isid/ms/2013/08
August 2, 2013
http://www.isid.ac.in/~statmath/eprints

A simple non-parametric estimator of the Gini index

Sudheesh K K and Isha Dewan

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

A SIMPLE NON-PARAMETRIC ESTIMATOR OF THE GINI INDEX

Sudheesh K K[†,*] and Isha Dewan[**]

[*] Indian Statistical Institute, Chennai, India..

[**] Indian Statistical Institute, New Delhi, India.

Abstract. Search for simple reliable estimators of the Gini index has lead to numerous publications. Most of these papers focus on the bias and the standard error of the estimator of the Gini index. In this paper, we propose a simple estimator of Gini index based on U-statistics. A simulation study shows that our estimator performs 'well' compared to other estimators in terms of bias and mean squared error.

1. Introduction

The celebrated Gini index is a widely used indicator of income inequality in a population. The estimation problem of Gini index mainly concentrated on finding plug-in estimators of the Gini index with reliable standard errors. This is achieved by expressing Gini index in different forms (see Yitzhaki (1998)) and then find the plug-in estimator with minimal distributional assumptions. Hence, the discussion of the sampling variance of the Gini index has produced vast amount of research in statistics and economics. Some recent references of interest are Bhattacharya (2007), Xu (2007), Davidson (2009), Peng (2011), Ceriani and Verme (2012) and Langel and Tille (2013).

Even though the inference on Gini index has been widely discussed, most of the estimators are either complicated or the bias and standard error are unreliable. Calculation of standard error is essential when confidence intervals or tests are to be constructed for this coefficient. Ogwang (2000) have

† *Corresponding author*, E-mail: `skkattu@yahoo.co.in`.

proposed the jackknife technique to get a large-sample approximation for the standard error.

However, as the Gini coefficient can be obtained from a simple ordinary least square regression based approach (see Ogwang (2000)), Giles (2004) noted that the adoption of the jackknife technique is unnecessary, and the construction of an appropriate standard error for the Gini coefficient is trivial. In fact, Giles (2004) claimed that the ordinary least square standard error from this regression could be used directly in order to compute the standard error of the Gini index itself. See Ogwang (2004, 2006) and Giles (2006) for further discussion and criticism on this point.

Modarres and Gastwirth (2006) hit a cautionary note on the use of Giles's approach, showing by simulation that the standard errors it produces are quite inaccurate. Also they recommend to use the complex or computationally intensive methods used previously. Based on sample empirical process theory and the functional delta method, Bhattacharya (2007) has developed techniques of asymptotic inference for Lorenz curves and the Gini index with stratified and clustered survey data. Davidson (2009) noted that these approach produce a formula for the variance of an estimated Gini index, which is very difficult to implement. Using delta method, Davidson (2009) showed how to compute an asymptotically correct standard error for an estimated Gini index. Davidson (2009) also discussed the use jackknife and bootstrap method for bias correction and variance estimation and noted that bootstrap can yield reasonably reliable inference compared to former.

The survey paper by Xu (2004) pointed out that there is a clear separation between publications from the field of statistics and publications in economic journals. He noted that the papers from one field are seldom cited in the other, it seems evident that researchers from these two fields do not necessarily read each others work. Xu (2007) has given overview of the use of U-statistics in inference for the Gini index. Ceriani and Verme (2012) have given an overview on the origin and development of Gini index and given complete review of all the expressions of the Gini index. Recently, Langel and Tille (2013) have discussed the variance estimation of the Gini index in historical point of view. They reviewed several linearization methods for approximating the variance of a non-linear statistic.

The estimation of the Gini index based on U-statistic has long history since Halmos (1946) and Hoeffding (1948). Hoeffding (1948) expressed the Gini index in terms of two U-statistics and then studied the asymptotic properties. This idea was adopted by Glasser (1962) and Gaswirth (1972) for studying both the Lorenz curve and the Gini index. Other papers on similar lines are due to Gail and Gastwirth (1978), Sandstrom et al. (1988), Bishop et al. (1997, 1998, 2001), Xu and Osberg (2002). Zheng et al. (2000) and Biewen (2002) considered statistical inference for the Gini index.

We discuss a non-parametric estimator of the Gini index using U-statistics with motivation from the papers of Xu (2007), Davidson (2009) and Langel and Tille (2013).

The rest of the paper is organized as follows. In Section 2, we express the Gini index in terms of the expectation of maximum of pairs of random variables. We use the expression to obtain a simple estimator of the Gini index based on U-statistics. The asymptotic properties of the estimator are studied in Section 3. In Section 4, we carry out a simulation study to find the bias and the standard error of the proposed estimator and compare it to the other estimators available in literature.

2. **A simple estimator of Gini index**

Gini index is usually defined either by the Lorenz curve or by covariance identity involving cumulative distribution function. Let X be a random variable representing the income of a member of the population under study, clearly $0 < X < \infty$. Let F be the absolutely continuous cumulative distribution function of X. Assume that X has finite mean μ, given by

$\mu = \int_0^\infty y dF(y)$. The Lorenz curve denoted by $L(.)$ is defined by the equation

$$L(F(x)) = \frac{1}{\mu}\int_0^x y dF(y). \tag{1}$$

Then Gini index is defined as

$$\begin{aligned} G &= 2\int_0^1 (z - L(z)))dz \\ &= 1 - 2\int_0^1 L(z)dz. \end{aligned} \tag{2}$$

In fact, from equation (2), it can be seen that the Gini index is the twice the area between the Lorenz curve and egalitarian line. By simple algebra, we can show that the expression given in (2) is same as

$$G = \frac{2}{\mu}\int_0^\infty \int_0^x yF(y)dF(y) - 1.$$

One can also write the above expression as follows:

$$G = \frac{2}{\mu} Cov(X, F(X)).$$

That is, for given F, the Gini index is simply the covariance between X and $F(X)$. We use this form for finding a simple estimator of Gini index. The choice of an estimator is based on the research objective. If one is mainly interested in studying the stochastic dominance, then the natural choice is the area based formula (Atkinson and Bourguignon (2000)). However, if the interest lies in the decomposition of the population Gini, then it is more convenient to use the covariance based estimator. Consider

$$\begin{aligned} G &= \frac{2}{\mu} Cov(X, F(X)) \\ &= \frac{2}{\mu} E((X - \mu)F(X))) \\ &= \frac{2}{\mu} E(XF(X)) - 1 \\ &= \frac{2}{\mu}\int_0^\infty yF(y)dF(y) - 1. \end{aligned} \tag{3}$$

Let $X_1, X_2, ..., X_n$ be n independent and identical copies of X. Then the distribution of $X_{(n)} = max(X_1, X_2, ..., X_n)$, is given by

$$F_{X_{(n)}}(x) = (F(x))^n.$$

Hence

$$E(X_{(n)}) = n \int_0^\infty y(F(y))^{n-1} dF(y).$$

In Particular, when $n = 2$

$$E(X_{(2)}) = 2 \int_0^\infty yF(y)dF(y). \tag{4}$$

Substituting (4) in (3), we find

$$G = \frac{E(X_{(2)})}{\mu} - 1. \tag{5}$$

The last expression is used to find a simple reliable estimator of the Gini index. Our approach is based on U-statistics. For a recent review on inference of Gini index based on U-statistics see Xu (2007).

Let $X_1, X_2, ..., X_n$ be i.i.d. random variables from F. Then, $U_1 = \frac{1}{n}\sum_{i=1}^n X_i$ is an unbiased and a consistent estimator of μ.

Consider the symmetric kernel $h(x_1, x_2) = max(x_1, x_2)$. Then a U-statistic estimate of $E(X_{(2)})$ is given by

$$U_2 = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} max(X_i, X_j).$$

Therefore a natural estimator of G is

$$\widehat{G} = \frac{U_2}{U_1} - 1. \tag{6}$$

After simplification, we obtain

$$\widehat{G} = \frac{2\sum_{i=1}^{n-1} ix_{(i+1)}}{(n-1)\sum_{i=1}^n X_i} - 1, \tag{7}$$

where $X_{(i)}$, $i = 1, 2, ...n$, is the ith order statistics based on i.i.d. observations $X_1, X_2, ..., X_n$.

Remark 2.1. The expression for the estimator given by (7) is used for calculation purposes. The expression (6) is convenient for studying the asymptotic properties of the estimator of Gini index.

Remark 2.2. When the mean income of the population is known, an unbiased estimator of the Gini index G is given by

$$\widehat{G}_u = \frac{U_2}{\mu} - 1.$$

3. Asymptotic Distribution of the estimator of Gini index

The proposed estimator of the Gini index given in (6) is based on U-statistics. Hence, we use the asymptotic theory of U-statistics (Lee (1990)) to discuss the limiting distribution of $\hat{G}$.

Theorem 3.1. *The distribution of $\sqrt{n}(U_2 - E(X_{(2)}))$, as $n \to \infty$, is Normal with mean zero and variance $4\sigma_1^2$ with*

$$\sigma_1^2 = (A) - (B)^2, \tag{8}$$

where

$$\begin{aligned} A &= \int_0^\infty x^2 F^2(x) dF(x) - 2\int_0^\infty xF(x)\Big\{\int_0^x ydF(y)\Big\}dF(x) \\ &\quad + \int_0^\infty \Big\{\int_0^x ydF(y)\Big\}^2 dF(x). \\ B &= \int_0^\infty xF(x)dF(x) - \int_0^\infty \Big\{\int_0^x ydF(y)\Big\}dF(x). \end{aligned}$$

Proof: We have

$$\sigma_1^2 = Var\big(h_1(X, X_2)\big),$$

where $h_1(x, X_2) = E(\max(x, X_2))$.

Consider

$$\begin{aligned} h_1(x,Y) &= E(h(x,Y)) = E(xI(Y<x) + YI(Y>x)) \\ &= xF(x) + \int_x^\infty y dF(y) \\ &= xF(x) + \mu - \int_0^x y dF(y). \end{aligned}$$

$$\begin{aligned} Var(h_1(X)) &= V\Big(XF(X) - \int_0^X y dF(y)\Big) \\ &= V(XF(X) - k(X)) \\ &= E(XF(X) - k(X))^2 - E^2(XF(X) - k(X)), \end{aligned} \tag{9}$$

where $k(x) = \int_0^x y dF(y)$. The first term in the equation (9) can be expanded as

$$E(XF(X) - k(X))^2 = E(X^2F^2(X)) - 2E(XF(X)k(X)) + E(k^2(X)). \tag{10}$$

Consider

$$E(X^2F^2(X)) = \int_0^\infty x^2F^2(x)dF(x) \tag{11}$$

$$E(XF(X)k(X)) = \int_0^\infty xF(x)\Big\{\int_0^x y dF(y)\Big\}dF(x) \tag{12}$$

and

$$E(k^2(X)) = \int_0^\infty \Big\{\int_0^x y dF(y)\Big\}^2 dF(x). \tag{13}$$

Substituting (11), (12) and (13) in (10), we obtain

$$\begin{aligned} E(XF(X) - k(X))^2 &= \int_0^\infty x^2F^2(x)dF(x) \\ &\quad -2\int_0^\infty xF(x)\Big\{\int_0^x y dF(y)\Big\}dF(x) \\ &\quad + \int_0^\infty \Big\{\int_0^x y dF(y)\Big\}^2 dF(x). \end{aligned} \tag{A}$$

Similarly,

$$E(XF(X) - k(X)) \quad = \quad \int_0^\infty xF(x)dF(x) - \int_0^\infty \Big\{ \int_0^x ydF(y) \Big\} dF(x) \quad (B)$$

Inserting (A) and (B) in (9), we find

$$\sigma_1^2 = Var(h_1(X)) = [(A) - (B)^2].$$

Then the result follows using the central limit Theorem of U-statistics.

Corollary 3.1.

$$\frac{\sqrt{n}}{\mu}(U_2 - E(X_{(2)})) \to N\big(0, \frac{4\sigma_1^2}{\mu^2}\big) \quad \text{as} \quad n \to \infty. \qquad (14)$$

The following lemma is useful for proving the asymptotic normality of $\widehat{G}$.

Lemma 3.1. *Suppose that V_n converges to V in distribution as $n \to \infty$ and $E(W_n - V_n)^2 \to 0$ as $n \to \infty$. Then W_n converges in distribution to V as $n \to \infty$.*

Theorem 3.2. $\sqrt{n}(\widehat{G} - G) \to N\big(0, 4\frac{\sigma_1^2}{\mu^2}\big)$ as $n \to \infty$, *provided* $P(U_1 = 0) = 0$.

Proof: Let $V_n = \frac{\sqrt{n}}{\mu}(U_2 - E(X_{(2)}))$ and $W_n = \sqrt{n}(\widehat{G} - G)$. Note that U_1 is a consistent estimator of μ. The proof is an immediate consequence of Lemma 3.1 and Theorem 3.1.

Theorem 3.3. $\widehat{\sigma_1^2} = \widehat{A} - \widehat{B}^2$, *where $\widehat{A}$ and $\widehat{B}$ as given in equations (15) and (16) below are consistent estimators of A and B, respectively.*

Proof: (A) can be written as

$$\frac{1}{3}\int_0^\infty x^2 dF^3(x) - \int_0^\infty x\Big\{ \int_0^x ydF(y) \Big\} dF^2(x) + \int_0^\infty \Big\{ \int_0^x ydF(y) \Big\}^2 dF(x).$$

Hence, a consistent estimator of A is

$$\begin{aligned}\widehat{A} &= \frac{1}{3}\sum_{i=1}^{n} x_{(i)}^2\Big[(\frac{i}{n})^3-(\frac{i-1}{n})^3\Big]-\sum_{j=1}^{n}\Big\{\frac{1}{n}\sum_{i=1}^{j}x_{(i)}\Big\}x_{(j)}\Big[(\frac{j}{n})^2-(\frac{j-1}{n})^2\Big]\\ &\quad +\frac{1}{n}\sum_{l=1}^{n}\Big\{\frac{1}{n}\sum_{i=1}^{l}x_{(i)}\Big\}\Big\{\frac{1}{n}\sum_{j=1}^{l}x_{(j)}\Big\}\\ &= \frac{1}{3n^3}\sum_{i=1}^{n}x_{(i)}^2(3i^2-3i+1)-\frac{1}{n^3}\sum_{j=1}^{n}\Big\{\sum_{i=1}^{j}x_{(i)}\Big\}x_{(j)}(2j-1)\\ &\quad +\frac{1}{n^3}\sum_{l=1}^{n}\Big\{\sum_{i=1}^{l}x_{(i)}\Big\}\Big\{\sum_{j=1}^{l}x_{(j)}\Big\}. \qquad (15)\end{aligned}$$

A consistent estimator of B is given by

$$\begin{aligned}\widehat{B} &= \frac{1}{2}\sum_{i=1}^{n}x_{(i)}\Big[(\frac{i}{n})^2-(\frac{i-1}{n})^2\Big]-\frac{1}{n^2}\sum_{j=1}^{n}\Big\{\sum_{i=1}^{j}x_{(i)}\Big\}\\ &= \frac{1}{2n^2}\sum_{i=1}^{n}x_{(i)}(2i-1)-\frac{1}{n^2}\sum_{j=1}^{n}\Big\{\sum_{i=1}^{j}x_{(i)}\Big\}. \qquad (16)\\ & \qquad (17)\end{aligned}$$

Remark 3.1. As $n \to \infty$, $\sqrt{n}(\widehat{G}-G)$ is a zero mean random variable, hence $\widehat{G}$ is asymptotically unbiased for G.

Remark 3.2. It is easy to see that $\widehat{G}$ is a consistent estimator of G.
In the next section we will compare these estimates in term of bias and mean square error (MSE) and noted that our estimate produce less bias comparing to other estimates available in literature.

4. **Simulation**

In this section, we carried out a simulation study to evaluate the performance of the estimator. We compare our estimator with that proposed by Davidson (2009) in terms of bias and MSE. As Davidson (2009) estimator outplays the other estimator in performance, we compare our estimator only with Davidson's estimator. The estimator obtained by Davidson (2009) is given by

$$\widehat{GD} = \frac{2}{n^2\widehat{\mu}} \sum_{i=1}^{n} x_{(i)}(i - \frac{1}{2}) - 1, \tag{18}$$

where $\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $x_{(i)}$, $i = 1, 2, ...n$, are the ith order statistics. Note that the above expression is equivalent to

$$\widehat{GD} = \frac{2\sum_{i=1}^{n} ix_{(i)}}{n\sum_{i=1}^{n} x_i} - \frac{(n+1)}{n}, \tag{19}$$

and Langel and Tille (2013) pointed out that the expression (15) can be found in Sen (1973).

In our study, we first generate random sample from the exponential distribution with cumulative distribution function $F(x) = 1 - \exp(-x), x \geq 0$. Note that the Gini index for this distribution is 0.5. To find the bias and the MSE, 10000 estimate of Gini index is obtained by taking the sample size 10, 25, 50, 75 and 100. The bias of our estimator when the sample size is 10 is 0.0009882803 and that of Davidson is -0.04911055. The MSE of our estimator is 0.009281472 and that of Davidson is 0.009929047. When the sample size is 100, the bias and MSE of our estimator are given by 0.0001007844 and 0.0008533806 respectively. The bias of Davidson estimator is -0.004900223 while the MSE is 0.0008604006. The similar behaviour shows in the case $n = 25, 50$ and 75. It is interesting to see that the bias and MSE are less in each of these cases so that our estimator perform better. The comparison is given in Table 1.

TABLE 1. Bias and MSE of estimate of Gini; Exponential distribution

Sample size	Bias	MSE	Bias(Davidson)	MSE(Davidson)
n=10	0.001364229	0.009411404	-0.04877219	0.01000046
n=25	0.001140615	0.003541781	-0.01890501	0.003620306
n=50	-0.0004959606	0.001677818	-0.01048604	0.001721097
n=75	0.0004012898	0.001132957	-0.006627073	0.001146863
n=100	0.0001007844	0.0008533806	-0.004900223	0.0008604006

The Pareto distribution is considered as the best model for income data as it capture heavy tail behaviour. For our study consider Pareto distribution with cumulative distribution function $F(x) = 1 - x^{\lambda}, x \geq 1, \lambda > 1$. The Gini index is $1/(2\lambda - 1)$. For different values of λ and for $n = 10$ and $n = 100$, the MSE and bias are given in Table 2 and 3. The bias of our estimator is less to that of Davidson estimator. For small values of λ MSE is also less.

Table 2. Bias and MSE of estimate of Gini; Pareto distribution

Value of λ	Bias	MSE	Bias (Davidson)	MSE(Davidson)
20	-0.0002880765	9.556528e-05	-0.002823371	8.531208e-05
10	-0.0007268856	0.0004351083	-0.005917355	0.0003870248
5	-0.003397838	0.002143008	-0.01416917	0.00192725
4	-0.006087871	0.003556633	-0.0197648	0.003241499
3	-0.01275642	0.007485633	-0.03148078	0.006922594
2	-0.04731666	0.01847661	-0.07591833	0.01891616
1.5	-0.1178946	0.04079824	-0.1561052	0.04615709

Table 3. Bias and MSE of estimate of Gini; Pareto distribution

Value of λ	Bias	MSE	Bias (Davidson)	MSE(Davidson)
20	-8.229567e-05	9.671476e-06	-0.000337883	9.586541e-06
10	-4.284207e-05	4.400358e-05	-0.0005687294	4.344956e-05
5	-0.0004339889	0.0002454435	-0.00154076	0.0002427485
4	-0.0008790373	0.000456129	-0.002298818	0.0004515792
3	-0.002024038	0.001153597	-0.004003797	0.001142656
2	-0.01104456	0.004570213	-0.01426745	0.004563271
1.5	-0.04198776	0.01207977	-0.04656788	0.01228006

Finally we compare our estimator when the sample came from lognormal distribution. We assume the mean is zero and the comparison is done for $\sigma = 0.5, 1, 1.5$. Corresponds to these values of σ, the true Gini index are 0.2763, 0.5205 and 0.7112. The Table 4 and 5 gives the comparison of bias

Table 4. Bias and MSE of estimate of Gini; Lognormal distribution

Value of σ	Bias	MSE	Bias (Davidson)	MSE(Davidson)
0.5	-0.003421048	0.004292636	-0.03070894	0.004410595
1	0.02509187	0.01302708	-0.07463268	0.015612
1.5	-0.06126468	0.01994048	-0.1262582	0.02905271

and MSE for $n = 10$ and $n = 100$ respectively. In this case also our estimator performs well in terms of bias and MSE.

Table 5. Bias and MSE of estimate of Gini; Lognormal distribution

Value of σ	Bias	MSE	Bias (Davidson)	MSE(Davidson)
0.5	-0.0004200164	0.0004135257	-0.003178816	0.0004152285
1	-0.003065953	0.001664766	-0.008240294	0.001690326
1.5	-0.01240101	0.003062635	-0.019389	0.003226897

5. Conclusion

In this paper we find a new expression for the Gini index in terms of the expectation of maximum of paired observations. This leads us to a simple estimator of the Gini index based on U-statistics. The theory of U-statistics is used to prove that the proposed estimator is consistent and asymptotically normal. We obtained an expression for asymptotic variance of the estimator of the Gini index. This asymptotic variance can easily be estimated. The proposed estimator has less bias and MSE compared to the other estimators available in literature.

References

Atkinson, A.B. and Bourguignon, F. (2000). Handbook of Income Distribution, vol. 1. Elsevier Science, Amsterdam.

Bhattacharya, D. (2007). Inference on inequality from household survey data. *Journal of Econometrics*, **137**, 674-707.

Biewen, M. (2002). Bootstrap inference for inequality, mobility, and poverty measurement. *Journal of Econometrics*, **108**, 317-342.

Bishop, J. A., Formby, J. P. and Zheng, B. (1997). Statistical inference and the Sen index of poverty. *International Economic Review*, **38** , 381-387.

Bishop, J. A., Formby, J. P., and Zheng B. (1998). Inference tests for Gini-based tax progressivity indexes. *Journal of Business and Economic Statistics*, **16**, 322-330.

Bishop, J. A., Formby, J. P., and Buhong Z. (2001). Sen measures of poverty in the United States: cash versus comprehensive incomes in the 1990s. *Pacific-Economic Review*, **6**, 193-210.

Ceriani, L. and Verme, P. (2012). The origins of the Gini index: extracts from Variabilit e Mutabilit (1912) by Corrado Gini. *Journal Economic Inequality*, **10**, 421-443.

Davidson, R. (2009). Reliable inference for the Gini index. *Journal of Econometrics*, **150**, 30-40.

Gail, M., and Gastwirth, J. L. (1978). A scale-free goodness-of-fit test for the exponential distribution based on the Gini statistics, *Journal of the Royal Statistical Society-Series B*, **40**, 350-357.

Gastwirth, J. L. (1972). The estimation of the Lorenz curve and the Gini Index. *Review of Economics and Statistics*, **54**, 306-316.

Giles, D. E. A. (2004). Calculating a standard error for the Gini coefficient: some further results. *Oxford Bulletin Economics and Statistics*, **66**, 425-433.

Giles,D. E. A. (2006). A cautionary note on estimating the standard error of the Gini index of inequality: comment. *Oxford Bulletin Economics and Statistics*, **68**, 395-396.

Glasser, G. J. (1962). Variance formulas for the mean difference and coefficient of concentration. *Journal of American Statistical Association*, **57**, 648-654.

Halmos, P. R. (1948). The theory of unbiased estimation. *Annals of Mathematical Statistics*, **17**, 34-43.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics*, **19**, 293- 325.

Langel, M and Tille, Y (2013). Variance estimation of the Gini index: revisiting a result several time published, *Journal of the Royal Statistical Society -Series A*, **176**, 521-540.

Lee, A.J (1990). *U-statistic: Theory and Practice.* CRC press, New York.

Modarres, R. and Gastwirth, J. L. (2006). A cautionary note on estimating the standard error of the Gini index of inequality. *Oxford Bulletin Economics and Statistics*, **68**,385-390.

Ogwang, T. (2000). A convenient method of computing the Gini index and its standard error. *Oxford Bulletin Economics and Statistics*, **62**, 123-129.

Ogwang, T. (2004). Calculating a standard error for the Gini coefficient: some further results: reply. *Oxford Bulletin Economics and Statistics*, **66**, 435-437.

Ogwang, T. (2006). A cautionary note on estimating the standard error of the Gini index of inequality: comment. *Oxford Bulletin Economics and Statistics*, **68**, 391-393.

Peng, L. (2011). Empirical likelihood methods for the gini index. *Australian and New Zealand Journal of Statistics*, **53**, 131139.

Sandstrom, A., Wretman, J. H. and Walden, B. (1988). Variance estimators of the Gini coefficientprobability sampling. *Journal of Business and Economic Statistics*, **6**, 113-119.

Schechtman, E. and Zitikis, R. (2006). Gini indices as areas and covariances: what is the difference between the two representations, *Metron- International Journal of Statistics*, **64**, 385-397.

Sen, A. (1973). *On Economic Inequality.* Oxford: Clarendon.

Xu, K. (2004). How has the literature on Ginis index evolved in the past 80 years? *Working Paper. Department of Economics, Dalhousie University*, Halifax.

Xu, K. (2007). U-statistics and their asymptotic results for some inequality and poverty measures. *Econometric Reviews*, **26**, 567-577.

Xu, K. and Osberg L. (2002). The Social Welfare Implications, Decomposability, and Geometry of the Sen Family of Poverty Indices. *Canadian Journal of Economics*, **35**, 138-152.

Yitzhaki, S. (1998). More than a dozen alternative ways of spelling Gini, *Research on Economic Inequality*, **8**, 13-30

Zheng, B., Formby, J.P. , Smith, W.J. and Chow, V.K. (2000). Inequality orderings, Normalized stochastic dominance and statistical inference. *Journal of Business and Economic Statistics*, **18**, 479-488.