METRON (2016) 74:141–143
DOI 10.1007/s40300-016-0096-5

EDITORIAL

The statistical legacy of Corrado Gini

Marco Alfò[1] · Antonio Lijoi[2] · Donata Marasini[3] · Giancarlo Ragozini[4]

Published online: 27 July 2016

On september 2015, the Italian Statistical Society (SIS) has organized an international statistical conference on the legacy of Corrado Gini, hosted by the Ca' Foscari University at the Treviso campus. This conference has been the occasion to share research ideas from different fields, Statistics, Economics, Demography, Biology, Sociology and Official Statistics, and to outline the impact of Corrado Gini on past and present research in theoretical and applied Statistics. Fifty-one years after his death, Metron is proud to host a special issue on "The Statistical Legacy of Corrado Gini".

The issue collects a selection of papers presented at the SIS2015 meeting that have undergone a thorough review process to meet the high quality standards of the journal. These papers discuss the links between the research developed by Corrado Gini and current research topics in theoretical and applied statistics.

In the following, we provide a brief introduction to the articles included in this Special Issue. These articles cover different areas: inequality theory, survey sampling, discriminant analysis, stochastic dominance, ordinal data modeling, just to mention a few.

The Special Issue begins with the historical note written by Prévost [13]. The author offers a detailed and contextual reconstruction of Gini's intellectual trajectory up to the end of the

✉ Marco Alfò
marco.alfo@uniroma1.it

Antonio Lijoi
lijoi@unipv.it

Donata Marasini
donata.marasini@unimib.it

Giancarlo Ragozini
giancarlo.ragozini@unina.it

1 Dipartimento di Scienze Statistiche, Sapienza Università di Roma, Rome, Italy

2 Dipartimento di Economia e Management, Università di Pavia, Pavia, Italy

3 Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa, Università di Milano-Bicocca, Milan, Italy

4 Dipartimento di Scienze Politiche, Università degli Studi di Napoli Federico II, Naples, Italy

Second World War. The first four decades of the last century were indeed crucial for the history of Italian statistics, due to methodological innovations and major developments in teaching, research and dissemination of statistics and, in general, to the great changes in the Italian political history, the involvement in two global wars and the period of fascist dictatorship. Taking into account the peculiarities of the historical context, the author provides an accurate insight into the impact Corrado Gini had on the scientific community.

Tillé [14] proceeds over the same path with a review on the main contributions Corrado Gini has given to the fields of Survey Sampling and Inequality Theory. The author starts from the work of the commission created in 1924 by the Bureau of the International Statistical Institute (ISI) to assess the potential of the representative method [8,9]. Corrado Gini was a member of this commission together with Arthur Bowley, Adolf Jensen, Lucien March, Verrijn Stuart, and Frantz Zizek. The report of the commission contains a list of recommendations on random and purposive selection of samples and, from this viewpoint, it is often considered as the founding text of survey sampling theory. The author offers a review on the Gini index, see [5], with a focus on the corresponding variance estimator. Further, he details how a sample can be calibrated using the Gini index and discusses how these two main contributions can be fruitfully combined.

Lando and Bertoli-Barsotti [11] discuss Lorenz dominance to compare non-negative distributions in terms of inequality. Since, in several cases, Lorenz curves may intersect, alternative criteria need to be to introduced, see e.g. [4,16]. For this purpose, the authors review second-degree Lorenz dominance, which proves to be suitable for ranking single-crossing Lorenz curves. Further, they introduce a new ordering, disparity dominance, which, due to its peculiarities, is suited to rank double-crossing Lorenz curves. The authors show that the two approaches are basically complementary, and that, in both cases, the value of the Gini index is crucial for the ranking.

Costa [3] discusses one of the most relevant features of the Gini index, namely its decomposition, that can be effectively used to evaluate the observed distribution and study potential causal relationship. When the population is heterogeneous, and K subgroups are present, the measurement of the traditional components of inequality is challenging. The author discusses the effects of overlapping components, see the concept of *transvariazione* introduced by [6], on the inequality within subgroups and the inequality between subgroups, and shows that their presence need to be carefully considered as it clearly leads to a more detailed framework for both measurement and interpretation of the inequality. In particular, it is shown that overlapping units should be analyzed separately from non overlapping units to properly evaluate and explain the total inequality.

Trendafilov and Gebru [15] discuss issues raising when *horizontal* data, i.e. data with more variables than observations, are analyzed. In such empirical situations, the classical Fisher's linear discriminant analysis (LDA) cannot be applied since the within-group sample covariance matrix is singular. Furthermore, the number of variables is usually large and the standard solutions (discriminant functions) may be complex to interpret. The aim is to develop fast and reliable algorithms for sparse LDA on horizontal data. The resulting discriminant functions depend on very few original variables, and this makes their interpretation quite easy. The main theoretical and numerical challenges entail coping with singularity of the sample covariance matrix. In this work, the authors classify the existing approaches according to the way they tackle the singularity issue, and try to suggest new ones; the authors notice that linear combinations can also be obtained in terms of Gini's concept of *transvariazione* [6], e.g. along the lines discussed by [1,12].

The paper by Capecchi and Iannario [2] entails sample surveys with responses measured on ordinal scales; frequently, in these cases, an inherent indecision is registered, that can

be generated by either objective and/or subjective causes. Therefore, such a component should be taken into account to avoid bias in estimation, interpretation and prediction. The authors show how the heterogeneity index proposed by [5] and its variants may be used to produce effective measures for such uncertainty; in particular, the index is commonly used to emphasize, for a given population, how evenly the subjects are distributed among categories. Starting from this perspective, the authors clearly exploit and discuss relationships between the Gini heterogeneity index and the parameters of a mixture model, the generalized CUB model, with a view towards exploratory procedures for covariates selection. They suggest to use the index as a non-parametric measure to detect uncertainty/heterogeneity in a model-based approach for the analysis of ordinal data.

We thank all the authors that have contributed to the realization of this Special Issue. Their enthusiasm has pushed us forward during all this period. Warm thanks go also to the reviewers, who have provided the authors with valuable comments and suggestions, and have greatly contributed to the overall quality of the volume. A special thank goes to Giovanni Maria Giorgi, editor-in-chief of Metron, who has invited and always supported us in our guest editor role. A further thank goes to Corrado Crocetta, program chair of the SIS2015 meeting, who has given us support throughout the development of the special issue. Last, but not least, thanks are due to the editorial board and the staff of the journal for giving us the chance of putting together this piece of work.

References

1. Calò, D.G.: On a transvariation based measure of group separability. J. Classification **23**, 143–167 (2006)
2. Capecchi, S., Iannario, M.: Gini heterogeneity index for detecting uncertainty in ordinal data surveys. Metron (2016). doi:10.1007/s40300-016-0088-5
3. Costa, M.: Overlapping component and inequality decomposition: a simulation study for the Gini index. Metron (2016). doi:10.1007/s40300-016-0092-9
4. Fullerton, D., Rogers, D.L.: Who bears the lifetime tax burden. The Brookings Institution, Washington D.C (1993)
5. Gini, C.: Variabilità e Mutabilità, Studi Economici-giuridici della Regia Facoltà di Giurisprudenza. Tipografia Paolo Cuppini, reprinted in Gini, Bologna (1939)
6. Gini, C.: Sulla misura della concentrazione e della variabilità dei caratteri, Atti Reale Istituto Veneto di Scienze, Lettere ed Arti, 73, 1203–1248 (English translation in. Metron **63**, 338 (2005)
7. Gini, C.: Measurement of inequality and incomes. Econ. J. **31**, 124–126 (1921)
8. Gini, C.: Une application de la methode représentative aux matériaux du dernier recensement de la population italienne (ler décembre 1921). Bull. Int. Statistical Inst. **23**, 198–215 (1928)
9. Gini, C., Galvani, L.: Di una applicazione del metodo rappresentativo al censimento italiano della popolazione 1. dic. 1921. Annali di Statistica **6**, 1–107 (1929)
10. Gini, C.: Memorie di Metodologia Statistica, vol. I. Giuffrè, Variabilità e concentrazione, Milano (1939)
11. Lando, T., Bertoli-Barsotti, L.: Weak orderings for intersecting Lorenz curves. Metron (2016). doi:10.1007/s40300-016-0087-6
12. Montanari, A.: Linear discriminant analysis and transvariation. J. Classification **21**, 71–88 (2004)
13. Prévost, J.-G.: Statistics, expertise and politics: Corrado Gini and the course of Italian history. Metron (2016). doi:10.1007/s40300-016-0085-8
14. Tillé, Y.: The legacy of Corrado Gini in survey sampling and inequality theory. Metron (2016). doi:10.1007/s40300-016-0090-y
15. Trendafilov, N.T., Gebru, T.G.: Recipes for sparse LDA of horizontal data. Metron (2016). doi:10.1007/s40300-016-0093-8
16. Zoli, C.: Intersecting generalized Lorenz curves and the Gini index. Soc Choice Welf. **16**, 183–196 (1999)