UNIVERSIDADE DA BEIRA INTERIOR, COVILHÃ - PORTUGAL

July, 3 - 4

2012

# BOOK OF ABSTRACTS

July, 3-4, 2012
Universidade da Beira Interior
Covilhã - PORTUGAL

**Chair**
Teresa Oliveira (DCeT, Universidade Aberta)
**Co-Chair**
Stanislaw Mejza (Poznan University of Life Sciences)
**Co-Chair**
Dário Ferreira (Universidade da Beira Interior)


**Organizing Committee**
Adérito Marcos (DCeT, Universidade Aberta)
Amílcar Oliveira (DCeT, Universidade Aberta
Célia Nunes (DM, Universidade da Beira Interior)
Dário Ferreira (DM, Universidade da Beira Interior)
Luís Grilo (UDMF, Instituto Politécnico de Tomar)
Rafael Sasportes (DCeT, Universidade Aberta)
Rosário Ramos (DCeT, Universidade Aberta)
Sandra Ferreira (DM, Universidade da Beira Interior)
Teresa Oliveira (DCeT, Universidade Aberta)

# Preface

Dear Participants, colleagues and friends

We are very pleased to welcome you to the $6^{th}$ Workshop on Statistics, Mathematics and Computation and $3^{rd}$ Portuguese-Polish Workshop on Biometry held in University of Beira interior.

This meeting, organized jointly by Aberta University and University of Beira Interior, provides an international forum for presentation and discussion of new results in the areas of statistics, mathematics and computation. This workshop offers a unique opportunity for extensive discussions and exchange of technical ideas in mathematical sciences, as well as explores the potential of their applications in areas such as Biostatistics, Biometry, Biomedicine and Biomathematics, considering computational issues.

The format of the Workshop involves plenary talks, parallel sessions with organized and contributed talks and posters. Special issues of the journals "Biometrical Letters" and "Discussiones Mathematicae" will be published, devoted to selected papers presented at the Workshop.

This workshop is dedicated to Professor Dinis Pestana who devoted his life to research, teaching and scientific collaboration. He promoted 18 Phds. His scientific contribution spanned the areas of Probability, Stable models, randomly stopped sums and extremes, Location/scale in non-gaussian populations, Functional Methods in Probability and Risk process. We know that nowadays such professors are very hard to come by, so we are very grateful for having had the opportunity and pleasure of knowing him.

We would like to take this opportunity to thank to all our speakers, session organizers and participants for coming to Covilhã and to all authors who submitted their work. Their time and valuable input are greatly appreciated.

Furthermore we would also like to acknowledge our Honour Committee, Rector of Aberta University, Rector of University of Beira interior, President of Instituto Politécnico de Tomar, President of Faculty of Sciences of the University of Beira Interior, President of Mathematics Department of University of Beira Interior, Coordinator of the Center of Statistics and Applications of University of Lisbon, Coordinator of the Center of Mathematics of the University of Beira Interior, President of Portuguese Statistical Society and all our sponsors, whose support has made this event possible.

We are most grateful to all the members of Organizing Committee and of Scientific Committee for making this Workshop a reality.

We hope you enjoy your visit to Covilhã,
Cordially yours

| Teresa Oliveira | Stanislaw Mejza | Dário Ferreira |
|:---:|:---:|:---:|
| Chair | Co-Chair | Co-Chair |

## WORKSHOP CHAIR
Teresa Oliveira - DCeT, Universidade Aberta

## SCIENTIFIC COMMITTEE
Adérito Marcos (DCeT, Universidade Aberta)
Amílcar Oliveira (DCeT, Universidade Aberta)
Ana Maria Portela (Faculdade de Psicologia, Universidade de Lisboa)
Angel A. Juan (Universitat Oberta de Catalunya)
Arminda Manuela Gonçalves (Universidade do Minho)
Áurea Sousa (Universidade dos Açores)
Célia Nunes (Departamento de Matemática, Universidade da Beira Interior)
Dário Ferreira (Departamento de Matemática, Universidade da Beira Interior)
Fátima Brilhante (Universidade dos Açores)
Fernando Branco (Universidade Lusófona)
Fernando Costa (DCeT, Universidade Aberta)
Giovani Silva (Instituto Superior Técnico, Universidade Técnica de Lisboa)
Ivette Gomes (Faculdade de Ciências da Universidade de Lisboa)
Iwona Mejza (Faculty Agronomy Bioeng., Poznan University of Life Sciences, Poland)
Jaime Remédios (DCeT, Universidade Aberta)
Javier Martín Vallejo (Facultad de Medicina, Universidad de Salamanca)
João Araújo (DCeT, Universidade Aberta)
Luís Grilo (Instituto Politécnico de Tomar)
Manuela Neves (ISA, Universidade Técnica de Lisboa)
Manuela Oliveira (Departamento de Matemática, Universidade de Évora)
Maria Antónia Huertas (Universitat Oberta de Catalunya, Barcelona, Spain)
Maria Antónia Turkman (DEIO, Faculdade de Ciências, Universidade de Lisboa)
Maria do Rosário Almeida (DCSH, Universidade Aberta)
Maria José Fernández Gómez (Facultad de Medicina, Universidad de Salamanca)
Marina Andrade (ISCTE, Instituto Universitário de Lisboa)
Miguel Pardal (Department of Life Sciences, University of Coimbra)
Purificación Galindo (Facultad de Medicina, Universidad de Salamanca)
Rafael Sasportes (DCeT, Universidade Aberta)
Roman Zmyślony (University of Zielona Góra, Poland
Rosário Ramos (DCeT, Universidade Aberta)
Sandra Ferreira (Departamento de Matemática da Universidade da Beira Interior)
Sónia Seixas (DCeT, Universidade Aberta)
Stergios Tzortzios (Faculty of Agricultural Sciences, University of Thessaly, Greece)
Teresa Oliveira (DCeT, Universidade Aberta)
Ulisses Azeiteiro (DCeT, Universidade Aberta)
Valter Vairinhos (CINAV - Escola Naval and CENTEC - IST)
Vítor Rocio (DCeT, Universidade Aberta)

## SCHEDULE OF EVENTS

---

### Tuesday (morning), July, 3, 2012

---

08:30 - 09:00 **Reception and Registration**

09:00 - 09:30 **Open Ceremony - Room 6.01**

**Invited Sessions 1,2,3 - Room 6.01**

**Chairman: Manuela Neves**

9:30 - 10:00 **Dinis Pestana** - *Population growth models tied to extensions of beta densities*

10:00 - 10:30 **Damodar Shanbhag** - *Identifiability Problems Associated with Damage Models*

10:30 - 11:00 **M. Ivette Gomes** - *MOP Reduced-Bias EVI-Estimation*

11:00 - 11:30 **Coffee Break and Poster Session 1**

**Invited Sessions 4,5,6 - Room 6.01**

**Chairman: Stanislaw Mejza**

11:30 - 12:00 **Christos Kitsos** - *On the properties of $\gamma$ ordered Generalized Normal Distribution*

12:00 - 12:30 **Carlos Agra Coelho** - *Instances of the Meijer G and Fox H functions and of the distribution of the product of independent beta random variables with finite representations*

12:30 - 13:00 **João Tiago Mexia** - *Linear and quadratic sufficiency in OBS*

13:00 - 14:30 **Lunch Break**

## Tuesday (afternoon), July, 3, 2012

### Invited Sessions 7,8,9 - Room 6.01

**Chairman: João Tiago Mexia**

14:30 - 15:00 **Stanislaw Mejza** - *Check Plots in Field Breeding Experiments*

15:00 - 15:30 **Iwona Mejza** - *Control Treatments in Split-Split-Plot Experiment Designs*

15:30 - 16:00 **Rosemary Bailey** - *Neighbour balance in a strip-block design for an experiment on irrigation*

### PARALLEL SESSIONS

### Organized Session 1 - Room 6.01

**Chairman: Rui Almeida**

16:00 - 16:20 **José C.M. Duque** - *Convergence of the Finite Element Method for the Porous Media Equation in 2D with Variable Exponent*

16:20 - 16:40 **Paulo Rebelo** - *An approximate solutions to a class of nonlinear partial differential equations*

16:40 - 17:00 **Rui J. Robalo** - *A Matlab implementation of the moving finite element method for the solution of time- dependent models*

### Organized Session 2 - Room 6.02

**Chairman: Amílcar Oliveira**

16:00 - 16:20 **J. António Macías** - *An Approach to Distribution of the Product of Two Normal Variables*

16:20 - 16:40 **Maria da Conceição Leal** - *Bootstrap: data speaks for itself*

16:40 - 17:00 **Domingos J. Lopes da Silva** - *Tables and percentiles graphics of some physical fitness indicators of secondary education students*

### Organized Session 3 - Room 6.03

**Chairman: Susana Mendes**

16:00 - 16:20 **Maria José Fernández Gómez** - *Canonical Correspondence Analysis vs Co-Inertia analysis to find patterns of response in spatio-temporal ecological relationships: an empirical comparison.*

16:20 - 16:40 **Sónia Cotrim Marques** - *Using multitable techniques for assessing the temporal variability of species-environment relationship in a copepod community from a temperate estuarine ecosystem*

16:40 - 17:00 **Valter Vairinhos** *Text Mining Using Cartesian and Cylindrical Biplots*

17:00 - 17:20 **Susana Mendes** *A Comparison of HJ-Biplot and Tucker3 Model: an Application in Environmental Data*

## Organized Session 4 - Room 6.04

**Chairman: Maria do Rosário Ramos**

16:00 - 16:20 **Bruno Fonseca** - *Statistical methods in the evaluation of inflammatory parameters in dialysis patients*

16:20 - 16:40 **Domingas Portela** - *Structural equation modeling and application to the study of scientific culture in high school students.*

16:40 - 17:00 **Mónica Ventura** - *Application of Path Analysis to Patient's Satisfaction with his Doctor*

17:00 - 17:30 **Coffee Break and Poster Session 1**

## Invited Sessions 10,11,12,13 - Room 6.01

**Chairman: M. Ivette Gomes**

17:30 - 18:00 **Manuela Neves** - *How resampling methodologies can help in extreme value estimation*

18:00 - 18:30 **Miguel Fonseca** - *Inference in Linear Models with Doubly Exchangeable Distributed*

18:30 - 19:00 **Ana Luísa Papoila** - *Age-Period-Cohort models to analyze temporal variation in cancer incidence*

19:00 - 19:30 **Mário Edmundo** - *Invariância da cohomologia o-minimal*

20:30 **Workshop Dinner**

## Wednesday(morning), July, 4, 2012

### Invited Sessions 14,15,16,17 - Room 6.01

**Chairman: Pedro Serranho**

09:00 - 09:30 **Angel A. Juan** - *An efficient, simple, and parameter-free algorithm for solving the permutation flow-shop problem*

09:30 - 10:00 **Vitor Cardoso** - *Using Real and Virtual Objects: educational feedbacks*

10:00 - 10:30 **João Cabral** - *Building bridges between Cluster Analysis and Dynamical Systems*

10:30 - 11:00 **José Gomes Requeijo** - *Monitorização de Processos com dados Auto-correlacionados*

11:00 - 11:30 **Coffee Break and Poster Session 2**

### Invited Sessions 18,19,20 - Room 6.01

**Chairman: Amílcar Oliveira**

11:30 - 12:00 **Dulce Pereira** - *A comparison between Joint Regression Analysis and AMMI model: a case study with barley*

12:00 - 12:30 **Paulo Canas Rodrigues** - *Statistical Methods to Detect and Understand Genotype-by-Environment Interactions and QTL-by-Environment Interactions*

12:30 - 13:00 **Pedro Serranho** - *3D Retinal Vascular Network from OCT data*

13:00 - 14:30 **Lunch Break**

**PARALLEL SESSIONS**

## Organized Session 5 - Room 6.01

**Chairman: Mário Bessa**

14:30 - 14:50 **Jorge Freitas** - *Extreme values from a dynamical perspective*

14:50 - 15:10 **João Lopes Dias**- *Hyperbolicity in dissipative polygonal billiards*

15:10 - 15:30 **Maria Joana Torres**- *Stability and Morse decompositions of non-deterministic systems*

15:30 - 15:50 **Glória Ferreira**- *Asymptotic behavior of infinite dimensional compact skew-product flows*

15:50 - 16:10 **Salete Esteves**- *Heterodimensional cycles: a model family*

16:10 - 16:30 **José J. Oliveira**- *Exponential stability for impulsive functional differential equations with infinite delay*

## Organized Session 6 - Room 6.02

**Chairman: Alberto M. Simões**

14:30 - 14:50 **Luís Castro** - *Operator viewpoint of general boundary value problems within wave diffraction*

14:50 - 15:10 **Anabela Silva** - *Fredholm characteristics of singular integral operators with a Carleman shift and almost periodic coefficients via factorization of matrix functions*

15:10 - 15:30 **Matteo Dalla Riva** - *Effective conductivity of a singularly perturbed periodic two-phase composite with imperfect thermal contact at the two-phase interface*

15:30 - 15:50 **M. Manuela Rodrigues** - *Heat kernel for the index Whittaker transform*

15:50 - 16:10 **Alberto M. Simões** - *Integral Equation Methods in Problems of Reactance Wave Diffraction by a Strip*

## Organized Session 7 - Room 6.03

**Chairman: Isabel Mendes**

14:30 - 14:50 **Ana Carapito** - *Existence of a contractive set of quadratic Lyapunov functions for a switched system*

14:50 - 15:10 **Henrique Cruz** *On (0, 1)-matrices with given line sums*

15:10 - 15:30 **Ilda Inácio Rodrigues** *Some theory about hypermaps*

15:30 - 15:50 **Celino Miguel** *Zero Division Graph of Quaternion Rings over $\mathbb{Z}_P$*

15:50 - 16:10 **Isabel Mendes** - *On completely \*-prime \*-ideals in involution rings*

## Contributed Session - Room 6.04

**Chairman: Miguel Fonseca**

14:30 - 14:50 **Adelaide Figueiredo** - *Selection of variables defined on the hypersphere*

14:50 - 15:10 **Fernanda Figueiredo** - *Bootstrap control charts to monitor skew-normal processes*

15:10 - 15:30 **Frederico Caeiro** *Further results on the extreme value index estimation: the maximum likelihood estimators of Feuerverger and Hall*

15:30 - 15:50 **Sandra Aleixo** *Strong and weak Allee effects, extinctions and bistability in Richards growth dynamics*

15:50 - 16:10 **Vadim Yurinsky** *Some Inequalities for Functions with Variable Exponents*

16:10 - 16:30 **Patrícia Gonçalves** *The Weakly Asymetric Exclusion Process with Open Boundaries*

16:30 - 17:00 **Coffee Break and Poster Session 2**

## Organized Session 8 - Room 6.01

**Chairman: Miguel Felgueiras**

17:00 - 17:20 **Dinis Pestana** - *Cost considerations in statistical work*

17:20 - 17:40 **João Paulo Martins** - *Meta-analysis techniques applied in prevalence rate estimation*

17:40 - 18:00 **Rui Santos** - *The number of bacteria in discrete quantitative compound analysis*

18:00 - 18:20 **Ricardo Sousa** *The prevalence rate estimation in compound tests with misclassification*

## Organized Session 9 - Room 6.02

**Chairman: Elisabete Carolino**

17:00 - 17:20 **Arlete Queirós** - *Ativação muscular na anca e joelho na variação do ângulo de Valgo durante a fase de apoio do salto vertical*

17:20 - 17:40 **Elisabete Carolino** - *Quantificação da perfusão da parede miocárdica em estudos Single Photon Emission Computed Tomography: Influência do número de contagens*

17:40 - 18:00 **Olga Valentim** - *Aplicabilidade da Escala de Satisfação com o Suporte Social (ESSS) a Pessoas com Síndrome de Dependência Alcoólica*

18:00 - 18:20 **Tiago Atalaia** *Análise da Influência da Aplicação de Kinesio Tape na Ativação Muscular Durante um Passe de Futsal*

**Organized Session 10 - Room 6.03**

**Chairman: Teresa Oliveira**

17:00 - 17:20 **Domingos Silva** - *Aptidão cardiorespiratória e índice de massa corporal em alunos do ensino secundário*

17:20 - 17:40 **Helga Correia** - *Planos em blocos incompletos equilibrados com repetições: uma aplicação na área da educação*

17:40 - 18:00 **Vitor Valente** - *Some considerations about models with application to data with hierarchical or multilevel structure*

18:00 - 18:20 **Ricardo Mendes** *Quadrados Latinos e Planos Projetivos: Revisão de problemas clássicos*

**Invited Sessions 21,22 - Room 6.01**

**Chairman: João Araújo**

18:20 - 18:50 **Michael K. Kinyon** - *Adventures in Automated Deduction and Algebra*

18:50 - 19:20 **Peter Cameron** - *Balanced and variance-balanced block designs*

19:20 **Closing Ceremony**

**Poster session 1**

- *Anabela Marques* - Variables' selection in Discrete Discriminant Analysis

- *Áurea Sousa* - Clustering of Symbolic Data based on Affinity Coefficient: Application to real data sets

- *Carla Noronha* - Crossover Designs: Revisão e Desafios

- *Conceição Leal* - Process Capability Analysis: an application to the orders management

- *Cristina Dias* - Models of Symmetric Stochastic Matrices

- *Dina Salvador* - Álgebras de Jordan Comutativas e Modelos Ortogonais

- *Fernando Afonso* - Modelling of symptoms in patients with pollinosis in the Alto Alentejo region

- *Osvaldo Silva* - Global Approach for the comparison of Clustering Results

**Poster session 2**

- *A. Manuela Gonçalves* - Online Water Quality Monitoring by using Kalman Filter Predictions

- *Andreia A. Monteiro* - Aplicações de Técnicas Estatísticas e Probabilísticas no Desenvolvimento e Análise de Algoritmos de Optimização Combinatória na Área da Bioinformática

- *Carla Noronha* - Blood Groups Gene Frequencies

- *Cláudia Roçadas* - Isomorphism of Markov Chains: Application to a Customer Population

- *Domingos J. Lopes da Silva* - Análise estatística da evolução da performance motora na flexibilidade ao longo do ensino secundário - Comparação Rapazes vs Raparigas

- *Maria Cristina Ferreira* - Modelo de regressão logística multivariada - Uma ferramenta na área da medicina

- *Marina Andrade* - Estimating Biological Age Through Dental and Skeletal Indicators

- *Sandra Nunes* - An Introduction to Item Response Theory

# Contents

ABSTRACTS — INVITED SPEAKERS

ABSTRACTS — ORGANIZED SESSIONS

ABSTRACTS — CONTRIBUTED PAPERS - ORAL

ABSTRACTS — CONTRIBUTED PAPERS — POSTERS

# ABSTRACTS

# INVITED SPEAKERS

# Age-Period-Cohort models to analyze temporal variation in cancer incidence

**Ana Luísa Papoila**
*Faculdade de Ciências Médicas da UNL, CEAUL, Portugal*

**Abstract:**Presentation and interpretation of temporal variation in disease incidence or mortality is undoubtedly a major concern of epidemiologists. This kind of variation may be due to secular (period) influences and/or to generational (cohort) influences and, so, age at occurrence of disease or death, time of occurrence of disease or death and year of birth are three important factors in the analysis of secular trends. Age-Period-Cohort models (APC models) are regression models that describe disease rates as a product of an age-effect, a period effect and a cohort effect. Several approaches have been proposed in the literature to model data obtained from cancer registries and Carstensen (2007) suggested a Generalized Additive Model with a Poisson response probability distribution where the desired effects are estimated through natural cubic splines. To illustrate these models, data on colon and rectum cancer obtained during the period 1998-2005, provided by Registo Oncológico Regional Sul, was analyzed.

**Keywords:** Age-Period-Cohort model; Poisson model; epidemiology; cancer.

**References**

Carstensen, B. (2007). *Age-period-cohort models for the Lexis diagram.* Statistics in Medicine, 26:3018-3045. DOI:10.1002/sim.2764.

# An efficient, simple, and parameter-free algorithm for solving the permutation flow-shop problem

**Juan, A.; Lourenço, H.; Mateo, M.; Castella, Q.; Barrios, B.**

**Abstract:** From a managerial point of view, the more efficient, simple, and parameter-free (ESP) an algorithm is, the more likely it will be used in practice for solving real-life problems. Following this principle, an ESP algorithm for solving the Permutation Flowshop Sequencing Problem (PFSP) is proposed in this article. Using an Iterated Local Search (ILS) framework, the so-called ILS-ESP algorithm is able to compete in performance with other well-known ILS-based approaches, which are considered among the most efficient algorithms for the PFSP. However, while other similar approaches still employ several parameters that can affect their performance if not properly chosen, our algorithm does not require any particular fine-tuning process since it uses basic 'common sense' rules for the local search, perturbation, and acceptance criterion stages of the ILS metaheuristic. Our approach defines a new operator for the ILS perturbation process, a new acceptance criterion based on extremely simple and transparent rules, and a biased randomization process of the initial solution to randomly generate different alternative initial solutions of similar quality -which is attained by applying a biased randomization to a classical PFSP heuristic. This diversification of the initial solution aims at avoiding poorly designed starting points and, thus, allows the methodology to take advantage of current trends in parallel and distributed computing. A set of extensive tests, based on literature benchmarks, has been carried out in order to validate our algorithm and compare it against other approaches. These tests show that our parameter-free algorithm is able to compete with state-of-the-art metaheuristics for the PFSP. Also, the experiments show that, when using parallel computing, it is possible to improve the top ILS-based metaheuristic by just incorporating to it our biased randomization process with a high-quality pseudo-random number generator.

**Keywords:** combinatorial optimization, flow-shop problem, biased randomization, meta-heuristics.

# Instances of the Meijer G and Fox H functions and of the distribution of the product of independent beta random variables with finite representations

**Carlos A. Coelho[1], Barry C. Arnold[2]**

[1]*Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Portugal, Centro de Matemática e Aplicações - Universidade Nova de Lisboa, Portugal - e-mail: cmac@fct.unl.pt*

[2]*Statistics Department, University of California at Riverside, U.S.A., e-mail: barry.arnold@ucr.edu*

**Abstract:**Through the use of some extended product expressions for the Gamma function the authors show how, by working on the characteristic functions of the negative logarithm of the product of particular independent Beta random variables, whose second parameters are rational and follow some amenable rules, it is possible to obtain in a single shot (i) simple expressions for both the probability density and cumulative distribution functions of this product, as well as, concomitantly (ii) easy to compute alternative finite form expressions for instances of the Meijer G and Fox H functions. These alternative expressions are based on the expressions for the probability density and cumulative distribution functions of the Generalized Integer Gamma distribution and are not used or recognized by the available softwares.

**Keywords:** Characteristic function, distribution of likelihood ratio statistics, Fox H function, Generalized Integer Gamma (GIG) distribution, sum of independent Gamma random variables.

# On the properties of $\gamma$ ordered Generalized Normal Distribution

Christos P. Kitsos

*Department of Mathematics, Technological Educational Institute of Athens, xkitsos@teiath.gr*

**Abstract:** The target of his paper is to discuss the $\gamma$ ordered Generalized Normal Distribution. This distribution generalizes the (multivariate) Normal distribution, introducing an extra shape parameter. This distribution emerged from the study of the Logarithm Sobolev Inequalities as an external for the generalized entropy type measure of information. For particular values of the $\gamma$ parameter, the order, of the distribution a number of well known distributions is produced as Uniform, Normal, Laplace, Dirac. The moments of this distribution are evaluated, as well as a number of properties. The generalized entropy type Fishers information is also evaluated for it, under the light of the generalization of Fishers entropy type information

**References**

1. Kitsos, C. P. and Toulias (2012) T. L Inequalities for the generalized entropy type Fisher's information. Int. Journal of Nonlinear Analysis & Applications.

2. Kitsos, C. P. (2011). On Fishers Information Measures. Digital Technologies 2011, Nov 10-11, 2011, Zilina, Slovakia, e-Prngs 178-183.

3. Kitsos, C. P , Toulias, T. L (2011). On the family of the $\gamma$-ordered normal distributions. Far East Journal of Theoretical Statistics, Vol 35(2), pp 95-14.

4. Kitsos, C. P., Tavoularis, K. N. (2009). Logarithmic Sobolev Inequalities for Information Measures. IEEE TRANSACTIONS ON INFORMATION THEORY, Vol 55, 6, June 2009, 2554-2561.

# IDENTIFIABILITY PROBLEMS ASSOCIATED WITH DAMAGE MODELS

**Damodar Shanbhag**

**Abstract:** The concept of damage models was introduced by C.R.Rao in 1963 in a conference paper; the paper mentioned here was reprinted later in 1965 in Sankhya, ser. A, 27, 311-324. The model considered by Rao applies to certain populations of items or individuals, which are affected by accidental damages or natural disasters. In mathematical terms, a damage model due to Rao can be described by a 2-component random vector (X, Y) where X and Y denote nonnegative integer-valued random variables with Y not larger than X. In this model, usually X refers to the original population size and Y and X - Y refer respectively, to its undamaged and damaged parts. Also, in this model, the conditional distribution of Y given X is frequently called the survival distribution. In his paper, Rao discussed various examples of damage models. An interesting property met by some of the damage models is that the marginal distribution of Y is the same as the conditional distribution of Y given that X is undamaged (i.e. X - Y = 0); this has been used as one of the criteria, in the literature, for identifying the distribution of X, given the survival distribution, and also for identifying the survival distribution, given the distribution of X. Multivariate versions and variations of such models have also been considered in the literature. In this talk, we review some important results and problems on damage models and reveal as to how these are related to certain problems on integral equations and moment sequences, met in applied probability or statistics.

# Population growth models tied to extensions of beta densities

**Maria de Fátima Brilhante[1], Maria Ivette Gomes[2], Dinis Pestana[3]**

[1]*Universidade dos Açores, Departamento de Matemática, and CEAUL Centro de Estatística e Aplicações da Universidade de Lisboa Portugal, fbrilhante@uac.pt*
[2]*CEAUL Centro de Estatística e Aplicações da Universidade de Lisboa, Portugal, ivette.gomes@fc.ul.pt*
[3]*CEAUL Centro de Estatística e Aplicações da Universidade de Lisboa, Portugal, dinis-pestana@fc.ul.ptl*

## Abstract

In the parabola $x(1 - x)$ the strict equilibrium of the growing factor $x$ and the retroaction factor $1 - x$ leads to the Verhults sustainable logistic growth model, an asymptotic stable model for maxima of geometriclly thinned sequences of iid random variables.

Considering $x$ and/or $1-x$ linear truncation of more complex functions, and powers of the above factors, we obtain other growth models, namely models appropriate to deal either with explosive growth or with very slow development, corresponding to dominant growth or retroaction factor.

**Keywords:** Verhulst logistic model, Beta and BeTaBoOp models, population dynamics, extreme value models, geometric thinning.

## References

Aleixo, S., Rocha, J.L., and Pestana, D. (2011). Probabilistic Methods in Dynamical Analysis: Population Growths Associated to Models Beta (p,q) with Allee Effect, in Peixoto, M. M; Pinto, A.A.; Rand, D. A. J., editors, Dynamics, Games and Science, in Honour of Maur´icio Peixoto and David Rand, vol II, Ch. 5, pages 7995, New York, Springer Verlag.

Brilhante, M. F., Gomes, M. I., and Pestana, D. (2011). BetaBoop Brings in Chaos. Chaotic Modeling and Simulation, 1: 3950.

Pestana, D., Aleixo, S., and Rocha, J. L. (2011). Regular variation, paretian distributions, and the interplay of light and heavy tails in the fractality of asymptotic models. In C. H. Skiadas, I. Dimotikalis and C. Skiadas, editors, Chaos Theory: Modeling, Simulation and Applications, pages 309316, Singapore, World Scientific.

# A comparison between Joint Regression Analysis and AMMI model: a case study with barley

**Dulce G. Pereira[1], Paulo C. Rodrigues[2], Stanislaw Mejza[3] and João T. Mexia[2]**

[1]*CIMA-UE and Department of Mathematics, Colégio Luís António Verney, University of Évora, 7000-671 Évora, Portugal;*
[2]*CMA and Department of Mathematics and CMA, Faculty of Sciences and Technology, Nova University of Lisbon, 2829-516 Caparica, Portugal;*
[3]*Department of Mathematical and Statistical Methods, Poznan University of Life Sciences, PL-60-637 Poznan, Poland*

**Abstract:** Joint regression analysis (JRA) and additive main effects and multiplicative interaction (AMMI) models are compared in order to (i) access the capacity of describing a genotype by environment interaction effects and (ii) evaluate the agreement between the winners of mega-environments obtained from the AMMI analysis and the genotypes in the upper contour of the JRA. An iterative algorithm is used to obtain the environmental indexes for JRA, and standard multiple comparison procedures are adapted for genotype comparison and selection. The study includes three data sets from spring barley (Hordeum vulgare L.) breeding program carried out between 2004 and 2006 in Czech Republic. The results from both techniques are integrated in order to advice plant breeders, farmers and agronomists for better genotype selection and prediction for new years and/or new environments.

### References

M.S. Kang and H.G. Gauch, Genotype -by- Environment Interaction, CRC Press, Boca Raton, 1996.

K.W. Finlay and G.N. Wilkinson, Analysis of Adaptation in a Plant-Breeding Programme, Aust. J. Agric. Res. 14 (1963), pp. 742-754.

H.G. Gauch, Statistical Analysis of Regional Yield Trials: AMMI Analysis of Factorial Designs, Elsevier, Amsterdam, 1992.

L. Gusmão, An Adequate Design for Regression-Analysis of Yield Trials, Theor. Appl. Genet. 71 (1985), pp. 314-319.

J.T. Mexia, D.G. Pereira and J. Baeta, L2 environmental indexes, Biometrical Letters 36 (1999), pp. 137-143.

J.T. Mexia, A.P. Amaro, L. Gusmão and J. Baeta, Upper contour of a Joint Regression Analysis, J. Genet. Breed. 51 (1997), pp. 253-255.

[7] H.G. Gauch, Model Selection and Validation for Yield Trials with Interaction, Biometrics 44 (1988), pp. 705-715.

H.G. Gauch and R.W. Zobel, Identifying mega-environments and targeting genotypes, Crop Sci. 37 (1997), pp. 311-326.

D.G. Pereira and J.T. Mexia, Comparing double minimization and zigzag algorithms in Joint Regression Analysis: the complete case, J. Stat. Comput. Sim. 80 (2010), pp. 133-141. [10] P.C. Rodrigues, D.G. Pereira and J.T. Mexia, A comparison between JRA and AMMI: the robustness with increasing amounts of missing data, Scientia Agricola 68 (2011), 679-686.

D.G. Pereira and J.T. Mexia, Selection proposal of cultivars of spring barley in the years from 2001 to 2004, using Joint Regression Analysis, Plant Breeding 127 (2008), pp. 452-458.

H. Scheffé, The analysis of variance, Wiley, New York,, 1959.

J. Paderewski, H.G. Gauch, W. Madry, T. Drzazga and P.C. Rodrigues, Yield Response of Winter Wheat to Spatial Conditions Using AMMI and Cluster Analysis, Crop Sci. 51 (2011), pp. 969-980.

D. Bradu and K.R. Gabriel, Biplot as a Diagnostic Tool for Models of 2-Way Tables, Technometrics 20 (1978), pp. 47-68.

P. Annicchiarico, F. Bellah and T. Chiari, Defining subregions and estimating benefits for a specific-adaptation strategy by breeding programs: A case study, Crop Sci. 45 (2005), pp. 1741-1749.

H.G. Gauch, MATMODEL version 3.0: Open source software for AMMI and related analyses, 2007, software available at http://www.css.cornell.edu/staff/gauch

H.P. Piepho, Robustness of statistical tests for multiplicative terms in the additive main effects and multiplicative interaction model for cultivar trials, Theor. Appl. Genet. 90 (1995), pp. 438-443.

P. Annicchiarico, Joint regression vs AMMI analysis of genotype-environment interactions for cereals in Italy, Euphytica 94 (1997), pp. 53-62.

[19] M. Brancourt-Hulmel, V. Biarnès-Dumoulin, and J.B. Denis, Points de repère dans l'analyse de la stabilité et de l'interaction génotype-milieu en amélioration des plants, Agronomie 17 (1997), pp. 219-246.

P. Annicchiarico, Genotype x Environment Interactions - Challenges and Opportunities for Plant Breeding and Cultivar Recommendations, FAO Plant Production and Protection Papers, FAO, Rome, 2002. Available at http://www.fao.org/docrep/005/y4391e/y4391e00.htm.

P. Annicchiarico, F. Bellah, and T. Chiari, Repeatable genotype x location interaction and its exploitation by conventional and GIS-based cultivar recommendation for durum wheat in Algeria, Eur. J. Agron. 24 (2006), pp. 70-81.

J.S. Ebdon and H.G. Gauch, Direct validation of AMMI predictions in turfgrass trials, Crop Sci. 51 (2011), pp. 862-869.

P. Annicchiarico, Coping with and exploiting genotype-by-environment interactions, in Plant Breeding and Farmer Participation, S. Ceccarelli, E.P. Guimarães. and E.Weltzien, eds., FAO, Rome, 2009, pp. 519-564.

# MOP Reduced-Bias EVI-Estimation

**M. Fátima Brilhante[1], M. Ivette Gomes[2], Dinis Pestana[3]**

[1] *Universidade dos Açores, DM, and CEAUL, Portugal, e-mail: fbrilhante@uac.pt*
[2] *Universidade de Lisboa, FCUL, DEIO and CEAUL, Portugal, e-mail: ivette.gomes@fc.ul.pt*
[3] *Universidade de Lisboa, FCUL, DEIO and CEAUL, Portugal, e-mail: dinis.pestana@fc.ul.pt*

**Abstract:** Given a random sample, $(X_1, \ldots, X_n)$, and the associated sample of ascending order statistics, $(X_{1:n} \leq \cdots \leq X_{n:n})$, the Hill estimator (Hill, 1975), the average of the log-excesses over a high threshold $X_{n-k:n}$, can be regarded as the logarithm of the geometric mean, or equivalently, the logarithm of the mean of order $p = 0$, of the set of statistics $U_{ik} := X_{n-i+1:n}/X_{n-k:n}$, for $1 \leq i \leq k < n$. Instead of such a geometric mean, it is thus sensible to consider the mean of order $p$ (MOP) of those statistics, $A_p(k) := \left( \sum_{i=1}^{k} U_{ik}^p / k \right)^{1/p}$, with $p \geq 0$, and the associated estimators of the *extreme value index* (EVI), studied in Brilhante *et al.* (2012). We now suggest a class of reduced-bias MOP EVI-estimators, built in the lines of the *minimum-variance reduced-bias* class of EVI-estimators in Caeiro *et al.* (2005), and derive their asymptotic and finite sample properties.

**Keywords:** Statistics of extremes, semi-parametric estimation, bias estimation, heavy right-tails.

**References**
Brilhante, M.F., M.I. Gomes, and D.D. Pestana (2012). *A Simple Generalization of the Hill Estimator.* Notas e Comunicações CEAUL 03/2012.
Caeiro, F., M.I. Gomes, and D.D. Pestana (2005). Direct reduction of bias of the classical Hill estimator. *Revstat* **3**:2, 111–136.
Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Annals Statistics* **3**, 1163-1174.

# CONTROL TREATMENTS IN SPLIT-SPLIT-PLOT EXPERIMENT DESIGNS

**Iwona Mejza**

*Department of Mathematical and Statistical Methods Poznan University of Life Sciences, Poland*

**Abstract:** Some methods of the construction of non-orthogonal split-split-plot designs for three or more factor experiments is provided. Their non-orthogonality with respect to whole plot treatments and sub-subplot treatments are considered. Additionally, some of these treatments as control treatments are treated. To generate a new treatment combination arrangement an orthogonally supplemented PEB block design with at most $(m + 1)$ - classes of efficiency is taken into account. Attention is paid to optimal statistical properties of the resulting designs with respect to the efficiency of estimation of some group of the contrasts.

# Building bridges between Cluster Analysis and Dynamical Systems

**João Cabral**

*University of Azores, Department of Mathematics, CMATI, Portugal, jcabral@uac.pt*

**Abstract:** A problem which arises quite often is that, given a set of n individuals, is to build a group of individual classes or subgroups so that every subgroup is internally homogeneous (ie is composed of individuals "similar") and that the various subgroups are heterogeneous among themselves (ie is, individuals from different subgroups are "dissimilar"). In methods of Discriminant Analysis we start with the assumption that such a subdivision is a known set of data that is available, and the objective is to seek directions in space that show the separation of these subgroups or determine a rule for future classifications. But often there is no such classification available, and the problem is to identify which (and how) are the different classes of individuals existing on the available data.

In this work we propose the introduction of tools derived from Dynamical Systems, from Kneading Theory and Symbolic Dynamics, to build the individual classes or subgroups, and more, helping the analysis of how much homogeneous this classes or subgroups can be.

**Keywords:** Cluster Analysis; Dynamical Systems; Symbolic Dynamics.

## References

Cabral, J. (2009). Estudo analítico e simbólico das bifurcações para a família das aplicações racionais de grau dois. Tese de Doutoramento, Universidade dos Açores, Ponta Delgada.

Carvalho, F.; Souza, R. (2009). Unsupervised pattern recognition models for mixed feature-type symbolic data. Pattern Recognition Letters, Volume 31, nº5, 430-443.

Kitchens, Bruce P. (1997) Symbolic Dynamics, One-sided, two-sided and Countable State Markov Shifts, Springer.

Silva O.; Bacelar-Nicolau, H.; Nicolau, F. (2009). Como Avaliar a Consistência dos Resultados de uma Análise Classificatória Hierárquica. In: Oliveira, I. et al. Eds. Actas do XVI Congresso Anual da Sociedade Portuguesa de Estatística, 2008, Edições S.P.E., 661-672.

Silva, O. (2012). Contributos para a Avaliação e Comparação de Partições em Análise Classificatória. Tese de Doutoramento, Universidade dos Açores, Ponta Delgada.

# Linear and quadratic sufficiency in OBS

**Augustyn Markiewicz[1], Francisco Carvalho[2], João Tiago Mexia[3]**

[1]*PULS, Department of Mathematical and Statistical Method, Poland, amark@au.poznan.pl*
[2]*IPT, ESGT, Portugal, fpcarvalho@ipt.pt*
[3] *Departamento de Matemática, Portugal, jtm@fct.unl.pt*

**Abstract:** Linear and quadratic sufficiency have been used to obtain BLUE and BQUE for models with only one variance component.
Now we extend This use to models with OBS. This is to models with variance-covariance matrix

$$V = \sum_{j=1}^{m} \gamma_j K_j$$

Where the $K_1, \ldots, K_m$, are pairewise orthogonal orthogonal projection matrices that add up to $I_n$.

**Keywords:** Linear sufficiency, quadratic sufficiency.

## References

Seely, J. (1971). Quadratic subspaces and completeness. Ann. Math. Statis. 42, 2, 710-721.

Zmyslony (1980). A characterization of best linear unbiades estimators in general linear model. Mathematical Statistics and Probability Theory, Proc. Sixth Internat. Conf. , Wisla/Pol. 1978, Lect. Notes Stat., 2, Springer, New York-Berlin, 365-373.

# Monitorização de Processos com dados Auto-correlacionados

**Requeijo, José Gomes**

*UNIDEMI, Departamento de Engenharia Mecânica e Industrial, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa Campus de Caparica, 2829-516 - Caparica, Portugal, jfgr@fct.unl.pt*

**Abstract:** O Controlo Estatístico do Processo (SPC) tem merecido uma atenção crescente por parte das organizações devido à eficácia que tem demonstrado na melhoria contínua do desempenho de qualquer processo produtivo. O SPC permite monitorizar o comportamento do processo através de técnicas estatísticas específicas, as cartas de controlo, reduzir a sua variabilidade e determinar se o processo tem capacidade para produzir de acordo com especificações estabelecidas. Um dos pressupostos fundamentais para a implementação das cartas de controlo tradicionais, habitualmente designadas por cartas de Shewhart, consiste em assumir que os dados são independentes e Normalmente distribuídos. No entanto, e particularmente em processos contínuos, é frequente a violação do pressuposto de independência dos dados, i.e., a existência de auto-correlação significativa no processo. A aplicação directa das cartas tradicionais aos dados, ignorando a existência da auto-correlação, pode aumentar drasticamente o número de falsos alarmes, o que afecta consideravelmente o desempenho destas técnicas estatísticas. Na literatura é possível identificar algumas abordagens distintas para a monitorização de processos com dados auto correlacionados; a primeira sugere a implementação de cartas com os limites de controlo modificados em função do tipo de auto-correlação evidenciado; uma segunda preconiza a aplicação de cartas de controlo dos resíduos/erros de previsão obtidos após modelação do processo, conseguida habitualmente através dos modelos ARIMA (Autoregressive Integrated Moving Average); uma terceira abordagem consiste na implementação de cartas de controlo MCEWMA (Moving Center-Line Exponentially Weighted Moving Average), e EWMAST (Exponentially Weighted Moving Average for Stationary Process) directamente aos dados originais do processo. A presente comunicação expõe de forma clara a implementação do SPC quando os dados são auto-correlacionados e compara o desempenho de diversas técnicas englobadas nas abordagens acima citadas, recorrendo a exemplos de aplicação.

**Keywords:** Controlo Estatístico do Processo, Cartas Tradicionais de Shewhart, Autocorrelação, Modelos ARIMA, carta MCEWMA, carta EWMAST

# How resampling methodologies can help in extreme value estimation

**M. Manuela Neves[1], M. Ivette Gomes[2], F. Figueiredo[3], D. Prata Gomes[4]**

[1] *CEAUL and UTL, ISA, Portugal, manela@isa.utl.pt*
[2] *CEAUL and UL, FC/DEIO, Portugal, migomes@fc.ul.pt*
[3] *CEAUL and UP, FE, Portugal, otilia@fep.up.pt*
[4] *CMA and UNL, FCT, Portugal, dsrp@fct.unl.pt*

**Abstract:** In this talk we work in our main field of research, Statistics of Extremes, dealing with the importance of resampling-based computer intensive methodologies, like the Generalized Jackknife and the Bootstrap for obtaining reliable semi-parametric estimates of parameters of extreme or even rare events. Among these parameters we can refer to the extreme value index, $\gamma$, the primary parameter in Statistics of Extremes, measuring the right tail-weight and the extremal index, $\theta$, a parameter of real interest for dependent samples (the common situation in practice), which can roughly be defined as the reciprocal of the expectation of the extremes duration. Most of the semi-parametric estimators of these parameters show the same type of behaviour: nice asymptotic properties, but a high variance for small values of $k$, the number of upper order statistics used in the estimation, a high bias for large values of $k$, and the need for the choice of $k$. After a brief introduction of some estimators of the aforementioned parameters and their asymptotic properties we propose, on the basis of bootstrap and jackknife computer-intensive methods, an algorithm for the choice of $k$ and the adaptive estimation of $\gamma$ and $\theta$. A simulation study as well as some applications to real data will be provided.

**Keywords:** Bootstrap, extreme value theory, jackknife, semi-parametric estimation.

# Invariância da cohomologia o-minimal

**Mário J. Edmundo**

**Abstract:** Nesta comunicação apresentaremos resultados recentes sobre a invariância da cohomologia o-minimal quando se muda de modelo ou se passa a expansões o-minimais. Como se sabe a o-minimalidade é uma generalização (no âmbito da teoria de modelos - lógica matemática) da geometria semi-algébrica e sub-analítica. Assim sendo estes resultados são uma generalização de resultados similares devidos a Delfs no caso particular da geometria semi-algébrica.

# Adventures in Automated Deduction and Algebra

**Michael Kinyon**

*Department of Mathematics, University of Denver*

**Abstract:** Except for singular events such as McCune's solution of the Robbins problem in Boolean algebras, there is a general perception that automated deduction software is not suitable for "real" mathematics and cannot handle anything more than toy problems. In keeping with the computational theme of this Workshop, in this talk I will attempt to convince the audience that automated deduction is a viable and useful tool in the mathematician's toolbox. In particular, I will show how automated deduction software is being used to tackle real unsolved problems in algebra. Examples will come from various areas such as quasigroup/loop theory, semigroup theory and lattice theory.

# Inference in Linear Models with Doubly Exchangeable Distributed Errors

**Miguel Fonseca[1], Anuradha Roy[2]**

[1]*Centro de Matemática e Aplicações, Universidade Nova de Lisboa, E-mail: fmig@fct.unl.pt*
[2]*Department of Management Science and Statistics, The University of Texas at San Antonio, E-mail: aroy@utsa.edu*

**Abstract:** We study the general linear model (GLM) with doubly exchangeable distributed error for $m$ observed random variables. The doubly exchangeable general linear model (DEGLM) arises when the $m-$dimensional error vectors are "doubly exchangeable" (defined later), jointly normally distributed, which is a much weaker assumption than the independent and identically distributed error vectors as in the case of GLM or classical GLM (CGLM). We estimate the parameters in the model and also find their distributions. We show that the testings of intercept and slope are possible in DEGLM as a particular case using parametric bootstrap as well as multivariate Satterthwaite approximation.

**Keywords:** Doubly exchangeable covariance structure; Linear model; Parametric bootstrap; Multivariate Satterthwaite approximation.

## References

Anderson, T. W. (2003). An introduction to multivariate statistical analysis, Third Ed. New Jersey: Wiley.

Arnold, S. F. (1979). Linear Models with Exchangeably Distributed Errors. J. of the American Statistical Association 74:(365) 194-199.

Gupta, A. K., Nagar, D. K. (2000). Matrix Variate Distributions, Chapman & Hall/CRC, Florida.

Khuri, A. I. (2010). Linear Model Methodology, Chapman & Hall/CRC, Florida.

Khuri, A. I., Mathew, T., Sinha, B. K. (1998). Statistical tests for mixed linear models, Wiley Series in Probability and Mathematical Statistics, Applied Section, New York: Wiley.

Kollo, T., von Rosen, D. (2005). Advanced multivariate statistics with matrices, Dordrecht, The Netherlands: Springer.

Harville, D. A. (1997). Matrix Algebra From a Statistician's Perspective, New York: Springer-Verlag.

Leiva, R. (2007). Linear Discrimination with Equicorrelated Training Vectors, J. Multivariate Analysis, 98(2):384-409.

Mitra, P. K., Sinha, B. K. (2007). A generalized $p$-value approach to inference on common mean, J. Stat. Plann. Inference, 137(11):3634-3642.

Pan, J. X., Fang, K. T. (2002). Growth Curva Models and Statistical Diagnostics, New York: Springer-Verlag.

Roy, A., Leiva, R. (2007). Discrimination with jointly equicorrelated multi-level multivariate data. Advances in Data Analysis and Classification, 1(3):175-199.

Tan, W. Y., Gupta, R. P. (1983). On approximating a linear combination of central Wishart matrices with positive coefficients. Commun. Stat., Theory Methods, 12:2589-2600.

Weerahandi, S. (1996). Exact statistical methods for data analysis, Berlin: Springer-Verlag.

# Statistical Methods to Detect and Understand Genotype-by-Environment Interactions and QTL-by-Environment Interactions

**Paulo Canas Rodrigues**
*Nova University of Lisbon, CMA - Department of Mathematics, Portugal, paulo-canas@gmail.com; paulocanas@fct.unl.pt*

**Abstract:** Genotype-by-environment interaction (GEI) is frequent in multi-environment trials, and represents differential responses of genotypes over a wide range of agro-ecological conditions and also over a number of years. The challenge of studying GEI is to access genotype stability and genotype adaptability, to establish relationships between environments, between genotypes and between both combined, and to make predictions for a give genotype in a given environment for genotype selection. For such a modeling, multi-environment trials (METs) are conducted in which a series of genotypes is evaluated over environmental conditions and over time, being the data summarized in a two-way table with genotypes in the rows and environments in the columns. When the differences are found in the two-way table of phenotypic, the data is not additive and GEI is present in the data.
With the development of molecular markers and mapping techniques, researchers can go one step further and analyze the whole genome to detect specific locations of genes which influence a quantitative trait such as yield. These locations are called quantitative trait locus (QTL), and when these QTLs have different expression across environments we talk about QTL-by-environment interactions (QEI), which is the base of GEI. Good understandings of these interactions enable researchers to select better genotypes across different environmental conditions and, consequently, to improve crops in developed and developing countries.
In this paper I intend to present standard and new statistical methods to detect and understand GEI and QEI in the context of multi-environment plant breeding trials. These methods range from the additive two-way analysis of variance to the Eco-physiological crop growth model with QTL information, where statistical and physiological models are combined for a better understanding of GEI and QEI.

**Keywords:** Joint regression analysis, AMMI models; Eco-physiological crop growth models, Factorial regression, Plant genetics.

# 3D Retinal Vascular Network from OCT data

**Pedro Serranho[1,2], Pedro Guimarães[2,3], Pedro Rodrigues[2,3], Rui Bernardes[2,3]**

[1]*Open University,Department of Science and Technology, Portugal, pserranho@uab.pt*
[2]*University of Coimbra, Faculty of Medicine, IBILI, Portugal*
[3]*AIBILI- Association for Innovation and Biomedical Investigation on Light and Image,CNTM, Portugal*

**Abstract:** The retinal vascular system can be directly observed non-invasively and its biometric properties correlate with vision, cardiac and brain diseases. In this way, the study of the biometric properties of 2D color fundus photographies and angiograms of human retinal system has been of interested in the past years.
In this talk we will show some recent results on extracting 3D information on the retinal vascular network from Optical coherence tomography (OCT) data, using a Support Vector Machine (SVM) based classification algorithm. This development allows taking into account the depth information of the vascular network, which is neglected in 2D approaches. Moreover, preliminary results on the computation of biometric properties of the vascular network in 3D will be discussed, since they are mathematically more challenging than in the 2D case. In this sense human retinal OCT data has the potential to become a window to the general health status of the subject and in this way it can be broadly used has a screening technique.

**Keywords:** Optical coherence tomography, Classification algorithms, Vascular network.

## References

Pedro Guimarães, Pedro Rodrigues, Pedro Serranho, Rui Bernardes (to appear), 3D Retinal vascular Network from Optical Coherence Tomography data, Proceedings of 9th annual International Conference on Image Analysis and Recognition, Aveiro, Portugal, 27-29th June 2012 (to appear).

# Balanced and variance-balanced block designs

Peter J. Cameron

**Abstract:** Mathematicians and statisticians agree on the importance of 2-designs (or balanced incomplete-block designs), even if they can't agree on what to call them. Much is known about their theory, including necessary and asymptotically sufficient conditions for their existence (for fixed $k$ and $\lambda$), but open problems remain; I will discuss some of these.

To a statistician, it is natural to allow blocks to be multisets rather than sets. (If a block is a collection of plots, we are allowed to apply the same treatment to more than one of these plots.) The natural generalization of BIBDs to this situation is the class of variance-balanced designs. These may in some cases be better than designs with no repeated treatments in a block (in the sense of minimizing the maximum variance of a normalized contrast), but comparatively little is known about them. I will discuss some of what is known and pose some questions.

# Neighbour balance in a strip-block design for an experiment on irrigation

**R.A. Bailey**[1] **and Silvio Zocchi**[2]

[1]*Queen University of London,*[2]*Universidade de São Paulo*

**Abstract:** I shall describe one particular experiment on irrigation of citrus plants because it demonstrates several interesting features of experimental design. The experiment was conducted in eight blocks. Each block was a four-by-four array, with four pots in each cell and one plant in each pot. Levels of one treatment factor were applied to whole rows, and levels of the other treatment factor were applied to whole columns. The design was neighbour-balanced for each treatment factor. The design was constructed by using a design key, which also helps to elucidate the confounding and hence show the appropriate skeleton analysis-of-variance.

# CHECK PLOTS IN FIELD BREEDING EXPERIMENTS

**Stanislaw Mejza and Iwona Mejza**

*Department of Mathematical and Statistical Methods, Poznan University of Life Sciences, PL-60-637 Poznan, Wojska Polskiego 28, Poland smejza@up.poznan.pl, imejza@up.poznan.pl*

**Abstract:** In field plant breeding trials, many times it is not possible to use a traditional experimental design that satisfies the requirement of replicating all the treatments. It results from the facts of the huge number of genotypes involved, the limited amount of seed and the low availability of resources. Then, the unreplicated designs can be used for testing genotypes. By unreplicated design we mean one in which examined genotypes are replicated only once. The use of unreplicated design may be very attractive and the only possible way to carry out an evaluation (inference) of the lines. Additionally, to control the real or potential heterogeneity of experimental units and to make statistical analysis available, control (check) plots are arranged in the trial.

Some plots (check plots) with a control variety are usually placed between the plots with the lines. The frequency of check plots with the control variety is usually between 5% and 20% of the total number of plots. There are two main problems that have to be considered in the experiment, i.e. density of check plots and arrangement of them, random or systematic.

In the paper the main tool of exploring information resulting from geometry of check plots will be based on a response surface methodology. At the beginning we will try to identify response surface characterizing the environment of experiment. The obtained (estimated) response surface will be then used to adjust the observations for genotypes. It is assumed that observation is a sum of an effect resulting from environment (soil fertility) and an effect of the genotype. The difference between forecast of the observation in the given places estimated by the response surface and the observation of the experimental field is then treated as the estimate of the genotype effect. Then the ranking of genotypes is made on the basis of those differences.

# Using Real and Virtual Objects: educational feedbacks

**Vitor Cardoso**

*Open University of Portugal, Department of Sciences and Technology, Portugal, UIED-Unidade de Investigação Educação e Desenvolvimento, FCT-UNL , Portugal, vcardoso@univ-ab.pt*

**Abstract:** Objects can incorporate *k*nowledge and used to represent and illustrate it in static or dynamic forms. Well used, objects can make complex subjects more easily handled and explained. They can, eventually, provide rich, effective and enjoyable experiences for students and educators. The experiments and applications we have assisted in recent decades show that the potential of "real" objects in education can be extended to virtual objects and this assumption has lead to a disseminated, often overrated, perception that virtual objects and environments are among the most significant tools in education. Yet, from primitive solids to interactive medical models, "real" objects have played well their role and continue to do so, despite the present race for the "all virtual" we are witnessing. Research shows that using real and virtual objects in schools can have a positive impact on learning yet the best tool is not necessarily virtual. A side effect of the virtual rush is a devaluation of the potential for growth on the side of real objects. Who says they cannot evolve? What about augmented reality? Are real and virtual objects evolving and incorporating characteristics of each other? Our presentation focuses on emergent experiences involving multimedia objects and 3D environments. While exploring new ways to make learning more effective and, at the same time, based on data gathered from research, we wonder about the "when?", "where?", "how?" to use virtual technology in *r*eal education.

**Keywords:** Virtual objects, Virtual Education, 3D objects.

# ABSTRACTS

# ORGANIZED SESSIONS

# Bootstrap: data speaks for itself

**Maria da Conceição Dias Leal[1], Amílcar Oliveira[2]**

[1]*Universidade Aberta, Departamento de Ciências e Tecnologia, Estudante do Mestrado em Estatística, Matemática e Computação, Secção de Matemática, Portugal, conceicao.leal2010@gmail.com*
[2]*Universidade Aberta and CEAUL, Portugal, aoliveira@uab.pt pt*

**Abstract:** *Bootstrap* methodology, introduced by Efron in 1979, consists of a set of intensive computational methods of simulation which, if connected to the statistics analysis, enable the calculation of bias, standard error, variance and several pontual estimations, percentiles, as well as confidence intervals and significance of hypothesis tests.

They are very versatile methods which can be applied both to data and to problems of varying nature. Their application may occur in any step of the modeling and it can be parametric, semi-parametric or completely non-parametric. These methods stand as a good alternative both to classical methods and to other resampling methods since their application is very wide, they are free from the strong presuppositions of the classical methodology, and they lead to considerable reliable results, if they are issued from a representative sample of the population.

Firstly, this work tries to frame the methodology according to the nature of the resampling, proceeding then to its description. An approach is then made of the *Bootstrap* distribution of an estimator of a population parameter and of the *Bootstrap* sampling standard error estimations, of the bias and of the variance of a sampling estimation. Reference is made to the *Bootstrap* pontual estimation of population parameter which appear to be of important notice whenever the population distribution is unknown, or when data cant show a good approximation to the normal distribution. A small reference is also made to the methods which have been developed to increase the efficiency of this methodology of simulation and to the precision of the Bootstrap estimations.

Given the variety of the applications of the *Bootstrap* methodology, three important applications have been selected: confidence intervals, ANOVA and linear regression. A simple description of its implementation was then made using this resampling technique. Along this work several examples concerning the application of this methodology are presented, using the first term classifications in Mathematics and Physics-Chemistry of secondary school students. These applications were implemented, mostly, by the use of the R program boot package, which was specially conceived for the implementation of the *Bootstrap* methodology.

**Keywords:** Bootstrap Methodology, Bootstrap Estimation, Bootstrap Applications

## References

Canty, A. J. (2002). Resampling Methods in R: The boot Package. The Newsletter of the R Project, Vol. 2/3, December.

Efron,B (1979). Bootstrap Methods: Another Look at the Jackknife. Ann. Statist. Volume 7, Number 1 (1979), 1-26.

Efron, B. (1982).The Jackknife, the Bootstrap and Other Resampling Plans. Society for industrial and Applied Mathematics Philadelphia, Pennsylvania.

Efron B. e Tibshirani R. J. (1998). An Introduction to the Bootstrap. CRC Press LLC.

Gentle J., Hardle W., Mori Y.(2004). Handbook of computational statistic: concepts and methods. Springer, pp. 467-495.

Logan , Murray (2010). Biostatistical Design and Analysis Using R A Practical Guide. John Wiley & Sons, Inc.

Perez, . G. 2005). Métodos Avanzados de Estatística Aplicada: Métodos Robustos e de Remuestreo. UNED. Madrid.

Sanchez, Alex (1990). Efficient Bootstrap Simulation: An Overview. Qüestiió, Vol 14,nº 1,2,3 pp.43-88.

R Development Core Team (2012). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0

# TABELAS E GRÁFICOS PERCENTÍLICOS DE ALGUNS INDICADORES DA APTIDÃO FÍSICA DE ALUNOS DO ENSINO SECUNDÁRIO

**Domingos J. Lopes da Silva[1], Amílcar Oliveira[2]**

[1]*Escola Secundária/3 de Barcelinhos (Barcelos) e MEMC/Universidade Aberta, domingosjlsilva@gmail.com*

[2]*Universidade Aberta e CEAUL, aoliveira@uab.pt*

**Abstract:** Este estudo teve como principal objectivo a criação de tabelas e gráficos percentílicos adequadas à avaliação de algumas componentes da aptidão física da população estudantil do ensino secundário, dos 10º, 11º e 12º anos de escolaridade. A amostra é constituída por 333 alunos da Escola Secundária de Barcelinhos (Barcelos) do ensino secundário, no ano lectivo 2009-2010. A avaliação motora incidiu sobre a flexibilidade (sit-and-reach), força explosiva dos membros inferiores (salto em comprimento sem balanço), força explosiva dos membros superiores (arremesso do peso de 2kg), força abdominal (sit up's) e a resistência cardiorespiratória (teste de corrida $\frac{1}{2}$ Cooper). Paralelamente, foram medidos o peso e a estatura. O estudo é de natureza transversal, tendo as avaliações sido efetuadas no início do ano letivo (Setembro). Em termos estatísticos, foi avaliada a normalidade das distribuições pelo teste de Komogorov-Smirnov e identificada a presença de outliers pelo diagrama de extremos e quartis, de forma a obter informação respeitante ao impacto que os mesmos poderiam ter no resultado percentílico. Os dados foram lançados no Microsoft Excel, onde foram construídos gráficos, no programa LMSChartmaker 2.54 foram construídas as cartas percentílicas (P3, P10, P25, P50, P75, P90, P97), e no package estatístico do SPSS 19.0 foram realizados o cálculo dos percentis, as cartas de controlo para valores individuais e as cartas para a amplitude móvel, por variável, género e ano de escolaridade. Foram calculados, para cada prova, género e ano de escolaridade, os percentis 3, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 97. Os principais resultados mostram um crescendo em praticamente todos os testes, nos dois géneros e em todos os anos de escolaridade, à medida que se evolui na tabela percentílica. As tabelas e gráficos criados servem como um importante auxiliar ao diagnóstico, prescrição e controlo do processo de monitorização da actividade física por parte dos professores de Educação Física.

**Keywords:** Aptidão Física, Tabelas e Gráficos Percentílicos, Educação Física, Ensino Secundário.

# An Approach to Distribution of the Product of Two Normal Variables.

**Seijas-Macías, A.**
*University of A Corunha, Faculty of Economics, Spain, jasm@udc.es*

**Amílcar Oliveira**
*Universidade Aberta, Portugal and Center of Statistics and Applications, University of Lisbon, aoliveira@uab.pt*

**Abstract:** The distribution of product of normal variables come from the first part of the XX Century. First works about this issue were Craig (1936) and Aroian (1947) showed that under certain conditions the product of normal variables could be considered as a normal distribution.
A more recent approach is Ware and Lad (2003) that studied approximation to density function of the product using three methods: numerical integration, Monte Carlo simulation and analytical approximation to the result using the normal distribution. They showed tat as the inverse variation coefficient $\frac{\mu}{\sigma}$ increases, the distribution of the product of two independent normal variables tends towards a normal distribution.
Our study is focused in Ware and Lad approaches. The objective was studying which factors have more influence in the presence of normality for the product of two independent normal variables. We have considered two factors: the inverse of the variation coefficient value $\frac{\mu}{\sigma}$ and the combined ratio (product of the two means divided by variance): $\frac{\mu_1\mu_2}{\sigma}$ for two normal variables with the same variance.
Our results showed that for low values of the inverse of the variation coefficient (less than 1) normal distribution is not a good approximation for the product of normal variables. Another one, influence of the combined ratio value is less than influence of the inverse of variation coefficients value.

**Keywords:** Product of Normal Variables, Inverse Variation Coeficient, Numerical Integration, Monte Carlo Simulation, Combined Ratio.

**References**
Aroian, L.A. (1947). The probability function of a product of two normal distributed variables. Annals of Mathematical Statistics, 18, 256-271.
Craig, Cecil C. (1936). On the Frequency Function of xy. Annals of Mathematical Society. 7, 1-15.
Ware, R. y F. Lad (2003). Approximating the Distribution for Sums of Product of Normal Variables, Research-Paper 2003-15. Department of Mathematics and Statistics, University of Canterbury, New Zealand.

# ATIVAÇÃO MUSCULAR NA ANCA E JOELHO NA VARIAÇÃO DO ÂNGULO DE VALGO DURANTE A FASE DE APOIO DO SALTO VERTICAL

Arlete Queiroz[1], Tiago Atalaia[2], Maria Isabel Coutinho[3]

[1]*Escola Superior de Tecnologias da Saúde de Lisboa, Portugal, arlete_queiroz@hotmail.com,*
[2]*Escola Superior de Saúde da Cruz Vermelha Portuguesa, Portugal,*
[3]*Escola Superior de Tecnologias da Saúde de Lisboa, Portugal*

**Abstract:** A co-contração muscular apresenta um papel crucial no controlo da estabilidade articular dinâmica, nomeadamente no caso do mecanismo de valgo de joelho. A fraqueza muscular dos rotadores externos da anca pode também contribuir para alterações deste mecanismo.

Realizou-se um estudo descritivo correlacional onde se pretendeu verificar a relação entre a ativação muscular medida por Eletromiografia de superfície e a variação do ângulo de valgo de joelho medida pela análise cinemática de vídeo. Para este efeito utilizou-se o salto vertical realizado em multisaltos. A amostra foi constituída por 4 executantes (dois do género feminino e dois do género masculino com idade média de 28 anos $\pm 6, 4$).

Foram incluídos 220 saltos (110 com cada membro inferior). Verificou-se a existência de correlação significativa (p=-0,257; p=-0,239) entre a ativação muscular do reto anterior na Fase Descendente do salto à direita e à esquerda, não existindo relação do ângulo na Fase Ascendente. O género influência a dinâmica do joelho, verificando-se que as mulheres apresentam estratégias de ativação diferentes dos homens.

A fraqueza muscular da anca influência os movimentos dinâmicos do joelho no plano frontal.

**Keywords:** ativação muscular , ângulo de valgo.

# Quantificação da perfusão da parede miocárdica em estudos Single Photon Emission Computed Tomography: Influência do número de contagens

**Joana Dias[1]; Inês Figueiredo[1]; Ana Alvernaz[1]; Elisabete Carolino[2], Lina Vieira[2]**

[1]*Estudantes do 4.º Ano de Licenciatura em Medicina Nuclear - ESTeSL/IPL,*
[2]*Escola Superior de Tecnologia da Saúde de Lisboa, Instituto Politécnico de Lisboa - Lisboa, lina.vieira@estesl.ipl.pt*

**Abstract:** A técnica de imagem Single Photon Emission Tomography (SPECT) permite quantificar a perfusão miocárdica do ventrículo esquerdo nas diferentes paredes/segmentos do miocárdio. De acordo com o protocolo aplicado (protocolo de um dia ou dois dias), e de acordo com as Guidelines da European Association of Nuclear Medicine, podem ser usadas, como referência, diferentes actividades e tempos de aquisição/projecção. Assim surge o presente estudo que, tem como objectivo: avaliar a influência do número de contagens na quantificação da parede do miocárdio, tendo em conta diferentes actividades administradas, e os diferentes tempos de aquisição por projeção.

**Métodos**: Dez estudos SPECT do miocárdio simulados pelo Método de Monte Carlo, com diferentes actividades no miocárdio (3MBq, 4,2MBq, 5,4MBq e 8,2MBq ) e com diferentes tempos por projecção (15 e 30 segundos, por cada actividade). Estes estudos foram processados com o *software* Quantification Gated SPECT/Quantification Perfusion SPECT (QGS/QPS), por três operadores distintos, em que cada um processou cada estudo cinco vezes. Foram recolhidos os valores em percentagens, referentes aos 20 segmentos dos mapas polares, das paredes do ventrículo esquerdo (ápex, lateral, inferior, septal e anterior) e dos vasos (artéria descendente anterior, artéria circunflexa esquerda e artéria coronária). Na análise estatística dos dados, para a comparação dos tempos em cada actividade, utilizou-se o teste de Wilcoxon.

**Resultados**: Relativamente às actividades de 3MBq e 4,2MBq no miocárdio foram detectadas diferenças estatisticamente significativas (valor de p ¡0,05) entre os diferentes tempos de aquisição por projeção (15seg e 30seg/projeção), a nível de alguns segmentos, não se verificando diferenças estatisticamente significativas para um valor de p ¿ 0,05, a nível dos vasos ou paredes. Nas actividades 5,4 MBq e 8,2MBq não foram detectadas diferenças significativas entre os tempos de aquisição, em quaisquer dos parâmetros avaliados. **Conclusão**: Tendo em conta os resultados da investigação concluímos que, para o primeiro estudo no protocolo de um dia, sempre que se utilizar atividades iguais ou superiores a 5,4MBq no miocárdio, o correspondente a uma atividade de 500MBq de corpo inteiro, poderá ser utilizado o menor tempo de aquisição/projecção, (15seg/projeção) sem comprometer a avaliação de quantificação de perfusão da parede do miocárdio. Estes resultados vêm de encontro às orientações das Guidelines da Associação Europeia de Medicina Nuclear, quando recomenda para os primeiros estudos do protocolo de um dia 400-500 MBq, como a atividade média de corpo inteiro.

**Keywords:** Single Photon Emission Tomograph, perfusão miocárdica do ventrículo esquerdo, paredes/segmentos do miocárdio.

# Aplicabilidade da Escala de Satisfação com o Suporte Social (ESSS) a Pessoas com Síndrome de Dependência Alcoólica

**Olga Sousa Valentim**[1]; **Célia Santos**[2]; **José Pais-Ribeiro**[3]

[1]*Centro Hospitalar Psiquiátrico de Lisboa, ommvalentim@gmail.com,*
[2]*Escola Superior de Enfermagem do Porto, celiasantos@esenf.pt,*
[3]*Faculdade de Psicologia e Ciências da Educação da Universidade do Porto, jlpr@fpce.up.pt*

**Abstract:** Este estudo teve como objectivo analisar a aplicabilidade da Escala de Satisfação com o Suporte Social (ESSS) de Ribeiro (1999) a uma amostra de 444 pessoas com Síndrome de Dependência Alcoólica, diagnosticada há pelo menos um ano. Pretendemos ainda conhecer as suas possíveis relações conceptuais com a percepção de qualidade de vida através do Questionário de Estado de Saúde (SF-36v2) de Ferreira (2000a; 2000b). A amostra é predominantemente do género masculino, com uma idade média de 45 anos e sete anos de escolaridade. Os participantes são maioritariamente casados ou vivem em união de facto, desempregados, diagnosticados há cerca de oito anos e com uma média de dois internamentos. A Análise de Componentes Principais seguida de confirmação através da validade convergente-discriminante, confirmou a estrutura conceptual apresentada pelo autor com 4 domínios (satisfação com amigos/amizades-SA; intimidade-IN; satisfação com a família-SF e actividades sociais-AS), embora a amostra documentasse uma tendência para associar a satisfação com os amigos/amizades à intimidade. Obteve-se para a escala global uma consistência interna (coeficiente alpha de Cronbach) de 0,83, e para as subescalas obteve-se valores considerados aceitáveis, variando entre 0,58 para a subescala IN e 0,87 para a subescala SF, contudo é a sub-escala da satisfação com amigos/amizades, que melhor explica a variância total da escala (correlação de 0,81 com a escala global). Apesar de apresentar correlações fracas, a escala mostrou ser sensível às inter-relações conceptuais com a qualidade de vida. É também sensível às características sócio-demográficas dos participantes em análise. Podemos assim considerar que a ESSS constitui um instrumento fiável, válido e adaptado para a população com Síndrome de Dependência Alcoólica.

**Keywords:** Síndrome de dependência alcoólica, suporte social, qualidade de vida.

# ANÁLISE DA INFLUÊNCIA DA APLICAÇÃO DE KINESIO TAPE NA ATIVAÇÃO MUSCULAR DURANTE UM PASSE DE FUTSAL

**Sérgio Nuno[1]; Pedro Rebelo[1]; Tiago Atalaia[2]**

[1]*Escola Superior de Tecnologias da Saúde de Lisboa, sergiomlnuno@gmail.com*
[2]*Escola Superior de Saúde da Cruz Vermelha Portuguesa, Portugal,*

**Abstract:** Foram estudados 40 passes por cada situação definida (sem banda de Kinesio Taping e com banda para o tibial anterior e longo peroneal, aplicadas em separado) perfazendo um total de 120 comportamentos distintos. O gesto técnico foi realizado por atletas sem historial de lesão. O vídeo foi apenas analisado no sentido de se obter o instante inicial do contacto no solo e o contacto com a bola. Após avaliação do sinal eletromiográfico, registado pelo sistema de electromiografia de superfície bioPLUX®, foram estudados os dois picos presentes por se ter verificado um padrão. Verificaram-se diferenças entre as três situações em estudo. A banda de Kinesio Taping não se revelou eficaz na antecipação e aumento de ativação muscular nas condições definidas no estudo. A indivualidade de cada comportamento técnico foi estudada e não se registaram melhorias na ativação muscular e tempo de resposta muscular. Pela ausência de história anterior de lesão e pelo gesto técnico proposto, o estímulo das bandas de Kinesio Taping pode não ter sido suficiente para produzir alterações nessas variáveis.

**Keywords:** Kinesio Taping, eletromiografia, ativação muscular, futsal, passe.

# Existence of a contractive set of quadratic Lyapunov functions for a switched system

**I. Brás[1], A. C. Carapito[2], P. Rocha[3]**

[1] *University of Aveiro, Department of Mathematics, Portugal, ibras@ua.pt,*
[2] *University of Beira Interior, Department of Mathematics, Portugal, carapito@ubi.pt,*
[3] *University of Porto, Faculty of Engineering, Portugal, mprocha@fe.up.pt*

**Abstract:** The existence of a contractive set of quadratic Lyapunov functions for a switched system guarantees the stability of the system under arbitrary switching rule. We consider that at the switching instants, the system state may exhibit discontinuous jumps. These jumps are determined by linear functions that depend on the switching rule. In this work, we show that the existence of a contractive set of quadratic Lyapunov functions is guaranteed if the switching bank is stable. Therefore, we show how to achieve stabil- ity of the switched system by carefully selecting those linear functions that determine the states jumps during switching.

**Keywords:** Lyapunov quadratic function, stability, switched linear systems.

**References**

Brás, I., A. C. Carapito, and P. Rocha (2010). Estabilidade de sistemas comutados com reset parcial, PhD Thesis. Universidade de Aveiro.

Hespanha, J. P. and A. S. Morse (2002). Switching between stabilizing controllers. Automatica 38, 19051917.

# On completely *-prime *-ideals in involution rings

**D.I.C. Mendes**

*Department of Mathematics, University of Beira Interior e-mail: imendes@ubi.pt*

**Abstract:** The notion of completely *-prime *-ideal of an involution ring is introduced and investigated. In particular, it is shown that involution rings in which the zero *-ideal is completely *-prime (called *-domains) are precisely those in which there are neither symmetric nor skew-symmetric zero divisors. It is shown that such involution rings may be embedded in involution rings with unity having the same property. Several characterizations of *-domains and, in particular, of Goldie *-domains are provided. Finally, this concept is used to describe the generalized nil radical of an involution ring and the Baer radical of an involution ring is characterized in terms of its *-prime *-biideals.

**Keywords:** involution ring, completely *-prime *-ideal, *-domain, generalized nil radical, Baer radical.

**References**

Andrunakievich, V.A., On completely prime ideals of a ring, Math. USSR Sbornik 49 (1984), 285-290.

Birkenmeier, G.F. and Groenewald, N.J., Prime ideals in rings with involution, Quaest. Math. 20 (1997), 591-603.

# On (0, 1)-matrices with given line sums

**Rosário Fernandes[1], Henrique F. da Cruz[2]**

[1]*Universidade Nova de Lisboa, Departamento de Matemática, Portugal, mrff@fct.unl.pt,*
[2]*Universidade da Beira Interior, Departamento de Matemática, Portugal, hcruz@ubi.pt*

**Abstract:** Given partitions $R$ and $S$ with the same weight and $S \preceq R^*$ let $A(R, S)$ be the class of the (0, 1)-matrices with row sum $R$ and column sum $S$. These matrices play an active role in modern mathematics and the applications extend from their natural context (Matrix Theory, Combinatorics or Graph Theory) to many other areas. The Robinson-Schensted-Knuth correspondence establishes a bijection between the class $A(R, S)$ and pairs of Young tableaux with conjugate shape $\lambda$ and $\lambda^*$ with $S \preceq \lambda \preceq R^*$. We give a canonical construction for matrices in $A(R, S)$ whose insertion tableau has a prescribed shape $\lambda$, with $S \preceq \lambda \preceq R^*$. This algorithm generalizes some recent constructions due to R. Brualdi for the extremal cases $\lambda = S$ and $\lambda = R^*$.

**Keywords:** (0, 1)-matrices; Row and column vectors; Young tableaux; RSK algorithm

**References**
R.A. Brualdi (1980). Matrices of zeros and ones with fixed row and column sum vectors, Linear Algebra Appl., 33: 159-231.
W. Fulton(1997). Young tableaux, London Mathematical Society Student Texts, vol. 35: Camb- drige University Press, Cambridge.
C. Greene(1974). An extension of Schensted´s theorem, Adv. Math., 14: 254-265.
C. Schensted (1961). Longest increasing and decreasing subsequences, Canada J. Math., 13: 179-191.

# Some theory about hypermaps

**António Breda DAzevedo**[1] **and** <u>**Ilda Inácio Rodrigues**</u>[2]

[1]*University of Aveiro, Department of Mathematics, Portugal, breda@ua.pt,*
[2]*University of Beira Interior, Department of Mathematics, Portugal, ilda@ubi.pt*

**Abstract:** This work is devoted to the study of regular hypermaps: orientable, non-orientable, oriented and pseudo-orientable. More specifically, we will present basic notions in theory of hypermaps, the classification of the bicontactual hypermaps (hypermaps with the property that each hyperface meets only two others) and we reclassify, using our algebraic method, the bicontactual non- orientable hypermaps.. In the seventies, S. Wilson classified the bicontactual maps and, in 2003, Wilson and Breda dAzevedo classified the bicontactual non-orientable hypermaps. When this property is transferred for hypermaps it gives rise to three types of bicontactuality, according as the two hyperfaces appear around a fixed hyperface.

**Keywords:** Hypermap, map, regular, bicontactual.

**References**

T. R.S. Walsh (1975). Hypermapsversus bipartite maps. Journal of Combinatorial Theory Series B 18, 155163.
Corn, D.; Singerman, D. (1988). Regular hypermaps. Europ. J. Comb., 9 , 337351.
Wilson, S. (1985). Bicontactual regular maps. Pacific Journal of Mathematics, v.120, n.2, 437451.

# ZERO DIVISION GRAPH OF QUATERNION RINGS OVER $\mathbb{Z}_P$

**Rogério Serôdio**[1]**, Celino Miguel**[2]

[1]*Departamento de Matemática, Universidade da Beira Interior,e-mail: rserodio@ubi.pt,*
[2]*Departamento de Matemtica, Universidade da Beira Interior, e-mail: celino@ubi.pt*

**Abstract:** In this work, we obtained a detailed descriptions of the structure of Zero division graph $\mathbb{Z}_p[i, j, k]$. The number of vertices, the in-degree and the out-degree of the graph $\Gamma(\mathbb{Z}_p[i, j, k])$. were determined. We also studied the graph $\overline{\Gamma}(\mathbb{Z}_p[i, j, k])$, namely the degree of each vertex. Finally, we found the diameter of $\Gamma(\mathbb{Z}_p[i, j, k])$ and the girth of $\overline{\Gamma}(\mathbb{Z}_p[i, j, k])$.

**Keywords:** quaternion ring, finite field, zero divisor, zero divisor graph.

# Statistical methods in the evaluation of inflammatory parameters in dialysis patients

**Bruno Gabriel Fonseca**[1], **M. Rosário Ramos**[2], **Olga Lourenço**[3]

[1]*Master in Statistics, Mathematics and Computation, Universidade Aberta, Portugal, bgabriel1977@gmail.com,*
[2]*Universidade Aberta, Portugal, CMAF- Center of Mathematics and Fundamental Applications, University of Lisbon, marosram@uab.pt,*
[3]*CICS - Centro de Investigação em Ciências da Saúde, Universidade da Beira Interior, olga@fcsaude.ubi.pt*

**Abstract:** This presentation aims to illustrate some of the potential of statistical methodologies and current software and to show how they can support decisions in a specific area of health that are the studies in patients undergoing hemodialysis.
We performed a characterization of the test sample, a group of patients who are under hemodialysis treatment in the Dialysis Center of Cova da Beira, giving particular relevance to inflammatory parameters and their relation to clinical characteristics of the population, such as the presence or absence of diabetes Mellitus and the type of vascular access used during treatment.
We used parametric and nonparametric statistical methods to explore relationships and / or find differences between groups and, in a more specific way, we used the binary logistic regression model to identify significant predictors that discriminate patients on the variable diabetes mellitus (has / has not).

**Keywords:** Hemodialysis, Diabetes Mellitus, Statistical tests, Logistic Regression.

# Structural equation modeling and application to the study of scientific culture in high school students

**Domingas Portela**[1], **M. Rosário Ramos**[2]

[1]*Master in Statistics, Mathematics and Computation, Universidade Aberta, Portugal, domingas.portela@gmail.com,*
[2]*Universidade Aberta, Portugal, CMAF- Center of Mathematics and Fundamental Applications, University of Lisbon, marosram@uab.pt*

**Abstract:** This presentation focuses on methods of structural equation modeling with variables related with the results of a survey based on Likert scales.
Structural equation modeling is a multivariate statistical technique that helps us to see how the different latent variables are operationalized and how they relate to each other, allowing to consider in the model measurement errors explicitly.
We are presently studying the conceptual model with a view to an application on data from a sample of young high school students who attended an internship in a scientific research institution during the summer holidays. The software to be used in the future for the analysis will be AMOS for SPSS.

**Keywords:** Structural Equation modelling, Latent Variables, Scientific culture of high school students.

# Application of Path Analysis to Patient's Satisfaction with his Doctor

**Mónica Ventura**[1]**, M. Rosário Ramos**[2]

[1]*Master in Statistics, Mathematics and Computation, Universidade Aberta, Portugal, moni15@sapo.pt,*
[2]*Universidade Aberta, Portugal, CMAF- Center of Mathematics and Fundamental Applications, University of Lisbon, marosram@uab.pt*

**Abstract:** Path Analysis is a technique that had in its origins the goal of modeling causal relationships between observed variable. This is one of the objectives of several studies in Psychology, where researches pretend to study the relations between the variables in observation and to look for variables that are crucial for others, exploring statistical models that attempt to translate conceptual models connecting important variables to a study. Thus, Path Analysis can be used to test previously idealized and studied models or to experiment new conceptions.

In this presentation will be shown an application of Path Analysis to modeling the level of patient's satisfaction with his doctor, conducted using data of an international survey on the general theme of patient's satisfaction with communication with health professionals. We will show the statistical techniques in which Path Analysis relies and what kind of statistical conclusions it's possible to draw.

**Keywords:** Path Analysis, Linear Regression, Structural Equation Modeling, Patient's Satisfaction

# Convergence of the Finite Element Method for the Porous Media Equation in 2D with Variable Exponent

**José C.M. Duque[1], Rui M.P. Almeida[1] and Stanislav N. Antontsev[2]**

[1]*Universidade da Beira Interior,Faculdade de Ciências,Departamento de Matemática,e-mail: jduque,ralmeida@ubi.pt, web:http://www.cmatubi.ubi.pt,*
[2]*Universidade de Lisboa, Faculdade de Ciências, Centro de Matemática e Aplicações Fundamentais, e-mail:anton@ptmat.fc.ul.pt, web:http://cmaf.ptmat.fc.ul.pt*

**Abstract:** In this work, we make a study of the convergence of the Finite Element Method (FEM) when applied to the following partial differential equation of parabolic type,

$$u_t = div(u^{\gamma(x)}\nabla u) + f(x,t).$$

Since the problem may be of degenerate type, we utilize an approximate pro- blem, regularized by introducing a parameter $\epsilon$. We prove, under certain conditions on $\gamma$ and $f$, that the weak solution of the approximate problem converges to the weak solution of the initial problem, when the parameter e tends to zero. Discrete solutions are built using the FEM and the conver- gence of these for the weak solution of the approximate problem is proven. Finally we present some numerical results of a MatLab implementation of the method.

**Keywords:** MEF, PME, Variable Exponent.

## References
S.N. Antontsev and S.I. Shmarev (2005). A model porous medium equation with variable exponent of nonlinearity: existence, uniqueness and localization properties of solutions. Journal of Nonlinear Analysis 60, 515545.
V. Thomée (1997). Galerkin Finite Element Methods for Parabolic Problems, Springer Series in Computational Mathematics, Vol. 25, Springer-Verlag, Berlin Heidelberg New York.

# An approximate solutions to a class of nonlinear partial differential equations

## Paulo Rebelo

*Universidade da Beira Interior, Departamento de Matemática, Covilhã, Portugal, rebelo@ubi.pt*

**Abstract:** The aim of this paper is to obtain an approximate solution to the initial boundary value problem to a class of nonlinear PDE. The Fourier Method is combined with the Adomians decomposition method in order to provide an approximate solution. One example of application is provided.

**Keywords:** Fourier Method, Initial boundary value problem. Adomian Decomposition Method, Nonlinear partial differential Equations, Approximate solution.

## References

Adomian, G. Nonlinear Stochastic Systems and Application to Physics. Kluwer, 1989.

Adomian, G. Solving Frontier Problems of Physics: The Decomposition Method. Kluwer, 1994.

A. Wazwaz. A new algorithm for calculating adomian polynomials for nonlinear operators. Appl. Math. Comput., 111(1):3351, 2001.

P. Rebelo, On the approximate solution to an initial boundary valued problem for the Cahn-Hilliard equation, Commun. Nonlinear Sci. Numer. Simul. 15(2) (2009), pp. 225-231.

P. Rebelo, An approximate solution to an initial boundary value problem to the one-dimensional Kuramoto-Sivashinsky equation, International Journal for Numerical Methods in Biomedical Engineering, 27(6) (2011), pp. 874-881.

P. Rebelo, An approximate solution to an initial boundary value problem: Rakib-Sivashinsky equation, International Journal of Computer Mathematics 89(7) (2012), pp. 881-889.

# A Matlab implementation of the moving finite element method for the solution of time- dependent models

**Rui J. Robalo[1], Rui M. Almeida[1], Maria do Carmo Coimbra[2], Alírio E. Rodrigues[2]**

[1]*Department of Mathematics, Faculty of Science, University of Beira Interior Rua Marquês DÁvila e Bolama, 6200-001 Covilhã, Portugal e-mail: rrobalo, ralmeida@ubi.pt, web: http://www.cmatubi.ubi.pt,*
[2]*LSRE - Laboratory of Separation and Reaction Engineering - Associate Laboratory LSRE/LCM, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, 4200- 465 Porto, Portugal e-mail: mcoimbra, arodrig@fe.up.pt, web: http://www.lrse.fe.up.pt*

**Abstract:** This paper provides a MATLAB implementation of the Moving Finite Element Method (MFEM) for the solution of systems of time-dependent Partial Differential Equations (PDEs). The numerical algorithm is based on the MFEM with polynomial approximations of any degree in each of the finite elements. These approximations are obtained by Lagrange interpolation. The MFEM is a continuously moving grid method, where the node movement and PDE integration are coupled. In our approach, a different dynamically-adaptive grid is associated to each one of the dependent variables. Some numerical tests to evaluate the performance of our Matlab code based on the MFEM are investigated. Numerical results demonstrate that the Moving Finite Element Method is an appropriate technique for computing the numerical solution of mathematical models of diffusion-convection and reaction problems.

**Keywords:** Matlab solver, Moving finite elements, Adaptive grids, Moving boundary problems, Time-dependent partial equations.

**References**

Miller, K.: "Moving finite elements, Part II", SIAM J. Numer. Anal., Vol. 18, nº 6 (1981), pp. 1033-1057.

Coimbra, M. C.; Sereno, C. A.; Rodrigues, A. E.: "Moving finite element method: applications to science and engineering problems", Computers and Chem. Engng., Vol. 28 (2004), pp. 597-603.

Robalo, R. J.; Almeida, R. M.; Coimbra, M. C.; Rodrigues, A. E.: "Implementação Matlab do Método dos Elementos Finitos Móveis para a Resolução de Problemas de Convecção-Difusão- Reacção", in Congresso de Métodos Numéricos em Engenharia 2011, A. Tadeu et al. Eds, Coimbra (2011), pp. 263.

# Cost considerations in statistical work

**Dinis Pestana**

*CEAUL   Center of Statistics and Applications of University of Lisbon, Portugal, dinis.pestana@fc.ul.pt*

**Abstract:** Data gathering, data cleaning, data storing and management, and data analysis strategies have a considerable impact in the budget of statistical work, and it is common knowledge that Walds sequential analysis and its applications in acceptance sampling has been classified information during World War II.
We briefly discuss issues concerning monetary or welfare costs and public interest concerns in public health programs and in quality control, with relevance to compound analysis.

**Keywords:** Compound analysis, sequential analysis, cumulative meta analysis.

# Meta-analysis techniques applied in prevalence rate estimation

**João Paulo Martins[1], Miguel Felgueiras[1,2], Rui Santos[1]**

*mfelg@ipleiria.pt, rui.santos@ipleiria.pt, jpmartins@ipleiria.pt,*
[1]*School of Technology and Management, Polytechnic Institute of Leiria, CEAUL Center of Statistics and Applications of University of Lisbon,*
[2]*CIIC Computer Science and Communications Research Centre of Polytechnic Institute of Leiria Portugal*

**Abstract:** In some cases, the estimators obtained in compound tests have better features than the traditional ones, obtained from individual tests, c.f. Sobel and Elashoff (1975), Garner et al. (1989) and Loyer (1983). The bias, the eficiency and the robustness of these estimators is investigated in several pa- pers, e.g. Chen and Swallow (1990), Hung and Swallow (1999) and Lancaster and Keller-McNulty (1998). Thus, the use of estimators based on compound tests not only allows a substantial saving of costs, but they also can (in some situations) be more accurate than the estimators based on individual tests.

Nevertheless, each laboratory produces estimates for the prevalence rate of a given infection using different methodologies, such as halving nested procedures (Sobel and Elashoff, 1975) and square array testing (Kim et al. , 2007). The logistic regression or the weighted least squares regression can be used in order to combine different prevalence rate estimates (Chen and Swallow, 1990). In this work are proposed some meta-analytical techniques (Kulinskaya et al. , 2008) as an alternative approach. This methodology has the advantage of being quite simple and flexible to account for the error source.

**Keywords:** Compound tests, estimation of prevalence, meta-analysis.

**References**

Chen, C.L., Swallow, W.H. (1990). Using group testing to estimate a proportion, and to test the binomial model, Biometrics 46 , 10351046.

Garner, F.C., Stapanian, M.A., Yfantis, E.A., Williams, L.R. (1989). Probability Estimation With Sample Compositing Techniques, Journal of Oficial Statistics 5 , 365374.

Hung, M., Swallow, W.H. (1999). Robustness of Group Testing in the Estimation of Proportions, Biometrics 55 , 231237.

Kim, H., Hudgens, M., Dreyfuss, J., Westreich, D., and Pilcher, C. (2007). Comparison of group testing algorithms for case identification in the presence of testing errors, Biometrics 63 , 11521163.

Kulinskaya, E., Morgenthaler, S., Staudte, R.G. (2008). Meta Analysis: a guide to calibrating and combining statistical evidence, Wiley, Chichester, England.

Lancaster, V.A., Keller-McNulty, S. (1998). A Review of Composite Sampling Methods, Journal of the American Statistical Association 93 , 1216-1230.

Loyer, M.W. (1983). Bad Probability, Good Statistics, and Group Testing for Binomial Estimation, The American Statistician 37 , 5759.

Sobel, K.M., Elashoff, R.M. (1975). Group testing with a new goal, estimation, Biometrika 62, 181193.

# The number of bacteria in discrete quantitative compound analysis

**Rui Santos[1], João Paulo Martins[1]**

[1]*School of Technology and Management, Polytechnic Institute of Leiria, CEAUL Center of Statistics and Applications of University of Lisbon, Portugal, rui.santos@ipleiria.pt, jpmartins@ipleiria.pt*

**Abstract:** Compound analysis, first used in qualitative blood testing (Dorfman, 1943), can be used also in quantitative tests, with considerable saving of resources in a wide variety of problems, cf. Boswell et al. (1996). In the present work it is assumed that the amount of the substance of interest (e.g. number of bacteria) in each individual is characterized by some known discrete distribution and that the tests are subject to some error sources (misclassification measured by sensitivity and specificity). Thus, for some commonly used count distributions (Johnson et al. , 2005) the bacteria number in a pooled sample is computed using hierarchical models. The sensitivity and specificity of a compound test in Dorfmans methodology is described as a function of the sensitivity and specificity of a single sampling unit test (Kim et al., 2007), in order to investigate the influence of dilution and rarefaction on misclassification in compound tests. Some suggestions to deal with subpopulations with different prevalence rates are also discussed (Finucan, 1964).

**Keywords:** Compound analysis, sensitivity, specificity, rarefaction, discrete distributions.

### References

Boswell, M.T., Gore, S.D., Lovison, G., Patil, G.P. (1996). Annotated bibliography of composite sampling, Part A: 193692, Environmental and Ecological Statistics **3** , 150.

Dorfman, R. (1943). The detection of defective members in large populations, Ann. Math. Statistics **14** , 436440.

Finucan, H.M. (1964). The blood testing problem, Applied Statistics **13** , 4350.

Johnson, N.L.; Kemp, A.W. and Kotz, S. (2005), Univariate Discrete Distributions, 3rd Edition, John Wiley & Sons.

Kim, H., Hudgens, M., Dreyfuss, J., Westreich, D., and Pilcher, C. (2007). Comparison of group testing algorithms for case identification in the presence of testing errors, Biometrics **63** , 11521163.

# The prevalence rate estimation in compound tests with misclassification

**Ricardo Sousa[1], Rui Santos[2]**

[1]*Higher School of Health Technology of Lisbon, Polytechnic Institute of Lisbon, CEAUL  Center of Statistics and Applications of University of Lisbon, ricardo.sousa@estesl.ipl.pt,*
[2]*School of Technology and Management, Polytechnic Institute of Leiria, CEAUL Center of Statistics and Applications of University of Lisbon, Portugal, mfelg@ipleiria.pt*

**Abstract:** The prevalence rate estimation in rare diseases using compound tests can create large benefits in order to minimize the costs of performing a large number of blood tests. The Dorfmans theory (1943) is rather limited for two main reasons: it doesnt take sensitivity and specificity into account, and its scope is limited to qualitative analysis. This work continues the investigation of Sousa (2005, 2006, 2008) which attempts to overcome the limitations of the original theory of Dorfman. Hence, the main goal is to compare, by simulation, different methodologies to estimate the prevalence rate using compound tests (Sobel and Elashoff, 1975; Loyer, 1983; Garner et al., 1989) with misclassification, i.e. with sensitivity and specificity taking values lower than one (Chen and Swallow, 1990; Hung and Swallow, 1999).

**Keywords:** Compound analysis, Dorfman theory, sensitivity, specificity, simulation.

**References**

Chen, C.L., Swallow, W.H. (1990). Using group testing to estimate a proportion, and to test the binomial model. Biometrics 46 , 10351046.

Dorfman, R. (1943). The detection of defective members of large populations, Ann. Math. Statist. 14 , 436440.

Garner, F.C., Stapanian, M.A., Yfantis, E.A., Williams, L.R. (1989). Probabil- ity Estimation With Sample Compositing Techniques. Journal of Oficial Statistics 5 , 365374.

Hung, M., Swallow, W.H. (1999). Robustness of Group Testing in the Estimation of Proportions. Biometrics 55 , 231237.

1Loyer, M.W. (1983). Bad Probability, Good Statistics, and Group Testing for Binomial Estimation. The American Statistician 37 , 5759.

Sobel, K.M., Elashoff, R.M. (1975). Group testing with a new goal, estimation. Biometrika 62 , 181193.

Sousa, R. (2005). Análises ao sangue conjuntas. Uma estratégia para redução de custos, Livro de actas da IV Conferência Estatística e Qualidade na Saúde, 190196.

Sousa, R. (2006). Análises conjuntas ao sangue: uma abordagem mais realista, Estatística e Qualidade na Saúde: Problemas e Temáticas, 129133

Sousa, R. (2008). O problema de Dorfman revisitado  Análises quantitativas, Estatística e Qualidade na Saúde: novos rumos, novos desafios, 140144.

# Canonical Correspondence Analysis vs Co-Inertia analysis to find patterns of response in spatio-temporal ecological relationships: an empirical comparison

**Fernández-Gómez, Mª José[1], Barrera-Mellado, Inmaculada[2], Mendes, Susana[3], Miranda, U.[4], Martín-Vallejo, Javier[5]**

[1] *Universidad de Salamanca, Departamento de Estadística, España, mjfg@usal.es,*
[2] *Universidad de Salamanca, Departamento de Estadística, España, ibm@usal.es,*
[3] *Instituto Politécnico de Leiria, Peniche, Portugal,*
[4] *ULISSES,*
[5] *Universidad de Salamanca, Departamento de Estadística, España, jmv@usal.es*

**Abstract:** Canonical Correspondence Analysis [CCA, ter Braak (1986)] and Co-Inertia Analysis [CoIA, Dolédec and Chessel, (1994); Dray et al. (2003)] are two ad-hoc techniques focused to analyze patterns of response in spatio-temporal relationships between species and environment, when the information is set as two data tables: the first one ($X_{nxp}$) containing environmental parameters (in columns) recorded in a set of sampling sites (in rows) and the second one (Ynxq) containing species data recorded at the same sites. These tables are called a "pair of ecological tables".

Both techniques fall within the set of multivariate direct gradient analysis methods, where the information provided by the two set of data is combined directly into a single analysis.

However, each technique has its advantages and disadvantages. While CCA is an asymmetrical technique and assumes that the species play the roles of response variables, CoIA is a symmetric one, and therefore more general, because it assumes no directionality in the relationship.

CCA is restricted to the case where the number of cases (sampling sites) is not less that the number of explanatory variables, in other case the analysis is just a Correspondence Analysis of the community composition data. CoIA, on the other hand, has the advantage that the researcher can handle more environmental variables than samples.

The advantages and disadvantages of the two methods in addressing the problem of coupling two tables are presented, showing how they can be used simultaneously to extract the patterns of the relationships between species-environmental variables. From our point of view the complementary use of both techniques would improve the ecological interpretation, which will be showed by applying both techniques to the same paired of ecological tables.

**Keywords:** Canonical Correspondence Analysis, Co-Inertia Analysis, Multivariate Direct Gradient Analysis. Paired of ecological tables.

### References

Dolédec, S.; Chessel, D. (1994). Co-Intertia analysis: an alternative method for studying species-environment relationships. Freshwater Biology, 31, 277-294.

Dray, S.; Chessel, D.; Thioulouse, J. (2003). Co-inertia analysis and the linking of ecological tables. Ecology, 84, 3078-3089. Ter Braak, C.J.F. (1986). Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology, 67, 1167-1179.

# USING MULTITABLE TECHNIQUES FOR ASSESSING THE TEMPORAL VARIABILITY OF SPECIES-ENVIRONMENT RELATIONSHIP IN A COPEPOD COMMUNITY FROM A TEMPERATE ESTUARINE ECOSYSTEM

**Sónia Cotrim Marques[1], Miguel Ângelo Pardal[2], Susana Mendes[3], Ana Ligia Primo[4], Ulisses Miranda Azeiteiro[5]**

[1] *CEF - Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Box 3046, 3001-401 Coimbra, Portugal. scotrim@ci.uc.pt.,*
[2] *CEF - Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Box 3046, 3001-401 Coimbra, Portugal. mpardal@ci.uc.pt.,*
[3] *GIRM - Marine Resources Research Group, School of Tourism and Maritime Technology, Polytechnic Institute of Leiria, 2520-641 Peniche, Portugal. susana.mendes@ipleiria.pt.,*
[4] *CEF - Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Box 3046, 3001-401 Coimbra, Portugal., ana.primo@student.zoo.uc.pt.,*
[5] *University Aberta, Department of Science and Technology, 4200-055 Porto, Portugal and CEF - Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Box 3046, 3001-401 Coimbra, Portugal. ulisses@uab.pt.*

**Abstract:** This study investigated copepod assemblages' distribution patterns and determined the main hydrological and physical factors structuring it at different temporal scales. Copepod and hydrographic surveys were conducted seasonally during a one-year period, at a fixed station in Mondego estuary. Hourly samples were collected, during a diel cycle from sub-surface and 1 m above the bottom, at spring tides. Temporal variability of copepod community was analyzed using the STATICO method. Seasonal variability of the environmental conditions, linked with periods of higher freshwater flow, were the main drivers responsible for copepod community structure observed. Seasonal changes related to environmental factors showed to be strongest in spring and in winter, which accounted with higher species density. Marine and estuarine copepods showed maximal abundance in spring and summer, while the period of higher river flow, coincidently with autumn and winter, imposed changes in the copepods composition, emphasizing the changes in saline zonation within the estuary due to regime of river discharge. STATICO also emphasized the effects of physical forcing at short-term temporal scales. The stable part of the species-environment relationships resulted from the combination of a salinity gradient linked to a fresh, brackish- and marine species gradient, which was dependent on tidal cycle. The STATICO method proved to be a valuable statistical tool to get a clear representation of temporal organization of the copepod assemblages.

**Keywords:** Temperate estuarine ecosystem, copepod assemblages, environmental forcing, STATICO.

# Text Mining Using Cartesian and Cylindrical Biplots

**Valter M.Vairinhos[1], M. Purificación Galindo[2]**

[1]*CINAV - Escola Naval e CENTEC - IST, valter.vairinhos@sapo.pt,*
[2]*Departmento de Estadística, Universidad de Salamanca, pgalindo@usal.es*

**Abstract:** In this work we present the concept of cylindrical biplot as opposed to Cartesian biplot, referring the representation of biplots in parallel coordinates and we show its usefulness as a text mining tool. We show that the natural equivalent, in a referential of parallel coordinates, of the usual biplot represented in an orthogonal referential system is obtained representing the results of singular value decomposition of the data matrix in a system of parallel coordinates having as axes not the observed variables but the singular vectors of such decomposition. This representation opens the way to see clusters of objects (individuals, variables and both variables and individuals) and other concepts in a space with any dimension. This concept allows a natural integration between the two kinds of representation, useful to interpret data, discover its structure, specifically in text analysis. Those ideas are implemented in a new version of the software BiplotsPMD and applied to the analysis of texts of Portuguese students answering questions of examinations.

**Keywords:** Biplot, Cartesian Biplot, Cylindrical Biplot, Text Mining.

**References**

GABRIEL, K. R. (1971). The biplot graphic of matrices with application to principal component analysis. Biometrika, 58 (3), 453-467.

GALINDO, M. P. (1985) Contribuciones a la Representación Simultanea de Datos Multidimensionales. Tesis Doctoral. Universidad de Salamanca.

GALINDO, M. P. (1986) Una Alternativa de Representación Simultánea: HJ-Biplot. Qestiió 10(1) pp 12-13.

INSELBERG, ALFRED (2009) Parallel Coordinates. Visual Multidimensional Geometry And Its Applications. Springer.

VAIRINHOS, V. M. (2003). Desarollo de un sistema de minería de datos basados en los metodos biplot. Tesis Doctoral. Universidade de Salamanca.

VAIRINHOS, V. M.; GALINDO, M. P. (2011) Graphical Text Mining Using HJ Biplots. XIIIth Spanish Biometry Conference and 3rd Ibero-American Biometry Meeting - 2011.

VAIRINHOS, V. M.; GALINDO, M. P. (2012) Biplots Cilindricos. XIX Jornadas de Classificação e Análise de Dados (JOCLAD2012), Livro de Resumos, ISBN: 978-972-9473-60-9pp 132-137.

# APTIDÃO CARDIORESPIRATÓRIA E ÍNDICE DE MASSA CORPORAL EM ALUNOS DO ENSINO SECUNDÁRIO

**Domingos J. Lopes da Silva[1], Teresa Oliveira[2]**

[1]*Escola Secundária/3 de Barcelinhos (Barcelos), MEMC/UAb, domingosjl-silva@gmail.com,*
[2]*Universidade Aberta e CEAUL toliveir@uab.pt*

**Abstract:** Este estudo teve como propósito central a descrição e comparação da aptidão cardiorespiratória, medida pela performance na corrida de resistência de 6 minutos (CR6min), entre alunos do ensino secundário, de ambos os sexos, dos 10º, 11º e 12º anos de escolaridade. A amostra é formada por 319 sujeitos, alunos do ensino secundário (10º ano: rapazes, n=35; raparigas, n=68; 11º ano: rapazes, n=43; raparigas, n=78; 12º ano: rapazes, n=33; raparigas, n=62). A resistência cardiorespiratória foi avaliada pelo teste de CR6min no final do ano lectivo. Paralelamente, foi avaliado o peso, a estatura e o índice de massa corporal. Entre os rapazes, o evoluir na escolaridade não é significativamente acompanhado por um crescendo nas medidas antropométricas de base e na aptidão cardiorespiratória. Entre as raparigas, apenas a performance na CR6min sofre um incremento com o evoluir dos anos de escolaridade, sendo significativamente mais elevada no 12º ano de escolaridade.

**Keywords:** corrida de resistência, teste $\frac{1}{2}$ Cooper, aptidão aeróbia, resistência cardiorespiratória, adolescentes.

# PLANOS EM BLOCOS INCOMPLETOS EQUILIBRADOS COM REPETIÇÕES: UMA APLICAÇÃO NA ÁREA DA EDUCAÇÃO

**Helga Correia[1]; Teresa Oliveira[2]**

[1] *MEMeC - Universidade Aberta,*
[2] *Universidade Aberta e CEAUL, toliveir@uab.pt*

**Abstract:** O principal objetivo de um Planeamento de Experiências reside essencialmente na procura de relações entre variáveis e na comparação de níveis de fatores, recorrendo ao tratamento estatístico dos dados recolhidos. A utilização de blocos no Planeamento de Experiências é fundamental, pois permite reduzir ou eliminar a variabilidade introduzida por fatores que podem influenciar a experiência mas que não interessam e/ou não foram explicitamente incluídos durante o planeamento.
Tendo como base o planeamento em blocos, apresentamos uma aplicação dos BIBDR na área da Educação com o objetivo de comparar cinco domínios do pensamento algébrico de uma amostra de alunos do 1º ano do ensino superior em Cabo Verde. Para a análise dos dados da amostra foi utilizado o software R, versão 2.12.1. Podemos constatar que existem diferenças significativas entre alguns dos domínios do pensamento algébrico, nomeadamente entre os domínios da Generalização da Aritmética e Tecnicismo Algébrico com os restantes domínios.
Recomendamos a escolha de uma amostra mais representativa constituída por alunos de todas as instituições superiores de Cabo Verde.

**Keywords:** Planeamento de Experiências; BIBD; BIBDR na Educação.

## References

BOSE, R. C. (1939): On the construction of Balanced Incomplete Block Designs, Annals of Eugenics, Vol. 9, 353-399

OLIVEIRA, T. (1994): Planos de Blocos Equilibrados Incompletos com repetições, Tese de Mestrado, Faculdade de Ciências da Universidade de Lisboa.

OLIVEIRA, T. (2000): Planeamento de Experiências: Novas perspectivas, Tese de Doutoramento, Universidade de Lisboa.

OLIVEIRA, T. A. (2010): BIB designs with repeated blocks: review and perspectives, The Islamic Countries Society of Statistical Sciences, Vol. I, 82-96.

PONTE, J. P. (2005): Álgebra no currículo escolar, Educação & Matemática: Revista da Associação de professores de matemática. nov-dez. 2005. p. 36-42.

YANG, J. (1985): An example of the application of a Balanced Incomplete Block Design. Assesssment and Evaluation in Higher Education, Vol. 10, Issue 3, pages 250-253.

# Some considerations about models with application to data with hierarchical or multilevel structure

**Vítor Valente[1], Teresa A. Oliveira[2]**

[1]*Universidade Aberta e CEAUL, Portugal, vitormfv@gmail.com,*
[2]*Universidade Aberta e CEAUL, Portugal, toliveir@univ-ab.pt*

**Abstract:** People tend to exist within organizational structures, such as families, schools, towns or countries. In education, students exist within a hierarchical social structure that can include family, peer group, classroom and school, among others. Hierarchical, or nested, data structures are common throughout many areas of science. By hierarchy we mean that units of a certain level (also referred micro units) are grouped into, or nested within, higher level (or macro) units. In these cases, the units within a cluster tend to have more differences than units from other clusters, i.e. they are correlated. Thus, unlike in the classical setting, where there exists a single source of variation between observational units, the heterogeneity between clusters introduces an additional source of variation and complicates the analysis.

For instance, students are not randomly assigned to classrooms from the population, (but rather are assigned to schools based on geographic factors) so they tend to be more similar to each other than students randomly sampled from the entire population. These students tend to share certain characteristics (environmental, demographic, or otherwise) and observations based on these students information are not fully independent.

Using classical regression, bias is produced in the typical error of measurement and an increased likelihood of committing errors of inference. The hierarchical linear or multilevel models are most suitable because they consider the hierarchical relationships and also provide estimates on the contextual variability of regression coefficients.

In practice, often the data structures are not hierarchical, but have more complex structures such as cross-classification (level 2 or macro) - for instance, students (level 1 or micro) to attend different courses at a school while in other schools there are students who attend the same courses, or multiple membership models - for instance, some students may attend more than one school.

Now, with several software programs available (MLwiN, HLM, SPSS, R) and some more readable texts and treatments on the topic, we need to be aware of the issue, and how it should be dealt with. The aim of this study is to discuss the problem and make some remarks how it can be dealt with appropriately.

**Keywords:** Hierarchical Linear Model, Multilevel Model, Cross-classified model, education.

**References**

Bryk, A. and Raudenbush, S., (1992). Hierarchical Linear Models: applications and data analysis methods, Sage Publications, London. Goldstein, H., (1995), Multilevel Statistical Models, 2nd Edition, Edward Arnold, London.

Hox, J., (2002), Multilevel Analysis: Techniques and Applications, Lawrence Erlbaum Associates, Mahwah, NJ.

Kreft, I. and de Leeuw, J., (1998), Introducing Multilevel Modeling, Sage Publications, London.

Longford, N. T., (1993), Random coefficient models, Oxford University Press, New York.

MLwiN, version 2.23 (2011), Centre for Multilevel Modelling, University of Bristol. Software authors: Rasbash, J.; Browne, W.; Healy, M.; Cameron, B. and Charlton, C.

Rasbash, J. and Browne, W., (2008), Non-Hierarchical Multilevel Models, in J. de Leeuw, E. Meijer (eds.), Handbook of Multilevel Analysis, Springer.

Rasbash, J.; Steele, F.; Browne, W. and Prosser, B., (2004), A User's Guide to MLwiN: version 2.0, Centre for Multilevel Modelling, Institute of Education, University of London, London.

Snijders, T. and Bosker, R., (1999), Multilevel Analysis: An introduction to basic and advanced multilevel modeling, Sage Publications, London.

Valente, V. and Oliveira, T.A. (2011). Application of HLM to data with multilevel structure. Discussiones Mathematicae Probability and Statistics 31 (2011) 87-101.

Valente, V. and Oliveira, T.A. (2009). Hierarchical Linear Models in Education Sciences: an Application. Biometrical Letters, vol. 46(1), 2009, pp. 71-86.

Valente, V. and Oliveira, T.A., (2006), "Diferentes áreas científicas, diferentes notas? Uma aplicação do MLH a alunos do 10º ano", IIWEMC - Workshop de Estatística, Matemática e Computação, Universidade Aberta, Lisboa.

# Quadrados Latinos e Planos Projetivos: Revisão de problemas clássicos

**Ricardo Mendes[1], Teresa A. Oliveira[2]**

[1]*UAb, MEMC, Portugal, rmendes@sapo.pt,*
[2]*Universidade Aberta e CEAUL, Portugal, toliveir@univ-ab.pt*

**Abstract:** Por Quadrado Latino de ordem n entende-se uma matriz quadrada de dimensão n onde cada uma das letras latinas ocorre uma única vez, quer em cada linha quer em cada coluna. Surgindo inicialmente como curiosidades matemáticas, os Quadrados Latinos assumem um papel fundamental em Planeamento de Experiências, nomeadamente em experiências em que dois tipos de fontes de variação afetam os dados, permitindo diferenciar estas fontes de variação do erro experimental.

São conhecidas aplicações dos Quadrados Latinos em áreas tão diversas como a Engenharia, a Medicina, a Agricultura e o Ensino, sendo também inúmeras as ligações dos Quadrados Latinos a outras áreas da Matemática: são exemplos os Quasigrupos, os Grafos e os Planos de Incidência. Encontram-se ainda na literatura aplicações dos Quadrados Latinos a aspetos puramente recreativos, como o Sudoku e o Skyscrapers Puzzle.

O presente trabalho explora o conceito de ortogonalidade entre Quadrados Latinos, ilustrando-o através de dois problemas clássicos: o Problema das 16 cartas e o Problema dos 36 oficiais. Será investigada a relação entre os Quadrados Latinos e os Planos Projetivos, com abordagem ao teorema de Bruck-Ryser-Chowla, sendo apontadas perspetivas de investigação futura.

**Keywords:** Quadrados Latinos, Ortogonalidade, O Problemas das 16 cartas, O Problema dos 36 oficiais, Planos Projetivos, Teorema de Bruck-Ryser-Chowla

## References

Alegri, Mateus (2006). Quadrados latinos e aplicações. Instituto de Matemática Estatística e Computação Científica, Universidade Estadual de Campinas. Dissertação de Mestrado.

Cameron, Peter J. (2007). Sudoku An alternative history. Talk to the Archimedeans.

Cardoso, Domingos Moreira (2004). Grafos e Combinatória. Departamento de Matemática, Universidade de Aveiro. Notas de aula.

Laywine, C.F. and Mullen, G.L. (1998). Discrete Mathematics Using Latin Squares. Wiley Interscience. Series in Discrete Mathematics and Optimization.

Mendes, Ricardo e Oliveira, Teresa (2011). Quadrados Latinos, Sudoku e Futebol. 1ªs Jornadas de Matemática da universidade Aberta.

Montgomery, Douglas C. (2001). Design and Analysis of Experiments. 5ª edição, John Wiley & Sons.

# Extreme values from a dynamical perspective

**Jorge Freitas**

*Faculty of Sciences - University of Porto*

**Abstract:** We will give conditions to prove the existence of Extreme Value Laws (EVL) for stochastic processes arising from deterministic dynamical systems. We will also establish a connection between EVL and Hitting Times Statistics (HTS) which have to do with the existence of a distributional limit for the waiting time before the orbits of the system hit asymptotically small sets in the phase space. Finally, we will give a link between the existence of an Extremal Index (EI) less than 1 and the occurrence of periodic phenomena. The EI essentially quantifies the amount of clustering of high values and, in many cases, it can be identified with the inverse of the average cluster size. No clustering means that the EI equals 1.

# Hyperbolicity in dissipative polygonal billiards.

**João Lopes Dias, G. del Magno, P. Duarte, J. P. Galvão and D. Pinheiro**
*ISEG - Technical University of Lisbon*

**Abstract:** A billiard is a mechanical system consisting of a point-particle moving freely inside a planar region and being reflected off the perimeter of the region according to some reflection law. The specular reflection law is the familiar rule that prescribes the equality of the angles of incidence and reflection. Billiards with this reflection law are conservative systems, and as such are models for physical systems with elastic collisions. For this reason and their intrinsic mathematical interest, conservative billiards have been extensively studied. Much less studied are dissipative billiards, which originate from reflection laws requiring that the angle of reflection is a contraction of the angle of incidence. These billiards do not preserve the Liouville measure, and therefore can model physical systems with non-elastic collisions. We will present the case of polygonal billiard tables, whose dynamics differs strikingly from the one of its conservative counterparts.

# Asymptotic behavior of infinite dimensional compact skew-product flows

**Glória Ferreira**

*Faculty of Sciences - University of Porto.*

**Abstract:** Consider a compact, connected, Hausdorff manifold $X$, an ergodic flow $\varphi^t : X \to X$ and a Borel $\varphi^t$-invariant probability measure $\mu$. Let $A : X L(H, H)$ be a continuous map, where $H$ is a Hilbert space, and we intend to study the asymptotic behavior of the solutions of $\dot{X} = A(\varphi^t(x))X$ when they are compact operators (for $t > 0$). The Bessa-Carvalho dichotomy for the discrete-time case establishes, in rough terms that, for most of the choices of the dynamical systems and most choices of the points $x \in X$, we have a weak form of hyperbolicity or else the limit operator along the orbit of $x$ is the null operator. The first step is understand the role of Lyapunov exponents in the study of the asymptotic behavior of the solutions and the definition of dominated splitting in the infinite dimensional case. Then, based on the work done in the discrete-time, understand the problems that appear due to continuous-time, try to perform a perturbation-scheme of our original system A and, finally, ask if the dichotomy still holds in our context.

# Exponential stability for impulsive functional differential equations with infinite delay.

**José J. Oliveira, Teresa Faria, Marta C. Gadotti**

*School of Sciences - University of Minho.*

**Abstract:** In this talk, we establish sufficient conditions for the exponential stability of an equilibrium point of the following general system of infinite delay differential equations with impulses:

$$\begin{cases} \dot{x}_i(t) = -a_i(x_i(t)[b_i(x_i(t)) + f_i(t, x_t)], 0 \le t \ne t_k, \quad i = 1, 2, ..., n, \\ (x_i(t_k)) = x_i(t_k^+) - x_i(t_k^-) = I_{ik}(x_i(t_k^-)), k \in N. \end{cases}$$

The main result is used to give a stability criterion for a very broad family of impulsive neural network model with both unbounded distributed delays and bounded time-varying discrete delays with impulses. Most of the impulsive neural network models with delay recently studied are included in the general framework presented and we give a example which illustrate the effectiveness of the result.

## Stability and Morse decompositions of non-deterministic systems.

**Maria Joana Torres, P. Duarte**
*School of Sciences - University of Minho*

**Abstract:** We introduce and study, from a combinatorial-topological viewpoint, some semigroups of continuous non-deterministic dynamical systems. Combinatorial stability, i.e. the persistence of the combinatorics of the attractors, is characterized and its genericity established. Some work in progress to prove a version of Smale's Spectral Decomposition Theorem for smooth non-deterministic systems and to analyze their stability is addressed.

## Heterodimensional cycles: a model family.

**Salete Esteves**
*EsACT - Polytechnic Institute of Bragança.*

**Abstract:** In this talk, we exhibit a two-parameters model family $(f_{a,t})_{t\in[-\epsilon,\epsilon]}$ and $a > 0$, such that for each a the corresponding one-parameter family unfolds a heterodimensional cycle at $t = 0$ associated to two fixed saddles $P$ and $Q$ having different indices (dimension of the unstable manifold), for which the central dynamic is given by the map $g_a : [-1, 2] \to R$ defined by

$$g_a(y) = \frac{e^a y}{2e^a y + (1 - 2y)}$$

For that, depending on the two parameters, we are able to identify scaled intervals of hyperbolicity and of non-hyperbolicity, and furthermore describe the secondary bifurcations associated to the transition from hyperbolicity to non-hyperbolicity. Moreover these families have a variety of behaviors that reproduces the ones obtained in several papers of J. Rocha and L. J. Díaz.

# Fredholm characteristics of singular integral operators with a Carleman shift and almost periodic coefficients via factorization of matrix functions.

**Anabela Silva**
*CIDMA  Centro de Investigação e Desenvolvimento em Matemática e Aplicações - Universidade de Aveiro.*

**Abstract:** : The talk will be devoted to the study of singular integral operators with almost periodic coefficients and a Carleman shift. Our aim is to obtain an operator factorization of those operators, from which such Fredholm characteristics as the kernel and cokernel can be described. In particular, the dimensions of their kernels and cokernels are obtained. This is done by considering appropriate properties of the related almost periodic elements and, in special, the partial indices of some of their relevant factorizations.

# Operator viewpoint of general boundary value problems within wave diffraction.

**Luís Castro**
*CIDMA  Centro de Investigação e Desenvolvimento em Matemática e Aplicações - Universidade de Aveiro.*

**Abstract:** : We will use toplinear isomorphisms and different types of operator relations to analyse boundary value problems with general conditions which can be seen as models for some wave diffraction problems. The so-called convolution type operators with symmetry and around the corner operators will be very important to use in the interpretation of the problems in study. Conclusions about the normal solvability, Fredholm property and (lateral) invertibility of the operators will allow corresponding characterizations for the problems.

# Heat kernel for the index Whittaker transform

**M. Manuela Rodrigues**

*CIDMA Centro de Investigação e Desenvolvimento em Matemática e Aplicações - Universidade de Aveiro.*

**Abstract:** : We introduce a notion of the heat kernel associated with the index Whittaker transform

$$W_\mu[f](\tau) = \int_0^\infty x^{-2} W_{\mu,i\tau}(x) f(x) dx,$$

where $W_{\mu,i\tau}(x)$ is the Whittaker function, with $x, \tau \in R_+$, and $\mu < \frac{1}{2}$. We will talk about its differential and mapping properties and a relationship with a family of the corresponding Weierstrass's type transforms.

# Effective conductivity of a singularly perturbed periodic two-phase composite with imperfect thermal contact at the two-phase interface

**Matteo Dalla Riva**

*CIDMA Centro de Investigação e Desenvolvimento em Matemática e Aplicações - Universidade de Aveiro.*

**Abstract:** We consider the effective thermal conductivity of a two-phase composite with imperfect thermal contact at the two-phase interface. The composite is obtained by introducing into an infinite homogeneous matrix a periodic set of inclusions of a different material. The radius of each inclusion is assumed to be proportional to a positive real parameter r. Then we show that the function which describes the effective conductivity can be continued real analytically in the parameter r around the value r=0 (in correspondence of which the inclusions collapse to points). The methods developed are based on functional analysis and potential theory and are alternative to asymptotic analysis. Based on a joint work with P. Musolino (Padova, Italy).

# Integral Equation Methods in Problems of Reactance Wave Diffraction by a Strip

**Alberto M. Simões**
*Department of Mathematics, University of Beira Interior, Portugal, asimoes@ubi.pt*

**L.P. Castro**
*Department of Mathematics, University of Aveiro, Portugal*

**Abstract:** A problem of reactance wave diffraction by a strip characterized with higher order boundary conditions will be analyzed from an integral equations viewpoint. The problem will be formulated as a boundary-transmission problem for the Helmholtz equation, in a Bessel potential space setting. Integral equations and operator theoretical methods are used to deal with the problem and, as a consequence, several convolution type equations are constructed and associated to the problem. The invertibility of these operators is obtained for certain smoothness space orders and a consequent mathematical analysis of the initial problem is derived in view of its well-posedness.

ABSTRACTS

CONTRIBUTED PAPERS - ORAL

# Selection of variables defined on the hypersphere

**Adelaide Figueiredo[1], Paulo Gomes[2]**

[1]*School of Economics, University of Porto and LIAAD-INESC TEC, Portugal, adelaide@fep.up.pt,*
[2]*Statistics Portugal and Nova University of Lisbon, ISEGI, Portugal, paulo.gomes@ine.pt*

**Abstract:** We discuss the identification of a mixture of Watson distributions defined on the hypersphere using the dynamic clusters algorithm and the EM algorithm (Estimation-Maximization). The identification of this mixture allows us to obtain homogeneous groups of variables and provides an useful contribution to the selection of variables. Then, we compare these algorithms by a simulation study and with real data.

**Keywords:** Dynamic clusters algorithm, EM algoritm, Selection of variables, Variable clustering.

## References

Figueiredo, A. (2000). Classification of variables in the context of a probabilistic model defined on the hypersphere, PhD thesis. School of Sciences, University of Lisbon.

Figueiredo, A., Gomes, P. (2006). Performance of the EM algorithm on the identification of a mixture of Watson distributions defined on the hypersphere. REVSTAT-Statistical Journal, vol. 4, no. 2, pp. 19.

Gomes, P. (1987). Distribution de Bingham sur la n-sphere: une nouvelle approche de lanalyse factorielle, Thése de doctorat DÉtat, Université de Montpellier.

# Bootstrap control charts to monitor skew-normal processes

**Fernanda O. Figueiredo[1] ; M. Ivette Gomes[2]**

[1]*Faculty of Economics, Porto University (FEP.UP) and CEAUL, Portugal, otilia@fep.up.pt,*
[2]*FCUL, DEIO and CEAUL, University of Lisbon, Portugal, ivette.gomes@fc.ul.pt*

**Abstract:** Highly-skewed distributions, possibly with heavy tails, are often used to model data from diversified areas of application, and in particular in the area of reliability. In this work we consider some classes of skew-normal distributions and we enhance some of their properties in order to motivate their use in applications. Then we develop bootstrap control charts to monitor such processes.

**Keywords:** Heavy tails, skewness, bootstrap control charts.

## References

Azzalini A. (1985). A Class of distributions which includes the normal Ones. Scandinavian J. of Statistics 12, 171-178.

Fernandez C., and Steel M.F.J. (1998). On Bayesian modeling of fat tails and skewness. J. Am. Stat. Assoc. 93, 359-371.

# Further results on the extreme value index estimation: the maximum likelihood estimators of Feuerverger and Hall

**Frederico Caeiro**[1],   **M. Ivette Gomes**[2]

[1] *Faculdade de Ciências e Tecnologia & CMA – Universidade Nova de Lisboa, fac@fct.unl.pt,*
[2] *Faculdade de Ciências & CEAUL – Universidade de Lisboa, ivette.gomes@fc.ul.pt*

**Abstract:** In this work we are interested in the semi-parametric estimation of the extreme value index, one of the most important parameters in statistics of extremes. Under a general third order condition for heavy tails, we study the asymptotic distributional properties of the maximum likelihood estimators of the extreme value index and the two second order parameters in Feuerverger and Hall (1999). Applications to a real data set and to simulated data are also provided.

**Keywords:** Statistics of extremes, semi-parametric estimation, extreme value index, heavy tail, log-spacings.

# References

Feuerverger, A., and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. Annals Statistics 27, 760781.

# Strong and weak Allee effects, extinctions and bistability in Richards growth dynamics

**J. Leonel Rocha and Sandra M. Aleixo**

*Instituto Superior de Engenharia de Lisboa, Mathematics Unit and CEAUL Rua Conselheiro Emídio Navarro, 1,1959-007 Lisboa, Portugal e-mails: jrocha@adm.isel.pt and sandra.aleixo@adm.isel.pt*

**Abstract:** In this paper we define and investigate populational growth models, with strong and weak Allee effects, from Richards Equation. The use of a parameter leads the presented generalization, which yields some more flexible models with variable extinction rates. An Allee limit is incorporated so that the models under study have different type of Allee effects. The complexity of the correspondent discrete dynamical systems is studied using symbolic dynamics techniques. Several populational dynamics behaviours are obtained when the intrinsic growth rates are modified: extinction, bistability, chaotic semistability and essential extinction.

## References

Aleixo, S.M. and J.L. Rocha (2012). Generalized Models from Beta(p,2) Densities with Strong Allee Effect: Dynamical Approach. To appear in Proc. ITI 2012, 34th Int. Conf. Information Technology Interfaces.

Boukal, D.S. and L. Berec (2002). Single-species models of the Allee effect: extinction boundaries, sex ratios and mate encounters. J. Theor. Biol. 218, 375-394.

Brassil, C.E. (2001). Mean time to extinction of a metapopulation with an Allee effect. Ecol. Model. 143, 9-16.

Lewis, M.A. and P. Kareiva (1993). Allee dynamics and the spread of invading organisms. Theor. Popul. Biol. 43, 141-158.

Rocha, J.L. and S.M. Aleixo (2012). Modeling Allee effect from Beta(p,2) densities. To appear in Proc. ITI 2012, 34th Int. Conf. Information Technology Interfaces.

Schreiber, S.J. (2003). Allee effects, extinctions, and chaotic transients in simple population models. Theor. Popul. Biol. 64, 201-209.

Tsoularis, A. (2001). Analysis of logistic growth models. Res. Lett. Inf. Math. Sci. 2, 23-46.

Wang, M. and M. Kot (2001). Speeds of invasion in a model with strong or weak Allee effects. Math. Biosci. 171, 83-97.

# Some Inequalities for Functions with Variable Exponents

**Vadim Yurinsky**[1]

[1] *Universidade da Beira Interior, Department of Mathematics, Portugal, e-mail: yurinsky@ubi.pt*

**Abstract:** The abstract deals with some inequalities that prove to be useful in the study of physical processes in continuous media with variable exponents [1,2]. Its results are mainly estimates expressed in terms of modulars [3]. For the sake of convenience the notation of these latter is adapted to the form

$$\left\langle \varphi^k \psi^{\mathcal{P}_*} \right\rangle \leq A \left\langle \varphi^k \right\rangle^{\mu} \left\langle \varphi^k \psi^{\mathcal{P}} \right\rangle^{1-\mu}, \quad \left\langle \varphi^{\mathcal{Q}} \right\rangle \leq B \langle \phi^{\mathcal{R}} \rangle^{\kappa} \langle 1 \rangle^{1-\kappa}.$$

The exposition is motivated by the study of solution´s decay rate for differential equations containing variable exponents.

**Keywords:** Variable exponent 1, modular 2.

## References

Kovacik, O.; Rakosnik, Z.; Rakosnik, J. (1991). On spaces $L^{p(x)}$ and $W^{k,p(x)}$. *Czechoslovak Math. J.*, v. 41 (116), 592-618.

Antontsev, S.; Shmarev, S. (2006). Elliptic equations with anisotropic nonlinearity and nonstandard growth conditions. *Handbook of Differential Equations — Stationary Partial Differential Equations*, v.3, Elsevier, pp. 1-100.

Diening, L.; Harjulehto, P.; Hasto, P.; Rusicka, M. (2011). Lebesgue and Sobolev Spaces with Variable Exponents, Lect. Notes Math., v.2017, Springer, Heidelberg.

# THE WEAKLY ASYMMETRIC EXCLUSION PROCESS WITH OPEN BOUNDARIES

**Patrícia Gonçalves, Cláudio Landim and Aniura Milanes**

*CMAT, Universidade do Minho*

**Abstract:** In this talk I introduce the weakly asymmetric simple exclusion process in contact with reservoirs. In the bulk particles jump, after a mean one exponential time, to the right neighboring site with probability $\frac{1}{2} + \frac{E}{N}$ or to the left neighboring site with probability $\frac{1}{2} - \frac{E}{N}$, where $E > 0$ is a constant. At the left boundary particles can enter (resp. leave) the system with rate $\alpha$ (resp. $1 - \alpha$) with $\alpha \in (0,1)$. At the right boundary particles can enter (resp. leave) the system with rate $\beta$ (resp. $1 - \beta$) with $\beta \in (0,1)$. It is known that its hydrodynamics is governed by the viscous Burgers equation with Dirichlets boundary conditions, namely:

$$\begin{cases} \partial_t \rho = \partial_u \rho - E \partial_u \rho (1 - \rho), & \\ \rho(t,0) = \alpha, & \rho(t-1) = \beta, \quad 0 \leq T, \\ \rho(0,u) = \rho_0(u), & 0 \leq u \leq 1 \end{cases} \qquad (1)$$

We prove that starting the system from an non-equilibrium state the fluctuations are Gaussian and given by the Ornstein-Uhlenbeck process $Y_t$ in $C(\mathbb{R}_+, C_0^2([0,1]))$ solution of

$$dY_t = \partial_u^2 Y_t dt + E(1 - 2\rho(t,u))\partial u Y_t dt + \sqrt{2\chi(\rho(t,u))}\partial u dW_t, \qquad (2)$$

where $W_t$ is a space-time white noise of unit variance and $\chi(\rho) = \rho(1 - \rho)$ and $C_0^2([0,1])$ is the set of functions $G : [0,1] \to \mathbb{R}$ twice-continuously differentiable on $(0,1)$ that vanish at the boundary: $G(0) = G(1) = 0$.

.

ABSTRACTS

CONTRIBUTED PAPERS - POSTERS

# Online Water Quality Monitoring by using Kalman Filter Predictions

**A. Manuela Gonçalves**
*University of Minho, Department of Mathematics and Applications, CMAT-Centro de Matem´atica, Portugal, mneves@math.uminho.pt*

**Marco Costa**
*University of Aveiro, Higher School of Technology and Managment of Águeda, CMAF-Centro de Matemática e Aplicações Fundamentais, Portugal, marco@ua.pt*

**Abstract:** The forecast accuracy is not always the most important characteristic of a model. Indeed, in water quality monitoring it is useful to analyse temporal evolution in order to identify possible factors or changes. This analysis can be performed in two ways: trends analysis by looking into the historical data and dynamic or online procedure. The first approach can be achieved largely through linear regression models by analysing trends parameters and it allows diagnosing any global tendency that can be statistically significant. Indeed, the state-space approach adds a dynamic component to the usual linear models that can be useful in the water quality monitoring procedure. In this work the discussion focuses on the dynamic monitoring procedure based on the state-space approach (associated to the Kalman filter algorithm), which the formulation presented, and can be interpreted as a calibration model of seasonal coeficients. As it will be shown, this formulation will allow making some useful interpretations. Furthermore, the state-space model associated to the Kalman filter allows predicting one-step forecasts and the filtered prediction of the calibration factor, as a measure of the discrepancy between the observed value of the quality variable and the excepted value (the seasonality mean value), in each time. The statistical modelling procedure was applied to a set of water monitoring sites grouped in homogeneous clusters in the River Ave hydrological basin located in the Northwest region of Portugal, and its main objective is to obtain an accurate forecast of a quality variable concentration.

**Keywords:** Water quality, state-space models, Kalman filter, accurate forecast.

**References**

Costa, M, Gonçalves, A.M. (2012). Combining Statistical Methodologies in Water Quality Monitoring in a Hydrological Basin - Space and Time Approaches. In Water Quality Monitoring and Assessment, ed. Kostas Voudouris and Dimitra Voutsa, 121142. ISBN:978-953-51-0486-5. InTech Published.

Costa, M., Gonçalves, A.M. (2011). Clustering and forecasting of dissolved oxygen concentration on a river basin. Stochastic Environmental Research and Risk Assessment, v.25, n.2, pp. 151163.

# Variables' selection in Discrete Discriminant Analysis

**Anabela Marques[1],Ana Sousa Ferreira[2] and Margarida Cardoso[3]**

[1]*Escola Superior de Tecnologia do Barreiro-IPS, Portugal, anabela.marques@estbarreiro.ips.pt*
[2]*Faculdade de Psicologia, Universidade de Lisboa, CEAUL e UNIDE, Portugal, asferreira@fp.ul.pt,*
[3]*Escola de Gestão do ISCTE, Dep. Métodos Quantitativos e UNIDE, Portugal, margarida.cardoso@iscte.pt*

**Abstract:** In Discrete Discriminant Analysis (DDA), namely in the social sciences, humanities and health domains, we often have to deal with dimensionality problems since even a moderate number of explanatory variables leads to an enormous number of possible states/outcomes when compared to the number of objects under study (Brito et al, 2006, Sousa Ferreira, 2000). As a consequence, models used in classification may exhibit poor performance due to the large number of parameters to be estimated.
In the present paper, we discuss variables selection techniques to address the issue of dimensionality in DDA (Benjamini et al. (1995), Marques et al. (2010)). We specifically perform DDA using a combined model approach, Marques et al. (2008). Since diverse models tend to yield classification errors for different objects the proposed approach tends to produce more efficient and stable results. In this setting, the variables selection techniques are particularly pertinent conquering degrees of freedom and reducing computational cost.

**Keywords:** Combining models; Discrete Discriminant Analysis, Variable selection.

## References

Benjamini, Y. e Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological), Vol.57, 1, pp. 289-300.

Brito, I., Celeux, G. e Sousa Ferreira, A. (2006) Combining method in supervised classification: A comparative study on discrete and continuous problems. REVSTAT - Statistical Journal, Vol. 4(3), 201-225.

Marques, A. Sousa Ferreira, A. e Cardoso, M. (2008). Uma proposta de combinação de modelos em Analise Discriminante Discreta . Estatística - Arte de Explicar o Acaso, em Oliveira, I. et al. (Eds.),Ciência Estatística, Edições SPE, pp. 393-403.

Marques, A. Sousa Ferreira, A. e Cardoso, M. (2010). Seleção de Variáveis em Análise Discriminante Discreta. Livro de resumos do XVIII Congresso Anual da Sociedade Portuguesa de Estatística (SPE 20120), pp. 50-52.

Marques, A. Sousa Ferreira, A. e Cardoso, M. (2011). Features selection in Discrete Discriminant Analysis. Book of Abstracts in CD-rom, 14th Conference of the ASMDA - Applied Stochastic Models and Data Analysis (ASMDA2011).

Sousa Ferreita, A. (2000). Combinação de modelos em Análise Discriminante sobre variáveis qualitativas. Tese de Doutoramento. Universidade Nova de Lisboa.

# Aplicações de Técnicas Estatísticas e Probabilísticas no Desenvolvimento e Análise de Algoritmos de Optimização Combinatória na Área da Bioinformática

**Andreia A. Monteiro[1], Teresa A. Oliveira[2] and Angel A. Juan[3]**
[1] *MEMeC -Universidade Aberta, Portugal, andreiaforte@portugalmail.pt*
[2] *Universidade Aberta e CEAUL, Portugal, toliveir@uab.pt*
[3] *UOC - Barcelona, Espanha, ajuanp@gmail.com*

**Abstract:** Os Problemas de Optimização Combinatória aparecem quando é necessário seleccionar de um conjunto discreto e finito de dados, o melhor subconjunto que satisfaz determinados critérios. A dificuldade é que mesmo para problemas de dimensão reduzida, o número de alternativas possíveis (soluções admissíveis) pode ser demasiado elevado para poderem ser todas examinadas.

Neste trabalho é apresentada uma revisão de métodos deterministas e métodos estocásticos, nomeadamente algoritmos probabilísticos, metaheurísticas e hiper-heurísticas, na resolução dos Problemas de Optimização. As técnicas Estatísticas e Probabilísticas têm um papel importante quer no desenvolvimento quer depois na fase de análise dos algoritmos, dando assim um contributo fundamental para o desenvolvimento de métodos cada vez mais eficazes.

Nos últimos anos, têm surgido cada vez mais problemas de optimização de difícil resolução e nos mais diversos campos da ciência, como por exemplo na Bioinformática, nas sequências de DNA e proteínas e na inferência de padrões. Urge desenvolver técnicas que sejam mais gerais e por isso, aplicáveis aos mais diversos problemas, as metaheurísticas. Na construção de algoritmos probabilísticos, é crucial o recurso a técnicas probabilísticas e a metodologias como testes de aleatoriedade e ajustamento, métodos de reamostragem e planeamento de experiências. Como proposta de resolução de problemas de optimização apresenta-se uma abordagem ao método de simulação de Monte Carlo e à optimização via simulação.

**Keywords:** Optimização Combinatória, Técnicas Estatísticas e Probabilísticas, Metaheurísticas, Simulação de Monte Carlo, Optimização via Simulação, Bioinformática.

## References

Aiex, R., Resende, M. and Ribeiro, C. (2002) "Probability Distribution of Solution Time in GRASP: An Experimental Investigation", Journal of Heuristics, v.8: 343-373

Banks, J., Carson J., Nelson B., and Nicol D. (2005). Discrete-event system simulation. 4th ed. Upper Sadder River, New Jersey. Prentice-Hall.

Cowling, P., Kendall, G. and Soubeiga, E. (2000). A hyperheuristic Approach to Scheduling a Sales Summit. In Proceedings of the Third International Conference of Practice and Theory of Automated Timetabling (PATAT 2000). Lecture Notes in Computer Science 2079: 176-190. Springer-Verlag.

# Clustering of Symbolic Data based on Affinity Coefficient: Application to real data sets

**Áurea Sousa**
*University of Azores, Department of Mathematics, CEEAplA, CMATI, Portugal,*
*aurea@uac.pt*

**Helena Bacelar-Nicolau**
*University of Lisbon, Faculty of Psychology, Laboratory of Statistics and Data Analysis, Portugal, hbacelar@fp.ul.pt*

**Fernando C. Nicolau**
*New University of Lisbon, FCT, Department of Mathematics, Portugal, geral@datascience.org*

**Osvaldo Silva**
*University of Azores, Department of Mathematics, CMATI, Portugal, osilva@uac.pt*

**Abstract:** The increasing use of databases, often large ones, in diverse areas of study makes it pertinent to summarise data in terms of their most relevant concepts. These concepts may be described by types of complex data, also known as symbolic data. In a symbolic data table, lines correspond to symbolic objects (SO's) and columns to symbolic variables, which may contain not just one value, as usual, but values, such as subsets of categories, intervals in real axes, or frequency distributions. Symbolic data arise in a number of different ways (for example, as the result of aggregation of large data sets to obtain a data set of manageable size, or as a result of some scientific question(s) of interest). The aim of Symbolic Data Analysis (SDA) is to extend classical data analysis techniques to these kinds of data.

Based on the affinity coefficient between two discrete probability distributions as defined by the pioneer work of Matusita, started in 1951, Bacelar-Nicolau suggested the use of the affinity coefficient as a basic similarity coefficient between the columns or the lines of a data matrix. Later on she extended that coefficient to different types of data, including complex data (symbolic data) and variables of mixed types (heterogeneous data), possibly with different weights.

We present some results from the Ascendant Hierarchical Cluster Analysis (AHCA) of symbolic objects described by interval data, in order to illustrate the effectiveness of the Ascendent Hierarchical Cluster Analysis based on the weighted generalized affinity coefficient, for symbolic data. The measure of comparison between the elements was combined with classical aggregation criteria and probabilistic ones. The probabilistic aggregation criteria used in this study belong to a parametric family of methods in the scope of the probabilistic approach of AHCA, named VL methodology and the validation of the clustering results is based on some validation measures. Finally, we compare the results achieved by our approach with the ones obtained by other authors.

**Keywords:**Ascendant Hierarchical Cluster Analysis, Symbolic Data, Interval Data, Affinity Coefficient, VL Methodology.

## References

Bacelar-Nicolau, H. (1988). Two Probabilistic Models for Classification of Variables in Frequency Tables. In: Classification and Related Methods of Data Analysis, H. -H. Bock (ed.), North Holland: Elsevier Publishers B.V, 181-186.

Bacelar-Nicolau, H. (2000). The Affinity Coefficient for Complex Data. In: Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information from Complex Data, H. -H. Bock, E. Diday (Eds.), Studies in Classification, Data Analysis, and Knowledge Organization, Berlin: Springer-Verlag, 160-165.

Bacelar-Nicolau, H. (2002). On the Generalised Affinity Coefficient for Complex Data. Biocybernetics and Biomedical Engineering, v. 22, n.1, 31-42.

Bacelar-Nicolau, H., Nicolau, F., Sousa, A., Bacelar-Nicolau, L. (2009). Measuring similarity of complex and heterogeneous data in clustering of large data sets. Biocybernetics and Biomedical Engineering, v. 29, n. 2, 9-18, 2009.

Bock, H. -H., Diday, E. (Eds.) (2000). Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information from Complex Data, Series: Studies in Classification, Data Analysis, and Knowledge Organization, Berlin: Springer-Verlag.

Matusita, K. (1951). On the Theory of Statistical Decision Functions. Ann. Instit. Stat. Math., v.III,1-30.

Nicolau, F. (1980). Critérios de Análise Classificatória Hierárquica baseados na Função de Distribuição. Tese de Doutoramento, FCL, Universidade de Lisboa.

Nicolau, F., Bacelar-Nicolau, H. (1998). Some Trends in the Classification of Variables. In: Data Science, Classification, and Related Methods, C. Hayashi, N. Ohsumi, K. Yajima, Y. Tanaka, H. -H. Bock, Y. Baba (Eds.), Springer-Verlag, 89-98.

Nicolau, F., Bacelar-Nicolau, H. (1999). Clustering Symbolic Objects Associated to Frequency or Probability Laws by the Weighted Affinity Coefficient. In: Applied Stochastic Models and Data Analysis. Quantitative Methods in Business and Industry Society, H. Bacelar-Nicolau, F. C. Nicolau and Jacques Janssen (Eds.), INE, Lisbon, Portugal, 155-158.

Sousa, A. (2005). Contribuições à Metodologia VL e Índices de Validação para Dados de Natureza Complexa. Tese de Doutoramento, Universidade dos Açores, Ponta Delgada.

Sousa, A., Nicolau, F., Bacelar-Nicolau, H., Silva, O. (2010). Weighted Generalised Affinity Coefficient in Cluster Analysis of Complex Data of the Interval Type. Biometrical Letters, v. 47, n. 1, 45-56.

# Blood Groups Gene Frequencies

**Carla Noronha**
*Universidade Aberta, MBB , Portugal, Carla.noronha6@gmail.com*

**Ana Cunha**
*Universidade Aberta, MBB , Portugal, anagoncunha@gmail.com*

**Ulisses Azeiteiro**
*Universidade Aberta, DCeT, Portugal, ulisses@univ-ab.pt*

**Abstract:** At a time of unprecedented expansion, the study of evolution transcends all areas of biology. All phenomena that are observable today in living organisms and their systems have a fantastic evolutionary component that can not be ignored! Evolution transcends mutations and natural selection of Darwin and proves to be in the first row of studying the mechanisms involved ... Life! Mathematical models have become allies in the biosciences. The probabilistic models become very useful for process analysis as it is possible to access the repetitive data, since they require a large initial sample data for forming, but the implementation of the evolutionary processes is often associated to the scarcity of data. Changes in allele frequencies over time and space are clues from which you can investigate the action of evolutionary processes. By understanding how "evolutionary forces" acting is also possible to construct mathematical models that approximate reality. These models are needed to understand the subtle interplay between "evolutionary forces" and genetic heritage, and from this knowledge, make inferences about the processes that acted in the past and which are the source of genetic diversity present.
The study of gene frequencies of blood groups is one of the areas where the mathematical models and probabilistic / stochastic theories stand for the progress achieved in recent decades.
In this work authors present and discuss such models.

**Keywords:** Fenotype, genotype, frequencies,evolution.

## References

Nowak, Martin A.,Evolutionary Dynamics, Exploring the Equations of Life, 2006,Cambridge, Massachusetts and London, England

# Crossover Designs: Revisão e Desafios

**Carla Noronha**
*Universidade Aberta, MBB , Portugal, Carla.noronha6@gmail.com*

**Teresa Oliveira**
*Universidade Aberta, CEAUL, toliveir@uab.pt*

**Abstract:** Na área do Planeamento de Experiências, que tem sido amplamente desenvolvida desde os trabalhos primeiros de Fisher nos anos 20 do séc. XX, são de primordial importância os estudos na área dos Crossover Designs, sobretudo pelas suas pertinentes aplicações em áreas como a Medicina, Ciências Farmacêuticas e as Ciências da Vida em geral. Neste tipo de planeamento, cada unidade experimental recebe diferentes tratamentos em diferentes períodos de tempo, o que contrasta com os estudos em paralelo em que a cada unidade é atribuído um tratamento e a ele permanece fiel até ao final do estudo. Neste trabalho investigamos os Crossover Designs, partindo de uma resenha histórica e abordando a metodologia usual, dando ênfase às áreas de aplicação mais relevantes e procurando ilustração em aplicações a dados reais.

**Keywords:** Crossover design, planeamento de experiências, drop-out, carry-over.

## References

Afsarinejad, K & Hedayat, A.S (2002): Repeated Measurements Designs for a model with Self and Simple Mixed Carryover Effects. Journal of Statistical Planning and Inference, 106(1-2), 449-459

# Isomorphism of Markov Chains: Application to a Customer Population

**Cláudia Roçadas**

*Center of Mathematics and Applications, Faculty of Sciences and Technology, Department of Mathematics, New University of Lisbon, claudia.rocadas@gmail.com*

**Teresa A. Oliveira**

*UAb-Universidade Aberta, Department of Sciences and Technology, CEAUL-Center of Statistics and Applications, University of Lisbon, toliveir@univ-ab.pt*

**João Tiago Mexia**

*Center of Mathematics and Applications, Faculty of Sciences and Technology, Department of Mathematics, New University of Lisbon, jtm@fct.unl.pt*

**Abstract:** Our research is centred on the stochastic structure of matched open populations, subjected to periodical reclassifications. These populations are divided into sub-populations. Two or more of such population are matched when there is a 1-1 correspondence between their sub-populations and the elements of one of them can go to another, if and only if the same occurs with elements from the corresponding sub-populations of the other.

The central objective of this investigation is the comparative study of paired open populations subject to periodic reclassifications, which led us to introduce the notion of isomorphism of Markov chains. These populations are divided into sub-populations in which the elements are placed. There will be pairing if there is a bijection between the sets of sub-populations and when the elements of one of those people can move between sub-populations, if and only if it is found for the corresponding sub-populations of the other population. Indeed, the presence of pairing involve pairing of sub-populations and consequently the state, must comply with this pairing relationships access. This notion is based on a finding pairing between populations. Considering this pairing we will study the action of one or more factors case populations corresponding to combinations of levels of these factors. In our case there is only a single factor at two levels: customers with and without account manager. The pairing allows us to analyze the effect of having an account manager on various aspects we consider relevant. We shall assume that the reclassifications are made in the early periods together with the classification of new elements. Thus Markov chains with discrete parameter, appear as the mathematical model for the study of these populations.

In our application we consider two populations of customers of a bank: with and without account manager. Besides these study connected with Markov chains we show how to carry out Analysis of Variance - like analysis of entries and departures to and from de populations of customers. This study was useful since it enabled as to length the model.

**Keywords:** Populations with periodic reclassification, isomorphism, Markov chains, limit distributions.

# References

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

Mexia, J.T.; Guerreiro, R. (2004), An Alternative Approach to Bonus Malus Discussiones Mathematicas, Probability and Statistics, 24, 197-213.

Norris, J.R.; Gill, R.; Ripley, B. D. and Ross S. (1998): Markov Chains, Cambridge Series in Statistical & Probabilistic Mathematics.

Norris, J. R. (1997): Markov Chains, Cambridge Series in Statistical & Probabilistic Mathematics.

Stewart, N. Ethier; Kurtz, Thomas G. (1986): Markov processes, New York, John Wiley.

# Process Capability Analysis: an application to the orders management

**Conceição Leal**

*Universidade Aberta, Departamento de Ciências e Tecnologia, Estudante do Mestrado em Estatística, Matemática e Computação, Secção de Matemática, Portugal, conceicao.leal2010@gmail.com*

**Amílcar Oliveira**

*Universidade Aberta and CEAUL, aoliveira@univ-ab.pt*

**Abstract:** Beyond the analysis of the process variability, the Process Statistical Control comprehends the analysis of process capability in responding to the established specifications. The process capability analysis may help in the quantification of its variability, in the analysis of this variability in relation to the specifications and in the definition of actions leading to the elimination or to the effective reduction of that variability. This work shows a set of statistical tools used to evaluate the capacity of a process. It was applied in analyzing the capacity of the daily prevision to correspond to the demand for bread in a school bar. The example illustrates the application of this step of the statistical control in a non-productive process.

**Keywords:** Process Capability, Process Capability Analysis, Orders Management.

## References

BOX, Hunter, Hunter (2005): Statistics for Experimenters, 2th Edition, Wiley and Sons, Inc.

CHANDRA, M.J. (2001): Statistical Quality Control, CRC Press LLC

KANE, V. E. (1986): Process Capability Indices. Journal of Quality Technology, Vol.18, No 1, pp. 41-52. January.

MITCHELL, Robert H., (1999): Process Capability Indices, ASQ Statistics Division Newsletter, Vol. 18, No. 1

MONTGOMERY, D. C. (2003, 2005): Introduction to Quality Control, 3th and 5th Editions, New York:Chapman & Hall.

OAKLAND, John S.(2003): Statistical Process Control, 5th Ed. Butterworth-Heinemann

SCRUCCA, Luca (2004): qcc: An R package for quality control charting and statistical process control, R News, Vol 4/1, Junho, URL http://www.stat.unipg.it/luca/Rnews_2004-11pag.11-17.pdf

# Models of Symmetric Stochastic Matrices

**Cristina Dias**
*Instituto Politécnico de Portalegre, Portugal, cpsd@estgp.pt.*

**Manuela M. Oliveira**
*CMA-Centro de Matemática e Aplicações, Universidade de Évora, Portugal,mmo@uevora.pt*

**João Tiago Mexia**
*CMA-Centro de Matemática e Aplicações, FCT-Universidade Nova de Lisboa, Portugal*

**Abstract:** Let $\mathbf{M}$ be a symmetric stochastic matrix with mean matrix $\mu$ then

$$\mathbf{M} = \sum_{j=1}^{k} \lambda_i \alpha_i \alpha_i^t + \overline{\mathbf{E}}$$

With $\mathbf{M} = \sum_{j=1}^{k} \lambda_i \alpha_i \alpha_i^t$ the spectral decomposition of the mean matrix of $\mathbf{M}$. We consider the adjustment of the structure vector

$$\beta_i = \lambda_i^{1/2} \alpha_i, i = 1, \ldots, k$$

and model validation. Moreover we consider the especial case of cross product matrices $\mathbf{X}\mathbf{X}^t$ and $\mathbf{X}^t\mathbf{X}$ availing ourselves of the identity of non null eigenvalues of the matrices.

**Keywords:** Stochastic matrices, Symmetric matrices.

## References

Areia A., Oliveira M. M., Mexia, J. T. (2008). Models for a series of studies based on geometrical representation. Statistical Methodology, Vol. 5, N.3, pp. 277288.

Oliveira M. M., Mexia J.(1998). Tests for the rank of Hilbert-Schmidt product matrices. Advances in Data Science and Classifications. (Rizzi, Vichi and Bock, Ed.), 619-625. Springer.

Schott J. R. (1997). Matrix Analysis for Statistics. Wiley Series in Probability and Statistics.

# Modelo de regressão logística multivariada - Uma ferramenta na área da medicina

**Maria Cristina Ferreira**
*MEMC- Universidade Aberta, Portugal, ferreiramcmail@gmail.com*

**Teresa Oliveira**
*CEAUL and Universidade Aberta, Portugal, toliveir@univ-ab.pt*

**J.A. Lobo Pereira**
*FMDUP e MBB da Universidade Aberta, lobopereiramail@gmail.com*

**Abstract:** A análise de regressão é uma técnica estatística utilizada para investigar e modelar a relação entre variáveis, sendo uma das mais utilizadas na análise de dados, com inúmeras aplicações na área da saúde. Neste trabalho procuramos estudar o impacto dos fatores sócio demográficos, médicos e comportamentais na saúde periodontal recorrendo à análise estatística multivariada.

Através dum modelo de regressão logística multivariada foram calculados os Odds Racio e os intervalos de confiança a 95% da extensão da destruição dos tecidos periodontais. Foi considerado o nível de aderência (NA) igual ou superior a 4mm em 50% dos sítios sondados, e as variáveis, diabetes, sexo, idade, escolaridade, tabaco, uso de prótese dentária, índice de massa corporal (IMC), lipoproteínas de alta densidade (HDL) e hemoglobina glicada (A1c).

O nosso modelo de regressão logística multivariada permitiu-nos relacionar a ocorrência de doença periodontal extensa com os níveis de HDL.

**Keywords:** Regressão Logística, Doença periodontal, Nível de aderência (NA) Lipoproteínas de alta densidade (HDL).

## References

Coelho-Barros, E. A., Simões, P. A., Achcar J. A., Martinez E. Z., Shimano, A. C.. (2008). Métodos de estimação em regressão linear múltipla: aplicação a dados clínicos. Revista Colombiana de Estadística ( 31), 111- 129

Maroco, João. (2007). Análise Estatística com utilização do SPSS. 3ª Ed., Edições Sílabo.

Menard, Scott (2002). Applied logistic regression analysis, 2nd Edition.

Montgomery, D. C.(2001). Design and Analysis of Experiments, 5th Ed, Wiley.

Allison, Paul D. (1999). Multiple Regression - A primer. Pine Forge Press.

Scott Menard. (2002). Applied Logistic Regression Analysis, 2nd Edition, Sage Publications, ISBN 0-7619-2208-3

# Álgebras de Jordan Comutativas e Modelos Ortogonais

**Dina Salvador**

*Escola Superior de Tecnologia de Setúbal - Instituto Politécnico de Setúbal, Estefanilha 2910-761 Setúbal, Portugal, e-mail: dina.salvador@estsetubal.ips.pt*

**Sandra Monteiro**

*Escola Superior de Ciências Empresariais - Instituto Politécnico de Setúbal, Estefanilha 2914-503 Setúbal, Portugal, e-mail: sandra.ines.monteiro@esce.ips.pt*

**João Tiago Mexia**

*Faculdade de Ciências e Tecnologia - Universidade Nova de Lisboa, Monte da Caparica 2829-516 Caparica, Portugal, e-mail: jtm@fct.unl.pt*

**Abstract:** Neste poster iremos apresentar os principais resultados associados às Álgebras de Jordan Comutativas bem como uma aplicação das Álgebras de Jordan Comutativas aos modelos ortogonais. As Álgebras de Jordan Comutativas de matrizes simétricas, são muito úteis para exprimir a estrutura desses tipos de modelos.

**Keywords:** Álgebras de Jordan Comutativas, Modelos Ortogonais.

## References

Covas,R.; Orthogonal mixed models and commutative Jordan algebras. Dissertação de Doutoramento em Matemática. Especialidade Estatística; Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa; 2007.

Fonseca, M.; Mexia, J. T. ; Zmyslony, R.; Binary Operations on Jordan Algebras and Orthogonal Normal Models; Linear Algebra and its Applications; 2006.

Rao, C. R.; Linear Statistical Inference and Its Applications. Second Edition; John Willey & Sons; New York; 1973. Schott, J. R.; Matrix Analysis for Statistics; Wiley - Interscience; 1997.

Seely, J. ; Quadratic Subspaces and Completeness. Ann. Math. Stat.; 42(2): 710-721; 1971.

# ANÁLISE ESTATÍSTICA DA EVOLUÇÃO DA PERFORMANCE MOTORA NA FLEXIBILIDADE AO LONGO DO ENSINO SECUNDÁRIO - COMPARAÇÃO RAPAZES VS RAPARIGAS

**Domingos J. Lopes da Silva**
*Escola Secundária/3 de Barcelinhos (Barcelos), MEMC/UAb, domingosjl-silva@gmail.com*

**Amílcar Oliveira**
*Universidade Aberta e CEAUL, aoliveira@uab.pt*

**Abstract:** Este estudo tem como propósito avaliar a evolução da flexibilidade ao longo do ciclo do ensino secundário (comparação intra-género), bem como comparar rapazes com raparigas do mesmo ano de escolaridade (comparação inter-géneros). A amostra é formada por um total de 692 alunos do ensino secundário (RAPAZES: 10º ano, n=35; 11º ano, n=44; 12º ano, n=35; RAPARIGAS: 10º ano, n=70, 11º ano, n=85; 12º ano, n=64), da Escola Secundária de Barcelinhos (Barcelos), que frequentaram assiduamente as aulas de Educação Física (EF), e que em cada ano de escolaridade realizaram os dois momentos de avaliação (M1: Setembro; M2: Junho). Entre cada momento foram leccionadas cerca de 66 aulas de EF. A flexibilidade foi avaliada pelo teste sit-and-reach. Os procedimentos estatísticos utilizados incluíram a média, desvio-padrão, valores mínimo e máximo, e erro de estimativa amostral. A dispersão foi, ainda, apreciada pelo coeficiente de variação. A normalidade das distribuições foi calculada pelo teste de Kolmogorov-Smirnov, com correcção de Lilliefors, ou pelo teste de Shapiro-Wilk. A homogeneidade das variâncias foi verificada pelo teste de Levene. O estudo comparativo intra-género e inter-momentos (M1 vs M2) foi realizado pelo teste t de medidas emparelhadas. O estudo comparativo rapazes vs raparigas, em cada ano de escolaridade e momento de observação, foi realizado pelo teste t de medidas independentes. Dentro de cada género, o estudo comparativo inter-anos de escolaridade (10º vs 11º vs 12º) foi efectuado pela one-way ANOVA. A comparação múltipla foi realizada pelo post-hoc de Bonferroni. Adoptou-se um nível de significância estatístico de 5% (p 0,05). Todos os cálculos foram efectuados no R 2.14, no editor RStudio. Em termos globais, pode-se afirmar que: i) as raparigas no início do ensino secundário e os rapazes na fase intermédia, evidenciam melhorias na flexibilidade, ii) as raparigas comparativamente aos rapazes, apenas apresentam melhores performances na flexibilidade no início do ano lectivo no 10º ano de escolaridade. Nos restantes anos e fases do ano lectivo, os desempenhos são equivalentes, iii) apenas os rapazes, melhoram substancialmente a performance no ensino secundário, sendo que a evolução ocorre na passagem do 10º para o 11º ano de escolaridade., iv) as raparigas não modificam o seu desempenho na flexibilidade com o tempo. Our individual classifers, stumps, are expected to present a large bias, due to their simplicity. Boosting is a bias reduction technique and starts by building in the frst iteration a stump to predict if a certain example belongs to class C1 (in our application ascorbic acid) or not.

Those observations which were not correctly classifed by this stump will have a larger weight in the next iteration; the opposite happens with the ones correctly classifed. In the end we have a set of M models (stumps in our case) and a weighted linear combination of them is found as the fnal model. The ada package (Culp & al. 2006) implements the original AdaBoost algorithm (Freund and Schapire 1996, 1997) with other extensions. Some important features incorporated in this package include the use of both regression and classifcation trees for boosting and also various useful plots that aid in assessing variable importance and relationships between subsets of variables. This last feature is particularly interesting for us because, as we have thousands of variables (genes), we want not only an accurate fnal prediction model but mainly a way of choosing amongst those thousands of genes, the ones that matter. The motivation for this work is to identify the genes involved in cell proliferation due to ascorbic acid (AA) and its stable form ascorbic acid 2-phosphate (AA2P). Ascorbic acid, also known as vitamin C, is a sugar acid with antioxidant properties. ROS (reactive oxygen species that appear due to lesions and cytotoxic molecules) which contain unpaired electrons may interact with nucleic acids, proteins or lipids destroying them.

# Modelling of symptoms in patients with pollinosis in the Alto Alentejo region

**Fernando Afonso**
*CIIEM, Centro de Investigao Interdisciplinar Egas Moniz, Portugal, d8718@alunos.uevora.pt*

**Elsa Guerra Caeiro**
*Sociedade Portuguesa Alergologia Imunologia, Portugal*

**Rui Brandão**
*ICAM, Inst. Ciencias Agr. Ambientais Mediterrânicas, Universidade de Évora, Portugal, ruibrand@uevora.pt*

**Maria Luísa Lopes**
*Hospital Espírito Santo de Évora*

**Luis Ramos**
*CMA-Centro de Matemática e Aplicações, FCT-Universidade Nova de Lisboa, Portugal*

**Manuela M. Oliveira**
*CMA-Centro de Matemática e Aplicações, Universidade de Évora, Portugal, mmo@uevora.pt*

**Abstract:** The symptoms of pollen allergy by patients with pollinosis are the target of the study presented. Patients with a diagnosis of asthma, rhinitis or asthma/rhinitis are from the region of Évora. The symptoms manifested were collected daily and report to the years 2001 - 2007 for the months of March to June. For data collection a 0 - 3 scale was used, 0 being no symptoms and 3 corresponding to severe symptoms, this scale was applied to a set of nine typical symptoms of asthma, rhinitis and conjunctivitis. In the present study, the pollen data are from the collection station of Évora and were analyzed by the laboratory of Palynology, Department of Biology, University of Évora and relate to 17 species of pollen. Based on the acquired data the independence between the various symptoms and variables such as sex, age, diagnosis, months was tested, using chi-square test. The odds ratio (OR) between the degree of severity and sex was determined for each symptom. The Spearman correlation was used to establish the degree of association between symptom severity and concentration of pollen. The original categorical variable, degree of severity of symptoms, measured on a scale 0 - 3 was transformed into a quantitative variable score. The values of the new score variable are shown having a background of the various pollen species concentration over time. Binomial and multinomial models are under development for predicting the presence/absence or the degree of severity of symptoms based on the variables studied, i.e. age, diagnosis, month, sex, concentration of pollen.

**Keywords:** Binomial, Multinomial, Odds Ratio.

**References**

Thompson LA, R (and S-Plus) Manual to Accompany Agrestis Categorical Data Analysis (2002) 2nd edition

D. Clayton & M. Hills, (1993) Statistical Models in Epidemiology, Oxford University Press

Hosmer, D. & Lemeshow, S., Applied Logistic Regression, 2000

# Estimating Biological Age Through Dental and Skeletal Indicators

**Marina Andrade**

*Instituto Universitário de Lisboa (ISCTE-IUL), BRU-IUL, Lisboa, Portugal,*
*marina.andrade@iscte.pt*

**Manuel Alberto M. Ferreira**

*Instituto Universitário de Lisboa (ISCTE-IUL), BRU-IUL, Lisboa, Portugal,*
*manuel.ferreira@iscte.pt*

**Abstract:** An important and interesting problem which has become to raise interest in the last years is the problem of age estimation in living individuals. The more important cases are related with age estimations in emigration cases and legal age estimation in criminal proceedings. Of these, the first situation mentioned demands much greater precision than the second. Therefore the most used and important methods, to estimate the age in human people, are presented and discussed in this work. The respective properties and problems are summarized. A brief statistical discussion, concerning its accuracy, is performed.

**Keywords:** Dental age, skeletal age, correlation, confidence regions.

## References

Niemeijer, M., Ginneken, B. van, Maas, C., Beek, F. J. A., and Viergever, M. A. (2003). Assessing the Skeletal Age From a Hand Radiograph: Automating the Tanner-Whitehouse Method. In SPIE Medical Imaging. M. Sonka, J.M. Fitzpatrick, Editors, SPIE, 5032, 1197-1205.

Mincer, H. H., Harris, E. F. and Berryman, H. E. (1993). The A.B.F.O. study of third molar development and its use as an estimator of chronological age. Journal of Forensic Science, 38 (2), 379-390.

Aykroyd, R. G., Lucy, D., Pollard, A. M. and Solheim, T. (1997). Regression analysis in adult a age estimation. American Journal of Physical Anthropology, 104 (2), 259-265.

Demish, A. and Wartman, P. (1956). Calcification of the mandibular third molar and its relation to skeletal and chronological age in children. Child Development, 27, 459-473.

# Global Approach for the comparison of Clustering Results

**Osvaldo Silva**
*University of Azores, Department of Mathematics, CMATI, Portugal, osilva@uac.pt*

**Helena Bacelar-Nicolau**
*University of Lisbon, FP - Laboratory of Statistics and Data Analysis, Portugal, hbacelar@fp.ul.pt*

**Fernando Nicolau**
*New University of Lisbon, Faculty of Science and Technology, Portugal, geral@datascience.org*

**Abstract:** The extraction of useful knowledge from a Hierarchical Cluster Analysis (HCA) is a complex process which depends on many factors, such as the applied clustering algorithms and the strategies developed in the initial stage of the HCA. We present a global approach for evaluating the quality of clustering results based on the comparison of partitions from the different clustering algorithms using the most relevant information available (e.g. stability, isolation and homogeneity of the clusters). In addition, we suggest a visual method to facilitate the evaluation of the quality of the partitions that allows us a quick perception of the similarities and the differences between the partitions, including the behaviour of the elements in the partitions. We illustrate our approach using a real data set (horse data). We applied HCA based on the weighted generalized affinity coefficient (similarity coefficient) to the case of complex data (symbolic data), combined with 26 clustering (classic and probabilistic) algorithms. Finally, we discuss the obtained results and the contribution of this approach to a better knowledge on the cluster structure of a data set.
**Keywords:** Cluster Analysis, VL Methodology, Affinity Coefficient, Comparing Partitions, Cluster Stability and Cluster Validation.

## References

Bacelar-Nicolau, H. (2000). The Affinity Coefficient. In: Analysis of Symbolic Data Exploratory Methods for Extracting Statistical Information from Complex Data, H.H. Bock, E. Diday (Eds.), Springer, 160-165.

Bacelar-Nicolau, H. (2002). On the Generalised Affinity Coefficient for Complex Data. Biocybernetics and Biomedical Engineering, Vol. 22, nº 1, 31-42.

Bacelar-Nicolau, H.; Nicolau, F.; Sousa, A.; Bacelar-Nicolau, L. (2009). Measuring similarity of complex and heterogeneous data in clustering of large data sets. Biocybernetics and Biomedical Engineering, Vol. 29, nº 2: 9-18.

Bacelar-Nicolau, H.; Nicolau, F.; Sousa, A.; Bacelar-Nicolau, L. (2010). Clustering Complex Heterogeneous Data Using a Probabilistic Approach. Proceedings of Stochastic Modeling Techniques and Data Analysis International Conference (SMTDA2010), Chania Crete Greece, 8-11 June 2010 - published on the CD Proceedings of SMTDA2010 (electronic publication).

Carvalho, F.; Souza, R. (2009). Unsupervised pattern recognition models for mixed feature-type symbolic data. Pattern Recognition Letters, Volume 31, nº5, 430-443.

Gordon, A D. (1999). Classification, 2nd.Chapman & Hall, London.

Lerman, I. C. (1981). Classification et Analyse Ordinale des Données. Paris, Dunod.

Nicolau, F. (1983). Cluster Analysis and distribution Function in Meth. Oper. Res., 45, 431-433.

Nicolau, F.; Bacelar-Nicolau, H. (1999). Clustering Symbolic Objects Associated to Frequency or Probability Laws by the Weighted Affinity Coefficient. In: Applied Stochastic Models and Data Analysis. Quantitative Methods in Business and Industry Society, H. Bacelar-Nicolau, F. Nicolau and Jacques Janssen (Eds.), INE, Lisboa, Portugal, 155-158.

Silva O.; Bacelar-Nicolau, H.; Nicolau, F. (2009). Como Avaliar a Consistência dos Resultados de uma Análise Classificatória Hierárquica. In: Oliveira, I. et al. Eds. Actas do XVI Congresso Anual da Sociedade Portuguesa de Estatística, 2008, Edições S.P.E., 661-672.

Silva, O.; Bacelar-Nicolau, H.; Nicolau, F. (2010). Global Approach for Evaluating the Quality of Clustering Results. In: Programme and Abstracts CFE 10 & ERCIM 10 (4th CSDA International Conference on Computational and Financial Econometrics and 3rd Conference of the ERCIM Working Group on Computing and Statistics), 40.

Silva, O. (2012). Contributos para a Avaliação e Comparação de Partições em Análise Classificatória. Tese de Doutoramento, Universidade dos Açores, Ponta Delgada.

# An Introduction to Item Response Theory

**Sandra Nunes**
*Escola Superior de Ciências Empresariais do Instituto Politécnico de Setúbal, Departamento de Economia e Gestão, Portugal, sandra.nunes@esce.ips.pt*

**Helena Penalva**
*Escola Superior de Ciências Empresariais do Instituto Politécnico de Setúbal, Departamento de Economia e Gestão, Portugal, helena.penalva@esce.ips.pt*

**Abstract:** The Item Response Theory (IRT) has become one of the most popular scoring frameworks for measurement data, frequently used in computerized adaptive testing, cognitively diagnostic assessment and test equating. The Item Response Theory can be defined (Andrade et al, 2000) as a set of mathematical models (Item Response Models - IRM) constructed to represent the probability of an individual giving the right answer to an item of a particular test.
The number of Item Responsible Models available to measurement analysis has increased considerably in the last fifteen years due to increasing computer power and due to a demand for accuracy and more meaningful inferences grounded in complex data. The developments in modeling with Item Response Theory were related with developments in estimation theory, most remarkably Bayesian estimation with Markov chain Monte Carlo algorithms (Patz and Junker, 1999). The popularity of Item Response Theory has also implied numerous overviews in books and journals, and many connections between IRT and other statistical estimation procedures such as factor analysis and structural equation modeling have been made repeatedly (Van der Lindem and Hambleton, 1997).
As stated before the Item Response Theory covers a variety of measurement models, ranging from basic one-dimensional models for dichotomously and polytomously scored items and their multidimensional analogues to models that incorporate information about cognitive sub-processes that influence the overall itemresponse process. The aim of this work is to introduce the main concepts associated with one-dimensional models of Item Response Theory, specify the logistic models with one, two and three parameters, discuss some properties of these models and present the main estimation procedures. In a future work we intent to build a performance scale for the discipline of Mathematics taught in courses of College of Business Administration of the Polytechnic Institute of Setubal.

**Keywords:** Item Response Theory (IRT), Item Response Models (IRM), Evaluation tool.

## References

Andrade, D.F., Tavares, H.R., and Valle, R.C. (2000). Teoria da resposta ao Item: Conceitos e Aplicações. Associação Brasileira de Estatística (SINAPE).

Baker, F.B. (2001). The Basics of Item Response Theory (2nd edition). ERIC Clearinghouse on Assessment and Evaluation, USA.

Hambleton, R.K., Swaminathan, H., and Rogers, H.J. (1991). Fundamentals of Item Response Theory. Sage Publications Inc.

Patz, R.J., and Jumker, B.W. (1999). A straightforward approach to markov chain Monte carlo methods for item response models. Journal of Educational and Behavioral Statistics, 24, 146-178.

Rupp, A.A. (2003). Item Response Modelling with BILOG_MG and MULTILOG for Windows. International Journal of Testing, 3(4), 365-384.

Van der Linden, W.J. and Hambleton, R.K. (1997). Handbook of Modern Item Response Theory. Springer.

**(Included in Organized Session 3)**

# A Comparison of HJ-Biplot and Tucker3 Model: an Application in Environmental Data

Susana Mendes[1], M. Purificación Galindo-Villardón[2], Mário Jorge Pereira[3], Ulisses Miranda Azeiteiro[4], M. José Fernández-Gómez[5]

[1]*School of Tourism and Maritime Technology, Marine Resources Research Group, Polytechnic Institute of Leiria, P-2520-641 Peniche, Portugal and Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Box 3046, 3001-401 Coimbra, Portugal, susana.mendes@ipleiria.pt.*

[2]*University of Salamanca, Department of Statistics, Spain pgalinfo@usal.es*

[3]*University of Aveiro, Department of Biology, Aveiro, Portugal, mverde@ua.pt*

[4]*University Aberta, Department of Science and Technology, 4200-055 Porto, Portugal and Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Box 3046, 3001-401 Coimbra, Portugal, ulisses@uab.pt.*

[5]*University of Salamanca, Department of Statistics, Spain, mjfg@usal.es*

**Abstract:** Multivariate analysis, and in particular, multivariate projection methods, have undergone an important development during the last few decades, and are widely applied in fields such as chemometrics, traditionally associated to arrays that have a wide number of potentially related variables compared to the number of observations. In these situations, the classical multivariate techniques do not seem appropriate (such as Principal Component Analysis (PCA), Canonical Correspondence Analysis (CCA) or STATIS methods family). These problems are usual when studying environmental problems and N-way methods, such as Tucker3 model (Tucker, 1966; Kroonenberg, 1983), had gained a lot of attention in the recent years in that field. The aim of this study is to compare the performance of Tucker3 model with one of the biplot methods, the HJ-Biplot (Galindo, 1986; Galindo and Cuadras, 1986), on the analysis of real data sets. Thus, the two techniques are applied to data (physical-chemical and bacteriological parameters) obtained from a study performed at Canal de Mira, Ria de Aveiro, Western Portugal from January 2002 to June 2003 (Mendes et al, 2012; Mendes et al, 2009; Resende et al, 2005). The environmental data consist of 7 variables measured in three sites, at two tide conditions and at the same dates. So, in this case the sites are the same at all dates and the data collected can therefore be arranged into three-way array as sites x environmental variables x time. The Tucker3 method was chosen due its flexibility at fitting data with different modes (ways), each of which can have a different number of latent factors (Smilde et al, 2004). On the other hand, HJ-Biplot (as extension of the classical biplot methods) had the advantage that it is a simultaneous representation, and it gets the maximum quality of representation both for the rows and for the columns of the data matrix.

**Keywords:** Three-way data, exploratory analysis, environmental variables, Biplot analysis, Tucker3.

# References

Kroonenberg, P.M. (1983). Three-mode principal component analysis. DSWO Press, Leiden, Germany.

Mendes S.; Fernández-Gómez, M.J.; Resende, P.; Pereira, M.J.; ; Galindo, M.P.; Azeiteiro, U.M. (2009). Spatio-temporal structure of diatom assemblages in a temperate estuary. A STATICO analysis. Estuarine, Coastal and Shelf Science, v.84, 637-644.

Mendes, S.; Fernández-Gómez, M.J.; Pereira, M.J.; Azeiteiro, U.M.; Galindo, M.P. (2012). An empirical comparison of canonical correspondence analysis and STATICO in the identification of spatio-temporal ecological relationships. Journal of Applied Statistics. v.39, n.5, 979-994.

Resende, P.; Azeiteiro, U.M.; Pereira, M.J. (2005). Diatom ecological preferences in a shallow temperate estuary (Ria de Aveiro, Western Portugal). Hydrobiologia, v.544, 77-88.

Tucker, L.R. (1966). Some mathematical notes on three-mode factor analysis. Psychometrika, v.31, 279-311.

Smilde, A.; Bro, R.; Geladi, P. (2004). Multi-way analysis. Applications in the chemical sciences. Chichester: Wiley.