

Contents

Part I Key Notes

Principal Component Analysis for Categorical Histogram Data: Some Open Directions of Research	3
Edwin Diday	
Factorial Conjoint Analysis Based Methodologies	17
Giuseppe Giordano, Carlo Natale Lauro, and Germana Scepti	
Ordering and Scaling Objects in Multivariate Data Under Nonlinear Transformations of Variables	29
Jacqueline J. Meulman, Lawrence J. Hubert, and Phipps Arabie	
Statistical Models to Predict Academic Churn Risk	41
Paolo Giudici and Emanuele Dequarti	
The Poisson Processes in Cluster Analysis	51
André Hardy	
TWO-CLASS Trees for Non-Parametric Regression Analysis	63
Roberta Siciliano and Massimo Aria	

Part II Classification and Discrimination

Efficient Incorporation of Additional Information to Classification Rules .	75
Miguel Fernández, Cristina Rueda, and Bonifacio Salvador	
The Choice of the Parameter Values in a Multivariate Model of a Second Order Surface with Heteroscedastic Error	85
Umberto Magagnoli and Gabriele Cantaluppi	

Mixed Mode Data Clustering: An Approach Based on Tetrachoric Correlations	95
Isabella Morlini	

Optimal Scaling Trees for Three-Way Data	105
Valerio A. Tutore	

Part III Data Mining

A Study on Text Modelling via Dirichlet Compound Multinomial	115
Concetto Elvio Bonafede and Paola Cerchiello	

Automatic Multilevel Thresholding Based on a Fuzzy Entropy Measure ..	125
D. Bruzzese and U. Giani	

Some Developments in Forward Search Clustering	135
Daniela G. Calò	

Spectral Graph Theory Tools for Social Network Comparison	145
Domenico De Stefano	

Improving the MHIST-p Algorithm for Multivariate Histograms of Continuous Data	155
Mauro Iacono and Antonio Irpino	

On Building and Visualizing Proximity Graphs for Large Data Sets with Artificial Ants	165
Julien Lavergne, Hanane Azzag, Christiane Guinot, and Gilles Venturini	

Including Empirical Prior Information in Test Administration	173
Mariagiulia Matteucci and Bernard P. Veldkamp	

Part IV Robustness and Classification

Italian Firms' Geographical Location in High-tech Industries: A Robust Analysis	185
Matilde Bini and Margherita Velucchi	

Robust Tests for Pareto Density Estimation	193
Aldo Corbellini and Lisa Crosato	

Bootstrap and Nonparametric Predictors to Impute Missing Data	203
Agostino Di Ciaccio	

On the Use of Boosting Procedures to Predict the Risk of Default	211
Giovanna Menardi, Federico Tedeschi and Nicola Torelli	

Part V Categorical Data and Latent Class Approach

Assessing Similarity of Rating Distributions by Kullback-Leibler Divergence	221
Marcella Corduas	

Sector Classification in Stock Markets: A Latent Class Approach	229
Michele Costa and Luca De Angelis	

Partitioning the Geometric Variability in Multivariate Analysis and Contingency Tables	237
Carles M. Cuadras and Daniel Cuadras	

One-Dimensional Preference Data Imputation Through Transition Rules .	245
Luigi Fabbri	

About a Type of Quasi Linear Estimating Equation Approach	253
Giulio D'Epifanio	

Causal Inference Through Principal Stratification: A Special Type of Latent Class Modelling	265
Leonardo Grilli	

Scaling the Latent Variable Cultural Capital via Item Response Models and Latent Class Analysis	271
Isabella Sulis, Mariano Porcu, and Marco Pitzalis	

Assessment of Latent Class Detection in PLS Path Modeling: a Simulation Study to Evaluate the Group Quality Index performance . . .	281
Laura Trinchera	

Part VI Latent Variables and Related Methods

Non-Linear Relationships in SEM with Latent Variables: Some Theoretical Remarks and a Case Study	293
Giuseppe Boari, Gabriele Cantaluppi, and Stefano Bertelli	

Multidimensional Scaling Versus Multiple Correspondence Analysis When Analyzing Categorization Data	301
Marine Cadoret, Sébastien Lê, and Jérôme Pagès	
Multidimensional Scaling as Visualization Tool of Web Sequence Rules ..	309
Antonio D'Ambrosio and Marcello Pecoraro	
Partial Compliance, Effect of Treatment on the Treated and Instrumental Variables	317
Antonio Forcina	
Method of Quantification for Qualitative Variables and their Use in the Structural Equations Models	325
C. Lauro, D. Nappo, M.G. Grassia, and R. Miele	
Monitoring Panel Performance Within and Between Sensory Experiments by Multi-Way Analysis	335
Rosaria Romano, Jannie S. Vestergaard, Mohsen Kompany-Zareh, and Wender L.P. Bredie	
A Proposal for Handling Categorical Predictors in PLS Regression Framework	343
Giorgio Russolillo and Carlo Natale Lauro	
Part VII Symbolic, Multivalued and Conceptual Data Analysis	
On the Use of Archetypes and Interval Coding in Sensory Analysis	353
Maria Rosaria D'Esposito, Francesco Palumbo, and Giancarlo Ragozini	
From Histogram Data to Model Data Analysis	363
Marina Marino and Simona Signoriello	
Use of Genetic Algorithms When Computing Variance of Interval Data ..	371
Jaromír Antoch and Raffaele Miele	
Spatial Visualization of Conceptual Data	379
Michel Soto, Bénédicte Le Grand, and Marie-Aude Aufaure	
Part VIII Spatial, Temporal, Streaming and Functional Data Analysis	
A Test of LBO Firms' Acquisition Rationale: The French Case	391
R. Abdesselam, S. Cieply and A.L. Le Nadant	

Kernel Intensity for Space-Time Point Processes with Application to Seismological Problems 401
Giada Adelfio and Marcello Chiodi

Summarizing and Mining Streaming Data via a Functional Data Approach 409
Antonio Balzanella, Elvira Romano, and Rosanna Verde

Clustering Complex Time Series Databases 417
Francesco Giordano, Michele La Rocca, and Maria Lucia Parrella

Use of a Flexible Weight Matrix in a Local Spatial Statistic 427
Massimo Mucciardi

Constrained Variable Clustering and the Best Basis Problem in Functional Data Analysis 435
Fabrice Rossi and Yves Lechevallier

Part IX Bio and Health Science

Plaid Model for Microarray Data: an Enhancement of the Pruning Step . 447
Luigi Augugliaro and Angelo M. Mineo

Classification of the Human Papilloma Viruses 457
Abdoulaye Baniré Diallo, Dunarel Badescu, Mathieu Blanchette, and Vladimir Makarenkov

Toward the Discovery of Itemsets with Significant Variations in Gene Expression Matrices 465
Mehdi Kaytoue, Sébastien Duplessis, and Amedeo Napoli