 and 2017

Raquel Oliveira[1,4], A. Manuela Gonçalves[2] and Rosa M. Vasconcelos[3]

[1]Universidade do Minho, CMAT- Centro de Matemática, Guimarães, Portugal
[2]Universidade do Minho, CMAT- Centro de Matemática, DMA- Departamento de Matemática e Aplicações, Guimarães, Portugal
[3]Universidade do Minho, 2C2T- Centro para a Ciência Têxtil e Tecnologia, Guimarães, Portugal
[4]Instituto Politécnico do Cávado e do Ave, Escola de Tecnologia, Departamento de Ciências, Portugal

E-mail: rmro_17@hotmail.com

Abstract

In this work we describe and characterize students' allocation satisfaction in the Portuguese public higher education system through the students' point of view, namely in the academic engineering programs, extending previous studies of the author's team. We compare the demand satisfaction index, which is the ratio provided by the Portuguese Education Ministry (estimated via the institutions' point of view) with the ratio we propose, called applicants' satisfaction index. The data used in this paper covers the years 2016 and 2017, and was provided by the Portuguese Education Ministry. Mann-Whitney and Kruskal-Wallis tests were performed in order to assess if there are significant differences between the tendencies of the results found for the other periods studied.

Keywords: students satisfaction index, applicants satisfaction index, higher education, hypothesis tests, non-parametric tests.

Acknowledgements

This work was partially supported by the Research Centre of Mathematics of the University of Minho with the Portuguese Funds from the FCT - Fundação para a Ciência e a Tecnologia, through the Project PEstOE-MAT-UI0013-2017. This work is also financed by FEDER funds through the Competitivity Factors Operational Programme - COMPETE and by national funds through FCT – Foundation for Science and Technology within the scope of the project POCI-01-0145-FEDER-007136

References

[1] OECD, Organization for economic co-operation and Development, (2006) Reviews of National policies for education: Tertiary education in Portugal, Examiner's Report, Available at http://www.dges.mctes.pt/NR/rdonlyres/8B016D34-DAAB- 4B5-0-ADBB-25AE105AEE88/2564/Backgoundreport.pdf.

[2] http://www.europa.eu.

[3] European Ministers of Education, The Bologna Declaration, 1999.

[4] R. Oliveira, A.M. Goncalves and Rosa M. Vasconcelos, (2015) A New Perspective of Students Allocation Satisfaction in Engineering Courses in Portugal, *AIP Conference Proceedings*, 41648, 840002.

[5] R. Oliveira, A.M. Goncalves and Rosa M. Vasconcelos, (2016) Index of satisfaction in engineering courses in Portugal based on the students perspective, *AIP Conference Proceedings*, 1738, 470003.

[6] http://www.dges.mctes.pt / DGES / pt / Estudantes / Acesso /Estatisticas /.

[7] http://www.dges.mctes.pt.

[8] http://www.dges.mctes.pt/DGES/pt/Estudantes/Acesso.

[9] R. Oliveira, A.M. Goncalves and Rosa M. Vasconcelos, (2013) An empirical study on the index of satisfaction of student allocation in the Portuguese undergraduate engineering courses, *Book of Abstracts ASMDA*, pp. 165.

[10] http://www.dados.gov.pt / PT/ CatalogoDados/ Dados.aspx?name =Class-Nacionaldeareasdeeducacaoeformacao.

[11] J.H. Higgins, (2004) Introduction to Modern Nonparametric Statistics, Thomson, Toronto.

[12] S. Siegel, N.J. Castellan, (1988) Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill, Second Edition.

A note on relationships between moments and cumulants

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

[1]Department of Mathematics and Center of Mathematics and Applications, University of Beira Interior, Covilhã, Portugal

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

E-mail: sandraf@ubi.pt

Abstract

In this presentation we will discuss the relationship between moments and comulants. It will be shown that the importance given to comulants is due to the fact of many properties of random variables can be better represented by comulants than by moments. It will also be presented different ways of understanding cumulants, showing their algebraic properties. Finally some examples of applications with comulants will be introduced.

Keywords: central moments, cumulants, cumulant generating function, moments.

Acknowledgements

This work was partially supported by national founds of FCT-Foundation for Science and Technology under UID-MAT-00212-2013 and UID-MAT-00297-2013.

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*, v.39, pp. 49–54.

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

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

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

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

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

Innovation and firm economic performance: evidence from Portuguese SME

Sérgio Nunes[1,2,3], Helena L. Grilo[1], Raul Lopes[2] and Oliva Martins[1]

[1]CIAEGT-IPT, Instituto Politécnico de Tomar, Portugal
[2]DINÂMIACET – Instituto Universitário de Lisboa, Portugal
[3]CIRIUS-ISEG, Universidade de Lisboa, Portugal

E-mail: spnunes@ipt.pt

Abstract

The problematic explored in this paper is the role of the innovation in the economic performance of firms. Portugal is considered a moderate innovator [1] and innovation is fundamental to change the profile of its competitiveness [2, 3], since it is widely recognized as the business strategy that drives firms to perform better and to explore more economic opportunities. The literature deals extensively with the relationship between innovation and economic performance of firms [4, 5, 6, 7], nevertheless there are few studies that address this relationship in a multidimensional way, highlighting the existent different complementarities and their role in the corresponding economic performances. On the other hand, firms have different possibilities regarding the competition type that they preferentially adopt in the markets [8, 9], which means that their choices affect their economic performance. This paper has two main objectives. The first is to analyse this general relationship through a multidimensional approach to the problem – the relationship between different innovation types (product, process and organizational) and different manifestations of business economic performance (sales, orders and exports) are analysed. The second, consists of testing this relationship differencing firms by the competition type that they preferentially use (price-cost or quality-innovation). Data were collected using a survey conducted in late 2010 and early 2011 to firms with at least 5% of turnover growth. A sample of 317 Portuguese SME, stratified by five levels of technological intensity and by three regions was used. The main descriptive statistics were calculated and a statistical model was constructed and tested using two ordinal regression models. Two clear results were obtained through the developed analysis. First, innovation assumes multiple typologies that, when combined, have positive impacts on the multiple dimensions of business performance. Second, this result is also confirmed taking the options of business competition into account. If these arguments are both useful then innovation must be analysed in academic, business and public policy terms, in a

multidimensional and path dependent way, understanding and strategically considering its various complementarities. From this point of view, conceptually, the main contribution of this work is to suggest that innovation is a complex process where the indeterminate nature of its results should not be confused with the nature of the rationality of the processes associated with its design and implementation. As it was not possible to include marketing innovations in this research, this might be seen as a limitation, since their inclusion would certainly make the results more general. Regarding future research, beyond the attempt to include marketing innovations, there are two promising lines of research, both theoretically and empirically. First, it would be important to relate these results to some determinant aspects of innovation as a dynamic process. Second, it would be important to test whether the different innovation typologies have positive impacts on other indirect dimensions associated with the firm's economic performance. Confirming these hypotheses, it might be suggested that the innovation process and its impacts on firm's economic performance should be understood as an evolutionary process along an innovation trajectory.

Keywords: firm performance, innovation types, ordinal regression model, Portugal.

References

[1] European Commission, (2017) European Innovation Scoreboard, Brussels.

[2] OECD, (2012) Innovation for Development, DCTI, OECD Publishing, Paris.

[3] Gibson, D. and Naquin, H., (2011) Investing in innovation to enable global competitiveness: The case of Portugal, *Technological Forecasting and Social Change*, 78(8), pp. 1299–1309.

[4] Nunes, S., Lopes, R. and Fuller-Love, N., (2017) Networking, Innovation, and Firms' Performance: Portugal as Illustration, *Journal of the Knowledge Economy*, DOI: 10.1007/s13132-017-0508-7.

[5] Azar, G. and Ciabuschi, F., (2017) Organizational innovation, technological innovation, and export performance: The effects of innovation radicalness and extensiveness, *International Business Review*, 26, pp. 324–336.

[6] Alves, M., Galina, S. and Dobelin, S., (2018) Literature on organizational innovation: past and future, *Innovation and Management Review*, DOI: 10.1108/INMR-01-2018-001.

[7] Rodil, O., Vence, X. and Sánches, M., (2016) The relationship between innovation and export behavior: The case of Galician firms, *Technologic Forecasting & Social Change* 113, pp. 248–265.

[8] Katila, R., Chen, E. and Piezunka, H., (2012) All the right moves: how entrepreneurial firms compete effectively, *Strategic Entrepreneurship Journal*, 6, pp. 116–132.

[9] Pertusa-Ortega, E., Molina-Azorín, J. and Claver-Cortés, E., (2009) Competitive Strategies and Firm Performance: A Comparative Analysis of Pure, Hybrid and Stuck in-the-middle Strategies in Spanish Firms, *British Journal of Management*, 20, pp. 508–523.

Relationship between models of program execution and trace similarity metrics in malicious attacks detection

Sergey Frenkel[1]

[1]Federal Research Center, Computer Science and Control, Russian Academy of Sc., Moscow, Russia

E-mail: fsergei51@gmail.com

Abstract

The main mechanism for detecting malicious code and malicious attacks is to determine whether a suspicious code belongs to a cluster of previously stored codes. In this case, either clusters of benign and malignant codes, or only malignant ones, can be used [1]. The quality of clustering and, accordingly, the reliability of detection depends on both the models of the programs, which determine the clusters content, and distance-similarity metrics used between clusters and cluster elements. We will consider the program models which are defined on the dynamic trace of system calls or instructions. The drawback of statistical based program models used for the malicious codes detection is that benign labels (from the machine learning point of view) are usually more imprecise than malware labels, in sense that a new applications that is originally labeled benign might later have its label changed to malware, that is we can never be certain that a benign application is either really benign or just this is an undetected malware. The presentation discusses the possibility of sharing two classes of approaches to the detection of malicious code, methodologically free of such uncertainty in the problem space, which are based on the representation of the execution traces of programs by both Markov Chains (MC) and Hidden Markov Models [2,3], and using different metrics of the distance between the traces and their clusters, namely, Edit (Levenshtein) Distance, that is the minimal number of edit operations (delete, insert and substitute of a single symbol) required to convert one of compared sequence to the other(for clusters of ordered sets ,e.g., strings of symbols), and Jaccard distance for non-ordering elements of clusters (e.g., just a heap of the system calls) [1,4]. The goal is to suggest some conceptual models enabling to connect the distance metrics with some parameters of the Markov models representing these traces. The models are analyzed from the point of view of their applicability to problems of classification of sets, compatibility and computational complexity. As for MC based model,

the calls trace can be represented with a MC structure with corresponding transition probability matrix (including, depending on the used model of malicious code generation, there may be a MC with two absorbing states [3]) such that each distinct system call corresponds to one unique state of the MC, and the transition matrix depends on the system calls parameters random updating. In the framework of HMM model (S, O, A, B, π), the system call sequences are compared to the observed sequences O will be either normal or a malicious attack, where S = (Normal, Attack, Intrusion) is the state space, represents that the system call sequence has the following three states, Normal (N), Attack (A) (that indicates an attack activity), Intrusion (I) state indicates that an attack corrupted the normal behavior. It is important that we investigate this approach in conjunction with the handling of parameters and with a clustering of system calls based on such parameters, and we show for these Markov approaches how to connect the state transition over all trace with similarity metrics which are used widely for semantic analysis, taking into account also our recent result regarding the relationship between different sets similarity metrics [4].

Keywords: distance metrics, Markov models, malware detection.

Acknowledgements

Research partially supported by the Russian Foundation for Basic Research under grants RFBR 18-07-00669 and 18-07-00576.

References

[1] Jacob, G., Debar, H. and Filiol, E., (2008) Behavioral detection of malware: from a survey towards an established taxonomy, *J. Comput Virol*, 4, pp. 251–266.

[2] Mazeroff, G., De Cerqueira, V., Jens G. and Thomason, M.G., (2003) Probabilistic trees and automata for application behavior modeling, *41st ACM Southeast Regional Conference Proceedings*, pp. 435–440.

[3] Frenkel, S. and Zakharov, V., (2017) Brief Announcement: A Technique for Software Robustness Analysis in Systems Exposed to Transient Faults and Attacks, *Cyber Security Cryptography and Machine Learning (CSCML 2017)*, LNCS 10332, pp. 196–199.

[4] Dolev S., Ghanayim M., Binun A., Frenkel S. and Sun Y.S., (2017) Relationship of Jaccard andEdit Distance in Malware Clustering and Online Identification of 16th IEEE, *Int. Symposium on Network Computing and Applications Proceedings*, pp. 369–373.

A scatter search algorithm for the uncapacitated facility location problem

Telmo Matos[1] and Dorabela Gamboa[1]

E-mail: tsm@estg.ipp.pt

Abstract

Facility Location Problems are widely studied problems in the literature with several practical applications, reaching areas such as telecommunications, design of a supply chain management, transport utilities and water distribution networks. A well-known variant of this problem is the Uncapacitated Facility Location Problem (UFLP). This problem can be formulated as:

$$\text{Minimize } \sum_{i=1}^{m}\sum_{j=1}^{n} C_{ij}x_{ij} + \sum_{i=1}^{m} F_i y_i$$

$$\text{s.t. } \sum_{i=1}^{m} x_{ij} = 1, \forall j = 1, \ldots, n$$

$$x_{ij} \leq y_i, \forall j = 1, \ldots, n, \forall i = 1, \ldots, m$$

$$x_{ij} \geq 0, \forall j = 1, \ldots, n, \forall i = 1, \ldots, m$$

$$y_i \in \{0, 1\}, \forall i = 1, \ldots, m$$

Where m represents the number of possible locations to open a facility and n the number of costumers to be served. F_i indicates the fixed cost for opening a facility at location i. C_{ij} represents the unit shipment cost between a facility i and a costumer j. The continuous variable x_{ij} represents the amount sent from facility i to costumer j and y_i indicates if facility i is open (or not). The objective is to locate a set of facilities in such way that the total sum of the costs for opening those facilities and the transportation costs for serving all costumers is minimized. The UFLP problem has been widely studied for the past 50 years with the development of exact and heuristics methods [4–7] and numerous surveys [1–3]. We propose a Scatter Search (SS) procedure to solve effectively the UFLP. The general procedure starts with a set of initial solutions (seeds). Then the algorithm tries to produce a large number of random solutions with different characteristics from the seeds (diversification generation method). A local search procedure (improvement method) is applied to each of the solutions (and seeds) to improve them.

These improved solutions (and seeds) form the population. Then, a small size population (reference set) consisting of elite and diversified solutions (to force diversity) is obtained (forming the reference set update method). A subset of solutions is defined (subset generation method) and combined with solutions (usually in pairs) of that smaller population (combination method), obtaining new solutions that are improved (improvement method) again and then considered to enter the reference set (reference set update method). The process is repeated until some termination criteria are achieved. The performance of the proposed algorithm was tested on well-known benchmark producing extremely competitive results.

Keywords: UFLP, scatter search, metaheuristics, facility location problem.

References

[1] Basu, S. et al., (2015) Metaheuristic applications on discrete facility location problems: a survey, *OPSEARCH*, 52(3), pp. 530–561.

[2] Drezner, Z., (1995) Facility location: a survey of applications and methods, Springer Verlag.

[3] Farahani, R.Z. et al., (2013) Hub location problems: A review of models, classification, solution techniques, and applications, *Comput. Ind. Eng.*, 64(4), pp. 1096–1109.

[4] Li, Q. et al., (2016) Hybrid ant colony algorithm for the uncapacitated facility location problem, Shanghai Ligong Daxue Xuebao, *Journal Univ. Shanghai Sci. Technol.*, 38(4).

[5] Michel, L., Hentenryck and P. Van, (2004) A simple tabu search for warehouse location, *Eur. J. Oper. Res.*, 157, pp. 576–591.

[6] Pullan, W., (2009) A population based hybrid meta-heuristic for the uncapacitated facility location problem, *Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation - GEC '09*, pp. 475, ACM Press, New York.

[7] Sun, M., (2006) Solving the uncapacitated facility location problem using tabu search, *Comput. Oper. Res.* 33(9), pp. 2563–2589.

Numerical analysis of solutions of neural field equations with oscillatory coupling functions

Weronika Wojtak[1,2], Flora Ferreira[3], Estela Bicho[2] and Wolfram Erlhagen[1]

[1]Center of Mathematics, University of Minho, Guimarães, Portugal
[2]Algoritmi Center, University of Minho, Guimarães, Portugal
[3]Laboratory of Artificial Intelligence and Decision Support, INESC TEC, Porto, Portugal

E-mail: w.wojtak@dei.uminho.pt

Abstract

We numerically investigate solutions of neural field equation, a nonlocal differential equation used to describe the large-scale spatio-temporal dynamics of neuronal populations [1]. This class of equations has been successfully used in the past as a mathematical framework for modeling a wide range of neurobiological phenomena, including multi-item working memory [2], which involves holding and processing of information on the time scale of seconds. Neural fields can exhibit a variety of spatiotemporal dynamics, for a recent review of neural field models and their solutions we refer the reader to [3]. For a working memory application of the neural field model, we are concerned with stationary, i.e., time–independent, localized solutions (bumps), that are initially triggered by brief sensory inputs and subsequently become self-sustained by recurrent interactions within the population of neurons [3,4]. In the original work of Amari, this inter-neuron connectivity function has a Mexican hat shape, with local excitation and distal inhibition, mimicking the interaction of excitatory and inhibitory neural sub-populations [1]. He proved the existence and stability of 1-bump solutions in this case. However, the Mexican hat connectivity does not generally support a stable pattern of two or more regions of high excitation. To allow such multi-bump solutions, a special class of oscillatory coupling functions was introduced by Laing et al. [2]. In [4], a specific version of this class of couplings was proposed, with an additional parameter α controlling the distance between consecutive zeros of the function. This parameter determines the spatial ranges of excitation and inhibition within the field, and consequently the shape of bump solutions. It therefore allows to control the maximum number of bumps that may exist in a field of a given finite length. With a working memory application of the dynamic field model in mind, it is important to understand how the shape

and spatial extension of multi-bump solutions change as the parameters of the field are varied. An important technique for investigating solutions of differential equations is numerical continuation. It involves finding solutions of interest, determining their stability and following them as parameters are changed. Here, we use pseudo-arclength continuation to find and follow solutions of neural field equations, based on the strategy outlined in [5]. More precisely, we carry out the bifurcation analysis using Fourier series, treating α as a bifurcation parameter, which complements the work in [4]. We show that for a specific choice of resting state, the width of both stable and unstable bump solutions decreases significantly when parameter α is increased. We then extend our numerical investigation to N-bump solutions showing that similar relationship between the excitation length and α also holds for solutions with two or more regions of high activation. Importantly from an application point of view, we investigate how the maximum number of bumps in a given finite interval is affected by changes in α.

Keywords: integro-differential equation, neural field, working memory, numerical continuation.

Acknowledgements

The work received financial support from FCT through the PhD fellowship PD-BD-128183-2016.

References

[1] Amari, S., (1977) Dynamics of pattern formation in lateral-inhibition type neural fields, *Biological Cybernetics,* 27(2), pp. 77–87.

[2] Laing, C.R., Troy, W.C., Gutkin, B. and Ermentrout, G.B., (2002) Multiple bumps in a neuronal model of working memory, *SIAM Journal on Applied Mathematics,* 63(1), pp. 62–97.

[3] Coombes, S., (2005) Waves, bumps, and patterns in neural field theories, *Biological Cybernetics,* 93(2), pp. 91–108.

[4] Ferreira, F., Erlhagen, W. and Bicho, E., (2016) Multi-bump solutions in a neural field model with external inputs, *Physica D: Nonlinear Phenomena,* 326, pp. 32–51.

[5] Laing, C.R., (2014) Numerical bifurcation theory for high-dimensional neural models, *The Journal of Mathematical Neuroscience,* 4(1), pp. 13.

Contributed Posters

Determinants of domestic water consumption: a case study in northern Portugal

A. Manuela Gonçalves[1], Cristina Matos[2,3], Ana Briga-Sá[2,3], Sandra Pereira[2,3], Isabel Bentes[2,3] and Diana Faria[2,3]

[1]CMAT – Centro de Matemática, DMA-Departamento de Matemática e Aplicações, Universidade do Minho, Portugal
[2]C-MADE – Centro de Materiais e Tecnologias Construtivas, Universidade da Beira Interior, Portugal
[3]ECT – Escola de Ciências e Tecnologia, Universidade de Trás-os-Montes e Alto Douro, Portugal

E-mail: mneves@math.uminho.pt

Abstract

Efficient management of water resources, in both rural and urban areas, requires a full understanding of existing patterns of water use. Water demand management has been mainly focused on meeting agriculture water demand, whereas domestic water demand is largely ignored, and household water consumption has not been thoroughly researched in the majority of the countries. The World Health Organization (WHO) defined domestic water as water used for all domestic purposes including consumption, bathing and food preparation [1], [5]. Information regarding domestic water consumption is vital but is still lacking. The success of domestic water demand management strategies depends on identifying the determinants, and their interaction, that influence water consumption at a household scale [3]. This paper presents an empirical analysis of domestic water consumption and factors influencing water consumption in Vila Real County, in Northern Portugal. Through a field survey, the data were collected from December 2016 to January 2017 from 245 urban and rural households in 20 parishes of Vila Real County, and determinants influencing domestic water consumption are studied. Data analysis was performed by descriptive statistics, non-parametric tests and ordinal regression, namely by comparing the two groups (urban and rural households) [2], [4].

Keywords: survey, domestic water consumption, non-parametric tests analysis, ordinal regression.

Acknowledgements

This work was partially supported by the Research Centre of Mathematics of the University of Minho with the Portuguese Funds from the FCT - Fundação para a Ciência e a Tecnologia, through the Project PEstOE-MAT-UI0013-2017. This work was partially funded by POCI-0-0145-FEDER-016730 Project (PTDC-AAG-REC-4700-2014) with the designation ENERWAT: Water to energy: characterization, modelling and measures for the reduction of urban and rural household consumption, financed by the Foundation for Science and Technology and co-financed by the European Regional Development Fund (FEDER) through the COMPETE 2020 - Programme Operational Competitiveness and Internationalization (POCI). This work was partially supported by the FCT (Portuguese Foundation for Science and Technology) through the project PEst-OE-ECI-UI4082-2013 (C-MADE).

References

[1] World Heatlth Organization, (2002) Guidelines for Drinking-Water Quality: Addendum Microbiological Agents in Drinking Water, 2nd ed.; WHO: Geneva, Switzerland.

[2] Higgins J.L., (2004) Introduction to Modern Nonparametric Statistics. Toronto, Canada: Thomson.

[3] Loureiro D., Pinheiro L., Rebelo M., Salgueiro A.R., Medeiros N., Covas D., Alegre H., (2008) Estudo dos fatores mais relevantes que influenciam o consumo doméstico de água: O caso de estudo do complexo de edifícios: Twin-Towers. LNEC.

[4] O'Connel A.A., (2006) Logistic Regression Models for Ordinal Response Variables, Series: Quantitative Applications in the Social Sciences, SAGE Publications, Inc., London.

[5] Vieira A.S., Ghisi E., (2016) Water-energy nexus in houses in Brazil: comparing rainwater and gray water use with a centralized system, *Water Science and Technology: Water Supply*, 16 (2), pp. 274–283.

Environmental stochasticity and woodland caribou populations: the effects of inter-annual variation in the number and size of fires

Carla Francisco[1], Teresa Oliveira[2] and Steve Cumming[1]

[1]Laval University, Department of Wood Science and Forestry, Faculty of Forestry, Geography and Geomatics, Québec, Canada
[2]Centro de Estatística e Aplicações (CEAUL), Portugal

E-mail: carla.francisco.1@ulaval.ca

Abstract

A new way of simulating fire variation and analyze the population viability, under a caribou population model, to determine what is the probability and how quickly the population can go extinct within the next years [3]. Fires that consume large areas of forest, affect Caribou populations [2], since for survival they depend on lichens, which are found in greatest abundance in old growth forests. In this context and due to its extension of damage it is crucial to improve fire prevention and detection models. This project aims to improve the existing Landscape Fire model, see Cumming, S. (2017) [1], parameterized from data, that changes the forest age structure, by including inter-annual variation in the fire regime in each year.

Keywords: landscape fire model, negative binomial count model, spatial simulation, wildfire.

Acknowledgements

Research partially funded by FCT Fundação para a Ciência e a Tecnologia, Portugal, through the projects PEst-OE-MAT-UI0006-2014, UID-MAT-00006-2013.

References

[1] Cumming, S., (2017) Parameterizing landscape fire models from data: Recent progress and remaining challenges. Cef, Université du Québec à Montréal, 11e Colloque Annuel du CEF.

[2] Environment Canada, (2012) Recovery Strategy for the Woodland Caribou (Rangifer tarandus caribou),Boreal population, in Canada. Species at Risk Act Recovery Strategy Series, Environment Canada, Ottawaxi +138pp.

[3] Sorensen et al, (2008) Determining sustainable levels of cumulative effects for boreal Caribou. Journal of Wildlife Management, 72(4), pp. 900–905.

Achieving COBS from a balanced mixed model

Carla Santos[1,2], Cristina Dias[1,3], Célia Nunes[4] and João T. Mexia[1,5]

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

E-mail: carla.santos@ipbeja.pt

Abstract

Models with orthogonal block structure, introduced by [4], are linear mixed models whose variance-covariance matrix is a linear combination of known pairwise orthogonal orthogonal projection matrices that add up to the identity matrix. Imposing a commutativity condition on them, we get models with commutative orthogonal block structure, COBS. COBS are such that the least square estimators, LSE, are the best linear unbiased estimators, BLUE, whatever the variance components. Resorting to the algebraic structure of COBS, we study the possibility of achieving COBS from the extension of a balanced mixed model.

Keywords: models with commutative orthogonal block structure, Jordan algebra, balanced mixed models, B-matrices.

Acknowledgements

This work was partially supported by national founds of FCT-Foundation for Science and Technology under UID-MAT-00297-2013 and UID-MAT-00212-2013.

References

[1] Nelder, J.A., (1965) 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, pp. 147–162.

[2] Nelder, J.A., (1965) The analysis of randomized experiments with orthogonal block structure II, Treatment structure and the general analysis of variance, *Proceedings of the Royal Society*, Series A 283, pp. 163–178.

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

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

[5] Zmyślony, R., (1978) A characterization of best linear unbiased estimators in the general linear model, *Mathematical Statistics and Probability Theory*, 2, pp. 365–373.

Completing a random effects model for an imbedded linear regression

Carla Santos[1,2], Cristina Dias[1,3], Célia Nunes[4] and João T. Mexia[1,5]

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

E-mail: carla.santos@ipbeja.pt

Abstract

To test the main effects and interactions in a random effects model with u factors we consider an orthogonal partition in subspaces associated with the sets of factors. Availing ourselves of the much larger dimension of one of the subspaces in the orthogonal partition we enrich the model. Through the example presented it is possible to see the advantage of the proposed model.

Keywords: factor crossing, linear regressions, orthogonal partition.

Acknowledgements

This work was partially supported by national founds of FCT-Foundation for Science and Technology under UID-MAT-00297-2013 and UID-MAT-00212-2013.

References

[1] F. Carvalho, J.T. Mexia, C. Santos, C. Nunes, (2015) Inference for types and structured families of commutative orthogonal block structures, *Metrika*, Vol.78, pp. 337–372.

[2] M. Fonseca, J.T. Mexia, R. Zmyslony, (2006) Binary Operations on Jordan algebras and orthogonal normal models, *Linear Algebra Appl.*, 117, Vol.1, pp. 75–86.

[3] M. Fonseca, J.T. Mexia, R. Zmyslony, (2008) Inference in normal models with commutative orthogonal block structure, *Acta et Commentationes Universitatis Tartuensis de Mathematica*, 12, pp. 3–16.

[4] A.I. Khuri, T. Mathew, B.K. Sinha, (1998) *Statistical Tests for Mixed Linear Models*, New York, Wiley.

[5] J.T. Mexia, R. Vaquinhas, M. Fonseca and R. Zmyslony, (2010) COBS: segregation, matching, crossing and nesting, *Latest Trends and Applied Mathematics, Simulation, Modelling*, 4-th International Conference on Applied Mathematics, Simulation, Modelling (ASM'10), pp. 249–255.

[6] M.M. Oliveira, (2002) Modelação de Séries Emparelhadas de Estudos com Estrutura Comum, PhD thesis, Universidade de Évora.

Structured families of models

Cristina Dias[1], Carla Santos[2] and João T. Mexia[3]

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

E-mail: cpsilvadias@gmail.com

Abstract

In this work we study structured families of models, whose matrices correspond to the treatments of a base design, see [3]. We also consider families of models divided into subfamilies that correspond to these treatments. Since the matrices have all the same order, we are in the balanced case where we have the same number of degrees of freedom for the error for each treatment. The ANOVA and related techniques are, in the balanced case, robust techniques for the case of heteroscedasticity and even more for the case of non-normality, see [4] and [2]. We are mainly interested in basic models with orthogonal structure. We present this structure and show how to apply these models in the study of structured families.

Keywords: ANOVA, symmetric matrices, ortogonal structure, base design.

Acknowledgements

This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID-MAT-00297-2013 (Centro de Matemática e Aplicações).

References

[1] Ito, P. K., (1980) Robustness of Anova and Macanova Test Procedures, P. R. Krishnaiah (ed) Handbook of Statistics1, Amsterdam: North Holland, 1, pp. 199–236.
[2] Dias, C., (2013) Modelos e Famílias de Modelos para Matrizes Estocásticas Simétricas, Ph.D. Thesis, Évora University.
[3] Scheffé, H., (1959) The Analysis of Variance, New York: John Wiley & Sons.

Modeling oil prices with non-stationary extreme values distribution: model ARMA-Gev, case Algeria (1973-2014)

Cheraitia Hassen[1]

[1]Department of Mathematics- JIJEL's University, Algeria

E-mail: hassen.stat@hotmail.fr

Abstract

In this paper we have studied non-stationary extreme values. In the traditional case, we have often focused on the independent and identically distributed variables (i.i.d.). This hypothesis is an unavoidable imposition in several fields of applications, for that it is advisable to replace it by a certain form of dependence, for example by introducing the notion of trend by considering the position parameter $\mu(t) = \beta_0 + \beta_1 t$ and the scale parameter $log(\sigma) = \beta_0 + \beta_1 t$ as functions of time t or a stochastic trend by considering that innovations are a random walk process. Some model of non-stationary extreme values have been tried on a series of the annual minimum price of Algerian oil (1973-2014), using methods of time series (Dicky Fuller tests) in order to make the series stationary. Once our series became stationary, we estimated the parameters of the generalized extreme values law (GEV) by the method of maximum likelihood (ML) and calculated levels of returns for some periods; as an example, we will have to wait about 100 years to see a monthly oil price of \$29.80 / barrel (that is, we have a probability of 0.01 to have an oil catch of \$29.80 each year).

Keywords: non-stationary series, generalized extreme values law (GEV), trend, maximum likelihood, level/return time.

References

[1] Brockwell, P.J. and Davis, R.A., (2009) Time series: Theory and Methods, Springer, 2nd edition.

[2] CHEN, H. and RAO, R., (2002) Testing hydrologic time series for stationarity, *J. Hydrol. Eng.*, 7, pp. 129–136.

[3] CLARKE, R.T., (2002) Estimating trends in data from the Weibull and a generalized extreme value distribution, *Water Resour. Res.*, 38, 25-1-25-10.

[4] COLES, G.S., (2001) An introduction to statistical modeling of extreme values, Springer (editeur).

[5] EMBRECHTS, P., RESNICK, S. and SAMORODNITSKY, G., (1999) Extreme value theory as a risk management tool, *North American Actuarial Journal*, 26, pp. 30–41.
[6] Embrechts P., Kluppelberg, C. and Mikosch, T.,(1997) Modelling Extremal Events for insurance and Finance, Springer, New York.
[7] Fisher, R.A. and Tippet, L.H.C., (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Lamb. Philos. Soc.*, 24, pp. 180–190.
[8] GENÇAY, R. and SELÇUK, F., (2004) Extreme value theory and value-at-risk: Relative performance in emerging markets, *International Journal of Forecasting*, 20, pp. 287–303.
[9] Gnedenko, B.V., (1943) Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.*, 44, pp. 423–453.
[10] Gumbel, E.J., (1958) Statistics of Extremes, Columbia University Press, New York.
[11] Jenkinson, A.F., (1955) The frequency distribution of the annual maximum (or minimum) of meteorological elements, *Quart. J. R. Met .Soc.*, 81, pp. 158–171.
[12] KATZ, R.W., (1999) Extreme value theory for precipitation: sensitivity analysis for climate change, *Adv. Water Resour.*, 23, pp. 133–139.
[13] KHARIN, V.V. and ZWIERS, F.W., (2005) Estimating extremes in transient climate change simulations, *J. Clim.*, 18, pp. 1156–1173.
[14] LONGIN, F.M., (1998) Value at risk: Une nouvelle approche fondée sur les valeurs extrêmes, *Annales d'économie et de statistique*, 52, pp. 23–51.
[15] LONGIN, F.M., (2000) From value at risk to stress testing: The extreme value approach, *Journal of Banking and Finance*, 24, pp. 1097–1130.
[16] McNEIL, A.J., (1998) Calculating quantile risk measures for financial time series using extreme value theory, Department of Mathematics, ETH, Swiss Federal Technical University E-Collection.
[17] ONOZ, B. and BAYAZIT, M., (2003) The power of statistical tests for trend detection, *Turkish J. Eng. Env. Sci.*, 27, pp. 247–251.
[18] SCARF, P.A., (1992) Estimation for a four parameter generalized extreme value distribution, *Comm. Statist. A Theory Methods*, 21, pp. 2185–2201.
[19] ZHANG, X., HARVEY, K.D., HOGG, W.D. and YUZY, T.R., (2001) Trends in Canadian stream flow, *Water Resour. Res.*, 37, pp. 987–999.
[20] Madsen, H., (2008) Time Series Analysis, Chapman & hall/CRC.

Evaluation of the compatibilization role of organoclays on PA6/PP nanocomposites

Fátima De Almeida[1], Eliana Costa e Silva[1] and Aldina Correia[1]

[1]Escola Superior de Tecnologia e Gestão, Instituto Politécnico do Porto, Portugal

E-mail: mff@estg.ipp.pt

Abstract

There is a large amount of on-going research that aims at developing nanocomposites materials for various industrial applications on different sectors, such as, automotive, electronics, food packaging, biotechnology, and biomedical. Several studies have shown that the addition until to 5% in weight of layered silicates can lead to a significant enhancement in stiffness and strength [1], flame retardancy [2], gas barrier properties [3], ionic conductivity [4], electrical properties [5], thermal stability [6], and biodegradability [7]. The characteristics required for this type of materials are usually obtained in polar matrices. However, it has proven to be a challenge in nonpolar matrices. In the case of polymer blends as a polar polyamide (PA) and a nonpolar polypropylene (PP), a compatibilizer must be introduced to improve the interactions between them [1, 8, 9]. Very often, a polyolefin grafted with maleic anhydride is used, leading to good interfacial properties of immiscible polymer pairs with a large decrease of the interfacial tension between the blend components and decreases in the coalescence of the dispersed phase [1, 8]. In order to reach those properties, layered silicates organophilic modified as alternative compatibilizer were used due to its lightweight, economic competitiveness, and wide availability. The present work aims at comparing the efficiency of two of these organoclays (C15A and C30B) in a PA6/PP polymeric blend as compatibilizer. PA/PP (70/30) nanocomposite was prepared successfully by melt processing technique using a twin-screw extruder with Cloisite 15A and C30B modified nanoclays having good diversity of CEC capacity, hydrophobicity, d-spacing, and dispersion capability [9]. Morphological analysis by Rheology, X-ray diffraction (XRD), Scanning Electronic Microscopy (SEM) and Transmission Electron Microscopy (TEM) revealed a good affinity of organophilic clays, C15A and C30B, with PA6/PP blend, even without the compatibilizer PP-g-MA, the uncompatibilized blend. In order to analyze the existence of differences between uncompatibilized and compatibilized polymer blend, when organoclays are added, the Kruskal-Wallis test. The use of this non-parametric alternative of the one-way ANOVA (analysis of variance technique), which is usually considered as ANOVA on ranks, over the

parametric ANOVA is justified by the small number of observations that is possible to obtain from the experimental test.

Keywords: blend nanocomposites, organoclays, compatibilization, Kruskal-Wallis test.

References

[1] R. Scaffaro, M. C. Mistretta, F. P. La Mantia, (2008) Compatibilized polyamide 6/polyethylene blend-clay nanocomposites: Effect of the degradation and stabilization of the clay modifier, *Polymer Degradation and Stability*, 93, pp. 1267–1274.

[2] Y. Hu, Y. Tang, R. Zhang, Z. Chen, and W. Fan, (2005) Study on the properties of flame retardant polyurethane/organoclay nanocomposite, *Polym Degrad Stab*, 87, pp. 111–116.

[3] J.H. Chang and Y.U. An, (2002) Nanocomposites of polyurethane with various organoclays: Thermomechanical properties, morphology, and gas permeability, *J. Polym. Sci. B: Polym. Phys.*, 40, pp. 670–677.

[4] J.D. Nam, S.D. Hwang, H.R. Choi, J.H. Lee, K.J. Kim, and Heo, (2005) Electrostrictive polymer nanocomposites exhibiting tunable electrical properties, *Smart Mater. Struct.*, 14, pp. 87–90.

[5] H.T. Lee and L.H. Lin, (2006) Waterborne polyurethane/clay nanocomposites: novel effect of the clay and its interlayer ions on the morphology and physical and electrical properties, *Macromolecules*, 39, pp. 6133–6141.

[6] Y.I. Tien and K.H. Wei, (2002) The effect of nanosized silicate layers from montmorillonite on glass transition, dynamic mechanical, and thermal degradation properties of segmented polyurethane, *J. Appl. Polym. Sci.*, 86, pp. 1741–1748.

[7] E.H. Jeong, J. Yang, H.S. Lee, S.W. Seo, D.H. Baik, J.Kim, and J.H. Youk, (2008) Effective preparation and characterization of montmorillonite/poly (ϵ-caprolactone)based polyurethane nanocomposites, *J. Appl. Polym. Sci.*, 107, pp. 803–809.

[8] M. Xanthos, S. S. Dagli, (1991) Compatibilization of polymer blends by reactive processing, *Polymer Engineering & Science*, 31, pp. 929–935.

[9] P.C. LeBaron, Z. Wang, and T.J. Pinnavaia, (1999) Polymer-layered silicate nanocomposites: an overview, *Applied Clay Science*, 15, pp. 11–29.

Bootstrap approaches to analyze the significance of variance components in Poisson mixed models

Susana Faria[1], A. Neco-Oliveira[2] and Raquel Menezes[1]

[1]Departamento de Matemática e Aplicações, Universidade do Minho, Portugal
[2]Instituto Federal Goiano, Campus Morrinhos, Brazil

E-mail: sfaria@math.uminho.pt

Abstract

Generalized linear mixed models (GLMMs) are used in situations where a nonnormal response is related to a set of predictors and the responses are correlated [1]. These models are useful for accommodating the overdispersion in Poisson regression models with random effects. The main difficulty of these models is the estimation of their parameters [2] and although this difficulty has to a large extent been overcome, there are still many unresolved problems. For example, analytical formulas for standard errors and confidence intervals for random effects are often not available. In this work, we consider both parametric and nonparametric bootstrap methods [3] for assessing the significance of variance components associated to the random effects in the context of poisson mixed models. Bootstrap methods are applied to both simulated and real data set consisting of 7 years of dengue incidence in the State of Goiás, Brazil. The results indicated that the coverage probabilities were above 90%, when the resampling process at the factor's level analysis under involves about 50% of the original data. The discussed bootstrap procedures prove to be an useful tool for the analysis of the significance of random effects in the context of poisson mixed models.

Keywords: bootstrap, poisson mixed models, simulation study, variance components.

References

[1] Breslow N.E. and Clayton D.G., (1993) Approximate inference in generalized linear mixed models, *Journal of the American Statistical Society*, 88, pp. 9–25.

[2] Casals M., Langohr K., Carrasco J.L. and Ronnegard L., (2015) Parameter estimation of Poisson generalized linear mixed models based on three different statistical principles: a simulation study, *SORT-Stat Oper Res T.*, 39, pp. 281–308.

[3] Efron B. and Tibshirani, R.J., (1994) An introduction to the bootstrap, CRC press.

Third-grade fluid model

Fernando Carapau[1], Paulo Correia[1] and Luís M. Grilo[2]

[1]Universidade de Évora, Departamento de Matemática e CIMA, Portugal
[2]Instituto Politécnico de Tomar e CMA/FCT/UNL, Portugal

E-mail: flc@uevora.pt

Abstract

Three-dimensional numerical simulations of non-Newtonian fluid flows are a challenging problem due to the particularities of the involved differential equations leading to a high computational effort in obtaining numerical solutions, which in many relevant situations becomes infeasible. Several models has been developed along the years to simulate the behavior of non-Newtonian fluids together with many different numerical methods. In this work we use a one-dimensional hierarchical approach to a proposed generalized third-grade fluid with shear-dependent viscoelastic effects model. This approach is based on the Cosserat theory related to fluid dynamics and we consider the particular case of flow through a straight and rigid tube with constant circular cross-section. With this approach, we manage to obtain results for the wall shear stress and mean pressure gradient of a real three-dimensional flow by reducing the exact three-dimensional system to an ordinary differential equation. This one-dimensional system is obtained by integrating the linear momentum equation over the constant cross-section of the tube, taking a velocity field approximation provided by the Cosserat theory. From this reduced system, we obtain the unsteady equations for the wall shear stress and mean pressure gradient depending on the volume flow rate, Womersley number, viscoelastic coefficients 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.

Keywords: one-dimensional model, generalized third-grade model, shear-thickening fluid, shear-thinning fluid, cosserat theory.

Acknowledgements

This work was supported by FCT - Fundação para a Ciência e a Tecnologia under the project UID-MAT-04674-2013.

References

[1] Carapau, F., Correia, P., Grilo, L.M. and Conceição, R., (2017) Axisymmetric Motion of a Proposed Generalized Non-Newtonian Fluid Model with Shear-dependent Viscoelastic Effects, *IAENG International Journal of Applied Mathematics*, accepted for publication.

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

[3] Fosdick, R.L. and Rajagopal, K.R., (1980) Thermodynamics and stability of fluids of third grade, *Proc. R. Soc. Lond. A.*, v.339, pp. 351–377.

Implementation of bootstrap methods for accuracy assessment of space-time data modelling

Gustavo Soutinho[1] and Raquel Menezes[2]

[1]University of Minho, Portugal
[2]Centre of Molecular and Environmental Biology, Department of Mathematics and Applications, University of Minho, Portugal

E-mail: gustavosoutinho@sapo.pt

Abstract

The large and small-scale variation of a spatio-temporal stochastic process can be separately estimated by using a two-stepwise approach. Firstly, a generalized linear model (GLM) is adopted, regarding the data distribution to approximate the trend and seasonality, by relaxing the assumption of non-correlated errors. This procedure provides point estimates of the regression parameters, specifying the large-scale variation. Secondly, the small-scale variation, imposed by the underlying dependence of the stationary residual, is estimated through a spatio-temporal variogram, such as the sum-metric model which accounts for the space-time interaction [1]. As a consequence of the assumption of independence of the residuals, the maximum likelihood estimates (MLEs) for the regression parameters usually give us over-optimistic standard errors [2] and, consequently, not enable us to identify whether an independent variable have or not a significant contribution into the response variable. Moreover, the common methods used to obtain the parameters' estimates of the spatio-temporal variogram (e.g. spatial, temporal and joint variances or ranges) do not provide the standard errors associated. To overcome these drawbacks, bootstrap approaches are integrated into the estimation of the large and small-scale variation components. This work aims to compare parametric and non-parametric bootstrap methods and to propose alternatives to assess the accuracy of estimates associated to the parameters of the large and small-scale variations. The parametric bootstrap approach considers replicates, drawn from a multivariate normal distribution, with expectation defined by the trend model and covariance matrix obtained from the sum-metric variogram. This parametric method may be used to analyse the significance of the parameters in the large-scale variation component. In this study the following independent variables were considered: location of the monitoring station (longitude and latitude); week reference (1 up to 212); type of site (background, industrial or traffic) and the type of environment

(urban, suburban or rural). To model the seasonality of data, we used an harmonic regression, assuming a period equal to 52-week. The non-parametric bootstrap approaches are based on *moving block* and *random block* bootstrap methods, two different ways of drawing the dependent time series observations, when taking fixed data in space dimension. As in environmental sciences, typically, the spatial resolution, defined by the monitoring stations, is smaller than the time resolution, these sequential time blocks procedures may prove to be useful. For both methods, in this work, the overlapping of blocks in the time dimension is considered. The main idea of the *moving block* bootstrap consists of dividing the temporal data, $X_1, ..., X_T$ into blocks of consecutive observations of length l. Each new block slides δ time units, allowing for a total of $k+1$ blocks, defined as $(X_1, \ldots, X_l)$, $(X_{1+\delta}, \ldots, X_{l+\delta})$, $\ldots$, $(X_{1+k\delta}, \ldots, X_{l+k\delta})$, such that $l + k * \delta \leq T$. Under the *random block* bootstrap, replicates are defined by M blocks of consecutive observations with the same length l randomly selected from the start time between 1 and $T-l+1$ [3]. The weekly average of the NO_2 between 1 January 2013 and 31 December 2016, from 50 monitoring stations located on Mainland of Portugal, are used to illustrate the differences among bootstrap methods. The NO_2 is a good marker for the exposure of the air quality and is among the main pollutants with significant impact on environmental and health problems. Results reveal that the bootstrap approaches are particularly appropriate to distinguish non-significant independent variables in GLM, when adopting the two-stepwise approach presented in [1], as well as, they are useful to analyse the significance of estimates of variogram parameters. In the former case, we can conclude that, among all the covariates initially considered in GLM, just type of environment and type of site have significant influence on the response variable (NO_2). The parametric bootstrap method is preferable when the data distribution is known but requires more computational costs, whereas a non-parametric method should be used when we wish to avoid distributional assumptions and is computationally faster. The random block bootstrap allows to improve the accuracy of the estimates due to the possibility of choosing a larger number of replicates.

Keywords: bootstrap methods, spatio-temporal data, geostatistics, air pollution.

Acknowledgements

This research was financed by Portuguese Founds through FCT - Fundação para a Ciência e Tecnologia.

References

[1] Menezes R., Piairo H., Garcia-Soidánand P. and Sousa I., (2016) Spatial temporal modellization of the NO2 concentration data through geostatistical tools, *Journal of*

Statistical Methods & Applications, issue1, pp. 107–124.

[2] Monteiro A., Menezes R. and Silva M.E., (2017) Modelling spatio-temporal data with multiple seasonalities: the NO2 Portuguese case, *Spatial Statistics journal*, Vol.22, Part 2, pp. 371–387.

[3] Kreiss J.P. and Paparoditis E. Rejoinder, (2011) Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society*, 40(4), pp. 393–395.

LQ-moments for right censored data

Ivana Malá[1]

[1]University of Economics, Prague, Czech Republic

E-mail: malai@vse.cz

Abstract

The L-moments, certain linear functions of the expectations of order statistics, were introduced by Hosking in 1990 (Hosking, 1990), for the right censored data in (Wang, at al., 2010). Then more analogues of these robust moments have been defined, such as TL-moments (trimmed L-moments). In the poster, LQ-moments are applied replacing the expectations of ordered statistics with selected quantiles or their weighted means (quick estimators). The LQ-moments always exist without any assumption on the probability distribution. It is easy to evaluate these moments from the known distributions or estimate them from a random sample (Mudholkar, Hutson, 1998). In the contribution, a problem of the estimation of these moments from the right censored data is treated. The Kaplan-Meier estimator (Kaplan, Meier, 1958) of the survival function is used to estimate quantiles and to evaluate the sample LQ-moments. Simulations are used to describe properties of this estimator for the lognormal distribution (the two-parametric version of the positively skewed distribution frequently applied in the survival analysis). The strong impact of the choice of probabilities of applied quantiles and weights of weighted means on the results is shown. Sample sizes from 50 to 500 and censoring rates up to 50 percent are used with 10 000 replications. All computations are performed in the program R (R Development Core Team, 2008).

Keywords: robust moments, LQ moments, right censored data.

Acknowledgements

The support of the funds of institutional support of a long-term conceptual advancement of science and research number IP 400 040 at the Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic is gladly acknowledged.

References

[1] Hosking, J.R.M., (1990) L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics, *Journal of the Royal Statistical Society. Series B (Methodological)*, 52, pp. 105–124.

[2] Kaplan, E.L. and Meier, P., (1958) Nonparametric Estimation from incomplete,*Journal of the American Statistical Association*, 53, pp. 457–481.

[3] Mudholkar G.S. and Hutson, A.D., (1998) LQ-moments: Analogs of L-moments, *Journal of Statistical Planning and Inference*, 71, pp. 191–208.

[4] R. Development Core Team, (2008) R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, URL: http://www.R-project.org.

[5] Shabri, A. and Jemain, A.A., (2007) LQ-moments for Statistical Analysis of Extreme Events, *Journal of Modern Applied Statistical Methods*, 6, pp. 228–238.

[6] Wang, D., Hutson, A.D. and Miecznikowski, J.C., (2010) L-moment estimation for parametric survival models given censored data, *Statistical Methodology*, 7, pp. 655–667.

Exploring decision making on hospitality brands: emotional or rational factors

Joana Campos[1], Fernanda A. Ferreira[1,2] and Mónica Oliveira[1]

[1]Superior School of Hospitality and Tourism of Polytechnic Institute of Porto, Portugal
[2]Applied Management Research Unit (UNIAG), Portugal

E-mail: faf@esht.ipp.pt

Abstract

The aim of this paper was to elaborate an investigation that allowed to correlate emotional and rational factors with the decision to purchase hotel services. Consequently, it led the authors to seek to know about loyalty to hotel brands, knowing in a superficial way this reality and influences of the factors mentioned above. When building a descriptive and correlational study, a questionnaire was developed based on the information gathered in the literature review that allowed the authors to accept/reject the research hypotheses: H1: there are differences regarding the degree of influence of emotional factors depending on the motivation for the purchase; H2: there are differences regarding the degree of influence of rational factors depending on the motivation for the purchase; H3: loyalty to a hotel brand is related to customer satisfaction with the hotel service; H4: loyalty to a hotel brand is positively related to the brand's prestige in the market. Characterizing the performance of the hotel unit as the ability to meet the preferences of the regular customer, ability to overcome a complaint in a satisfactorily way and quality of its employees, it was possible to conclude that it explains the consumer's loyalty to the brand in a statistically significant way. On the other hand, in a multivariate logistic regression analysis, we can draw significant conclusions regarding the brand's prestige in the market, the awards it has received, the ability to meet the habits and preferences of its regular customer and the ability to resolve a complaint. These results are based on the bivariate analysis performed in the test of hypothesis four and the information found in the literature [1], [2], [3]. Thereby, the results from the test of hypothesis four meets the theory of Kim & Yu [4]: the prestige of the brand is factor of perceived quality and may influence buying decision, which may be applied to hotel brands as this investigation confirmed. In conclusion, winning customer loyalty to service brands is more challenging than to a product brand. This may happen because the intangibility is associated with a greater risk in the purchase of a hotel service. Thus, building a strong brand on the market can

become a facilitator of economic profitability by simplifying the consumer decision-making process by providing guarantees of good service. Finally, the literature suggests that loyalty to hotel brands is strongly influenced by factors related to the assessment of consumer experience, summarizing that perceived value, quality service and brand confidence influence their purchasing decision [4].

Keywords: consumer behavior, emotional factors, rational factors, hotel management, brand-consumer relationship, brand loyalty.

Acknowledgements

Research partially funded by UNIAG, R&D unit funded by the FCT – Portuguese Foundation for the Development of Science and Technology, Ministry of Science, Technology and Higher Education, under the Project UID-GES-04752-2016.

References

[1] Cohen, S. Moita, M. and P., Girish, (2014) Consumer Behaviour in Tourism: concepts, influences and opportunities, *Current Issues in Tourism*, 17(10), pp. 872–909.

[2] Hwang, J., Lee, J. and Ok, C., (2016) An emotional labor perspective on the relationship between customer orientation and job satisfaction, *International Journal of Hospitality Management*, 54, pp. 139–150.

[3] Manhas, P. and Tukamushaba, E., (2015) Understanding service experience and its impact on brand image in hospitality sector, *International Journal of Hospitality Management*, 4, pp. 77–87.

[4] Kim, T. and Yu, J., (2010) The Differential Roles of Brand Credibility and Brand Prestige in a consumer brand choice, *Psychology & Marketing*, 2(7), pp. 662–678.

Spectral clustering tools for e-Learning analitics

Juan Luis García Zapata[1], Maria Clara Grácio[2] and Irene Rodrigues[3]

[1]Departamento de Matemáticas, Universidad de Extremadura, Espanha
[2]Universidade de Évora, Departamento de Matemática e CIMA, Portugal
[3]Universidade de Évora, Departamento de Informática e LISP, Portugal

E-mail: mgracio@uevora.pt

Abstract

The study of complex systems through has proved to be very useful, especially in the analysis of social networks. Using clustering techniques, communities are detected in networks of friendship or shared interests. It is done also in networks of scientific collaboration or networks of employment and professional services, see [1]. In this work we study a network of this second type, formed by the students and the disciplines that they have cursed in the e-learning system of the University of Évora. We apply a spectral clustering tool that we have developed, based on the second eigenvector of the Laplacian matrix of the graph. This technique allows to avoid the high cost of combinatorial algorithms using numerical methods of linear algebra, well established in scientific computation, see [2]. In the case under study, the detection of communities identifies trends (such as training profiles that are frequently chosen) and to compare these data with the usual metrics in learning analytics such as performance, study leaving, or repetition rates. In addition to this trajectory detection, our technique can help to the university manager to decide on the investment of resources (mainly attention, guidance and tutoring) over students according to their community profile needs.

Keywords: social network graphs, clustering, e-Learning analitics, weighted graphs, spectral clustering.

Acknowledgements

This work has been partially supported by Centro de Investigação em Matemática e Aplicações (CIMA) through the grant UID-MAT-04674-2013, by Laboratório de Informática, Sistemas e Paralelismo (LISP) through the grant UID-CEC-4668-2016, both research centers are supported by FCT (Fundação para a Ciência e a Tecnologia, Portugal) and, also, by Departamento de Matemáticas, y Escuela Politécnica de Cáceres, de la Universidad de Extremadura, Spain.

References

[1] Newman, M., Barabasi, A.L. and Watts, D.J., (2011) The structure and dynamics of networks, Princeton University Press.

[2] Kannan, R., Vempala, S. and Vetta, A., (2004) On Clusterings: Good, Bad and Spectral, *Journal of the ACM*, v.51, pp. 497–515.

A critical analysis of the variables that affect the reverse logistics

Maria Varadinov[1,2], Cristina Dias[3,4] and Carla Santos[5,4]

[1]Instituto Politécnico de Portalegre, Departamento de Ciências Económicas e das Organizações, Portugal
[2]Coordenação Interdisciplinar para a Investigação e a Inovação (C3i), Portugal
[3]Instituto Politécnico de Portalegre, Departamento de Tecnologias, Portugal
[4]Centro de Matemática e Aplicações da Universidade Nova de Lisboa (CMA), Portugal
[5]Instituto Politécnico de Beja Departamento de Matemática e Ciências Físicas, Portugal

E-mail: mjvaradinov@iportalegre.pt

Abstract

Knowing the importance that global organizations attach to environmental protection and food quality, the wine and olive industries need to follow specific procedures at strategic and operational level. The return of the bottled wine, having reached the expiration date or change the quality, which influences the quality perceived by retail customers, especially the HORECA channel distribution (consisting of hotels, restaurants and cafes), requires adoption of a RLS -reverse logistics systems. A reverse logistics system defines a supply chain that is redesigned to efficiently manage the flow of products and parts intended for reprocessing, recycling or disposal. In that way, the wine and olive oil producers have an interest in finding a centralized solution that adds value to these products through the implementation of a reverse logistics system. This study intended to analyze the logistics activities of companies that implement or not a reverse logistics system and understand the economic, social and legislative factors that significantly determine that adoption in the companies with production facilities for wine and olive oil in a specific region. A critical analysis of the variables that affect the reverse logistics as well as their interactions can be quite valuable as an important source of information for decision makers.

Keywords: reverse logistics, wine industry.

References

[1] Beamon, B.M., (1999) Designing the Green Supply Chain, *Logistics Information Management*, 12(4), pp. 332–342.

[2] Bollen, K., (1989) Structural Equations with Latent Variables, John Wiley and Sons, Inc., New York.

[3] González-Benito, J. et al., (2006) The role of stakeholder pressure and managerial values in the implementation of environmental logistics practices, *International Journal of Production Research*, 44(7), pp. 1353–1373.

[4] Murphy, P. and Poist, R., (2003) Green perspectives and practices: a comparative logistics study, *Supply Chain Management: An International Journal*, 8(2), pp. 122–131.

[5] Wu, H. and Dunn, S., (1994) Environmentally responsible logistics systems, *International Journal of Physical Distribution & Logistics Management*, 25(2), pp. 20–38.

Fourth central moment

Patrícia Antunes[1], Sandra S. Ferreira[1,2], Célia Nunes[1,2] and Dário Ferreira[1,2]

[1]Center of Mathematics and Applications, Universidade da Beira Interior, Portugal
[2]Department of Mathematics, Universidade da Beira Interior, Portugal

E-mail: antunes.patricia03@gmail.com

Abstract

We present some mathematical formulas dealing with Central Moments and Cumulants. These relationships are useful to determine some central measures. While formulas involving the first, second and third central moment are easy to follow, mathematical complications arise with the fourth central moment, the aim of our study.

Keywords: central moments, cumulant generating function, descriptive statistics, moments.

References

[1] Darlington, R. B., (1970) Is Kurtosis Really Peakedness? *The American Statistician*, 24(2), pp. 19–22.
[2] Halmos, P. R., (1946) The Theory of Unbiased Estimation, *The Annals of Mathematical Statistics*, 17, pp. 34–43.
[3] Moors, J. J. A., (1986) The Meaning of Kurtosis: Darlington Reexamined, *The American Statistician*, 40(4), pp. 283–284.
[4] Kendall, M. G. and Stuart, A., (1963) *The Advanced Theory of Statistics*, vol. 1, 2nd ed, London: Charles Griffin.
[5] Rousson, V., (1995) Axiomatique et Statistique Descriptive, *Student*, 1, pp. 17–34.
[6] Ruppert, D., (1987) What is Kurtosis? An Influence Function Approach, *The American Statistician*, 41, pp. 1–5.
[7] Dodge Y. & Valentin R., (1999) The Complications of the Fourth Central Moment, *The American Statistician*, 53(3), pp. 267–269.

Comparison of time series forecasting methods: an application for retail sales

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

[1]DMA-Departamento de Matemática e Aplicações, Universidade do Minho, Portugal
[2]CMAT – Centro de Matemática, DMA-Departamento de Matemática e Aplicações, Universidade do Minho, Portugal
[3]CIDMA - Centro de Investigação e Desenvolvimento em Matemática e Aplicações, Escola Superior de Tecnologia e Gestão de Águeda, Universidade de Aveiro, Portugal

E-mail: mneves@math.uminho.pt

Abstract

Time series forecasting is an important area of forecasting in which past observations of the same variable are collected and analyzed to develop a model describing the underlying relationship. The model is then used to extrapolate the time series into the future. Forecasting methods are a key tool in decision-making processes in many areas, such as economics, management, finance or environment and over the past several decades much effort has been devoted to the development and improvement of time series forecasting models. There are several forecasting methods supported by different statistical methodologies: Box-Jenkins models and their extensions, the classical decomposition time series associated to multiple linear regression models, artificial Neural networks models, Holt-Winters models, among others [1]. In this work, we compare 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. According to the GfK study European Retail 2016 that analyzes 33 countries in Europe, the retail segment in private consumption in Portugal in 2015 accounted for 33.4 per cent. The study of the economic variables associated with this area is essential and quite useful, both to characterize the recent past and to anticipate trends. A set of retail time series in the Eurostat is modelled considering different approaches. In this sense, the main propose of this work is to compare the accuracy of various models for forecasting retail sales and to bring new insights about the methods used throughout this approach [2].

Keywords: retail sales, time series modeling, forecast accuracy.

Acknowledgements

A. Manuela Gonçalves was supported by the Research Centre of Mathematics of the University of Minho with the Portuguese Funds from the FCT - Fundação para a Ciência e a Tecnologia, through the Project PEstOE-MAT-UI0013-2017. Marco Costa was partially supported by Portuguese funds through the CIDMA - Center for Research and Development in Mathematics and Applications, and the Portuguese Foundation for Science and Technology (FCT– Fundação para a Ciência e a Tecnologia), within project UID-MAT-04106-2013.

References

[1] Gooijer J.G.D., Hyndman R.J. (2006) 25 years of time series forecasting, *International Journal of Forecasting* 22, pp. 443–473.

[2] Ramos P., Santos N., Rebelo R. (2015) Performance of state space and ARIMA for consumer retail sales forecasting, *Robotics and Computer-Integrated Manufacturing* 34, pp. 151–163.

Index of authors

Index of authors

CÂMARA MUNICIPAL
DE FELGUEIRAS

Terras de Felgueiras
Caves de Felgueiras, C.R.L

MUNICÍPIO
Felgueiras
+turismo

ISBN 978-989-98447-6-6