Imprint

Book of Abstracts of the 20th Applied Stochastic Models and Data Analysis International Conference with the Demographics 2023 Workshop
Heraklion, Crete, Greece: 06-09 June, 2023
Published by: ISAST: International Society for the Advancement of Science and Technology.
Editor: Christos H Skiadas

e-book (PDF)
Preface

It is our pleasure to welcome the guests, participants and contributors to the International Conference (ASMDA 2023) on Applied Stochastic Models and Data Analysis and (DEMOGRAPHICS2023) Demographic Analysis and Research Workshop.

The main goal of the conference is to promote new methods and techniques for analyzing data, in fields like stochastic modeling, optimization techniques, statistical methods and inference, data mining and knowledge systems, computing-aided decision supports, neural networks, chaotic data analysis, demography and life table data analysis.

ASMDA Conference and DEMOGRAPHICS Workshop aim at bringing together people from both stochastic, data analysis and demography and health areas. Special attention is given to applications or to new theoretical results having potential of solving real life problems.

ASMDA 2023 and DEMOGRAPHICS 2023 focus in expanding the development of the theories, the methods and the empirical data and computer techniques, and the best theoretical achievements of the Applied Stochastic Models and Data Analysis field, bringing together various working groups for exchanging views and reporting research findings.

We thank all the contributors to the success of these events and especially the authors of this Abstracts Book. Special thanks to the Plenary, Keynote and Invited Speakers, the Session Organisers, the Scientific Committee, the ISAST Committee, Yiannis Dimotikalis, Aristeidis Meletiou, the Conference Secretary Eleni Molfesi, and all the members of the Secretariat.

Our sincere thanks and appreciation to Prof. Christos Floros, Department of Accounting & Finance, Hellenic Mediterranean University, Greece and all the members of LAFIM Laboratory.
May 2023

Christos H. Skiadas
Conference Chair

Professor Christos Floros
Conference Co-Organiser

ASMDA Conferences and Organizers


4th ASMDA 1988 Nancy, France. J. Janssen and Jean-Marie Proth

5th ASMDA 1991 Granada, Spain. Mariano J. Valderrama

6th ASMDA 1993 Chania, Crete, Greece. Christos H Skiadas

7th ASMDA 1995 Dublin, Ireland. Sally McClean

8th ASMDA 1997 Anacapry, Italy. Carlo Lauro

9th ASMDA 1999 Lisbon, Portugal. Helena Bacelar-Nicolau

10th ASMDA 2001 Compiègne, France. Nikolaos Limnios
11th ASMDA 2005 Brest, France. Philippe Lenca

12th ASMDA 2007 Chania, Crete, Greece. Christos H Skiadas

13th ASMDA 2009 Vilnius, Lithuania. Leonidas Sakalauskas

14th ASMDA 2011 Rome, Italy. Raimondo Manca

15th ASMDA 2013 Mataró (Barcelona), Spain. Vladimir Zaiats

16th ASMDA 2015 Piraeus, Greece. Sotiris Bersimis

17th ASMDA 2017 London, UK. Christos H Skiadas

18th ASMDA 2019, Florence, Italy

19th ASMDA 2021, Athens, Greece

20th ASMDA 2023, Heraklion, Crete, Greece
SCIENTIFIC COMMITTEE

Jacques Janssen, Honorary Professor of Universite’ Libre de Bruxelles, Honorary Chair
Alejandro Aguirre, El Colegio de México, México
Alexander Andronov, Transport and Telecom. Institute, Riga, Latvia
Vladimir Anisimov, Center for Design & Analysis, Amgen Inc., London, UK
Dimitrios Antzoulakos, University of Piraeus, Greece
Soren Asmussen, University of Aarhus, Denmark
Robert G. Aykroyd, University of Leeds, UK
Narayanaswamy Balakrishnan, McMaster University, Canada
Helena Bacelar-Nicolau, University of Lisbon, Portugal
Paolo Baldi, University of Rome “Tor Vergata”, Italy
Vlad Stefan Barbu, University of Rouen, France
S. Bersimis, University of Piraeus, Greece
Henry W. Block, Department of Statistics, University of Pittsburgh, USA
James R. Bozeman, American University of Malta, Malta
Mark Brown, Department of Statistics, Columbia University, New York, NY
Ekaterina Bulinskaya, Moscow State University, Russia
Jorge Caiado, Centre Appl. Math., Econ., Techn. Univ. of Lisbon, Portugal
Enrico Canuto, Dipart. di Automatica e Informatica, Politec. di Torino, Italy
Mark Anthony Caruana, University of Malta, Valletta, Malta
Erhan Çinlar, Princeton University, USA
Maria Mercè Claramunt, Barcelona University, Spain
Marco Dall’Aglio, LUISS Rome, Italy
Guglielmo D’Amico, University of Chieti and Pescara, Italy
Pierre Devolder, Université Catholique de Louvaine, Belgium
Giuseppe Di Biase, University of Chieti and Pescara, Italy
Yiannis Dimotikalis, Technological Educational Institute of Crete, Greece
Dimitris Emiris, University of Piraeus, Greece
N. Farmakis, Aristotle University of Thessaloniki, Greece
Lidia Z. Filus, Dept. of Mathematics, Northeastern Illinois University, USA
Jerzy K. Filus, Dept. of Math. and Computer Science, Oakton Community College, USA
Leonid Gavrilov, Center on Aging, NORC at the University of Chicago, USA
Maria Gavrilova, Center on Aging, NORC at the University of Chicago, USA
A. Giovannis, Technological Educational Institute of Athens, Greece
Valerie Girardin, Université de Caen Basse Normandie, France
Joseph Glaz, University of Connecticut, USA
Maria Ivette Gomes, Lisbon University and CEAUL, Lisboa, Portugal
Gerard Govaert, Universite de Technologie de Compiegne, France
Alain Guennoche, University of Marseille, France
Y. Guermeur, LORIA-CNRS, France
Montserrat Guillen University of Barcelona, Spain
Steven Haberman, Cass Business School, City University, London, UK
Diem Ho, IBM Company
Emilia Di Lorenzo, University of Naples, Italy
Aglaios Kalamianou, Panteion Univ. of Political Sciences, Athens, Greece
Udo Kamps, Inst. fur Stat. und Wirtschaftsmath., RWTH Aachen, Germany
Alex Karagiorgiou, Department of Mathematics, University of the Aegean, Greece
Lino Sant, University of Malta, Valletta, Malta
José M. Sarabia, Department of Economics, University of Cantabria, Spain
Sergio Scarlatti, University of Rome “Tor Vergata”, Italy
Hanspeter Schmidli, University of Cologne, Germany
Dmitrii Silvestrov, University of Stockholm, Sweden
P. Sirirangsi, Chulalongkorn University, Thailand
Christos H. Skiadas, ManLab, Technical University of Crete, Greece (Co-Chair)
Charilaos Skiadas, Hanover College, Indiana, USA
Dimitrios Sotiropoulos, Techn. Univ. of Crete, Chania, Greece
Fabio Spizzichino, University of Rome “La Sapienza”, Italy
Gabriele Stabile, University of Rome “La Sapienza”, Italy
Valeri Stefanov, The University of Western Australia
Anatoly Swishchuk, University of Calgary, Canada
R. Szekli, University of Wroclaw, Poland
T. Takine, Osaka University, Japan
Andrea Tancredi, University of Rome “La Sapienza”, Italy
P. Taylor, University of Melbourne, Australia
Cleon Tsimbos, University of Piraeus, Greece
Mariano Valderrama, University of Granada, Spain
Panos Vassiliou, Department of Statistical Sciences, University College London, UK
Larry Wasserman, Carnegie Mellon University, USA
Wolfgang Wefelmeyer, Math. Institute, University of Cologne, Germany
Shelly Zacks, Binghamton University, State University of New York, USA
Vladimir Zaitsev, Universitat de Vic, Spain
K. Zografos, Department of Mathematics, University of Ioannina, Greece
Scientific committee continued
Prof. Georgios Agiomirgianakis, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Prof. Georgios Iatrakis, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Prof. Dimitrios Kirikos, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Prof. Christos Floros, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Associate Professor Stavros Arvanitis, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Associate Professor Ioannis Chatziantoniou, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Assistant Professor Georgios Galifianakis, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Assistant Professor Efthalia Tampouratzis, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Assistant Professor Dimitrios Vortelinos, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Assistant Professor Evanthis Zervoudi, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Lecturer Ioannis Yiannoulis, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Local Committee

Professor Christos Floros, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Associate Professor Ioannis Chatziantoniou, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Assistant Professor Efthalia Tampouratzi, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Dr. Georgios Matalliotakis, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.

Organising Committee

Christos Floros, Department of Accounting & Finance, Hellenic Mediterranean University, Greece.
Christos H Skiadas, ISAST Chair
Anthi Katsirikou, Secretary
Aris Meletiou, Webex and computing
Eleni Molfesi, Secretariat
Plenary / Keynote / Invited Talks

Nikolaos Apergis, Dept of Banking and Financial Management, University of Piraeus, Greece

Narayanaswamy Balakrishnan, McMaster University, Canada

Lars Fadnes
Department of Global Public Health and Primary Care & Bergen Addiction Research, Department of Addiction Medicine, University of Bergen & Haukeland University Hospital

Revaz Kakubava, Georgian Technical University, Tbilisi, Georgia

Sally McClean, Ulster University, Northern Ireland, UK

Claude Lefèvre, Université Libre de Bruxelles, Belgium

Ying Ni, Mälardalen University, Sweden

Peter Pflaumer, Department of Statistics, Technical University of Dortmund, Germany

Jean-Marie Robine, INSERM, the French National Institute of Health and Medical Research

Fabrizio Ruggeri, CNR, IMATI, Milano, Italy

Gilbert Saporta, Conservatoire National des Arts et Métiers, Paris, France

Dmitrii Silvestrov, University of Stockholm, Sweden
xi
### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>iii</td>
</tr>
<tr>
<td>ASMDA Conferences and Organizers</td>
<td>v</td>
</tr>
<tr>
<td>Committees</td>
<td>vi</td>
</tr>
<tr>
<td>Plenary / Keynote / Invited Talks</td>
<td>viii</td>
</tr>
<tr>
<td>Abstracts</td>
<td>1</td>
</tr>
</tbody>
</table>
Plenary and Keynote Sessions

Climate Change and Monetary Policy
(Honorary Speech)
Nikolaos Apergis
Department of Banking & Financial Management University of Piraeus

Climate change poses an important policy challenge for policymakers around the world. The critical questions have to do with the role of central banks, along with how climate change is similar to or distinct from other natural disasters and how they can cope with it without compromising their other goals, such as price and financial stability. Their credibility is also expected to be further enhanced by avoiding the temptation to exaggerate the treatment of climate change.

Accelerated Life Testing of One-Shot Devices: Data Collection and Analysis
N. Balakrishnan
Fellow of the Royal Society of Canada
Fellow of the American Statistical Association
Fellow of the Institute of Mathematical Statistics
Distinguished University Professor, Department of Mathematics and Statistics
McMaster University, Hamilton, Ontario, Canada

In this talk, I will first introduce the problem of modeling lifetimes of one-shot devices. I shall then elaborate on parametric, Bayesian and robust inferential methods for one-shot device data. Finally, I will describe methods that are suitable for devices with competing risks, dependent risks, and model validation. Through-out, I will use many practice datasets to illustrate all the methods discussed.

Comparison of Life Expectancy Gains from Various Dietary Changes: Modelling in Seven Countries under the Food4HealthyLife Framework
Lars Fadnes
Professor & Research Group Leader
Department of Global Public Health and Primary Care & Bergen Addiction Research, Department of Addiction Medicine, University of Bergen & Haukeland University Hospital

Reducing mortality and morbidity from unhealthy diets is essential to reduce premature mortality from non-communicable diseases. We estimate life expectancy gains from sustained dietary changes from nationally specific dietary patterns in for China, France, Germany, Iran, Norway, the United Kingdom, and the United States to a health-optimised diet, a vegetarian/vegan diet, or feasible dietary changes. Our models
use data from meta-analyses with dose-response relationships data between intake of fifteen food groups and their associations with all-cause mortality, mortality data from the Global Burden of Disease study. Model adjustment accounted for correlations between food groups, estimation of expected time delay, and life-table methodology, while taking age, sex, height, weight, and physical activity level into account. Estimated life expectancy with uncertainty intervals gains from a sustained change among 40-year-old females/males from the typical dietary patterns in the respective countries to an optimised diet were in the range of 5.7-8.3 years. The largest gains would be achieved by eating more legumes, whole grains and nuts, and less processed meat. The relative order and magnitude differ slightly between settings. Among 40-year-old adults, a sustained change from a nationally typical dietary pattern to optimised vegetarian/vegan diets, estimated gains were in the range of 4.8-7.9 years. Estimated gains from a sustained change among 40-year-old females/males from a nationally specific typical diets to feasible beneficial dietary changes were in the range of 1.7-4.3 years. A sustained dietary change from the current typical diet to a health-optimised diet, a vegetarian/vegan-diet, or feasible beneficial dietary changes, all provide substantial gains in life expectancy for female and male adults across ages and countries. The Food4HealthyLife could be useful for clinicians, policy makers, and lay-people in understanding the health impact of dietary choices.

Markovization and Marking in semi-Markov Queuing systems
Revaz Kakubava
Georgian Technical University, Georgia, Tbilisi

On Partially Schur-constant Models and their Associated Copula
Claude Lefèvre
Université Libre de Bruxelles, Département de Mathématique, Bruxelles

Schur-constant vectors are used to model duration phenomena in various areas of economics and statistics. They form a particular class of exchangeable vectors and, as such, rely on a strong property of symmetry. To broaden the field of applications, partially Schur-constant vectors are introduced which correspond to partially exchangeable vectors. First, their copulas of survival, said to be partially Archimedean, are explicitly obtained and analyzed. Next, much attention is devoted to the construction of different partially Schur-constant models with two groups of exchangeable variables. Finally, partial Schur-constancy is briefly extended to the modeling of nested and multi-level dependencies. This research work was recently published in Lefèvre (2021). An interesting partly related article is Ressel (2018).

References:
Using Markov and Related Models for Characterising and Monitoring Patients in Smart Homes

Sally McClean
Ulster University, School of Computing, Northern Ireland, UK

Older people, in particular, often experience difficulties in performing activities of daily living (ADLs), such as grooming, showering, or taking medication. Such difficulties can lead to the loss of independence for the patients with associated increased cost of care and distress. This has led to the developments in smart home technology that use sensors or imagery to provide a home environment that offers assisted living, through the use of probability models to describe the home environment, patient behavioural patterns and the potential impact of interventions if such behaviours become unhealthy, unsafe, or dangerous.

Markov, or more expressive models such as semi-Markov or nth-order Markov, have been previously used in healthcare in a wide range of contexts, including patient monitoring in smart homes. For example, sensors have been used in smart homes to collect information about a patient's activities, such as their movements, sleep patterns, and medication adherence. Markov models can help process this data to provide insights into the current and long-term health and well-being status of a patient and provide useful information for the healthcare professionals or prevent the patient engaging in harmful activities. By collecting data over time, a Markov, or related, model can be used to predict the patient's future activity levels and detect any changes in their health status. In addition, we can use such models to detect anomalies which may be indicative of deterioration or danger for the patient.

We will describe the use of Applied Stochastic Models and Data Analysis for patients living in sensorised smart homes, for different such scenarios.

Option Pricing and Hedging using Artificial Intelligence (AI) Technology

Ying Ni
Associate Professor in Mathematics/Applied mathematics, Mälardalen University, Sweden

This talk provides an overview about applications of AI technology in option pricing/hedging and presents results on two joint collaborative projects with Bjørn André Aaslund, Johannes Berge, and Rita Pimentel from Norwegian University of Science and Technology. In the first project, we price Bermudan basket options with maximum and minimum payoffs on a substantial amount of underlying assets. In the literature, neural network architectures have shown great advantage when applied to option pricing problems. This motivates us to adopt a neural network approach to approximate the continuation values for the Bermudan options. The underlying asset processes for the basket options are often assumed to be geometric Brownian motions with constant volatilities. However, in financial markets one observe mean-reverting stochastic volatility process, the novelty here is to implement a neural network architecture for high-dimensional max and min-basket options under such stochastic volatilities. We also propose, as a by-product, a simple synthetic regression method that can always reduce the problem into a single asset problem requiring a low computing budget. Both the neural network and the synthetic regression methods are evaluated by comparing our results with the literature for the same type of options, when such a comparison is possible. In the second project, a novel hedging framework
called X-hedging is proposed. The previously used neural network method makes good performance in mapping complex non-linear relationships. However, this popular AI technology also has a drawback of lacking explainability. In our X-hedging framework, we use instead gradient boosted decision trees to increase the explainability of the state-of-the-art frameworks without sacrificing performance. With this increased explainability, X-hedging complies with the newly proposed guidelines and regulations related to the modern concept of Explainable Artificial Intelligence. We validate our X-hedging framework via numerical experimental studies, comparing to benchmark methods including the delta hedging in Black-Scholes model and the more recent deep hedging methods. We observe that X-hedging achieves comparable performance.

Leonhard Euler’s Study on the Reproduction of the Human Race

Peter Pflaumer
Department of Statistics, Technical University of Dortmund, Germany

The famous Swiss mathematician Leonhard Euler developed three variants of a simple population projection model that assumes a couple with different fertility behavior and different lifespans. One of these projection variants was published and made known by the German demographer Johann Peter Süßmilch in his book “The Divine Order”. The other two variants are fairly unknown in today's literature. In this paper, the variants of Euler's population projection are compared and reanalyzed using matrix algebra. Diagrams and tables of the population time series and their growth rates are shown. Age structures of selected years are presented. The solutions of the projection equations are derived. It is shown that the projection models can be described by a geometric trend model overlaid by cyclical components. In the long run, the population time series can be explained quite well by the sum of only two components, the trend component and a long-term component with explosive cycles. Gradually, the influence of the cyclical component diminishes, and the series can be explained by its geometric trend component alone, as already described by Euler.

Change Over Time in Human Longevity
(summary of a study by Camarda and Robine, 2023)

Jean-Marie Robine

Our study suggests a major break in the history of human longevity around 1950 with the establishment of a new demographic regime that witnessed an increase in the limits of human longevity. Traditionally measured by life expectancy at birth, the history of longevity in Western countries seemed to be well known. The first statistical series available from the 18th century onward indicate values that fluctuate considerably from one year to the next, averaging around 40 years. These values were again confirmed in the second half of the 19th century when a first increase in life expectancy took place, coinciding globally with the Industrial Revolution. Life expectancy reached about fifty years on the eve of the First World War, then sixty years after the Second World War. Women’s life expectancy reaches seventy years in 1980 and eighty years in the 2010s. Proposed by Wilhelm Lexis, the modal age at death (M) has been advocated as a better indicator of human longevity. This measure did not vary significantly before the First World War fluctuating between 72 and 75 years of age. A clear increase was
visible from the 1930s and nowadays the most frequent age at death is close to 90 years for women. Both life expectancy at birth and $M$ are central indicators of human longevity. However, whereas life expectancy at birth is modified by changes in mortality at all ages of life, including juvenile ages, modal age at death is only influenced by mortality at old ages. Our study heavily relies on an alternative and empirical indicator focusing on the tail of the age-at-death distribution: the oldest age still providing 30 deaths in a year. Labelled HAPaL30, this indicator was constructed by trial and error and the number 30 is necessarily a subjective choice which, however, does not impact our analysis. HAPaL30 has been developed looking for a statistical indicator that would come as close as possible to the age of the oldest person dying each year (MRAD) without being impacted by natural fluctuations that characterize extreme values. In other words, this indicator aims to reveal trends that could be hidden by annual fluctuations in MRAD (Robine and Herrmann, 2020). Trend of HAPaL30 in Sweden suggest a quasi-stagnation of longevity since vital statistics have been available going back to the 18th century. Additional data from other countries in the 19th century validate this pattern until 1950. Afterwards this indicator starts to increase linearly until the last available years, without any sign of acceleration or slowing down in its increase, going, for example, from 99 years in France in 1946 to 109 years in 2016, i.e. an increase of 10 years in 70 calendar years for French women. This first visual inspection may seem consistent with the considerable increase in the number of centenarians observed since the 1950s as well as with the increase in the modal age at death since the 1930. However, mortality changes are not the sole mechanism behind an increase in number of survivors at very high ages in general, and of HAPaL30 in particular. This led us to plan the second part of our study: decompose the increase in HAPaL30 between what was due to larger numbers of survivors at age 90 (whether due to an increase in the size of birth cohorts, a positive migration balance or the fall in mortality before age 90) and what was due to a possible decrease in mortality at age 90 and beyond. We chose age 90 based on the observed mortality trends. Whereas all analyzed countries experience a sharp decline in mortality at age 85 over the period studied, trends beyond this age become increasingly blurry and mortality rates at age 95 or 100 do not present any clear pattern. By means of the proposed decomposition, we contrast a scenario where mortality at age 90 and beyond remains identical to that observed from 1900 in the cohort born in 1810 versus the actual observation of mortality trends for each cohort. This allowed us to show that while, on average for females, 61% of the increase in HAPaL30 was due to the increase in the number of survivors at age 90 for the 11 Western countries studied, 39% of this increase was indeed due to the mortality improvement beyond age 90. Slightly different outcomes are found for males: 68% and 32% for residual and old-age contributions, respectively. What surprised us most was that highlighted changes in human longevity history date back to the 1950s; that is, more than 70 years ago, and decades before the still ongoing debate about what the “true” limit to human longevity. This urged us to investigate historical sources of the debate about upper limits to human longevity. Likely due to an enduring stagnated mortality interrupted by cyclical crises and to a constant absence of reliable data, scholars could hardly believe that limit to the human life span may have been pushed back. This perception do not even change in the first half of the 20th century when limits started being broken by the effect of mortality improvement that has been progressively affecting higher and higher ages, at least since the 1920s or 1930s in Western countries. In the absence of statistical information, ancient thinkers imagined almost everything. They dreamt of immortality and eternal youth on a par with gods. Above all, they theorized about the existence of limits to longevity which were set by god or transcendental powers. Aristotle told us that they were characteristic of each species and Buffon that they seemed unchangeable on a time scale. Expected for the Marquis de Condorcet in the 18th century and until the 20th century, no one proposed that limits to human longevity could vary over time, and in particular increase. Buffon did suggest the existence of
exceptions to the limits of common longevity, but he also stated them in immutable
terms: some privileged individuals can have their life extended “almost to the double
of the ordinary term”. It is easier to understand from all these old texts why, in the
1970s, statisticians and demographers had difficulty projecting past increases in life
expectancy into the future and why, still at the end of the 20th century, so many
biologists continued to strongly advocate no more centenarians today than in the
Neolithic period.

Unsupervised Statistical Tools for Anomaly Detection: The
Case of Healthcare Frauds
Fabrizio Ruggeri
CNR IMATI, ITALY

The research is motivated by the increased interest in detecting possible frauds in
healthcare systems. We propose some unsupervised statistical tools (Lorenz curve,
concentration function, sum of ranks, Gini and Pietra indices to provide efficient and
easy-to-use methods aimed to signal possible anomalous behaviours. A more
sophisticated method, based on Bayesian co-clustering, is presented as well.

Optimal Scaling: New Insights into an Old Problem
Gilbert Saporta
Cedric Lab, Conservatoire national des arts et métiers, Paris, France

Processing qualitative variables with a very large number of categories in Machine
Learning is an opportunity to revisit the theory of optimal scaling and its applications.
Coding (or scoring) a qualitative variable consists in assigning numerical values to its
modalities, thus transforming it into a discrete numerical variable. Scoring qualitative
variables has a long history going back to K. Pearson, R.A. Fisher, L. Guttman,
C.Hayashi, etc. It was the origin of correspondence analysis. The 1970s and early
1980s were the years of the search for optimal scoring (aka optimal scaling) in
supervised and unsupervised contexts, performed with alternating least squares. For
nearly 30 years, the topic did not generate much research; applications became
routine, such as risk scores in banking and insurance. With the availability of massive
data, machine-learning researchers and practitioners were confronted with categorical
data, ill-suited to neural networks with moreover a large number of categories (eg zip
codes). Generally ignoring the work of statisticians, dozens of encoding methods have
flourished like Hash encoding, or methods where the encoding only depends on the
response variable (conditional average) as well as the One-Hot Encoding which is
nothing else than the disjunctive form with as many indicators as modalities. The large
size of some categorical data raises problems of stability and overfitting that were
neglected in classical statistical applications where the number of modalities is small
and where the learning-testing approach was not frequent. The confrontation of these
two worlds allows us to consider a renewal of the coding methods and how
 correspondence analysis may be reinterpreted from a non-linear point of view.
Keywords: Optimal scaling, One hot encoding, Correspondence analysis.
Perturbed Semi-Markov Type Processes and Applications

Dmitrii Silvestrov
Department of Mathematics, Stockholm University, Stockholm, Sweden

The lecture presents results from a two-volumes monograph [1, 2], which is devoted to the study of limit and ergodic theorems for regularly and singularly perturbed Markov chains, semi-Markov processes, and multi-alternating regenerative processes with semi-Markov modulation. The first volume presents phase space reduction algorithms and limit theorems on weak convergence of distributions and convergence of expectations of hitting times for regularly and singularly perturbed semi-Markov processes. These results have their own theoretical and applied significance and also play an important role in obtaining ergodic theorems for perturbed semi-Markov type processes. The second volume presents new super-long-, long- and short-time ergodic theorems for perturbed multi-alternating regenerative processes modulating by regularly or singularly perturbed finite semi-Markov processes. The theoretical results are illustrated by applications to risk processes and queuing systems.

References:

Keywords: Limit theorem, Hitting time, Ergodic theorem, Perturbation, Phase space reduction, Semi-Markov process, Regeneration process, Multi-alternating regenerative process, Risk process, Queuing system.
The objective of this paper is to propose a topological approach of clustering in evolutionary data analysis. We are interested in clustering resulting from exploratory methods of joint analysis of several data tables, methods applied more particularly to temporal data.

The clustering is one of the most widely used approaches to exploring multidimensional data. The two common unsupervised clustering strategies are Hierarchical Ascending Clustering (HAC) and k-means partitioning used to identify groups of similar objects in a dataset to divide it into homogeneous groups. The proposed approach, called Topological Clustering on Evolutionary Data (TCED), is based on the notion of neighborhood graphs in an evolutionary data context. It makes it possible to simultaneously explore several tables of data collected at different times on the same individual-rows, the variables possibly being different according to the tables considered. The columns-variables of each table are more-or-less correlated or linked according to whether the variable type. It analyzes in each table the structure of the correlations or associations observed between the variables according to their quantitative, qualitative type or a mixture of both.

The proposed TCED approach is presented and illustrated here using a real dataset with quantitative variables. Its results are compared with those resulting from the methods in evolutionary data analysis STATIS and AFM.

**References:**

**Keywords:** Evolutionary data cluster, proximity measure, neighborhood graph, adjacency matrix, hierarchical clustering.
On the First-Passage Area of a One-Dimensional Diffusion Process with Stochastic Resetting

Mario Abundo
Dipartimento di Matematica, Università Tor Vergata, Roma, Italy

We consider a one-dimensional diffusion process in the presence of stochastic resetting, obtained from an underlying diffusion $X(t)$. We suppose that a reset to the position $x_R > 0$ can occur according to a homogeneous Poisson process with rate $r > 0$, so obtaining a process $X(t)$, called Reset Diffusion (RD) process. Our aim is to study the statistical properties of the first-passage time (FPT) through zero of the process $X(t)$, starting from $x > 0$, and its FPA, that is the random area swept out by $X(t)$, till its FPT through zero. In particular, we find explicit expressions for the Laplace transform of the FPT and FPA, and their single and joint moments; moreover, we study the distribution of the maximum displacement of $X(t)$. Thus, we extend to RD processes the results about FPA of diffusions without resetting (see e.g. [1-3]). We generalize the results which hold when $X(t)$ is Brownian motion, obtained in [4] by using special functions. Instead, we use only elementary functions and the properties of jump-diffusion processes on an interval; we are able to express the various quantities as solutions of suitable ODEs. FPT and FPA of a RD process is a very interesting topic in several applied fields, e.g. in biological models for neuronal activity with resetting, in percolation models and queuing theory. In this ambit, the FPT through zero can be identified with the busy period, that is the time until the queue is first empty, whilst FPA represents the cumulative waiting time experienced by all the “customers” during a busy period.

References:

Keywords: First-passage time, First-passage area, Diffusion process.

The Crossovers between Births and Deaths caused by COVID-19: The Mexican case

Alejandro Aguirre$^1$ and Eliud Silva$^1$

$^1$Universidad Anahuac, Mexico

In this paper we describe the monthly behaviour of both Births and Deaths from 2008 to 2020 in Mexico. To do that, we consider the mother’s ages available in the observed fertility datasets taken from the Mexican statistical office (INEGI). We also take into account the evidence from other countries according to the information provided by the Max-Planck Institute. We identify that this kind of crossover is rather unusual around the world and a long time, especially for 2020. In fact, for the Mexican case [it is undoubtedly stronger than for the other countries]. A multivariate time series model is employed to forecast future trends according with mother’s ages. In terms of statistics, it is appreciated that the fertility patterns change mainly in the older ages (more than 30), contrasting with the younger groups (less than 30).

Keywords: cross overs, COVID-19, births, deaths, Mexico.
A General Growth Equation, Its Generalizations and Related First Passage Time problems

Giuseppina Albano¹, A. Barrera¹, V. Giorino¹, P. Roman-Roman¹, and F. Torres-Ruiz¹

¹Dipartimento di Studi Politici e Sociali, Università di Salerno, Italy

The talk focuses on a general growth curve able to unify the classical cases of Malthusian, Richards, Gompertz, Logistic and several their generalizations. Here two stochastic genralizations of the deterministic growth are obtained by introducing a multiplicative and an additive noise, respectively, to the deterministic equation. We show that the resulting processes are lognormal and gaussian respectively, and their mean is the deterministic trend of the curve. Since the distributions of the two processes are known, the problem of estimating the parameters of the model is analyzed by means of the maximum likelihood method. Further, due to the parametric structure of the processes, the resulting systems of equations are quite complex and numerical solutions have to be searched. In particular estimating procedures for the involved parameters make use of specific metaheuristic algorithms. A second focus of the talk is the first passage and of passage of processes across suitable boundaries and the first exit time of the process throgh a region. In particular, by suitable choice of parameters, we show that the thresholds can be expressed as a percentage of the average size of the population and this is often of interest in applications in population growth. For such cases we provide explicit solutions for the probability density function of the random variable First Passage Time and of the First Exit Time.

Keywords: Inference in diffusion processes, First Passage Time, First Exit Time.

Critical Analysis of ENT Statistical Battery

Elena Almaraz Luengo

Department of Statistic and Operational Research, Faculty of Mathematical Science, Universidad Complutense de Madrid (UCM), Spain

In many areas of knowledge, it is necessary to work with sequences of random numbers or at least that behave as such (pseudo-random numbers). There are different sequence generators among which we can highlight the true random number generators (TRNGs) and the pseudo-random numbers generators (PRNGs). Randomness testing is a crucial instrument for evaluating the quality of random and pseudo-random number generators. There is a wide variety of tests designed for this purpose. These tests are grouped in different sets called suites or batteries. One of the best-known batteries is ENT (from Entropy), which provides four measures (Entropy, Arithmetic Mean, Monte Carlo estimation of π and Serial Correlation coefficient) and the value of a statistic (of the Chi-square goodness-of-fit test).

In this work a detailed analysis of this battery will be carried out and its vulnerabilities will be shown in order to present improvements in its design.

Keywords: Cryptography, Hypothesis testing, Pseudo-random numbers, Random numbers, Statistical test suite.
Food Waste Comparative Statistical Analysis

Panagiotis Andreopoulos¹, Fragkiskos G. Bersimis², Kleomenis Kalogeropoulos³, Alexandra Tragaki⁴, Katerina Synani⁵, and Konstantinos Abeliotis⁶

¹Department of Geography, Harokopio University, Athens, Greece
²Department of Informatics and Telematics, Harokopio University, Tavros, Greece
³Department of Surveying and Geoinformatics Engineering, University of West Attica, Egaleo, Greece
⁴Department of Geography, Harokopio University, Athens, Greece
⁵Department of Geography, Harokopio University, Athens, Greece
⁶Department of Economics & Sustainable Development, Harokopio University, Athens, Greece

The increasing amount of food waste and its reflection on the consequences entailed, are points of interest for the academic community and beyond. Several studies have focus on measuring quantities, as it is not possible to take prevention practices for something that has not been measured. It is understood that knowledge of the total quantity of food waste is not sufficient and requires more detailed knowledge of the quantities by food waste stream. The present study aims at the interregional comparison of data obtained from previous academic studies. Specific surveys were selected, which provide quantitative data on specific food waste flows at a household level. Four European countries: Greece, the Netherlands, Spain, Italy, as well as for the European Union as a whole were selected. Statistical tools (Pareto analysis and t-test) compare both countries and the average prices per category of food waste for the whole of the European Union. The results showed that the eating habits of the people also affect the type of waste produced, as expected. This information can be a key tool both locally and collectively to take appropriate measures to mitigate food waste generation.

Keywords: Food waste, Compare, Data analysis, Pareto analysis.

Advances in Forecasting and Optimally Designing Patient Enrolment in Multicenter Clinical Trials under Various Restrictions

Vladimir Anisimov¹ and Matthew Austin²

¹Data Science, Center for Design & Analysis, Amgen, London, U.K.
²Data Science, Center for Design & Analysis, Amgen, Thousand Oaks, CA, U.S.

Forecasting and optimal designing patient enrolment is among the greatest challenges that the clinical research enterprise faces today, as an inefficient enrolment can be a major cause of clinical trials delays. The talk is describing further innovative developments in this area. The underlying technique uses a Poisson-gamma enrolment model developed by Anisimov & Fedorov in the previous publications and is extended here to analytic modelling of the enrolment on country/region level. For this purpose, a new analytic technique based on the approximation of the enrolment processes in countries by Poisson-gamma processes with aggregated parameters is developed. Another innovative direction is the development of the analytic technique for modelling and forecasting the restricted enrolment (enrolment caps in countries).

Some discussion on using historic trials for better prediction of the enrolment rates in the new trials is provided. These results are used for solving the problem of the optimal cost-efficient enrolment design: find an optimal allocation of sites/countries that
minimizes the global trial cost given that the probability to reach an enrolment target in time is no less than a given probability. Different techniques to find an optimal solution for high dimensional optimization problem for the cases of unrestricted and restricted enrolment and for a small and large number of countries that use non-linear constraint optimization techniques and also evolution algorithms are discussed.

**Keywords:** Patient enrolment, Poisson-gamma model, Forecasting enrolment, Restricted enrolment, Optimal enrolment design.

**High Performance Computing in the Digital Economy**

Emmanuel Apergis$^1$ and Andreas Markoulakis$^2$

$^1$Huddersfield Business School, University of Huddersfield, U.K.
$^2$University of Warwick, U.K.

Quantum computers had been prophesised as instruments for the attainment of even greater heights of scientific achievement. Quantum performance computing is aspiring to spur the development of advanced airframes for space exploration in computer-aided design by introducing digital-twin imaging and replacing wind tunnels, solving aerodynamic problems, expanding “what if” statements by expanding on finite element analysis (necessary for deformation, stress and physical responses tests) and solving environmental issues (e.g., predicting natural disasters and climate events) to pharmaceutical research (e.g., vaccines for upcoming pandemics), seismic exploration by improved signal-to-noise ratios and data analysis of acoustic impedance prior to drilling recommendation using immersive three-dimensional depth imaging, solving the Schrodinger equations faster in the chemical industry, where assumptions question the validity and accuracy of the calculations in molecular dynamics (interactions between atoms in the molecule) and promise the introduction of new military doctrines (e.g., for multi-domain operations — land, air, sea, space and cyber) following the war in Ukraine. Using a system GMM for dynamic panel methodology we aim to capture the role of high-computing capacity on scientific development and progress. The methodology will help us tackle issues of endogeneity and cross-country autocorrelations. Building on a dataset ranging from 1993 till 2023, the data aspire to provide us with unexplored findings aimed at shedding useful light into an area with vast opportunities in artificial intelligence and cybernetic prioritisation.

**An International Comparison of the Impact of Covid-19 on Mortality Rates using Graduation Techniques**

Giovanna Apicella$^1$, Eliseo Navarro$^2$, Pilar Requena$^2$, and Marilena Sibillo$^3$

$^1$Department of Economics and Statistics, University of Udine, Udine, Italy
$^2$Department of Economics and Business, Universidad de Alcalá, Madrid, Spain
$^3$Department of Economics and Statistics, University of Salerno, Fisciano, Italy

There have been several studies that use different biometric indicators to measure the effect of Covid-19 on population mortality such as life expectancy or the number of years of life lost due to the pandemic. Usually, these studies show a dependence of mortality on age and sex indicating that the older the population the greater the impact of Covid-19 and also a greater impact of the pandemic on the male population with respect to women. We apply mortality curve graduation techniques to estimate
mortality rates in order to eliminate the random fluctuations of a stochastic phenomenon such as the number of deaths during a given period of time. This procedure makes it possible to obtain an age by age picture of the changes experienced by mortality rates from the period 2018-19 to 2020 for different European countries. The availability of graduated mortality rates age by age and distinguishing by sex, has revealed mortality patterns that would otherwise have gone unnoticed. The use of graduation techniques allows a detailed analysis of the behavior of mortality rates showing complex changes in some tranches of the mortality curve due to the impact of Covid-19 and the measures implemented to fight the pandemic. Likewise, the fact of focusing on relative changes in mortality rates instead of the number of deaths or the life expectancy reveals different patterns of the impact of the pandemic on the main European countries.

Keywords: Covid-19, graduation techniques, mortality rate.

Longevity Comparison by Gender: Exploring the Future through an Evidence-based Approach

Giovanna Apicella¹, Emilia Di Lorenzo², Giulia Magni³, and Marilena Sibillo⁴

¹Department of Economics and Statistics, University of Udine, Italy
²Department of Economic and Statistical Sciences, University of Naples Federico II, Italy
³University of Milan, Milan (UNIMI), Italy
⁴Department of Economics and Statistics, University of Salerno, Italy

Our demographic study provides a detailed picture of the short-term, weekly, mortality fluctuations charactering females and males over more than a decade. We follow an evidence-based approach, since we use data from the Human Mortality Database to detect stylized empirical evidence about the behaviour of mortality, during normal times and in the most recent time span, highly affected by the COVID-19 pandemic disease. Our study relies on time-series mortality data collected at a finer scale than traditionally done in the actuarial literature and encompasses different age groups, countries, and gender. The empirical evidence represents the starting point for exploring the future mortality patterns. Our quantitative analysis, namely based on stochastic mortality modelling, indeed, answers the question whether accounting for higher frequency mortality data delivers more reliable mortality projections, namely projections that approximate more closely the realized mortality phenomenon.

Keywords: Gender gap, Survival forecasting, Backtesting methods.

The Functional Clustering of the Mortality Gender Gap: A Multi-country Analysis

Giovanna Apicella¹, Emilia Lorenzo², Gabriella Piscopo², and Marilena Sibillo³

¹Department of Economics and Statistics, University of Udine, Italy
²University of Naples Federico II Di, Naples, Italy
³Department of Economics and Statistics, University of Salerno, Italy

Demographic literature is rich of empirical analyzes showing how women have historically experienced lower mortality rates than men. In this paper, we consider a
measure of the gender gap in mortality rates, the Gender Gap Ratio, across a wide range of populations collected in the Human Mortality Database. With the aim of highlighting similarities and differences between the countries considered, we apply a functional clustering method to the multivariate time series of Gender Gap. We reconstruct the functional form of the trends from the available discrete observations and derive the curves through non-parametric smoothing. Results for sixty years olds from 1965 to 2014 are presented and discussed.

**Keywords:** Functional Clustering, Gender Gap Ratio, multi-country mortality trend.

---

**Hotel Sales during COVID-19: Evidence from the US**

Alexandros Apostolakis¹, Christos Floros², Evanthia Zervoudi², and Dimitrios Vortelinos²

¹Department of Business Administration and Tourism, Hellenic Mediterranean University, Heraklion, Greece
²Department of Accounting and Finance, Hellenic Mediterranean University, Heraklion, Greece

The objective of this study is to examine the hotel sales in the US during COVID-19 using recent (quarterly) data. We consider several variables such as: sales price, number of rooms as well as the LARC score. Our aim is to see the effect of location and rooms on the price; further, we test if there is any economic impact. The results are recommended to both the real estate and financial managers dealing with hotel sales.

**Keywords:** Hotel sales, location, rooms, LARC score, economics, US, COVID-19.

---

**fdalomocca: An R Package for Model-based Clustering for Functional Data with Covariates**

Per Arqvist¹, Sara Sjöstedt de Luna¹, and Natalya Pya Arqvist¹

¹Department of Mathematics and Mathematical Sciences, Umeå University, Umeå, Sweden

Global concern about climate change and the impact of human activities on the environment and global heating has surged for knowledge of the climate variations over the last centuries and millennia. Varved lake sediment has the potential to play an important role in understanding past climate with its inherent annual time resolution and within-year seasonal patterns. This paper presents an R package fdalomocca that provides routines for model-based functional cluster analysis for functional data with optional covariates. The idea is to cluster functional subjects (often called functional objects) into homogenous groups by using spline smoothers (for functional data) together with scalar covariates. The spline coefficients and the covariates are modelled as a multivariate Gaussian mixture model, where the number of mixtures corresponds to the number of clusters. The parameters of the model are estimated by maximizing the observed mixture likelihood via an EM algorithm. The clustering method is used to analyze annual lake sediment from lake Kassjo’n (Northern Sweden) which cover more than 6400 years and can be seen as historical records of weather and climate.

**Keywords:** Clustering, Functional data analysis, Mixed effect, Climate, EM-algorithm.
Applications of Text Mining and Natural Language Processing in Health

G. Arvanitopoulos, P. Economou$^2$, and S. Bersimis$^3$

$^2$Department of Civil Engineering, University of Patras, Greece
$^3$Department of Business Administration, University of Piraeus, Greece

As technology evolves rapidly, and big data keeps producing more and more, the need for artificial intelligence in the health domain is becoming greater and greater. The current paper aims to cover a wide range of applications of text mining and natural language processing in healthcare from the recent past. The applications have been categorized based on the tasks they aim to solve and in the methods that are used for the implementation of those tasks.

Keywords: Conference, CMSIM, ASMDA, Text mining, healthcare, natural language processing.

On the Uniform Ergodicity of a Class of Semi-Markov Processes

Giacomo Ascione

Scuola Superiore Meridionale, Napoli, Italy

Consider a Continuous-Time Markov Chain (CTMC) with finite state space that is uniformly ergodic and whose generator admits nontrivial negative eigenvalues. Then, the Total Variation (TV) distance between the distribution at a certain time and the limit distribution decays exponentially fast with a rate dependent only on the first (in absolute value) nontrivial eigenvalue. This behavior completely changes in the Semi-Markov context. Here, we will consider a class of Semi-Markov processes constructed by time-changing the aforementioned CTMC by means of the inverse of an independent subordinator. Precisely, the limit distribution remains unchanged, but we can identify (at least) three different asymptotic behaviors: in the first case, we observe a non-exponential decay that depends on all the eigenvalues at the same time; in the second case, we still observe an exponential decay, but the rate is lower than the first nontrivial eigenvalue (and all the eigenfunctions play a role); in the third one, we obtain again an exponential decay whose rate depends on the first nontrivial eigenvalue. Each behavior is observed by means of a suitable class of subordinators, in terms of the properties of their Laplace exponent. These results underline some unexpected asymptotic behaviors for Semi-Markov processes, as, for instance, in the fractional evolution of random walks on a graph, where the algebraic connectivity was expected to play a fundamental role.

Keywords: Subordinators, Semi-Markov processes, ergodicity in total variation, Continuous-Time Markov Chains.
The Use of the Refined Descriptive Sampling to Reduce the Estimation Variance of the Survival Function

Mahdia Azzal¹ and Megdouda Ourbih-Tari²
¹UMBB, Boumerdès, Algérie
²Centre universitaire de Tipaza, Tipaza, Algérie
²Laboratoire de Mathématiques Appliquées, Faculté de Sciences exactes, Université de Bejaia, Bejaia, Algérie

In this work, we will introduce a methodology to reduce the variance in non-parametric estimation of the survival function in the case of right censored data. The Kaplan Meier (KM) and Fleming Harrington (FH) estimators are selected for such work. Therefore, we will use the Refined Descriptive Sampling (RDS) by developing an efficient algorithm that we call Non-parametric Adaptive RDS (NARDS). The obtained results show that the RDS method reduces the estimation variance of KM and FH estimators.

Optimal Stopping of Gauss-Markov Bridges via Time-space Transformations

Abel Azze, Bernardo D’Auria¹, and Eduardo García-Portugúes
University of Padova, Padova, Italy

We study the optimal stopping problem of a general Gauss Markov bridge, having as objective function its expected value. This model is non-homogeneous and shows unbounded drift near the expiration date and includes the Brownian Bridge and the Ornstein-Uhlenbeck as special cases. We show that a feasible way to deal with the optimal stopping problem is to use a time-space transformation that allows to reduce it to an infinite horizon problem about a Brownian motion. The analysis includes a numerical study that allows to visualise the optimal stopping barrier for specific cases. Keywords: Optimal stopping, Gauss – Markov Bridge, Time-inhomogeneous.

Mixedupness of children’s productions of consonant clusters and how to measure it

Elena Babatsouli
University of Louisiana at Lafayette

During speech development, children’s productions of consonant clusters undergo several stages: deletion, reduction, vowel epenthesis, substitution, and adult-like production. These stages, however, are not exclusive of each other; one or more mix up. In the present study, the mixedupness of cluster stages is measured by adapting Shannon entropy of information theory in which the probability of each cluster stage that is present in the productions at a specific age is included. The entropy of cluster stages is shown to be bounded above and below by the proportion of whole clusters correctness. The measure is applied to both specific consonant clusters as well as classes of consonant clusters. Detailed examination is made for the developmental productions of a female child’s consonant clusters from age two years and six months to age four. In addition, cluster productions of other children with typical and atypical development are also utilized. Results reveal that the entropy measure captures mixedupness of cluster stages longitudinally in a bell-shaped curve with the highest
entropy value, halfway whole cluster correctness. Conclusively, this study shows that Shannon entropy is a reliable measure to be applied in practice in evaluating children’s productions of consonant clusters complementing other measures that have been used (PCC, PCIC, MCP) in the literature which do not address stage mixedupness. **Keywords:** consonant clusters, children, speech, mixedupness, stages, entropy.

**Analytics of Restaurant Ratings in Crete Island by Greek and Foreigner Users of TripAdvisor**

Yiannis Bakirlis & Yiannis Dimotikalis  
Hellenic Mediterranean University,  
Dept of Management Science & Technology, Agios Nikolaos, Crete, Greece

In this work we compare tourist destinations of Crete Island, in particular, restaurant ratings per Regional Unit from the ratings and reviews written in Greek (by Greeks) and not written in Greek (by foreigners) on the TripAdvisor website. The ratings data was collected from the TripAdvisor website, in total ratings of Greek and foreign TripAdvisor travelers for 400 restaurants in Crete. Hypothesis Test was performed on the 2 samples of data (Greek and foreign) from 100 restaurants for each Regional Unit. A Binomial Distribution was also fitted to the ratings of each restaurant to check the fitting quality of the real ratings from the theoretical Binomial probability distribution. Finally, the conclusions on the differences in evaluations by Greeks and foreigners in the restaurants for the 4 Regional Units of Crete are discussed.  
**Keywords:** Crete Island, TripAdvisor, Restaurants Ratings, Hypothesis Test, Binomial Distribution, Distribution Fitting, Five Stars Rating

**On The Arbitrage in Models Based On Fractional Brownian Motion**

Bilal Barakeh  
Associate Professor, Lebanese University, Faculty of Sciences, Dpt. of Mathematics, Lebanon

In this paper we discuss the arbitrage problem using the fractional Brownian motion (FBM) in the Black-Scholes model. We show that the (FBM) cannot be considered as a volatility but a stochastic trend. Therefore, it is important to add a Brownian motion as a diffusion part. The two cases where long memory results of volatility clustering or of deterministic trends will be treated. Several empirical data show that the stock price distribution has usually properties as “fat tails”, “volatility clustering”, “self-similarity”, “long range dependence” and all these facts contradict the Black-Scholes (BS) assumptions that the natural logarithm of the stock price follows a random walk. Actually, the idea that stock returns could exhibits long range dependence was first suggested by Mandelbrot [1]. Mandelbrot and Von Ness [2] introduce the (FBM), a non-semi martingale Gaussian process which has a long range dependence a dependent increment and a parameter called Hurst parameter (H) commonly used as a measure of market predictability which may result from the dependence of the returns. Peters in [3] suggests that if a stock time series has a high Hurst exponent, then the stock will be lessrisky and noise in the data set will occur. The use of (FBM) in applications to Financewasneglected, since this process is not a semi martingale and this is leading to two major difficulties: The first is in defining a stochastic calculus
relative to the (FBM) and the second is that arbitrage opportunities cannot be excluded in this model. In fact, several methods can be used to construct a stochastic calculus for this process. One can cite the Malliavin calculus [4] which is based on a probabilistic and functional analytic notion of derivation and integration and allows a very efficient stochastic calculus for Gaussian processes which has already been successfully applied to the calibration of mathematical models in finance. As for the second difficulty, arbitrage strategies are defined as strategies which realizes a possibly positive gain with a zero initial capital. But the arbitrage depends on the set of admissible strategies $A$ and on the definition of the stochastic integral. Arbitrage pricing theory is based on a fundamental result known as the Fundamental Theorem of Asset Pricing which shows in full generality that “No free lunch with vanishing risk” is equivalent to the existence of an equivalent local martingale measure. In other terms, there is no arbitrage if and only if there exists a martingale measure under which the discounted prices are martingales. A model in which the log price is described by a fractional Brownian motion cannot be arbitrage-free, that is there exists a strategy $\emptyset \in A$ realizing a possibly non-zero gain by starting from a zero initial capital. Rogers in [5] has formulated the mathematical argument for the existence of arbitrage opportunities for this process. Several attempts have been proposed in order to exclude arbitrage opportunities in the context of non-semimartingale models like fractional Brownian motion. Patrick Cheridito [6] proposed to forbid high frequency trading, and Paolo Guasoni [7] suggests to introduce transaction costs.

References:
Stochastic Processes Associated with Fully Nonlinear Parabolic Equations in Financial Mathematics

Yana Belopolskaya\textsuperscript{1,2} and Andrey Chubatov\textsuperscript{1}
\textsuperscript{1}Sirius University of Science and Technology, Krasnodarsky kraj, Sirius, Russia
\textsuperscript{2}PDMI RAS, St. Petersburg, Russia

Fully nonlinear parabolic equations arise in financial mathematics as mathematical models of option prices for markets with illiquidity and transaction costs as well as in optimization problems. We derive stochastic differential equations (SDE) describing stochastic processes which allow to construct stochastic processes underlying the considered problem and construct its probabilistic representation. To this end two approaches are suggested. The first is based on a possibility to include the original problem into a system of semi linear parabolic equations and applying the approach from the book by Dalecky and Belopolskaya (1990). The second is based on a combination of this approach with the theory of backward SDEs developed by Pardoux and Peng (1990). Note that in the first approach one needs a differential prolongation of higher order than in the second. In our report, we consider an optimization problem of an investment portfolio in a couple of financial market models. We discuss the advantages and drawbacks of these approaches and apply them to derive numerical schemes and obtain numerical solutions. The work is supported by RSF Grant 22-21-00016.

Keywords: Fully Nonlinear PDEs, BSDEs, Financial Markets.

Non Parametric Estimation for Time Varying Integer Valued Processes

Sara Bendjeddou
University of Science and Technology Houari Boumediene, Bab-Ezzouar, Algeria

This paper aims to study the Non parametric Negative Binomial Quasi-Maximum Likelihood Estimation NBQMLE for locally stationary integer valued processes models. In particular, we consider the locally stationary negative binomial integer-valued GARCH (p, q) processes. This estimation is based on a kernel function and the considered kernels are the Epanishnikov kernel, the Uniform kernel and the Gaussian kernel. First, we prove the uniform consistency and asymptotic normality of such kernel based estimator then we confirm the performance of this kernel estimator through applications to real and simulated series.

Keywords: Time varying integer-valued processes, Local stationnarity, Non parametric estimation, Negative Binomial QMLE.
A Non-Parametric Monitoring Procedure for Monitoring Multivariate Processes Based on Convex Hulls

S. Bersimis¹, F. Bersimis², S. Chakraborti³, P. Economou⁴

¹Department of Business Administration, University of Piraeus, Greece
²Department of Statistics and Insurance Science, University of Piraeus, Greece
³Department of Information Systems, Statistics and Management Science, University of Alabama, USA
⁴Department of Civil Engineering, University of Patras, Greece

Multivariate Statistical Process Monitoring (MSPM) is one of the most rapidly developing areas of statistical process control in the last decade due to the enormous number of practical problems involving multiple dependent variables. Hotelling's T2 control chart (CC) is the most well-known tool used to monitor a multivariate process. However, the Hotelling's T2 CC makes strict assumptions about the distribution of the random vector representing the stability of the process, i.e. assumes multivariate normality. Furthermore, today, in the big data era where multiple data streams sources should be monitored such a strict assumption is unrealistic. Thus, robust non-parametric monitoring procedures are of great importance. In this work, we introduce a non-parametric control scheme based on convex hulls. The proposed non-parametric control chart is using bootstrap for estimating the kernel of the multivariate distribution and then appropriate statistics based on convex hull are monitored. The performance of the proposed control chart is very promising.

Keywords: Statistical Process, Monitoring, Convex Hulls.

BSDE Method for Pricing Options under Heston-type Models

Abigail Berta¹, Marko Dimitrov², Achref Bachouch², Christian Ewald¹, and Ying Ni²

¹Department of Mathematics and Mathematical Statistics, University Umeå, Umeå, Sweden
²Division of Mathematics and Physics, Mälardalen University, Västerås, Sweden

The paper’s aim is to price options under Heston-type stochastic volatility models using numerical Forward-Backward Stochastic Differential Equations (FBSDEs) methods. Namely, the focus is on approximating the solution to the corresponding FBSDE using methods based on deep learning techniques. Such problem formulation is crucial for pricing derivatives as it considers market imperfections, and the method allows quick computation in higher dimensions. First, the paper gives FBSDE representations of the option prices under the considered Heston-type models. Then, the FBSDE representation is used to derive a numerical scheme that allows an approximation of the option prices. Finally, a comparison with existing methods is performed.

Keywords: Forward Backward Stochastic Differential Equations (FBSDE), Option Pricing, Deep BSDE Method.
Simulation of a Non-Parametric Regression Estimator with Dependent and Incomplete Data

Siham Bey¹, Zohra Guessoum and Abdelkader Tatachak
¹University of Science and Technology Houari Boumediene, Algiers, Algérie

A proposed nonparametric estimator for the classical regression function is calculated using a simulated sample for dependent and incomplete data. The minimization of the mean squared error makes it possible to find the optimal bandwidth which will be implemented in the calculation of our estimator. The simulation results will be compared to other methods for choosing the bandwidth.

Keywords: Non-parametric regression, Kernel estimator, Truncated-censored data, association, optimal bandwidth.

Semiparametric Statistical Calibration of the Numerical Air Pollution Model Outputs

Marek Brabec
Institute of Computer Science, Czech Academy of Sciences. Praha, Czech Republic

In this paper, we will present an approach to large-scale spatial statistical modeling of selected air pollutants with the special complication of known spatial heterogeneity brought by the patchwork of background and urban areas with substantially different behavior of the spatial field not only in terms of expected values, but also in terms of stochastic part of the model. The statistical model we propose is, in fact, blending two sources of available information, namely i) empirical information coming from air pollution concentration measurements obtained from existing professional monitoring network, and ii) information coming from numerical modeling (based on chemical transport models). The systematic part of the model can then be seen as a calibration of the numerical model output (or, alternatively, as theoretically motivated downscaling of the measured data). Stochastic part of the model acknowledges heteroscedasticity and also vastly different ranges of the pollutant concentration spatial processes in the background and within city limits. We illustrate the procedure of the model identification and estimation of the spatial field of concentrations on a fine grid over the Czech Republic. Furthermore, we show how the structural information about the concentration behavior we obtain from data can be used for several design problems motivated by practical considerations related to measurement network changes (expansion and reduction tasks). The work has been supported by the ARAMIS project SS02030031 of the Technology Agency of the Czech Republic.

Spread of Branching Random Walk with Periodic Branching Sources

Ekaterina Vl. Bulinskaya
Lomonosov Moscow State University, Moscow, Russia

The model of a branching random walk (BRW) with periodic branching sources was studied in a series of papers by M.V.Platonova and K.S.Ryadovkin (see, e.g., [1]). It is a natural extension of BRW with single or several sources of branching. The main feature of these models is a combination of a random walk of a particle on an integer
lattice $\mathbb{Z}^d$ and its splitting at a source of branching. In our present work the number of the branching sources is infinite and they are located periodically on the lattice. Earlier the propagation of the particles population in supercritical BRW with finite number of sources was investigated (see, e.g., [2]). It was proved in [3] that almost surely the particles population spreads asymptotically linearly in time under condition of light tails of the random walk jump. Here our goal is to obtain the similar results for supercritical BRW with periodic branching sources and light tails of the random walk jump. The proposed approach involves certain results for a space-homogeneous BRW ([4]) where the Crump-Mode-Jagers branching processes were employed and a point process described the locations of the offspring.

References:

Keywords: Branching Random Walk, Periodic Branching Sources, Spread of Population, Light Tails, Supercritical Regime.

Stochastic Insurance Models with Investment and Reinsurance

Ekaterina Bulinskaya
Department of Probability Theory, Faculty of Mathematics and Mechanics of Lomonosov Moscow State University, Moscow, Russia

Insurance has a long and interesting history containing 4 periods, see e.g. [1]. The main feature of modern period is employment of sophisticated mathematical tools for investigation of complex systems and frequent use of computers for their numerical analysis. Thus, discrete-time models providing approximation for continuous-time ones became popular in the 21st century. It is worth mentioning that sometimes such models describe more precisely the real-life situations. Hence, we consider several discrete-time insurance models. The aim is to establish the optimal control providing the extremum of a chosen objective function (or risk measure). We use the cost approach and reliability one. In the first case we are interested in the minimization of costs associated with the system functioning or maximization of profit. In the framework of reliability approach the system ruin probability is minimized or the time of uninterrupted work is maximized. Investment, bank loans, dividend strategy and reinsurance are used as controls. The models' stability and the limit behavior of the company surplus are also considered. Probability metrics are used for this purpose. It is explained as well how the obtained results can be extended for application in other domains, such as finance, queuing, reliability, inventory, biology or medicine.

References:

Keywords: Insurance, Investment, Reinsurance.
Some Generalizations and Applications of Rényi’s Theorem

Alexander Bulinski
Faculty of Mathematics and Mechanics of the Lomonosov Moscow State University, Moscow, Russia

The theory of random sums of random variables has various applications, e.g., in queuing, insurance and reliability. Special attention in the literature is paid to geometric sums (i.e., the number of summands follows the geometric law). The famous Rényi theorem describing the convergence of appropriately normalized geometric sums to an exponential law plays the role analogous to the CLT for sums of random variables. Recently in [1] the sharp estimates for the proximity of geometric and related sums distributions to limit laws were established by means of the ideal probability metric of the second order. There are no assumptions that the summands are positive and identically distributed. The development of Stein’s techniques was employed essentially as well as the integration with respect to sign measures. We mention also the introduction of the inverse to the equilibrium transformation of the probability measures. On this way it is possible to provide the new approximation of the Pareto distribution by an exponential one (such result is used in signal processing). We consider also stochastic models described by dependent random variables. For exchangeable random variables an extension of the Rényi theorem is obtained with explicitly indicated limit law. Some other advances are also tackled.

References:

Keywords: Geometric Sums, Stein’s Method, Probability Metrics, Exchangeable Random Variables, Generalization of the Rényi Theorem.

Quantifying Non-Markovianity of Evolution Processes in Open Quantum Systems

Andrey Bulinski
Department of Higher Mathematics, Moscow Institute of Physics and Technology (National Research University), Institutsky Lane, Dolgoprudny, Russia

The concept of (nonisolated) open quantum system interacting with its environment permeates the contemporary research in Physics, Chemistry, Biology and Quantum Information Theory. By analogy with classical stochastic processes, the notion of quantum process related to a system dynamics has emerged and is still subject to scrutinized studies. Despite huge efforts to qualify and quantify a variety of loss of memory effects in open systems dynamics, a universal and comprehensive treatment of Non-Markovian quantum evolution process is absent as yet. Thus, it makes sense to investigate diverse measures of quantum Non-Markovianity for grasping its different aspects. We report on the recent progress in analysis of leak and backflow of information, due to System-Environment partition and accounting for the open dynamics Non-Markovianity, in terms of certain informational characteristics and various distances in quantum states spaces. This permits to distinguish processes by degree of their Non-Markovianity.

Keywords: Open Quantum System, System-Environment Correlations, CP - (non)divisibility, Information Backflow, Quantum Process, State Space, Bounds on Distances.
Prosocial Behaviour of Students in Maltese Schools: A Multilevel Model

Liberato Camilleri
Maths and Physics Building, Faculty of Science, University of Malta, Malta

Prosocial behaviour refers to a range of positive behaviors including positive interactions altruism and behaviors that reduce stereotypes. Prosocial behaviour in schools include comforting a peer, complimenting classmates, helping others with schoolwork, making sure everyone is included, settling disagreements, and sharing. It is the opposite of 'antisocial' behaviour such as bullying, denigrating and cheating. The main objective of this paper is to identify a number of student, class and school-related factors that are significantly related to prosocial behaviour and develop new ways of understanding and enhancing prosocial behaviour in Maltese schools. A proper methodology for analyzing hierarchical structured data where observations are nested within groups is multilevel modelling. This paper presents a three-level random intercept model that accommodates random effect within each level of nesting and examines the contribution of a number of predictors in explaining variations in the prosocial behaviour scores elicited from 5300 students attending primary and secondary state, church and independent schools. The model identifies student related variables, particularly engagement, diagnosis and intervention, as better predictors of prosocial behaviour than class and school related variables.

Modelling ESG greenwashing reputational risks in the insurance market

Maria Carannante¹ and Valeria D'Amato¹
¹Difarma Department, University of Salerno, Italy

In this paper, we propose a metric based on subjective probabilities to evaluate the effects of sustainability reputation on the overestimation or underestimation of a financial or actuarial product. In particular, from the perspective of an investment, we consider the greenwashing phenomenon as a revaluation of its value in terms of loss of value as no longer beneficial of a sustainability premium, as the company or economic sector to which it refers loses value in terms of ESG score. Insurance companies play an important role in the composition of pension funds or in the choice of sectors to be insured as "green", determining an indirect reputational risk resulting from the effective sustainability of the products proposed.

Keywords: ESG, Greenwashing, Reputational Risk, Subjective probability.

A Conditional Embedding Problem for Finite Homogeneous Markov Chains

Philippe Carette¹ and Marie-Anne Guerry²
¹Ghent University, Department of Economics, Ghent, Belgium
²Vrije Universiteit Brussel (VUB), Department of Business Technology and Operations, Brussels, Belgium

The embedding problem of Markov chains is a long standing problem where a given stochastic matrix is examined as the 1-step transition matrix of some continuous-time homogeneous Markov chain (CTHMC). This problem boils down to characterizing the
empirical transition matrix $P$ as the exponential of some matrix $Q$ with all non-negative off-diagonal entries and zero row-sums, called an intensity matrix. If such a $Q$ exists, $P$ is said to be embeddable. It turns out that the embedding problem is a formidable one in a number of respects. First, $P$ may not be embeddable. In that case, a regularization algorithm can be used to find an intensity matrix $Q$ for which $||P - \exp(Q)||$ is minimized. Next, no embeddability criteria in terms of the matrix elements, which are easily verifiable in practice, seem at hand when the number of states exceeds 3. Lastly, for an embeddable $P$, there may not be a unique solution to the equation $\exp(Q) = P$ in the set of intensity matrices. The identification aspect of the embedding problem deals with the selection of the suitable intensity matrix reflecting the nature of the system under study. We propose the conditional embedding approach where the empirical 1-step transition matrix $P$ corresponds with the conditional 1-step transition matrix of the CTHMC given the event that at most one jump has occurred during a time interval of unit length. For a Markov model the unit time interval can be defined in such a way that the empirical 1-step transition matrix meets this condition. Moreover, this condition is inherent in some applications. For example, in credit rating migration models the credit ratings are typically based on slowly varying characteristics, such that they do not tend to change more than once within the baseline time interval (e.g. a quarter). We found that, regardless the number of states, exactly one intensity matrix solves this conditional embedding problem when $P_{ii} > 0$ for all $i$. Our approach results in an easy embeddability criterium and does not require identification neither regularization.

**Keywords:** Embedding problem; Markov chain; transition matrix.

**Media Bias and Polarization through the Lens of a Markov Switching Latent Space Model**

Roberto Casarin$^1$, Antonio Peruzzi$^2$, and Mark Steel$^3$

$^{1,2}$Department of Economics, Università Ca’ Foscari, Venezia, Italy
$^3$University of Warwick, Coventry, UK

The news consumption landscape has changed quickly in recent years. The advent of web 2.0 and social media allowed to increase news production and speed up news diffusion at low costs. This new environment brought -- along with many benefits -- also some risks. News outlets are more than ever incentivized to provide their audience with slanted news, while the intrinsic hemophilic nature of online social media may exacerbate polarized opinions. For this reason, developing tools that offer insights about the degree of media bias and polarization in a given information environment may be beneficial for consumers. In this work, we propose a new dynamic latent space model for timevarying online audience duplication networks which exploits social media content to provide a way of measuring news outlets’ political leaning and polarization regimes. In our model, the latent positions of news outlets have a proper interpretation in terms of political ideology by means of an observable proxy. The work has the twofold aim of making advancements both concerning LS models and concerning the estimation of media ideology and social-media polarization. The developed model is applied to a set of data on the online activity of national and local news outlets from four European countries (France, Germany, Italy and Spain) in the years 2015 and 2016. We find evidence of strong positive correlation between our media slant measure and a well-grounded external source of media bias (~75% correlation with the PEW Research index). We further provide insights on the in-platform polarization regimes across the four countries considered.
Keywords: Latent Space Models, Bayesian Inference, Markov-Switching, Political Learning, Polarization.

Exploring Measurement Invariance: Political Trust in National Institutions across Social Groups

Anastasia Charalampi¹, Catherine Michalopoulou², and Clive Richardson³

¹,²Department of Social Policy, Panteion University of Social and Political Sciences, Athens, Greece
³Department of Economic and Regional Development, Panteion University of Social and Political Sciences, Athens, Greece

A prerequisite for meaningful cross-national comparisons of scale scores is establishing the scale’s measurement invariance. In parallel, scaling theory presupposes that the structure of the construct (scale) is determined and its psychometric properties are assessed before its application. In this study, the psychometric properties and measurement invariance of a five-item scale measuring political trust in national institutions from the 2018 European Social Survey (ESS) were investigated for four European countries: France, Poland, Spain and the UK. First, factor analytic theory indicated that only a unidimensional model could be tested. Therefore, only Confirmatory Factor Analysis (CFA) was applied on the full samples of all countries. Then, based on reliable and valid scales of acceptable or adequate model fit, multiple-groups CFA was performed for testing the scales’ measurement invariance across countries. Measurement invariance was confirmed for the social groups defined by the level of educational attainment and the left/right self-placement scales for a combined sample of those four countries. Based on the mean scores, the findings showed that the higher the level of educational attainment, the greater the political trust in national institutions, with the exception of Poland, and higher levels of political trust in national institutions were observed for those inclined to the right end of the left/right self-placement scale, with the exception of France. This work could be extended to cover more countries of this or other Rounds of the ESS and other socio-demographic groups.

Keywords: measurement invariance, multiple-groups Confirmatory Factor Analysis, European Social Survey, political trust, social groups.

A Methodology for Evaluating Human Resources in the Maritime Industry Using Analytics

Paraskevas-Marios Chavatzopoulos¹, Pavlos Karamitros², and Ioannis Theotokas³

¹,²,³University of Piraeus, Greece

In this paper, a new methodology is presented in order to assess and reorganize the personnel in a shipping company using analytics. The main feature is the use of both quantitative and qualitative data; quantitative data can be easily recovered from a company’s database whereas qualitative data are more difficult to be documented. For that reason, the authors propose the use of self-assessment questionnaires and non-technical skills’ assessment forms. This large number of data tuples makes analytics an invaluable and necessary part of the process so as to be able to correlate the characteristics of the employees. The paper examines the two distinct categories of employees of a shipping company; the employees that work in the offices and those
who work aboard the ships. There is a big gap between those two categories that have to be addressed resulting in different forms, questionnaires, and necessary data needed to complete the puzzle. Last, but not least, a different approach for examining the accident is employed; the accident is not only happening at sea but also at the office when an employee takes a wrong decision which leads to loss of freight or when a machine breaks down due to lack of supervision and the ship cannot carry out her mission. 

**Keywords:** Human resources analytics, Maritime hr, Analytics.

---

**Neural Network for First-hitting-time Based Threshold Regression Model**

Yiming Chen, Paul Smith, Mei-Ling Ting Lee

*University of Maryland, College Park, MD, USA*

In this paper, we propose a neural network expansion (TRNN) of the first hitting time (FHT) based threshold regression (TR) model. The FHT-based model conceptualizes a random process for subjects’ latent health status. The time-to-event outcome is modeled as the first time the random process hits a pre-specified threshold. Due to the flexibility of neural networks, TRNN is capable of handling high-dimensional inputs while providing clinically meaningful interpretations. The proposed TRNN can further be used in causal survival analysis such as G-computation. Due to the algorithm's robustness, more efficient causal estimations are expected. The performance of our proposed neural network algorithm is illustrated using real data from a large observational study.

**Keywords:** Causal inference, G computation, High-dimensional data, Inverse-Gaussian distribution, Latent health process, Wiener processes.

---

**Health State Function and Red Herring Hypothesis**

Chia Ngee Choon¹ and Lim Kia Yee²

¹National University of Singapore, Faculty of Arts & Social Sciences, Department of Economics, Singapore

²PhD Student, National University of Singapore, Faculty of Arts & Social Sciences, Department of Economics, Singapore

In health economics literature, the red herring hypothesis states that increasing proximity to death is the key driver of healthcare expenditure increases, rather than ageing itself. This paper seeks to test two different versions of the red herring hypothesis: a weak version in which proximity to death has a significant impact on healthcare expenditure, as well as a strong version in which ageing has no significant impact on healthcare expenditure after controlling for proximity to death. We motivate the econometric strategy used in Zweifel et al. (2004) by modifying Janssen and Skiadas (1995)’s stochastic health state model, and in the process derive an alternative method of estimating stochastic health state functions from a panel dataset. We then test these versions of the red herring hypothesis using data from the Singapore Retirement and Health Survey, and in doing so are able to reaffirm the weak red herring hypothesis. However, evidence for the strong red herring hypothesis remains inconclusive.
Multi-way MANOVA and General Linear Model
Likelihood Ratio Tests for the High-Dimensional Case

Carlos A. Coelho

Mathematics Department and NOVA Math (Research Center in Applied Mathematics)
NOVA School of Science and Technology, NOVA University of Lisbon, Portugal

A Likelihood ratio test (LRT) to test the fit of the multi-way MANOVA model in the case of high-dimensionality, that is, when the number of observations is smaller than the number of variables under analysis, is introduced, as well as LRTs to test for the significance of factors and interactions. These tests are actually introduced as particular cases of LRTs for the test of independence of two groups of variables in a high-dimensional setting, and LRTs to test between an original multivariate model and one of its submodels, in a high-dimensional setting. The exact and asymptotic distributions for the test statistics are derived and it is shown how a quite intuitive and very functional Normal asymptotic distribution may be quite easily derived. Also, it is shown that in a number of situations the test statistics do have finite closed form expressions for their probability density and cumulative distribution functions. Moreover, through the implementation of quite extensive numerical simulations it is shown that the tests proposed exhibit a better preservation of the Type I error rates than other tests, even the ones more recently developed, as well as they also show values of power which are much larger than those exhibited by other available tests.

Improved Shewhart-type Control Charts based on Weak Runs in Multistate Trials

Spiros D. Dafnis\textsuperscript{1}, Theodoros Perdikis\textsuperscript{2}, and Markos V. Koutras\textsuperscript{1}

\textsuperscript{1}Department of Statistics and Insurance Science, University of Piraeus, Piraeus, Greece
\textsuperscript{2}Department of Statistics, Athens University of Economics and Business, Athens, Greece

In this work we introduce two new classes of Shewhart-type control charts with supplementary runs rules using a three or four region partition of the observed characteristic's range. The study of the control charts' characteristics (run length distribution, average run length) is carried over by considering an extension of the concept of r-weak runs (which was recently introduced for sequences of binary trials) under a multistate framework. Our numerical experimentation provides useful hints for the implementation of the new control charts by a practitioner and reveals that the proposed charts significantly improve the detection power (as compared to other existing charts) especially for small shifts in the process mean.

\textbf{Acknowledgment:} This research is co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme “Human Resources Development, Education and Lifelong Learning 2014-2020” in the context of the project “Innovative Titanium Nanoparticles for Development of Autocleaning and Auto antibacterial Application” (MIS 5131364).

\textbf{Keywords:} Average run length, Markov chain, Runs rules, Waiting time distribution.
The Cost of Advertising: Real Option Estimation through Markov Chains

Guglielmo D’Amico¹, Riccardo De Blasis², and Veronica Vigna³
¹Department of Economics, University “G. d’Annunzio” of Chieti-Pescara, Italy
²Department of Management, Marche Polytechnic University, Ancona, Italy
³Department of Neuroscienze, Imaging e Scienze Cliniche, University “G. d’Annunzio” of Chieti-Pescara, Italy

In this paper, we propose the valuation of a real option in the telecommunications industry. According to the probabilistic present worth approach (Carmichael, 2016), we estimate the value of a contract between a television network and a company willing to advertise its business on this network. We assume that the value of the contract depends on a time-dependent variable, i.e., the number of viewers tuned into the network, which behaves like a Markov process. After discretizing and converting this number into a monetary value through a specific function, we compute the nth-order moment of the payoff function, measured as discounted earnings minus the strike price. Finally, we apply the proposed model to the real television ratings data.

References:

Dividend Based Stock Risk Measures: A Markov Chain Approach

Guglielmo D’Amico¹ and Riccardo De Blasis²
¹Department of Economics, University “G. d’Annunzio” of Chieti-Pescara, Italy
²Department of Management, Marche Polytechnic University, Ancona, Italy

In this paper we propose further advancements in the Markov chain stock model. The model performs the stock valuation according to the well-known relation existing between the prices and the dividends (see, e.g., Gordon (1962) and subsequent literature). The fundamental valuation analysis in a random setting states that \( p(k) = E_k[P(k)] \) obeys the equation

\[
p(k) = \frac{E_k[D(k+1) + p(k+1)]}{r},
\]

where \( E_k \) is the conditional expectation given the information available at time \( k \), \( r \) is one plus the required rate of return on the stock, \( D(k+1) = G(k+1)D(k) \), with the dividend growth factor, \( \{G(k)\} \), described by a Markov chain.

We provide a formula for the nth-order moment of the fundamental price process with transversality conditions that avoid the presence of speculative bubbles. Moreover, we determine the specific probability function of the price process related to those moments applying the principle of maximum entropy. It means that we find the probability distribution which maximises the entropy under the constraint of the first n-order moments. Therefore, we are able to compute the standard risk measures, such as value-at-risk, expected shortfall and other quantile-based measures. The advantage of this methodology is that the computed risk measures are based on the financial fundamentals, i.e., the dividend process. Thus, any difference existing between the proposed risk measures and the corresponding ones computed only using the market prices, highlights that the latter might not reflect entirely the effect of the fundamentals on the risk, either overestimating or underestimating the risk. Finally, we apply the proposed measures to the real dividend data.

Keywords: Markov chain, moment problem, fundamental analysis.
Drawdown-Based Measures for Financial Volumes Via Semi-Markov Model

Guglielmo D’Amico¹, Bice Di Basilio², and Filippo Petroni³
¹,²Department of Economics, University of Chieti-Pescara, Pescara, Italy
³Department of Management, Marche Polytechnic University, Ancona, Italy

In this research paper we assess the suitability of weighted indexed semi-Markov chains (WISMC) to analyze drawdown-based risk measures on high-frequency financial volumes. The considered risk measures are a generalization of the drawdown of fixed level, the time to crash, the speed of crash, the recovery time, and the speed of recovery, already introduced in previous work [1]. The novelty is the insertion of a general starting time, not necessarily coinciding with the beginning time of the trading day, which permits the consideration of sub-intervals of the trading day to calculate all the measures under study. In this way, it is not regarded the sudden initial fluctuations in financial volumes, probably due to both overlaps with the European Stock Exchanges and huge numbers of transactions [2]. The implementation is realized on high-frequency financial volumes of three different stocks listed on the Nasdaq stock Exchange in the period April-September 2022. We chose Tesla, Apple, and Netflix assets which belong to the Automotive and Renewable, Information Technology, and Technology & Entertainment sector, respectively. The results obtained by implementing the WISMC model are compared with those based on the real data and also with those achieved by EGARCH and GJR models. All the investigation is realized considering right censoring which affected the drawdown of fixed level, the time to crash, and the recovery time.

References:

Keywords: Drawdown-based measures, High-Frequency Financial Data, WISMC model, Right censoring.

Research on the Degree of Use of New Technologies by Households and their Mempers. E-Commerce and the Security and Protection of their Personal Data

Stamatina Daskalaki¹ and George Matalliotakis²
¹,²Hellenic Mediterranean University

Technology has become an integral part of our lives, and its influence on the family is undeniable. From desktops to smartphones, the digital age has offered a variety of tools and resources that have changed the way people see the world. The use of new technologies within households largely depends on the characteristics of each family member. Households that actively use technologies are more likely to be receptive to new technologies, and it has been observed that new technologies are introduced more easily when the level of use is high. Younger Internet users are more likely to say that the Internet has had a positive impact on society than older Americans who use the Internet. The use of new technology has created a dramatic change in the way households operate and the way family members interact with each other. As technology has advanced and become part of everyday life, it has had a huge impact on family dynamics. E-commerce is a fast-growing business sector that enables
consumers and businesses to transact online. It has become a popular way to buy, sell and exchange goods and services with the help of digital technology. The emergence of e-commerce has revolutionized the way households shop for goods and services. E-commerce is a business model that allows businesses and consumers to shop or sell things online. Retail e-commerce apps use online marketing techniques to attract customers to their platforms. Ecommerce includes a wide variety of activities, including B2B, B2C, C2C, and C2B transactions. It allows buyers and sellers to make transactions with ease, without the need for a physical presence. The three most common ecommerce business models are business-to-business (B2B), business-to-consumer (B2C), and consumer-to-consumer. The increasing share of e-commerce transactions may also be due to the increasing number, quality and affordability of products available online versus in-person shopping. With the emergence of online shopping giants like Amazon and Alibaba, traditional retailers have had to adjust their business models to stay competitive. Smartphones and other devices have never made it easier for customers to connect with businesses and make purchases. Now, more and more households are taking advantage of the convenience and affordability of e-commerce technology to buy goods. Therefore, it is clear that households are increasingly turning to e-commerce technology to buy goods.

General information measures for loss models and survival models involving truncated and censored random variables

Silvia Dedu¹, Vasile Preda², and Muhammad Sheraž³

¹ Bucharest University of Economic Studies and “Costin C. Țiriac” National Institute of Economic Research, Bucharest, Romania
² “Gheorghe Mihoc-Caius Iacob” Institute of Mathematical Statistics and Applied Mathematics and “Costin C. Țiriac” National Institute of Economic Research and University of Bucharest, Romania
³ Institute of Business Administration Karachi, Pakistan

Risk assessment is an important topic in various fields, since it allows designing the optimal strategy in real-world problems. Information measures can be used to evaluate the uncertainty degree corresponding to phenomena described by random variables. This paper aims to develop an entropy-based approach to risk assessment for actuarial models involving truncated and censored random variables. The research is focused on studying the effect of partial insurance models, such as inflation, truncation and censoring from above and truncation and censoring from below upon the entropy of losses using general information measures. Analytic expressions for the per-payment and per-loss entropies are derived, and the relationship between these measures is investigated. The combined effect of a deductible and a policy limit is also studied. By considering the residual and past entropies for survival models, the entropies of losses corresponding to the proportional hazard and proportional reversed hazard models are derived. The properties of the resulting entropies, such as residual loss entropy and past loss entropy, are investigated as a result of using a deductible or a policy limit. The entropy of losses is computed and illustrated for different distribution models. Numerical results on real data are provided. The detected behavior proves that the general information measures approach for actuarial models involving truncated and censored random variables provides a new and relevant perspective, since it allows a higher flexibility and improves the modeling accuracy. Acknowledgements: This work was supported by a grant of the Romanian Ministry of Education and Research, CNCS–UEFISCEDI, project number PN-III-P4-ID-PCE-2020-1112, within PNCDI III. Keywords: Information measures; Risk assessment; Actuarial models.
Energy-Sharing Policy: Influence of Demand

Riccardo De Blasis¹, Graziella Pacelli¹, and Salvatore Vergine¹
¹Department of Management, Marche Polytechnic University, Ancona, Italy

Energy management has gradually become a complex problem because of the intermittent nature of renewable power sources, such as photovoltaic power, and the large use of energy storage systems. The latter are able to store energy when exceeding power generation happens and supply energy during low generation time. In this context, a fundamental role is played by the demand power characteristics which strongly influence the benefits brought by this energy management scheme. Therefore, this paper investigates the influence of the variability of power demand from different users. In particular, the demand power is either simulated using the stochastic model presented in [1] or analyzed through real data. Moreover, our model investigates the energy demand when photovoltaic generation and storage systems are considered. The obtained results show how the variability of the demand affects the energy-sharing policies.

References:

Keywords: Energy Community, Demand variability, Energy storage.

Applying Entropy Criterion to Input Allocation: Fertilizer Cost Estimates for European Countries

Dominique Desbois
UMR Economie publique, INRAE-AgroParisTech, Paris-Saclay University

The decision to adopt one or another of the sustainable land management alternatives should not be based solely on their respective benefits in terms of climate change mitigation but also on the performances of the productive systems used by farm holdings, assessing their environmental impacts through the cost of specific resources used. This communication uses the entropy criterion in order to estimate the fertilizer costs of specific productions in agriculture, as a replacement proxy for internal soil erosion costs. After recalling the conceptual framework of the estimation of agricultural production costs (Desbois, 2015), we present the empirical data model, the entropy regression approach and the tools used to obtain typologies of European countries based on the conditional distributions of fertilizer cost empirical estimates. The comparative analysis of econometric results for main products between European countries illustrates the relevance of this approach for international comparisons based on their input specific productivity.

Following Léon et al (1999), we use the generalized maximum entropy (GME) method (Gollan, Judge and Miller, 1996) to estimate the following empirical model:

\[ x^i_l = \sum_{k=1}^{K} \beta^i_k y^k_l + \epsilon^i_l \]  

for input \( i \) and countries \( l = 1, \ldots, L \) where \( x^i_l \) denotes country \( l \)'s expenditure on input \( i \) and \( y^k_l \) denotes the value of good \( k \) produced by country \( l \), the regression coefficient \( \beta^i_k \) denotes the intermediate consumption of input \( i \) in order to produce one unit of output value of good \( k \), with the term \( \epsilon^i_l \) being an input-\( i \) and country-\( l \) specific hazard.

GME estimation allows the introduction of restrictions such as:

\[ \sum_{i=1}^{I} \beta^i_k = 1 \]  

(2)
which derives from the accounting identity balancing expenditure and income for each good $k$ produced in a country $l$ ($\sum_{i=1}^{I} x_i^l = \sum_{i=1}^{I} \sum_{k=1}^{K} \beta_{ik}^l y_k^l = \sum_{k=1}^{K} (\sum_{i=1}^{I} \beta_{ik}^l) y_k^l = \sum_{k=1}^{K} y_k^l$),

as well as the non-negativity of the regression coefficients ($\beta_{ik}^l \geq 0$) regardless of the positive or zero input expenditures ($x_i^l \geq 0$) as conditioning for the estimation.

The reparametrization of the coefficients $\beta_{ik}^l$ and the hazards $\varepsilon_{il}^l$, is carried out as follows:

i) for the technical coefficients

$$\beta_{ik}^l = \sum_{m=1}^{M} z_m p_{lk}^m, \text{ for } i = 1, ..., I \text{ and } k = 1, ..., K,$$

where $z_m$ denotes the points on the $M$-dimensional support for the regression coefficients $\beta_{ik}^l$ and $p_{lk}^m$ the associated probabilities.

(ii) for hazards

$$\varepsilon_{il}^l = \sum_{n=1}^{N} v_n w_{il}^n, \text{ for } i = 1, ..., I \text{ and } l = 1, ..., L$$

where $v_n$ denotes the $N$-dimensional grid points for the random variable $\varepsilon_{il}^l$, and $w_{il}^n$ are the associated probabilities.

The coefficients $\beta_{ik}^l$ and the hazards $\varepsilon_{il}^l$ are estimated as the optimal solution of the equation

$$\max_{(p,w)} \left\{ H = -\sum_{m=1}^{M} p_{lk}^m \ln p_{lk}^m - \sum_{n=1}^{N} w_{il}^n \ln w_{il}^n \right\} \text{, for any triplet } (i, k, l)$$

under the following constraints:

$$x_i^l = \sum_{k=1}^{K} \beta_{ik}^l y_k^l + \varepsilon_{il}^l = \sum_{k=1}^{K} (\sum_{m=1}^{M} z_m p_{lk}^m + \sum_{n=1}^{N} v_n w_{il}^n) \text{ for all } i \text{ and } l$$

(5.1)

$$\sum_{i=1}^{I} \beta_{ik}^l = 1 \text{ for all } k \text{ and } m$$

(5.2)

$$\sum_{m=1}^{M} p_{lk}^m = 1 \text{ for all } i \text{ and } k$$

(5.3)

$$\sum_{n=1}^{N} w_{il}^n = 1 \text{ for all } i \text{ and } l$$

(5.4)

Equation 5.1 ensures that the optimal values for the coefficients $\beta_{ik}^l$ and the hazards $\varepsilon_{il}^l$ follow model (1).

Equation 5.2 represents the accounting constraint linking the regression coefficients. Equations 5.3 and 5.4 are derived from the probabilistic definition of the support and grid weights.

For the coefficients $\beta_{ik}^l$, the normalized entropy indicator $S(\hat{p})$ is defined as follows:

$$S(\hat{p}) = -\sum_{m=1}^{M} (p_{lk}^m \ln p_{lk}^m)/(KI \ln M) \text{ whatever } i \text{ and } k$$

(6)

where $S(\hat{p}) \in [0,1]$.

For hazards $\varepsilon_{il}^l$, the normalized entropy indicator $S(\hat{w})$ is defined as follows:
\[ S(\hat{w}) = -\sum_{n=1}^{N} (w_n^\text{IL} \ln w_n^\text{IL}) / (IL \ln N) \text{ whatever i and l} \quad (7) \]

where \( S(\hat{w}) \in [0, 1] \).

The selection of the optimal model is made for the coefficients corresponding to the maximum values of the standardised entropy indicator. The comparative analysis of the econometric results for the main products between European countries illustrates the relevance of this approach for international comparisons based on their specific input productivity.

**References:**

**Psychological Distress Surveillance Using Google Trends: Data Storytelling for the “New” Mental Health Pandemic**

Fotini Diamanti¹ and Christina Parpoula²

¹(cced, student that will deliver the oral talk) Department of Psychology, Panteion University of Social and Political Sciences, Athens, Greece
²Assistant Professor of Applied Statistics & Research Methodology Department of Psychology, Panteion University of Social and Political Sciences, Athens, Greece

A pandemic is a public health emergency situation of international concern. Pandemics are not only life-threatening; they disrupt a community’s normal functioning, thus posing a huge burden on mental health of the general population. The long-term psychosocial burden of COVID-19 remains to be explored and evaluated in order to gain an overall understanding of its impact and implement effective management strategies. Monitoring the population’s mental health status will inform public health decision makers of any potential deterioration in mental health on a national level, and allow them to recognize issues of substantive importance, identify possible mechanisms of change, design tailor made interventions, and intervene sooner. Toward this end, the goal of this study is to use Google Trends to track and compare pre- and post-pandemic population-level mental health-related Google searches in Greece. The findings indicate a significant increase - following the pandemic declaration - in Google searches related to “loneliness”, “worry”, “anxiety”, and certain therapeutic techniques commonly used in the treatment of anxiety, such as “psychotherapy” and “meditation”. As it can be seen, a near-real-time surveillance tool such as Google Trends could be used by mental health experts and stakeholders to identify emerging mental health needs sooner and develop action plans for mental health promotion and prevention through population-based interventions. Although Google Trends is not a “window to the soul”, people’s searches on the Internet reflect their uncensored desires and needs for information, and therefore reduce some of the bias of the traditional self-report surveys. As other discernable population-level changes in mental health have yet to emerge, continued surveillance is warranted, and Google Trends may be proved a powerful data exploration and visualization tool for “storytelling” in regards to the “new” mental health pandemic.
Keywords: Decision-making, Mental health, Pandemic, Search engine, Surveillance.

Does Annual Report Readability Affect Investment Efficiency?

Dimitrios Konstantios\(^1,2\), and Maria Grydaki\(^3,4\)

\(^1\)Department of Economics, University of Piraeus, Piraeus, Greece
\(^2\)Department of Maritime Studies, University of Piraeus, Piraeus, Greece
\(^3\)Laboratory of Accounting and Financial Management (LAFIM), Hellenic Mediterranean University, Heraklion, Crete, Greece
\(^4\)Department of Economics, University of Crete, Crete, Greece

A considerable volume of literature, both theoretical and empirical, has witnessed the crucial role of information revealed from annual firms’ prospectuses on the cost of equity financing and firm’s performance. There is evidence that complex textual reporting hinder investor’s understanding about the firm’s economic situation leading to an increase in information risk, which eventually is associated with higher cost of equity financing. However, it is not only the quantitative information that affects the cost of equity capital but also the narrative disclosure. If low readability increases estimation risk and information asymmetry, then firms with hard-to-read annual reports should have higher cost of equity. Moreover, findings show that there is a negative relationship between information asymmetry and firm’s performance. More specifically, it has been detected that the annual reports of firms with lower earnings are harder to read, and firms with annual reports that are easier to read have more persistent positive earnings. The paper focuses on the relationship of annual reports readability and investment efficiency. More analytically, if there is asymmetric information extracted from public disclosures then managers have more incentives to hide information when firm performance is bad. With this practice, investors are not able to evaluate accurately the credibility of the financial statements, and thus the investment decisions they make are far from the optimal ones. To examine the relation between annual report readability and over/under investment we focus on U.S. firms over the period 2000-2022, retrieving financial data from Compustat and considering firm-specific variables as additional control variables. As robustness checks, we use alternative measures of annual report readability along with multiple estimation techniques accounting for endogeneity and other econometric concerns.

Limit Theorems for a Network Evolution Model describing N-interactions

Istvan Fazekas\(^1\), Attila Barta\(^1\), and Bettina Porvazsnyik\(^1\)

\(^1\)Faculty of Informatics, University of Debrecen, Debrecen, Hungary

We study a network evolution model based on collaborations inside small teams. A team is described by a clique, i.e. a non-directed graph having n vertices so that any two vertices are connected with one edge. The size of a clique is the number of its vertices.

We assume that the sizes of the cliques are 1, 2, ..., N, where N is a fixed number. The evolution of the network starts with one clique having size n, where n can be 1, 2, ..., N.

This clique is the ancestor. The ancestor has its own Poisson process which gives the reproduction times. When this Poisson process jumps, then a new member, that is a new vertex joins to the ancestor. The new vertex and some vertices of the ancestor create a new clique having size 1, 2, ..., or n+1, but not larger than N. This new clique...
is a child of the ancestor. The children of the ancestor start their own reproduction processes. Then their children also start their reproduction, etc. The evolution of the network can be described by a multi-type branching process so that the individuals of the branching process are the cliques. In our model, the life length of a clique depends on the number of its children. We prove limit theorems for the number of cliques being born and the number of cliques alive. We also obtain results for the probability of extinction. For the proof, we apply known limit theorems for multi-type branching processes. We also present some simulation results.

References:

Keywords: Network evolution, Multi-type branching process, Malthusian parameter, Poisson process.

An Efficient Distributed Fusion Smoothing Algorithm for T-proper Hypercomplex Multisensor Systems with Uncertain Observations

Rosa M. Fernández-Alcalá, José D. Jiménez-López, Jesús Navarro-Moreno, and Juan C. Ruiz Molina

1Department of Statistics and Operations Research, University of Jaén, Jaén, Spain

The distributed fusion smoothing problem for Multisensor systems with uncertain observations is addressed by using hypercomplex techniques. Specifically, the uncertainty in the observations is assumed to be modelled by a sequence of independent Bernoulli random variables. Then, taking advantage of the recent advances in tessarine algebra, the problem is solved in this domain where, under Tk-properness conditions, the dimension of the problem is considerably reduced and, consequently, a significant computational saving is implied. In this framework, a distributed fusion methodology is used to provide an efficient smoothing estimation algorithm with a lower computational cost than its real-field counterpart. The effectiveness of the proposed solution is experimentally validated by means of a simulation example. This paper has been supported in part by the Project PID2021124486NB-100 of the Plan Estatal de I+D+i", Ministerio de Educación y Ciencia, Spain, the I+D+i project with reference number 1256911, under ‘Programa Operativo FEDER Andalucía 2014--2020’, Junta de Andalucía, and Project EI-FQM2-2021 of ‘Plan de Apoyo a la Investigación 2021--2022’ from the University of Jaén.

Keywords: Multisensor distributed fusion estimation, Smoothing algorithm, Tessarine algebra, Tk-properness, Widely linear processing.

Equitable Normal Pension Age Adjusted to Fertility and Migration

Tomáš Fiala1, Jitka Langhamrová2, and Jana Vrabcová3

1,2,3Department of Demography, Faculty of Economics and Statistics, Prague University of Economics and Business, Czechia

The logical response to expected continual raise in life expectancy is to increase the retirement age. This increase should be gradual, comprehensible and intergenerationally fair, and the retirement age should be linked to mortality trends. One such option is the idea of equitable normal pension age (the ENPA), which
assumes for all cohorts the same relation between the average years of receiving a pension and the number of expected years of economic activity. In many countries. The old-age pension system is based not on the cohort, but on the PAYG principle. A very simple and rough indicator of the financial burden of this system is the adjusted old-age dependency ratio (AOADR) defined as the ratio of the population at retirement age to the population at productive age using the actual retirement age threshold instead of standard 65 years. The ENPA would guarantee the stability of AOADR only in the case of stationary type population, i.e. under fertility at replacement level or adequate compensation of lower fertility by foreign immigration. Otherwise the financial burden of pension system using ENPA would grow in time. The paper shows calculation of ENPA and corresponding AOADR for selected European countries and possible adjustment of ENPA to stabilize the values of AOADR under various variants of fertility and migration development.

Keywords: Population ageing, equitable normal pension age, old-age dependency ratio, fertility, migration.

The Joiner Representation of Multivariate Survival Functions. The r-Dependence Analysis

Jerzy Filus¹ and Lidia Filus²
¹Illinois Institute of Technology, Chicago, Illinois, U.S.A.
²Prof. Northeastern Illinois University, Chicago, Illinois, U.S.A.

Given k-variate random vectors and their univariate marginal survival functions we provide the construction and some theory of multivariate survival functions in their general universal form. In that form every multivariate survival function is represented as the product of all the univariates and the joiner function which in the case of independence reduces to the constant 1. This approach is, in a way, competitive to the copula methodology. Even though the joiner representation is universal this time we restrict our attention to the negative dependence only and analyze, more closely than in our previous papers, the emerging from the underlying theory notion of r-dependence (2 ≤ r ≤ k). The case of bi-dependence (r = 2), which, for example, is the only case for the classical multivariate normal distributions, was already investigated in our previous works. In this presentation we will consider the possibility of existence and basic properties of r-dependence for 3 ≤ r ≤ k. If r = k = 3 then, roughly speaking, the 3-dependence takes place when bi-dependences are (already) present, and their strengths (correlation coefficients, for example) depend on realization of remaining (third) random variable. The so described fact, if it takes place, determines an interesting and often useful stochastic relationship between any 3 random quantities. An attempt for generalization of 3-dependence to the case of r-dependence (4 ≤ r ≤ k) and eventual analysis of the resulted k-variate probability distributions with negative stochastic dependence will be suggested too.

Keywords: Probability, Multivariate stochastic models, their joiner representation, 3-dependence, r-dependence generalization.

The Distribution of the Maximum Likelihood Estimates of the Change Point and their Relation to Random Walks

Stergios B. Fotopoulos and Alexandros Paparas
Washington State University, Washington, U.S.

The problem of estimating the change point in a sequence of independent observations is considered. Hinkley (1972) demonstrated that the maximum likelihood estimate of
the change point is associated with a two-sided random walk in which the ascending and descending epochs and heights are the key elements for its evaluation. The aim here is to expand the information generated from the random walks and from fluctuation theory and applied to the change point formulation. This permits us to obtain computable expressions for the asymptotic distribution of the change point with respect to convolutions and Laplace transforms of the likelihood ratios. Further, if moment expressions of the likelihood ratios are known, explicit representations of the asymptotic distribution of the change point become accessible up to the second order with respect to its likelihood ratio moments. Finally, the rate of convergence between the finite and infinite distribution of the change point distribution is established and it is shown to be of polynomial order.

**Keywords:** Ascending and descending epochs and heights, Independent increments, Maximum likelihood ratio, Regular variation, Maximum processes, Conditional limit theorems.

**SEM and the Analysis of Mexican Poverty Data**

James Freeman¹ and Jola Nika

¹Honorary Fellow, Alliance Manchester Business School, University of Manchester, Manchester, UK

The study focuses on poverty in Mexico - and the potential impact of a range of cultural and social indicators on the phenomenon. Three different levels of poverty (“feeding”, “capability” and “asset”) are considered in the analysis. And two datasets from the literature – for 2000 and 2010 respectively - provide the basis of the SEM (Structural Equation Modelling) results that follow. SEM proved especially well-suited to this application – the final models obtained yielding a wealth of significant insights.

**Keywords:** Factor analysis, Mexico, Poverty, SEM.

**The Influence of Socio-economical Variables on Fertility in Selected European Countries**

Tereza Frömmelová¹, Filip Hon, and Jitka Langhamrová

¹Prague University of Economics and Business

Department of Demography, Czech Republic

The article aims to examine the influence of various socio-economical variables on the development of the total fertility rate in European countries with the highest level of total fertility rate, namely France, Romania, and Czechia. The analysis is performed [...]

**Optimal Age-replacement Policy for Series Systems Considering Parameter Uncertainty**

Kentaro Fujioka¹, Ying Ni², and Lu Jin¹

¹Department of Informatics, University of Electro-Communications, Tokyo, Japan

²Division of Mathematics and Physics, Mälardalen University, Västerås, Sweden

An age-replacement policy provides an optimal preventive replacement interval to minimize the maintenance costs. Most of the previous research proposed an age-replacement policy based on lifetime distribution with certain parameters. However, the parameters of the lifetime distributions might be subject to uncertainties. In
addition, the time required to maintain a system is likely to vary from the expected time due to various factors (e.g., location, weather, availability of replacement parts and maintenance personnel). This research considers series systems consisting of multiple types of units and proposes a model that optimally manages the uncertainty in the parameters of both the lifetime and the maintenance time distributions for an age-replacement policy. The proposed policy makes it possible to obtain the optimal replacement interval according to the priorities of the company providing maintenance for the series systems. The priority, for example, can be minimizing the risk of failure with a limited maintenance budget or minimizing maintenance costs with an upper limit of acceptable failure risk.

Keywords: Age-replacement policy, series system, parameter uncertainty.

Constrained Maximum Likelihood Estimation of Conditional Migration Probabilities and Regimes-switching Matrices

Timon Gaertner, Serguei Kaniovski, and Yuri Kaniovski

Faculty of Economics and Management
Free University of Bozen-Bolzano, Bolzano, Italy

Given a model of dependent credit rating migrations, several constrained likelihood maximization problems are proposed for estimating the conditional migration probabilities and the regime-switching matrices. Defining a feasible set, inequalities between the corresponding conditional migration probabilities account for the higher (lower) frequency of downgrading (upgrading) during downturns. The frequency patterns typical for upturns are incorporated using analogous means. The approach is tested on a S&P's dataset. Seven non-default credit classes are considered.

Keywords: Conditional, Multinomial, Mixture, Maximum likelihood, Constraint, Business cycle, Regime-switching.

A Metric based on the EDC Criterion

Jesús E. García, V.A. González-López and J.I. Gomez Sanchez
State University of Campinas, Campinas, Brazil

The Efficient Determination Criterion (EDC), proposed by Zhao et al. (2001) [1] generalizes the Bayesian Information Criterion (BIC) (see Schwarz (1978) [2]). The EDC criterion generalizes the BIC by substituting the penalty sequence term \( \{\ln(n)\} \) by the sequence \( \{c(n)\} \) that under certain conditions preserves the consistency regarding the estimation of Partition Markov Models (see García and González-López (2017) [3], Pereira (2021) [4]). In this paper, we explore the EDC criterion to generalize the metric proposed in García and González-López (2017) [3] and based on the BIC criterion. We propose the EDC-metric and we conduct a simulation study on the behavior of the EDC-metric regarding some choices of sequences \( \{c(n)\} \).

References:
Independence Test for Continuous Random Variables based on the Shape of the Young Tableau

Jesús E. García, V.A. González-López, and M.M. Kcala Álvaro
State University of Campinas, Campinas, Brazil

Given a bi-variate sample of a continuous random vector, the scatter plot of the sample defines a permutation (see [1] and [4]). From the permutation, the corresponding Young tableau is computed. The profile of the Young tableau depends on the copula defining the dependence between the random variables (see [3]). We present an independence test based on the shape of the Young tableau of the permutation. The typical shape of the Young tableau for the independence case is retrieved. A distance between the typical shape for independence and the shape computed from the sample is used as the base for the test statistics. This procedure expands the proposal introduced in [1] and [2], which used only the length of the first row and first column from the Young diagram to build an independence test. We show the procedure's efficiency in detecting dependence in real cases and through simulations.

References:

Keywords: Young diagram; Statistical independence; Hypothesis tests.

Sustainable Finance, Herding Behavior, and Risk Aversion under Uncertainty

Nektarios Gavrilakis¹ and Christos Floros²
¹PhD candidate and ²Professor of Finance
Hellenic Mediterranean University,
Department of Accounting and Finance, Heraklion, Crete, Greece

The objective of this study is to evaluate, compare and examine co-movements and dynamic correlations of herding and risk aversion under uncertainty to a sample of selected sustainability indices (DJS US index, DJS Europe index, DJS Australia index, DJS Nordic index, and DJS Emerging Markets index). Our empirical results document the existence of herding toward the risk aversion indicator. The benchmark model of herding (CSAD) influences the conditional variance of all selected sustainable markets positively except Emerging Markets, while the risk aversion indicator influences the conditional variance negatively for US and Europe. We find long-run contagion effects
in response to cross-sectional absolute deviation of returns, in the selected sample of sustainable markets except for Emerging Markets. Furthermore, we explore a spillover mechanism of risk aversion to sustainable investing in the short and long run. By using Mixed Data Sampling methods we conclude, for the first time in the financial literature, that CSAD can predict the movements of long-term volatilities of Europe and Nordic, while the risk aversion indicator can predict the movements of long-term volatilities of all sustainable markets expect Nordic. Finally, we presume that CSAD indicator and risk aversion are dynamic and time-varying, while we noticed that the short-term volatility observations do not have heavier weights than more distant past observations. These novelty results have implications for investors' behavior, risk management, portfolio construction, and hedging strategies regarding sustainable finance.

**Keywords:** Sustainable finance, Cross-sectional absolute deviation of returns, Risk aversion index, Volatility, uncertainty.

---

**Assessment of Nutritional Status and Quality of Life in Cancer Patients**

Marianna Gekou¹ and George Matalliotakis²

¹,²Hellenic Open University

The health-related quality of life (HRQoL) of oncology patients appears deteriorated due to the multiple functional disorders they experience. Additionally, these patients often face cancer cachexia, of up to 80% in metastatic disease. Cancer cachexia seems to worsen patient's HRQoL. Purpose: The aim of this study is to evaluate the nutritional status and the HRQoL of patients with gastrointestinal cancer, using the most appropriate nutritional screening tools that were well established in literature and try to find correlations between nutritional status and HRQoL. Methodology: The study involved 127 patients with gastrointestinal cancer, who underwent chemotherapy in the Day Clinic Unit of the General Oncology Hospital "Agioi Anargyroi". The nutritional status of these patients was assessed by the Body Mass Index, the White criteria, and the MNA and PG-SGA questionnaires. HRQoL was assessed with the EORTC QLQ-C30 questionnaire. Statistical analysis was performed in order to determine the effect of certain demographic characteristics and nutrition status, as measured by PG-SGA and MNA-SF, on the patients' HRQoL, using the IBM SPSS v.25 statistical software. Results: MNA-SF detected 26% malnourished patients and 36.2% at risk of malnutrition. According to PG-SGA, 72.4% of patients require immediate intensive medical and nutritional intervention, while 23.6% require dietary intervention. The BMI of patients with cancer in the upper gastrointestinal tract was statistically lower in all time periods. End-stage patients had a lower HRQoL score in QLQ-C30 compared to the rest of the patients. The HRQoL was positively correlated with the overall score of MNA-SF and negatively with the score of PG-SGA. Conclusions: The majority of patients with gastrointestinal cancer appeared to suffer from malnutrition. In the case of cancer in the upper gastrointestinal tract, the transition to a higher degree of malnutrition seems to be faster. Malnutrition appears to be associated with worse HRQoL in this population, highlighting the need to intensify the nutritional approach.
Healthcare Risk Management: Machine Learning in Myocardial Infarction

Vasileios Georgakis¹ and Panos Xenos²
¹,²University of Piraeus, Piraeus, Greece

Acute myocardial infarction is a particularly urgent medical condition that entails adverse consequences for the population, essentially necrosis of part of the heart muscle that benefits in severe stenosis or in sudden blockage of the coronary arteries. Although sufferers have an increased chance of ending up, the increase in survival is constantly increasing with the continuous progress of medical science and the fact that adversely affects the insurance industry, a fact that turns the interest of scientists to identify and modify the relative risks.

The purpose of this study is to investigate the contribution of statistical machine learning to Healthcare Risks related to myocardial infarction. Specifically, attempts to answer the following inquiry question "How can we predict risk factors of the problem with the help of new statistical methods". According to the literature, there are many machine learning methods used by researchers most common are the logistic regression, Neural networks, Decision trees, K-Nearest Neighbors, Random Forest etc. consequently this study provides beneficial related information.

References:

Keywords: Acute myocardial infarction, Healthcare Risk Management, Machine Learning, Logistic Regression.

The Financial Ramifications of the Pandemic of COVID -19 on primary Healthcare Sectors of the 7th health Region of Crete and the Management for the Treatment

Troka Gersilda¹ and George Matalliotakis²
¹,²Hellenic Open University

The Covid-19 pandemic caused significant implications by affecting and putting severe pressure on the health care systems even of the most developed countries. In this thesis, the ramifications of the pandemic on Greece and other countries, and the measures taken by the government of the relief of the affected people because of the pandemic are initially presented. What's more, the thesis focuses on the primary health care system which received the greatest pressure when it comes to the handling of the situation the country has ever been in. This means that the structure, the operation mode and the chronic problems the health sector copes with, is analysed, which is the cause of the lack of strategic measures on the part of the competent body. In the research part of the thesis, the economic ramifications of the COVID-19 pandemic on the 7th health region of Greece and the health centers of the region are analysed. More specifically, the expenditure for the payroll, the expenditure for the extra material that
was needed for the handling of the pandemic are analysed, as well as the expenditure of the rapid tests which trace the virus and the vaccines which were used.

**Approximate Formula for Adjustment Coefficient of a Non-linear Risk Model**

**Basak Gever¹ and Tahir Khaniyev²,³**

¹Department of Industrial Engineering, University of Turkish Aeronautical and Association, Bahcekapi Mah., Etimesgut, Ankara, Turkey

²Department of Industrial Engineering, TOBB University of Economics and Technology, Sogutozu, Ankara

³The Center of Digital Economics, Azerbaijan State University of Economics, Baku, Azerbaijan

This study examines a non-linear Cramér-Lundberg risk model, in order to determine the adjustment coefficient (r) when the claims have Weibull distribution. Similar insurance models have been studied in the literature, usually when premiums are linear. However, in some real-world problems, the increase in revenue may not be linear. In such cases, it is beneficial to consider nonlinear risk models. Accordingly, in this study, a nonlinear Cramér-Lundberg risk model is mathematically constructed and investigated when the premium function is . As is known, the investigation of the ruin probability is directly related to the analysis of the adjustment coefficient (r). Thus, a detailed examination of this coefficient is important. However, it is a difficult process to derive the exact value of r from an integral equation when the claims have Weibull distribution. For this reason, in this study, the adjustment coefficient is investigated using approximation methods, and a relatively simple approximate formula for the adjustment coefficient is proposed.

**Keywords:** Cramér-Lundberg Risk Model, Adjustment Coefficient, Weibull Distributed Claims, Approximate Formula.

**New Bicluster Algorithm for Trading**

**Gloria Gheno**

Ronin Institute, Montclair, New Jersey, U.S.

The choice of the rules to follow for a profitable investment in stocks is essentially guided by the trend in their price on the global market. Understanding the trend of stocks in advance determines a significant advantage for the success of investors: therefore, establishing rules to formulate buying and selling decisions is really essential. Indeed, in recent years many scholars have thought of combining financial technical indicators with biclusters [1,2,3,4] to determine their rules. By combining the methodologies, more technical indicators can be considered in order to improve the trading rules. They are usually based on historical stock values, but there are many other factors which can cause a stock price to change. Investor feelings about, for example, changes in political scenarios due to telematics or ethical choices, new business projects or even natural disasters greatly influence the stock market. In the most recent studies, sentimental analysis has been introduced in price forecasting to consider the psychological factors of investors. Following this new perspective, I propose a modified version of the bicluster proposed by [4] in which I introduce sentimental analysis so as to improve the possibility of buying or selling a stock with a significant economic gain for investors.

**References:**


Keywords: Bicluster, Sentimental analysis, Technical indicators Trading rule.

A Flexible Generalization of the Latent Dirichlet Distribution

Alice Giampino¹, Roberto Ascari¹, and Sonia Migliorati¹
¹Department of Economics, Management and Statistics, University of Milano-Bicocca, Milan, Italy

Text modeling techniques have been used in a variety of applications in the past years, among which the analysis of documents for latent topics detection. One of the most commonly used tools for topic modeling is the Latent Dirichlet Allocation (LDA). The LDA allows to obtain a representation of a document in terms of latent topic structure. The stiffness of the topic prior distribution, i.e. the Dirichlet, is one of the main limitations of this approach. To overcome this restriction, we take advantage of a generalization of the Dirichlet distribution, the so-called flexible Dirichlet (FD). The FD distribution still maintains some appealing properties of the Dirichlet distribution, e.g. interpretability, as well as conjugacy to the multinomial model, while the flexibility is guaranteed by the enriched parametrization. This work proposes a Collapsed Gibbs sampling estimation procedure for obtaining estimates of the relevant quantities. The performances of the flexible LDA are illustrated through extensive simulation studies.

Keywords: Topic models, Latent variables, Mixture model, Collapsed gibbs sampling, Textual data.

Asymptotic Results for Sums and Extremes

Rita Giuliano¹, Claudio Macci², and Barbara Pacchiarotti²
¹Dipartimento di Matematica, Università di Pisa, Pisa, Italia
²Dipartimento di Matematica, Università di Roma Tor Vergata, Roma, Italia

The term moderate deviations is often used in the literature to mean a class of large deviation principles that, in some sense, fills the gap between a convergence in probability of some random variables to a constant, and a weak convergence to a centered Gaussian distribution (when such random variables are properly centered and rescaled). We talk about non-central moderate deviations when the weak convergence is towards some non-Gaussian weak limit. In this paper we prove a non-central moderate deviation result for the bivariate sequence of sums and maxima of i.i.d. random variables. Moreover, we prove a moderate deviation result for sums of partial minima of i.i.d. exponential random variables.

Keywords: Central Limit Theorem, Fisher Tippett Theorem, Joint distribution of sum and maxima, Large Deviations, Moderate Deviations, sums of partial minima.
Extreme Value Parameters’ Estimation: An overview of Simulation Procedures

Dora Prata Gomes¹ and M. Manuela Neves²
¹Faculdade de Ciências e Tecnologia and CMA/FCT, Universidade Nova de Lisboa, Portugal
²Instituto Superior de Agronomia, and CEAUL, Universidade de Lisboa, Portugal

The main objective of statistics of extremes is the estimation of parameters of rare events. The extreme value index, \( \xi \), and the extremal index, \( \theta \), play a central role by themselves and because their influence in the estimation of other parameters, such as high quantiles. Several estimators of \( \xi \) and \( \theta \) have appeared in the literature. However those estimators depend on tuning parameters that need to be adequately chosen. Computational procedures have been considered and present nice results. They will be revisited and large simulation procedures illustrate their performance.

Acknowledgments: This work is funded by national funds through the FCT - Fundação para a Ciência e a Tecnologia, I.P., under the scope of the project UIDB/00297/2020 (CMA) and the project UIDB/00006/2020 (CEA/UL).

Keywords: Computational procedures, extremal index, extreme value index, Monte Carlo simulation.

How Vulnerable Public Health Financing is to Military Expenditure Fluctuations: Evidence from NATO Members in the Foreground of a New Cold War Era

Nikolaos Grigorakis¹,² and Georgios Galyfianakis¹
¹Hellenic Mediterranean University, Dept. of Accounting and Finance, Lab. of Accounting and Financial Management (LAFIM), Heraklion, Greece
²National Organization for the Provision of Health Services (EOPYY). Heraklion Regional Division, Greece

One year post the outbreak of Russian-Ukrainian War the European political climate appears to fundamentally be transformed. Many European members of NATO, committed to proceed on momentous increases in military expenditure for offsetting Russian hostility, by abandoning their exclusive focus on economy, trade and welfare of a long-lasting peace period on the continent. Some other showed willingness to join the military alliance by faithfully meeting its requirements and obligations. Whilst, it had already been literately examined the crowd-out hypothesis of military spending on public health financing for several group of countries, nevertheless there is no empirical evidence on the impact of military expenditure on public health financing for NATO allies. In this context, this paper examines whether military spending crowds-out public expenditure on health by using a panel data of 30 countries (i.e., 28 member states and 2 forthcoming memberships) from 2000 to 2021 and applying the Generalised Method of Moment (GMM) estimator. After a series of several econometric tests and robustness checks that had been implemented, the findings of this study present that military re-armaments deprive government fiscal resources allocated to public expenditure on health. On the foreground of a new European political economy scenery based on an unprecedented military spending impetus, this study enriches the literature, as well it provides useful insights to policy-makers that strong commitments for sharp increases in military spending confine welfare policies.

Keywords: public health financing, military expenditure, panel data, GMM model, NATO

JEL classification: H51; H56; I13; I18
Effects Selection via Likelihood-based Boosting in Cox Frailty Models

Andreas Groll¹, Trevor Hastie², Thomas Kneib¹, and Gerhard Tutz³

¹Georg-August-University Goettingen, Germany; presenting author
²University of Stanford, U.S.A.
³Ludwig-Maximilians-University Munich, Germany

In all sorts of regression problems it has become more and more important to deal with high dimensional data with lots of potentially influential covariates. A possible solution is to apply estimation methods that aim at the detection of the relevant effect structure by using regularization methods. In this talk, the effect structure in the Cox frailty model, which is the most widely used model that accounts for heterogeneity in time-to-event data, is investigated.

Since in time-to-event modeling one has to account for possible variation of the effect strength over time, the selection of the relevant features has to distinguish between several cases, covariates can have time-varying effects, can have time-constant effects or be irrelevant. Regularization approaches are discussed that are able to distinguish between these types of effects to obtain a sparse representation that includes the relevant effects in a proper form. This idea is applied to a real world data set, illustrating that the complexity of the influence structure can be strongly reduced by using such a regularization approach.

The Influence of the Asymptomatic Transmission on the Number Symptomatic Cases within a Modified SIR model

Flavius Guias
Dortmund University of Applied Sciences and Arts, Germany

We consider a compartmental model of SIR-type which considers two type of infectious states: symptomatic and asymptomatic. We consider a deterministic ODE —setting as well as a stochastic one, using a model based on Markov processes. The usual approaches in the literature to describe the asymptomatic transmission are either to assign certain probabilities of a susceptible individual to get into each of these two infectious states after having contact with an infected one (regardless of the type), or within a SEIR model by going first through the state „exposed“, after which a similar mechanism as before is considered. Or, finally, we can find the approach where an infected individual passes first always through the asymptomatic state, after which it either recovers, or it moves into the symptomatic state. In the approach presented here, we consider a slightly different transmission mechanism. The states „symptomatic“ or „asymptomatic“ are denoted for convenience, but in fact have in mind a high, respectively low viral load of the infected individual. We take into account the contacts of susceptible individuals with these two categories of infected ones, while the new state after the contact can be again one of these two. The relevant parameters are the type-specific reproduction numbers $R_{aa}$, $R_{as}$, $R_{sa}$, $R_{ss}$, where $R_{ij}$ denotes the expected number of secondary infections of type $j$ originating from contacts with an infected individual of type $i$. By considering different average times for staying in the two infective states, our model turns out to be different than the categories described above. Our focus is on the early dynamics of the epidemic, until the peak of the infectious wave has been reached. The main goal is to perform a parameter analysis and corresponding numerical simulations in order to demonstrate that it is possible to have peaks of observed cases of magnitudes like those observed
in the COVID-19 pandemics (assuming that virtually all symptomatic cases and only a certain fraction of asymptomatic cases are detected) by having a dominant transmission only on the asymptomatic level. This turns out to be possible even if the total reproduction number of symptomatic infections (taking into account all types of secondary infections) is less than 1, due to their quick isolation, and at the same time when the number of average number of symptomatic infections after contacts with asymptomatic individuals is small, of the magnitude order 0.01. We discuss the results by comparing the relevant reproduction numbers from our simulations with the values estimated in the literature on modeling the COVID-19 pandemics.

**High Dimensional Portfolio Selection with Cardinality Constraints**

Yifeng Guo¹, Jin-hong Du, and Xueqin Wang
¹The University of Hong Kong, Hong Kong, China

The expanding number of assets offers more opportunities for investors but poses new challenges for modern portfolio management (PM). As a central plank of PM, portfolio selection by expected utility maximization (EUM) faces uncontrollable estimation and optimization errors in ultrahigh-dimensional scenarios. Past strategies for high-dimensional PM mainly concern only large-cap companies and select many stocks, making PM impractical. We propose a sample-average-approximation-based portfolio strategy to tackle the difficulties above with cardinality constraints. Our strategy bypasses the estimation of mean and covariance, the Chinese walls in high-dimensional scenarios. Empirical results on S&P 500 and Russell 2000 show that an appropriate number of carefully chosen assets leads to better out-of-sample mean-variance efficiency. On Russell 2000, our best portfolio profits as much as the equally-weighted portfolio but reduces the maximum drawdown and the average number of assets by 10% and 90%, respectively. The flexibility and the stability of incorporating factor signals for augmenting out-of-sample performances are also demonstrated. Our strategy balances the trade-off among the return, the risk, and the number of assets with cardinality constraints. Therefore, we provide a theoretically sound and computationally efficient strategy to make PM practical in the growing global financial market.

Keywords: Portfolio management, Expected utility maximization, Sample average approximation, Fenchel-Rockafellar duality, Safe screening.

**Twitter Sentiment Analysis:**
Exploring Users’ Perceptions on Health and Well-Being in Europe

Qi Guo¹, Aurea Grané² and Giancarlo Manzi³
¹,²Department of Statistics, Universidad Carlos III de Madrid, Spain
³Department of Economics, Management and Quantitative Methods University of Milan, Milan, Italy

The main goal of this study is to identify the users’ content on the Twitter social network for information related to health and well-being. The pandemic has changed people’s thoughts, and people pay more attention to personal health and the national medical system. To track the evolution, we used an API tool in Python to scrap data from Twitter based on keywords such as #long-term care #pension, #insurance, and #expectations for the future during a given period (before and after March 2019). Then, we apply sentiment analysis to selected tweets by different dimensions, including timeline,
countries, languages, and related to local restriction policies. We also focus on retweets and use statistical learning models to detect the potential pattern of keywords in tweets. Keywords can also identify users’ attitudes. We introduce a global indicator WBDI, which was designed to evaluate the health and well-being status of EU residents over 50 years old. Regarding the WBDI indicator, we have mapped health and well-being status levels by country. It presents that Northern EU countries have the best general status in 2018 and 2019; we compare the distribution of WBDI as a baseline to the performance of tweets.

**Keywords:** Care Economy, Indicator, Sentiment Analysis, Statistical Tools, Well-being.

**From Prediction to Interpretation: Neural Networks vs Classical Statistics. Application to Dementia Prediction**

**Catherine Huber**

MAP5 a CNRS 8145 and University of Paris, Paris, France

We are interested in predicting the onset of a disease D [4], based on several risk factors. For that purpose, two classes of techniques are available, whose properties are quite different in terms of interpretation, which is the focus of this paper:

(i) Classical Statistics (for example: Generalized Linear Models (GLM) or First Hitting Time Models (FHT) [5])

(ii) Neural Networks (NN) (or more generally Artificial Intelligence (AI) [3]).

Both methods are rather good at prediction, with a preference for Neural Networks when the dimension of the potential predictors is high. But the advantage of the classical statistics is cognitive: the role of each factor is generally summarized in the value of a coefficient which is highly positive for a harmful factor, close to 0 for an irrelevant factor, and highly negative for a beneficial one. While the underlying model in a neural network approach mixes repeatedly all factors together so that it is rather difficult to summarize the effect of each factor.

However, we can reach some insight into interpretation of the respective impact of each risk factor. We propose to distort the data set, doing sequential permutations of the risk factors. If the prediction performance of the neural network is stable, this means that the corresponding factor is irrelevant. Conversely, if the quality of the prediction decreases, the impact of the corresponding risk factor may be considered as proportional to this decrement.

Several algorithms were developed in order to derive interpretation of the impact of each risk factor [7]. Some of them are local explanations and others global ones, while the most recent ones aim at unifying both concepts [6,1,2].

**References:**


Some Properties on Optimal Maintenance Policies for k-out-of-n:G Systems with Imperfect Repair

Sota Ikenoya¹ and Lu Jin¹

¹Department of Informatics, University of Electro-Communications, Tokyo, Japan

Most of the recent research on condition-based maintenance has assumed that repair can restore the system to as-new-as-good state. However, repairs sometimes are imperfect, which means the repair restores the system to a state between as-good-as-new and as-bad-as-old. This research investigates optimal maintenance policies for k-out-of-n:G systems considering imperfect repair with controllable repair levels. In such a system, there are totally n components. The system functions if and only if k or more of the n components function. The component deteriorates in accordance with a continuous stochastic process, and the deterioration state can only be inspected at some equally spaced time interval. Based on the inspection results, decision maker decides whether to perform a preventive maintenance of the whole system/part of components or to operate till the next inspection. This research formulates the optimal decision-making problem over an infinite horizon using a Markov decision process, where the optimal maintenance policy can minimize the total expected discounted operating and maintenance cost. Under certain conditions, we obtained the property of the optimal maintenance policy that the optimal action changes at some control limits of the state. Furthermore, a numerical example is presented to illustrate the applicability of the proposed model.

Keywords: Control limit policy, Markov decision process, Multi-component system, Preventive maintenance, Stochastic process.

Some Limit Theorems for the Total Progeny of a Positive Recurrent Q-processes

Azam Imomov¹ and Zuhriddin A. Nazarov²

¹Prof, Head of Department, Physics and Mathematics faculty, Karshi State University, Karshi city, Uzbekistan
²V.I.Romanovskiy Institute of Mathematics of Uzbekistan Academy of Sciences, Tashkent, Uzbekistan

We examine the population growth system called Q-processes. This is defined by the Galton-Watson Branching system conditioned on non-extinction of its trajectory in the remote future. In this paper we observe the total progeny up to time $n$ in the Q-process. By analogy with branching systems, this variable is of great interest in studying the deep properties of the Q-process. We find that the sum total progeny as a random variable approximates the standard normal distribution function under a second moment assumption for the initial Galton-Watson system offspring law. We estimate the speed rate of this approximation.
On Asymptotical Insensitivity of Reliability Function of a k-out-of-n System to the Shape of Its Repair Time Distribution

Nika Ivanova¹,²
¹Department of Applied Probability and Informatics, Peoples’ Friendship University of Russia (RUDN University), Moscow, Russia
²V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

The study of a k-out-of-n system's reliability is interesting both from theoretical and practical points of view. There are many investigations devoted to reliability-oriented analysis of a k-out-of-n systems, i.e., systems consisting of n components which are operable while at least k of system's components are functioning. Another problem dealing with the previous issue is the sensitivity of k-out-of-n system's reliability characteristics to the shape of their components' life and repair times distribution. In some author's previous works, closed-form representations for reliability characteristics of k-out-of-n system with an exponential distribution of components lifetime and general distribution of their repair time have been found. In this investigation the Markovization method is applied for calculation of the main reliability characteristics.

In the current paper reliability function is considered for a repairable k-out-of-n system in the same assumptions about life and repair time distributions. The problem of its sensitivity to the repair time distribution and its coefficient of variation is discussed on an example of 2-out-of-6 system. The case of rare failures of system's components is considered. It is shown that reliability function in scale of its mean time to failure is an exponential one.

The work has been supported by RSF Grant No. 22-49-02023 and the RUDN University Strategic Academic Leadership Program.

Keywords: Reliability Function, k-out-of-n System, Sensitivity Analysis, Rare Failures.

Inspection and Replacement Policy with a Flexible Inspection Schedule under Delay Time Model

Ryu Izuwaki¹ and Lu Jin¹
¹Department of Informatics, University of Electro-Communications, Tokyo, Japan

This research considers a system with three states: normal, defective, or failed states. The defective state is not self-announced and must therefore be detected by inspections, while the failed state can be known immediately. It is assumed that the system deteriorates following a delay time model. Inspections are performed periodically, and repairs will be carried out if the system is found to be in either a defective or failed state. In addition, a replacement that is carried out at a pre-specified time epoch will restore the system and end the cycle from the start of operation. An inspection and replacement policy with a flexible inspection schedule is proposed to minimize the average cost per unit time. Some numerical examples are presented to illustrate the proposed policy. Furthermore, the relationship between the optimal inspection schedule and the effects of the parameters is discussed.

Keywords: time-based maintenance, preventive maintenance, inspection policy, Weibull Distribution, cost rate.
Bermudan Option Pricing Using Almost-Exact Scheme under Heston-type Models

Mara Kalicanin Dimitrov¹, Marko Dimitrov¹, Anatoliy Malyarenko¹, and Ying Ni¹
Division of Mathematics and Physics, Mälardalen University, Västerås, Sweden

A Bermudan option is an American-style financial option that can only be early-exercised on a set of predetermined dates prior to maturity. While using the Monte-Carlo simulation method for Bermudan option pricing, it is efficient to use an exact simulation scheme so that the simulation is done only for the dates of early exercising. Such an exact scheme exists for the classical Black-Scholes model since the corresponding stochastic differential equation has an analytical solution. For the more recent Heston model with one stochastic volatility process, there is a so-called Almost Exact Scheme (AES) which uses the non-central chi-square probability distribution for the variance process. This paper focuses on pricing Bermudan options under Heston-type stochastic volatility models using an AES scheme for simulations. The models under consideration are the traditional Heston model with one stochastic volatility and the Double Heston model with two stochastic volatilities. We derive the AES scheme analytically for the Double Heston model and investigate the advantages of using the AES scheme under both the Heston and the Double Heston model numerically. The paper shows that the AES scheme works well when the number of simulated steps equals the number of exercise dates, making it efficient. Keywords: Almost Exact Scheme, Monte Carlo, Bermudan Options, Least Squares Monte Carlo, CIR, Heston Model, Double Heston Model, Stochastic Volatility.

Statistical Analysis of Groundwater Level Regime in Slovakia

Jana Kalická and Dominika Sónak Ballová
Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, Bratislava, Slovakia

In our analysis we deal with the statistical evaluation of the groundwater level in the Great Corn Island. Groundwater is one of the most important sources of drinking water. The Great Corn Island is the largest river island in Europe and also the largest reservoir of drinking water in Central Europe. Naturally, it is very important to monitor and measure the state of groundwater here. In this paper, we used statistical methods to determine sudden and extreme changes in the groundwater level, but we also monitored the gradual development of the groundwater level on selected probes. Another aim was to compare the development of the groundwater level in these probes to each other. This includes, in addition to detecting changes, a comparison of the appropriate time series model, the length of the most significant period and the number of significant periods. All results were obtained by analyzes in statistical software R. Acknowledgement: This work was supported by Slovak Research and Development Agency under contract No. APVV-18-0052 and contract No. APVV-17-0066. Keywords: Hydrometeorology, Trend, Change-point, Groundwater.
Stochastic Orders and Reliability Properties for Compound Geometric Convolutions with Applications in Perturbed Risk Processes

Lazaros Kanellopoulos  
Department of Statistics and Insurance Science, Piraeus, Greece

It is well known that the survival probability in the perturbed risk processes can be deduced to the form of a compound geometric convolution. We give a general result comparing the maximal aggregate losses of two different perturbed risk processes by an integral stochastic order, which has the property of convolution closure. Also, we derive reliability properties for the ruin probability and the deficit at ruin at risk model perturbed by an independent diffusion process, and study stochastic orders related to the Laplace transform for that risk model.  
Keywords: Laplace transform order, Ruin probability, Compound geometric, Compound geometric convolution, Diffusion.

An Exponentiality Test against Heavy & Light Tailed Alternatives

A. Karagrigoriou\textsuperscript{1}, I. Mavrogiannis\textsuperscript{1}, G. Papasotiriou\textsuperscript{2}, and I. Vonta\textsuperscript{2}  
\textsuperscript{1}Laboratory of Statistics & Data Analysis, Univ. of the Aegean, Greece  
\textsuperscript{2}Department of Mathematics, National Technical of Athens, Greece

This work is filling up the gap in the literature regarding the verification of the log-concavity property which is a widely studied topic due to the fact that it provides desirable estimating properties. At the same time, log-concavity together with log-convexity are vital in reliability, engineering and stochastic modeling for distinguishing between an exponential, a light-tailed and a heavy-tailed distribution. In this work we propose an exponentiality test of fit to be used for distinguishing between exponential and log-concave or long-convex distributions. The proposed test statistic is based on the conspiracy and catastrophe principles through which a characterization for the (tail part of the) exponential distribution is established. The details of the formulation of the test are provided, an extended simulation study showcasing the performance of the proposed test statistic is given, and some concluding remarks are given.  
Keywords: Exponentiality test, Heavy-tailed distributions, Log-concavity, Log-convexity.

Models and Applications of Reinforcement Learning

Michael N. Katehakis  
Rutgers University, New Jersey, USA

This talk provides a survey of the basic models of reinforcement learning algorithms and their applications. It starts with a review of the origins including connections with the multi-armed bandit models. It introduces UCB and Optimistic-based algorithms and applications in online and offline settings. We outline the limitations that arise due to very large state-action spaces and computational complexities. We then review Neural networks and Deep Learning approaches that enable the efficient application of Reinforcement Learning to problems that were previously intractable.
Factors Explaining Perceptions towards Information and Communication Technologies (ICT) with Data Drawn from the ESS

Aggeliki Kazani¹, Eva Tsouparopoulou¹, Dimitris Parsanoglou², and Maria Symeonaki¹

¹Department of Social Policy, Panteion University of Social and Political Sciences, Athens, Greece
²Department of Sociology, National and Kapodistrian University of Athens, Athens, Greece

Information and Communication Technologies (ICT) evidently play an important role in the everyday lives of individuals invading their personal sphere, leisure time, working activities and civic participation, whether at work, at home or on the move. What has led to a renewed focus on this matter turning it recently into an urgent policy concern was the COVID-19 pandemic and its aftermath, although the changing nature and use of Internet has long been identified as a mirror of social inequalities and a determinant of transformations in social stratification. The motivation behind this study is to systematically explore and present indicators that aim to capture the attitudes of individuals, towards digital communications. Raw data are drawn from the latest round of the European Social Survey (ESS, round 10) and analysis is performed for the participating countries for which raw data are released. Discrepancies in attitudes are measured with the aid of the ESS's rotating module included in this round 'Digital social contacts in work and family life'. The aim is to measure general attitudes towards digital mobile and online communication of individuals indicating differences amongst different socio-economic groups. Questions analysed will give an insight of digital positivity and/or disbelieve which will also be used as a composite indicator of digital scepticism and optimism in the different participating countries. This study will document whether alarming discrepancies in attitudes towards ICT use exist among European citizens.

Keywords: ICT, digital skills, digital divide.

Asymptotic Upper Bounds for Overflow Probability in the Case of Heterogeneous Traffic

Yury Khokhlov¹, Oksana Sidorova², and Lev Suslov¹

¹Lomonosov Moscow State University, Faculty of Computational Mathematics and Cybernetics, Department of Mathematical Statistics, GSP-1, Moscow, Russia
²Tver State University, Faculty of Applied Mathematics and Cybernetics, Department of Mathematical Statistics and System Analysis, Tver, Russia

One of the most important tasks in analyzing of telecommunications system is to assess the quality of service of this system. For this purpose, the probability of a system buffer overflow is often used. It is rarely possible to calculate it explicitly, so one or another estimate of this characteristic is used. In our report we consider a system whose input load is the sum of some average load with intensity μ and the sum of independent fractional Brownian motion and stable Levy motion. The system has one service device with service intensity C>0. When r=C-μ>0 there is a stationary mode. Let Q be the maximum load in stationary mode, b be the buffer size. We are interested in the value P(Q>b) for large b. We propose some upper asymptotic estimates. These upper bounds have a power order. The proof method is based on

This research was carried out in accordance with scientific program of the Moscow Center for Fundamental and Applied Mathematics and Faculty of Computational Mathematics and Cybernetics in Lomonosov Moscow State University.

References:

Keywords: Quality of Service, Asymptotic Buffer Bounds for Buffer Overflow Probability.

Determination of Moods in Turkish Sentences with Sentiment Analysis Method

Seda Kilicer1,2 and Ruya Samli2
1Department of Computer Engineering, Beykent University, Beykent University Ayazaga Campus, Istanbul, Türkiye
2Department of Computer Engineering, Istanbul University-Cerrahpasa, Istanbul University, Istanbul, Türkiye

People want to use the developments in technology more actively in their daily lives in order to make their lives easier more comfortable. Thanks to the internet, people can do many things that need to be done during the day quickly from wherever they are. The active use of the Internet leads to the formation of rapidly increasing data. However, it is desirable to obtain useful information from these data in a short time. For this reason, it is tried to obtain results by analyzing the data quickly by using different methods. Sentiment analysis is used to determine the emotional states in the texts written in order to express ideas about a subject. In sentiment analysis, which is used to process the data and transform it into useful information, it is aimed to determine the general point of view on the subject examined by determining the opinion and emotional state in the texts by using various methods. In this study, sentiment analysis was performed in sentences belonging to the Turkish language. It is aimed to reveal which emotional state these sentences belong to from different emotional states such as neutral, positive and negative. The dataset used in this study was obtained from Turkish comments shared publicly on the Kaggle platform. This dataset was created from comments made about different products in various electronics stores. In this study, the emotional state of 11429 Turkish interpretations was examined. Weka program was used to determine the emotional states of Turkish comments. Many algorithms have been experimented with using this program. Success rates obtained as a result of training and testing were compared using Decision Tree (J48), K-Nearest Neighbor (IBk), KStar, Logistics, Naive Bayes, Naive Bayes Multinominal, Random Forest and Support Vector Machine algorithms. In order to determine the emotional states of the sentences, all data are divided as 70% training data and 30% test data. The training data set was loaded into the system and the training process was carried out using these algorithms, respectively. Then, the test data were loaded into the system and it was tried to predict which of the comments belonged to neutral, positive or negative emotional states with this algorithm. The highest accuracy rate with the training data set was obtained as 98% using Random Forest and K-Nearest Neighbor algorithms. The highest accuracy rate with the test data set was obtained as 60% using Naive Bayes Multinominal and Support Vector Machine algorithms. In addition, 59% accuracy was obtained by using the logistic algorithm with the test data set.
Keywords: Sentimental Analysis, Weka, Decision Tree (J48), K- Nearest Neighbor (IBk), KStar, Logistics, Naive Bayes, Naive Bayes Multinominal, Random Forest, Support Vector Machine.

**Prognosis of COVID-19 Patients by the Underlying Diseases and Drug Treatment**

Ho Kim¹ and Taerim Lee²

¹Institute of Health and Environment and Department of Public Health Sciences, Graduate School of Public Health, Seoul National University, Seoul, Korea
²Department of Data Science & Statistics, Korea National Open University, Seoul, Korea

Certain underlying diseases such as diabetic mellitus and hypertension are a risk factor for the severity and mortality of coronavirus disease (COVID-19) patients. Furthermore, both angiotensin converting enzyme inhibitors (ACEi) and angiotensin II receptor blockers (ARBs) are controversial at role in the process of COVID-19 cases. The aim of the study was to investigate whether underlying diseases and taking ACEi/ARBs, affect the duration of hospitalization and mortality in patients with confirmed COVID-19. Medical usage claims data for the past three years until 15 May, 2020, from the “CORONA-19 International Cooperation Research” project was used. We analyzed the medical insurance claims data for all 7590 coronavirus (COVID-19) patients confirmed by RTPCR tests nationwide up to May 15, 2020. Among the comorbidities, a history of hypertension (hazard ratio [HR], 1.51; 95% confidence interval [CI], 1.056–2.158) and diabetes (HR, 1.867; 95% CI, 1.408–2.475) were associated significantly with mortality. Furthermore, heart failure (HR, 1.391; 95% CI, 1.027–1.884), chronic obstructive pulmonary disease (HR, 1.615; 95% CI, 1.185–2.202), chronic kidney disease (HR, 1.451; 95% CI, 1.018–2.069), mental disorder (HR, 1.61; 95% CI, 1.106–2.343), end stage renal disease (HR, 5.353; 95% CI, 2.185–13.12) were also associated significantly with mortality. The underlying disease has increased the risk of mortality in patients with COVID-19. Diabetes, hypertension, cancer, chronic kidney disease, heart failure, and mental disorders increased mortality. Controversial whether taking ACEi/ARBs would benefit COVID-19 patients, in our study, patients taking ACEi/ARBs had a higher risk of mortality.

**References:**


Keywords: COVID-19; underlying disease; medical treatment.

Regional Mortality and Health Status in Greece (1990-2020)

Spyridon A. Kontos1 and George S. Atsalakis2

1Technical University of Crete, School of Production Engineering and Management, Crete, Greece
2Technical University of Crete, Crete, Greece

In this paper we investigate regional mortality and health status in Greece, using data from the European Statistical Office (Eurostat) for the period 1990-2020. Firstly, a logistic function was fitted to the data series of each period life table, in order to estimate and smooth the age specific death rate in older ages. Using these age-specific death rates and through a dynamic model with stochastic simulation, we estimated the health state function across the 13 regions of Greece for each year of study and sex. Furthermore, a hierarchical cluster analysis was performed on our results of the previous analysis and a heatmap was created to determine possible groupings in the years and regions of study.

A New Family of Continuous Univariate Distributions with Applications in Actuarial Science

Markos V. Koutras1 and Spiros D. Dafnis1

1Department of Statistics and Insurance Science, University of Piraeus, Piraeus, Greece

In this work we introduce a wide family of continuous univariate distributions with support \((0,\infty)\). The new family, denoted by \(D^+_\theta(h)\), contains distributions with cumulative distribution functions of the form \(F(x; \theta) = g^{-1}(h(x; \theta))\) where \(g\) is a function free of parameters, \(h\) is a function carrying the new distribution’s parametric vector \(\theta\) and \(g, h\) satisfy specific conditions. Several properties of \(D^+_\theta(h)\) are studied,
including unimodality, aging and probability bounds. Special attention is drawn to members of \( D_g^h (h) \) which have heavy tails and an illustration is provided how the new family could be exploited for fitting real data from the actuarial field.

**Acknowledgment:** This research was financially supported by the MSc program in Actuarial Science & Risk Management of the University of Piraeus through the grant “C.908: Development of new methodologies in insurance and risk management”.

**Keywords:** Generators, Aging properties, Modes, probability bounds, Heavy tailed distributions.

---

**On the Bivariate Polya-Aeppli Counting Process and its Applications**

Meglena Lazarova\(^1\), Leda Minkova\(^2\), and Michail Todorov\(^3\)

\(^1,3\) Faculty of Applied Mathematics and Informatics, Technical University of Sofia, Sofia, Bulgaria

\(^2\) Faculty of Mathematics and Informatics, Sofia University “St. Kliment Ohridski”, Sofia, Bulgaria

In this paper an extended bivariate Polya-Aeppli process (BPAP), the joint probability mass function and a Bivariate Polya-Aeppli process as birth process are introduced. We consider a bivariate risk model with BPAP counting process. The ruin probability and a case of exponentially distributed claims are also discussed. For this case we derive a partial differential equation for the non-ruin probability and we discuss the achieved solution of the obtained partial differential equation.

**Keywords:** Bivariate Polya-Aeppli counting process, ruin probability, exponentially distributed claims.

---

**Simple form of Probability Density Functions via Sampling**

Maira Ledaki\(^1\) and Myrto Papageorgiou\(^2\)

\(^1\) Inter-Faculty Master Program “Networks & Complexity”, Aristotle University of Thessaloniki, Thessaloniki, Greece

\(^2\) PhD Candidate Dept of Mathematics Aristotle University of Thessaloniki, Thessaloniki, Greece

A critical point in Statistical Analysis is the determination of the Probability Density Function (PDF) of the random variables entered by the problem under study. The data to be used is sampling data. The use of classic methods to determine the PDF is often extremely time-consuming. Over the last two decades, the Coefficient of Variation (CV) has been a useful tool for the above purpose, as it reduces the time of calculations by far. In this work we study the case where the graph of the PDF does not pass through the origin of the axes, i.e. it is \( f(0)=m>0 \); X axis represents the random variable values and Y axis represents their frequencies. In this paper we deal with continuous random variables and we present some examples.

**References:**


The paper [3] addresses the asymptotic analysis of sojourn functionals of spatiotemporal Gaussian random fields with long-range dependence (LRD) in time also known as long memory. Specifically, reduction theorems are derived for local functionals of nonlinear transformation of such fields, with Hermite rank \( m \geq 1 \), under general covariance structures. These results are proven to hold, in particular, for a family of non–separable covariance structures belonging to Gneiting class. For \( m = 2 \), under separability of the spatiotemporal covariance function in space and time, the properly normalized Minkowski functional, involving the modulus of a Gaussian random field, converges in distribution to the Rosenblatt type limiting distribution for a suitable range of the long memory parameter. For spatiotemporal isotropic stationary fields on sphere similar results obtained in Marinucci et al. [5]. Some other related results can be found in Makogin and Spodarev [4]. For short-memory random fields the asymptotic analysis of sojourn functionals can be done using the Mallivin-Stein technique, fourth-moment limit theorems, Breuer-Major type theorems (see [1,2,6,7,8] and the references therein).

This is joint results with M.D.Ruiz-Medina (Granada University, Spain).

References:
An Alternative Approach to Causes of Death Prediction using Support Vector Machines

Susanna Levantesi¹, Andrea Nigri², and Damiano Ticconi³

¹Department of Statistics, Sapienza University of Rome, Rome, Italy
²Department of Economics, Management and Territory, University of Foggia, Foggia, Italy
³Assicurazioni Generali, Rome, Italy

The analysis of the causes of death provides valuable insight into changes in the overall mortality trends hidden in population-level data. During the past two centuries, the causes of death have shifted from infectious diseases to chronic diseases, and, consequently, the mortality reduction has moved from childhood to old age. In addition, the evolution of mortality by causes, monitored only for a small set of countries, may present different classifications over time due to regular revisions of classification rules and consequent disruption in long-term cause-of-death analysis. We develop and apply a novel statistical framework leveraging a machine learning algorithm to model cause-specific mortality. Our approach consists of using Support Vector Machines to classify mortality rates into specific macro classes of death causes. The proposed model is formulated by exploiting specific features of mortality dynamics such as age, gender, and central death rates, and could be even used to classify data for countries where causes of death are not registered. We use internationally classified cause-of-death categories and data obtained from Human Cause-of-Death Database, considering all the available countries.

Keywords: Mortality, Causes of death, Support Vector Machine.

Optimising Financial Trading Strategies using Dynamic Bayesian Networks

Karl Lewis¹, Mark Anthony Caruana², and David Paul Suda³

¹,²,³Department of Statistics and Operations Research, Faculty of Science, University of Malta

Devising a trading strategy that allows for long-term gains is a very common problem in finance. In this paper, we aim to formulate a mathematically rigorous framework for the problem, study various trading strategies, and compare and contrast the results obtained. The main approach we consider is based on Dynamic Bayesian Networks (DBN). Within the DBN setting, we will consider long-term as well as short-term trading strategies, applied on different sets of equities. We conclude that both the long-term and the short-term strategies proposed in this paper outperform the benchmark buy-and-hold trading strategy. Despite the clear advantages of the former trading strategies, we will also discuss the limitations of this model as well as a number of possible improvements which can be implemented.

Keywords: Finance, Dynamic Bayesian Networks, Trading Strategies, Equities.
The Use of the Refined Descriptive Sampling to Reduce the Estimation Variance of the Survival Function

Azzal Mahdia¹ and Ourbih-Tari Megdouda²
¹UMBB, Boumerdès, Algérie
²Centre universitaire de Tipaza, Tipaza, Algérie
Laboratoire de Mathématiques Appliquées, Faculté de Sciences exactes, Université de Bejaia, Bejaia, Algérie

In this work, we will introduce a methodology to reduce the variance in non-parametric estimation of the survival function in the case of right censored data. The Kaplan Meier (KM) and Fleming Harrington (FH) estimators are selected for such work. Therefore, we will use the Refined Descriptive Sampling (RDS) by developing an efficient algorithm that we call Non-parametric Adaptive RDS (NARDS). The obtained results show that the RDS method reduces the estimation variance of KM and FH estimators.

Drawdown Risk Measures for Asset Portfolios with High-Frequency Data

Giovanni Masala¹ and Filippo Petroni²
¹Dipartimento di Scienze Economiche e Aziendali, Università di Cagliari, Italy
²Dipartimento di Management, Università Politecnica delle Marche, Ancona, Italy

In this paper, we analyze Drawdown-based risk measures for an equity portfolio with high-frequency data. The returns of individual stocks are modeled through multivariate weighted-indexed semi-Markov chains with a copula dependence structure. Through this recently published model, we show that the estimate of Drawdown-based risk measures is more faithful than that obtained with the application of classic econometric models.

References:

Keywords: Drawdown risk measure; Weighted-indexed semi-Markov models; Asset portfolio; High-frequency data; Right censoring; GARCH models.

Probability weighted Moment Method: An Approach to Estimating the Pareto Distribution

Ayana Mateus¹ and Frederico Caeiro
¹NOVA School of Science and Technology (FCT NOVA) and CMA, Largo da Torre, Caparica, Portugal

The estimation of Pareto model parameters is a popular field of study due to the fact that the tail of many probabilistic models exhibits Pareto behavior. This makes estimating Pareto parameters more crucial than for many other distributions. Estimators for the Pareto distribution can also estimate tail characteristics such as the tail index, which is a key topic in the statistics of extreme values. In this research, we
propose to investigate and compare the performance of a class of estimators for the shape parameter of a Pareto type I distribution obtained by extending the probability-weighted moment technique. This class of estimators depends on a control parameter that must be properly chosen. We present here a procedure for choosing this parameter.

**Keywords:** Pareto distribution, Probability weighted moment, Parameter estimation.

---

**Portfolio Time Consistency and Utility Weighted Discount Rates**

Oumar Mbowdji¹ and Traian A Pirvu²

¹Independent Researcher

²Math and Stats Department, McMaster University, Hamilton, Canada

Merton portfolio management problem is studied in this paper within a stochastic volatility model and non constant time discount rate framework. This problem is time inconsistent and the way out of this predicament is to consider the sub game perfect strategies. The later are characterized through a value function. A two stage approach is developed to find the value function: in a first step the utility weighted discount rate is introduced and characterized as the fixed point of a certain operator; in the second step, given the utility weighted discount rate, the value function is found through solving a parabolic linear equation. Numerical experiments explore the effect of the time discount rate on the sub game perfect and pre commitment strategies.

---

**Updating the Proposal for Assessing the Quality of Census Count of Centenaries**

Juliana Barbosa Medeiros¹ and Neir Antunes Paes²

¹,²Health and Decision Modelling Postgraduate Course. Federal University of Paraíba, Brazil

The main goal of this study was to review and updating the proposal for assessing the quality of the census data counting of centenaries elderly people (RELAP - Revista Latinoamericana de Población Vol. 16: e202207). For the construction of the proposal, data were originally collected from individuals aged 100 years or older of the population census of 193 countries belonging to the United Nations, based on the last censuses of each country (mostly carried out in 2010 and 2011). Therefore, the need to review the parameters used for the construction of this proposal is justified, as well as to update it with more recent census data. Following the original proposal, this study was based on a group of indicators for regional population groups in the world. Three demographic indicators were constructed: Centenarians’ Sex Ratio; individually, for male and female population aged 100 years and over, the ratio with the elderly aged 85 years or over was calculated. These indicators were calculated for the 85 countries selected among the initial 193, and related to the Human Development Index and life expectancy at birth applying linear regression. Once a high relationship was proven, the multivariate classification technique cluster analysis was applied to select groups of countries with similar characteristics and formulate ranges of values for each indicator. Five scores were designated that determined the final classification. The revised and updated proposal showed good internal consistency, with results
consistent with the reality of the countries studied, satisfactory levels of application, scalability and reproducibility for different locations. Furthermore, the proposed classification is easy to understand and simple, and uses few indicators. Therefore, the use of the revised proposal may favor the understanding of the true panorama in the distribution of longlived elderly of the world.

**Keywords:** Aged, Data Accuracy, Census Data, Classification.

---

**A Shock Model with Time Randomized via a Multivariate Subordinator**

**Alessandra Meoli**  
*Dipartimento di Matematica, Università degli Studi di Salerno, Fisciano (SA), Italy*

We study a shock model for a system subject to two types of shocks that occur as events in an underlying bivariate counting process. This is of interest for many applications in reliability theory and survival analysis. The novelty in this work consists in the construction of the bivariate counting process. Indeed, previous approaches considered for the occurrence of the shocks: a bivariate Poisson process [1], a bivariate space-fractional Poisson process, which is a two-dimensional vector of independent Poisson processes time-changed by an independent stable subordinator [2]. In this last case, the random time is common for both components. Alternatively, we now time-change a bivariate Poisson process by taking a bidimensional gamma subordinator. Therefore, the random time is not common to both components. This is fundamental when different time scales need to be considered. Under different regimes for the failure, we aim at (i) providing in closed and tractable form the exact expressions for the hazard rates, the failure densities, and the survival function, and (ii) investigating the related ageing notions. This is a joint work with Antonio Di Crescenzo, Università degli Studi di Salerno, Italy.

**References:**

**Keywords:** Shock model, Poisson process, Gamma subordinator, Failure distribution, Ageing notions.

---

**From Zero Clearing to Opening Management: The Estimation of the Confirm Cases and Mortality Cases In Post-Covid-19 Pandemic Governance (Nov.2022-Feb.2023)**

**Hong Min** and **Yixiao Wei**

1,2Workshop of Population Big Data and Policy Simulation, School of Public Affairs, Zhejiang University, Hangzhou, China

With the release of a series of China policies on COVID-19 Prevention and Control Measures by the end of 2022, the mode of COVID-19 prevention and control in China has gradually shifted from zero clearing to opening management. During the same
period, China’s health Commission suspends releasing daily numbers of confirmed cases and deaths. Specially, the ratio of China to receive at least one dose of COVID-19 vaccine is 92.6%, while representative OECD countries that have already adopted fully opening management such as Britain, France, Germany and Japan are below 90%, of which data analyses may not appropriate for China's post-COVID-19 assessment. On account of a variety of reasons including insufficient availability of data on the number of infections and deaths of COVID-19, there has been no specific estimate and analysis of infection data in post-Covid-19 pandemic. In this paper, by constructing the small-world network model and the three-input-parameter model life table of Developing Countries Mortality Database (DCMD, software copyright No. 2023R11S0021930), the confirmed cases and deaths during the period from November 2022 to February 2023 are estimated in post-Covid-19 pandemic period in China; combined with the characteristics of population flow and age distribution of infected population, the spatial and age distribution characteristics of infected population and death cases are analyzed; the impact of the pandemic on future changes in Chinese life expectancy is reassessed, especially for the aged over 65 years. The results show that during the period from November 2022 to February 2023: (1). The cumulative death population of COVID-19 in China is about 1.5 million or so, including the death population die of COVID-19 and die with COVID-19. (2) The total infection rate in China is about 80%. (3) The infected population in rural areas is relatively lower than that in urban areas. (4) The infection rate of minor is relatively lower than that of adults. This paper can provide scientific evidence for the post-Covid-19 pandemic governance. This study is supported by Natural Science Key-Project of Zhejiang Provence, P.R.C., “Turning Crisis into Opportunity: Monitoring and Early Warning and International Public Health Strategy System Simulation under the Background of COVID-19 Pandemic” (No.LD21G030001), and Key-Project of National Social Science Foundation of China, “Research on the development of multi-level and multi-pillar pension insurance system " (No.21ZDA101).

Previous Relative Results:
Yichao Li, Nan Li, Hong Mi. Life Expectancy and Different Parameter Identification in Chinese Retirement Plan. ASMDA 2021 International Conference.
Hong Mi, Yuan Li, Qiyini Ma. The Belt and Road“ International Migration of Asia. China Renmin University Press, 2022.
A software for Intelligent generation of indicator groups of life table of national models along the Belt and Road. Copyright Protection Center of China: No. 2023R11S0021930.

Keywords: Post-Covid-19 Pandemic, Small-world network, DCMD.

Using Surrogate Models and Variable Importance to better Understand Random Forests Regression Fitting

Manlio Migliorati\(^1\) and Anna Simonetto\(^1\)

\(^1\)University of Brescia, Department Civil, Environmental, Architectural Engineering and Mathematics (DICATAM), Brescia, Italy

Interpretability mechanisms helping users in better understanding machine learning models are crucial for Artificial Intelligence acceptance. In this manuscript, our experience in interpretation of random forest regression via surrogate models, i.e.
models trying to replicate in an interpretable framework an original fitting difficult to understand, is reported. It is shown how, beyond classical R2 analysis, adequacy of surrogate models can be assessed via variable importance analysis.  

**Keywords:** Interpretability, Surrogate models, Variable importance.

“Non-Heaping” in Age Data: A New Technique  
Barun Kumar Mukhopadhyay  
*Population Studies Unit, Indian Statistical Institute, India (Retired)*

Turner (1958) proposed a hypothesis that people while estimating age (numerical data) ending in digits with preferences for multiples of ten, five and two. Out of the remaining odd digits, one, three, seven and nine as one and nine were between the most attracted digit, ‘0’ hence they were at a most disadvantaged position compared to three and seven surrounding even digits, ‘2’, ‘4’, ‘6’ and ‘8’. He graphically had shown the pattern of heaping for all the ten digits and precisely gave a table of hypothetical ranking for overall countries. Using census age data of a country, he calculated the actual ranking of those digits and compared the observed rank order to the expected one by Spearman’s rank correlation coefficient (ρ). Higher the coefficient greater the extent of error in the raw data. After Turner’s work in 1950s, a long time already elapsed. Since the quality of single year age data much improved in most of the developed countries over this long period of time, Turner’s ranking for them is invalid. However, Turner’s ranking is still valid for most of the developing world particularly South, South East Asia, and Sub Saharan African countries etc., where huge peaks and troughs are still observed in the single year raw age data in censuses. As a result the present paper tries to develop a new theoretical ranking only for developed countries only based on life table stationary age distribution by ten digits of age declining monotonically within a broad age band. The rest part of the work is similar to Turner. To test the new technique and the Turner’s one a kind of case study using some developed and developing countries’ census single year age data are analysed and the result is consistently good.  

**Keywords:** Turner’s method, Non-heaping in age data, Digit Preference error, Myers’ Index, Whipple’s index.

The Mean-Variance Problem Considering Investor’s Views:  
Black-Litterman model on a Swedish Portfolio  
Jean-Paul Murara  
*Department of Mathematics and Physics, School of Education, Culture and Communication, Mälardalen University – Västerås, Sweden*

In this article, we build an investment portfolio in the Swedish market based on mean-variance problem introduced by Markowitz, and we use OMXS30 as our benchmark. We make modifications of the problem inserting the investor’s views inspired by Black-Litterman model. This model provides an efficient way to combine investor’s view with implied equilibrium return and these two factors are both used to make an improved prediction on the stock performance. We test how effective is the Black-Litterman model is and perform the empirical results on different types of portfolios. We observe
that the quality of the investor’s views influences the performance of the model. Our portfolio outperforms the benchmark and the CAPM portfolio.

**Keywords:** Portfolio Optimization, Black-Litterman, Mean-Variance, CAPM.

**Probability equivalent level of Value at Risk and higher-order Expected Shortfalls**

*Fanni K. Nedényi*

*ELKH-SZTE Analysis and Applications Research Group, Bolyai Institute, University of Szeged, Szeged, Hungary*

The Basel Accords require banks to set aside capital in line with their levels of risk. Currently, Value-at-Risk (VaR) is the applied risk measure of the potential loss in the value of a portfolio. In particular, 99% VaR is utilized, which is the loss that is likely to be exceeded only 1% of the time. While VaR is widely used and easy to compute, it has no information on the magnitude of the biggest 1% of losses. Moreover, it is not a coherent risk measure. Indeed, it is not subadditive, which means that VaR of a portfolio can be higher than the sum of the VaRs of the individual assets in the portfolio. The Fundamental Review of the Trading Book is expected to make a complete revision of the approach to calculating risk-based capital requirements for investments. The 99% VaR is supposed to be replaced by 97.5% Expected Shortfall, which is the average of VaR(x) for x between 0.975 and 1. Li and Wang [2] studied the effect of this proposed change. Similarly to their work, we are examining higher-order Expected Shortfalls as potential alternative risk measures. We define PELVE(n), which basically tells us what level n-th-order Expected Shortfall corresponds to a certain level VaR. We investigate its properties and calculate PELVE(2) for some important distributions including ones with heavy tail. Moreover, for PELVE(2), we present some simulation results along with real data analysis.

Supported by the ÚNKP-22-4 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.

**References:**


**Keywords:** Value at Risk, higher-order Expected Shortfall, Gini Shortfall, PELVE.

**Multiple Trajectory Analysis in Finite Mixture Models**

*

*Cédric Noel and Jang Schiltz*

*University of Luxembourg, Department of Finance, Luxembourg*

Finite Mixture Models, as defined by D. Nagin (2005) and extended by J. Schiltz (2015) are used with diverse underlying distributions, like normal, ZIP, logistic and Beta distributions. These models allow to analyze several indicators of interests. Rather than studies each indicator separately it is often most interesting to studies them jointly. Nagin (2005) and Burckhardt et al. (2016) propose an approach to deal with this situation. But their method has the drawback that the many parameters do not have a practical interpretation.
We present another methodology which allows for multiple trajectory analysis with parameters that all have an easy interpretation. The model is illustrated by numerical examples and can be fitted with the R package trajeR developed by C. Noel (2023).

**References:**


**Keywords:** Finite Mixture Models, Multiple trajectory analysis, R package.

---

**Small Time LDP and Stochastic Volatility Models**

*Barbara Pacchiarotti*

*Dipartimento di Matematica. Università degli Studi di Roma, Roma, Italia*

We investigate the large deviation behavior in small time of continuous Gaussian processes. We introduce a general procedure allowing to derive large deviation principles in small time starting from the well understood context of large deviation principles with a small parameter, going beyond the self-similar case. Starting from these results we provide a short-time large deviation principle for stochastic volatility models, where the volatility is expressed as a function of a Volterra process. This large deviation principle holds under suitable conditions, but does not require any self-similarity assumption on the Volterra process. As examples we consider models where the volatility is given as a function of a log-modulated fractional Brownian motion and models where it is given as a function of a fractional Ornstein-Uhlenbeck process.

**References:**


**Keywords:** Large Deviations, small time, stochastic volatility.

---

**Discover Hidden Trends and Patterns using Text Visualization**

*Grigoris Papageorgiou¹, Polychronis Economou¹, and Sotiris Bersimis²*

¹Department of Civil Engineering, University of Patras, Greece

²Department of Business Administration, University of Piraeus, Greece

Text visualization is a useful technique to understand and gain insights to the characteristics and the hidden information of textual data since their volume and nature makes the traditional data visualization techniques not suitable. Moreover, using dynamic text visualization approaches to monitor text data (by taking into account the time dependency of the text - for example date of publication, or date an email is received) can identify seasonal patterns and hidden trends that draw attention to important or exceptional situations and reveal key aspects of process dynamics. More
specifically, monitoring a company’s text data, as for example in the shipping sector, might anticipate future observations, highlight market inconsistencies and spot significant shifts in market patterns which may affect the company’s decisions. In the current work, various strategies are investigated and used to visualize and monitor text data on executed orders from the shipping industry across trading regions. The scope of the paper is to identify hidden trends and seasonal patterns as well as potential anomalies, such as outliers and break points. 

**Keywords:** Text visualization, trends, patterns.

**Assessing Pollution Risk Using Asymmetric GARCH Models and Dynamic Correlation**

Giuliana Passamani¹ and Paola Masotti  
¹University of Trento, Trento, Italy

Exposure to high levels of pollution represents a major risk factor, as it may produce adverse health effects documented by numerous studies (e.g. Fuller et al. 2022). In order to work towards reaching SDG target 3.9.1, that aims to reduction in illnesses and deaths attributed to ambient air pollution, we need to define a model to measure risk associated to it and to predict this risk. Given that ambient air pollution results from complex mixtures of air pollutants emitted from various activities, an approximation of their combined effects and impacts on health is possible if we could assume some form of independence and little correlation between the pollutants. However, there are some limitations in estimating these joint effects given nonlinear interactions among pollutants and with their impacts. For this purpose, we consider the extension of time-varying volatility models of time series data, to dynamic multivariate autoregressive models, where a dynamic conditional correlation analysis method is proposed (Engle, 2002) that was developed for measuring and hedging financial risk. The data set consists of daily standardized concentrations, over two years, on three pollutants, PM10, NO2 and O3, which are interrelated and represent the so-called photochemical pollution factor. The three variables are observed at multiple monitoring sites within a geographical area. When necessary, the stochastic trends in the observed nonstationary variables are estimated using a smooth-trend unobserved component model and we use the estimated trends to de-trend the variables to make them stationary. As observed, pollutants concentrations show the presence of asymmetric GARCH effects and we explore whether the use of a multivariate asymmetric GARCH dynamic conditional correlation model can lead to a more accurate risk prediction for air pollution. In particular, we aim to determine how positive shocks to the observed pollutants can increase health risk. Interesting results emerge for particulate matter and ozone, both of which have great effects on human health. 

**Keywords:** Air pollution, Asymmetric heteroskedastic model, Dynamic conditional correlation.

**Statistically Explainable Artificial Intelligence: Influential Features & Leverage Instances**

P. Pefanis¹, A. Karagrigoriou¹, A. Effraimis², and V. Georgiou²  
¹Laboratory of Statistics & Data Analysis, Univ. of the Aegean, Greece  
²IRI, Analytics, Athens, Greece

In recent years the need to find ways to explain the effects and consequences of artificial intelligence (AI) has been given considerable attention. As a result, the
Explainable Artificial Intelligence (EAI) has surfaced the goal of which is to reveal the rationale behind the output of machine learning algorithmic procedures, like for instance, the well-known Random Forest. To make this possible, various methods like model-agnostic methods have been developed in an attempt to explain the outcomes of AI models. In this work we briefly present the most widely used machine learning algorithms and describe the classical model-agnostic methods discussing their pros and cons. In an attempt, to ensure the statistical significance of model-agnostic methods we propose in this work, extensions of the Break Down (BD) and Permutation Feature Importance (PFI) methods and discuss their advantages and their implementations in real case scenarios. The methods rely on permutations and bootstrap samples from the original data set and on the concept of leverage for the identification of both influential features as well as influential instances (points).

**Keywords:** Artificial Intelligence, Bootstrap, Break Down, Explainable Artificial Intelligence, Leverage, Permutations, Permutation Feature Importance.

The Impact of the COVID-19 Pandemic on the Financing and Financial Management of Public Hospitals in Greece

*Dimitrios Petrakis¹ and George Matalliotakis²*

¹²Hellenic Open University

The COVID-19 pandemic, in addition to being a catastrophic global health crisis, has caused disharmony in the operation of Greek public hospitals, especially at an administrative and financial level. Specifically, due to the excessive increase in demand for hospital services, challenges arose related to meeting the needs of hospitals in logistical equipment and staffing, which threatened the smooth operation and sustainability of an already tested health system. To address this threat, measures and policies were implemented to support the Greek National Health System. The purpose of this thesis is to describe the effects of the pandemic on the efficiency and effectiveness of Greek public hospitals. The method used to outline the degree to which Greek hospitals were affected by the health crisis and the subsequent support measures is the analysis, using financial ratios, of the published financial statements of reference hospitals by Health Region, while the main sources of economic data are the databases of the Ministry of Health, the Diaygeia program, the B.I.forms system and the Hellenic Statistical Service (ELSTAT). In addition, data was gathered on the nursing data of the hospitals in order to correlate the treatment burden and the change in costs and expenses. During the ongoing COVID-19 pandemic, public hospital revenues are heavily dependent on government funding in order to cope with the high costs that have incurred. The successful response of hospitals to future crises depends on the adoption of structural changes of the national healthcare system. Furthermore, the two fundamental weaknesses of the Greek hospitals: delay in the collection of hospital claims and increased expenses should be addressed with great care. The healthcare stakeholders in Greece should ponder on the chronic phenomenon of the inability to collect hospital claims from the main insurance company (EOPYY), the adoption of a more realistic reimbursement system such as the DRG’s but also the cutdown of expenses with the use of biosimilar drugs and the use of telehealth for non-emergency chronic patients.
The Cox Model and Beyond in the Prediction of Business Failures

Francesca Pierri\textsuperscript{1} and Chrys Caroni\textsuperscript{2}
\textsuperscript{1}Department of Economics, University of Perugia, Italy
\textsuperscript{2}Department of Mathematics, National Technical University of Athens, Greece

The basic form of lifetime data consists of durations from a starting point until the occurrence of the event under study; that is, it represents the first passage time from an initial state to a final absorbing state. The well-known extension to a situation of competing risks arises when there are two or more mutually exclusive absorbing states and the occurrence of any one of these terminates the data generating process. Also encountered in applications are multistage models, in which the lifetime process can be represented as a one-way progression through a sequence of transient states to the final outcome. More generally, multistate models may be employed, including competing risks and multistage models as special cases. In all these cases, assuming proportional hazards may play a central part in the analysis. Cox’s semi-parametric proportional hazards regression model has now been part of statistical methodology for 50 years. Even though Cox himself has stated that the concept of proportional hazards lacks a physical and substantive basis, the model’s simplicity and apparently easy interpretation means that it has achieved a dominant position in the analysis of lifetime data (time-to-event data), especially in biomedical applications. However, in some respects, statistical modelling is now a parallel activity to machine learning. Several studies indicate that machine learning techniques tend to outperform Cox regression as far as predictions are concerned. The purpose of this talk is to review, apply and discuss modern applications of the Cox model, in the context of prediction of the failure of commercial enterprises.

On a Fractional Stochastic Risk Model with a Random Initial Surplus and a Multi-Layer Strategy

Enrica Pirozzi
Department of Mathematics and Applications "R.Caccioppoli", University of Naples Federico II, Italy

The paper deals with a fractional time-changed stochastic risk model, including stochastic premiums, dividends and also a stochastic initial surplus as a capital derived from a previous investment. The inverse of a $\nu$-stable subordinator is used for the time-change. The submartingale property is assumed to guarantee the net-profit condition. The long-range dependence behavior is proven. The infinite-horizon ruin probability, a specialized version of the Gerber–Shiu function, is considered and investigated. In particular, we prove that the distribution function of the infinite horizon ruin time satisfies an integral-differential equation. The case of the dividends paid according to a multi-layer dividend strategy is also considered.
Deep Learning and Surplus Production Models for Multivariate Autoregressive Modelling and Simulation of the Jack Mackerel Fishery Associated with Environmental Conditions

Francisco Plaza¹, ², Eleuterio Yáñez³, Pierre Freón⁴, Héctor Araya⁵, and Antonio Aranis¹

¹Instituto de Fomento Pesquero, Chile
²Universidad de Valparaíso, Chile
³Profesor Titular, Pontificia Universidad Católica de Valparaíso, Chile
⁴Fisheries Scientific, France
⁵Universidad Adolfo Ibáñez, Chile

This work studies jack mackerel fishery in Southeastern Pacific (SEP) from 1973-2020, under single-stock unit assumption, the following variables: i) total SEP catch (C); ii) catch per standard unit of effort (CPUE) as abundance index of industrial purse-seine fleet off central-southern Chile; iii) standard fishing effort (E = C/CPUE); and iv) NOAA satellite sea surface temperature (SST) between 32° - 42°S and 71° - 80°W. The approach proposes modelling and simulation procedures for both surplus production models and deep-learning-based models with environmental variability integration. The implementation of surplus models was performed using CLIMPROD software, proposing several ad-hoc models; the deep learning approximation, considers sets of convolutional neural network architectures. Selected models were used for simulation of environmental and fisheries management scenarios resulting in different maximum sustainable yields (MSY). Results show a proposed surplus model that considers E and SST as explaining variables: \( CPUE = (-a + b \, TSM) \exp(-cE) \), assuming landings containing 5-year age classes, recruitment at 2 years age, and environment affecting abundance, with good performance (\( R^2 = 0.86 \)). Two environmental scenarios are then proposed, using SST average from 1999 to 2014, and a period not-warmer-than 13°C, obtaining a MSY of 563,476 and 945,446 tons, for both scenarios, respectively. Regarding the deep learning approach, using monthly data, an autoregressive multivariate convolutional neural network is implemented that considers 36-months lagged SST and fishing effort, obtaining good performance (\( R^2 = 0.84 \)), and capture simulations that range from 481,656 to 1,167,933 tons, on different scenarios. The latter models consider longer autoregressive periods and long memory properties of time series, this approach could be used for short to long-term forecasts, helping with national to international fisheries sustainability-based policies.

Keywords: deep learning, surplus models, MSY, jack mackerel, Southeast Pacific.

Investigation of Renewal-Reward Process with Dependent Components

Aynura Poladova¹, Salih Tekin¹, and Zulfiye Hanalioglu²

¹Department of Industrial Engineering, TOBB University of Economics and Technology, Ankara, Turkey
²Department of Actuarial Sciences, Karabuk University, Demir Celik Campus, Karabuk, Turkey

A renewal-reward process with dependent components and a discrete interference of chance is addressed and investigated in this study. This process is also a mathematical representation for a stochastic inventory model of type (s, S). In this study, it is assumed that the discrete chance interference is expressed with a truncated
normal distribution. In this case, the exact expressions for the ergodic distribution and moments of ergodic distribution of this process are attained. Moreover, the weak convergence theorem is proved for ergodic distribution of the process. In addition, the asymptotic behavior of all ergodic moments of this process is analyzed by using Laplace and Laplace-Stiltjes transforms. As a result of the analysis, two-term asymptotic expansions for ergodic moments are revealed. Finally, in order to show that the asymptotic expansions are close enough to the exact formulas, a special example is considered and the accuracy of the asymptotic results is examined on this example. **Keywords:** Renewal-Reward Process, Dependent Components, Normally Distributed Interference of Chance, Asymptotic Expansion.

**Some Relations between the Renewal and the Harmonic Renewal Function**

**Konstadinos Politis**  
*University of Piraeus, Greece*

A special form of the so-called generalized renewal functions is the harmonic renewal function. Although the latter lacks the applicability of the (ordinary) renewal function in a wide range of applied probability models, it bears a close connection with random walks. In particular, various identities exist that link the behaviour of the harmonic renewal function with the distribution of the (strictly) ascending ladder epoch in a random walk with nonnegative drift. However, the relation between the renewal and the harmonic renewal function appears not to have been studied in detail. In the present paper, we obtain some formulae which link these two functions. We also obtain two-sided bounds for the harmonic renewal function and, in the absolutely continuous case, we discuss the behaviour of its density. This is a joint work with Vaios Dermitzakis and Sotiris Losidis  
**Keywords:** Renewal function, Harmonic renewal function, First passage time.

**Studying the European Union Innovation Index and its Affecting Factors**

**Georgios Pozios**¹, **Menita Masoura**², and **Sonia Malefaki**²  
¹Hellenic Open University, School of Science and Technology, Patra, Greece  
²Department of Mechanical Engineering and Aeronautics, School of Engineering, University of Patras Patra, Greece

The European Innovation Scoreboard (EIS) is a tool developed at the initiative of the European Commission to provide a comparative assessment of the innovation performance of European Union Member States on an annual basis through the Summary Innovation Index (SII). The assessment is based on a wide range of indicators covering four key innovation activities. The first category examines the main components of innovation related to the external environment of an organization, the second category captures public and private investments made in Research and Innovation, the third category outlines different types of innovations and intellectual property rights, and the fourth category captures the impact of innovation on social, economic, and environmental level. In the current work, we attempt to assess the level as well as the evolution of innovation in the European Union (EU) by studying the four main activity categories of the SII for the years 2014-2021. Additionally, our aim is to group EU countries according to their performance in the four main categories by applying the K-means clustering method. EU member states are then classified in two
different groups in terms of performance in the four categories presenting one group with High and one group with Low performance respectively, since the optimal number of clusters is two. The evolution of each member country and the possible transitions from one group to another during the years 2014-2021 is also a point of interest. The grouping of EU member states into the two groups showed that socio-economic factors may affect the overall SII. Linear Mixed Effect Models confirm the effect of Gross Domestic Product per capita, the average number of weekly working hours, renewable energy consumption, the public expenditure on education and the Digital Economy and Society Index on the SII, but this is not the case for unemployment, CO2 emissions and conventional energy consumption.

**Keywords:** Innovation, Summary Innovation Index (SII), Repeated Measures ANOVA, Clustering, K-means, Linear Mixed Effect Models.

Silver Economy in the Age of Seniors’ Activity

Klaudia Przybysz¹ and Agnieszka Stanimir²

¹,²Wroclaw University of Economics and Business, Wroclaw, Poland

In our study, we attempted to identify the features of seniors that may have a direct impact on their preferences. This, in turn, may provide an answer to the question of which activities in the area of the silver economy can be projects that satisfy these preferences and at the same time bring financial success for companies. The analysis is based on a questionnaire - pilot study conducted on a sample of 57 seniors (60+) from large, medium and small towns South-West Poland (on the first part of the 2020 year). The respondents were selected by the snowball method. The pilot study was meant to introduce more extensive and detailed research, but the time of the pandemic made it impossible. The results obtained so far have allowed, however, a preliminary description of the life situation of seniors in Poland and constituted a point of reference for comparisons and analyzes conducted on a wider basis of European research (Eurostat database). Seniors are rather active. The diverse structure of this social group should be considered. Our research has shown that it is possible to identify areas where the activity of seniors can translate into the directions of development of the silver economy. Based on our research questions, we can conclude that the majority of older people perceive the concept of “senior” rather negatively and identify it with people older than themselves. This discovery may be crucial for the silver economy in terms of constructing offers for seniors.

**Keywords:** Seniors’ activity, Pilot study, Silver economy, Comparative analysis.

The Moderating Impact of Human Resource on ESG Performance of Financial Companies in Europe

Konstantina Ragazou¹,², Konstantinos Spinthiropoulos⁵, Alexandros Garefalakis²,³, Christos Papademetriou², and George Sklavos⁴

¹Department of Accounting and Finance, University of Western Macedonia, Kozani, Greece
²Department of Economics and Business, Neapolis University Pafos, Pafos, Cyprus
³Department of Business Administration and Tourism, Hellenic Mediterranean University, Heraklion, Greece
⁴Department of Business Administration, University of Thessaly, Larissa, Greece
⁵Department of Management Science and Technology, University of Western Macedonia, Kozani, Greece

Environmental, social, and governance (ESG) aspects are frequently associated with investor relations, although ESG encompasses all stakeholders. Companies are
becoming more aware of the critical role that human resources (HR) play in ESG strategy. Although the role of the HR departments as a whole has evolved over the course of the previous two years due to the pandemic, heightened societal awareness, and the Great Resignation. In this study, we investigate how human resources affect financial businesses' performance in the areas of governance, the environment, and social responsibility (ESG). Based on the large dataset, the current study offers the opportunity to more clearly understand this influence. Findings imply that the value of a corporation is considerably and favorably correlated with its entire human resource. However, the nomination of employees to the board of directors has a greater impact on how the market views an organization's environmental, social, and governance (ESG) performance.

Keywords: Board of director, Efficiency, ESG, Financial sector, Human Resources, Performance

Non-standard Computational Analysis of the Stochastic COVID-19 Pandemic Model: An Application of Computational Biology

Ali Raza¹, Chanda Shahzadi, and Rashid Javed
¹Department of Mathematics, Govt. Maulana Zafar Ali Khan Graduate College Wazirabad, Punjab Higher Education Department (PHED), Lahore, Pakistan

The present study is conducted to analyse the computational dynamical analysis of the stochastic susceptible-infected-recovered pandemic model of the novel coronavirus. We adopted two ways for stochastic modelling like as transition probabilities and parametric perturbation techniques. We applied different and well-known computational methods like Euler Maruyama, stochastic Euler, and stochastic Runge Kutta to study the dynamics of the model mentioned above. Unfortunately, these computational methods do not restore the dynamical properties of the model like positivity, boundedness, consistency, and stability in the sense of biological reasoning, as desired. Then, for the given stochastic model, we developed a stochastic non-standard finite difference method. Following that, several theorems are presented to support the proposed method, which is shown to satisfy all of the model's dynamical properties. To that end several simulations are presented to compare the proposed method's efficiency to that of existing stochastic methods.

Keywords: Coronavirus pandemic model, Stochastic ordinary differential equations, Numerical methods, Convergence analysis.

Network of Adaptive Frequency Oscillators in a Ballistic (i.e., non-Gaussian) Noisy Environment

Julio Rodriguez¹,²
¹School of Engineering and Computer Science, Bern University of Applied Sciences (Berner Fachhochschule - BFH), Biel, Switzerland
²Swissgrid AG, Aarau, Switzerland

Frequency adaptation (i.e., the “plasticity” of oscillators to adapt their own frequency) has been broadly studied in the realm of deterministic dynamics. However, random perturbations can strongly impact their capability to adapt – hence the interest to investigate this phenomenon of adaptation in noisy environments specifically. Here, based on the well-known Kuramoto model, a collection of coupled phase oscillators
with adapting frequencies is considered. Analogous to other unsupervised learning mechanisms, adaptation happens depending on the state of their respective neighbors’ phases. The global dynamics is set in a fluctuating environment with a noise source emanating from non-Gaussian, super-diffusive (i.e., ballistic component), stochastic processes. For the linearized system, one can analytically calculate the mean and covariance matrix of the resulting stochastic process and explicitly observe the influence of the non-Gaussian noise source. Looking to the real world, these types of models can for example be used to investigate the influence of stochastic environments impacting the network frequency of electrical grids, where the ballistic input corresponds to random run-of-rivers and the equations can be generalized to incorporate the typical class of “Generator” and “Load” models.

**Keywords:** Coupled phase oscillators, Adaptive mechanism, Kuramoto model, Super-diffusive noise source, Laplacian matrices.

---

**Business Density and Economic Growth in Portugal**

Cátia Rosário\(^1\) and Ana Lorga da Silva\(^2\)

\(^1\)Escola de Ciências Económicas e das Organizações; Lusófona University, Campo Grande Lisboa, Portugal & Governance, Competitiveness and Public Policies (GOVCOPP), Universidade de Aveiro, Campus Universitário de Santiago, Aveiro, Portugal

\(^2\)Departamento de Gestão Portuária, Escola Superior Náutica Infante D. Henrique, Paço de Arcos, Portugal & ECEO, Lusófona University, Campo Grande, Lisboa, Portugal

According to Eurostat, the Portuguese business fabric is dominated by small companies. Out of a total of 833,028 small and medium-sized enterprises (SMEs) in the country, 99.3% are small and the rest medium. These companies seek to be strategically located both to capture customers, establish links with suppliers and distribution chains, attract qualified labour and create synergies with other stakeholders. Several business associations in the interior of the country warn of the number of companies that are unable to attract workers. There are many vacancies to be filled and more and more. In some cases, it is already holding back investments, preventing business expansion and delaying deliveries of orders and works.

This study analyzes the business concentration by area (km\(^2\)) and by inhabitants (100 inhabitants), differentiated by NUTS II, in order to distinguish the business density of the 7 regions. The analysis is also carried out for the unemployed population, individuals’ gross income, the difference between the national minimum wage and the monthly remuneration of employees and gross added value. Finally, an analysis using structural equation models is presented on the impact that business density has on value creation, measured through the indicators.

**Keywords:** Business density, Economic growth, Value creation, Structural equation models.

---

**Finding a Population having the Better Distribution**

Sheldon Ross and Tianchi Zhao

*University of Southern California, U.S.*

Suppose there are an infinite number of populations, with each population having an associated probability distribution, and that sampling from a population incurs a cost 1 and yields the value of a random variable having the population distribution. Suppose
there are only 2 possible population distributions, one having density (or mass) function \( f \) and the other having density (or mass) function \( g \). Suppose that \( f \) is the density of population \( i \) with probability \( \alpha_i \) when \( i = 1, \ldots, k \) and with probability \( \alpha \) when \( i > k \). At each stage one selects one of the populations to sample from, with this continuing until there is a population whose posterior probability of having distribution \( f \) is at least some specified value. The objective is to find a sampling strategy that minimizes the expected cost until this occurs. We also consider the problem where one may stop at any time and choose a population, with a reward of \( R \) earned if that population has distribution \( f \). The optimal policy is explicitly found in the former case and properties of it in the latter case.

### Impact of COVID-19 Pandemic on Mortality of Elderly in Gender Aspect

**Tamara Sabgayda\(^1\)\(^2\), Alexander Zubko\(^1\)\(^2\), and Victoria Semyonova\(^1\)**

\(^1\)Institute for Demographic Research – Branch of the Federal Center of Theoretical and Applied Sociology of the Russian Academy of Sciences, Moscow, Russia

\(^2\)Russian Research Institute of Health, Moscow, Russia

The pandemic has led to an increase in mortality which is more pronounced among elderly. When analyzing SDR for men and women aged 60+ living in urban areas, data of official statistics of ROSSTAT for 2019-2021 were used. Impact of access to medical care and medicines was analyzed by comparing changes in mortality between the capital and the Russia as a whole. From 2019 to 2021, mortality increased by 36.6% among men and by 45.4% among women in Moscow and by 35.2% and 46.0% in the country. The greatest differences in increase rates in mortality in Moscow and the country are noted for residents aged 85+. During two years of the pandemic, reduction in LE60 among Muscovites is greater than the national average (4.0 years vs 2.7 for men and 3.5 years vs 3.3 for women). From 2020 to 2021, COVID-19 deaths increased 2.0 times for men and 2.9 times for women in Moscow and 2.6 and 3.7 times in the country. Mortality from other causes decreased by 9.1% among men and by 6.5% among women in Moscow and by 5.2% and 0.38% in the country, remaining above 2019 level. Mortality trends for various causes do not coincide for men and women in Moscow and in the country as a whole, only mortality from nervous diseases has increased in all cases. Changes in the structure of death causes are largely due to different practices in choosing underlying cause in case of COVID-19. For pathologies that led to death after COVID-19 in long term, mortality trends in Moscow coincide with national average, but not always for chronic diseases, which is determined by availability of medical care. It can be assumed that the consequences of COVID-19 are more dangerous for women.

**Keywords:** Structure of causes of death, Long-term consequences, Urban population, Trends in mortality, Life expectancy at the age of 60.

### Comparative Analysis of Hybrid MIDAS-SARIMA Models for GDP Forecasting: Empirical Evidence from Palestine

**Samir K. Safi\(^1\) and Olajide Idris Sanusi\(^1\)**

\(^1\)United Arab Emirates University, UAE

In this research, the focus is on predicting Palestine's quarterly gross domestic growth (GDP) using monthly consumer price index (CPI) and monthly exchange rate as explanatory variables. However, since the data are observed at different frequencies,
aligning them in time presents a challenge. To address this issue, the researchers use a Mixed-data sampling (MIDAS) regression model, specifically the mixed-data sampling-seasonal autoregressive integrated moving average (MIDAS-SARIMA) model. The data used cover the period from January 2000 to December 2020. The study employed three different functions, namely Unrestricted MIDAS, Beta function MIDAS, and Normalized exponential Almon lag polynomial MIDAS, and the results indicate that the fitted hybrid MIDAS-SARIMA model produced better results in terms of log-likelihood values compared to the hybrid MIDAS-ARIMA model due to the seasonality effect. Additionally, the forecast accuracy measures between the three functions of the hybrid MIDAS-SARIMA models considered indicate a smaller mean square error (MSE) value for the hybrid MIDAS-SARIMA with Beta function than the MIDAS-SARIMA with Normalized exponential Almon lag polynomial and Unrestricted hybrid MIDAS-SARIMA model. Overall, this research highlights the potential of hybrid MIDAS-SARIMA models in forecasting GDP using monthly data with different frequencies. 

**Keywords:** MIDAS, Hybrid, SARIMA-MIDAS, Forecasting.

**Assessing Path-dependence in Socioeconomic Processes: The Contribution of Spatial Econometrics and Exploratory Multivariate Analysis**

Luca Salvati

Department of Methods and Models for Territory, Economics and Finance, Faculty of Economics, Sapienza University of Rome, Italy

Although economic forces have demonstrated to influence the geographical distribution of resident population, non-economic (historical, social, environmental) drivers were documented to significantly impact spatial direction and intensity of urban expansion. In this perspective, analysis of path-dependent population growth may refine the comprehension of socioeconomic processes underlying urban expansion in advanced economies. Despite a long settlement history, empirical investigations of the role of path dependence in the long-term evolution of human populations are rather scarce in Europe, and especially in the Mediterranean countries. Our study discusses the empirical evidence stemming from a statistical analysis of population growth rates in 115 districts of metropolitan Athens, Greece, discriminating path-dependent growth from the impact of other socioeconomic forces on long-term urban expansion (1920-2020), using spatially explicit (global and local) econometric approaches. The empirical findings of this study clarify how path-dependent regulation of population growth was heterogeneous over time and space, depending on the specific stage of the city life cycle. After an initial period with path-independent population expansion reflecting the inherent impact of exogenous shocks, path-dependent growth was particularly evident under compact urbanization – with population settling in central locations and economic dynamics governed by agglomeration and scale advantages. Path-dependent growth was instead less intense with suburbanization – when population sprawled over larger areas. An improved understanding of path-dependent mechanisms of population growth contributes to rethink spatial planning and development strategies adapting to transitional contexts and heterogeneous socioeconomic conditions at local scale.
Efficient Calibration of a Markov Regime Switching Model: An Application to Variable Annuity Pricing

Mohammad Sarfraz\textsuperscript{1}, Guglielmo D'Amico\textsuperscript{2}, and Dharmaraja Selvamuthu\textsuperscript{3}
\textsuperscript{1,3}Department of Mathematics, IIT Delhi, New Delhi, India
\textsuperscript{2}Department of Economics, University G. d'Annunzio of Chieti-Pescara, Pescara, Italy

In view of the increasing complexity of financial markets and the need to better monitor and manage risk, the pricing of variable annuity guarantees under regime-switching hidden Markov models has become an important research topic. This study demonstrates the importance of efficient calibration methods for financial models and provides a practical approach to calibrating the Markov regime switching model for variable annuity pricing. The methodology seeks to capture variations in asset price correlation dynamics driven by various market conditions, such as bull and bear markets, that can lead to non-linear movements in asset prices. We present our model for pricing variable annuity guarantees under regime-switching model and use an empirical test on historical S&P500 Index data to illustrate its practical implications. Overall, this paper proposes a new approach for pricing variable annuity guarantees linked to complex indices like the S&P500, while carefully considering various market situations that might affect dependence between assets in the index basket. Also, this approach significantly reduces the computational cost caused by the introduction of independent regimes.

\textbf{Keywords:} Variable Annuity, Markov Regime-switching, Calibration, Independent Regimes.

Graphical Estimation of Multivariate Count Time Series

Vurukonda Sathish\textsuperscript{1}, Debraj Chakraborty\textsuperscript{2}, and Siuli Mukhopadhyay\textsuperscript{3}
\textsuperscript{1}Department of Electrical Engineering Indian Institute of Technology Bombay, Powai, India
\textsuperscript{2}Professor, Department of Electrical Engineering Indian Institute of Technology Bombay, Powai, India
\textsuperscript{3}Professor, Department of Mathematics Indian Institute of Technology Bombay, Powai, India

The problems of selecting partial correlation and causality graphs for count data are considered. A parameter driven generalized linear model is used to describe the observed multivariate time series of counts. Partial correlation and causality graphs corresponding to this model explain the dependencies between each time series of the multivariate count data. In order to estimate these graphs with tunable sparsity, an appropriate likelihood function maximization is regularized with an $\ell_1$-type constraint. A novel MCEM algorithm is proposed to iteratively solve this regularized MLE. Asymptotic convergence results are proved for the sequence generated by the proposed MCEM algorithm with $\ell_1$-type regularization. The algorithm is first successfully tested on simulated data. Thereafter, it is applied to observed weekly dengue disease counts from each administrative block of an Indian metro city. The interdependence of various blocks in the proliferation of the disease is characterized by the edges of the inferred partial correlation graph. On the other hand, the relative roles of various blocks as sources and sinks of dengue spread is quantified by the number and weights of the directed edges originating from and incident upon each
block. From these estimated graphs, it is observed that some special blocks act as epicentres of dengue spread even though their disease counts are relatively low.

Social Consequences of the COVID-19 Pandemic in Moscow and St. Petersburg: Similarities and Differences

V.G. Semyonova1,2, A.E Ivanova2, T.P. Sabgaida1,2, A.V. Zubko1,2, and G.N. Evdokushkina1,2
1Institute for Demographic Research – Branch of the Federal Center of Theoretical and Applied Sociology of the Russian Academy of Sciences, Moscow, Russia
2Russian Research Institute of Health, Moscow, Russia

Rising incidence of COVID-19 suggests that the pandemic is yet to shrink into insignificance, substantiating the research topic. The pandemic-induced issues remain mainly unsolved, social aspects in particular. The purpose of the study is to compare social consequences of the 2-year pandemic in Moscow and St. Petersburg by age and nosology (suicides with latent component, alcohol and drug-related deaths). The analysis is based on adult mortality Rosstat data in both capitals in 2019-2021, calculated in FAISS-Potential. The analysis shows that, first, in St. Petersburg, suicide mortality increased in almost all age/gender groups, while in Moscow— in females only against decreased male deaths. The rates of mortality growth in St. Petersburg mainly exceeded those in Moscow. Second, the cumulative alcohol-related mortality in Moscow increased in all age/gender groups, except for population aged 75+, against a steady decline in St. Petersburg with most significant disparities among young people. Third, both capitals have registered increased drug-related mortality among young and middle-aged population with unprecedentedly high growth in negative trends in Moscow: the indicators increased 4.3 in males and 3.4 times in females aged 15-29 versus 6.4% and 59% increase in St. Petersburg. However, the noted ultra–high rates of mortality growth in Moscow are rather due to a large-scale systematic underestimation of “Symptoms, signs and ill-defined conditions”, namely “R99–unknown cause of mortality” in the previous period than actual negative shifts. Fourth, the main loss from socially determined causes was formed in the first year of the pandemic (except for suicide-associated mortality in Moscow and St. Petersburg. Thus, it can be stated that social consequences of the pandemic in both capitals with a shared increase in drug-related deaths had pronounced specifics: in Moscow they are associated with increased alcohol-related mortality, while in St. Petersburg – with increased suicide-related mortality, with the first year of the pandemic being crucial for both capitals.

Keywords: social consequences of the COVID-19 pandemic, suicides, alcohol and drug-related deaths, suicides with latent component, alcohol and drug-related deaths.

Assessing Depression Risk using Subjective and Objective Health Indicators: A Cox Regression Analysis

Eleni Serafetinidou1 and Christina Parpoula1
1Department of Psychology, Panteion University of Social and Political Sciences, Athens, Greece

Depression is a prevalent mental illness that seriously affects physical and mental health. Gender differences in depression first emerge in adolescence and continue into
old age. The limited findings in adulthood are inconsistent with respect to both the magnitude and direction of the gender difference in depression. Additionally, estimates of gender difference in depression in older adults suggest marked variability. Thus, recent studies use a developmental perspective to elucidate patterns of gender differences across the lifespan. In this context, the aim of the study was twofold: (1) to estimate depression risk among older adults (aged 50 and older) having poorer physical health and (2) to explore the effect of various subjective and objective health indicators on depression risk (focusing on gender differences). For this purpose, data provided form the Survey of Health, Aging and Retirement in Europe (SHARE, fourth wave) were used (multiple values were imputed for missing data), and a COX regression analysis was performed to assess the study explanatory variables’ effect on survival (occurrence) time of depression. Our findings suggest that hazard (depression) rate is 48% higher for females compared to males, and for an additional year of age, the depression hazard is reduced by 7.4% in total. Physical health measures (both subjective and objective) put people (especially males) at higher risk for depression. Further, lower self-rated writing skills slightly increase depression risk for males, whereas better orientation scores (objective factor) reduce depression risk for females. Cognitive self-rated reading skills and behavioral risk of smoking status were not found to affect depression hazard ratio for both genders. Physical inactivity was found to affect only men (decreasing the depression risk contrary to previous findings). Finally, self-reported life satisfaction was found to decrease the most depression burden for both genders (with a slightly higher decrease being observed for males)

**Keywords:** Cox regression, Depression, Objective health, Risk factors, Self-reported health status, SHARE.

---

**Analysis of Stochastic Anthrax Epidemic Model**

**Huma Shabir**\(^1\) and **Ali Raza**\(^2\)

\(^1\)School Education Department, Government of Punjab, Pakistan

\(^2\)Department of Mathematics, Govt. Maulana Zafar Ali Khan Graduate College Wazirabad, Punjab and Higher Education Department (PHED) Lahore, Pakistan

Anthrax is the most contagious disease around the world. The mathematical modeling of an anthrax disease includes the four-compartment of the population as susceptible animals (s), infected animals (i), carcasses animals (c), and grams spores of animals in the environment (a). The continuous model analysis is studied rigorously. The stochastic model is based on the ideas of transition probabilities and parametric perturbation technique. The fundamental properties of the model with standard computational methods are studied. Unfortunately, these methods are time-dependent and even valid for short period analysis of the disease. Then, a non-standard computational along with the dynamical properties (positivity, boundedness, and dynamical consistency) is studied thoroughly. In the end, numerical results of the proposed method with the existing techniques are provided.

**Keywords:** Anthrax disease, SIAC model; stochastic differential equations; stochastic methods; convergence analysis.
Dynamical Analysis of Discrete Stochastic SIR Model

Chanda Shahzadi\textsuperscript{1} and Ihtasham\textsuperscript{1}

\textsuperscript{1}National College of Business Administration and Economics Lahore, Pakistan

The paper presents a dynamic analysis of a discrete stochastic SIR model, a mathematical model used to study the spread of infectious diseases in a population. The model is based on the classical SIR model, which divides the population into three groups: susceptible, infected, and recovered. The stochastic SIR model is an extension of the classical SIR model that considers random fluctuations in the transmission and recovery rates of the disease. The model is described by a set of differential equations that describe the evolution of the number of individuals in each of the three groups over time. The authors first derive the primary reproduction number $R_0$, which measures the potential for the disease to spread in a population. They then use a combination of analytical and numerical methods to analyze the dynamic behavior of the model. The authors show that the stochastic SIR model exhibits complex dynamical behavior, including multiple steady states and limit cycles. They also show that the initial conditions and the values of the model parameters strongly influence the dynamics of the model. The authors conclude that the stochastic SIR model provides a valuable framework for studying the spread of infectious diseases in a population. They suggest that future research should develop more sophisticated models considering additional factors, such as spatial structure and the effects of interventions such as vaccination and quarantine.

\textbf{Keywords.} Stochastic SIR model, dynamical analysis, infectious diseases, basic reproduction number, difference equations, steady states, limit cycles, initial conditions, model parameters, interventions, vaccination, quarantine.

An Efficient Tensor Regression for High-dimensional Data

Yuefeng Si\textsuperscript{1}, Yingying Zhang\textsuperscript{2}, and Guodong Li\textsuperscript{1}

\textsuperscript{1}Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong

\textsuperscript{2}Academy of Statistics and Interdisciplinary Sciences, East China Normal University, Shanghai, China

Most currently used tensor regression models for high-dimensional data are based on Tucker decomposition, which has good properties but loses its efficiency in compressing tensors very quickly as the order of tensors increases, say greater than four or five. However, for the simplest tensor autoregression in handling time series data, its coefficient tensor already has the order of six. This paper revises a newly proposed tensor train (TT) decomposition and then applies it to tensor regression such that a nice statistical interpretation can be obtained. The new tensor regression can well match the data with hierarchical structures, and it even can lead to a better interpretation for the data with factorial structures, which are supposed to be better fitted by models with Tucker decomposition. More importantly, the new tensor regression can be easily applied to the case with higher order tensors since TT decomposition can compress the coefficient tensors much more efficiently. The methodology is also extended to tensor autoregression for time series data, and non-asymptotic properties are derived for the ordinary least squares estimations of both tensor regression and autoregression. A new algorithm is introduced to search for estimators, and its theoretical justification is also discussed. Theoretical and computational properties of the proposed methodology are verified by simulation studies, and the advantages over existing methods are illustrated by two real examples.
Comparison and Evaluation of Partitions in Hierarchical Cluster Validation

Osvaldo Silva¹, Áurea Sousa², and Helena Bacelar-Nicolau³
¹Universidade dos Açores, Department of Mathematics and Statistics, CICSNOVA.UAc, Portugal
²Universidade dos Açores, Department of Mathematics and Statistics, CEEAplA, Portugal
³Universidade de Lisboa (UL), Faculdade de Psicologia and Institute of Environmental Health (ISAMB/FM-UL), Portugal

In the context of Hierarchical Cluster Analysis, it is most relevant to evaluate the quality of the obtained clustering structures, since that clusters into the resulting partitions should be internally homogeneous and externally well separated. Internal homogeneity and external separability of clusters depend in large amount on the descriptive model of the elements to classify (e.g., on the nature of the variables) and on the clustering algorithms. For instance, aggregation criteria can sometimes impose a classification on the data that is not its true structure. Therefore, it is important to assess the quality of the partitions, especially those corresponding to the best cut-off levels of the dendrograms, with suitable validation indices.

This study, based on a dataset collected during the pandemic crisis of Covid-19, aims to analyze the clustering results quality of a typology of anxiety symptoms in students during their examination’s periods. Data analysis was carried out based on the affinity coefficient combined with classical and probabilistic aggregation criteria (these ones issued from a VL parametric family). The obtained results are compared, and the most significant partitions are chosen using several validation indices.

Keywords: Hierarchical cluster analysis, Affinity coefficient, VL methodology, Validation indices.

Optimal Investment-disinvestment Choices in Health-dependent Variable Annuity

Shakti Singh¹, Guglielmo D'Amico², and Dharmaraja Selvamuthu³
¹Department of Mathematics, IIT Delhi, New Delhi, India
²Department of Economics, University G. d'Annunzio of Chieti-Pescara, Pescara, Italy
³Department of Mathematics, IIT Delhi, New Delhi, India

The aim of this study is to evaluate a new insurance product, i.e., the variable annuity product, accompanied by the health status and the guaranteed lifelong withdrawal benefit (GLWB). Due to specific problems, the insurance sector is now one of the riskiest industries. The ageing of the population and rising medical service costs as a result of technological advancements are to blame for this. Thus, one of the most basic needs in the health insurance sector is to design an innovative product. Recently lot of work has been done in the field of variable annuity (VA) product, however, VA products doesn’t have the option of investment and disinvestment according to the will of the policyholder. Only classical investment criteria, such as the net present value is considered. Lack of this feature put the policyholder at a disadvantage of extracting the maximum benefit out of the product. So, we provide the benefit of initiating the
withdrawal regime according to the will and health status of the policyholder. Moreover, a mixed discrete-continuous time model is proposed in this article, which is used to determine the investment trigger and disinvestment trigger according to the market conditions.

**Keywords:** Variable Annuity, Guaranteed Lifetime Withdrawal Benefit, Health Status, Investment, Disinvestment.

---

**Detect Outliers, Break Points and Level Shifts: An Application in Shipping Industry**

Ekaterini Skamnia¹, Polychronis Economou¹, and Sotiris Bersimis²

¹Department of Civil Engineering, University of Patras, Greece  
²Department of Business Administration, University of Piraeus, Greece

Time series analysis is a crucial task in many scientific fields and monitoring such data can provide useful insights of the behavior of the process under study. For instance, the existence of trends, seasonal patterns or even the presence of outliers and/or level shifts could be affirmed. The identification of such characteristics can reveal important aspects of the dynamic of the process and highlight important or special instances. Concerning financial time series of a corporation, monitoring time series can not only provide predictions of the upcoming months, days or years but also reveal inconsistencies in their financial records. In this work, different approaches for detecting outliers, break points and/or level shifts in time series are discussed. These approaches are then applied in time series data originated from shipping industry concerning orders in a trading region, in order to identify possible outliers and change points when act complementary.

**Keywords:** Level Shift, break points, outliers.

---

**Detecting change points in multivariate data streams. A use case in shipping**

Kyriakos Skarlatos¹, Petros E. Maravelakis¹, and Sotiris Bersimis¹

¹Department of Business Administration, University of Piraeus, Greece

Data science relies heavily on changing point and outlier identification, which is an often phenomenon in multivariate statistical process control. The identification of shifting points or outliers is crucial in many applications, such as sound and image analysis, health status monitoring, and environmental analysis, in order to prevent data noise and identify dynamic system changes. The analysis of cases and scenarios where outliers or state changes take place at random times in time-ordered data is the main goal of this study which compares the performance of the well-known Mahalanobis distance with that of the Dunn Index and a new index. Moreover, the separation between two consecutive samples is used to determine the clustering identification (previous and current state).

**Keywords:** change points, mean shift, separability.
Expanding the Life Tables to Include the Healthy Life Expectancy. The Case of Norway

Christos H. Skiadas¹ and Yiannis Dimotikalis²
¹ManLab, Technical University of Crete, Chania, Crete, Greece
²Assistant Professor at Hellenic Mediterranean University, Crete, Greece

We provide a method to expand the Life Tables to include the Healthy Life Years Lost to Disability (HLYL), to estimate the Proportion with Disability and then apply the Sullivan Method to estimate the Healthy Life Expectancy (HLE) and the Healthy Life Years Lost (HLYL). One Template is provided. The Template is based on seven columns added to the right of life tables provided from Population Division of the United Nations for all countries expanded by adding another seven columns for estimating the proportion of disability and applying the Sullivan method for calculating the HLE and the HLYL. In the Template only the mx column is needed or alternatively the qx column. Applications Norway apply.

Keywords: Life Expectancy, Healthy Life Expectancy, Life Tables, Health Parameters, Healthy Life Years Lost.

Distribution of Gender Correspondence and Phonological Similarity between French, Greek, and Spanish Nouns

Maria-Sofia Sotiropoulou and Stuart Cornwell
Episcopal School of Acadiana, Broussard, Louisiana, USA

About half of the world languages have grammatical gender. For example, French has two genders, masculine and feminine, while German has three genders, masculine, feminine, and neuter. When we translate from one language to another, gender can either be kept or change. Investigating the correspondence of grammatical gender between languages contributes to the understanding of language processing in the bilingual mind and to guiding multilingual teaching and learning methods. In the present study, we examine the bilingual and trilingual distribution of gender correspondence and phonological similarity between French, Greek and Spanish nouns. The trilingual gender distribution comprises twelve gender corresponding triples as Greek has neuter gender in addition to masculine and feminine in French and Spanish. We find that a discrete Poisson distribution successfully models the distribution of gender corresponding triples of frequently spoken nouns in the three languages, as indicated by the p-value of 0.95 in chi-squared testing. 50% of feminine gender is kept as opposed to only 20% of masculine gender. This is partly attributed to Greek neuter, 50% of which becomes masculine gender in French and Spanish, 30% becomes feminine, and 20% becomes a gender that differs between French and Spanish. In contrast, we find that when there is phonological similarity as is the case for loan nouns, Greek neuter becomes predominantly masculine in French and Spanish, at 85%. It is known that in many cases noun phonological ending can predict gender in a language. This is true for French, Greek, and Spanish. For example, predominant are: [-a] for feminine in Spanish, [-o(n)] for neuter in Greek, and [-o, ō] for masculine in French. In our study, we find that there is also a considerable interlanguage gender similarity for specific phonological endings. 75% of French masculine [-o] (e.g. chapeau) are neuter in Greek (καπέλο [kapelō]), with masculine and feminine about equally much less, as this phonological ending is predominantly neuter in Greek. In contrast, French masculine [-ō] (e.g., accent, plant) are mostly feminine (προφορά [proforā]) (45%) and neuter (φυτό [fitō]) (45%) in Greek, as the nearest phonologically
similar ending in Greek is [-a], dominated by feminine and neuter. 80% of Greek feminine [-ia] are also feminine in French and Spanish, whose endings are predominantly /-ie/ [i] and /-ia/-[-ia], respectively, when they are loans, like φωνολογία [fonolojía] in Greek, phonologie in French, and fonología in Spanish. Our findings put in perspective experimentally obtained results on the acquisition and processing of grammatical gender in bilingualism and multilingualism. We hope that this will encourage investigations on interlanguage gender correspondence and phonological similarity between other languages that are not necessarily Indo-European languages or of the same origin, as was the case here.

**Keywords:** grammatical gender, phonological similarity, nouns, correspondence, French, Greek, Spanish.

**Comparison Between two Algorithms for Computing the Weighted Generalized Affinity Coefficient in the Case of Interval Data**

Áurea Sousa¹, Osvaldo Silva², and Helena Bacelar-Nicolau³  
¹Universidade dos Açores, Department of Mathematics and Statistics, CEEAplA, Portugal  
²Universidade dos Açores, Department of Mathematics and Statistics, CICSNOVA.UAc, Portugal  
³Universidade de Lisboa, Faculdade de Psicologia and Institute of Environmental Health (ISAMB/FM-UL), Lisboa, Portugal

In this study, we refer to the most significant partitions obtained with the hierarchical cluster analysis (h.c.a.) of two well-known data sets (macro-data), taken from the literature on complex (symbolic) data analysis. The h.c.a. is based on the weighted generalized affinity coefficient for the case of interval data, and on probabilistic aggregation criteria from a VL parametric family (Lerman, 2016; Bacelar-Nicolau et al., 2009; Sousa et al., 2010). Two alternative algorithms to calculate the values of this coefficient were used and compared. Both algorithms were able to detect clusters of macro-data that were consistent and consonant with those reported in the literature, but one performs better than the other in some specific cases. Moreover, both approaches allow the treatment of large microdata bases (non-aggregated data).

**Keywords:** Interval data, Hierarchical cluster analysis, Weighted generalized affinity coefficient.

**First Passage Time Problems for a Random Walk on the L-lattice**

Serena Spina  
Department of Mathematics, University of Salerno, Fisciano (SA), Italy

Random walks on graphs play a relevant role in the theory of stochastic processes, since they allow to study a variety of mathematical problems like diffusions on manifolds, harmonic analysis, infinite graph theory, group theory, etc (see, e.g., [3] and [4]). In particular, random walks on lattices and networks are used as simple models of physical systems. For instance (see [1] and [2]), the Manhattan lattice and the L-lattice are used to study the quantum and classical localization and ordinary integer quantum Hall transition, respectively.
The considered L-lattice is an unbounded square lattice consisting of directed edges and nodes, such that a particle that starts at a vertex \((x,y)\) moves to an adjacent vertex following an appropriate transition probability. In this framework, we study the random walk on the L-lattice defined according to the following rules. Let \(V_0\) (resp., \(V_1\)) be the set of points \((x,y)\) such that \(x+y\) is even (resp., odd), for \(x\) and \(y\) integer. Moreover, (i) if the particle is located in a vertex of \(V_0\), then it can reach one of the two adjacent positions on the right (with probability \(q\)) or on the left (with probability \(1-q\)); (ii) if the particle is located in a vertex of the set \(V_1\), then it can reach one of the two adjacent positions on the top (with probability \(p\)) or on the bottom (with probability \(1-p\)).

We firstly determine the probability generating functions related to the sets \(V_0\) and \(V_1\), the transition probabilities and the main moments of the described random walk. Then, we investigate the first-passage-time problem of the random walk through straight line boundaries. We also focus on the ‘taboo probabilities’ concerning transitions on the nodes of the L-lattice that avoid the boundary. Thanks to suitable symmetry properties, we can determine the taboo probabilities and, in turn, of the corresponding first-passage-time probabilities.

**References:**


**Keywords:** First passage Time, random walk, L-lattice.

---

**Some Results on First Passage Time Problems for Markov Processes**

Serena Spina

*Department of Mathematics, University of Salerno, Fisciano (SA), Italy*

**Intergenerational social mobility in Europe: Findings from the European Social Survey**

Glykeria Stamatopoulou\(^1\) and Maria Symeonaki\(^2\)

\(^1,^2\)Department of Social Policy, Panteion University of Social and Political Sciences, Athens, Greece

The present study displays patterns of intergenerational educational mobility in Europe using data drawn from the latest round of the European Social Survey (ESS)\(^1\) implemented during 2022. Due to the impact of the COVID-19 pandemic at Round 10, a total of 9 countries switched to a self-completion (web and paper) approach, while 22 countries used ESS’s usual face-to-face fieldwork approach. Our motivation stems from the fact that although social justice is theoretically promoted in the prevailing system of economic and social organization, studies on social mobility prove the difficulties of the lower working classes to move upward in the social ladder. Using raw data drawn from the ESS survey we investigate the relationship between the individual’s and the parents’ probability distributions and quantify the transitions of
Penalised Regression Adaptations of the Longstaff-Schwartz Algorithm for Pricing American Options

David Suda¹, Monique Borg Inguanez¹, and Lara Cilia²

¹Department of Statistics and Operations Research University of Malta, Malta
²Central Bank of Malta, Malta

One of the most popular techniques for evaluating the American put option is the Longstaff-Schwartz algorithm, where orthogonal polynomials are typically used to estimate the expected future payoff given the current value of the American option. An optimal exercise strategy then ensues for each of these paths, where the average payoff over all paths becomes equivalent to the fair price of the American option. Convergence results have been proven over the years which show that, under certain regularity conditions and using a least squares estimation approach, this average payoff converges to the true price as the sample size of the paths and the order of the orthogonal polynomial go simultaneously to infinity. A number of alternative modelling and estimation approaches have been attempted to make the Longstaff-Schwartz algorithm more accurate and computationally efficient. However, studies on the use of penalized regression approaches are scarce. In this paper, we conduct an empirical assessment of the benchmark least squares method in comparison with the Ridge, LASSO and Elastic Net estimation to see which of these methods is the best in terms of accuracy, under different sample path and polynomial order settings. We compare these methods on three staple processes in finance, namely geometric Brownian motion, the Heston stochastic volatility and the Meixner jump processes, while convergence properties for the standard Longstaff-Schwarz approach is used to determine a benchmark for accuracy. Across most settings in our simulation experiment design, we find that LASSO resulted in the best precision across the four algorithm variations, followed closely by Elastic Net. Ridge regression often produced results which were less accurate than LASSO and Elastic Net, however, these were also often more precise than the least squares approach. We close off the paper with a brief discussion on the computational aspect, which could be investigated more deeply in further studies.

Keywords: American option pricing, Penalised regression, Mathematical finance.

Analysis of Health Status in the Czech Republic: 2019-2022 Trends

Kornélia Svačinová¹ and Elena Říhová²

¹, ²Department of Quantitative Methods, ŠKODA Auto University, Czechia

Background: This study will monitor the health status in the Czech Republic during the last four years, including the impact of the Covid-19. Aims: To review the existing health status measures. To develop a new indicator of the health status focusing on
Czech data. An innovative aggregate health indicator will be calculated. Materials and methods: Literature review was conducted. The main data sources were primarily the Czech Statistical Office Database and the Institute of Health Information and Statistics of the Czech Republic. Results: The modified health status indicator was enumerated. The results for Czechia were compared with other EU countries. Conclusion: Measurement results may serve as recommendations towards social and health sector. Findings in this study may be considered in future research by analyzing further aspects of the population health status in Czechia.

**Keywords:** Czechia, Covid-19, Health, Health Status.

**Modeling and Forecasting Stochastic Volatility of Cryptocurrencies**

Maria Tantoula¹, Konstantinos Gkillas¹, Dimitrios Vortelinos¹, and Christos Floros¹  
¹Department of Accounting and Finance, Hellenic Mediterranean University, Greece

We propose novel non-parametric estimators for stochastic volatility and volatility of volatility of cryptocurrencies, based on traditional asset price models assumptions. We further proceed in the analysis, by applying the Heterogeneous Autoregressive (HAR) model and several machine learning methods, to forecast the volatility and volatility of volatility of five major cryptocurrencies. As an alternative modeling approach to the HAR model, we compute forecasts of both estimators proposed, by employing several non-linear machine learning techniques, as alternative to the OLS methodology, to accurately forecast in a one step ahead and multi-step ahead forecasting horizon (1 day, 5 days and 25 days respectively) cryptocurrency volatility and volatility of volatility. Thus, we compare and analyze the ability of the selected models for both short term and long-term forecast of the volatility and volatility of volatility of five leading cryptocurrencies.

**Keywords:** stochastic volatility, HAR, machine learning, cryptocurrencies, Bitcoin.

**Portuguese Prison Statistics from a Gender Perspective**

Karla Tayumi Ishy¹ and Ana Lorga da Silva²  
¹Collaborating Researcher at University of Coimbra Institute for Legal Research, Portugal  
²Lusófona University and Escola Superior Nautica Infante D. Henrique, Portugal

In the year 2021, it was observed that between 10.8 and 11.5 million people were deprived of their liberty in penal institutions around the world, more than half of which were in the USA (more than 2 million), China (1.69 million, plus people in pre-trial detention and other forms of detention not accounted for), Brazil (811 thousand), India (478 thousand), Russia (471 thousand), Thailand (309 thousand) and in Turkey (291 thousand). In Portugal, on December 31, 2021, 12.649 people were serving sentences and custodial measures in prison and non-prison institutions. The prison population in Portugal is mostly male (93%). On December 31, 2021, there were 10.774 men and 814 women deprived of their liberty in prison and non-prison facilities. The proportion of imprisoned women in Portugal (7%) is higher than the European median rate which, on January 31, 2021, was 4.7%. As for nationality, the prison population was composed of 80.0% Portuguese men, 5.6% Portuguese women, 12.9% foreign men, and 1.4% foreign women. When comparing the data for the female and male prison population, it is observed, for example, that the procedural situation and the nature of the crimes for which they are arrested vary considerably by gender and nationality.
Over the entire period from 2001 to 2021, the rates of people in preventive custody were lower among Portuguese citizens than among foreign citizens, and within these groups, the rates were lower among men than among women. On December 31, 2021, considering gender and nationality, 16.3% of Portuguese men, 20.2% of Portuguese women, 29.8% of foreign men and 35.2% of foreign women were in preventive custody. Concerning the nature of the crimes, while drug-related crimes were the main reason for women’s prison sentences (35.7% of women, 2021), crimes against the person were the main crimes committed by men sentenced to prison (31.8% of men, 2021). The total of crimes by type and by gender from 2015 to 2021 is also analyzed. These and other comparative analyses will be presented in order to verify if, in the Portuguese prison population, there are gender-related differences and to what extent they might be the result of gender specificities.  

**Keywords:** Criminology, Prison, Gender, Data Analysis, Comparative Study.

---

**Effects of Uncertain Input Data on Sample Size and Test Duration in Reliability Demonstration Tests**

**Leopold Tobias**  
*University of Applied Sciences Esslingen, Germany*

**Overview of Content.** During the development of products, the proof of product reliability is often a challenging activity. Basic planning activities have to be executed in an early design stage of the product to show the endurance of the specimen. A popular method is the so-called zero-failure testing, which is also known as reliability demonstration testing (rdt). The procedure allows a statement to be made about the minimum reliability of a product if all test specimens have survived the test procedure without any failure. The theory and its practical application seem to be quite simple at first glance. However, for most practical applications the assumption of distribution parameters of the failure behavior of the product has to be considered, especially when testing with a so-called lifetime ratio. A lifetime ratio describes the relationship between test duration and lifetime requirement. The influence of uncertain input parameters must be taken into account both in the planning of tests and in the evaluation of test results. The paper will show the correct application of the statistics of rdt and will especially focus on practical aspects under consideration of uncertainties of the assumed failure distribution of the product. This enables both test planning and the evaluation of zero-failure tests to be accompanied by statistically target-oriented considerations. Practical challenges are highlighted and approaches to solutions are outlined. Some central questions are for example:  
- Which number of specimen and testing duration is necessary?  
- What is the probability of a successful rdt?  
- Why does knowledge about the failure behavior of the product need to be available for a reliability demonstration test?  
- What are the consequences of uncertain input parameters of a failure behavior on the required sample size (see Fig. 1) and test duration?  
- Which uncertainties have to be regarded and how to consider in the evaluation phase.
Keywords: Reliability proof, Reliability Demonstration Testing, Zero-Failure-Testing, Binomial Testing.

Robust Pitman Type Estimators for Moment Condition Models

Aida Toma$^{1,2}$, Amor Keziou$^3$, Luiza Bădin$^{1,2}$, and Silvia Dedu$^1$

1Department of Applied Mathematics, Bucharest University of Economic Studies, Bucharest, Romania
2“Gh. Mihoc - C. Iacob” Institute of Mathematical Statistics and Applied Mathematics of the Romanian Academy, Bucharest, Romania
3Laboratoire de Mathematiques de Reims, Universite de Reims, Champagne Ardenne, UFR Sciences, Reims, France

We present robust minimum empirical divergence estimators for moment condition models, based on truncated orthogonality functions and dual forms of divergences. For moment condition models invariant with respect to additive or multiplicative transformations groups, these estimators are also equivariant. For models invariant with respect to additive groups, Pitman type estimators are proposed. Approximations of the Pitman estimator are given and it is shown that these approximations represent robust estimators for the model parameter. Some examples based on Monte Carlo simulations illustrate the performance of the estimation methods. (This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI – UEFISCDI, project number PN-III-P4-ID-PCE-2020-1112, within PNCDI III.).

Continuous Increasing Probability Density Functions: An Approach Through Sampling

Maria Tsifoutidou$^1$ and Nikolaos Farmakis$^1$

1Master Educational Programme Department of Mathematics, Aristotle University of Thessaloniki, Greece

A critical point in Statistical Analysis is the determination of the Probability Density Function (p.d.f) of the random variables entered by the problem under study. Many times the effort for the determination is extremely time-consuming. For a time saving process we need the help of sampling techniques and the use of the concept of Coefficient of Variation (Cv) for a random variable X. Sampling can contribute to the
study of the distribution of a random variable with very good results in terms of accuracy and more especially of speed. So the data to be used in this paper is sampling data. From early the 2000 A.D. the Coefficient of Variation (Cv) has been used for the above determination. It is already a tool that reduces the time to determine the p.d.f. by far, thousands of times often. There is a lot of publications from 2003 and so on, basically for continuous random variables. In this work we study the case where the graph of an increasing p.d.f. passes through the origin of the axes of X and Y, i.e. it is \( f(0)= 0 \). Here X is for the random variable values and Y = f(x) for the frequencies of those values. We deal with continuous random variables in this paper. Some illustrating examples are given.

References:
Farmakis N. (2016), «Introduction to Sampling», Edit Kyriakidis Bros S.A. Thessaloniki. (in Greek)

Keywords: Sampling, Probability density function, Mean, Variance, Standard deviation, Coefficient of variation.

Anonymization with no Gender Information Loss: The Demographic Profile of Diligent Greek Students Belonging to Vulnerable Social Groups of the 2019-20 Academic Year

Eva Tsouparopoulou¹ and Maria Symeonaki²

¹Department of Primary Education, University of the Aegean, Rhodes, Greece, and Department of Social Policy, Panteion University of Social and Political Sciences, Athens, Greece
²Department of Social Policy, Panteion University of Social and Political Sciences, Athens, Greece

Efficient data anonymization has become a matter of great concern nowadays since sharing data and microdata without revealing sensitive information is imperative. National public organizations collect data usually containing personally identifiable information (PII). Although they are treasuries of rare data accessible under special permissions for their strictly research use, applying data anonymization techniques to such data frequently leads to information loss. The motivation behind this study is to present two different algorithmic techniques inspired by optimizing language models to obtain gender information before applying data anonymization. The analysis was based on raw data of the diligent Greek students belonging to vulnerable social groups of the 2019-20 academic year who applied for a financial support scholarship program offered by the State Scholarships Foundation (IKY) and co-funded under the European Social Fund. The anonymized data derived from the two different applied techniques were analyzed, and the demographic profiles of the applicants and the beneficiaries of
the diligent Greek students belonging to vulnerable social groups of the 2019-20 academic year program are to be presented. This methodological study aims to contribute to the continually growing field of data anonymization techniques applied to large microdata datasets with less information loss in harmony with the PII data protection principles.

**Keywords:** data anonymization, information loss, large microdata datasets.

---

**Bayesian Birth-Death Skyline Model - A Case Study on Heterochronous Maltese SARS-CoV-2 Genomic Data**

Gianluca Ursino¹, Monique Borg Inguanez¹, David Suda¹, Joseph Borg², and Graziella Zahra³

¹Department of Statistics and Operations Research University of Malta, Malta
²Department of Applied Biomedical Science, Faculty of Health Sciences, University of Malta, Malta
³Molecular Diagnostics – Infectious Diseases Mater Dei Hospital, Malta

When studying viral genome sequence data the Bayesian framework has the advantage that it can simultaneously construct phylogenetic trees, which allow us to analyse the relationships between different genomes (phylogenetic analysis), and at the same time infer viral dynamic across time (phylodynamic analysis). This requires the specification of three models: (i) the transmission model (ii) the substitution model and (iii) the molecular clock model used to infer the root of a phylogenetic tree. In this study as transmission model we consider the Bayesian birth-death skyline (BDSKY) model which can be applied for both homochronous and heterochronous genome datasets under specific formulations of it. When it comes to defining a substitution model we resort to the bModelTest method. As a case study we consider 681 heterochronous genome sequences of COVID-19 sampled by the Molecular Diagnostics-Infectious Diseases at the Mater Dei Hospital in Malta between 19/8/2020 and 5/1/2022. For this purpose, we consider both serial BDSKY and the multi-rho BDSKY models, with the former being more suitable for genome sequences individually sampled at several time points, and the latter being more suitable for genome sequences sampled in batches at multiple time points. For each transmission model we consider four different molecular clock models which are a combination of the strict and relaxed molecular clock setups, and two settings for the number of intervals over which the reproductive number is considered constant (m=15 and m=30). The relaxed molecular clock option with m=15 intervals resulted to be the preferred molecular clock model of the data being studied. In general the serial BDSKY and the multi-rho BDSKY transmission models gave considerably similar results yet some discrepancies were observed and these will be discussed.

**Keywords:** Bayesian Birth-Death Skyline Models, Phylogenetic tree, Genomics, SARS-CoV-2.
Essay of a Multinomial Logistic Model for the Thematic Classification of a Dental Journal Based on the Degree of Adherence to Quality Guidelines

Pilar Valderrama\textsuperscript{1} and Mariano J. Valderrama\textsuperscript{2}
\textsuperscript{1}Department of Information and Communication. University of Granada, Spain
\textsuperscript{2}Department of Statistics and Operations Research. University of Granada, Spain

Publication standards and quality guidelines in Health Science journals are very helpful tools for improving the scientific level of articles published. Nevertheless, among the broad array of journals published nowadays, guidelines are not heeded to an equal extent. Furthermore, Open Access publication has grown in recent years, and in fact many journals publish only under this modality. In this context, the journals of a given field can be grouped into categories according to their specialized field. The aim of this paper was to analyse, by means of a multinomial logistic model, how the classification of a journal in a certain category may be influenced by adhesion to the norms for publication upheld by the ICMJE, the CONSORT statement and the Clinical Trials Register. Besides, the publication under Open Access system was considered as an additional factor. This methodology was applied in the field of Dentistry in view of the listing of Journal Citation Reports corresponding to 2021. As a conclusion, the most influential variables were: adherence to the CONSORT guideline and to the Clinical Trials Register.

**Keywords:** ICMJE, CONSORT, Clinical Trials Register, open access, Journal Citation Indicator, multinomial logistic model.

Comparative Performance Analysis of YOLOv4 and YOLOv5 Algorithms on Dangers Objects

Ayşe Berika Varol Malkoçoğlu\textsuperscript{1} and Ruya Samli\textsuperscript{2}
\textsuperscript{1}Department of Computer Programming, Beykoz University, Istanbul, Turkey
\textsuperscript{2}Department of Computer Engineering, Istanbul University–Cerrahpasa, Istanbul, Turkey

It is very important that the legal order is carried out without interruption and that people can live without fear. This can only occur in a high security environment. For this reason, today's security mechanisms are constantly evolving in order for life in society to be able to live safely. In particular, many locations have surveillance systems with cameras in order to ensure environmental safety in recent years. Images obtained from these cameras can be analyzed using computer vision techniques to create auxiliary systems. One of the sub-branches of artificial intelligence, the visual technique is inspired by the ability of people to understand, identify and detect objects, and is the solution to many problems today, thanks to the development of hardware components and data storage systems. The main purpose of the computer vision technique, which provides solutions to problems such as pedestrian detection, autonomous driving systems etc., is to enable visual data to be understood, identified and classified with artificial intelligence techniques. Many deep learning-based algorithms are used to perform these operations. The operating principles and performance of these different algorithms vary. Especially among the latest versions of YOLO, one of the popular object detection algorithms performance comparison varies from various sources. Based on this situation, within the scope of the study, performance comparison was performed between YOLOv4 and YOLOv5 models. To the models are tasked with identifying the weapon and the person carrying it that could threaten security when
used by unauthorized persons. A data set of weapons and people was used to make this test possible. Since the quality and labeling of the dataset to be used is known to be very important to the success of the model, the data is collected by us and tagged one by one. The obtained data were used in YOLOv4 and YOLOv5 models with the same hyper-parameters, and the results produced by both models were compared and analyzed. According to the results obtained, we confirm that these algorithms can be used to detect dangerous objects, and we report the performance values generated for this application with details. **Keywords:** Image processing, Gun Detection, YOLOv4, YOLOv5.

**Board Ethnic Diversity and ESG Performance: Empirical Evidence from US Banks**

Evangelos G. Varouchas¹, Stavros E. Arvanitis¹, George M. Agiomirgianakis¹,²,³ and Christos Floros¹

¹Department of Accounting and Finance, Hellenic Mediterranean University, Heraklion, Greece
²School of Social Sciences, Hellenic Open University, Patras, Greece
³Hellenic Scientific Institute on the Economics of Education & Life-Long Learning, Research and Innovation, Athens, Greece

In recent years, the issue of ethnic equality on corporate boards has attracted the interest of scholars, regulators, and lawmakers. The purpose of this research is to investigate the impact of board ethnic diversity on the ESG performance of 304 US publicly traded banks over the period from 2010-2021. For this purpose, we use the two-step system GMM estimator which deals efficiently with endogeneity issues that have plagued many previous studies. Our main empirical finding supports that board ethnic diversity positively influences ESG performance. In addition, we find that there is an inverted U-shaped relation between the proportion of foreign board members and ESG performance. To the best of our knowledge, this is the first empirical study that examines the impact of ethnic board heterogeneity on the ESG performance of US banks. Therefore, our empirical results are an important source of knowledge for regulatory bodies and policymakers and enrich the limited governance literature concerning the influence of board attributes on ESG performance. **Keywords:** Corporate governance, Board ethnic diversity, ESG performance, Banks, US.

**The Importance of the Initial Selection of Suppliers in the Food Service Divisions of Hotels, the Current Situation in the Supply Chain of Greece**

Konstantinos Vasilakakis, Efthalia Tabouratzi¹, and Despoina Sdrali

¹Assist.Professor, Department of Accounting and Finance, Laboratory of Accounting and Financial Management (LAFIM), Heraklion, Crete, Greece

Strategic supply chain management has become more important with the impact of globalisation. The main idea of supply chain management is to establish good relationships between the members of the chain to serve customers accurately. The success of the supply chain depends largely on strong cooperation among the chain members, where the selection of supplier is a strategic decision-making process that
affects all supply chain performance. In this paper, the current situation of the Greek Hotel Industry in terms of supply chain is analyzed mainly from the perspective of initial supplier selection as some stages and their factor analysis, correlation of factors and influence of environmental variables on supply chain management are analyzed with degree of importance. The sample concerns 653 Greek hotel companies and was conducted in the first quarter of 2019. The findings of the study highlight the importance of the function in the initial selection of suppliers, the problems and the process of initial selection of suppliers.

**Keywords:** Suppliers, hotel industry, food service divisions, supply chain management, Greece.

**Probabilistic Characteristics of a Two-channel Retrial Queue Operating in a Semi-Markov Environment**

Viacheslav Vavilov

*Department of Software Engineering, National Research Tomsk State University, Tomsk, Russian Federation*

The paper proposes the mathematical model of a two-channel retrial queue with operating in a random semi-Markov-type environment. We consider incoming fresh calls arrive at the system according to a Poisson process. The device of this system has two service channels and can be in one of four states:
1) both channels are idle;
2) the first channel is idle and the second channel is busy;
3) the first channel is busy and the second channel is idle;
4) both channels are busy.

Upon arrival, an incoming call either with $r$-probability trying to occupy the first channel and with $(1-r)$-probability trying to occupy the second channel. The call begins to be serviced immediately if a channel is idle or joins an orbit if a channel is busy. From the orbit call retries to occupy the device and behaves the same as a fresh incoming call. After the service, the call leaves the system. The time intervals of calls in orbit and the time intervals of call service on the device lines are distributed exponentially with different parameters.

Random external factors affect the call service times and determine the values of the considered parameters of the service time distribution. The model of a random environment is a semi-Markov process. The mathematical model of the system under consideration is investigated by the method of asymptotic analysis under conditions of increasing duration of calls delays in orbit. As a result of applying this method, we have obtained the average value of the number of calls in the orbit, the deviation from the average value, the probability distribution of the device states and the probability distribution of a number of calls in the system.

**Keywords:** retrial queue, random environment, semi-Markov process, incoming and outgoing calls, method of asymptotic analysis, orbit.
Primary Health Care during the Period of Covid - 19 pandemic.  
The Effect of the Pandemic on the Cost and the Quality of Services rendered at Health Care Centers of Fthiotida Prefecture

Ilias Velaoras¹ and George Matalliotakis²
¹,²Hellenic Open University

The Covid-19 pandemic caused a series of changes in the function of medical services. A literature review was conducted aiming at describing the facts of the covid-19 pandemic and highlighting the role of First-degree Health Care (F.H.C.) especially during the pandemic. Additionally, the present paper reports whether the pandemic affected the cost and the quality of the services provided, and foregrounds issues of social justice and equality in the accessibility of health services in F.H.C. Firstly, an analysis of the terminology of health, the pandemic and F.H.C. is provided and there is a detailed description of the function of F.H.C. in Greece. Following that, I present the development and the contemporary evidence related to the Covid-19 pandemic in Greece and in other European countries. The purpose of that is to propose practices for the improvement of the health services in this level as well as limiting inequalities in its accessibility. Finally, I present the effects that the role of the Covid-19 pandemic had on the cost and quality of First-degree Health Care in County Fthiotis.

Maintainability and Attainability for Discrete-time Homogeneous Semi-Markov models

Brecht Verbeken¹ and Marie-Anne Guerry²
¹,²Department of Business Technology and Operations, Vrije Universiteit Brussel (VUB), Brussels, Belgium

In previous research the importance of both Markov and semi-Markov models in manpower planning is highlighted. Maintainability and attainability of personnel structures for different types of personnel strategies (i.e. under control by promotion and control by recruitment) is extensively investigated for various types of Markov models (homogeneous as well as non-homogeneous) ([1], [2]). Semi-Markov models are extensions of Markov models that account for length of stay in the states. Less attention is paid to the study of maintainability and attainability for semi-Markov models. Although, some interesting maintainability results were obtained for non-homogeneous semi-Markov models ([3]). The current paper focuses on discrete-time homogeneous semi-Markov models, and explores concepts of maintainable and attainable personnel structures in this setting. Various types of personnel strategies are presented for which the set of maintainable and attainable structures is examined. The obtained insights can easily be converted to various application domains.

References:
The Distribution of the Amount of Rain. A Renewal Stochastic Process Model

Javier Villarroel\textsuperscript{1} and Juan A. Vega\textsuperscript{2}

\textsuperscript{1}Department of Statistics and Inst. Univ. Fis. Matem, Universidad de Salamanca, Salamanca, Spain
\textsuperscript{2}Inst. Univ. Fis. Matem. Universidad de Salamanca, Salamanca, Spain

With this work we intend to study the distribution of the amount of rain and other related statistics in the environment of Lake of Sanabria (Spain). To do this we will develop a model using a compound renewal process in which we will not assume that the initial time is a renewal time. We will deduce the resulting integral equations and solve them for various distributions of both, waiting times and jumps. We will then apply them for the study of the distribution of the amount of rain and the remaining statistics.

Keywords: Compound renewal process, non-renewal time, amount of rain, Lake of Sanabria.

Data Analysis of Discrete-Valued Time Series Based on Exponential Family

Valeriy Voloshko\textsuperscript{1} and Yuriy Kharin\textsuperscript{2}

\textsuperscript{1,2}Belarusian State University, Research Institute for Applied Problems of Mathematics and Informatics, Minsk, Belarus

We consider a new parsimonious model DCNAR(s,d) for discrete-valued stationary s-order Markov chain whose conditional distributions subject to fixed s-prehistory lie in some d-parametric exponential family. Canonic d-dimensional parameter of conditional distribution depends on s-prehistory in a special way similar to GLM models: linear combination of m fixed base functions from s-prehistory is transformed by some fixed nonlinear link function. Vector of m coefficients of linear combination is a model parameter of DCNAR(s,d). We construct a family of statistical estimators of explicit form for parameter of model DCNAR(s,d), prove their consistency and asymptotic normality, find the asymptotically efficient estimator within the constructed family. Model DCNAR(s,d) is universal because it does not depend on structure of states space of time series and therefore it may be applied to vector time series or spatio-temporal data.

Keywords: High-Order Markov Chain, Parsimonious Model, Statistical Estimator, Exponential Family, Asymptotic Properties.

An Exponentiality Test against Heavy & Light Tailed Alternatives

Filia Vonta

Department of Mathematics, National Technical University of Athens, Athens, Greece

This work is filling up the gap in the literature regarding the verification of the log-concavity property which is a widely studied topic due to the fact that it provides desirable estimating properties. At the same time, log-concavity together with log-
convexity are vital in reliability, engineering and stochastic modeling for distinguishing between an exponential, a light-tailed and a heavy-tailed distribution. In this work we propose an exponentiality test of fit to be used for distinguishing between exponential and log-concave or long-convex distributions. The proposed test statistic is based on the conspiracy and catastrophe principles through which a characterization for the (tail part of the) exponential distribution is established. The details of the formulation of the test are provided, an extended simulation study showcasing the performance of the proposed test statistic is given, and some concluding remarks are given.

Note: this is a joint work with Alex Karagrigoriou, Ioannis Mavrogiannis, Georgia Papasotiriou

**Which Socio-demographic Indicators Influence the Health Rating? Example of EHIS on Czech data**

Jana Vrabcová¹, Markéta Majerová², Tomáš Fiala³, and Jitka Langhamrová⁴

*Department of Demography, Prague University of Economics and Business, Prague, Czech Republic*

In most developed countries of the world, population live at an older age. With increasing age, there is a greater probability of a certain health restriction, and the proportion and number of people with health restrictions, and disabilities increases. At the level of society, disability increases the demand for formal and informal care. The health prevalence model is designed on individual anonymized data from the European health interview survey (EHIS) in 2008, 2014 and 2019 and enters selected socio-demographic indicators. The paper is examined and described with mutual relations among the indicators of the Minimum European health module (MEHM): self-perceived health, chronic/long-term morbidity, and activity limitations (Global Activity Limitation Indicator). Within binary logistic regression, the influence of sex, age groups, marital status, education, presence of long-term illness and long-term limitation of activities on self-rated health has been examined. A statistically significant influence on self-rated health has been demonstrated in age, education, long-term illnesses, and long-term limitation of activities. As part of the sensitivity analysis, the assignment of the middle category (fair) of self-rated health is chosen for poor assessment of your health (bad, very bad). New dichotomic variable called “healthy vs ill” was designed, which can be further used to calculate the new population health indicator and compare it with life expectancy indicators based on MEHM. This new variable is unique it combines answers to subjective and objective health issues. With the help of the Brass relational method, prevalence curves are modelled for the newly designed dichotomic variable and examined the potential trend of development over time, but it has not been clearly proven. The standard in the relational method is chosen a combination of three EHIS surveys in 2008, 2014 and 2019.

**Keywords:** Self-rated health, Chronic disease, Global Activity Limitation Indicator, binary logistic regression, Brass relational method.
A Generalization of the k-means Method for Trends of Time Series

Norio Watanabe
Chuo University, Japan

The clustering is one of the important methods in multivariate analysis. The clustering of time series is also important and several clustering methods are available. Recently, a k-means type method was proposed for trends of time series. In this method each object for clustering consists of univariate time series. In this study, a generalization of this method is proposed for the case where each object consists of multivariate time series. The applicability of the proposed method is examined by simulation studies and an application using real data.

Compliance with IUU Fisheries of Manila Clam in Tagus Estuary

Margarida Xavier¹, Paula Chainho², and Ana Lorga da Silva³

¹Auchan Retail Portugal, Estrada de Paço de Arcos, Portugal
²MARE, Lisbon, Portugal; Department of Animal Biology, Faculty of Sciences, University of Lisbon, Lisbon, Portugal & Instituto Politecnico de Setubal, Portugal
³Departamento de Gestão Portuária, ENIDH, Paço de Arcos, Portugal & ECEO, Lusófona University, Campo Grande, Lisboa, Portugal

The Tagus estuary has been facing IUU fishing activities due to the illegal catching of Manila clam, which ultimately undermines the sustainable management of this resource. There is consensus on the importance of understanding the motives and behavioral factors that lead fishermen to adopt IUU fishing activities, in order to provide policy information on what measures should be applied to encourage compliance and combat violations. The objectives of this study were to characterize and quantify IUU fishing activities in the Tagus estuary, to analyze the factors influencing compliance with the daily kg quotas for catching Manila clams and to analyze the factors influencing compliance with the fishing equipment gears allowed for catching clams. The static ordered probit model was used to understand the factors influencing the daily kg quota. The binary probit model was used to understand the factors influencing the probability of violating the fishing equipment/crew gear allowed for the catching of Manila clams. This study was applied in the Tagus estuary in 2015 and the results suggest that the most significant IUU activities are fishing without a license and not declaring catches. It was found that 91.5% of the catchers do not have a license. Relative to the determinants influencing the probability of violating daily kg quotas, the results suggest that having fewer years of experience, practicing the activity accompanied and considering the regulations respected are incentives to cooperate with the regulations. Regarding the determinants influencing the probability of choosing a legal or illegal fishing gear it was found that individuals whose main activity is fishing in the Tagus estuary are more likely to choose an illegal fishing gear. The provision of alternative income or employment opportunities outside fishing would be important to encourage the use of legal fishing gears.

Keywords: Manila clam, Tagus estuary, IUU fishing, Compliance, Probit Model.
Limit Theorems for Continuous-Time Branching Random Walks

Elena Yarovaya
Department of Probability Theory, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Moscow, Russia

In modeling the evolution of a particle population, the main feature of the population is the ability of its particles to die, produce copies, or migrate. Such processes are usually referred to as branching random walks and are described in terms of birth, death, and transport of particles. Nowadays, it is generally accepted that branching random walks are of crucial use in the study of complex stochastic systems involving birth, death, and migration of their elements. Examples include various natural processes and phenomena, such as population dynamics (Molchanov et al.), problems in the theory of homopolymers (Cranston et al.), etc. In some of these branching random walks, the phase space in which the particles 'live' can be naturally assumed to be continuous; in other cases, this phase space may be a finite or infinite set of discrete points. In this talk, the main focus will be on the properties of branching random walks on multidimensional integer lattices. We will be mainly interested in the problems related to the existence of phase transitions when various parameters are changed, the properties of the limiting distribution of the particle population, etc. Of course, the answer to these and other questions depends on numerous factors affecting the properties of a branching random walk. We shall therefore attempt to describe in some fair detail how the properties of a branching walk depend on the fact of randomness (or determinacy) of the branching medium and its homogeneity, on the number and mutual arrangement of branching sources and also on such properties of a branching walk as its symmetry and the finiteness or infiniteness of the variance of the jumps. Some evolutionary models of multi-type particle population will be discussed. We also present some results of simulation of branching random walks and discuss how they can be applied to numerical estimation on bounded time intervals of various characteristics describing the properties of phase transitions.

Keywords: Branching Random Walks, Limit Theorems, Simulation.

Modeling Religion Variables on Self-perceived Social Class: Evidence from the 7th Wave of the World Values Survey, 2017-2020

Aggeliki Yfanti1 and Catherine Michalopoulou2
1STADA Lab, Department of Social Policy, Panteion University of Social and Political Sciences, Athens, Greece
2Department of Social Policy, Panteion University of Social and Political Sciences, Athens, Greece

In the literature, although social class is considered important in all aspects of life, its relation to religion has not been adequately researched. In this respect, the purpose of this study is to explore how religious values, beliefs and practices are impacting social class perceptions as measured in the World Values Survey (WVS) questionnaire that classifies respondents into five class categories: upper, upper middle, lower middle, working and lower class. The analysis is based on the 2017-2020 7th wave WVS datasets for six countries: Brazil, Canada, Germany, Greece, Turkey and the USA. First, at country level, each of the five social class categories was modeled separately using linear probability models. The variables included in these analyses decided the
common model that was used so as to allow for comparisons cross nationally and among social class categories. The main findings of the regression analyses showed that respondents in all six countries differentiated in their social class perceptions according to their beliefs on the outcome of the science-religion conflict, the acceptance of one’s religion, church attendance, praying frequency, identifying oneself as religious and the basic meaning of religion. Certainly, within each country, social class perceptions differed. This study shows that, although religious values, beliefs and practices have an impact on social class identification, in future research additional social indicators should be considered so as to provide a more coherent interpretation. **Keywords:** subjective social class, religious values, religious beliefs, religious practices, World Values Survey (WVS).

The Expectation of a Mixed Moving Average Process Subject to Ambiguous Lévy Basis

Hidekazu Yoshioka¹ and Yumi Yoshioka²

¹,²Faculty of Life and Environmental Sciences, Shimane University, Matsue, Japan

A mixed moving average process (MMA process) is a superposition of infinitely many affine stochastic differential equations. It is a tractable mathematical model of jump-driven processes having a subexponential memory. The probability measure for the superposition is often ambiguous due to the data limitation in applications, motivating us to operate the model accounting for the ambiguity. We present a suitable optimization framework for bounding nominal statistics based on the Tsallis divergence, which is a proper risk measure for dealing with the subexponential memory structure. An application example of the proposed framework to the streamflow time series data that have qualitatively different memory structures.

**Keywords:** Superposition of Affine Processes, Markovian Lifts, Streamflow.

On the Use of Copula for Quality Control based on an AR(2) Model

Timothy M. Young¹ and Ampalavanar Nanthakumar²

¹School of Natural Resources, University of Tennessee, Knoxville, TN, U.S.
²Department of Mathematics, State University of New York at Oswego, Oswego, NY, U.S.

This study addresses the significant problem of positive autocorrelation in data collected from online sensors, which may impair assessment of statistical control. Sensor data collected at short time intervals typically have significant autocorrelation, and traditional statistical process control (SPC) techniques cannot be deployed. There is a plethora of literature on techniques for SPC in the presence of positive autocorrelation. This paper contributes to this area of study by investigating the performance of ‘Copula’ based control charts by assessing the average run length (ARL) when the subsequent observations are correlated and follow the AR(1) or AR(2) model. The conditional distribution of $y_t$ given $y_{t-1}$ is used in deriving the control chart limits for three different categories of Copulas: Gaussian, Clayton, and Farlie-Gumbel-Morgenstern Copulas. Results suggest that the overall performance of the Clayton Copula and Farlie-Gumbel-Morgenstern Copula is better compared to other Archimedean Copulas for the AR(1) situation. The Clayton Copula is the more robust with respect to changes in the process standard deviation as the correlation coefficient
increases. The Gaussian Copula seems appropriate if the Copulas are to be considered in Quality Control, especially in an AR(2) situation.

**Keywords:** Copula, Model, Autocorrelation, Statistical Process Control.

### Nonlinear Dynamics of US Inflation, Money Supply, and Growth Time Series

Tai Young-Taft

1 Levy Economics Institute of Bard College, Blithewood, Bard College, Annandale-on-Hudson, New York, 12504-5000, USA

The market at large may be conceived of as a network, and the global shock of the pandemic as a significant and multi-level pruning of the market (a sledge hammer to a mycelium network, if you will).

In this paper we consider nonlinear time series analysis with respect to inflation, money supply, and growth in the US, as individual series and in tandem. In particular we consider the maximum characteristic Lyapunov exponent, embedding dimension with principal component analysis, embedding time delay using average mutual information and the minimum sufficient embedding dimension using the false nearest neighbour method, effects of the Theiler window using space-time separation plots, thresholded and unthresholded recurrence plots, and recurrence quantification analysis.

**Keywords:** Economics, Inflation, Nonlinear Dynamics, Time Series.

### A Note on the Convergence of Euler Contributions, Depending on the Underlying Risk Measure

Georgios C. Zachos

Dept. of Statistics and Actuarial-Financial Mathematics, University of the Aegean, Karlovassi, Samos, Greece

This paper is devoted to the Convergence of the so called Euler Risk Contributions when the underlying Risk Measures differ. To that end, our discussion is in regard of Euler contributions in a Risk Measure environment. In addition, we proceed by defining some conditions where the rate of convergence of Euler Risk Contributions in a Value at Risk regulation environment and Distortion Risk Measure regulation environment coincide. Finally, we generalize our findings in regard of the Expected Shortfall case.

**Keywords:** Systemic Risk, Euler Contributions, Distortion Risk Measures, Rate of Convergence.
Comparing Healthy Life Expectancy in Uganda and Kenya

Konstantinos N. Zafeiris¹ and Christos H. Skiadas²
¹Laboratory of Physical Anthropology, Department of History and Ethnology, Democritus University of Thrace, Komotini, Greece
²ManLab, Technical University of Crete, Chania, Greece

Uganda and Kenya, two well-known African Countries, two ex-colonies located in the central-eastern part of Africa, faced several political, socioeconomic and health crises in their recent history. Both constitute a unique paradigm of Africa’s historical momentum; thus, this paper aims to comparatively analyse the temporal trends of average longevity and healthy life expectancy from a comparative perspective. The results of the analysis indicate the several periods of crisis and development during the 21st century, pointing to the urged need for national and international intervention programmes to benefit their people’s health.

Keywords: Life expectancy at birth, healthy life expectancy, Africa, Uganda, Kenya.

Mortality and Health Developments in Bosnia-Herzegovina and Montenegro

Konstantinos N. Zafeiris¹, Byron Kotzamanis², and Christos H. Skiadas³
¹Laboratory of Physical Anthropology, Department of History and Ethnology, Democritus University of Thrace, Komotini, Greece
²Laboratory of Social and Demographic Analyses. University of Thessaly, Bolos, Greece
³ManLab, Technical University of Crete, Chania, Greece

The scope of this paper is to analyse the mortality transition in two newly established countries that emerged after Yugoslavia’s dissolution, i.e. Bosnia & Herzegovina and Montenegro. Data come from the World Population Prospects and are used to allow the calculation of several mortality indicators. Therefore measures of the mean duration of life, probabilities of death, and the Arriaga method describe major aspects of the mortality transition that occurred in the two countries. Additionally, measures relating the old age heap and the rectangularisation process occurring on the survival curves give additional information on the mortality regimes in the two genders and countries and their differentiation over time. Results indicate that significant deviations, convergences and divergences occurred between the two countries per gender. Mortality transition is evident in both of them, but it takes place with different rhythms, timetables and general characteristics, described in detail in this paper.

Keywords: Bosnia-Herzegovina, Montenegro, Life expectancy at birth, Arriaga method, Old age heap, Gini coefficient, Inflexion points of the death density curves, Modal age at death, Rectangularisation of survival curves, Euclidean distance.

Improved Applications of Arak's Inequalities to the Littlewood-Offord Problem

Andrei Zaitsev

We discuss the results of joint papers of the speaker and Friedrich Götze (Bielefeld University). In a new paper, a supplement to the authors' paper [1] is given. In the Littlewood-Offord problem we study the behavior of concentration functions of weighted sums of independent identically distributed random variables. We use a new general inequality [2] showing that the Littlewood–Offord problem can be reduced to
estimation of the concentration functions of some symmetric infinitely divisible distributions whose spectral measures are concentrated on the set of plus-minus weights. We show that the dependence of constants on the underlying distributions in [1] may be clarified.

References:

Posters

An Efficient Distributed Fusion Smoothing Algorithm for Tk-proper Hypercomplex Multisensor Systems with Uncertain Observations

Rosa M. Fernández-Alcalá, José D. Jiménez-López, Jesús Navarro-Moreno, and Juan C. Ruiz Molina

Department of Statistics and Operations Research, University of Jaén, Jaén, Spain

The distributed fusion smoothing problem for Multisensor systems with uncertain observations is addressed by using hypercomplex techniques. Specifically, the uncertainty in the observations is assumed to be modelled by a sequence of independent Bernoulli random variables. Then, taking advantage of the recent advances in tessarine algebra, the problem is solved in this domain where, under Tk-properness conditions, the dimension of the problem is considerably reduced and, consequently, a significant computational saving is implied. In this framework, a distributed fusion methodology is used to provide an efficient smoothing estimation algorithm with a lower computational cost than its real-field counterpart. The effectiveness of the proposed solution is experimentally validated by means of a simulation example.

Keywords: Multisensor distributed fusion estimation, Smoothing algorithm, Tessarine algebra, Tk-properness, Widely linear processing.
Special Session

Gender Inequalities in Educational Contexts
Organized by Adele H. Marshall and Mariangela Zenga

First talk
Authors: Aglaia Kalamatianou and Fotios S. Milienos
Title: Modelling the share of women in university education
Abstract. Historically, women's education began to grow in the second half of the 19th century, mainly in Western Europe and the US, while previously it was virtually non-existent. Nowadays the bare statistics of international organizations' publications are impressive in themselves. In many countries, gender parity has been approached, regarding attainment and graduation rates, or even more, the gender hierarchy has been inverted in favor of women. The interest in this paper is on modeling the evolution of women's share in the bodies of new entrants/enrolments and of graduates in universities. Motivated by the experience of a long series of data, we focus on specific classes of stochastic models, such as growth models. Some theoretical aspects of these classes are discussed before a numerical study of the proposed statistical methodology and the real-data analysis.
Keywords: University Education, Women's Share in Universities, Growth Models, Time Series.

Second talk
Authors: Michele La Rocca, Marcella Niglio, and Marialuisa Restaino
Title: An analysis of gender gap in students' performance: preconception or reality?
Abstract: In recent years, exploring the determinants that may influence students' achievement has received much attention. Empirical studies have found that the most important factors which affect student performance are students' characteristics, family background, school attended, and regional residence. This paper aims at investigating the differences in students' performances among Science, Technology, Engineering and Mathematics (STEM) courses by using regression models, in order to capture the gender gap. To measure performance (university success), we focus on the number of ECTS credits earned during the first year, since it represents an important moment in the students' path at university. The analysis concerns students enrolled at 3-year STEM degrees in a university located in the South of Italy during the last 5 years, with a focus on the number of university credits earned during the first year (a good predictor of the regularity of the career). In particular, the main purposes are i) to estimate the probability of getting at least a certain number of credits at the end of the first year and identify the factors which might affect it, and ii) to capture if there are any differences in performance between men and women, and which factors might affect these differences.
Keywords: gender, university performance, regression analysis, STEM.

Third talk
Authors: Adele H. Marshall and Mariangela Zenga
Title Gender and ICT inequalities among adolescents: a comparative analysis in PISA
Abstract: ICT literacy has become an important key driver regarding the competence for the adults, but even more for young people. As it was pointed out in 2003, ICT literacy is defined as “the interest, attitude, and ability of individuals to appropriately use digital technology and communication tools to access, manage, integrate, and evaluate information, construct new knowledge, and communicate with others in order to participate effectively in society” (Lennon et al., 2003). Using data from Programme for International Student assessment (PISA), this work will show the gender differences in ICT literacy and discover differences in OECD countries.
Keywords: gender, ICT, multilevel analysis, PISA.