

Book of Short Papers

SIS 2021



Editors: **Cira Perna, Nicola Salvati and Francesco Schirripa Spagnolo**



Distribuzione Software | Formazione Professionale
Statistica | Economia | Finanza | Biostatistica | Epidemiologia
Sanità Pubblica | Scienze Sociali
www.tstat.it | www.tstattraining.eu

Copyright © 2021

PUBLISHED BY PEARSON

WWW.PEARSON.COM

ISBN 9788891927361

Contents

Preface	XIX
1 Plenary Sessions	1
1.1 Citizen data, and citizen science: a challenge for official statistics. <i>Monica Pratesi</i>	2
2 Specialized Sessions	8
2.1 A glimpse of new data and methods for analysing a rapidly changing population	9
2.1.1 The diffusion of new family patterns in Italy: An update. <i>Arnstein Aassve, Letizia Mencarini, Elena Pirani and Daniele Vignoli</i>	10
2.1.2 Causes of death patterns and life expectancy: looking for warning signals. <i>Stefano Mazzucco, Emanuele Aliverti, Daniele Durante and Stefano Campostrini</i>	16
2.2 Advances in ecological modelling	22
2.2.1 A Bayesian joint model for exploring zero-inflated bivariate marine litter data. <i>Sara Martino, Crescenza Calculli and Porzia Maiorano</i>	23
2.3 Advances in environmental statistics	29
2.3.1 Bayesian small area models for investigating spatial heterogeneity and factors affecting the amount of solid waste in Italy. <i>Crescenza Calculli and Serena Arima</i>	30
2.3.2 A spatial regression model for for predicting abundance of lichen functional groups. <i>Pasquale Valentini, Francesca Fortuna, Tonio Di Battista and Paolo Giordani</i>	36
2.4 Advances in preference and ordinal data theoretical improvements and applications	42
2.4.1 Boosting for ranking data: an extension to item weighting. <i>Alessandro Albano, Mariangela Sciandra and Antonella Plaia</i>	43
2.4.2 An Extended Bradley-Terry Model For The Analysis Of Financial Data. <i>Alessio Baldassarre, Elise Dusseldorp and Mark De Rooij</i>	49

2.5	Business system innovation, competitiveness, productivity and internationalization	55
2.5.1	An analysis of the dynamics of the competitiveness for some European Countries. <i>Andrea Marletta, Mauro Mussini and Mariangela Zenga</i>	56
2.5.2	National innovation system and economic performance in EU. An analysis using composite indicators. <i>Alessandro Zeli</i>	62
2.6	Challenges for observational studies in modern biomedicine	68
2.6.1	Data integration: a Statistical view. <i>Pier Luigi Conti</i>	69
2.6.2	Exploring patients' profile from COVID-19 case series data: beyond standard statistical approaches. <i>Chiara Brombin, Federica Cugnata, Pietro E. Cippà, Alessandro Ceschi, Paolo Ferrari and Clelia di Serio</i>	75
2.6.3	On the statistics for some pivotal anti-COVID-19 vaccine trials. <i>Mauro Gasparini</i>	81
2.7	Data Science for Industry 4.0 (ENBIS)	87
2.7.1	Sample selection from a given dataset to validate machine learning models. <i>Bertrand Iooss</i>	88
2.7.2	Reliable data-drive modelling and optimisation of a batch reactor using bootstrap aggregated deep belief networks. <i>Changhao Zhu and Jie Zhang</i>	94
2.8	Integration of survey with alternative sources of data	100
2.8.1	A parametric empirical likelihood approach to data matching under nonignorable sampling and nonresponse. <i>Daniela Marella and Danny Pfeffermann</i>	101
2.8.2	Survey data integration for regression analysis using model calibration.	107
2.8.3	Latent Mixed Markov Models for the Production of Population Census Data on Employment. <i>Daniela Filippini, Ugo Guarnera and Roberta Varriale</i>	112
2.9	Media, social media and demographic behaviours	118
2.9.1	Monitoring the Numbers of European Migrants in the United Kingdom using Facebook Data. <i>Francesco Rampazzo, Jakub Bijak, Agnese Vitali, Ingmar Weber and Emilio Zagheni</i>	119
2.10	New developments in ensemble methods for classification	125
2.10.1	An alternative approach for nowcasting economic activity during COVID-19 times. <i>Alessandro Spelta and Paolo Pagnottoni</i>	126
2.10.2	Assessing the number of groups in consensus clustering by pivotal methods. <i>Roberta Pappadà, Francesco Pauli and Nicola Torelli</i>	132
2.10.3	Clustering of data recorded by Distributed Acoustic Sensors to identify vehicle passage and typology. <i>Antonio Balzanella and Stefania Nacchia</i>	138
2.11	New developments in latent variable models	144
2.11.1	A Hidden Markov Model for Variable Selection with Missing Values. <i>Fulvia Pennoni, Francesco Bartolucci, and Silvia Pandolfi</i>	145
2.11.2	Comparison between Different Likelihood Based Estimation Methods in Latent Variable Models for Categorical Data. <i>Silvia Bianconcini and Silvia Cagnone</i>	151
2.11.3	A Comparison of Estimation Methods for the Rasch Model. <i>Alexander Robitzsch</i>	157

2.12	New issues on multivariate and univariate quantile regression	163
2.12.1	Directional M-quantile regression for multivariate dependent outcomes. <i>Luca Merlo, Lea Petrella and Nikos Tzavidis</i>	164
2.13	Semi-parametric and non-parametric latent class analysis	170
2.13.1	Stepwise Estimation of Multilevel Latent Class Models. <i>Zsuzsa Bakk, Roberto di Mari, Jennifer Oser and Jouni Kuha</i>	171
2.13.2	Distance learning, stress and career-related anxiety during the Covid-19 pandemic: a students perspective analysis. <i>Alfonso Iodice D'Enza, Maria Iannario, Rosaria Romano</i>	177
2.13.3	A Tempered Expectation-Maximization Algorithm for Latent Class Model Estimation. <i>Luca Brusa, Francesco Bartolucci and Fulvia Pennoni</i>	183
2.14	Statistics for finance high frequency data, large dimension and networks	189
2.14.1	The Italian debt not-so-flash crash. <i>Maria Flora and Roberto Reno'</i>	190
3	Solicited Sessions	197
3.1	Advances in social indicators research and latent variables modelling in social sciences	198
3.1.1	A composite indicator to measure frailty using administrative healthcare data. <i>Margherita Silan, Rachele Brocco and Giovanna Boccuzzo</i>	199
3.1.2	Clusters of contracting authorities over time: an analysis of their behaviour based on procurement red flags. <i>Simone Del Sarto, Paolo Coppola and Matteo Troia</i>	205
3.1.3	An Application of Temporal Poset on Human Development Index Data. <i>Leonardo Salvatore Alaimo, Filomena Maggino and Emiliano Seri</i>	211
3.1.4	The SDGs System: a longitudinal analysis through PLS-PM. <i>Rosanna Cataldo, Maria Gabriella Grassia and Laura Antonucci</i>	217
3.2	Changes in the life course and social inequality	223
3.2.1	Heterogeneous Income Dynamics: Unemployment Consequences in Germany and the US. <i>Raffaele Grotti</i>	224
3.2.2	In-work poverty in Germany and in the US: The role of parity progression. <i>Emanuela Struffolino and Zachary Van Winkle Z.</i>	230
3.2.3	Parenthood, education and social stratification. An analysis of female occupational careers in Italy. <i>Gabriele Ballarino and Stefano Cantalini</i>	236
3.3	Composition in the Data Science Era	242
3.3.1	Can we Ignore the Compositional Nature of Compositional Data by using Deep Learning Approaches? <i>Matthias Templ</i>	243
3.3.2	Principal balances for three-way compositions. <i>Violetta Simonacci</i>	249
3.3.3	Robust Regression for Compositional Data and its Application in the Context of SDG. <i>Valentin Todorov and Fatemah Alqallaf</i>	255

3.4	Evaluation of undercoverage for censuses and administrative data	261
3.4.1	Spatially balanced indirect sampling to estimate the coverage of the agricultural census. <i>Federica Piersimoni, Francesco Pantalone and Roberto Benedetti</i>	262
3.4.2	Next Census in Israel: Strategy, Estimation and Evaluation. <i>Danny Pfeffermann</i>	268
3.4.3	Administrative data for population counts estimations in Italian Population Census. <i>Antonella Bernadini, Angela Chieppa, Nicola Cibella and Fabrizio Solari</i>	274
3.4.4	LFS non response indicators for population register overcoverage estimation. <i>Lorella Di Consiglio, Stefano Falorsi</i>	279
3.5	Excesses and rare events in complex systems	285
3.5.1	Space-time extreme rainfall simulation under a geostatistical approach. <i>Gianmarco Callegher, Carlo Gaetan, Noemie Le Carrer and Ilaria Prosdocimi</i>	286
3.6	Hierarchical forecasting and forecast combination	292
3.6.1	Density calibration with consistent scoring functions. <i>Roberto Casarin and Francesco Ravazzolo</i>	293
3.6.2	Forecasting combination of hierarchical time series: a novel method with an application to CoVid-19. <i>Livio Fenga</i>	298
3.7	Household surveys for policy analysis	304
3.7.1	Did the policy responses to COVID-19 protect Italian households' incomes? Evidence from survey and administrative data. <i>Maria Teresa Monteduro, Dalila De Rosa and Chiara Subrizi</i>	305
3.8	Learning analytics methods and applications	311
3.8.1	Open-Source Automated Test Assembly: the Challenges of Large-Sized Models. <i>Giada Spaccapanico Proietti</i>	312
3.8.2	How Much Tutoring Activities May Improve Academic Careers of At-Risk Students? An Evaluation Study. <i>Marta Cannistra, Tommaso Agasisti, Anna Maria Paganoni and Chiara Masci</i>	318
3.8.3	Composite-based Segmentation Trees to Model Learners' performance. <i>Cristina Davino and Giuseppe Lamberti</i>	324
3.8.4	Test-taking Effort in INVALSI Assessments. <i>Chiara Sacco</i>	330
3.9	Light methods for hard problems	336
3.9.1	Fast Divide-and-Conquer Strategies to Solve Spatial Big Data Problems. <i>Michele Peruzzi</i>	337
3.9.2	Application of hierarchical matrices in spatial statistics. <i>Anastasiia Gorshechnikova and Carlo Gaetan</i>	343
3.10	Management and statistics in search for a common ground (AIDEA)	349
3.10.1	Customer Segmentation: it's time to make a change. <i>Fabrizio Laurini, Beatrice Luceri and Sabrina Latusi</i>	350
3.10.2	Multivariate prediction models: Altman's ZScore and CNDCEC's sectoral indicators. <i>Alessandro Danovi, Alberto Falini and Massimo Postiglione</i>	356
3.10.3	Comparing Entrepreneurship and Perceived Quality of Life in the European Smart Cities: a "Posetic" Approach. <i>Lara Penco, Enrico Ivaldi and Andrea Ciacci</i>	362

3.10.4	The Relationship between Business Economics and Statistics: Taking Stock and Ways Forward. <i>Amedeo Pugliese</i>	368
3.11	Mathematical methods and tools for finance and insurance (AMASES)	373
3.11.1	On the valuation of the initiation option in a GLWB variable annuity. <i>Anna Rita Bacinello and Pietro Millosovich</i>	374
3.11.2	Modern design of life annuities in view of longevity and pandemics. <i>Annamaria Olivieri</i>	380
3.11.3	Risk Management from Finance to Production Planning: An Assembly-to-Order Case Study. <i>Paolo Brandimarte, Edoardo Fadda and Alberto Gennaro</i>	386
3.11.4	Some probability distortion functions in behavioral portfolio selection. <i>Diana Barro, Marco Corazza and Martina Nardonthors</i>	392
3.12	Multiple system estimation	398
3.12.1	Multiple Systems Estimation in the Presence of Censored Cells. <i>Ruth King, Oscar Rodriguez de Rivera Ortega and Rachel McCrea</i>	399
3.12.2	Bayesian population size estimation by repeated identifications of units. A semi-parametric mixture model approach. <i>Tiziana Tuoto, Davide Di Cecco and Andrea Tancredi</i>	405
3.13	Network sampling and estimation	411
3.13.1	Targeted random walk sampling. <i>Li-Chun Zhang</i>	412
3.13.2	Estimation of poverty measures in Respondent-driven sampling. <i>María del Mar Rueda, Ismael Sánchez-Borrego and Héctor Mullo</i>	418
3.13.3	Sampling Networked Data for Semi-Supervised Learning Algorithms. <i>Simone Di Zio, Lara Fontanella, Francesco Pantalone and Federica Piersimoni</i>	423
3.13.4	A sequential adaptive sampling scheme for rare populations with a network structure. <i>Emilia Rocco</i>	429
3.14	New perspectives on multidimensional child poverty	435
3.14.1	Estimating uncertainty for child poverty indicators: The Case of Mediterranean Countries. <i>Ilaria Benedetti, Federico Crescenzi and Riccardo De Santis</i>	436
3.14.2	Child poverty and government social spending in the European Union during the economic crisis. <i>Angeles Sánchez and María Navarro</i>	442
3.14.3	The Children's Worlds Study: New perspectives on children's deprivation research. <i>Caterina Giusti and Antoanneta Potsi</i>	448
3.14.4	The impact of different definition of "households with children" on deprivation measures: the case of Italy. <i>Laura Neri and Francesca Gagliardi</i>	454
3.15	Perspectives in social network analysis applications	460
3.15.1	A comparison of student mobility flows in Erasmus and Erasmus+ among countries. <i>Kristijan Breznik, Giancarlo Ragozini and Marialuisa Restaino</i>	461
3.15.2	Network-based approach for the analysis of LexisNexis news database. <i>Carla Galluccio and Alessandra Petrucci</i>	467
3.15.3	A multiplex network approach to study Italian Students' Mobility. <i>Ilaria Primerano, Francesco Santelli and Cristian Usala</i>	473
3.15.4	Ego-centered Support Networks:a Cross-national European Comparison. <i>Emanuela Furfaro, Elvira Pelle, Giulia Rivellini and Susanna Zaccarin</i>	479

3.16	Statistical analysis of energy data	485
3.16.1	Machine learning models for electricity price forecasting. <i>Silvia Golia, Luigi Grossi, Matteo Pelagatti</i>	486
3.16.2	The impact of hydroelectric storage in the Italian power market. <i>Filippo Beltrami</i>	492
3.16.3	Jumps and cojumps in electricity price forecasting. <i>Peru Muniain, Aitor Ciarreta and Ainhoa Zarraga</i>	498
3.17	Statistical methods and models for the analysis of sports data	507
3.17.1	Football analytics: a Higher-Order PLS-SEM approach to evaluate players' performance. <i>Mattia Cefis and Maurizio Carpita</i>	508
3.17.2	Bayesian regularized regression of football tracking data through structured factor models. <i>Lorenzo Schiavon and Antonio Canale</i>	514
3.17.3	A dynamic matrix-variate model for clustering time series with multiple sources of variation. <i>Mattia Stival</i>	520
3.17.4	Evaluating football players' performances using on-the-ball data. <i>David Dandolo</i>	526
3.18	The social and demographic consequences of international migration in Western societies	532
3.18.1	Employment and job satisfaction of immigrants: the case of Campania (Italy). <i>Alessio Buonomo, Stefania Capecchi, Francesca Di Iorio and Salvatore Strozza</i>	533
3.18.2	Social stratification of migrants in Italy: class reproduction and social mobility from origin to destination. <i>Giorgio Piccitto, Maurizio Avola and Nazareno Panichella</i>	539
3.19	Well-being, healthcare, integration measurements and indicators (SIEDS)	545
3.19.1	A Composite Index of Economic Well-being for the European Union Countries. <i>Andrea Cutillo, Matteo Mazziotta and Adriano Pareto</i>	546
3.19.2	Poverty orderings and TIP curves: an application to the Italian regions. <i>Francesco M. Chelli, Mariateresa Ciommi and Chiara Gagliarano</i>	552
4	Contributed Sessions	558
4.1	Advances in clinical trials	559
4.1.1	Quantitative depth-based [18F]FMCH-avid lesion profiling in prostate cancer treatment. <i>Lara Cavinato, Alessandra Ragni, Francesca Ieva, Martina Sollini, Francesco Bartoli and Paola A. Erba</i>	560
4.1.2	Modelling longitudinal latent toxicity profiles evolution in osteosarcoma patients. <i>Marta Spreafico, Francesca Ieva and Marta Fiocco</i>	566
4.1.3	Information borrowing in phase II basket trials: a comparison of different designs. <i>Marco Novelli</i>	572
4.1.4	Q-learning Estimation Techniques for Dynamic Treatment Regime. <i>Simone Bogni, Debora Slanzi and Matteo Borrotti</i>	578
4.1.5	Sample Size Computation for Competing Risks Survival Data in GS-Design. <i>Mohammad Anamul Haque and Giuliana Cortese</i>	584

4.2	Advances in neural networks	590
4.2.1	Linear models vs Neural Network: predicting Italian SMEs default. <i>Lisa Crosato, Caterina Liberati and Marco Repetto</i>	591
4.2.2	Network estimation via elastic net penalty for heavy-tailed data. <i>Davide Bernardini, Sandra Paterlini and Emanuele Taufer</i>	596
4.2.3	Neural Network for statistical process control of a multiple stream process with an application to HVAC systems in passenger rail vehicles. <i>Gianluca Sposito, Antonio Lepore, Biagio Palumbo and Giuseppe Giannini</i>	602
4.2.4	Forecasting air quality by using ANNs. <i>Annalina Sarra, Adelia Evangelista, Tonio Di Battista and Francesco Bucci</i>	608
4.3	Advances in statistical methods	614
4.3.1	Robustness of Fractional Factorial Designs through Circuits. <i>Roberto Fontana and Fabio Rapallo</i>	615
4.3.2	Multi-objective optimal allocations for experimental studies with binary outcome. <i>Alessandro Baldi Antognini, Rosamarie Frieri, Marco Novelli and Maroussa Zagoraiou</i>	621
4.3.3	Analysis of three-way data: an extension of the STATIS method. <i>Laura Bocci and Donatella Vicari</i>	627
4.3.4	KL-optimum designs to discriminate models with different variance function. <i>Alessandro Lanteri, Samantha Leorato and Chiara Tommasi</i>	633
4.3.5	Riemannian optimization on the space of covariance matrices. <i>Jacopo Schiavon, Mauro Bernardi and Antonio Canale</i>	639
4.4	Advances in statistical methods and inference	645
4.4.1	Estimation of Dirichlet Distribution Parameters with Modified Score Functions. <i>Vincenzo Gioia and Euloge Clovis Kenne Pagui</i>	646
4.4.2	Confidence distributions for predictive tail probabilities. <i>Giovanni Fonseca, Federica Giummolè and Paolo Vidoni</i>	652
4.4.3	Impact of sample size on stochastic ordering tests: a simulation study. <i>Rosa Arboretti, Riccardo Ceccato, Luca Pegoraro and Luigi Salmaso</i>	658
4.4.4	On testing the significance of a mode. <i>Federico Ferraccioli and Giovanna Menardi</i>	664
4.4.5	Hommel BH: an adaptive Benjamini-Hochberg procedure using Hommel's estimator for the number of true hypotheses. <i>Chiara G. Magnani and Aldo Solari</i>	670
4.5	Advances in statistical models	676
4.5.1	Specification Curve Analysis: Visualising the risk of model misspecification in COVID-19 data. <i>Venera Tomaselli, Giulio Giacomo Cantone and Vincenzo Miracula</i>	677
4.5.2	Semiparametric Variational Inference for Bayesian Quantile Regression. <i>Cristian Castiglione and Mauro Bernardi</i>	683
4.5.3	Searching for a source of difference in undirected graphical models for count data – an empirical study. <i>Federico Agostinis, Monica Chiogna, Vera Djordjilovic, Luna Pianesi and Chiara Romualdi</i>	689
4.5.4	Snipped robust inference in mixed linear models. <i>Antonio Lucadamo, Luca Greco, Pietro Amenta and Anna Crisci</i>	695

4.6	Advances in time series	701
4.6.1	A spatio-temporal model for events on road networks: an application to ambulance interventions in Milan. <i>Andrea Gilardi and Riccardo Borgoni and Jorge Mateu</i>	702
4.6.2	Forecasting electricity demand of individual customers via additive stacking. <i>Christian Capezza, Biagio Palumbo, Yannig Goude, Simon N. Wood and Matteo Fasiolo</i>	708
4.6.3	Hierarchical Forecast Reconciliation on Italian Covid-19 data. <i>Andrea Marcocchia, Serena Arima and Pierpaolo Brutti</i>	714
4.6.4	Link between Threshold ARMA and tdARMA models. <i>Guy M�elard and Marcella Niglio</i>	720
4.7	Bayesian nonparametrics	726
4.7.1	Bayesian nonparametric prediction: from species to features. <i>Lorenzo Masoero, Federico Camerlenghi, Stefano Favaro and Tamara Broderick</i>	727
4.7.2	A framework for filtering in hidden Markov models with normalized random measures. <i>Filippo Ascolani, Antonio Lijoi, Igor Pr�unster and Matteo Ruggiero</i>	733
4.7.3	On the convex combination of a Dirichlet process with a diffuse probability measure. <i>Federico Camerlenghi, Riccardo Corradin and Andrea Ongaro</i>	739
4.7.4	Detection of neural activity in calcium imaging data via Bayesian mixture models. <i>Laura D'Angelo, Antonio Canale, Zhaoxia Yu and Michele Guindani</i>	745
4.8	Clustering for complex data	751
4.8.1	Clustering categorical data via Hamming distance. <i>Edoardo Filippi-Mazzola, Raffaele Argiento and Lucia Paci</i>	752
4.8.2	Penalized model-based clustering for three-way data structures. <i>Andrea Cappelozzo, Alessandro Casa, and Michael Fop</i>	758
4.8.3	Does Milan have a smart mobility? A clustering analysis approach. <i>Nicola Cornali, Matteo Seminati, Paolo Maranzano and Paola M. Chiodini</i>	764
4.8.4	A Fuzzy clustering approach for textual data. <i>Irene Cozzolino, Maria Brigida Ferraro and Peter Winker</i>	770
4.8.5	Valid Double-Dipping via Permutation-Based Closed Testing. <i>Anna Vesely, Livio Finos, Jelle J. Goeman and Angela Andreella</i>	776
4.9	Data science for complex data	782
4.9.1	Text mining on large corpora using Taltac4: An explorative analysis of the USPTO patents database. <i>Pasquale Pavone, Arianna Martinelli and Federico Tamagni</i>	783
4.9.2	Emotion pattern detection on facial videos using functional statistics. <i>Rongjiao Ji, Alessandra Micheletti, Natasa Krklec Jerinkic and Zoranka Desnica</i>	789
4.9.3	The spread of contagion on Twitter: identification of communities analysing data from the first wave of the COVID-19 epidemic. <i>Gianni Andreozzi, Salvatore Pirri, Giuseppe Turchetti and Valentina Lorenzoni</i>	795
4.9.4	Composition-on-Function Regression Model for the Remote Analysis of Near-Earth Asteroids. <i>Mara S. Bernardi, Matteo Fontana, Alessandra Menafoglio, Alessandro Pisello, Massimiliano Porreca, Diego Perugini and Simone Vantini</i>	801
4.9.5	Determinants of football coach dismissal in Italian League Serie A. <i>Francesco Porro, Marialuca Restaino, Juan Eloy Ruiz-Castro and Mariangela Zenga</i>	805
4.10	Data science for unstructured data	810
4.10.1	Identification and modeling of stop activities at the destination from GPS tracking data. <i>Nicoletta D'Angelo, Giada Adelfio, Antonino Abbruzzo and Mauro Ferrante</i>	811

4.10.2	A generalization of derangement. <i>Maurizio Maravalle and Ciro Marziliano</i>	817
4.10.3	Analysis of clickstream data with mixture hidden markov models. <i>Furio Urso, Antonino Abbruzzo and Maria Francesca Cracolici</i>	823
4.10.4	Using Google Scholar to measure the credibility of preprints in the COVID-19 Open Research Dataset (CORD-19). <i>Manlio Migliorati, Maurizio Carpita, Eugenio Brentari</i>	829
4.10.5	Mobile phone use while driving: a Structural Equation Model to analyze the Behavior behind the wheel. <i>Carlo Cavicchia and Pasquale Samacchiaro</i>	835
4.11	Demographic analysis	841
4.11.1	Life expectancy in the districts of Taranto. <i>Stefano Cervellera, Carlo Cusatelli and Massimiliano Giacalone</i>	842
4.11.2	Family size and Human Capital in Italy: a micro-territorial analysis. <i>Gabriele Ruii, Marco Breschi and Alessio Fornasin</i>	848
4.11.3	Estimate age-specific fertility rates from summary demographic measures. An Indirect Model Levering on Deep Neural Network. <i>Andrea Nigri</i>	854
4.11.4	Patterns in the relation between causes of death and gross domestic product. <i>Andrea Nigri and Federico Crescenzi</i>	860
4.11.5	Locally sparse functional regression with an application to mortality data. <i>Mauro Bernardi, Antonio Canale, Marco Stefanucci</i>	866
4.12	Environmental statistics	871
4.12.1	A Distribution-Free Approach for Detecting Radioxenon Anomalous Concentrations. <i>Michele Scagliarini, Rosanna Gualdi, Giuseppe Ottaviano, Antonietta Rizzo and Franca Padoani</i>	872
4.12.2	Ecosud Car, a novel approach for the predictive control of the territory. <i>Giacomo Iula, Massimo Dimo, Saverio Gianluca Crisafulli, Marco Vito Calciano, Vito Santarcangelo and Massimiliano Giacalone</i>	878
4.12.3	Effect of ties on the empirical copula methods for weather forecasting. <i>Elisa Perrone, Fabrizio Durante and Irene Schicker</i>	884
4.12.4	Spatio-temporal regression with differential penalization for the reconstruction of partially observed signals. <i>Eleonora Amone and Laura M. Sangalli</i>	890
4.12.5	Sea Surface Temperature Effects on the Mediterranean Marine Ecosystem: a Semiparametric Model Approach. <i>Claudio Rubino, Giacomo Milisenda, Antonino Abbruzzo, Giada Adelfio, Mar Bosch-Belmar, Francesco Colloca, Manfredi Di Lorenzo and Vita Gancitano</i>	895
4.13	Functional data analysis	901
4.13.1	Remote Analysis of Chapas Stops in Maputo from GPS data: a Functional Data Analysis Approach. <i>Agostino Torti, Davide Ranieri and Simone Vantini</i>	902
4.13.2	A Conformal approach for functional data prediction. <i>Jacopo Diquigiovanni, Matteo Fontana and Simone Vantini</i>	907
4.13.3	Block testing in covariance and precision matrices for functional data analysis. <i>Marie Morvan, Alessia Pini, Madison Giacomci and Valerie Monbet</i>	911
4.13.4	Analysing contributions of ages and causes of death to gender gap in life expectancy using functional data analysis. <i>Alessandro Feraldi, Virginia Zarulli, Stefano Mazzuco and Cristina Giudici</i>	917
4.13.5	Supervised classification of ECG curves via a combined use of functional data analysis and random forest to identify patients affected by heart disease. <i>Fabrizio Maturo and Rosanna Verde</i>	923

4.14	Mixture models	929
4.14.1	Alternative parameterizations for regression models with constrained multivariate responses. <i>Roberto Ascari, Agnese Maria Di Brisco, Sonia Migliorati and Andrea Ongaro</i>	930
4.14.2	Spatially dependent mixture models with a random number of components. <i>Matteo Gianella, Mario Beraha and Alessandra Guglielmi</i>	936
4.14.3	Finite mixtures of regression models for longitudinal data. <i>Marco Alfò and Roberto Rocci</i>	942
4.14.4	Mixtures of regressions for size estimation of heterogeneous populations. <i>Gianmarco Caruso</i>	948
4.14.5	Finite mixtures of regressions with random covariates using multivariate skewed distributions. <i>Salvatore D. Tomarchio, Michael P.B. Gallagher, Antonio Punzo and Paul D. McNicholas</i>	954
4.15	New applications of regression models	960
4.15.1	The Shapley-Lorenz decomposition approach to mitigate cyber risks. <i>Paolo Giudici and Emanuela Raffinetti</i>	960
4.15.2	A spatially adaptive estimator for the function-on-function linear regression model with application to the Swedish Mortality dataset. <i>Fabio Centofanti, Antonio Lepore, Alessandra Menafoglio, Biagio Palumbo and Simone Vantini</i>	967
4.15.3	POSetR: a new computationally efficient R package for partially ordered data. <i>Alberto Arcagni, Alessandro Avellone and Marco Fattore</i>	972
4.15.4	Multi Split Conformal Prediction. <i>Aldo Solari and Vera Djordjilović</i>	978
4.15.5	Changes in the consumption of fruits and vegetables among university students during master courses: an analysis of data automatically collected from cashier transactions. <i>Valentina Lorenzoni, Giuseppe Turchetti and Lucio Masserini</i>	984
4.16	New challenges in clustering and classification techniques	990
4.16.1	A Dynamic Stochastic Block Model with infinite communities. <i>Roberto Casarin and Ovielt Baltodano López</i>	991
4.16.2	Cross-Subject EEG Channel Selection for the Detection of Predisposition to Alcoholism. <i>Michela Carlotta Massi and Francesca Ieva</i>	997
4.16.3	Some Issues on the Parameter Selection in the Spectral Methods for Clustering. <i>Cinzia Di Nuzzo and Salvatore Ingrassia</i>	1003
4.16.4	The link-match tale: new microdata from unit level association. <i>Riccardo D'Alberto, Meri Raggi and Daniela Cocchi</i>	1009
4.17	New developments in Bayesian methods	1015
4.17.1	Spatio-temporal analysis of the Covid-19 spread in Italy by Bayesian hierarchical models. <i>Nicoletta D'Angelo, Giada Adelfio and Antonino Abbruzzo</i>	1016
4.17.2	Modelling of accumulation curves through Weibull survival functions. <i>Alessandro Zito, Tommaso Rigon and David B. Dunson</i>	1021
4.17.3	Model fitting and Bayesian inference via power expectation propagation. <i>Emanuele Degani, Luca Maestrini and Mauro Bernardi</i>	1026
4.17.4	Bayesian quantile estimation in deconvolution. <i>Catia Scricciolo</i>	1032
4.17.5	Bayesian inference for discretely observed non-homogeneous Markov processes. <i>Rosario Barone and Andrea Tancredi</i>	1038

4.18	New developments in composite indicators applications	1044
4.18.1	Building composite indicators in the functional domain: a suggestion for an evolutionary HDI. <i>Francesca Fortuna, Alessia Naccarato and Silvia Terzi</i>	1045
4.18.2	Small Area Estimation of Inequality Measures via Simplex Regression. <i>Silvia De Nicolò, Maria Rosaria Ferrante and Silvia Pacei</i>	1051
4.18.3	Relational Well-Being and Poverty in Italy Benessere relazionale e povertà in Italia. <i>Elena Dalla Chiara and Federico Perali</i>	1057
4.18.4	A composite indicator to assess sustainability of agriculture in European Union countries. <i>Alessandro Magrini and Francesca Giambona</i>	1063
4.18.5	Interval-Based Composite Indicators with a Triplex Representation: A Measure of the Potential Demand for the “Ristori” Decree in Italy. <i>Carlo Drago</i>	1069
4.19	New developments in GLM theory and applications	1075
4.19.1	Variational inference for the smoothing distribution in dynamic probit models. <i>Augusto Fasano and Giovanni Rebaudo</i>	1076
4.19.2	Interpretability and interaction learning for logistic regression models. <i>Nicola Rares Franco, Michela Carliotta Massi, Francesca Ieva and Anna Maria Paganoni</i>	1082
4.19.3	Entropy estimation for binary data with dependence structures. <i>Linda Altieri and Daniela Cocchi</i>	1088
4.19.4	A Comparison of Some Estimation Methods for the Three-Parameter Logistic Model. <i>Michela Battauz and Ruggero Bellio</i>	1094
4.19.5	A statistical model to identify the price determinations: the case of Airbnb. <i>Giulia Contu, Luca Frigau, Gian Paolo Zammarchi and Francesco Mola</i>	1100
4.20	New developments in social statistics analysis	1106
4.20.1	Data-based Evaluation of Political Agents Against Goals Scheduling. <i>Giulio D'Epifanio</i>	1107
4.20.2	Local heterogeneities in population growth and decline. A spatial analysis for Italian municipalities. <i>Federico Benassi, Annalisa Busetta, Gerardo Gallo and Manuela Stranges</i>	1113
4.20.3	The assessment of environmental and income inequalities. <i>Michele Costa</i>	1119
4.20.4	Household financial fragility across Europe. <i>Marianna Brunetti, Elena Giarda and Costanza Torricelli</i>	1125
4.20.5	Refugees' perception of their new life in Germany. <i>Daria Mendola and Anna Maria Parroco</i>	1131
4.21	New perspectives in clinical trials	1137
4.21.1	Improved maximum likelihood estimator in relative risk regression. <i>Euloge C. Kenne Pagui, Francesco Pozza and Alessandra Salvan</i>	1138
4.21.2	Development and validation of a clinical risk score to predict the risk of SARS-CoV-2 infection. <i>Laura Savaré, Valentina Orlando and Giovanni Corrao</i>	1144
4.21.3	Functional representation of potassium trajectories for dynamic monitoring of Heart Failure patients. <i>Caterina Gregorio, Giulia Barbatì and Francesca Ieva</i>	1150
4.21.4	Effect of lung transplantation on the survival of patients with cystic fibrosis: IMaCh contribution to registry data. <i>Cristina Giudici, Nicolas Brouard and Gil Bellis</i>	1156
4.21.5	Categories and Clusters to investigate Similarities in Diabetic Kidney Disease Patients. <i>Veronica Distefano, Maria Mannone, Claudio Silvestri and Irene Poli</i>	1162

4.22	New perspectives in models for multivariate dependency	1168
4.22.1	Parsimonious modelling of spectroscopy data via a Bayesian latent variables approach. <i>Alessandro Casa, Tom F. O'Callaghan and Thomas Brendan Mur</i>	1169
4.22.2	Bias reduction in the equicorrelated multivariate normal. <i>Elena Bortolato and Euloge Clovis Kenne Pagui</i>	1175
4.22.3	Some results on identifiable parameters that cannot be identified from data. <i>Christian Hennig</i>	1181
4.23	Novel approaches for official statistics	1187
4.23.1	Web data collection: profiles of respondents to the Italian Population Census. <i>Elena Grimaccia, Gerardo Gallo, Alessia Naccarato, Novella Cecconi and Alessandro Fratoni</i>	1188
4.23.2	Trusted Smart Surveys: architectural and methodological challenges at a glance. <i>Mauro Bruno, Francesca Inglese and Giuseppina Ruocco</i>	1194
4.23.3	On bias correction in small area estimation: An M-quantile approach. <i>Gaia Bertarelli, Francesco Schirripa Spagnolo, Raymond Chambers and David Haziza</i>	1200
4.23.4	The address component of the Statistical Base Register of Territorial Entities. <i>Davide Fardelli, Enrico Orsini and Andrea Pagano</i>	1206
4.23.5	A well-being municipal indicator using census data: first results. <i>Massimo Esposito</i>	1212
4.24	Prior distribution for Bayesian analysis	1218
4.24.1	On the dependence structure in Bayesian nonparametric priors. <i>Filippo Ascolani, Beatrice Franzolini, Antonio Lijoi, and Igor Prünster</i>	1219
4.24.2	Anisotropic determinantal point processes and their application in Bayesian mixtures. <i>Lorenzo Ghilotti, Mario Beraha and Alessandra Guglielmi</i>	1226
4.24.3	Bayesian Screening of Covariates in Linear Regression Models Using Correlation Thresholds. <i>Ioannis Ntzoufras and Roberta Paroli</i>	1232
4.25	Recent advances in clustering methods	1238
4.25.1	Biclustering longitudinal trajectories through a model-based approach. <i>Francesca Martella, Marco Alfò and Maria Francesca Marino</i>	1239
4.25.2	Monitoring tools for robust estimation of Cluster Weighted models. <i>Andrea Cappozzo and Francesca Greselin</i>	1245
4.25.3	Co-clustering Models for Spatial Transcriptomics: Analysis of a Human Brain Tissue Sample. <i>Andrea Sottosanti and Davide Risso</i>	1251
4.25.4	Graph nodes clustering: a comparison between algorithms. <i>Ilaria Bombelli</i>	1257
4.26	Social demography	1263
4.26.1	Childcare among migrants: a comparison between Italy and France. <i>Eleonora Trappolini, Elisa Barbiano di Belgiojoso, Stefania Maria Lorenza Rimoldi and Laura Terzera</i>	1264
4.26.2	Employment Uncertainty and Fertility in Italy: The Role of Union Formation. <i>Giammarco Alderotti, Valentina Tocchioni and Alessandra De Rose</i>	1270
4.26.3	Determinants of union dissolution in Italy: Do children matter? <i>Valentina Tocchioni, Daniele Vignoli, Eleonora Meli and Bruno Arpino</i>	1276
4.26.4	Working schedules and fathers' time with children: A Sequence Analysis. <i>Annalisa Donno and Maria Letizia Tanturri</i>	1282
4.26.5	Correlates of the non-use of contraception among female university students in Italy. <i>Annalisa Busetta, Alessandra De Rose and Daniele Vignoli</i>	1288

4.27	Social indicators applications and methods	1294
4.27.1	A logistic regression model for predicting child language performance. <i>Andrea Briglia, Massimo Mucciardi and Giovanni Pirrotta</i>	1295
4.27.2	Subject-specific measures of interrater agreement for ordinal scales. <i>Giuseppe Bove</i>	1301
4.27.3	A Tucker3 method application on adjusted-PMRs for the study of work-related mortality. <i>Vittoria Carolina Malpassuti, Vittoria La Serra and Stefania Massari</i>	1307
4.27.4	Two case-mix adjusted indices for nursing home performance evaluation. <i>Giorgio E. Montanari and Marco Doretti</i>	1313
4.27.5	The ultrametric covariance model for modelling teachers' job satisfaction. <i>Carlo Cavicchia, Maurizio Vichi and Giorgia Zaccaria</i>	1319
4.28	Some recent developments in compositional data analysis	1325
4.28.1	A Robust Approach to Microbiome-Based Classification Problems. <i>Gianna Serafina Monti and Peter Filzmoser</i>	1326
4.28.2	What is a convex set in compositional data analysis? <i>Jordi Saperas i Riera, Josep Antoni Martín Fernández</i>	1332
4.28.3	Compositional Analysis on the Functional Distribution of Extended Income. <i>Elena Dalla Chiara and Federico Perali</i>	1338
4.28.4	Evaluating seasonal-induced changes in river chemistry using Principal Balances. <i>Caterina Gozzi and Antonella Buccianti</i>	1344
4.28.5	Compositional Data Techniques for the Analysis of the Ragweed Allergy. <i>Gianna S. Monti, Maira Bonini, Valentina Ceriotti, Matteo Pelagatti and Claudio M. Ortolani</i>	1350
4.29	Spatial data analysis	1356
4.29.1	Spatial multilevel mixed effects modeling for earthquake insurance losses in New Zealand. <i>F. Marta L. Di Lascio and Selene Perazzini</i>	1357
4.29.2	Weighted distances for spatially dependent functional data. <i>Andrea Diana, Elvira Romano, Claire Miller and Ruth O'Donnell</i>	1363
4.29.3	Spatial modeling of childcare services in Lombardia. <i>Emanuele Aliverti, Stefano Campostrini, Federico Caldura and Lucia Zanotto</i>	1369
4.29.4	On the use of a composite attractiveness index for the development of sustainable tourist routes. <i>Claudia Cappello, Sandra De Iaco, Sabrina Maggio and Monica Palma</i>	1375
4.30	Statistical applications in education	1381
4.30.1	Does self-efficacy influence academic results? A separable-effect mediation analysis. <i>Chiara Di Maria</i>	1382
4.30.2	Statistics Knowledge assessment: an archetypal analysis approach. <i>Bruno Adabbo, Rosa Fabbricatore, Alfonso Iodice D'Enza and Francesco Palumbo</i>	1388
4.30.3	Exploring drivers for Italian university students' mobility: first evidence from AlmaLaurea data. <i>Giovanni Boscaino and Vincenzo Giuseppe Genova</i>	1394
4.30.4	Can Grading Policies influence the competition among Universities of different sizes? <i>Gabriele Lombardi and Antonio Pio Distaso</i>	1400
4.30.5	The class A journals and the Italian academic research outcomes in Statistical Sciences. <i>Maria Maddalena Barbieri, Francesca Bassi, Antonio Irpino and Rosanna Verde</i>	1406
4.31	Statistical methods for finance	1412
4.31.1	Hypotheses testing in mixed-frequency volatility models: a bootstrap approach. <i>Vincenzo Candila and Lea Petrella</i>	1413

4.31.2	Quantile Regression Forest with mixed frequency Data. <i>Mila Andreani, Vincenzo Candila and Lea Petrella</i>	1419
4.31.3	Higher order moments in Capital Asset Pricing Model betas. <i>Giuseppe Arbia, Riccardo Bramante and Silvia Facchinetti</i>	1425
4.31.4	When Does Sentiment Matter in Predicting Cryptocurrency Bubbles? <i>Arianna Agosto and Paolo Pagnottoni</i>	1431
4.32	Statistical methods for high dimensional data	1437
4.32.1	Virtual biopsy in action: a radiomic-based model for CALI prediction. <i>Francesca Ieva, Giulia Baroni, Lara Cavinato, Chiara Masci, Guido Costa, Francesco Fiz, Arturo Chiti and Luca Viganò</i>	1438
4.32.2	Functional alignment by the “light” approach of the von Mises-Fisher-Procrustes model. <i>Angela Andreella and Livio Finos</i>	1444
4.32.3	A screening procedure for high-dimensional autologistic models. <i>Rodolfo Metulini and Francesco Giordano</i>	1450
4.32.4	Covariate adjusted censored gaussian lasso estimator. <i>Luigi Augugliaro, Gianluca Sottile and Veronica Vinciotti</i>	1456
4.32.5	Ranking-Based Variable Selection for ultra-high dimensional data in GLM framework. <i>Francesco Giordano, Marcella Niglio and Marialuisa Restaino</i>	1462
4.33	Statistical methods in higher education	1468
4.33.1	Effects of remote teaching on students' motivation and engagement: the case of the University of Modena & Reggio Emilia. <i>Isabella Morlini and Laura Sartori</i>	1469
4.33.2	A random effects model for the impact of remote teaching on university students' performance. <i>Silvia Bacci, Bruno Bertaccini, Simone Del Sarto, Leonardo Grilli and Carla Rampichini</i>	1475
4.33.3	Multinomial semiparametric mixed-effects model for profiling engineering university students. <i>Chiara Masci, Francesca Ieva and Anna Maria Paganoni</i>	1481
4.33.4	Evaluating Italian universities: ANVUR periodic accreditation judgment versus international rankings. <i>Angela Maria D'Uggento, Nunziata Ribecco and Vito Ricci</i>	1487
4.33.5	Women's career discrimination in the Italian Academia in the last 20. <i>Daniele Cuntrera, Vincenzo Falco and Massimo Attanasio</i>	1493
4.34	Statistical methods with Bayesian networks	1499
4.34.1	Statistical Micro Matching Using Bayesian Networks. <i>Pier Luigi Conti, Daniela Marella, Paola Vicard and Vincenzina Vitale</i>	1500
4.34.2	Modeling school managers challenges in the pandemic era with Bayesian networks. <i>Maria Chiara De Angelis and Flaminia Musella and Paola Vicard</i>	1506
4.34.3	Structural learning of mixed directed acyclic graphs: a copula-based approach. <i>Federico Castelletti</i>	1512
4.34.4	Inference on Markov chains parameters via Large Deviations ABC. <i>Cecilia Viscardi, Fabio Corradi, Michele Boreale and Antonietta Mira</i>	1518
4.34.5	A propensity score approach for treatment evaluation based on Bayesian Networks. <i>Federica Cugnata, Paola M.V. Rancoita, Pier Luigi Conti, Alberto Briganti, Clelia Di Serio, Fulvia Mecatti and Paola Vicard</i>	1524
4.35	Statistical modelling for the analysis of contemporary societies	1530
4.35.1	Social Network Analysis to analyse the relationship between 'victim-author' and 'motivation' of violence against women in Italy. <i>Alessia Forciniti</i>	1531
4.35.2	Satisfaction and sustainability propensity among elderly bike-sharing users. <i>Paolo Maranzano, Roberto Ascari, Paola Maddalena Chiodini and Giancarlo Manzi</i>	1537

4.35.3	Media and Investors' Attention. Estimating analysts' ratings and sentiment of a financial column to predict abnormal returns. <i>Riccardo Ferretti and Andrea Sciandra</i>	1543
4.35.4	Predictions of regional HCE: spatial and time patterns in an ageing population framework. <i>Laura Rizzi, Luca Grassetti, Divya Brundavanam, Alvisa Palese and Alessio Fornasin</i>	1549
4.36	Surveillance methods and statistical models in the Covid-19 crisis	1555
4.36.1	The Italian Social Mood on Economy Index during the Covid-19 Crisis. <i>Alessandra Righi and Diego Zardetto</i>	1556
4.36.2	Modeling the first wave of the COVID-19 pandemic in the Lombardy region, Italy, by using the daily number of swabs. <i>Claudia Furlan and Cinzia Mortarino</i>	1562
4.36.3	Analysing the Covid-19 pandemic in Italy with the SIPRO model. <i>Martina Amongero, Enrico Bibbona and Gianluca Mastrantonio</i>	1568
4.36.4	Intentions of union formation and dissolution during the COVID-19 pandemic. <i>Bruno Arpino and Daniela Bellani</i>	1574
4.37	Time series methods	1580
4.37.1	Bootstrap-based score test for INAR effect. <i>Riccardo Ievoli and Lucio Palazzo</i>	1581
4.37.2	Evaluating the performance of a new picking algorithm based on the variance piecewise constant models. <i>Nicoletta D'Angelo, Giada Adelfio, Antonino D'Alessandro and Marcello Chiodi</i>	1587
4.37.3	Conditional moments based time series cluster analysis. <i>Raffaele Mattera and Germana Scepi</i>	1593
4.37.4	On the asymptotic mean-squared prediction error for multivariate time series. <i>Gery Andrés Díaz Rubio, Simone Giannerini, and Greta Goracci</i>	1599
4.37.5	Spherical autoregressive change-point detection with applications. <i>Federica Spoto, Alessia Caponera and Pierpaolo Brutti</i>	1605
5	Posters	1611
5.1	A method for incorporating historical information in non-inferiority trials. <i>Fulvio De Santis and Stefania Gubbiotti</i>	1612
5.2	Optimal credible intervals under alternative loss functions. <i>Fulvio De Santis and Stefania Gubbiotti</i>	1618
5.3	Statistical learning for credit risk modelling. <i>Veronica Bacino, Alessio Zoccarato, Caterina Liberati and Matteo Borrotti</i>	1624
5.4	Evaluating heterogeneity of agreement with strong prior information. <i>Federico M. Stefanini</i>	1630
5.5	Analysis of the spatial interdependence of the size of endoreduplicated nuclei observed in confocal microscopy. <i>Ivan Sciascia, Andrea Crosino, Gennaro Carotenuto and Andrea Genre</i>	1636
5.6	A Density-Peak Approach to Clustering Graph-Structured Data. <i>Riccardo Giubilei</i>	1642
5.7	The employment situation of people with disabilities in Tuscany, A Survey on the workplace. <i>Paolo Addis, Alessandra Coli and Gianfranco Francese</i>	1648
5.8	Robustness of statistical methods for modeling paired count data using bivariate discrete distributions with general dependence structures. <i>Marta Nai Ruscone and Dimitris Karlis</i>	1654

6	Satellite events	1660
6.1	Measuring uncertainty in key official economic statistics	1661
6.1.1	Uncertainty in production and communication of statistics: challenges in the new data ecosystem. <i>Giorgio Alleva and Piero Demetrio Falorsi</i>	1662
6.1.2	Uncertainty and variance estimation techniques for poverty and inequality measures from complex surveys: a simulation study. <i>Riccardo De Santis, Lucio Barabesi and Gianni Betti</i>	1668
6.1.3	Pandemics and uncertainty in business cycle analysis. <i>Jacques Anas, Monica Billio, Leonardo Carati, Gian Luigi Mazzi and Hionia Vlachou</i>	1674
6.2	Covid-19: the urgent call for a unified statistical and demographic challenge	1680
6.2.1	Environmental epidemiology and the Covid-19 pandemics	1681
6.2.1.1	The Covid-19 outbreaks and their environment: The Valencian human behaviour. <i>Xavier Barber, Elisa Espín, Lucía Guevara, Aurora Mula, Kristina Polotskaya and Alejandro Rabasa</i>	1682
6.2.2	Estimation of Covid 19 prevalence	1686
6.2.2.1	Optimal spatial sampling for estimating the SARS-Cov-2 crucial parameters. <i>Piero Demetrio Falorsi and Vincenzo Nardelli</i>	1687
6.2.2.2	Survey aimed to estimate the seroprevalence of SARS-CoV-2 infection in Italian population at national and regional level. <i>Stefano Falorsi, Michele D'Alò, Claudia De Vitiis, Andrea Fasulo, Danila Filippini, Alessio Guandalini, Francesca Inglese, Orietta Luzi, Enrico Orsini and Roberta Radini</i>	1693
6.2.3	Measuring and modeling inequalities following the Covid-19 crisis	1699
6.2.3.1	COVID-19 impacts on young people's life courses: first results. <i>Antonietta Bisceglia, Concetta Scolorato and Giancarlo Ragozini</i>	1700
6.2.3.2	Exploring Students' Profile and Performance Before and After Covid-19 Lock-down. <i>Cristina Davino and Marco Gherghi</i>	1705
6.2.4	Nowcasting the Covid-19 outbreaks methods and applications	1711
6.2.4.1	Modeling subsequent waves of COVID-19 outbreak: A change point growth model. <i>Luca Greco, Paolo Girardi and Laura Ventura</i>	1712
6.2.4.2	The second wave of SARS-CoV-2 epidemic in Italy through a SIRD model. <i>Michela Baccini and Giulia Cereda</i>	1718
6.2.5	The impact of Covid-19 on survey methods	1724
6.2.5.1	Collecting cross-national survey data during the COVID-19 pandemic: Challenges and implications of data collection for the 50+ population in the Survey of Health, Ageing and Retirement in Europe (SHARE). <i>Michael Bergmann, Arne Bethmann, Yuri Pettinicchi and Borsch-Supan</i>	1725
6.2.5.2	Adapting a Long-Term Panel Survey to Pandemic Conditions. <i>Peter Lynn</i>	1731
6.2.6	Young contributions in Covid-19 statistical modelling	1737
6.2.6.1	Statistical communication of COVID-19 epidemic using widely accessible interactive tools. <i>Marco Mingione and Pierfrancesco Alaimo Di Loro</i>	1738
6.2.6.2	Modelling COVID-19 evolution in Italy with an augmented SIRD model using open data. <i>Vincenzo Nardelli, Giuseppe Arbia, Andrea Palladino and Luigi Giuseppe Atzeni</i>	1744

Effect of lung transplantation on the survival of patients with cystic fibrosis: IMaCh contribution to registry data

Trapianto di polmone e sopravvivenza in fibrosi cistica: il contributo di IMaCh all'analisi di dati di registro

Cristina Giudici, Nicolas Brouard and Gil Bellis

Abstract Using the Interpolated Markov Chain (IMaCh) software, we analyse transitions from different degrees of pulmonary function and mortality with and without transplantation, and compute life expectancy for cystic fibrosis patients, starting from different level of their respiratory health. We use data from the French Cystic Fibrosis Registry. The period of the study is 2008–2013 (7112 patients). Individuals enter the analysis in different years and their health status is regularly monitored. Health refers to the *volume* of air that can forcibly be blown out in first 1 second, after full inspiration (FEV₁). Globally, we found higher mortality for those who have been transplanted but also lower probability of transition towards critical respiratory functions.

Abstract Lo studio analizza i dati del registro francese della fibrosi cistica utilizzando il software IMaCh - Interpolated Markov Chain. Vengono stimate le transizioni di salute e mortalità in presenza/assenza di trapianto di polmone e calcolata la speranza di vita a partire da diversi livelli di funzionalità polmonare. Quest'ultima è misurata attraverso il volume di aria espirata nel corso del primo secondo di una espirazione massima forzata (FEV₁). Lo studio è svolto su 7112 pazienti presenti nel Registro tra il 2008 e il 2013. I pazienti che hanno subito un trapianto di polmone mostrano una mortalità più alta, ma anche una minore probabilità di transizione verso livelli critici di funzionalità polmonare.

Key words: Interpolated Markov Chain (IMaCh), Registry data, Cystic fibrosis

¹ Cristina Giudici, Sapienza University of Rome; email: cristina.giudici@uniroma1.it
Nicolas Brouard, Institut National d'Etudes Démographiques (INED); email: brouard@ined.fr
Gil Bellis, Institut National d'Etudes Démographiques (INED); email: bellis@ined.fr

1 Background

Cystic fibrosis (CF) is a multiorgan genetic disease that affects primarily the lungs and often leads to progressive respiratory insufficiency and premature death. CF is the most common hereditary disease among children in Europe (12,3 per 100.000). Median age of survival with CF has increased in several countries, and nowadays a steadily growing number of patients are adults.

Actually, among CF patients, morbidity and mortality is mostly caused by bronchiectasis, small airways obstruction, and progressive respiratory impairment (Nkam (2017)).

In the last decades, remarkable improvements in quality of life and clinical outcomes in patients with cystic fibrosis have been achieved thanks to innovative therapies. Although, for those patients who failed to respond to standard therapy lung transplantation remains the only treatment option with the potential to ameliorate symptoms, preserve quality of life, and extend life (Thabut et al. (2013)). In particular, bilateral lung transplantation has been shown to be an important therapeutic option for end-stage CF pulmonary disease (Hirche (2014)).

Several statistical models have been developed to identify prognostic factors in CF patients. One of the most significant predictors for survival is the Forced Expiratory Volume in 1 second (FEV1). Patients with low FEV1 may be referred for lung transplantation with the aim of improving their life expectancy and their quality of life (Nkam (2017)). Although, predicting life and health expectancy with or without transplantation is still a major issue.

In France, a national cystic fibrosis Registry (Registre français de la mucoviscidose) was created in 1992, and managed by the Institut National d'Etudes Démographiques (INED) since 1998. In October 2001 the Ministry of Health introduced systematic neonatal screening for cystic fibrosis on a national scale, and universal screening of newborn babies was introduced in 2003.

A number of studies have been carried out on this kind of data using classical statistical model (Nkam (2017)). Using the latest version of Interpolated Markov Chain (IMaCh) approach we analyse transitions from different degrees of pulmonary function and mortality and we compute life expectancy at different ages with and without transplantation.

2 Data: the French Cystic Fibrosis Register

Data are collected via questionnaires sent once a year to the healthcare centres cooperating with the Registry, in mainland France and Réunion Island. Individuals enter the Registry in different years and patient health status is regularly monitored. For each annual survey, each participating center reports on the patients seen at least once in the year. In particular, patients may visit their health center either a few times each year, or only once, or not every year according to their state of health. Globally,

the number of patients treated in the healthcare centres cooperating with the Registry increased from 2,168 in 1994 to 6,408 in 2013. Actually the Registry contains longitudinal data on more than 8,000 patients, which represents approximately 90% of all CF patients in France (Bellis et al. (2015)).

Among other analysis, the Forced Expiratory Volume in 1 second, after full inspiration (FEV_1) is measured in patients aged 6 years old or over. In case of multiple measures during the same year, only the best value of FEV_1 is recorded. This measure is considered good when it is greater than 80% of the predicted value. On the contrary, we consider that patients having a FEV_1 lower than 40% of the predicted value may be referred for lung transplantation, with the aim of improving their life expectancy and their quality of life.

The data examined in this article concern the deaths of patients included in the Registry database from 2008 to 2013. The database provides up to 6 measures for each individual, one for each year. A total of 7,112 patients were registered between 2008 and 2013. As 1,308 patients do not have any measure of the FEV_1 (they are mostly children under the age of 6), only 5,804 patients have been considered for the analysis. During the period, the median age of patients increased, the transplantation rate increased as well and the death rate decreased (see table 1). In 2013 a total of 668 lung transplanted patients was still alive.

Table 1: *Characteristic of the Register Population*

	2008	2009	2010	2011	2012	2013
Patients (N)	5,419	5,700	5,685	6,077	6,277	6,408
Median age of patients (years)	17.2	18.2	18.7	19.2	19.7	20.3
Max age of patients (years)	75	77	80	87	88	87
Adults ≥ 18 (%)	44.8	46.2	47.9	48.7	49.8	51.0
Mean FEV_1	71.0	71.0	71.8	75.9	77.0	77.7
std FEV_1	26.9	26.7	26.3	25.8	25.5	25.1
Death (N)	56	64	56	67	51	45
Mortality rate (x1000)	10.3	11.2	9.9	11.0	8.1	7.0
Median age at death (years)	29.7	25.2	29.1	26.5	32.3	34.6
Transplantations (N)	61	70	72	96	94	94
Transplantations (x1000)	11.3	12.3	12.7	15.8	15.0	15.0
Age at transplantation (year)	26.2	26.1	27.6	27.3	29.4	27.6

3 Methods

We estimated the age-specific flows of entry into and exit from critical respiratory functions, and the matrix of the transition probabilities between good (coded 1) and

severe (coded 2) FEV₁ and death (coded 3), using version 0.99 (r19) of the software IMaCh (Interpolation of Markov Chains)¹.

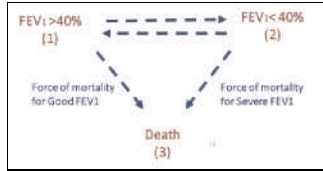


Figure 1: States and possible transitions among states

The probability for an individual aged x , observed in the health state i to find him/herself in state j after one year is indicated by p_{ij}^x and the transition probabilities are estimated based on a series of 3×3 matrices:

$$P_{ij}^x = \begin{pmatrix} p_{11}^x & p_{12}^x & p_{13}^x \\ p_{21}^x & p_{22}^x & p_{23}^x \\ 0 & 0 & 1 \end{pmatrix}$$

The first and the second rows represent transitions for individuals who begin the interval respectively in good and severe FEV₁. The third row represents the absorbing state of death. The probabilities of transition are then parameterized using the following multinomial logit model:

$$\ln \frac{p_{ij}^x}{p_{ii}^x} = \alpha_{ij} + \beta_{ij}x + \gamma_{ij}z$$

Where z is a time varying covariate which could influence the transition from one state to another or to death. This covariate corresponds to the answer to the question: “Was the patient already transplanted (1) at the time of each FEV₁ measure or not (0)?”.

A complicating factor was that, if patients have had the opportunity to be transplanted in year t , the FEV₁(t) might be measured before or after transplantation, but this information was not recorded. Moreover, some patients might have no measure at all during the year of transplantation.

¹ IMaCh is a publicly available computer program introduced by Lièvre, Brouard and Heatcote (2003) and mostly used for the estimation of Health Expectancy from longitudinal surveys. It allows to estimate transition probabilities by the method of maximum likelihood, using a discrete time embedded Markov chain approach. Transitions are supposed to occur at any time and death is always an additional competing risk. See Brouard (2019) for theory and applications. Several applications can also be found in literature, see for example Molla and Madans (2008) and Giudici et al. (2013).

To overcome this problem, we set up a decision rule based on the expected increase in the ventilation capacity just after the surgery.

4 Results

Figure 2 shows the transition probabilities from different initial state of respiratory function for non-lung transplanted ($z = 0$) and lung transplanted ($z = 1$) patients.

As expected, the probability of dying is always higher among those with severe FEV₁ (p_{23} vs p_{13}) and for transplanted patients ($z=1$ vs $z=0$), and the probability of recovering (p_{21}) is decreasing with age.

The transplantation modifies mainly the transition rate towards severe FEV₁ (p_{12}) which is lower for those who have been transplanted.

On the basis of transition probabilities estimates, IMaCh computes life expectancy for patients in state 1 and 2 by age, given that they were in that state initially. Globally, total life expectancy at birth for the analysed population is 54 years and, at any age, patients may expect to live 11 years in critical respiratory health. Life expectancy for lung transplanted patients is almost totally free from severe respiratory distress. At the age of 27 (median age of transplantation) they may expect to live 21 years (Figure 3).

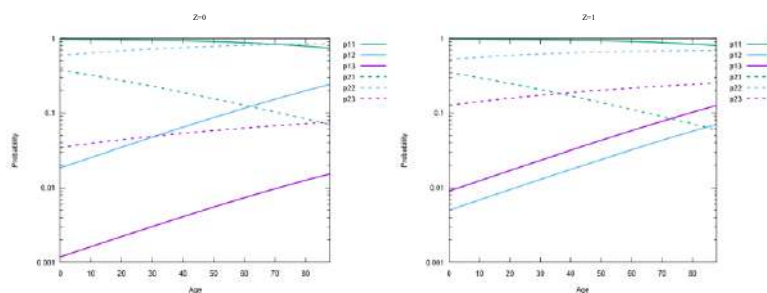


Figure 2: Conditional probabilities to be observed in state j being in state i , 12 months before ($z = 0$) and after ($z = 1$) lung transplantation

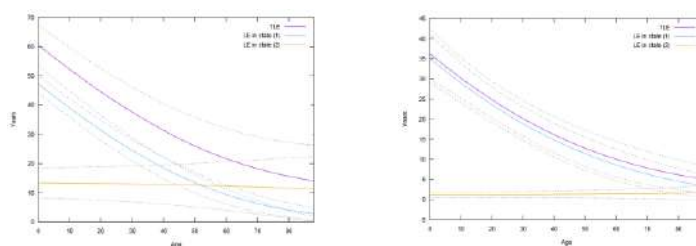


Figure 3: Life expectancies according to the state of health for non-transplanted (on the left) and transplanted (on the right) patients

5 Conclusions

The aim of our analysis was to compute life and health expectancy for CF patients with and without lung transplantation, starting from Registry data and using the latest version of IMaCh software.

For a given year, registry data do not necessarily provide the exact date at which respiratory health is measured and neither the date of the surgery in case of transplantation. To overcome this lack of information we set up a decision rule based on the expected increase in the ventilation capacity just after the surgery, and create a time dependent covariate which indicates, for each year, if the patient was already transplanted at the time of the respiratory measure or not.

Globally, we found higher mortality for those who have been transplanted but also lower probability of transition towards critical respiratory functions.

Our approach allows a more complex life expectancy analysis than classical approach on median age at death and contribute to exploit CF registry data. Nonetheless, the study is not without limitation. Indeed, CF prognostic factor are very complex. In particular, female gender is recognized in several studies as having a negative impact on survival. Moreover, in addition to compromised respiratory function, mortality in CF may be also related to liver complications or compromised nutritional status (Buzzetti et al. (2008)). Finally, we did not distinguish single lung and double or heart-lung transplantation.

To conclude, the French Registry of Cystic Fibrosis provides an excellent basis for measuring and monitoring mortality and survival of CF patient in France and our study suggests a useful approach to the registry data analysis.

References

1. Bellis, G., Dehillotte C., Lemonnier L.: *Registre Français De La Mucoviscidose. Bilan Des Données 2014. Vaincre la Mucoviscidose et Institut national d'études démographiques, INED, Paris (2016)*
2. Brouard N.: *Theory and Applications of Backward Probabilities and Prevalences in Cross-Longitudinal Surveys*. In Rao, A.S.R.S., Rao, C.R. (eds). *Handbook of Statistics*, Elsevier, Vol. 40, pp. 435-486 (2019) doi 10.1016/bs.host.2018.11.009
3. Buzzetti R. et al.: *An overview of international literature from cystic fibrosis registries:1. Mortality and survival studies in cystic fibrosis*. *J. Cyst. Fibros.* 8, 229-237 (2009)
4. Giudici, C., Arezzo, M. F., & Brouard, N.: *Estimating health expectancy in presence of missing data: an application using HID survey*. *Stat. Methods Appt.*, 22(4), 517-534 (2013)
5. Hirche, T.O. et al.: *Practical guidelines: lung transplantation in patients with cystic fibrosis*. *Pulm. Med.* (2014)
6. Lievre, A., Brouard, N., & Heathcote, C.: *The estimation of health expectancies from cross-longitudinal surveys*. *Math. Popul. Stud.* 10(4), 211-248 (2003)
7. Molla, M. T., & Madans, J. H.: *Estimating healthy life expectancies using longitudinal survey data: methods and techniques in population health measures*. *Vital Health Stat.*2 (146), 1-24 (2008)
8. Nkam, L., Lambert, J., Latouche A., Bellis G., Burgel P.R., Hocine M.N.: *A 3-year prognostic score for adults with cystic fibrosis*. *J. Cyst. Fibros.* 16 (2017)
9. Thabut, G., Christie J.D., Mal H., Fournier M., Brugiere O., Leseche G., et al.: *Survival benefit of lung transplant for cystic fibrosis since lung allocation score implementation*. *Am J Respir Crit Care Med.* 187 (2013)