Journ@l Electronique d'Histoire des Probabilités et de la Statistique

Electronic Journ@l for History of Probability and Statistics

Vol 1, n°1; Mars 2005

Early Gini's Contributions to Inequality Measurement and Statistical Inference

ANTONIO FORCINA[1] and GIOVANNI MARIA GIORGI[2]

Résumé

Dans ce papier, nous examinons les contributions originales à la mesure d'inégalité des revenus et de richesse que Gini a publiées au début de sa carrière scientifique. Nous montrons aussi que dans son travail sur probabilité et inférence statistique, il a anticipé des notions de surdispersion et d'échangeabilité.

Abstract

In this paper we review the seminal contributions to the measurement of inequality in income and wealth that Gini made towards the beginning of his scientific career. We also show that in his early work on probability and inference contained a clear anticipation of the notions of overdispersion and exchangeability.

1. Introduction

Corrado Gini was born in 1884 in Motta di Livenza (Province of Treviso) in the North East of Italy, he graduated at Bologna University in 1905 with a degree in Law, the only Faculty where Statistics was taught as a subject at that time. His dissertation, a throughout investigation into the distribution of the sex ratio in human populations, later became a book [Gini, 1908a] and is an early example of the wide range of his scientific interests in Demography, Genetics and Biology. Soon his interests extended into Sociology and Economics (see [Giorgi, 2001] and [Regazzini, 1997] for some biographical notes). As far as statistics is concerned, his contributions pertained mainly to descriptive measures of dispersion, concentration, association as well as to foundational aspects of Statistical Inference (see [Forcina,1982] for a critical assessment).

In 1920 Gini founded Metron, an international statistical journal: the first issue had a paper by Edgeworth on entomological statistics, the third issue had a paper by L. J. Reed who introduced

[1] Department of Statistics, University of Perugia (Italy), forcina@stat.unipg.it

[2] Address for correspondence: "La Sapienza" University , Department of Statistics, Probability and Applied Statistics, Piazzale Aldo Moro n.5, 00185 Roma (Italy), giovanni.giorgi@uniroma1.it

principal components in two dimensions and the fourth issue a paper by R. A. Fisher on the "probable error" of the correlation coefficient. In 1934 he founded Genus, a journal devoted to demography. He was editor and owner of both journals until his death. Gini held several public positions; in particular he was President of the Central Italian Statistical Institute from 1926 to 1932; during these years he reorganized the national statistical services. When he died in 1965, his scientific production contained over 800 titles (see [Castellano, 1966] for the complete bibliography). In this paper we limit our attention to his early contributions.

2 Inequality measurement

Gini's interest in economic problems, and in particular into the measurement of inequality in the distribution of income and wealth, probably dates back to his university education. The political and economic debate on the way to reach a more equal distribution of income and wealth was particularly alive at the beginning of the last century, partly stimulated by the contribution made by Vilfredo Pareto.

2.1 The δ index

Pareto [1895 and 1897], inspired by the analysis of fiscal data, proposed the following formulation. Having denoted with x a given income and with N the number of taxpayers with an income greater or equal to x, he noted that a plot of the points ($\log x$, $\log N$) had a clear tendency to cluster around a straight line. Therefore, according to Pareto, the distribution of total income was well represented by the equation

(2.1.1) $$\log N = \log A - \alpha \log x$$

that is

(2.1.2) $$N = Ax^{-\alpha}$$

where the parameter α is the well known Pareto index.

Because in the large number of empirical applications that he made the estimated values of α varied very little, he reached the conclusion that the distribution of income was more or less the same in all the countries and for all the years he considered. Such an assertion seemed quite unrealistic and in Gini's opinion it was based on results obtained using indices that were not very sensitive. In order to overcome such an inconvenience Gini [1909] proposed a new index of inequality based on the following: given n incomes ranked in non-decreasing order ($x_{i-1} \leq x_i$), the average for the last m incomes must be greater than the overall average as long as $n > m$, thus

(2.1.4) $$\frac{1}{m}\sum_{i=n-m+1}^{n} x_i > \frac{1}{n}\sum_{i=1}^{n} x_i$$

which implies

(2.1.5) $$\left[\sum_{i=n-m+1}^{n} x_i \Big/ \sum_{i=1}^{n} x_i\right] > \frac{m}{n}$$

thus it must exist $\delta < 1$ such that

(2.1.6) $$\left[\sum_{i=n-m+1}^{n} x_i \Big/ \sum_{i=1}^{n} x_i\right]^{\delta} = \frac{m}{n},$$

δ may be interpreted as the concentration index which, in the case of incomes, Gini [1911a, p.16] defined as "the exponent to which the fraction of assessed and taxed incomes possessed by the taxpayers with the highest income should be raised so as to obtain the fraction of taxpayers that possess it". He denoted the global income of a taxpayer with *x*, the frequency of people with an income greater than *x* with N and the total amount of their incomes with C; when the values of $\log N$ were plotted against those of $\log C$, using Pareto's previous data, Gini obtained points which could conveniently be fitted by the straight line

(2.1.7) $$\log N = -\log K + \delta \log C$$

or, equivalently,

(2.1.8) $$N = \frac{C^{\delta}}{K}.$$

Gini emphasised that δ, contrarily to α, grew as the concentration grew. Gini [1911a, p.41-43] also derived the theoretical relation between α and δ

(2.1.9) $$\delta = \frac{\alpha}{\alpha - 1}$$

by which he claimed that the "distribution of wealth was enormously different from one country to another and from time to another", thereby pointing out how Pareto's conclusions on the uniformity of the distribution of wealth were only due to the fact that his index was rather insensitive. However Gini [1911a, p.44] asserted that the relation (2.1.9) was only theoretical inasmuch as in practice it was rarely true and he hypothesized that difference between theoretical and empirical results was due to the influence of tax evasion on the values of α and δ.

In general, Gini [1911a, pp.48-49] thought that δ had several advantages with respect to α and, in particular that the index δ was more sensitive than α and that the influence of inflation was less strongly felt by δ than by α. In the years that followed there was a heated debate in the literature on the advantages and disadvantages of the two indices but as Arnold [1983, p.5] rightly pointed out "both α and δ are questionable summary measures of inequality since they both are only meaningful when the parent distribution is Paretian".

2.2 The Gini concentration ratio

Gini himself was not satisfied with the results on inequality measurement obtained with δ and, even though he considered the latter better than α, he continued to study the topic and few years later derived an inequality measure which is still of topical interest (see [Giorgi, 1990,1993,1999]). In particular, Gini [1914] proposed the *concentration ratio* R and showed how it was linked to the Lorenz [1905] curve and the mean difference [Gini,1912]. He also provided formulas for calculating R from grouped data and examined the influence which the truncation of the variable distribution in its lower part could have on such an index. Naturally, he accompanied everything with appropriate empirical applications.

In more detail, let be $x_1 \le x_2 \le ... \le x_n$ the set of households (or individuals) incomes arranged in non-decreasing order and

(2.2.1) $$C_n = x_1 + x_2 + ... + x_n = \sum_{i=1}^{n} x_i = n\mu$$

the total income, where μ is the mean income. Gini considered the fraction of the total income received altogether by the i poorest units

(2.2.2) $$q_i = \frac{C_i}{C_n} = \frac{1}{n\mu}\sum_{j=1}^{i} x_j \qquad i = 1,2...,n$$

against the fraction of the corresponding frequency

(2.2.3) $$p_i = \frac{i}{n} \qquad i = 1,2...,n.$$

Obviously, since $i < n$, the first i individuals (or households) in the income ranking are considered, the concentration must be higher when their total income share is smaller. Or, equivalently, as Gini [1914, p.1204] put it, "the concentration of the variable (income) is stronger as the inequality $p_i > q_i$ increases for the $n-1$ values of i". This is so because, when $i = n$, we always have $p_n = q_n = 1$. He also noted that, in relative terms, income inequality must increase when the ratios

(2.2.4) $$R_i = \frac{p_i - q_i}{p_i}$$

increase and proposed [Gini,1914, p.1207], as concentration measure, "the weighted mean of the $n-1$ values of R_i, where each R_i has a weight proportional to the value of p_i", in this way he obtained the index

(2.2.5) $$R = \frac{\sum_{i=1}^{n-1}(p_i - q_i)}{\sum_{i=1}^{n-1} p_i}$$

which he called *concentration ratio*. In the case of perfect equality of incomes, that is when $p_i = q_i (\forall i)$, $R = 0$ while the maximum value $R = 1$ is achieved only when $q_i = 0$ for $i = 1,2,...,n-1$.

Remembering that $p_i = i/n$ and

(2.2.6) $$\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$$

(2.2.7) $$\sum_{i=1}^{n-1}\sum_{j=1}^{i} x_j = \sum_{i=1}^{n-1}(n-i)x_i$$

after some algebra, the Gini concentration ratio may be written as

(2.2.8) $$R = \frac{2}{n(n-1)m}\sum_{i=1}^{n} i\,x_i - \frac{n+1}{n-1}.$$

Gini [1914, p.1229-1233] also showed how to arrive at the concentration ratio trough the Lorenz diagram. More precisely he suggested that, for large *n*, one could measure the inequality through the index

(2.2.9) $$R = \frac{A}{\max A} = \frac{A}{1/2} = 2A$$

where, in a Lorenz diagram, *A* is the area between the diagonal (egalitarian line) of the unit square and the Lorenz curve (*LC*). In the extreme case where all incomes are equals, *LC* coincides with the diagonal and $A = 0$ and, in the other extreme case, if the overall income is "concentrated" on the hands of a single subject, $A = 1$.

Further, Gini [1914, p.1237-1238] demonstrated that the concentration ratio *R* can be expressed in terms of the mean difference (Δ), that is

(2.2.10) $$R = \frac{\Delta}{\max \Delta} = \frac{\Delta}{2m}$$

where

(2.2.11) $$\Delta = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \left| x_i - x_j \right|$$

and $2m$ is, for large *n*, the maximum of Δ.

Gini, in his seminal paper of 1914, considered the concentration ratio in the discrete case; a few months later, two papers by Pietra [1915] were published in which a simple geometrical procedure for computing the concentration area was proposed. Furthermore, in these papers, the continuous Lorenz curve in explicit form was defined for the first time (see also [Gastwirth, 1971]) and the graduation function, also known as the inverse cumulative distribution function, was introduced.

The Gini concentration ratio was, for a long while, at the heart of the debate on the measurement of income inequality. It stimulated fruitful discussion as well as sharp disagreement among economists, statisticians, sociologists and econometricians. The relationship with the mean difference, in particular, and the authorship of resulting theories was the subject of bitter disputes. An article by Bresciani Turroni [1916] questioned Gini's preference for the mean difference rather than other summary indices; Pietra [1917] motivated his preference for the mean difference and harshly criticized assertions and results by Ricci [1916]. The quarrel was taken up again at the 19th Session of the International Statistical Institute held in Tokyo in 1930 when a paper by Bortkiewicz [1931a] who, while analyzing some summary indices, among which the mean difference and the relative mean difference suitable for measuring income inequality, presented as new results which had been obtained several years before by Gini [1912] and Pietra [1915]. This provoked resentful reaction from the latter [Gini, 1931a,b; Pietra, 1931a,b]. Bortkiewicz [1931b] reacted to their criticisms, and complained that he had been unable to access Gini's [1912] book on the subject.

The early 70s witnessed a novel interest on Gini ratio that lead to explore new, more modern and sophisticated aspects which have gone well beyond Gini's initial framework; for the genesis, evolution and some interesting interpretations of the Gini index, see also Giorgi [1990, 1993, 1999].

3 On the concept of probability

Gini tried to come to terms with the foundational aspects of probability in two early papers [Gini,1907, 1908b]. In the following only the first of these two papers are discussed because the second seems to be mainly a simplified exposition of the same ideas without going as much in depth. He first discussed the notion of *equally likely* events upon which the classical definition rests. He criticized the widely accepted opinion according to which two events could be regarded as equally likely whenever there was no reason against such a belief. He opposed this not simply because, as was usually said, "one cannot extract knowledge from ignorance", but on the ground that sets of events can often be grouped into a single event in an arbitrary way (p.4); thus the number of equally likely events is not determined uniquely. This argument is similar to the objection against *uniform priors*, a notion which is not invariant to non linear transformations. He stressed the fact that it is not the lack of knowledge, but, possibly, specific features of symmetry and homogeneity in the physical construction, which may be used as the base for judging that a set of events are equally likely.

Gini's main objection to the definition of probability as the frequency in the long run, is that it treats probably as an abstract property of events, irrespective of the "circumstances of time and place" which are related to the events of interest (p.7). He seems to say that, though events are never equal, we can still learn from experience and base our assessment of the probability of a future event of interest on the frequency of occurrences in a suitable collection of similar events. This is reasonable if we feel that we have no reason to expect that, by selecting a smaller and more specific subset of cases (that is conditioning to additional specific features), our assessment will not be improved. In taking the observed frequency as the probability, he seems to say, we are optimising our behaviour in the sense that the overall frequency within the selected set of cases, being the arithmetic mean of the frequencies computed within any additional partition of the same set, minimizes the sum of expected errors (p.8). In other words, when people place the event of interest within a given collection of similar events in order to compute its probability, they do so because they believe that the probability of the single events in the class should differ only randomly, otherwise we would gain by referring to a smaller and more specific collection. Clearly, once it is known which event has occurred, people may discover that they would have gained by measuring the probability of the event of interest with the frequency computed in a narrower collection. However Gini claims that, in the long run, the proportion of those who would have lost by behaving differently is going to be much bigger (p.10).

The interesting aspects in this discussion are both the notion that events should always be conditional to *informative features* specific to the event of interest as well as the attempt to formulate a decision framework with the intuitive reference to the mean square error. However it is not clear whether the choice of the relevant features upon which to construct a suitable collection of similar events is subjective or not. He discusses the issue again with reference to his intuitive notion of mean square error as follows (p.11). Gini observed that the knowledge of the features which may affect the occurrence of the event "varies widely depending on the information and intuition of each individual". In addition, because the sum of squares from the overall mean are always larger than the sum of the squares computed from several arithmetic means, each one computed within a separate sub-collections, Gini was convinced that it was convenient for each individual to behave according to his own assessment of the probability. The conclusion seems to be that, in a situation of shared information, the probability of an event has an objective value, however in practice it can only be determined on a subjective base.

3.1 Over and under-dispersion with binary data

The subject is discussed with some detail in Gini [1908a, Chapter V]: the issue is whether from the fact that the frequencies of, say, a male birth varies across regions (or subsections of the population or a sequence of years) in such a way that the resulting distribution has "normal dispersion" may be taken as evidence that the probability of a male birth is not affected by the region (or the subsection of population considered or by the frequency in previous years). *Normal dispersion* in the language of that period meant that the variance of the empirical distribution was equal to the binomial variance where the unknown probability was replaced by the sample estimate.

Gini's point here is not only to stress that an apparent normal dispersion could hide substantial variations, a discovery which he attributes to Poisson and von Bortkiewicz, (see Stigler, 1986, p. 229-234 for an assessment of the contribution by Lexis). In this passage Gini presents a larger list of possible models leading to normal, more than normal and less than normal dispersion which are of some complexity even for a modern statistician. These models he describes in words and then by a simulation device based on drawing balls from sets of urns having different proportions of black balls, in an effort to define precisely, but without formal notations (which probably he did not have), the different models. The most relevant models are described below, translated in modern words, to make the exposition simpler to follow.

1. Suppose that the probability of a male birth is different across regions; if we ignore the region, the overall resulting distribution will exhibit over-dispersion.
2. Suppose that the probability of a male birth depends on the age of the mother and that in each region we take a stratified sample of births with mothers of different ages; the resulting overall distribution will exhibit under-dispersion.
3. In the same context as above, if, instead of taking a stratified sample, we take a simple random sample within each region, so that the proportion of mothers of different ages can fluctuate at random across regions, the resulting distribution will have normal dispersion.

4. Consider the proportion of male births in families with *k* children; if the probability that the first born is a male was the same for all families but the probability of successive births depended on the proportion of males in previous births, the presence of positive correlation would imply over-dispersion while negative correlation would imply under-dispersion.
5. Suppose that the probability of a male birth within a given community depends on the proportion of male births in the previous years; then the proportion of male births in a sequence of years will exhibit over-dispersion irrespective of the fact that association is positive or negative.

Basically, in model 1. there are two components of variance: the binomial variances within regions and the variation across regions. In model 2. under dispersion is produced by mixing; to take the extreme case, suppose that younger mothers had only males and older ones had only females, then the overall variance would be exactly 0. In model 3. the effect of mixing, which reduces the variance, is compensated by the multinomial fluctuation of the mixing weights which, instead, were constant within the stratified sampling. The result of model 4. may be derived as the variance of a sum of correlated observations while model 5. may be seen as a special case of model 1. where over-dispersion is produced by the fact that the probability of a male birth is not constant, irrespective of the sign of the serial correlation.

As a general comment, it can be noted that some of these models are rather sophisticated even for a modern statistician. The book contains extended analyses of several different datasets in order to test a wide range of possible underlying assumptions; for instance, on p. 367-370 he considered the time series of the sex ratio in Sweden and in several Italian regions and classified successive pairs of years in a two-way table according to the joint distribution of the sex ratio and found that the correlation was positive and around 0,5 which explains the amount of observed over-dispersion within each series.

3.2 Empirical exchangeability

After 1908 Gini must have thought very deeply about using the binomial model and its modifications for inference. The results of this investigation, which are surprisingly innovative, are contained in a single paper [Gini, 1911b] which has been reprinted several times. In the following reference to pages are based on the version printed in *Metron*. Forcina [1993] noted that in this paper Gini provides an anticipation of the notion of empirical exchangeability.

Since the beginning he stated clearly his aim: how to assign a value to the probability that a given event occurs in the future given that the event has occurred with a given frequency in the past; this is what modern Bayesians call *predictive probability*. Having derived this probability through Bayes' theorem (p.135), he noted that this formula is not applicable in practice as we do not know most of the quantities required, unless we make arbitrary assumptions about uniform prior probabilities, a solution which he considered to be arbitrary (p. 136).

In order to summarize Gini's approach it is useful to introduce some notations which, however, are different from his own notations, nonetheless the line of the argument is exactly the same (p. 137-139). Let S_n, S_1, S_t denote, respectively, the number of times that the event has occurred in a set of n trials, in a single next trial and in a final set of t trials. Let also $P(S_n) = m$ be the probability that the event occurs m times in n trials. His argument then goes like this:

(3.2.1) $$P(S_1 = 1 \mid S_n = m) = \frac{P(S_n + S_1 = m+1)\ P(S_1 = 1 \mid S_n + S_1 = m+1)}{P(S_n = m)}.$$

Because he wanted to apply this formula to compute the probability of a future male birth for a couple who had *m* males in *n* previous births, he was confident to be able to estimate the first factor in the numerator from the large datasets available, so he remained with the problem of computing the conditional probability in the numerator and the denominator. Under the assumption that the same "system of causes" is operating across the $n+1$ trials and that, conditionally on this, events are independent, the second probability in the numerator is simply given by the hypergeometric formula. He also expanded $P(S_n = m)$ first as $P(S_n = m,\ S_1 = 0) + P(S_n = m,\ S_1 = 1)$ and then wrote each of these terms as in the numerator above, so that the hypergeometric expression can be used again. Now, the assumptions that the same "system of causes" is operating, once translated into modern language means that there the $n+1$ trials are made under the same context and, given this, they are independent. Gini's treatment of Bayes formula (p.135) makes it clear that he is conditioning to a give context which remains constant throughout the whole sequence of trials, but the context is itself a random event with a prior probability. This assumption, which leads to compute the probability $P(S_1 = 1 \mid S_n + S_1 = m+1)$ as in the hypergeometric distribution, is equivalent to the assumption required for *finite exchangeability* within the given sequence, though the language with which this is expressed is very different from modern language.

The purpose of this computation is not simply predictive. Gini was convinced that the probability of a male birth did not remain constant for a given couple and he was searching for a way to test his null hypothesis that the probability does remain constant, which is equivalent to his intuitive notion of exchangeability. Then he goes on like this: suppose we have data about families with $m+1+t$ births and suppose we assume that the same assumptions about constant probability of a male birth within a given family hold throughout the $n+1+t$ births. Then we may expand $P(S_n + S_1 = m+1)$ and $P(S_n + S_1 = m+0)$ as fallows

$$(3.2.2) \qquad P(S_n + S_1 = m + x) = \sum_h P(S_n + S_1 = m + x,\ S_t = h)$$

$$= \sum_h P(S_n + S_1 + S_t = m + x + h) P(S_t = h \mid S_n + S_1 + S_t = m + x + h)$$

so that the assumption of exchangeability, extended to the whole sequence of $n+1+t$ trials, again leads to replace the conditional probability with the hypergeometric function while the remaining marginal probabilities may be estimated from overall observed frequencies. In this way he has a sequence of estimates of the same probability and he expects to be able to discern whether the sequence of his estimates evolves at random or according to well defined pattern, providing evidence against the null.

References

[Arnold, 1983] B.C. Arnold: Pareto distributions, Int. Publishing House, Burtonsville, 1983.

[Bortkiewicz, 1931a] L. Bortkiewicz: Die Disparitätsmasse der Einkommensstatistiks, Bulletin de l'Institut International de Statistique, XIX Session (Tokyo,1930), 25(3), 189-298, 1931.

[Bortkiewicz, 1931b] L. Bortkiewicz: Erwiderung, Bulletin de l'Institut International de Statistique, XIX Session (Tokyo,1930), 25(3), 311-316, 1931.

[Bresciani Turroni, 1916] C. Bresciani Turroni: Recenti pubblicazioni: <Gini C.(1914). Sulla misura della concentrazione e della variabilità dei caratteri>, Giornale degli Economisti e Rivista di Statistica, 27, 68-70, 1916.

[Castellano, 1966] V. Castellano: Corrado Gini. A memoir, Metron, 24, 3-84, 1966.

[Forcina, 1982] A. Forcina: Gini's contributions to the theory of inference, International Statistical Review, 50, 65-70, 1982.

[Forcina, 1993] A. Forcina: Discussion of "Exchangeability and data analysis" by Draper, Hodges, Mallows and Pregibon, Journal of the Royal Statistical Society, A156, 33, 1993.

[Gastwirth, 1971] J.L. Gastwirth: A general definition of the Lorenz curve, Econometrica, 34, 1037-1039, 1971.

[Gini, 1907] C. Gini: Contributo alle applicazioni statistiche del calcolo delle probabilità, Giornale degli Economisti, 32, 3-41, 1907.

[Gini, 1908a] C. Gini: Il sesso dal punto di vista statistico, Sandroni, Napoli, 1908.

[Gini, 1908b] C. Gini: Sul concetto di probabilità. IN: Questioni filosofiche, Biblioteca di Filosofia e Pedagogia, Formaggini, Bologna, 1908.

[Gini, 1909] C. Gini: Il diverso accrescimento delle classi sociali e la concentrazione della ricchezza, Giornale degli Economisti, 20 (series 2), 27-83, 1909.

[Gini, 1911a] C. Gini: Indici di concentrazione e di dipendenza, Biblioteca dell'Economista, 20 (series 5), Utet, Torino, 1911.

[Gini, 1911b] C. Gini: Considerazioni sulla probabilità a posteriori e applicazioni al rapporto dei sessi nelle nascite umane, Studi Economici e Giuridici, Università di Cagliari, 3, 1-37, 1911. (reprinted in Metron, 15, 133-171, 1949).

[Gini, 1912] C. Gini: Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche, Studi Economico-Giuridici della Facoltà di Giurisprudenza della Regia Università di Cagliari, 3 (part 2), Cuppini, Bologna, 1912.

[Gini, 1914] C. Gini: Sulla misura della concentrazione e della variabilità dei caratteri, Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti, 73, 1203-1248, 1914.

[Gini, 1931a] C. Gini:. Observations à la communication du Prof. L. von Bortkiewicz, Bullettin de l'Institut International de Statistique, XIX Session (Tokyo,1930), 25(3), 299-306, 1931.

[Gini, 1931b] C. Gini:. Note, Bullettin de l'Institut International de Statistique, XIX Session (Tokyo,1930), 25(3), 317-320K, 1931.

[Giorgi, 1990] G.M. Giorgi: Bibliographic portrait of the Gini concentration ratio, Metron, 48(1-4), 183-221, 1990.

[Giorgi, 1993] G.M. Giorgi: A fresh look at the topical interest of the Gini concentration ratio, Metron, 51(1-2), 83-96, 1993.

[Giorgi, 1999] G.M. Giorgi: Income inequality measurement: the statistical approach. IN: Handbook on income inequality measurement (Silber J. editor), Kluwer Academic Publishers, Boston, 245-260, 1999.

[Giorgi, 2001] G.M. Giorgi: Corrado Gini (1884-1965). IN: Statisticians of the Centuries (Hyde C. and Seneta E. editors), Springer, New York, 364-368, 2001.

[Lorenz, 1905] M.O. Lorenz: Methods of measuring the concentration of wealth, Quarterly Publications of the American Statistical Association, 9 (new series), 209-219, 1905.

[Pareto, 1895] V. Pareto: La legge della domanda, Giornale degli Economisti, 10, 58-69, 1895.

[Pareto, 1897] V. Pareto: Aggiunta allo studio sulla curva delle entrate, Giornale degli Economisti, 14, 15-26, 1897.

[Pietra, 1915] G. Pietra: Delle relazioni tra indici di variabilità (I, II), Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti, 74, 775-804, 1915.

[Pietra, 1917] G. Pietra: Recenti pubblicazioni di metodologia statistica: <Ricci U.(1916). L'indice di variabilità e la curva dei redditi>, Rivista Italiana di Sociologia, 21(2-3), 310-319, 1917.

[Pietra, 1931a] G. Pietra: Observations à la communication du Prof. L.von Bortkiewicz, Bulletin de l'Institut International de Statistique, XIX Session (Tokyo,1930), 25(3), 310, 1931.

[Pietra, 1931b] G. Pietra: Note, Bulletin de l'Institut International de Statistique, XIX Session (Tokyo,1930), 25(3), 320L-320M, 1931.

[Regazzini, 1997] E. Ragazzini: Gini Corrado. IN: Leading personalities in statistical sciences (Johnson N.L. and Kotz S. editors), J. Wiley & Sons, New York, 291-296, 1997.