



On the Gini measure decomposition

P. Giudici, E. Raffinetti

► To cite this version:

P. Giudici, E. Raffinetti. On the Gini measure decomposition. Statistics and Probability Letters, 2010, 81 (1), pp.133. 10.1016/j.spl.2010.10.005 . hal-00698845

HAL Id: hal-00698845

<https://hal.science/hal-00698845>

Submitted on 18 May 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

On the Gini measure decomposition

P. Giudici, E. Raffinetti

PII: S0167-7152(10)00281-6
DOI: [10.1016/j.spl.2010.10.005](https://doi.org/10.1016/j.spl.2010.10.005)
Reference: STAPRO 5817

To appear in: *Statistics and Probability Letters*

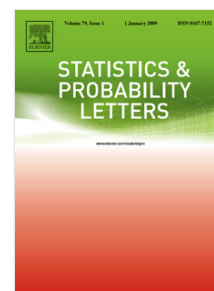
Received date: 23 April 2010

Revised date: 4 October 2010

Accepted date: 6 October 2010

Please cite this article as: Giudici, P., Raffinetti, E., On the Gini measure decomposition.
Statistics and Probability Letters (2010), doi:10.1016/j.spl.2010.10.005

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



On the Gini measure decomposition

P. Giudici^{a,*}, E. Raffinetti^{b,**}

^a*Statistics and Applied Economics Department, University of Pavia, Strada Nuova 65,
27100 Pavia, Italy*

^b*Statistics and Applied Economics Department, University of Pavia, Strada Nuova 65,
27100 Pavia, Italy*

Abstract

The purpose of this research is to introduce a new approach to the decomposition of the Gini measure in terms of concordance and discordance shares: a new kind of dependence, the *Gini rank dependence* (*GRD*), and its formal definition are provided.

Keywords: Gini measure, concordance curve, Gini rank dependence.

1. Introduction

The Gini measure decomposition always assumed the role of partitioning the total inequality of a population into two components, concerning the inequality between and within subpopulations. Theil (1967), partitioning the total population into h subpopulations, decomposed the total inequality T into the inequality *within* (T_w) and *between* (T_b) the h subpopulations, such that $T = T_w + T_b$ (where T_b is Theil inequality between the income means of the h subpopulations weighted by the subpopulations sizes). This decomposition approach stimulated further research: for example, researchers concentrated on the Gini ratio, deriving important transformations to capture the idea of decomposability, see, for example, Rao (1969) and Yitzhaki (1994). More recently Dagum (1997) suggested to decompose the Gini ratio into three components: the Gini inequality *within* the subpopulations (G_w), the Gini inequality *between* subpopulations (G_b) and the intensity of transvariation between subpopulations. According to Gini (1959) two

*Corresponding author

**Principal corresponding author

Email addresses: giudici@unipv.it (P. Giudici), emanuela.raffinetti@unipv.it (E. Raffinetti)

groups G_1 and G_2 are said to transvary if there exists at least one pair of units, belonging to different groups, such that the sign of the difference between their values is opposite to that of $m_1 - m_2$, where m_1 and m_2 are the corresponding group mean values.

Our research aim consists in proposing a new approach devoted to employ the Gini measure decomposition in terms of concordance and discordance. Section 2 describes the main statistical tools useful in characterizing this new approach. Section 3 is focused on defining a new kind of dependence, the *Gini rank dependence* (*GRD*), whose formal definition is provided. Finally, Section 4 is devoted to the conclusion and further reaserch developments.

2. Background

The aim of this section is introducing the main topics concerning the statistical tools needed in obtaining the Gini measure decomposition in terms of multivariate concordance and discordance.

Muliere et al. (1992) have illustrated these topics in order to define a new multidimensional concordance index through the employment of the Lorenz curves: in this context of analysis, we now recall all the background elements that allow to establish the decomposition of the Gini measure through the so called concordance curve.

Let us suppose to consider a k -variate random vector $(Y, X_1, X_2, \dots, X_{k-1})$, on which one can apply a model able to describe the relation among a response variable Y and the explanatory variables X_1, X_2, \dots, X_{k-1} . For this purpose one can recur, for example, to the linear regression function

$$E(Y|X_1, X_2, \dots, X_{k-1}) = \hat{Y}. \quad (1)$$

Given the response variable Y , our starting point is based on building its Lorenz curve, denoted with L_Y , and its dual, denoted with L'_Y : the former is characterized by the set of ordered pairs $(i/n, S_i/(nM_Y))$, $i = 1, \dots, n$, where $S_i = \sum_{j=1}^i y_{(j)}$, denoting the sum of the y_i ordered in an increasing sense and M_Y is the Y variable mean. The latter is characterized by the set of ordered pairs $(i/n, S'_i/(nM_Y))$, where $S'_i = \sum_{j=1}^i y_{(n+1-j)}$ denotes the sum of the y_i ordered in a decreasing sense. Once computed the estimated Y values, \hat{y}_i , through (1), one can proceed by the construction of the concordance curve based on ordering the Y values with respect to the ranks assigned by its estimated values \hat{y}_i . Let us denote this ordering with $(y_i|r(\hat{y}_i))$ and, more specifically, by y_i^* . The set of pairs $(i/n, (1/(nM_Y)) \sum_{j=1}^i y_j^*)$

defines the concordance curve, denoted with $C(Y|r(\hat{y}_i))$.

Through a direct comparison between the set of points that represent the Y Lorenz curve, L_Y , and the set of points that represent the concordance curve, $C(Y|r(\hat{y}_i))$, one can show that a perfect "overlap" is provided only if

$$\sum_{j=1}^i y_{(j)} = \sum_{j=1}^i y_j^* \text{ for every } i = 1, 2, \dots, n, \quad (2)$$

that is, if and only if $r(y_i) = r(\hat{y}_i)$. The further comparison between the set of points that represent the Y dual Lorenz curve, L'_Y , and the set of points that represent the concordance curve, $C(Y|r(\hat{y}_i))$, allows to conclude that there is a perfect "overlap" if and only if

$$\sum_{j=1}^i y_{(n+1-j)} = \sum_{j=1}^i y_j^* \text{ for every } i = 1, 2, \dots, n. \quad (3)$$

that is, if and only if $r(y_{(n+1-i)}) = r(\hat{y}_i)$

Note that the egalitarian line, which represents the bisector of the unit side square in which the Lorenz and the dual Lorenz curves lie, splits the Gini measure in two equal parts: we call the upper area, located between the egalitarian line and the Y dual Lorenz curve, the *discordance area* while the lower area, located between the egalitarian line and the Y Lorenz curve, the *concordance area*.

3. The Gini measure decomposition: a proposal

This section is devoted to the decomposition of the Gini measure. The role of the Gini measure consists in describing the statistical dispersion and, therefore, one can consider it as a variability measure.

3.1. The concentration curve, its dual and the Gini measure

Let us start our analysis providing the definition of the Gini measure. To help our illustration, Figure 1, represents an example of a Lorenz and a dual Lorenz curve.

The Gini measure is defined as the ratio of the areas on the Lorenz curve diagram. If the area between the egalitarian line and the Lorenz curve is A (see Figure 1), and the area under the Lorenz curve is B , then the Gini measure is $A/(A + B)$. Since $A + B = 0.5$, the Gini measure can be defined

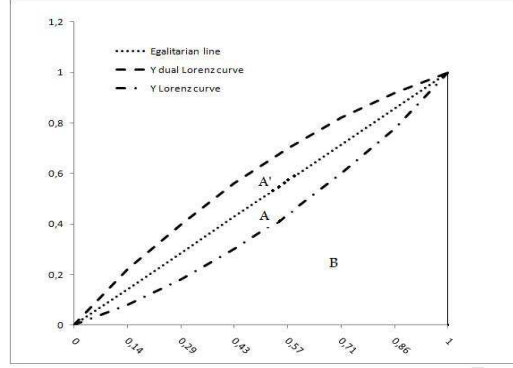


Figure 1: The Gini measure diagram: Gini measure = $A' + A$.

as $G = A/(0.5) = 2A = 1 - 2B$. If the Lorenz curve of the Y variable is denoted with $L_Y(t)$, the value of B can be found by integration and

$$G = 1 - 2 \int_0^1 L_Y(t) dt, \text{ with } 0 \leq t \leq 1. \quad (4)$$

It is trivial to establish that if the concordance curve $C(Y|r(\hat{y}_i))$ is close to the Y dual Lorenz curve, the relationship between the response variable Y and the estimated response \hat{Y} based on the explanatory variables, points towards discordance, in the other case towards concordance. For this reason we call the area between the Y dual Lorenz curve and the egalitarian line "discordance area" and the area between the Y Lorenz curve and the egalitarian line "concordance area".

Let us now provide the measures associated to these two different areas. The concordance area (CA) can be defined as follows:

$$CA = \frac{1}{2} - \int_0^1 L_Y(t) dt \text{ with } 0 \leq t \leq 1; \quad (5)$$

on the other hand, being the Y dual Lorenz curve, denoted with $L'_Y(t)$, equivalent to $1 - L_Y(1 - t)$ (see e.g. Koshevoy et al. (1996)), with $0 \leq t \leq 1$, the discordance area (DA) can be computed as follows:

$$\begin{aligned} DA &= \int_0^1 L'_Y(t) dt - \frac{1}{2} = \int_0^1 [1 - L_Y(1 - t)] dt - \frac{1}{2} \\ &= 1 - \int_0^1 L_Y(1 - t) dt - \frac{1}{2} = \frac{1}{2} - \int_0^1 L_Y(1 - t) dt. \end{aligned} \quad (6)$$

The relationship between the Gini measure and the discordance and concordance areas can thus be expressed as

$$\begin{aligned} G &= CA + DA = \frac{1}{2} - \int_0^1 L_Y(t)dt + \frac{1}{2} - \int_0^1 L_Y(1-t)dt \\ &= 1 - \int_0^1 L_Y(t)dt - \int_0^1 L_Y(1-t)dt = 1 - 2 \int_0^1 L_Y(t)dt, \end{aligned} \quad (7)$$

since $\int_0^1 L_Y(1-t)dt = \int_0^1 L_Y(t)dt$, for all $0 \leq t \leq 1$. Note that, in particular, the maximum concordance and discordance areas are equal to each other and assume value $G/2$.

A direct implication of the previous definition is that the Gini measure corresponds to the area between the Y variable Lorenz curve and the Y variable dual Lorenz curve, that is the area given by the sum of A and A' in Figure 1.

3.2. The Gini rank dependence

Our main purpose is now decomposing the Gini measure in terms of concordance and discordance. More precisely, we define the share of the Gini measure which corresponds to a concordance or to a discordance situation between Y and \hat{Y} . This kind of proceeding implies the study of a new form of dependence that we call *rank dependence*: the measure associated to this dependence form will be called *Gini rank dependence* and denoted with *GRD*.

Our proposal can be explained considering three different cases, illustrated in Figure 2, Figure 3 and Figure 4.

Case 1: the concordance curve $C(Y|r(\hat{y}_i))$ completely lies in the concordance area.

Our aim is to measure the "concordance" degree: in other terms, to define the measure of the Y variable concentration "explained" by the Y rank dependence with respect to the explanatory variables.

The Gini measure, explained by the rank dependence (*GRD*), assumes the following expression

$$GRD = \frac{1}{2} - \int_0^1 C(Y|r(\hat{y}_i))dy : \quad (8)$$

which can be normalized by dividing it by $2G$.

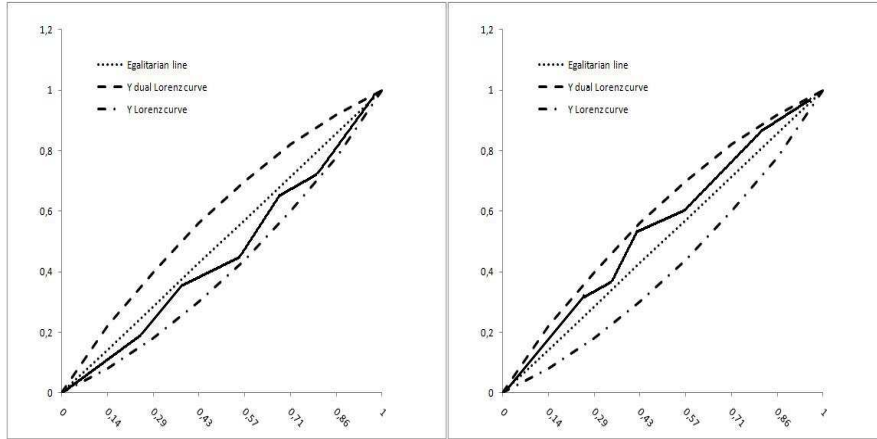


Figure 2: a) Case 1 and b) Case 2. The concordance curve is the continuous curve.

Case 2: the concordance curve $C(Y|r(\hat{y}_i))$ completely lies in the discordance area.

In this case the aim is the mirror image of the previous one, in fact one has to define the measure of the Y variable concentration "explained" by the Y rank dependence with respect to the explanatory variables, so

$$GRD = \int_0^1 C(Y|r(\hat{y}_i))dy - \frac{1}{2}, \quad (9)$$

which can be normalized by dividing it by $2G$.

Case 3: the concordance curve $C(Y|r(\hat{y}_i))$ partially lies in the concordance area and partially in the discordance area meaning that between the concordance curve and the egalitarian line there are one or more intersection points. This case is a little more complex than the first two and deserves a further development: we need to describe the main conditions that can occur in relation to the number of intersection points, between the concordance curve and the egalitarian line, and the concordance curve initial position with respect to the concordance or discordance area. More precisely, in order to measure the concordance and discordance areas, one has to consider the following steps:

1. identifying the number of intersection points between the concordance curve and the egalitarian line;

2. defining the concordance and discordance "extent" recurring to a series of integrals whose integration extremes are represented by the intersection points x -axis values.

Through some examples we can get the "size" of Gini measure explained by the rank dependence in terms of concordance and discordance, distinguishing between two main situations, as follows.

Even number of intersection points. Let us start considering an even number of intersection points: for instance let us suppose that the concordance curve intersects the egalitarian line in two points, A and B , respectively of x -axis values a_1 and a_2 .

The subsequent step concerns the identification of the the first segment position of the concordance curve with respect to the concordance or discordance area: let us denote with Con the concordance area share and with Dis the discordance area share, whose bounds are defined by the concordance and the egalitarian curves. Two subcases have to be taken into account.

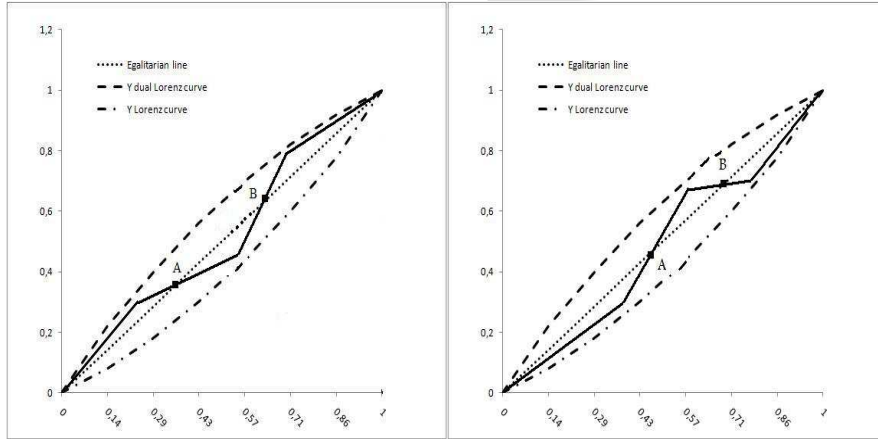


Figure 3: a) Subcase 1: even number of intersection points and first segment of the concordance curve located in the discordance area and b) Subcase 2: even number of intersection points and first segment of the concordance curve located in the concordance area.

Subcase 1: the first segment of the concordance curve lies in the discordance area.

$$\text{Since } Con = \int_{a_1}^{a_2} t dt - \int_{a_1}^{a_2} C(Y|r(\hat{y}_i)) dy, \text{ with } 0 \leq t \leq 1, \text{ and } Dis =$$

$\left[\int_0^{a_1} C(Y|r(\hat{y}_i))dy - \int_0^{a_1} tdt \right] + \left[\int_{a_2}^1 C(Y|r(\hat{y}_i))dy - \int_{a_2}^1 tdt \right]$ the measure of the Y variable concentration "explained" by the Y rank dependence with respect to the explanatory variables is obtained by the following expression:

$$GRD = Con + Dis = \int_0^{a_1} C(Y|r(\hat{y}_i))dy - \int_{a_1}^{a_2} C(Y|r(\hat{y}_i))dy + \int_{a_2}^1 C(Y|r(\hat{y}_i))dy + \left[-a_1^2 + a_2^2 - \frac{1}{2} \right].$$

Subcase 2: the first segment of the concordance curve lies in the concordance area.

The GRD is obtained by the following expression:

$$GRD = Con + Dis = - \int_0^{a_1} C(Y|r(\hat{y}_i))dy + \int_{a_1}^{a_2} C(Y|r(\hat{y}_i))dy + \int_{a_2}^1 C(Y|r(\hat{y}_i))dy + \left[a_1^2 - a_2^2 + \frac{1}{2} \right],$$

being $Con = \left[\int_0^{a_1} tdt - \int_0^{a_1} C(Y|r(\hat{y}_i))dy \right] + \left[\int_{a_2}^1 tdt - \int_{a_2}^1 C(Y|r(\hat{y}_i))dy \right]$ and $Dis = \int_{a_1}^{a_2} C(Y|r(\hat{y}_i))dy - \int_{a_1}^{a_2} tdt$.

Odd number of intersection points. Let us continue considering an odd number of intersection points: for instance let us suppose that the concordance curve intersects the egalitarian line in three points, A , B and C respectively of x -axis values a_1 , a_2 and a_3 . Then, as described previously, we have to proceed to the identification of the first segment position of the concordance curve with respect to the concordance or discordance area. We have to take into account two subcases, as before.

Subcase 1: the first segment of the concordance curve lies in the discordance area.

The measure of the Y variable concentration "explained" by the Y rank dependence with respect to the explicative variables is obtained by the following expression:

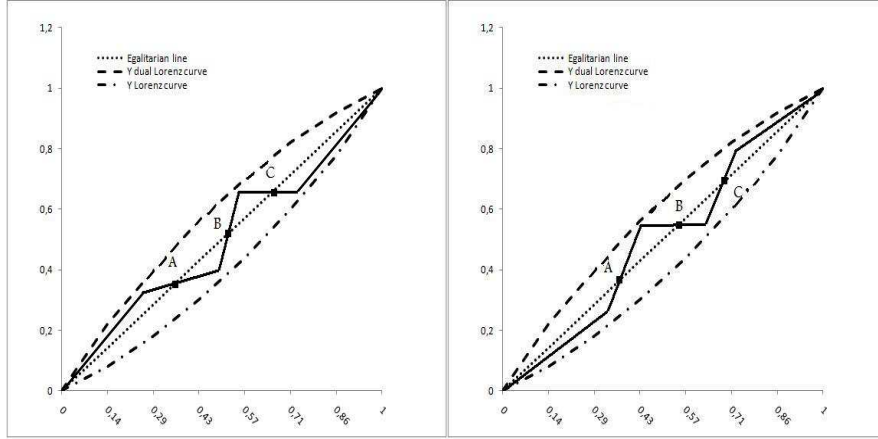


Figure 4: a) Subcase 1: odd number of intersection points and first segment of the concordance curve located in the discordance area, b) Subcase 2: odd number of intersection points and first segment of the concordance curve located in the concordance area.

$$GRD = Con + Dis = \int_0^{a_1} C(Y|r(\hat{y}_i))dy - \int_{a_1}^{a_2} C(Y|r(\hat{y}_i))dy + \int_{a_2}^{a_3} C(Y|r(\hat{y}_i))dy - \int_{a_3}^1 C(Y|r(\hat{y}_i))dy + \left[-a_1^2 + a_2^2 - a_3^2 + \frac{1}{2} \right],$$

where $Con = \left[\int_{a_1}^{a_2} tdt - \int_{a_1}^{a_2} C(Y|r(\hat{y}_i))dy \right] + \left[\int_{a_3}^1 tdt - \int_{a_3}^1 C(Y|r(\hat{y}_i))dy \right]$ and $Dis = \left[\int_0^{a_1} C(Y|r(\hat{y}_i))dy - \int_0^{a_1} tdt \right] + \left[\int_{a_2}^{a_3} C(Y|r(\hat{y}_i))dy - \int_{a_2}^{a_3} tdt \right]$.

Subcase 2: the first segment of the concordance curve lies in the concordance area.

The GRD is obtained by the following expression:

$$GRD = Conc + Dis = - \int_0^{a_1} C(Y|r(\hat{y}_i))dy + \int_{a_1}^{a_2} C(Y|r(\hat{y}_i))dy - \int_{a_2}^{a_3} C(Y|r(\hat{y}_i))dy + \int_{a_3}^1 C(Y|r(\hat{y}_i))dy + \left[a_1^2 - a_2^2 + a_3^2 - \frac{1}{2} \right],$$

being $Con = \left[\int_0^{a_1} tdt - \int_0^{a_1} C(Y|r(\hat{y}_i))dy \right] + \left[\int_{a_2}^{a_3} tdt - \int_{a_2}^{a_3} C(Y|r(\hat{y}_i))dy \right]$ and $Dis = \left[\int_{a_1}^{a_2} C(Y|r(\hat{y}_i))dy - \int_{a_1}^{a_2} tdt \right] + \left[\int_{a_3}^1 C(Y|r(\hat{y}_i))dy - \int_{a_3}^1 tdt \right]$.

3.3. The GRD generalization formula

The previous construction inductively suggests to find a general formulation of *GRD*. Through the computations related to the *GRD* definition, one can obtain a recursive form consisting in alternate signs in integral terms and in constant terms, represented by the intersection points x -axis values: this recursive formula depends either on the nature of the intersection points number as well as on the first segment position of the concordance curve with respect to the concordance or discordance area.

Let us suppose to have p intersection points (whose x -axis values can be denoted with a_1, a_2, \dots, a_p).

Before proceeding we have to define some conditions:

- $a_j = 1$ with $j = p + 1$;
- $a_{j-1} = 0$ with $j = 1$.

In order to achieve a general formulation, we need to take into account the difference between the even or odd number of the intersection points but also the first segment concordance curve position with respect to the concordance/discordance area. We thus need to consider a further term in the expression characterizing *GRD*. This term is a multiplicative factor equivalent to $(-1)^s$, where s can assume only two values 0 or 1: in particular

- if $s = 0$, then the first segment of the concordance curve is located in the discordance area;
- if $s = 1$, then the first segment of the concordance curve is located in the concordance area.

In conclusion, *GRD* can be defined as:

$$GRD = (-1)^s \left\{ \sum_{j=1}^{p+1} (-1)^{j+1} \left[\int_{a_{j-1}}^{a_j} C(Y|r(\hat{y}_i)) dy - a_j^2 \right] + \frac{1}{2} \right\}, \quad (10)$$

for p even and

$$GRD = (-1)^{s+p} \left[\sum_{j=1}^{p+1} (-1)^{j+1} \left(a_j^2 - \int_{a_{j-1}}^{a_j} C(Y|r(\hat{y}_i)) dy \right) \right] - \frac{1}{2} (-1)^s, \quad (11)$$

for p odd.

Note that the extreme cases, characterized by a null number of intersection points, occur when the concordance curve completely lies in the concordance area (Case 1) or in the discordance area (Case 2). In order to conduct these cases under the general formulation, we have to determine the nature of the 0 number. Pitaghora considers this number neither even nor odd even if the number 0 satisfies some typical even numbers conditions. Indeed to support considering the number 0 as even, in our context, note that the existence of an even number of intersection points implies that neither the first nor the last segments of the concordance curve lie in the same concordance or discordance area. Since our current discussion regards the cases where the concordance curve does not present any position reversal (in particular any reversal in the first or in the last segment) we will thus recur to the generalized expression concerning p even.

Now if we let:

- $a_{j-1} = 0$ if $j = 1$, then $a_0 = 0$;
- $a_j = 1$ if $j = p + 1$, then $a_1 = 1$,

we can show that the two extreme cases are special cases of the general formulation:

Case 1: the concordance curve $C(Y|r(\hat{y}_i))$ completely lies in the concordance area $\Rightarrow s = 1$.

$$\begin{aligned} GRD &= (-1)^1 \left\{ \sum_{j=1}^1 (-1)^{j+1} \left(\int_0^1 C(Y|r(\hat{y}_i)) dy - a_1 \right) + \frac{1}{2} \right\} \\ &= (-1)^1 \left\{ \int_0^1 C(Y|r(\hat{y}_i)) dy - 1 + \frac{1}{2} \right\} \\ &= \frac{1}{2} - \int_0^1 C(Y|r(\hat{y}_i)) dy \quad \square \end{aligned}$$

Case 2: the concordance curve $C(Y|r(\hat{y}_i))$ completely lies in the discordance area $\Rightarrow s = 0$

$$\begin{aligned}
 GRD &= (-1)^0 \left\{ \sum_{j=1}^1 (-1)^{j+1} \left(\int_0^1 C(Y|r(\hat{y}_i)) dy - a_1^2 \right) + \frac{1}{2} \right\} \\
 &= (-1)^2 \left(\int_0^1 C(Y|r(\hat{y}_i)) dy - 1 \right) + \frac{1}{2} = \int_0^1 C(Y|r(\hat{y}_i)) dy - \frac{1}{2} \quad \square
 \end{aligned}$$

4. Conclusions

This study has introduced a new approach to the decomposition of the Gini measure into a concordance share and a discordance share. This in order to obtain the “quota” of the the response variable Y concentration explained by \hat{Y} , function of the explanatory variables X_1, \dots, X_{k-1} .

We remark that our proposal can be applied to the model assessment context, particularly when one wants to compare alternative classifiers (e.g. logistic regression, tree models). In this context, a frequent model performance measure is the area under the ROC curve (AUC). It can be shown that the ROC curve is the equivalent to the Lorenz dual curve (see e.g. Hand (2009)). AUC then corresponds to the Gini measure defined above (see e.g. Bradley (1997), Fawcett (2006), Hand (2009), Krzanowski et al. (2009)).

In this respect, in the paper we have provided a way to decompose the contribution to the AUC measure in terms of the contributions of the different observed units.

Another application of our approach can be as an useful measure of fit when the relevant dependent variables are of qualitative nature (nominal or ordinal): in the classical literature one can recur to the measure of fit represented by R^2 and residual analysis, but when qualitative variables are involved and prevalent in the explanation of the response variable, R^2 , based on the euclidean distance, may not be appropriate. Instead one can recur to the Gini measure in place of R^2 , and to its decomposition, in place of residual analysis. In particular, one can split the response variable values in deciles (obtained by ordering these values in an increasing sense), one can establish the position of each value, with respect to the ranks assigned by its corresponding regression estimates, in order to define the single contribution to the concordance or discordance. If the position of the observed variable values are different with respect to the fitted ones, one can conclude that the predictive selected model does not allow a good fitting.

Finally we would like to remark that our approach is nonparametric and

can thus be applied, generally, for all predictive models. A relevant further research topic could concern the definition of the concordance curve analytical expression, characterized as a functional of order statistics, under alternative parametric assumptions.

5. Acknowledgements

The authors acknowledge financial support from the European grant EU-IP MUSING (contract number 027097).

The paper is the result of the close collaboration among the authors, however, it has been written by Emanuela Raffinetti with the supervision of Paolo Giudici.

The authors wish to thank the Associate Editor for his helpful comments.

References

- [1] Bradley, A. P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30, 1145–1159.
- [2] Dagum, C., 1997. A new Approach to the Decomposition of the Gini Income Inequality Ratio. *Empirical Economics* 22, 515–531.
- [3] Fawcett, T., 2006. An introduction to ROC analysis. *Pattern Recognition*, 27, 861–874.
- [4] Gini, C., 1959. *Transvariation*, Rome (in Italian). Libreria Goliardica.
- [5] Hand, D. J., 2009. Measuring classifier performance: a coherent alternative to the area under the ROC curve. *Mach Learn*, 77, 103–123.
- [6] Koshevoy, G., Mosler, K., 1996. The Lorenz Zonoids of a Multivariate Distribution. *Journal of the American Statistical Association*, 91, No. 434, Theory and Methods.
- [7] Krzanowski, W. J., Hand, D. J., 2009. *ROC curves for continuous data*. London: Chapman and Hall.
- [8] Muliere, P., Petrone, S., 1992. Generalized Lorenz curve and monotone dependence orderings. *Metron*, L, n. 3–4.
- [9] Rao, V. M., 1969. Two decompositions of concentration ratio. *Journal of the Royal Statistical Society, CXXXII, A*, 418–425.

- [10] Theil, H., 1967. Economics and information Theory. North-Holland Publishing Company, Amsterdam.
- [11] Yitzhaki, S., 1994. Economic distance and overlapping of distributions. Journal of Econometrics, 61, 147–159.