PROGRAMME AND ABSTRACTS

11th International Conference on

Computational and Financial Econometrics (CFE 2017)

http://www.cfenetwork.org/CFE2017

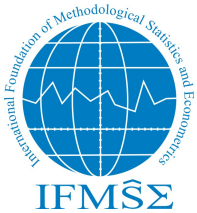
and

10th International Conference of the
ERCIM (European Research Consortium for Informatics and Mathematics) Working Group on

Computational and Methodological Statistics (CMStatistics 2017)

http://www.cmstatistics.org/CMStatistics2017

Senate House & Birkbeck University of London, UK
16 – 18 December 2017



ISBN 978-9963-2227-4-2

Technical Editors: Gil Gonzalez-Rodriguez and Marc Hofmann.

International Organizing Committee:
Ana Colubi, Erricos Kontoghiorghes, Marc Levene, Bernard Rachet, Herman Van Dijk.

CFE 2017 Co-chairs:
Veronika Czellar, Hashem Pesaran, Mike Pitt and Stefan Sperlich.

CFE 2017 Programme Committee:

Knut Are Aastveit, Alessandra Amendola, Josu Arteche, Monica Billio, Roberto Casarin, Gianluca Cubadda, Manfred Deistler, Jean-Marie Dufour, Ekkehard Ernst, Jean-David Fermanian, Catherine Forbes, Philip Hans Franses, Marc Hallin, Alain Hecq, David Hendry, Benjamin Holcblat, Jan Jacobs, Degui Li, Alessandra Luati, Richard Luger, J Isaac Miller, Claudio Morana, Bent Nielsen, Juan-Pablo Ortega, Michael Owyang, Gareth Peters, Cesare Robotti, Paulo Rodrigues, Jeroen Rombouts, Olivier Scaillet, Peter Siklos, Robert Taylor, Carlos Velasco, Martin Wagner and Jean-Michel Zakoian.

CMStatistics 2017 Co-chairs:
Bertrand Clarke, Steven Gilmour, Michele Guindani and Leopold Simar.

CMStatistics 2017 Programme Committee:

Peter Buehlman, Guang Cheng, Jeng-Min Chiou, Stephane Chretien, Harry Crane, Christophe Croux, Michael Daniels, Xavier De Luna, Herold Dehling, Fabrizio Durante, John Einmahl, Piotr Fryzlewicz, Yulia Gel, Armelle Guillou, Christopher Hans, Siegfried Hormann, Stephan Huckemann, Marie Huskova, Timothy Johnson, Ivan Kojadinovic, Tatyana Krivobokova, Clifford Lam, Carlos Lamarche, Christophe Ley, Marloes Maathuis, Paul McNicholas, Thomas Valentin Mikosch, Domingo Morales, Davy Paindaveine, Marco Riani, Juan Romo, Holger Rootzen, Jason Roy, Armin Schwartzman, Ana-Maria Staicu, Francisco Stingo, Stefan Van Aelst, Ingrid Van Keilegom, Tim Verdonck, Mattias Villani, Andrew T. Walden, Jane-Ling Wang, Nanny Wermuth and David van Dyk.

Local Organizer:
Birkbeck University of London.
CFEnetwork and CMStatistics.

E1073: **Dissimilarity functions for copula-based hierarchical clustering of continuous variables**
Presenter: **Sebastian Fuchs**, Free University of Bozen-Bolzano, Italy
Co-authors: F Marta L Di Lascio, Fabrizio Durante
A copula–based notion of dissimilarity between continuous random variables is introduced and formalized. Such a concept aims at detecting rank–invariant dependence properties among random variables and, as such, it will be defined as a functional on the collection of all copulas. We show how the provided definition includes previous dissimilarity measures considered in the literature like those derived from measures of association and tail dependence but also those of agglomerative hierarchical type. In the latter case, it turns out that the related clustering procedure does not consider the higher-dimensional dependencies among the involved random variables; for instance, they cannot correctly group variables that are pairwise independent but not globally independent. Finally, we compare novel proposed clustering algorithms (taking into account higher-dimensional dependencies) with classical agglomerative clustering methods.

E1577: **On Sklar's theorem with multivariate marginals**
Presenter: **Piotr Jaworski**, University of Warsaw, Poland
The well-known Sklar's theorem gives a characterization of a family of multivariate distribution functions with given univariate marginals. We will characterize a family of multivariate distribution functions with given multivariate marginals. Furthermore, we will present an algorithm how to construct such distribution functions.

E0795: **Exogenous shock models: Characterization and hierarchical construction**
Presenter: **Henrik Sloot**, Technical University of Munich, Germany
A new characterization theorem for the survival-functions of multivariate failure-times arising in exogenous shock models with non-negative, continuous, and unbounded shocks is presented. These survival-functions are the product of their ordered and individually transformed arguments — where the transformations may depend on the specific order and must fulfill a monotonicity condition. Conversely, every survival-function of that very form can be attained using an exogenous shock model. Furthermore, every extendible exogenous shock model has an alternative stochastic representation as a frailty-model involving some additive subordinator. Finally, examples based on hierarchical additive-frailty models are presented.

EO232 Room MAL 538 CLUSTERING/CLASSIFICATION AND MIXTURES I Chair: Geoffrey McLachlan

E0300: **Advances in robust estimation of skew normal mixtures**
Presenter: **Francesca Greselin**, University of Milano Bicocca, Italy
Co-authors: Luis Angel Garcia-Escudero, Agustin Mayo-Iscar, Geoffrey McLachlan
To extend the applicability of classic mixtures to dataset presenting asymmetry, mixtures of skew normal distributions are considered. A robust approach for their estimation allow us to analyze also data affected by outliers or deviations from model assumptions. Some issues related to truncated moment estimation are presented, and methods to overcome their awkward estimation will be discussed, also through applications to real data.

E1126: **Flexible mixtures of factor models based on skew component distributions**
Presenter: **Sharon Lee**, University of Queensland, Australia
Flexible mixtures of skew factor analyzers are gaining increasing attention, being exploited as powerful tools for the modelling, clustering, and dimension reduction of high-dimensional data that exhibit non-normal distributional features. These models have emerged as robust generalizations of the traditional mixtures of factor analyzers, where the assumption of normality for the latent factors is relaxed to cater for skewness in the observed data. We discuss several different formulations of skew factor models and propose to adopt a very flexible form of skew distribution as the density for the component latent factors. This allows the model to accommodate various types of skewness and asymmetry in the data, including multiple arbitrary directions of skewness. As such, it encompasses a number of commonly used models as special cases, such as some versions of the skew normal and skew t-factor analyzers. Parameter estimation can be carried out by maximum likelihood via an EM-type algorithm. The usefulness and potential of the proposed model are demonstrated using both real and simulated datasets.

E0708: **The importance of being clustered: Uncluttering the trends of statistics from 1970 to 2015**
Presenter: **Laura Anderlucci**, University of Bologna, Italy
Co-authors: Angela Montanari, Cinzia Viroli
The recent history of statistics is retraced by analyzing all the papers published in five prestigious statistical journals since 1970, namely: Annals of Statistics, Biometrika, Journal of the American Statistical Association, Journal of the Royal Statistical Society, series B and Statistical Science. The aim is to construct a kind of "taxonomy" of the statistical papers by organizing and clustering them in main themes. In this sense being identified in a cluster means being important enough to be uncluttered in the vast and interconnected world of the statistical research. Since the main statistical research topics naturally born, evolve or die during time, we will also develop a dynamic clustering strategy, where a group in a time period is allowed to migrate or to merge into different groups in the following one. Results show that statistics is a very dynamic and evolving science, stimulated by the rise of new research questions and types of data.

E1452: **Mixtures of locally-mapped support vector machines**
Presenter: **Hien Nguyen**, La Trobe University, Australia
Co-authors: Geoffrey McLachlan, Geoffrey McLachlan
Support vector machines (SVMs) have been highly successfully in application to classification problems of all sizes. However, the usual construction of SVMs only allow for a singular mapping between input and classification, even if that mapping is nonlinear, such as in the case of kernel SVMs. Via recent approaches to local-mapping of mixture regressions, we construct an approach to SVM classification that allows for different mappings in different parts of the input domain. We demonstrate how our mixtures of locally-mapped SVMs can be estimated via maximum quasi-likelihood estimation and the MM (minorization-maximization) algorithm. Furthermore, we present an online algorithm for its estimation in the face of big data. Some theory regarding the estimator and algorithms are presented, and example applications of the methodology are provided.

E1121: **Distributed and private model-based clustering**
Presenter: **Kaleb Leemaqz**, University of Queensland, Australia
Privacy is becoming increasingly important in collaborative data analysis, especially those involving personal or sensitive information commonly arising from health and commercial settings. The aim of privacy preserving statistical algorithms is to allow inference to be drawn on the joint data without disclosing private data held by each party. We propose a privacy-preserving EM (PPEM) algorithm, a novel scheme for training mixture models for clustering in a privacy-preserving manner. We focus on the case of horizontally distributed data among multiple parties for which cooperative learning is required. More specifically, each party wishes to learn the global parameters of the mixture model while preventing the leakage of party-specific information, including any intermediate results that may potentially be traced to an individual party. Another advantage of PPEM is that it does not involve a trusted third party, unlike most existing schemes that implement a master/slave hierarchy. This helps prevent

