



**Grupo de Probabilidades e Estatística**  
CENTRO DE INVESTIGAÇÃO E DESENVOLVIMENTO EM MATEMÁTICA E APLICAÇÕES  
Departamento de Matemática  
Universidade de Aveiro



Celebrating 40th anniversary of Statistics Portuguese Society (SPE) – “12 months/12 initiatives”:

VIII Workshop of Probability and Statistics group — Interdisciplinarity and applications —  
Sousa Pinto room, Department of Mathematics, University of Aveiro  
25<sup>th</sup> March 2020

**Program:**

14:00 – 14:10 **Celebrating 40th anniversary of Statistics Portuguese Society (SPE)**

*Isabel Pereira (Vice-President of SPE)*

14:10 – 14:50 **Tax professionals' profiles concerning tax noncompliance and tax complexity**

*Carlos Ferreira*

DEGEIT/IEETA, University of Aveiro

14:50 – 15:30 **Choosing among alternative histories of a tree**

*Rui Costa*

DFis, University of Aveiro

15:30 – 15:50 Coffee break

15:50 – 16:30 **The HJ-Biplot Visualization of the Singular Spectrum Analysis Method**

*Alberto Oliveira da Silva*

CIDMA, University of Aveiro

16:30 – 17:10 **Different regression models applied to chemical fields: is robust regression a better choice?  
A simulation study**

*M. Cristina Miranda*

ISCA & CIDMA, University of Aveiro



---

This workshop is free for anyone who is interested in applied statistics. Only a quick registration is required. For registration please go to <https://sites.google.com/view/workshops-ps-cidma> and provide your name and affiliation. Official languages: English and Portuguese (all slides will be in English)

---



**Grupo de Probabilidades e Estatística**  
CENTRO DE INVESTIGAÇÃO E DESENVOLVIMENTO EM MATEMÁTICA E APLICAÇÕES  
Departamento de Matemática  
Universidade de Aveiro



**Tax professionals' profiles concerning tax noncompliance and tax complexity \***

Carlos Ferreira

**Abstract**

In recent decades, tax systems, reacting to the encouraging environment for the use of sophisticated mechanisms to reduce tax payments, as a consequence of the increasing complexity and globalization of business, have systematically increased their levels of complexity, namely the legislative component. In self-reporting systems, and due to the general lack of tax-payers fiscal literacy, the role of accountants and other tax professionals is crucial in advising or endorsing compliance with their tax obligations. In the Portuguese case this role is paramount as companies, entrepreneurs and professionals under the accounting regime must hire CAs (Chartered Accountants) to represent them before the AT (Tax and Customs Authority) and thus these professionals have a deep knowledge of the issues resulting from the fiscal complexity. Thus their characterization is most relevant in the scope of this subject.

This talk will present a characterization of the CA profile in the context of the perception of complexity and fiscal noncompliance, in Portugal. Data collected through a questionnaire was analysed using non-parametric techniques and multivariate analysis. As an example, contingency tables allowed to detect trends in socio-demographic and professional characteristics as a foreseeable predominance of women. The research hypotheses included, for instance, the relationship between the perception and dimensions of the causes of tax complexity and, age, gender, level of experience, tax knowledge and the size of the companies. Chi-squared, Mann-Whitney and Kruskal-Wallis tests were used for the nominal and ordinal data. Moreover, to measure relationships among variables the Spearman correlation was used. Results suggest that most CAs considered the Portuguese tax system complex and that this perception impacts their activity. Multiple-Correspondence Analysis allowed identifying different CA profiles concerning how prone they are to intentional and unintentional tax non-compliance. Using Principal Component Analysis allowed variable reduction leading to new synthetic indexes to be used in explicative models of intentional and unintentional tax non-compliance through Logistic Regression.

\*A joint work with Ana Clara Borrego (ESTGP/CICF/C3i, Instituto Politécnico de Portalegre) and Cidália Lopes (ISCAC/CICF, Instituto Politécnico de Coimbra).

**Carlos Manuel dos Santos Ferreira** tem uma licenciatura em Engenharia Eletrónica e de Telecomunicações, Mestrado em Estatística e Investigação Operacional, Doutoramento em Matemática e Agregação em Economia e Gestão. É membro integrado do Instituto de Engenharia Eletrónica e Informática de Aveiro (IEETA) e Professor Associado com Agregação na Universidade de Aveiro. É, atualmente, o Coordenador da Comissão de Análise do Sistema de Garantia de Qualidade (SGQ) do Processo de Ensino-Aprendizagem do DEGEI. Integrou comissões de organização e de programa de cerca de 50 conferências nacionais e internacionais. Tem como áreas de interesse de investigação: Análise de Dados, Investigação Operacional, Logística, Apoio à Decisão e Gestão da Informação. É autor de mais de 150 publicações em conferências internacionais, capítulos de livro e em revistas.



# Grupo de Probabilidades e Estatística

CENTRO DE INVESTIGAÇÃO E DESENVOLVIMENTO EM MATEMÁTICA E APLICAÇÕES  
Departamento de Matemática  
Universidade de Aveiro



## Choosing among alternative histories of a tree

Rui Costa

### Abstract:

The structure of an evolving network contains information about its past. Extracting this information efficiently, however, is, in general, an extremely difficult challenge. We formulate a fast and efficient method to estimate the most likely history of growing trees, based on exact results on root finding. We show that our algorithm produces the exact step-wise most probable history in a broad class of tree growth models. Our formulation is able to treat very large trees and therefore allows us to make reliable numerical observations regarding the possibility of root inference and history reconstruction in growing trees. We obtain a general formula for the size-dependence of the mean logarithmic number of possible histories of a given tree, a quantity that largely determines the reconstructability of tree histories. We also reveal an uncertainty principle: a relationship between the inferrability of the root and that of the complete history, indicating that there is a trade-off between the two tasks; the root and the complete history cannot both be inferred with high accuracy at the same time.

*Rui Costa* é investigador júnior no Departamento de Física da Universidade de Aveiro e membro do Lab. /UI: I3N/Theoretical and Computational Physics Group. As áreas de investigação são: Complex Networks and Systems, Statistical Physics, Dynamical Systems, Percolation Theory, Deep Learning, Lossy Compression, Complex Topology



**Grupo de Probabilidades e Estatística**  
CENTRO DE INVESTIGAÇÃO E DESENVOLVIMENTO EM MATEMÁTICA E APLICAÇÕES  
Departamento de Matemática  
Universidade de Aveiro



**The HJ-Biplot Visualization of the Singular Spectrum Analysis Method**

Alberto Oliveira da Silva

**Abstract**

The extraction of essential features of a real-valued time series is crucial for exploring, modeling, and forecasting. Taking advantage of the trajectory matrix constructed applying Singular Spectrum Analysis, as well as of its decomposition through Principal Component Analysis via Partial Least Squares, we implement a graphical display employing the biplot methodology to visualize the separation of the components of time series. In this work, we discuss the so-called HJ-biplot, which yields a simultaneous representation of both rows and columns of a matrix with maximum quality. Interpretation of this type of biplot on trajectory matrices is discussed from a real-world time series data.

*Alberto Oliveira da Silva* is currently a Ph.D. student of the Doctoral Programme in Mathematics of the Department of Mathematics at the University of Aveiro (DMat – UA) and a member of the Center for Research & Development in Mathematics and Applications (CIDMA). He holds a Master's degree in Mathematics and Applications from University of Aveiro, and his scientific activity has been mainly in the area of multivariate statistical analysis, with particular focus on partial least squares (PLS) and biplot methods.



**Grupo de Probabilidades e Estatística**  
CENTRO DE INVESTIGAÇÃO E DESENVOLVIMENTO EM MATEMÁTICA E APLICAÇÕES  
Departamento de Matemática  
Universidade de Aveiro



Different regression models applied to chemical fields: is robust regression a better choice? A simulation study

Maria Cristina Souto Miranda

**Abstract**

Different scientific communities usually communicate using their own languages. Sometimes the same statistic methodology is baptized with a different name for each field of application where it is used. After identifying similar designation, the aim is to apply the most adequate methodology to a given problem. In clinical chemistry research, it is usual to have different methodologies, treatments, chemical compounds or devices that should give the same measure. How to compare them? An adequate regression model usually presents a hint. For example, it can highlight proportional estimates between two kinds of procedures or between two measures of the same quantity. Ordinary Least Squares (OLS) Linear Regression, Orthogonal Regression, Deming Regression or Passing-Bablok Regression are among the most used methodologies in clinical literature. Different approaches often lead to different results. Which is the most accurate process of validation? It is not a simple answer, namely, when the required assumptions for the OLS seem to fail, as it happens most frequently with real data. Results might be misleading and this suggests that robust methods should be used. In this work, a simulation study is performed, whose main goal is to evaluate what is the best type of regression for comparing different validation methods, in the presence of outliers. Some of the most spread methods over the clinical chemistry community are compared with robust regression methods.

**M. Cristiana Miranda** graduated from University of Aveiro in mathematics and geometry in 1988, received a master degree in applied mathematics with a thesis in probability and statistics from the faculty of Sciences of University of Lisbon in 1992. In 2005 obtained a PhD in the same faculty, also from the University of Lisbon in probability and statistics with the thesis: *Statistics of extreme values tail index and extremal index estimation*. She is currently a Professor at Higher Institute for Accountancy and Administration of University of Aveiro in Portugal. Her research interests include statistics of extremes, semi-parametric inference, applied statistics and computational statistics. She is a member of the Portuguese Statistical Society, Mathematics Portuguese Society and International Statistical Institute.