Contents

Part I Clustering Large Data-Sets

Clustering Large Data Set: An Applied Comparative Study 3
Laura Bocci and Isabella Mingo

Clustering in Feature Space for Interesting Pattern Identification of Categorical Data 13
Marina Marino, Francesco Palumbo and Cristina Tortora

Clustering Geostatistical Functional Data 23
Elvira Romano and Rosanna Verde

Joint Clustering and Alignment of Functional Data: An Application to Vascular Geometries 33
Laura M. Sangalli, Piercesare Secchi, Simone Vantini, and Valeria Vitelli

Part II Statistics in Medicine

Bayesian Methods for Time Course Microarray Analysis: From Genes' Detection to Clustering 47
Claudia Angelini, Daniela De Canditiis, and Marianna Pensky

Longitudinal Analysis of Gene Expression Profiles Using Functional Mixed-Effects Models 57
Maurice Berk, Cheryl Hemingway, Michael Levin, and Giovanni Montana

A Permutation Solution to Compare Two Hepatocellular Carcinoma Markers 69
Agata Zirilli and Angela Alibrandi

Bibliografische Informationen
http://d-nb.info/1011288257

digitalisiert durch DEUTSCHE NATIONAL BIBLIOTHEK

Part III Integrating Administrative Data

Statistical Perspective on Blocking Methods When Linking Large Data-sets 81
Nicoletta Cibella and Tiziana Tuoto

Integrating Households Income Microdata in the Estimate of the Italian GDP 91
Alessandra Coli and Francesca Tartamella

The Employment Consequences of Globalization: Linking Data on Employers and Employees in the Netherlands 101
Fabienne Fortanier, Marjolein Korvorst, and Martin Luppes

Applications of Bayesian Networks in Official Statistics 113
Paola Vicard and Mauro Scanu

Part IV Outliers and Missing Data

A Correlated Random Effects Model for Longitudinal Data with Non-ignorable Drop-Out: An Application to University Student Performance 127
Filippo Belloc, Antonello Maruotti, and Lea Petrella

Risk Analysis Approaches to Rank Outliers in Trade Data 137
Vytis Kopustinskas and Spyros Arsenis

Problems and Challenges in the Analysis of Complex Data: Static and Dynamic Approaches 145
Marco Riani, Anthony Atkinson and Andrea Cerioli

Ensemble Support Vector Regression: A New Non-parametric Approach for Multiple Imputation 159
Daria Scacciatelli

Part V Time Series Analysis

On the Use of PLS Regression for Forecasting Large Sets of Cointegrated Time Series 171
Gianluca Cubadda and Barbara Guardabascio

Large-Scale Portfolio Optimisation with Heuristics 181
Manfred Gilli and Enrico Schumann

Detecting Short-Term Cycles in Complex Time Series Databases 193
F. Giordano, M.L. Parrella and M. Restaino

Assessing the Beneficial Effects of Economic Growth: The Harmonic Growth Index 205
Daria Mendola and Raffaele Scuderi

Time Series Convergence within I(2) Models: the Case of Weekly Long Term Bond Yields in the Four Largest Euro Area Countries 217
Giuliana Passamani

Part VI Environmental Statistics

Anthropogenic CO_2 Emissions and Global Warming: Evidence from Granger Causality Analysis 229
Massimo Bilancia and Domenico Vitale

Temporal and Spatial Statistical Methods to Remove External Effects on Groundwater Levels 241
Daniele Imparato, Andrea Carena, and Mauro Gasparini

Reduced Rank Covariances for the Analysis of Environmental Data 253
Orietta Nicolis and Doug Nychka

Radon Level in Dwellings and Uranium Content in Soil in the Abruzzo Region: A Preliminary Investigation by Geographically Weighted Regression 265
Eugenia Nissi, Annalina Sarra, and Sergio Palermi

Part VII Probability and Density Estimation

Applications of Large Deviations to Hidden Markov Chains Estimation 279
Fabiola Del Greco M.

Multivariate Tail Dependence Coefficients for Archimedean Copulae 287
Giovanni De Luca and Giorgia Rivieccio

A Note on Density Estimation for Circular Data 297
Marco Di Marzio, Agnese Panzera, and Charles C. Taylor

Markov Bases for Sudoku Grids 305
Roberto Fontana, Fabio Rapallo, and Maria Piera Rogantin

Part VIII Application in Economics

Estimating the Probability of Moonlighting in Italian Building Industry 319
Maria Felice Arezzo and Giorgio Alleva

Use of Interactive Plots and Tables for Robust Analysis of International Trade Data 329
Domenico Perrotta and Francesca Torti

Generational Determinants on the Employment Choice in Italy 339
Claudio Quintano, Rosalia Castellano, and Gennaro Punzo

Route-Based Performance Evaluation Using Data Envelopment Analysis Combined with Principal Component Analysis 351
Agnese Rapposelli

Part IX WEB and Text Mining

Web Surveys: Methodological Problems and Research Perspectives 363
Silvia Biffignandi and Jelke Bethlehem

Semantic Based DCM Models for Text Classification 375
Paola Cerchiello

Probabilistic Relational Models for Operational Risk: A New Application Area and an Implementation Using Domain Ontologies 385
Marcus Spies

Part X Advances on Surveys

Efficient Statistical Sample Designs in a GIS for Monitoring the Landscape Changes 399
Elisabetta Carfagna, Patrizia Tassinari, Maroussa Zagoraiou, Stefano Benni, and Daniele Torreggiani

Studying Foreigners' Migration Flows Through a Network Analysis Approach 409
Cinzia Conti, Domenico Gabrielli, Antonella Guarneri, and Enrico Tucci

Estimation of Income Quantiles at the Small Area Level in Tuscany 419
Caterina Giusti, Stefano Marchetti and Monica Pratesi

The Effects of Socioeconomic Background and Test-taking Motivation on Italian Students' Achievement 429
Claudio Quintano, Rosalia Castellano, and Sergio Longobardi

Part XI Multivariate Analysis

Firm Size Dynamics in an Industrial District: The Mover-Stayer Model in Action 443
F. Cipollini, C. Ferretti, and P. Ganugi

Multiple Correspondence Analysis for the Quantification and Visualization of Large Categorical Data Sets 453
Alfonso Iodice D'Enza and Michael Greenacre

Multivariate Ranks-Based Concordance Indexes 465
Emanuela Raffinetti and Paolo Giudici

Methods for Reconciling the Micro and the Macro in Family Demography Research: A Systematisation 475
Anna Matysiak and Daniele Vignoli