IES 2022 Innovation & Society 5.0: Statistical and Economic Methodologies for Quality Assessment

BOOK OF SHORT PAPERS

Editors: Rosaria Lombardo, Ida Camminatiello and Violetta Simonacci

Book of Short papers
10th International Conference **IES 2022**
Innovation and Society 5.0: Statistical and Economic Methodologies for Quality Assessment

Department of Economics, University of Campania "L. Vanvitelli",
January 27th - 28th 2022

Scientific Committee of the group of the Italian Statistical Society on Statistics for the Evaluation and Quality of Services – SVQS

Pietro Amenta -University of Sannio
Matilde Bini -European University of Roma
Luigi D'Ambra -University of Naples "Federico II"
Maurizio Carpita -University of Brescia
Paolo Mariani -University of Milan "Bicocca"
Marica Manisera -University of Brescia
Monica Palma -University of Salento
Pasquale Sarnacchiaro -University of Rome Unitelma Sapienza

Program Committee of the conference IES 2022

Chair: Rosaria Lombardo -University of Campania "L. Vanvitelli"
Fabio Bacchini -ISTAT
Laura Baraldi -University of Campania "L. Vanvitelli"
Eric Beh -University of Newcastle, Australia
Wicher Bergsma -The London School of Economic and Political Science, UK
Enrico Bonetti -University of Campania "L. Vanvitelli"
Eugenio Brentari -University of Brescia
Clelia Buccico -University of Campania "L. Vanvitelli"
Rosalia Castellano -University of Naples "Parthenope"
Ida Camminatiello -University of Campania "L. Vanvitelli"
Carlo Cavicchia -University of Rotterdam, The Netherlands
Enrico Ciavolino -University of Salento
Corrado Crocetta -University of Foggia
Claudio Conversano -University of Cagliari
Antonello D'Ambra -University of Campania "L. Vanvitelli"
Antonio D'Ambrosio -University of Naples "Federico II"
Alfonso Iodice D'Enza -University of Naples "Federico II"
Tonio Di Battista -University of Chieti "G. D'Annunzio"
Michele Gallo -University of Naples "L'Orientale"
Francesco Gangi -University of Campania "L. Vanvitelli"
Michele La Rocca -University of Salerno
Amedeo Lepore -University of Campania "L. Vanvitelli"
Riccardo Macchioni -University of Campania "L. Vanvitelli"
Filomena Maggino -University of Rome "La Sapienza"
Angelos Markos -Demokritus University of Thrace, Greece
Lucio Masserini -University of Pisa
Stefania Mignani -University of Bologna
Nicola Moscariello -University of Campania "L. Vanvitelli"
Francesco Palumbo -University of Naples "Federico II"
Alessandra Petrucci -University of Florence
Alessio Pollice -University of Bari

Donato Posa -University of Salento
Luca Secondi -University of Tuscia
Amalia Vanacore -University of Naples “Federico II”
Michel van de Velden -University of Rotterdam, The Netherlands
Rosanna Verde -University of Campania “L. Vanvitelli”
Donatella Vicari -University of Rome “La Sapienza”
Grazia Vicario -Polytechnic University of Turin
Maurizio Vichi -University of Rome “La Sapienza”

Organizing Committee

Ida Camminatiello -University of Campania “L. Vanvitelli”
Antonello D’Ambra -University of Campania “L. Vanvitelli”
Rosaria Lombardo -University of Campania “L. Vanvitelli”
Elvira Romano -University of Campania “L. Vanvitelli”
Luca Rossi -University Cusano
Violetta Simonacci -University of Naples “L’Orientale”

SIS Società Italiana di Statistica

2022

pke PROFESSIONAL KNOWLEDGE EMPOWERMENT

Editors

Rosaria Lombardo - University of Campania "L. Vanvitelli", Italy
Ida Camminatiello - University of Campania "L. Vanvitelli", Italy
Violetta Simonacci - University of Naples "L'Orientale", Italy

PKE - Professional Knowledge Empowerment s.r.l.
Sede legale: Villa Marelli - Viale Thomas Alva Edison, 45 - 20099 Sesto San Giovanni (MI)
Sede operativa: Villa Marelli - Viale Thomas Alva Edison, 45 - 20099 Sesto San Giovanni (MI)
Ufficio Di Rappresentanza: Via Giacomo Peroni, 400 - 00131 Roma (RM)
CF / P.I. 03167830920 — www.pke.it; e-mail info@pke.it — Privacy

February 2022 PKE s.r.l.
ISBN 978-88-94593-35-8 on print
ISBN 978-88-94593-36-5 online

Contents

Solicited Session SS1 – *Compositional Data Analysis*
Organizer and Chair: Gianna Monti **1**
Rieser C. and Filzmoser P.. *Compositional Data and graph theory* 2
Egozcue J. J., Pawlowsky-Glahn V. and Buccianti A. *Compositional deviations from linear and non-linear equilibria* 8
Monti G. S. and Filzmoser P. *The knockoff filter for FDR control in robust ZeroSum regression in microbiome analysis* 14

Solicited Session SS2 - *Monitoring progress towards SDGs: statistical approaches and methods for measuring poverty, inequalities and food insecurity*
Organizer and Chair: Luca Secondi **18**
Marchetti S., Giusti C., Pratesi M. and Biggeri L. *Poverty indicators adjusted using local price indexes* 19
Tonutti G., Bertarelli G., Giusti C. and Pratesi M. *Assessing the targeting of the anti-poverty measure "Reddito di Cittadinanza" using Small Area Estimation methods* 25
Vargas-Lòpez A. and Secondi L. *Household consumption and food insecurity in Mexico: COVID19 and sustainable development* 31

Session of free contributes SCL1 – *Big Data, Proximity data, Multi way data*
Chair: Donatella Vicari **37**
Metulini R. and Carpita M. *Forecasting Traffic Flows with Complex Seasonality using Mobile Phone Data* 38
Simonacci V., Menini T. and Gallo M. *CP decomposition of 4th-order tensors of compositions* 44
Bove G. *A strategy of analysis of symmetry and skew-symmetry in asymmetric relationships* 50

Session of free contributes SCL2 –*Industry and Society*

Chair: Germana Scepi 56

Crosato L., Domenech J. and Liberati C. *Toward an early detection of SME's default with websites' indicators* . . . 57

Angelone R. *Italians' Culture and Values after two years of pandemic* . . 63

Aria M., Cuccurullo C., D'Aniello L. and Spano M. *Thematic evolution of Academic Medical Centers' research: a focus on Italian public owned AOUs in metropolitan areas* . . . 67

Session of free contributes SCL3 – *Society and Tourism*
Chair: Emma Zavarrone 73

Rondinelli R., Palazzo L. and Ievoli R. *Local clustering coefficient to measure intra-regional tourism in Italy* . . . 74

Zavarrone E., M. and Forciniti A. *Local clustering coefficient to measure intra-regional tourism in Italy* . . . 80

Firza N. *Sustainable tourism: The case of Albania* . . . 86

Romano M., Zammarchi G. and Conversano C. *Threshold-based Naïve Bayes Classifier Customer Satisfaction evaluation* . . . 90

Session of free contributes SCL4 –*Society and Innovation*
Chair: Michelangelo Misuraca 95

Seri E., Alaimo L. S., di Bella E., Cataldo R. and Piscitelli A. *Migrant Integration Policy Index (MIPEX): an analysis of countries via Gaussian mixture modelbased clustering* . . . 96

Rossi L. and Daddi S. *The measure of the BES: a proposal for the aggregation of the indicator education and training* . . . 102

Marino M., Mazza R., Misuraca M. and Stavolo A. *Monitoring consumer sentiment using control charts* . . . 108

Session of solicited contributes SS3 – *Statistical methods for the assessment of student careers in higher education*
Organizer and Chair: Maria Prosperina Vitale 114

La Rocca M., Niglio M. and Restaino M. *Predicting university students' churn risk* . . . 115

Primerano I., Santelli F. and Usala C. *Discovering archetypal universities in Italian higher education mobility flows* . . . 121

Cascella C. and Ragozzini G. *Measuring quality of students' careers in Higher Education: a systematic literature review* . . . 127

Porcu M., Sulis I. and Usala C. *Estimating the peers effect on students' university choices* . . . 134

Session of solicited contributes SS4 – *Conformity assessment and quality predictions- itENBIS*
Organizer and Chair: Amalia Vanacore **140**
Vanacore A., Pellegrino M. S. and Ciardiello A. *Testing the predictive performance of multi-class classifiers* . . . 141
Borgoni R., Gilardi A. and Zappa D. *Optimal Subgrids from Spatial Monitoring Networks* . . . 148
Pennecchi F. and Kuselman I. *Extension of the JCGM 106:2012 - Conformity assessment of multicomponent items and finite statistical samples* . . . 153

Session of solicited contributes SS5 – *Big Data and Large-dimensional Data*
Organizer and Chair: Stefania Mignani **160**
Farné M. and Vouldis A. *ROBOUT: a conditional outlier detection methodology for large-dimensional data* . . . 161
Camillo F. *Behaviours, emotions and opinions in modern citizen or customer relationship systems: a correct integration of small and big data for hyper-targeting, personal advertising and look-alike* . . . 168
Camminatiello I. and Lucadamo A. *A model for assessing sea environmental quality* . . . 172

Session of solicited contributes SS6 – *Health Quality*
Organizer and Chair: Paolo Mariani **178**
Bartolini B., Bertoldi S., Benedan L., Galeone C., Mariani P., Sofia F. and Zenga M. *The uneasiness index in a patient-designed quality of life questionnaire* . . . 179
Benedan L., Hachem M. E., Galeone C., Mariani P., Pilo C. and Tadini G. *Assessing the Quality of Life of patients with Epidermolysis Bullosa (EB): Development of a patient-centered questionnaire* . . . 184
Marletta A. and Morandi M. *Survival analysis in a business context: how to control the abandons of my subscribers* . . . 190

Session of free contributes SCL5 – *Assessing Performance*
Chair: Cristina Davino **194**
di Trapani G. *Political performance measuring and tracking through a system based on the Political Performance Indicator (Iep): Naples 2021 case* . . . 195
Montanari G. E. and Doretti M. *A class of case-mix adjusted probability-based indices for performance evaluation* . . . 202

Cavicchia C., Sarnacchiaro P., Vichi M. and Zaccaria G. *An ultrametric model to build a Composite Indicators system* 208

Session of free contributes SCL6 – *Statistical Learning*
Chair: Massimo Aria **212**

Migliorati M. and Brentari E. *Feature definition for NBA result prediction through Deep Learning* 213

Levantesi S., Lizzi M. and Nigri A. *An application of contrast trees for mortality models diagnostic and boosting* 219

Aria M., Gnasso A. and D'Aniello L. *Twenty Years of Random Forest: preliminary results of systematic literature review* 225

Park H., Hong J., Shin Y. and Park J-S. *A study on the GEV activation function for classification of class imbalance data* 231

Session of free contributes SCL7 – *Society and Disparity*
Chair: Pasquale Sarnacchiaro **236**

Alaimo L. S., D'Urso P. and Nigri A. *The gender gap in lifespan disparity as a social indicator of international countries: A fuzzy clustering analysis approach* 237

Iannario M. and Tarantola C. *Modelling scale effects via a Bayesian approach: an application to decision making in public sector* 243

Gangi F. , Daniele L. M. and Coscia M. *Board Gender Diversity and Social engagement: evidence from the banking industry* 249

Session of free contributes SCL8 – *Education*
Chair: Matilde Bini **256**

Davino C. and Lamberti G. *Assessing heterogeneity in students' performance. The case of the Massive Open Online Courses* 257

Primerano I., Catone M. C., Giordano G. and Vitale M. P. *Assessing undergraduate students' perceptions of distance learning during the COVID-19 pandemic* 263

Crisci A., Lucadamo A. and Amenta P. *PhD satisfaction analysis in Italian University via Classification tree, Bagging and Random Forest* 269

Cervellera S., Cusatelli C. and Giacalone M. *Comparative Analysis of Student Learning: Technical, Methodological and Result Assessing of PISA-OECD and INVALSI-Italian Systems* 275

Session of solicited contributes SS7 – *SEM with PLS: Theory and Applications*
Organizer and Chair: Enrico Ciavolino **281**

Wang S., Cheah J. and Roldan J. L. *PLS-SEM basics and its potential applications: A quick journey* 282
Cefis M. and Carpita M. *A PLS-SEM confirmatory composite analysis for football goalkeeper's performance validation* 288
Pasca P., Misuraca M., Meloni A. and Ciavolino E. *Text-mining and PLS-SEM combination to measure food satisfaction with Google Review: When the gut (re)counts!* 294

Session of solicited contributes SS8 – *Applications of non standard statistical tools to real-life*
Organizer and Chair: Antonio D'Ambrosio **300**
Iorio C. and Pandolfo G. *A robust strategy for building a financial portfolio* 301
Ruscone M. N. and De Luca G. *Conditional copula a financial application* 307
Ortu M., Frigau L. and Contu G. *Explaining Student Satisfaction Assessments: A Natural Language Processing Approach* 313

Session of solicited contributes SS9 – *Innovation and Value Co-creation in Society*
Organizers and Chairs: Alessandra De Chiara & Anna D'Auria **319**
Mauro S. *The smart working towards a Society 5.0* 320
Del Vacchio E., Carignani F., Laddaga C. and Bifulco F. *Innovative interaction in Society 5.0: insights from the cultural sector* 326
D'Auria A. and De Chiara A. *Society 5.0: a bibliometric analysis* 333

Session of solicited contributes SS10 – *Statistical learning for mainstream press, health and fiscal data*
Organizer and Chair: Claudio Conversano **339**
Baldassarre A. and Carullo D. *The regression trunk model for partitioning Italian municipalities based on their fiscal capacities and its determinants* 340
Pandolfo G. and Iorio C. *An analysis of Italian healthcare mobility through a depth-based clustering procedure* 346
Dossou B. F. P. and Wilhelm A. F. X. *Automatic Fake News Detection to Ensure Quality of News Articles* 352

Session of solicited contributes SS11 – *Multi-way Methods for Evaluation Service*
Organizer and Chair: Michele Gallo **358**
Simonacci V., Marino M., Grassia M. G. and Gallo M. *Multiple factor analysis with external information on PISA survey data* 359

Bocci L. and D. Vicari D. *A three-way analysis of well-being in Italy over time* . . . 365
Cerqueti R., Mattera R. and Scepi G. *Multiway approach for clustering time series with time varying parameters* . . . 371

Session of solicited contributes SS12 – *Assessment of Management Quality*
Organizer and Chair: Clelia Fiondella **377**
Belfiore A., Cuccurullo C. and Aria M. *Using partial triadic analysis for depicting the temporal evolution of Italian private healthcare organizations* . . . 378
Bollani L., Celegato A., Barbero F. and Fontemaggi F. *Management of the human factor into the company. An experience from the aeronautic sector.* . . . 385

Session of solicited contributes SS13 – *New technologies for students learning assessment and evaluation*
Organizer and Chair: Alfonso Iodice D'Enza **391**
Themelis E. and Markos A.*A non parametric cognitive diagnostic method in classroom assessment conditions* . . . 392
Iannario M., Iodice D'Enza A. and Romano R. *Hybrid unfolding models to Likert-scale data to assess distance learning perception in higher education* . . . 398
Pacella D., Fabbricatore R., Galluccio C. and Palumbo F. *Classification of Statistics learners using multi-dimensional latent class IRT model and archetypal analysis: the ALEAS app* . . . 404

Session of solicited contributes SS14 – *Labor Market and Enterprises*
Organizer and Chair: Lucio Masserini **408**
Maggioni G., Mariani P., Marletta A. and Zenga M. *Searching for new trends and dynamics in Labour Market: a statistical approach for the recruiting process* . . . 409
Bruttini P., Mariani P., Marletta A., Masserini L. and Zenga M. *A new definition of the professional figure Open Manager* . . . 413

Session of solicited contributes SS15 – *Statistical Approaches to Environmental Sustainability*
Organizer and Chair: Alfonso Piscitelli **417**
Aspinall R. *Measuring sustainability as an emergent property of whole system dynamics* . . . 418

Alaimo L. S. and Finocchiaro G. *Tourism sustainability in the Italian regions: a fuzzy approach* 424
D'Uggento A. M. *How young people perceive environmental issues, react to ecological concerns and commit themselves to sustainable behaviours* 430

Session of solicited contributes SS16 – *Statistical Methods for Environmental, Natural Resources and Health Assessment*
Organizer and Chair: Alessio Pollice **437**
Ferretti A., Ippoliti L. and Valentini P. *Spatial-ARFIMA models for the statistical analysis of environmental lattice processes* 438
Arima S., Pasculli G. and Polettini S. *A Bayesian non parametric approach for bias correction for underreported data* 444
Lasinio G. J., Mastrantonio G., Pollice A., Ventura D., Mancini G. and Ardizzone G. *Assessment of the impact of anthropic pressures on the Giglio island meadow of Posidonia oceanica* 450

Session of solicited contributes SS17 – *Functional Data Analysis Methodologies for Quality Assessment*
Organizer and Chair: Elvira Romano **456**
Fortuna F., Maturo F. and Di Battista T. *Improving the quality of questionnaires via the combined use of functional outlier detection and Item Response Theory* 457
Naccarato A., Fortuna F. and Terzi S. *Assessing government effectiveness over time: a functional data analysis approach* 462
Balzanella A. and Verde R. *Mining Distributed Acoustic Sensing data for vehicle traffic monitoring* 467

Session of solicited contributes SS18 – *Evaluation and assessment of cognitive and learning processes*
Organizer and Chair: Francesco Palumbo **473**
Ponticorvo M. and Argiuolo T. *Neuropsychological Assessment supported by Technology: the E-BTT case* 474
Milano N. and Gigliotta O. *Mining Introducing OpenAi-ES in Interactive Data. Clustering with R-EVOK* 480
Davino C., Gherghi M., Palumbo F. and Vistocco D. *Modeling heterogeneity in student's satisfaction during the Covid-19 pandemia* . . . 484

Session of solicited contributes SS19 – *Substainability and Environment*
Organizer and Chair: Ida Camminatiello **490**

Lombardo R. and Beh E. J. *Partitioning the Cressie-Read divergence statistic for three-way contingency tables: a study on environmental sustainability data* 491
Prahadchai T., Hong J., Busababodhin P. and Park J-S.. *Analysis of maximum precipitation in Thailand using non-stationary extreme value models* 498
Tregua M. and Scaglione M. *Assessing citizens' participation to urban transformation: a review of quantitative methods* 504

Session of solicited contributes SS20 – *Statistical methods for health and environmental impact assessment*
Organizer and Chair: Fabrizio Maturo **512**
Gattone S. A. and Di Battista T. *Density estimation via Functional Data Analysis* 513
Evangelista A., Acal C., Aguilera A. M., Sarra A., Di Battista T. and Palermi S. *A new multivariate functional ANOVA approach for assessing air quality data amid COVID-19 pandemic* 517
Diana A., Romano E. and Irpino A. *Conformal Prediction for Geographically Weighted Functional Regression models: an application for environmental impact assessment.* 523

Session of solicited contributes SS21 – *Local sustainability assessment: challenges in building quality indicators*
Organizers and Chairs: Francesca Fortuna & Alessia Naccarato **529**
Grimaccia E. *Urban Sustainability Assessment: A Proposal for an Index Based on SDGs' Indicators* 530
Liberati P. and Resce G. *Between and Within Country Inequality in Regional Well Being* 536
Di Battista T., Nissi E. and Sarra A. *Equitable and sustainable well-being over time: a functional approach* 542

Session of free contributes SCL9 – *Health and Covid-19*
Chair: Maria Sole Pellegrino **548**
Cao N., Calcagní A. and Finos L. *Twitter about COVID-19: An application of Structural Topic Models to a sample of Italian tweets* 549
Parretti C., Tartaglia R., Sbrana G., Mandó M. and Pacchi S. *Assessing the quality of a health service through the risk profile number (RPN)* 555
Di Lorenzo G., Franchetti G. and Politano M. *The insurance premium structure for a covid-19 insurance policy* 562

Session of solicited contributes SS22 – *Statistics, culture and tourism*
Organizer and Chair: Marica Manisera **568**
Cristiani P. *How data can influence the promotion and the consumption of cultural experience* 569
Carpita M., Manisera M. and Zuccolotto P. *Mobile phone data to monitor the impact of social and cultural events of Brescia* 575
Capecchi S., Quaranta G. and Salvia R. *Analysing opinions on sustainable tourism in the Vallo di Diano area, Campania, Italy* 582

Session of free contributes SCL10– *Time Series data, Panel data and Circular Economy*
Chair: Anna Crisci **588**
Scaccabarozzi D., Toninelli D., Zurlo D., Bacchini F. and Iannaccone R. *Testing the Participation Gap Inclusion within the Wage Phillips Curve* 589
Fusco D., Liguori M. A. and Moretti V. *Multisource approach for trends evaluation. An application at the agricultural sector* 596
Bonnini S. and Borghesi M. *A permutation test on the relationship between Circular Economy and firm size* 601
Mele S., Izzo F. and Tomnyuk V. *Circular economy and business models: a literature review* 607

Session of free contributes SCL11 –*Modelling Extreme Values, High dimensional, time series data*
Chair: Violetta Simonacci **611**
Shin Y., Busababodhin P. and Park J-S. *Modeling extreme values using the r-largest four parameter distribution* 612
Sabri M., Maturo F., Verde R., Riffi J. and Yahyaouy A. E. *Classification of ECG signals based on functional data analysis and machine learning techniques* 618
Park J-S., Shin Y., Shin Y. and Hong J. *Determining shape parameters in a climate multi-model ensemble* 624

THE INSURANCE PREMIUM STRUCTURE FOR A COVID-19 INSURANCE POLICY.

Struttura di un premio assicurativo per una polizza COVID-19.

Giovanna Di Lorenzo, Girolamo Franchetti and Massimiliano Politano

Abstract In the context of the Sars-CoV-2 virus pandemic, this paper deals with the analytical framework that involves a stochastic model to describe the probability of contagion and, therefore, of its outcome for a subject; subsequently, an actuarial model for an insurance policy against the risk of contracting the virus is proposed and the quantification of the related premium. It is assumed that the insurance coverage lasts for one year and that during the coverage it could happen the infection. The theoretical distribution of the contagion probability is of geometric type, in which every coverage day is a Bernoulli distribution of infection event. Four outcomes of the infection are considered below: hospitalization in home isolation, in hospital Medic Area, in intensive care and, finally, death. The Gamma distribution is taken into account as the theoretical distribution of number of days for each trajectory of recovery regarding the outcome of the infection, whereas for the outcome of death a lump-sum payment is defined to be paid as a single solution. A payment variable will be obtained whose mathematical expectation is the expected value of the expected benefits, assuming that in the event of death it remains the capital value. For each day of coverage, the expected payment is calculated and then weighted by the probability of infection on that given day; then, the expected payment is discounted to the effective date of coverage and, finally, it is calculated

[1]Giovanna Di Lorenzo, University of Naples “Federico II”; email:giodilor@unina.it
Girolamo Franchetti, University of Naples “Fe; email:gi.franchetti@studenti.unina.it
Politano Massimiliano*, University of Naples “Federico II”; email:politano@unina.it
*Corresponding Author

the fair premium of the policy. For this paper it is used the software R for statistical evaluation

Abstract *Nel contesto della pandemia Sars-CoV-2, questo lavoro affronta il quadro analitico che prevede un modello stocastico per descrivere la probabilità di infezione e, quindi, del suo esito per un soggetto; successivamente viene realizzato un modello attuariale per una polizza assicurativa contro il rischio di contrarre il virus e la quantificazione del relativo premio. Si ipotizza che la copertura assicurativa abbia una durata di un anno e che durante la copertura possa verificarsi il contagio. La distribuzione teorica della probabilità di contagio è di tipo geometrico, in cui ogni giorno di copertura è una distribuzione Bernoulliana dell'evento di infezione. Di seguito vengono considerati quattro esiti del contagio: il ricovero in isolamento domiciliare, in Area Medica ospedaliera, in terapia intensiva e, infine, il decesso. La distribuzione Gamma viene presa in considerazione come distribuzione teorica del numero di giorni per ogni traiettoria di ricovero in relazione all'esito della guarigione, mentre per l'esito del decesso si definisce un'indennità una tantum da erogare in un'unica soluzione. Si otterrà una variabile di pagamento la cui aspettativa matematica è il valore atteso dei benefici attesi, assumendo che in caso di morte rimanga il valore del capitale. Per ogni giorno di copertura viene calcolato il pagamento previsto e poi pesato per la probabilità di infezione in quel dato giorno; quindi, il pagamento previsto viene attualizzato alla data di decorrenza della copertura e, infine, viene calcolato il premio equo della polizza. Per questo lavoro viene utilizzato il software R per la valutazione statistica.*

Key words: Sars-Cov2, Covid-19 insurance, fair premium, vaccinated ad unvaccinated risk profile

1 The Model

The model is composed by two different sections strictly linked each other:
1 A stochastic model for the description of COVID-19 epidemic dynamics
2 An actuarial model to quantify a fair premium of a COVID-19 insurance policy.

1.1 Stochastic model

The stochastic model takes in account the data presented by the Instituto Superiore della Sanità (ISS) in his weekly report about the COVID-19 epidemics during the observation period from 14/07/2021 to 17/11/2021. It is calculated the probability of infection using the classic approach, taking in account the entire reference population, and the probability of a specific outcome of the infection itself taking in account infected people as the reference population:

Table 1: Infection and related outcomes probabilities

Non infection	*Infection*	*Home quarantine*	*Hospitalization*	*Intensive care*	*Death*

The Insurance premium structure for a Covid-19 insurance policy

$$p^S = \frac{S}{N} \quad p^I = \frac{I}{N} \quad p^C = \frac{C}{I} \quad p^O = \frac{O}{I} \quad p^{TI} = \frac{TI}{I} \quad p^D = \frac{D}{I}$$

Where

$p^S + p^I = 1$

$p^C + p^O + p^{TI} + p^D = 1$

Where *S* stands for people who are susceptible to the infection but not infected. Calculating these probabilities for every observation data it is possible to see the time series of this probabilities.

1.2 Actuarial model

Using the probabilities from the stochastic model seen before it is possible to quantify the fair premium of a COVID-19 insurance policy.

It is assumed that the policy coverage period is of 365 days and that the infection is represented by a bernoulli variable for every day of this period. Then, it is calculated the probability of being infected for every single day assuming that in the previous days the infection event has not happened.

$$\mathcal{B}(p^I) = \begin{cases} 0 & 1 - p^I \\ 1 & p^I \end{cases}$$

$$p^I(t) = \mathcal{G}(p^I) = (1 - p^I)^{t-1} * p^I$$

So, it is used a geometric distribution for representing this probability for every coverage day and, considering every single observation data, it is possible to obtain the dynamic of this probability during the observation period.

For every infection outcome (except death) it is calculated the expected benefit calculating the expected value of recovery days multiplied for the benefit paid every single recovery day. For death case it is paid a lump sum capital.

$$E[P] = E[rC] * p^C + E[rO] * p^O + E[rTI] * p^{TI} + 300 * p^D$$

Table 2: Benefit related to the infection outcomes (first three paid for every recovery day)

Home quarantine	***Hospitalization***	***Intensive care***	***Death***
1	1.5	2	300

It is used the probabilities related to the four possible outcomes to calculate the expected benefit of the policy by calculating the expected value of the expected benefits.

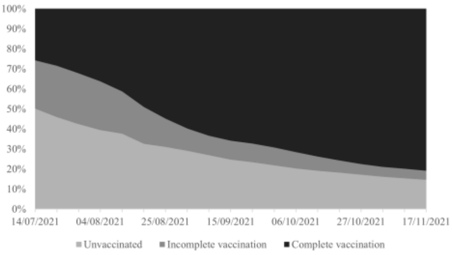
Finally, the expected payment is assumed the same for every day of the coverage period, weighted by the related infection probability and it is discounted at a technical interest rate of 1%.

$$Fair\ premium = \sum_{t=1}^{365} V_0(t) = \sum_{t=1}^{365} (1 + i)^{-\frac{t}{365}} * p^I(t) * E[P]$$

3 Data

It is used the data present in the weekly report about COVID-19 epidemic published by the Istututo Superiore della Sanità (ISS) about the infection and the related outcomes. This data are shown for three vaccination status: Unvaccinated, incomplete vaccination, complete vaccination.

Graph 1: Population partition by vaccination status



Regarding the recovery days, related to non-death outcomes, it is used the mean of the hospitalization and intensive care days presented by Azienda Sanitaria Toscana (ARS Toscana) and the mean of recovery days related to the home quarantine presented by the Health Ministry. So, the gamma distribution is considered as the theoretical distribution of number of days for each trajectory of recovery regarding the outcome of the infection with shape parameter equal to 2.

Graph 2: Theoretical gamma distribution of recovery days

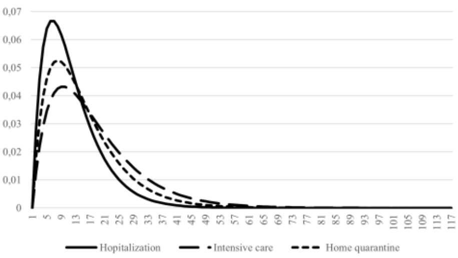


Table 3: Means of the recovery days

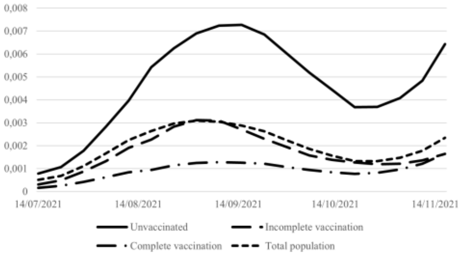
Home quarantine	***Hospitalization***	***Intensive care***
14	11	17.3

4 Analysis

In the analysis it is computed the time series of infection probability and the time series of the expected benefits by vaccination status. Then, it is calculated the fair premium for each class of people differentiated by vaccination status.

The Insurance premium structure for a Covid-19 insurance policy

Graph 3: Time series of infection probability by vaccination

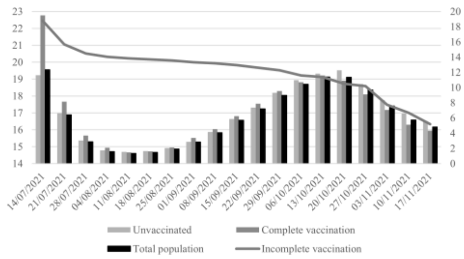


So, it is possible to see that during August and September the probability of infection increased due to the wave of infections and the same is happening during November.
Another statement is that the infection probability is higher than the other vaccination status and higher than the probability calculated on entire population. This data shows that the class of unvaccinated is the one with the highest risk about infection event.
Regarding the expected benefits related the infection outcome, they are shown in the graph except the lump sum capital in death case. The highest is the intensive care due to the mean of recovery days and the payment in each one that are the highest. The lowest is the home quarantine for the same reason presented before but the difference that in this case they are the lowest.

Table 3: Expected benefits for the infection outcome

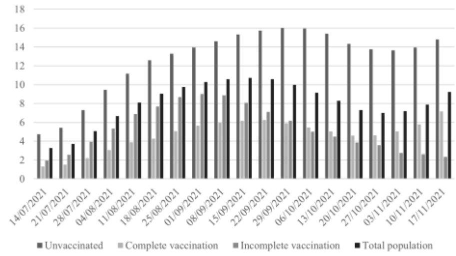
Home quarantine	*Hospitalization*	*Intensive care*
14,02368661	16,54542718	34,03564112

Graph 4: Expected payment of the policy by vaccination



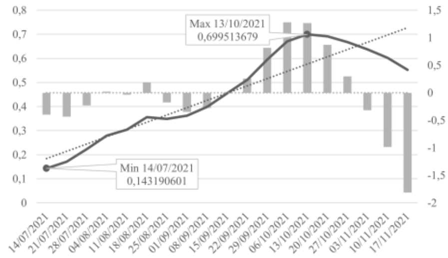
It is possible to state that during the wave of the infections, most of them are people that have the home quarantine as the infection outcome. So, this brings to see a reduction of expected payment during the wave, because it is more expected that it is paid the lowest benefit. Only by 20/10/2021 the expected benefit for the unvaccinated is higher than the expected for the entire population. Applying the actuarial model, it is possible to see that the class of unvaccinated is the one with the highest fair premium compared to the other vaccination status, confirming that this is the class of the highest risk related to the infection event and to the possible benefit. It is possible to see that the premium increases during the wave, as expected, and decreases outside it. Observing this result, it is possible to state that the vaccination status has a relevant impact on the quantification of fair premium.

Graph 5: Time series of the fair premium by vaccination



In fact, people who is unvaccinated has a higher risk to get infected and have a higher risk to be in more serious infection outcome.

Graph 6: Time series of the fair premium gap between unvaccinated people and total



In conclusion, if an insurance company sells a COVID-19 insurance policy to the entire population has a relevant risk to incur in losses due to the serious mismatch between the fair premium of the policy and the one that unvaccinated people should pay that is much higher that the one of the insurance policy offered to the entire population.
So, there is the necessity to apply a safety loading on the probability of infection and on the expected payment to cover the risk that the company could provide a payment higher that expected in the offered policy if the insured person who incurs in the infection has the unvaccinated status.

References

1. F. Gemmi, F., L. Bachini, L., S. Forni S.: I ricoveri per Covid-19 in Toscana. Aggiornamento "terza ondata" (2021). www.ars.toscana.it/2-articoli/4649-ricoveri-per-covid-19-in-toscana-aggiornamento-terza-ondata.html
2. Greenwood, P. E., Gordillo, L. F.: Stochastic Epidemic Modeling. In: Chowell, G., Hyman, J. M., Bettencourt, L. M. A., Castillo-Chavez, C. (eds) Mathematical and Statistical Estimation Approaches in Epidemiology, pp. 31-52. Springer (2009).
3. He, S., Hang, S., Rong, L.: A discrete stochastic model of the COVID-19 outbreak: Forecast and control. Mathematical Biosciences and Engineering **17**, 2792-2804 (2020)
4. Kermack, W.O., McKendrick, A. G.: A Contribution to the Mathematical Theory of Epidemics. Bulletin of Mathematical Biology **53**, 33-55 (1991).
5. Istituto Superiore di Sanità: Aggiornamento dei dati: i bollettini della sorveglianza integrata COVID-19 in Italia (2021). www.epicentro.iss.it/coronavirus/aggiornamenti
6. Pitacco, E.: Elementi di matematica delle assicurazioni. Lint Editoriale. Trieste (2016)
7. Zhou, Y., Ma, Z.: A Discrete Epidemic Model for SARS Transmission and Control in China. Mathematical and Computer Modelling **40** (13), 1491-1506 (2004).

Editors

Rosaria Lombardo - University of Campania "L. Vanvitelli", Italy
Ida Camminatiello - University of Campania "L. Vanvitelli", Italy
Violetta Simonacci - University of Naples "L'Orientale", Italy

PKE - Professional Knowledge Empowerment s.r.l.
Sede legale: Villa Marelli - Viale Thomas Alva Edison, 45 - 20099 Sesto San Giovanni (MI)
Sede operativa: Villa Marelli - Viale Thomas Alva Edison, 45 - 20099 Sesto San Giovanni (MI)
Ufficio Di Rappresentanza: Via Giacomo Peroni, 400 - 00131 Roma (RM)
CF / P.I. 03167830920 — www.pke.it; e-mail info@pke.it — Privacy

February 2022 PKE s.r.l.
ISBN 978-88-94593-35-8 on print
ISBN 978-88-94593-36-5 online