

INDEX

Preface	1
----------------	----------

Keynote Speakers

Jean-Michel Loubes

Optimal transport methods for fairness in machine learning	5
---	----------

Peter Rousseeuw, Jakob Raymaekers and Mia Hubert

Class maps for visualizing classification results	6
--	----------

Robert Tibshirani, Stephen Bates and Trevor Hastie

Understanding cross-validation and prediction error	7
--	----------

Cinzia Viroli

Quantile-based classification	8
--------------------------------------	----------

Bin Yu

Veridical data science for responsible AI: characterizing V4 neurons through deepTune	9
--	----------

Plenary Session

Daniel Diaz

A simple correction for COVID-19 sampling bias	14
---	-----------

Jeffrey S. Morris

A seat at the table: the key role of biostatistics and data science in the COVID-19 pandemic	15
---	-----------

Bhramar Mukherjee

Predictions, role of interventions and the crisis of virus in India: a data science call to arms	16
---	-----------

Danny Pfeffermann

Contributions of Israel's CBS to rout COVID-19	17
---	-----------

Invited Papers

Claudio Agostinelli, Giovanni Saraceno and Luca Greco

Robust issues in estimating modes for multivariate torus data	21
--	-----------

Emanuele Aliverti

Bayesian nonparametric dynamic modeling of psychological traits	25
--	-----------

<i>Andres M. Alonso, Carolina Gamboa and Daniel Peña</i> Clustering financial time series using generalized cross correlations	27
<i>Raffaele Argiento, Edoardo Filippi-Mazzola and Lucia Paci</i> Model-based clustering for categorical data via Hamming distance	31
<i>Antonio Balzanella, Antonio Irpino and Francisco de A.T. De Carvalho</i> Mining multiple time sequences through co-clustering algorithms for distributional data	32
<i>Francesco Bartolucci, Fulvia Pennoni and Federico Cortese</i> Hidden Markov and regime switching copula models for state allocation in multiple time-series	36
<i>Michela Battauz and Paolo Vidoni</i> Boosting multidimensional IRT models	40
<i>Matteo Bottai</i> Understanding and estimating conditional parametric quantile models	44
<i>Niklas Bussmann, Roman Enzmann, Paolo Giudici and Emanuela Raffinetti</i> Shapley Lorenz methods for eXplainable artificial intelligence	45
<i>Andrea Cappelozzo, Ludovic Duponchel, Francesca Greselin and Brendan Murphy</i> Robust classification of spectroscopic data in agri-food: first analysis on the stability of results	49
<i>Andrea Cerasa, Enrico Checchi, Domenico Perrotta and Francesca Torti</i> Issues in monitoring the EU trade of critical COVID-19 commodities	53
<i>Marcello Chiodi</i> Smoothed non linear PCA for multivariate data	54
<i>Roberto Colombi, Sabrina Giordano and Maria Kateri</i> Accounting for response behavior in longitudinal rating data	58
<i>Claudio Conversano, Giulia Contu, Luca Frigau and Carmela Cappelli</i> Network-based semi-supervised clustering of time series data	62
<i>Federica Cugnata, Chiara Brombin, Pietro Cippà, Alessandro Ceschi, Paolo Ferrari and Clelia Di Serio</i> Characterising longitudinal trajectories of COVID-19 biomarkers within a latent class framework	64
<i>Silvia D'Angelo</i> Sender and receiver effects in latent space models for multiplex data	68
<i>Anna Denkowska and Stanisław Wanat</i> DTW-based assessment of the predictive power of the copula-DCC-GARCH-MST model developed for European insurance institutions	71
<i>Roberto Di Mari, Zsuzsa Bakk, Jennifer Oser and Jouni Kuha</i> Two-step estimation of multilevel latent class models with covariates	75
<i>Marie Du Roy de Chaumaray and Matthieu Marbac</i> Clustering data with non-ignorable missingness using semi-parametric mixture models	79

<i>Pierpaolo D'Urso, Livia De Giovanni and Vincenzina Vitale</i> Spatial-temporal clustering based on B-splines: robust models with applications to COVID-19 pandemic	83
<i>Leonardo Egidi, Roberta Pappadà, Francesco Pauli and Nicola Torelli</i> PIVMET: pivotal methods for Bayesian relabelling in finite mixture models	87
<i>Tahir Ekin and Claudio Conversano</i> Cluster validity by random forests	91
<i>Luis Angel García-Escudero, Agustín Mayo-Iscar and Marco Riani</i> Robust estimation of parsimonious finite mixture of Gaussian models	92
<i>Silvia Facchinetti and Silvia Angela Osmetti</i> A risk indicator for categorical data	93
<i>Matteo Fasiolo</i> Additive quantile regression via the qgam R package	97
<i>Michael Fop, Dimitris Karlis, Ioannis Kosmidis, Adrian O'Hagan, Caitriona Ryan and Isobel Claire Gormley</i> Gaussian mixture models for high dimensional data using composite likelihood	98
<i>Carlo Gaetan, Paolo Girardi and Victor Muthama Musau</i> On model-based clustering using quantile regression	102
<i>Carlotta Galeone</i> Socioeconomic inequalities and cancer risk: myth or reality?	106
<i>Michael Gallagher, Christophe Biernacki and Paul McNicholas</i> Parameter-wise co-clustering for high dimensional data	107
<i>Francesca Greselin and Alina Jędrzejczak</i> Quantifying the impact of covariates on the gender gap measurement: an analysis based on EU-SILC data from Poland and Italy	108
<i>Alessandra Guglielmi, Mario Beraha, Matteo Giannella, Matteo Pegoraro and Riccardo Peli</i> A transdimensional MCMC sampler for spatially dependent mixture models	112
<i>Christian Hennig and Pietro Coretto</i> Non-parametric consistency for the Gaussian mixture maximum likelihood estimator	116
<i>Yinxuan Huang and Natalie Shlomo</i> Improving the reliability of a nonprobability web survey	120
<i>Maria Iannario and Claudia Tarantola</i> A semi-Bayesian approach for the analysis of scale effects in ordinal regression models	124
<i>Jayant Jha</i> Best approach direction for spherical random variables	128

<i>Maria Kateri</i>	
Simple effect measures for interpreting generalized binary regression models	129
<i>Shogo Kato, Kota Nagasaki and Wataru Nakanishi</i>	
Mixtures of Kato–Jones distributions on the circle, with an application to traffic count data	133
<i>John Kent</i>	
How to design a directional distribution	137
<i>Simona Korenjak-Černe and Nataša Kejžar</i>	
Identifying mortality patterns of main causes of death among young EU population using SDA approaches	141
<i>Fabrizio Laurini and Gianluca Morelli</i>	
Robust supervised clustering: some practical issues	142
<i>Daniela Marella and Danny Pfeffermann</i>	
A nonparametric approach for statistical matching under informative sampling and nonresponse	146
<i>Mariagiulia Matteucci and Stefania Mignani</i>	
Investigating model fit in item response models with the Hellinger distance	150
<i>Matteo Mazziotta and Adriano Pareto</i>	
PCA-based composite indices and measurement model	154
<i>Marcella Mazzoleni, Angiola Pollastri and Vanda Tulli</i>	
Gender inequalities from an income perspective	158
<i>Yana Melnykov, Xuwen Zhu and Volodymyr Melnykov</i>	
Transformation mixture modeling for skewed data groups with heavy tails and scatter	162
<i>Luca Merlo, Lea Petrella and Nikos Tzavidis</i>	
Unconditional M-quantile regression	163
<i>Jesper Møller, Mario Beraha, Raffaele Argiento and Alessandra Guglielmi</i>	
MCMC computations for Bayesian mixture models using repulsive point processes	167
<i>Keefe Murphy, Cinzia Viroli and Isobel Claire Gormley</i>	
Infinite mixtures of infinite factor analysers	168
<i>Stanislav Nagy, Petra Laketa and Rainer Dyckerhoff</i>	
Angular halfspace depth: computation	169
<i>Yarema Okhrin, Gazi Salah Uddin and Muhammad Yahya</i>	
Nonlinear Interconnectedness of crude oil and financial markets	173
<i>M. Rosário Oliveira, Ana Subtil and Lina Oliveira</i>	
Detection of internet attacks with histogram principal component analysis	174
<i>Sally Paganin</i>	
Semiparametric IRT models for non-normal latent traits	178

<i>Giuseppe Pandolfo</i>	
A graphical depth-based aid to detect deviation from unimodality on hyperspheres	182
<i>Panos Pardalos</i>	
Networks of networks	186
<i>Xanthi Pedeli and Cristiano Varin</i>	
Pairwise likelihood estimation of latent autoregressive count models	187
<i>Mark Reiser and Maduranga Dassanayake</i>	
A study of lack-of-fit diagnostics for models fit to cross-classified binary variables	191
<i>Giorgia Riveccio, Jean-Paul Chavas, Giovanni De Luca, Salvatore Di Falco and Fabian Capitanio</i>	
Assessing food security issues in Italy: a quantile copula approach	195
<i>Nicoleta Rogovschi</i>	
Co-clustering for high dimensional sparse data	199
<i>Massimiliano Russo</i>	
Malaria risk detection via mixed membership models	203
<i>Paula Saavedra-Nieves and Rosa M. Crujeiras</i>	
Nonparametric estimation of the number of clusters for directional data	207
<i>Shuchismita Sarkar, Volodymyr Melnykov and Xuwen Zhu</i>	
Tensor-variate finite mixture model for the analysis of university professor remuneration	208
<i>Florian Schuberth</i>	
Specifying composites in structural equation modeling: the Henseler-Ogasawara specification	209
<i>Jarod Smith, Mohammad Arashi and Andriette Bekker</i>	
Network analysis implementing a mixture distribution from Bayesian viewpoint	210
<i>Paul Smith, Peter van der Heijden and Maarten Cruyff</i>	
Measurement errors in multiple systems estimation	211
<i>Valentin Todorov and Peter Filzmoser</i>	
Robust classification in high dimensions using regularized covariance estimates	215
<i>Salvatore Daniele Tomarchio, Luca Bagnato and Antonio Punzo</i>	
Clustering via new parsimonious mixtures of heavy tailed distributions	216
<i>Agostino Torti, Marta Galvani, Alessandra Menafoglio, Piercesare Secchi and Simone Vantini</i>	
A general bi-clustering technique for functional data	217
<i>Laura Trinchera</i>	
Developing a multidimensional and hierarchical index following a composite-based approach	220

<i>Rosanna Verde, Francisco T. de A. De Carvalho and Antonio Balzanella</i> A generalised clusterwise regression for distributional data	223
<i>Marika Vezzoli, Francesco Doglietto, Stefano Renzetti, Marco Fontanella and Stefano Calza</i> A machine learning approach for evaluating anxiety in neurosurgical patients during the COVID-19 pandemic	227
<i>Isadora Antoniano Villalobos, Simone Padoan and Boris Beranger</i> Prediction of large observations via Bayesian inference for extreme-value theory	231
<i>Maria Prosperina Vitale, Vincenzo Giuseppe Genova, Giuseppe Giordano and Giancarlo Ragozini</i> Community detection in tripartite networks of university student mobility flows	232
<i>Ernst Wit and Lucas Kania</i> Causal regularization	236
<i>Qiuyi Wu and David Banks</i> Minimizing conflicts of interest: optimizing the JSM program	240

Contributed Papers

<i>Antonino Abbruzzo, Maria Francesca Cracolici and Furio Urso</i> Model selection procedure for mixture hidden Markov models	243
<i>Roberto Ascari and Sonia Migliorati</i> A full mixture of experts model to classify constrained data	247
<i>Luigi Augugliaro, Gianluca Sottile and Angelo Mineo</i> Sparse inference in covariate adjusted censored Gaussian graphical models	251
<i>Simona Balzano, Mario Rosario Guarracino and Giovanni Camillo Porzio</i> Semi-supervised learning through depth functions	255
<i>Lucio Barabesi, Andrea Cerasa, Andrea Cerioli and Domenico Perrotta</i> A combined test of the Benford hypothesis with anti-fraud applications	256
<i>Chiara Bardelli</i> Unbalanced classification of electronic invoicing	260
<i>Claudia Berloco, Raffaele Argiento and Silvia Montagna</i> Predictive power of Bayesian CAR models on scale free networks: an application for credit risk	264
<i>Marco Berrettini, Giuliano Galimberti and Saverio Ranciati</i> Semiparametric finite mixture of regression models with Bayesian P-splines	268

<i>Giuseppe Bove</i>	
A subject-specific measure of interrater agreement based on the homogeneity index	272
<i>Antonio Calcagni</i>	
Estimating latent linear correlations from fuzzy contingency tables	276
<i>Andrea Cappozzo, Alessandro Casa and Michael Fop</i>	
Model-based clustering with sparse matrix mixture models	280
<i>Andrea Cappozzo, Luis Angel Garcia Escudero, Francesca Greselin and Agustín Mayo-Iscar</i>	
Exploring solutions via monitoring for cluster weighted robust models	284
<i>Maurizio Carpita and Silvia Golia</i>	
Categorical classifiers in multi-class classification problems	288
<i>Gianmarco Caruso, Greta Panunzi, Marco Mingione, Pierfrancesco Alaimo Di Loro, Stefano Moro, Edoardo Bompiani, Caterina Lanfredi, Daniela Silvia Pace, Luca Tardella and Giovanna Jona Lasinio</i>	
Model-based clustering for estimating cetaceans site-fidelity and abundance	292
<i>Carlo Cavicchia, Maurizio Vichi and Giorgia Zaccaria</i>	
Model-based clustering with parsimonious covariance structure	296
<i>Francesca Condino</i>	
Clustering income data based on share densities	300
<i>Paula Costa Fontichiarì, Miriam Giuliani, Raffaele Argiento and Lucia Paci</i>	
Group-dependent finite mixture model	304
<i>Salvatore Cuomo, Federico Gatta, Fabio Giampaolo, Carmela Iorio and Francesco Piccialli</i>	
A machine learning approach in stock risk management	308
<i>Cristina Davino and Giuseppe Lamberti</i>	
Pathmix segmentation trees to compare linear regression models	312
<i>Houyem Demni, Davide Buttarazzi, Stanislav Nagy and Giovanni Camillo Porzio</i>	
Angular halfspace depth: classification using spherical bagdistances	316
<i>Agostino Di Ciaccio</i>	
Neural networks for high cardinality categorical data	320
<i>F. Marta L. Di Lascio, Andrea Menapace and Roberta Pappadà</i>	
Ali-Mikhail-Haq copula to detect low correlations in hierarchical clustering	324
<i>Maria Veronica Dorgali, Silvia Bacci, Bruno Bertaccini and Alessandra Petrucci</i>	
Higher education and employability: insights from the mandatory notices of the ministry of labour	328
<i>Lorenzo Focardi Olmi and Anna Gottard</i>	
An alternative to joint graphical lasso for learning multiple Gaussian graphical models	332

<i>Francesca Fortuna, Alessia Naccarato and Silvia Terzi</i>	
Functional cluster analysis of HDI evolution in European countries	336
<i>Sylvia Frühwirth-Schnatter, Bettina Grün and Gertraud Malsiner-Walli</i>	
Estimating Bayesian mixtures of finite mixtures with telescoping sampling	340
<i>Chiara Galimberti, Federico Castelletti and Stefano Peluso</i>	
A Bayesian framework for structural learning of mixed graphical models	344
<i>Andrea Gilardi, Riccardo Borgoni, Luca Presicce and Jorge Mateu</i>	
Measurement error models on spatial network lattices: car crashes in Leeds	348
<i>Carmela Iorio, Giuseppe Pandolfo, Michele Staiano, Massimo Aria and Roberta Siciliano</i>	
The L^P data depth and its application to multivariate process control charts	352
<i>Petra Laketa and Stanislav Nagy</i>	
Angular halfspace depth: central regions	356
<i>Michele La Rocca, Francesco Giordano and Cira Perna</i>	
Clustering production indexes for construction with forecast distributions	360
<i>Maria Mannone, Veronica Distefano, Claudio Silvestri and Irene Poli</i>	
Clustering longitudinal data with category theory for diabetic kidney disease	364
<i>Laura Marcis, Maria Chiara Pagliarella and Renato Salvatore</i>	
A redundancy analysis with multivariate random-coefficients linear models	368
<i>Paolo Mariani, Andrea Marletta and Matteo Locci</i>	
The use of multiple imputation techniques for social media data	372
<i>Federico Marotta, Paolo Provero and Silvia Montagna</i>	
Prediction of gene expression from transcription factors affinities: an application of Bayesian non-linear modelling	376
<i>Francesca Martella, Fabio Attorre, Michele De Sanctis and Giuliano Fanelli</i>	
High dimensional model-based clustering of European georeferenced vegetation plots	380
<i>Ana Martins, Paula Brito, Sónia Dias and Peter Filzmoser</i>	
Multivariate outlier detection for histogram-valued variables	384
<i>Giovanna Menardi and Federico Ferraccioli</i>	
A nonparametric test for mode significance	388
<i>Massimo Mucciardi, Giovanni Pirrotta, Andrea Briglia and Arnaud Sallaberry</i>	
Visualizing cluster of words: a graphical approach to grammar acquisition	392

<i>Marta Nai Ruscone and Dimitris Karlis</i> Robustness methods for modelling count data with general dependence structures	396
<i>Roberta Paroli, Luigi Spezia, Marc Stutter and Andy Vinten</i> Bayesian analysis of a water quality high-frequency time series through Markov switching autoregressive models	400
<i>Mariano Porcu, Isabella Sulis and Cristian Usala</i> Detecting the effect of secondary school in higher education university choices	404
<i>Roberto Rocci and Monia Ranalli</i> Semi-constrained model-based clustering of mixed-type data using a composite likelihood approach	408
<i>Annalina Sarra, Adelia Evangelista, Tonio Di Battista and Damiana Pieragostino</i> Antibodies to SARS-CoV-2: an exploratory analysis carried out through the Bayesian profile regression	412
<i>Theresa Scharl and Bettina Grün</i> Modelling three-way RNA sequencing data with mixture of multivariate Poisson-lognormal distribution	416
<i>Luca Scrucca</i> Stacking ensemble of Gaussian mixtures	420
<i>Rosaria Simone, Cristina Davino, Domenico Vistocco and Gerhard Tutz</i> A robust quantile approach to ordinal trees	424
<i>Venera Tomaselli, Giulio Giacomo Cantone and Valeria Mazzeo</i> The detection of spam behaviour in review bomb	428
<i>Donatella Vicari and Paolo Giordani</i> Clustering models for three-way data	432
<i>Gianpaolo Zammarchi and Jaromir Antoch</i> Using eye-tracking data to create a weighted dictionary for sentiment analysis: the eye dictionary	436