Afrika Statistika
Vol. 18 (03), 2023, pages 3603 - 3621.
In: Special issue of the statistical applications of the fep and the *lrfep* to nonparametric estimation
DOI: https://dx.doi.org/10.16929/as/2023.3603.320.sFEP

Afrika Statistika

ISSN 2316-090X

Asymptotic representation of S-Gini index in the fep-lrfep and application

Esossolam Kpatcha [1,*], **Tchilabalo Abozou Kpanzou** [1,3], **Diam Ba** [2,3]

[1] LaMMASD, University of Kara, Kara, Togo
[2] Cheikh Anta Diop University, Dakar, Senegal
[3] Imhotep International Mathematical Center (imh-imc), https://imhotepsciences.org/

Received on January 01, 2023; Accepted on July 1, 2023; Published on September 01, 2024

Abstract In this paper we establish a general asymptotic representation (*GAR*) of the S-Gini index of inequality $J(\nu)$, $\nu \geq 2$, using the functional empirical process (*fep*) and the Lo's residual functional empirical process introduced by Lo and Sall (2010a) and Lo and Sall (2010b), abbreviated *lrfep*.

Résumé. Dans cet article, nous établissons une représentation asymptotique générale (*GAR*) de l'indice d'inégalité S-Gini, noté $J(\nu)$, $\nu \geq 2$, en utilisant le processus empirique fonctionnel (*fep*) et le processus empirique fonctionnel résiduel de Lo introduit par (Lo et Sall (2010a)), et (Lo et Sall (2010b)), abrégé *lrfep*.

Key words: functional empirical function; functional empirical process; index asymptotic representation; Functional Brownian Bridge.
AMS 2010 Mathematics Subject Classification Objects : 62E20, 62F05

* Correspondence Author: Esossolam Kpatcha (kpatcharodrigue2@gmail.com),
Tchilabalo Abozou Kpanzou: kpanzout@gmail.com
Diam Ba: diam.ba@ucad.edu.sn

Presentations of authors.

Esossolam Kpatcha, M.Sc., is preparing a Ph.D. degree at the LaMMASD, University of Kara, under the supervision of Tchilabalo Abozou Kpanzou.

Tchilabalo Abozou Kpanzou, Ph.D., is a Professor of Statistics and Mathematics at the University of Kara, Togo. He is a lead researcher at: Imhotep International Mathematical Center (imh-imc), https://imhotepsciences.org.

Diam Ba, Ph.D., is a Senior Lecturer of Statistics and Mathematics at the University of Cheikh Anta Diop, Senegal. He is a lead researcher at: Imhotep International Mathematical Center (imh-imc), https://imhotepsciences.org.

1. Introduction

It is common in Statistics to estimate parameters and to study the asymptotic behavior of the estimators. In that sense, approaches based on the functional empirical processes (*fep*) can be applied to a number of statistics and indexes, that are used in Mathematical Statistics and in many other disciplines (see, e.g., Zeng (1997); Foster *et al.* (1984); Kakwani (1980) and Sen (1976)).

The first attempt of having a common frame for the asymptotic theory of those statistics was done in Lo (2013), where the real-valued empirical processes are used. Later, the combination of the *fep* with the so-called residual empirical process - that has been later turned into a functional form named *residual functional empirical process* - started a new and global approach that provides general asymptotic representations (*GAR*) in the *fep* for a huge class of indices arising in many fields (see, e.g., Mergane *et al.* (2018a) for Gini's inequality index). More details can be found in Lo et al. (2023).

In this paper, we use the described approach to provide a *GAR* for the S-Gini based on the functional empirical process (*fep*) and the Lo's residual functional empirical process introduced by Lo and Sall (2010a) and Lo and Sall (2010b), abbreviated *lrfep*.

The layout of the paper is as follows. We define and discretize the S-Gini index of inequality in Section 2. In that section, we also give some reminders on empirical functions. The results on the GAR are given in Section 3 and their proofs in Section 4. We conclude the paper in Section 5.

2. Definition and discretization of the S-Gini index

The S-Gini index is a variation of the traditional Gini coefficient, which is a measure of inequality. While the Gini coefficient is widely used to measure income or wealth inequality within a population, the S-Gini coefficient introduces a parameter that allows for more sensitivity to changes in different parts of the income distribution.

Let $X \geq 0$ be an income variable with cumulative distribution function *cdf* F, $F(0) = 0$ and let $Q(\circ) = F^{-1}(\circ)$ be the quantile function and let $\nu > 0$. The S-Gini inequality index with parameter is defined us follows in Zitikis and Gastwirth (2002):

Let us suppose that we have the income observations $X_1, X_2, \ldots$ on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$, independent with $\mathbb{P}_X$ common probability characterized by us *cdf* $F_X = F$.

Let $\mu_n = \frac{1}{n}\sum_{j=1}^{n} X_j,\ n \geq 1$, be the sample $X_1, \ldots, X_n$ and $\mu = \mathbb{E}(X)$ the exact mean that is assumed to be finite. The most used empirical expression of the Gini index (GI) from the sample of size $n \geq 1$ is:

$$GI_n = \frac{1}{n\mu_n}\sum_{j=1}^{n}\left(\frac{2j-1}{n} - 1\right) X_{j,n}, \tag{1}$$

where $0 < X_{1,n} < \ldots < X_{n,n}$ is the order statistic associated with the sample. From Zitikis and Gastwirth (2002), a theoretical expression is:

$$GI = \frac{1}{2\mu}\mathbb{E}(|X' - X''|), \tag{2}$$

where X' and X'' are two independent copies of X. From the asymptotic as in Mergane *et al.* (2018b) and Lo et al. (2023) (in the current issue of this journal), we also define it as:

$$GI = \frac{\mathbb{E}(X(2F(X) - 1))}{\mathbb{E}(X)}$$

In Gisbert and *al.* (2010), we can found a remarkable study of properties, the strengths and weaknesses of the GI alongside interesting ways to improve it as well a significant list of related references is given therein.

Here we want to quickly go to our point by remarking, as Gisbert and *al.* (2010) recalled from Donalson and Weymark (1983) that the GI can be written as:

$$GI_n = 1 - \frac{1}{\mu \sum_{j=1}^{n} a_j}\sum_{j=1}^{n} a_j X_{n-j+1,n}, \tag{3}$$

(which is an equivalent writing of Formula 9 is Gisbert and *al.* (2010)) for

$$a_j = 2j - 1, \quad 1 \leq j \leq n.$$

The statistic at the right hand of (3) is defined as given in Gisbert and *al.* (2010), page 6. In the frame of current project, we are concerned with the asymptotic representation through the fep and $lrfep$ of $J(\nu)$.

The asymptotic theory of the $J(\nu)$ seems to be investigated very recently, in particular in Xu (2000) for the computation of the asymptotic variance and Zitikis and Gastwirth (2002) for a general central limit theorem using the expression:

$$J(\nu) = 1 - \frac{\nu}{\mu} \int_0^1 F^{-1}(u)(1-u)^{\nu-1} du \tag{4}$$

(see also Yitzhaki (1983)).

Let us now discretize the index in (4) by plugging in the empirical quantile function in $J(\nu)$ as below. Let $X_1, X_2, \ldots,$ be independent copies of X and for $n \geq 1$, we consider the order statistics of the sample

$$X_{0,n} = 0 < X_{1,n} < \ldots < X_{n,n} < X_{n+1,n} = +\infty.$$

The empirical mean is

$$\mu_n = \frac{1}{n} \sum_{j=1}^{n} X_{j,n}$$

and the empirical quantile is defined by $Q_n(0+) = X_{1,n}$ and for $0 < s < 1$,

$$Q_n(s) = X_{j,n}, \text{ if } \frac{j-1}{n} < s \leq \frac{j}{n}, \ j \in \{1, \ldots, n\}.$$

We have

$$\begin{aligned}
J_n(\nu) &= 1 - \frac{\nu}{\mu_n} \int_0^1 Q_n(u)(1-u)^{\nu-1} du \\
&= \frac{\mu_n}{\mu_n} - \frac{\nu}{\mu_n} \sum_{j=1}^{n} \int_{\frac{j-1}{n}}^{\frac{j}{n}} Q_n(u)(1-u)^{\nu-1} du \\
&= \frac{1}{\mu_n} \left\{ \mu_n - \nu \sum_{j=1}^{n} \int_{\frac{j-1}{n}}^{\frac{j}{n}} X_{j,n}(1-u)^{\nu-1} du \right\} \\
&= \frac{1}{\mu_n} \left\{ \mu_n - \nu \sum_{j=1}^{n} X_{j,n} \int_{\frac{j-1}{n}}^{\frac{j}{n}} (1-u)^{\nu-1} du \right\}.
\end{aligned}$$

We have

$$\int_{\frac{j-1}{n}}^{\frac{j}{n}} (1-u)^{\nu-1} du = -\frac{1}{\nu}\left\{\left(1-\frac{j}{n}\right)^{\nu} - \left(1-\frac{j-1}{n}\right)^{\nu}\right\}. \quad (5)$$

Thus,

$$J_n(\nu) = \frac{1}{n\mu_n}\sum_{j=1}^{n}\left(1+n\left\{\left(1-\frac{j}{n}\right)^{\nu} - \left(1-\frac{j-1}{n}\right)^{\nu}\right\}\right)X_{j,n}. \quad (6)$$

Our objective is to obtain the General Asymptotic Representation *GAR* using the functional empirical process (*fep*) and the Lo residual empirical process (*lrfep*).

Before we go further, we need some recalls for stating this objective. Next, we will describe our results and comment one the signification of ν as suggested by our study.

Recalls on the empirical functions

In this part, we complete the notations we already gave and specify our probability space.

The notation below concerns the univariate random distributions. We are going to describe the general Gaussian field in which we present our results. Indeed, we use a unified approach when dealing with the asymptotic theories of the welfare statistics. It is based on the Functional Empirical Process (*fep*) and its Functional Brownian Bridge (*fbb*) limit. It is laid out as follows.

When we deal with the asymptotic properties of one statistic or index at a fixed time, we suppose that we have a non-negative random variable of interest which may be the income or the expense X whose probability law on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, the Borel measurable space on $\mathbb{R}$, is denoted by $\mathbb{P}_X$. We consider the space $\mathcal{F}$ of measurable real-valued functions f defined on $\mathbb{R}$ such that

$$V_X(f) = \int (f-\mathbb{E}_X(f))^2 d\mathbb{P}_X = \mathbb{E}(f(X)-\mathbb{E}(f(X))^2 < +\infty,$$

where

$$\mathbb{E}_X(f) = \mathbb{E}f(X).$$

On this functional space $\mathcal{F}$, which is endowed with the L_2-norm

$$\|f\|_2 = \left(\int f^2 d\mathbb{P}_X\right)^{1/2},$$

we define the Gaussian process $\{\mathbb{G}(f), f \in \mathcal{F}\}$, which is characterized by its variance-covariance function

$$\Gamma(f,g) = \int^{2} (f - \mathbb{E}_X(f))(g - \mathbb{E}_X(g))d\mathbb{P}_X, (f,g) \in \mathcal{F}^2. \tag{7}$$

This Gaussian process is the asymptotic weak limit of the sequence of functional empirical processes (fep) defined as follows. Let $X_1, X_2, \ldots$ be a sequence of independent copies of X. For each $n \geq 1$, we define the functional empirical process associated with X by

$$\mathbb{G}_n(f) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (f(X_j) - \mathbb{E}f(X_j)), f \in \mathcal{F},$$

and denote the integration with respect to the empirical measure by

$$\mathbb{P}_n(f) = \frac{1}{n} \sum_{i=1}^{n} f(X_i), \ f \in \mathcal{F},$$

Let us denote by $\ell^\infty(T)$ the space of real-valued bounded functions defined on $T = \mathbb{R}$ equipped with its uniform topology. In the terminology of the weak convergence theory, the sequence of objects $\mathbb{G}_n$ weakly converges to $\mathbb{G}$ in $\ell^\infty(\mathbb{R})$, as stochastic processes indexed by $\mathcal{F}$, whenever it is a Donsker class. The details of this highly elaborated theory may be found in Billingsley (1968), Pollard (1984), van der Vaart and Wellner (1996) and similar sources.

Here, we only need the convergence in finite distributions which is a simple consequence of the multivariate central limit theorem, as described in Chapter 3 in Lo *et al.* (2016).

We will use the Rényi's representation of the random variable X_i's of interest by means of the (*cdf*) F of X as follows:

$$X =_d F^{-1}(U),$$

where U is a uniform random variable on $(0, 1)$, $=_d$ stands for the equality in distribution and F^{-1} is the generalized inverse of F, defined by

$$F^{-1}(s) = \inf\{x, F(x) \geq s\}, \ s \in (0,1).$$

Based on these representations, we may and do assume that we are on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ holding a sequence of independent $(0, 1)$-uniform random variables U_1, U_2, ..., and the sequence of independent observations of X are given by

$$X_1 = F^{-1}(U_1), \quad X_2 = F^{-1}(U_2), \quad etc. \tag{8}$$

For each $n \geq 1$, the order statistics of $U_1, ..., U_n$ and of $X_1, ..., X_n$ are denoted respectively by $0 \equiv U_{0,n} < U_{1,n} \leq \cdots \leq U_{n,n} < U_{n+1,n} \equiv 1$ and

$-\inf \equiv X_{0,n} < X_{1,n} \leq \cdots \leq X_{n,n} < X_{n+1,n}$ ($0 = -\inf \equiv X_{0,n}$ for a non-negative random variable X).

We also associate the sequence of $(U_n)_{n\geq 1}$ with the sequence of real-argument empirical functions

$$\begin{aligned}\mathbb{U}_n(s) &= \frac{1}{n}\#\{j, 1 \leq j \leq n,\ U_j \leq s\},\ s \in (0,1),\ n \geq 1 \qquad (9)\\ &= \sum_{j=1}^{n} \frac{j}{n} 1_{(U_{j,n} \leq s < U_{j+1,n})}, \qquad (10)\end{aligned}$$

and the sequence of real-argument uniform quantile function

$$\mathbb{V}_n(s) = U_{1,n}1_{(s=0)} + \sum_{j=1}^{n} U_{j,n}1_{((j-1)/n<s\leq(j/n))},\ s \in (0,1),\ n \geq 1 \qquad (11)$$

and next, the sequence of real-argument uniform empirical processes

$$\alpha_n(s) = \sqrt{n}(\mathbb{U}_n - s), \qquad (12)$$

for $s \in (0,1)$ and $n \geq 1$, and the sequence of real-argument uniform quantile processes

$$\gamma_n(s) = \sqrt{n}(s - \mathbb{V}_n),\ s \in (0,1),\ n \geq 1. \qquad (13)$$

The same can be done for the sequence $(X_n)_{n\geq 1}$, and we obtain the associated sequence of real empirical processes as

$$\mathbb{G}_{n,r}(x) = \sqrt{n}\left(\mathbb{F}_n(x) - F(x)\right),\ x \in \mathbb{R},\ n \geq 1, \qquad (14)$$

where

$$\mathbb{F}_n(x) = \frac{1}{n}\#\{j, 1 \leq j \leq n,\ X_j \leq x\},\ x \in \mathbb{R},\ n \geq 1 \qquad (15)$$

is the associated sequence of empirical functions. We also have the associated sequence of quantile processes

$$\mathbb{Q}_n(x) = \sqrt{n}\left(\mathbb{F}_{(n)}^{-1}(s) - F^{-1}(s)\right),\ s \in (0,1),\ n \geq 1 \qquad (16)$$

where, for $n \geq 1$,

$$\mathbb{F}_n^{-1}(s) = X_{1,n}1_{(0\leq s\leq 1/n)} + \sum_{j=1}^{n} X_{j,n}1_{((j-1)/n\leq s\leq(j/n))},\ s \in (0,1), \qquad (17)$$

is the associated sequence of quantile processes.

By passing, we recall that $\mathbb{F}_n^{-1}$ is actually the generalized inverse of $\mathbb{F}_{(n)}$ and for the uniform sequence, we have

$$\mathbb{V}_n = \mathbb{U}_n^{-1}. \tag{18}$$

In virtue of Representation (8), we have the following remarkable relations

$$\mathbb{G}_{n,r}(x) = \alpha_n(F(x)),\ x \in \mathbb{R} \tag{19}$$

and

$$\mathbb{Q}_n(x) = \sqrt{n}\left(F^{-1}(\mathbb{V}_n(s)) - F^{-1}(s)\right)\ s \in (0,1),\ n \geq 1. \tag{20}$$

We also have the following relations between the empirical functions and quantile functions

$$\mathbb{F}_n(x) = \mathbb{U}_n(F(x)),\ x \in \mathbb{R} \tag{21}$$

and

$$\mathbb{F}_n^{-1}(s) = F^{-1}(\mathbb{V}_{(n)}(s)),\ s \in (0,1),\ n \geq 1. \tag{22}$$

As well, the real-argument and function-argument empirical processes are related as follows: for $n \geq 1$,

$$\mathbb{G}_{n,r}(x) = \mathbb{G}_n(f_x^*),\ \alpha_n(s) = \mathbb{G}_n(\tilde{f}_s),\ s \in (0,1),\ x \in \mathbb{R}, \tag{23}$$

where for any $x \in \mathbb{R}$, $f_x^* = 1_{]-\infty,x]}$ is the indicator function of $]-\infty,x]$ and for $s \in (0,1)$, $f_s = 1_{[0,s]}$ and $\tilde{f}_s = 1_{]-\infty,F^{-1}(s)]}$.

To finish the description, a result of Kiefer-Bahadur (See Bahadur (1966)) that says that the addition of the sequences of uniform empirical processes and quantile processes (12) and (13) is asymptotically, and uniformly on $[0,1]$, zero in probability, that is

$$\sup_{s\in[0,1]} |\alpha_n(s) + \gamma_n(s)| = o_{\mathbb{P}}(1) \text{ as } n \to +\infty. \tag{24}$$

3. General Asymptotic Representation of the S-Gini index

3.1. Our Objective

Based on the notation above, we may say that we search the *GAR* in two steps. First, we want to put the statistic $J_n(\nu)$ in the form

$$J_n(\nu) = \frac{1}{\mu_n}\left\{\frac{1}{n}\sum_{1\leq j\leq n} h(X_j) + \frac{1}{n}\sum_{1\leq j\leq n}\Big(\mathbb{F}_n(X_j) - F(X_j)\Big)\, q(X_j)\right\} + o_{\mathbb{P}}(n^{-1/2}), \tag{25}$$

From there, as in Lo et al. (2023), we need the following conditions:
(*CRe1*) $\mathbb{E}q(X) < +\infty$;

(*CRe2*) and, as $n \to +\infty$,

$$\textbf{\textit{(H2G)}} \quad \int_0^1 \sqrt{n}\,(s - \mathbb{V}_n(s))\,\Delta_n(q,s)\,ds \to 0$$

with

$$\Delta_n(q,s) = \left\{q\Big(F^{-1}\left(\mathbb{V}_n(s)\right)\Big) - q\Big(F^{-1}(s)\Big)\right\}, \quad s \in (0,1).$$

Then under (*(CRe1)*) and (*CRe2*), the *GAR* is turned into

$$J_n(\nu) = \frac{\mathbb{E}H(X)}{\mu} + n^{-1/2}\mathbb{G}_n(h) + n^{-1/2}\int_0^1 \mathbb{G}_n(\tilde{f}_s)\;\ell(s)\;ds + o_{\mathbb{P}}(n^{-1/2}), \tag{26}$$

with

$$\tilde{f}_s = 1_{]-\infty, F^{-1}(s)]}, \;\; s \in (0,1). \tag{27}$$

Let us describe our results. Here, we focus the study on integer values of $\nu > 0$. Since $J(1) = 0$, our study holds only for $\nu \geq 2$. We study the case $\nu = 2$ and the case $\nu > 2$. The first case corresponds to the Gini index et we obtain the results in Lo et al. (2023) (of this series in particular). Finally for $\nu \geq 2$, and under specified conditions, mainly on the moments, we obtain the **GAR** (26) with the associated parameters and functions:

$$
\begin{aligned}
H(x) &= \left\{1-\nu\,(1-F(x))^{\nu-1}\right\}x & (28)\\
h(x) &= \frac{H}{\mu} - \mu^{-2}\mathbb{E}H(X)h_1(x)\\
h_p(x) &= x^p, \quad p>0\\
q(x) &= \frac{\nu(\nu-1)}{\mu}(1-F(x))^{\nu-2}x\\
\ell(s) &= -q(F^{-1}(s)) = \frac{-\nu(\nu-1)}{\mu}(1-s)^{\nu-2}F^{-1}(s).
\end{aligned}
\tag{29}
$$

Besides, the results remain valid for real values $\nu \geq 2$. We obtain the S-Gini as expressed in terms of *cdf* and not in terms of the quantile function. Generally, the method **fep-lrfep** provide an expression in terms of the *cdf*, what may be called the direct expression as opposed to the indirect expression that uses the quantile function.

Here, the direct expression allows us to easily see the impact and the effect of the parameter ν in

$$J(\nu) = \mathbb{E}\Big(X(1-\nu\,(1-F(X))^{\nu-1})\Big). \tag{30}$$

We can see that

$$
\begin{aligned}
J(\nu) &= \int_0^{+\infty} x(1-\nu(1-F(x))^{\nu-1})\,dF(x)\\
&= \int_0^{+\infty} x\,dF(x) - \nu\int_0^{+\infty} x(1-F(x))^{\nu-1})\,dF(x)\\
&= \mu - \nu\int_0^{+\infty}\frac{-x}{\nu}\Big(d(1-F(x))^{\nu}\Big)\\
&= \mu - \nu\Big\{\Big[\frac{-x}{\nu}\Big((1-F(x))^{\nu}\Big)\Big]_0^{+\infty} + \int_0^{+\infty}\frac{1}{\nu}(1-F(x))^{\nu}\,dx\Big\}.
\end{aligned}
$$

For distributions in the Gumbel domain of attraction, for example, we have $[x(1-F(x))^{\nu}]_{x=+\infty} = 0$. For example, this holds for the exponential case with $1-F(x) = e^{-\lambda\ x}$. For Pareto-type distribution $1-F(x) = x^{-1/\gamma}$, $\gamma>0$. The mean μ exists (and then $J(\nu)$) only if $\gamma<1$. In that case, we also have

$$[x(1-F(x))^{\nu}]_{x=+\infty} = [x^{-(\nu-\gamma)/\gamma}]_{x=+\infty} = 0,\ for\ \nu>1.$$

In such a case, we have

$$J(\nu) = \mu - \int_0^{+\infty} (1 - F(x))^{\nu} \, dx.$$

Interpretation. Now we clearly that $J(\nu)$ gets greater with ν. Since the inequality measure increases with ν, that latter parameter can be interpreted as an aversion of inequality due to its role in emphasizing the inequality with a more noticeable value.

Here, we did not study the case $\nu \in]0,1[$ that should be interpreted as an aversion to equality by undermining the inequality measure value. We will focus in this case in a next work.

3.2. Our results

Results of $\nu = 2$.

First let us re-discover the *GAR* of the Gini's index, i.e. for $\nu = 2$. Let $H(x) = (2F(x) - 1)x$ for $x \in \mathbb{R}$ and $\ell(s) = -2F^{-1}(s)$ for $s \in [0,1]$. We need the following assumptions:

(H0) $\mathbb{E}|X| < \infty$.

(H2) $\mathbb{E}X^2 < \infty$.

(H22) As $n \to +\infty$,

$$\int_0^1 \gamma_n(s) \Big(F^{-1}(\mathbb{V}_n(s)) - (F^{-1}(s) \Big) ds \to 0.$$

Theorem 1. *If the hypotheses **(H0) and (H22)** hold, then we have*

$$J_n(2) = \frac{\mathbb{E}H(X)}{\mu} + n^{-1/2}\mathbb{G}_n(h) + n^{-1/2} \int_0^1 \mathbb{G}_n(\tilde{f}_s)\ell(s) \, ds + o_{\mathbb{P}}(n^{-1/2}), \tag{31}$$

where

$$\begin{aligned} h(x) &= \frac{\mu H(x) - \mathbb{E}H(X)h_1(x)}{\mu^2}, \qquad (32) \\ h_1(x) &= x \\ H(x) &= (2F(x) - 1)x \\ \ell(s) &= \frac{-2F^{-1}(s)}{\mu}, \quad x \in \mathbb{R}, \quad s \in (0,1). \end{aligned}$$

Special issue of the statistical applications of the fep and the *lrfep* to nonparametric estimation. E. Kpatcha, T.A. Kpanzou and D. Ba, Afrika Statistika, Vol. 18 (3), 2023, pages 3603 - 3621.

Next, we have

Theorem 2. *Define*

$$\begin{aligned} H(x) &= \left\{1-\nu\,(1-F(x))^{\nu-1}\right\}x \qquad (33)\\ h(x) &= \frac{H}{\mu}-\mu^{-2}\mathbb{E}H(X)h_1(x)\\ h_p(x) &= x^p,\quad p>0\\ q(x) &= \frac{\nu(\nu-1)}{\mu}(1-F(x))^{\nu-2}x\\ \ell(s) &= -q(F^{-1}(s))=\frac{-\nu(\nu-1)}{\mu}(1-s)^{\nu-2}F^{-1}(s). \end{aligned} \qquad (34)$$

Suppose that hypotheses ***H0, (H22) and (H2G)*** *hold. For* $\nu>0$*, we require* ***(H2)****. Then we have*

$$J_n(\nu)=\frac{\mathbb{E}H(X)}{\mu}+n^{-1/2}\biggl(\mathbb{G}_n(h)+\int_0^1\mathbb{G}_n(\tilde{f}_s)\ell(s)\,ds+o_{\mathbb{P}}(1)\biggr). \qquad (35)$$

4. Proofs of the results

Let us begin the proofs by general computations for all cases $\nu>0$. Let us follow Lo et al. (2023) in deriving the *GAR* of $J_n(\nu)$ as defined in (6). We recall that

$$\forall 1\leq j\leq n,\ F_n(X_{j,n})=\frac{j}{n}.$$

Let us transform $J_n(\nu)$ as follows, where we apply the mean value theorem *(MVT)* in Lines (La), (Lb) and (Lc),

$$
\begin{aligned}
J_n(\nu) &= \frac{1}{n\mu_n}\sum_{j=1}^{n}\left\{1-n\left[-\left(1-F_n(X_{j,n})\right)^{\nu}+\left(1-F_n\left(X_{j,n}\right)+\frac{1}{n}\right)^{\nu}\right]\right\}X_{j,n}\\
&= \frac{1}{n\mu_n}\sum_{j=1}^{n}\left\{1+\nu\left(1-F_n(X_{j,n})+\frac{\theta_n}{n}\right)^{\nu-1}\right\}X_{j,n}\\
&= \frac{1}{n\mu_n}\sum_{j=1}^{n}\left\{1+\nu\left(1-F_n(X_{j,n})\right)^{\nu-1}\right\}X_{j,n}\\
&+ \frac{1}{n\mu_n}\sum_{j=1}^{n}\left\{-\nu\left(\left(1-F_n(X_{j,n})+\frac{\theta_n}{n}\right)^{\nu-1}-\left(1-F_n(X_{j,n})\right)^{\nu-1}\right)\right\}X_{j,n} \quad (La)\\
&= \frac{1}{n\mu_n}\sum_{j=1}^{n}\left\{1-\nu\left(1-F(X_{j,n}\right)^{\nu-1}\right\}X_{j,n}\\
&+ \frac{1}{n\mu_n}\sum_{j=1}^{n}\left\{-\nu\left(\left(1-F_n(X_{j,n})\right)^{\nu-1}-\left(1-F(X_{j,n})\right)^{\nu-1}\right)\right\}\ X_{j,n}\\
&+ \frac{1}{n\mu_n}\sum_{j=1}^{n}\left\{-\nu\left(\left(1-F_n(X_{j,n})+\frac{\theta_n}{n}\right)^{\nu-1}-\left(1-F_n(X_{j,n})\right)^{\nu-1}\right)\right\}\ X_{j,n}. \quad (G1)\\
&= \frac{1}{n\mu_n}\sum_{j=1}^{n}\left\{1-\nu\left(1-F(X_{j,n})\right)^{\nu-1}\right\}X_{j,n}\\
&+ \frac{\nu(\nu-1)}{n\mu_n}\sum_{j=1}^{n}\left((F_n(X_{j,n}-F(X_{j,n})(1-F(X_{j,n})+\overline{\theta}_n\delta_{j,n})^{\nu-2}\right)\ X_{j,n} \quad (Lb)\\
&- \frac{\nu(\nu-1)}{n\mu_n}\sum_{j=1}^{n}\frac{\theta_n}{n}\left(1-F_n(X_{j,n})+\frac{\theta_n}{n}\right)^{\nu-2}\ X_{j,n}. \ (Lc)
\end{aligned}
$$

$$\begin{aligned}
&= \frac{1}{n\mu_n}\sum_{j=1}^{n}\Big(1-\nu\,(1-F(X_{j,n}))^{\nu-1}\Big)X_{j,n}\\
&+ \frac{\nu(\nu-1)}{n\mu_n}\sum_{j=1}^{n}\Big(F_n(X_{j,n})-F(X_{j,n})\Big)\Big(1-F(X_{j,n})\Big)^{\nu-2}X_{j,n}\\
&+ \frac{\nu(\nu-1)}{n\mu_n}\sum_{j=1}^{n}\Big(F_n(X_{j,n})-F(X_{j,n})\Big)X_{j,n}\\
&\times \left\{\Big(1-F(X_{j,n})+\overline{\theta}_n\delta_{j,n}\Big)^{\nu-2}-\Big(1-F(X_{j,n})\Big)^{\nu-2}\Big)\right\}\\
&- \frac{\nu(\nu-1)}{n\mu_n}\sum_{j=1}^{n}\frac{\theta_n}{n}\left(1-F_n(X_{j,n})+\frac{\theta_n}{n}\right)^{\nu-2}X_{j,n}.
\end{aligned}$$

where θ_n and $\overline{\theta}_n$ are random numbers in $]0,\ 1[$ and

$$\delta_{j,n}=(F_n(X_{j,n})-F(X_{j,n})),\quad 1\leq j\leq n.$$

Now for $q_1(x)=\nu(\nu-1)(1-F(x))^{\nu-2}x$, and $H(x)=(1-\nu(1-F(x))^{\nu-1})x$, for $x\in\mathbb{R}$, we get

$$J_n(\nu)=:\frac{1}{\mu_n}\Big(\mathbb{P}_n(H)+\hat{R}e_n(q_1)+R_n(1)+R_n(2)\Big), \tag{36}$$

with

$$\begin{aligned}
\mathbb{P}_n(H) &= \frac{1}{n}\sum_{j=1}^{n}H(X_{j,n}),\\
\hat{R}e_n(q_1) &= \frac{1}{n}\sum_{j=1}^{n}\Big(F_n(X_{j,n})-F(X_{j,n})\Big)q_1(X_{j,n}),\\
R_n(1) &= \frac{\nu(\nu-1)}{n}\sum_{j=1}^{n}\Big(F_n(X_{j,n}-F(X_{j,n})\Big)X_{j,n}\\
&\times \left\{\Big(1-F(X_{j,n})+\overline{\theta}_n\delta_{j,n}\Big)^{\nu-2}-\Big(1-F(X_{j,n})\Big)^{\nu-2}\Big)\right\}
\end{aligned}$$

and

$$R_n(2) = -\frac{\nu(\nu-1)}{n}\sum_{j=1}^{n}\frac{\theta_n}{n}\left(1-F_n(X_{j,n})+\frac{\theta_n}{n}\right)^{\nu-2}X_{j,n}.$$

Now, let us divide the proofs into two cases, based on the value of $\nu \in [2,\ +\infty[$.

(C1) Case $\nu = 2$. Back to the Gini index. Proof of Theorem 1. Here, we are trying to join the computations in Lo et al. (2023), page 3547 and hence get the same conclusions. For $\nu = 2$, we have

$$H(x) = (2F(x) - 1)x, \qquad q_1(x) = 2x$$

and

$$J_n(2) = \frac{1}{\mu_n}\left(\frac{1}{n}\sum_{j=1}^{n} H(X_{j,n}) + \frac{1}{n}\sum_{j=1}^{n}(F_n(X_{j,n}) - F(X_{j,n})q_1(X_{j,n})\right) - \frac{1}{n\mu_n} \times \frac{\theta_n}{n}\sum_{j=1}^{n} q_1(X_{j,n}).$$

It clear that the absolute value of

$$-\frac{\theta_n}{n}\sum_{j=1}^{n} q_1(X_{j,n}).$$

is less than

$$\frac{1}{n}\sum_{j=1}^{n} |q_1(X_{j,n})|.$$

Using the weak law of large numbers with Assumption (H0), and by the fact that $\frac{1}{n\mu_n} = \frac{1}{n}(\mathbb{E}(X) + o_{\mathbb{P}}(1))^{-1} = O_{\mathbb{P}}(n^{-1})$, we get

$$J_n(2) = \frac{1}{\mu_n}\left(\frac{1}{n}\sum_{j=1}^{n} H(X_{j,n}) + \frac{1}{n}\sum_{j=1}^{n}(F_n(X_{j,n}) - F(X_{j,n})q_1(X_{j,n})\right) + O_{\mathbb{P}}(n^{-1/2}).$$

From here, we join the computations in Lo et al. (2023) in page 3547. Hence we conclude as in it under (H0) and (H22), ending the proof of Theorem 1.

(C2) Case $\nu > 2$. Proof of Theorem 2. Consider Equation (36). Let us begin to handle $R_n(2)$. The function in $j \in \{1, \cdots, n\}$

$$1 - F_n(X_{j,n}) + \frac{\theta_n}{n}$$

is decreasing and takes the values $1 - (1 - \theta_n)/n$ at $j = 1$ and θ_n/n at $j = n$, which are both in $[0, 1]$. Since $\nu - 2 > 0$, we have that

$$0 \leq \left(1 - F_n(X_{j,n}) + \frac{\theta_n}{n}\right)^{\nu-2} \leq 1$$

and hence

$$|R_n(2)| \leq \frac{\nu(\nu-1)}{n}\left(\frac{1}{n}\sum_{j=1}^{n}|X_j|\right).$$

So, given (H0), we have that $R_n(2) = O_{\mathbb{P}}(1/n)$. Now to treat $R_n(1)$, we recall the Kolmogorov-Smirnov law that applies to

$$D_n = \sup_{x\in\mathbb{R}} \sqrt{n}|F_n(x) - F(x)| \rightarrow Z,$$

we have, for $z \geq 0$,

$$\mathbb{P}(Z \leq z) = 1 - 2\sum_{h=1}^{\infty}(-1)^{h-1}\exp(-2h^2z^2).$$

We conclude that

$$L_n = \sup_{1\leq j\leq n} |\max(\overline{\theta}_n\delta_{j,n}, \delta_{j,n})| = O_{\mathbb{P}}(n^{-1/2}). \tag{37}$$

Let us denote

$$\begin{aligned} V_n &= \max_{1\leq j\leq n}\left|(1-F(X_{j,n})+\overline{\theta}_n\delta_{j,n})^{\nu-2} - (1-F(X_{j,n})^{\nu-2}\right| \quad (38)\\ &= \max_{1\leq j\leq n}\left|(1-F_n(X_{j,n})+\underline{\theta}_n\delta_{j,n})^{\nu-2} - (1-F_n(X_{j,n})+\delta_{j,n})^{\nu-2}\right| \quad (39) \end{aligned}$$

with $\underline{\theta}_n = (1+\overline{\theta}_n)$. Now we have, for any $\varepsilon > 0$, for $\lambda > 0$,

$$\begin{aligned} \mathbb{P}(V_n > \varepsilon) &= \mathbb{P}\Big(\{V_n > \varepsilon\} \cap \{2L_n \leq \lambda\sqrt{n}\}\Big) + \mathbb{P}\Big(\{V_n > \varepsilon\} \cap \{2L_n > \lambda\sqrt{n}\}\Big) \quad (40)\\ &\leq \mathbb{P}\Big(\{V_n > \varepsilon\} \cap \{2L_n \leq \lambda\sqrt{n}\}\Big) + \mathbb{P}\Big(\{2L_n > \lambda\sqrt{n}\}\Big) \quad (41)\\ &=: A_n(1) + A_n(2) \quad (42) \end{aligned}$$

Now, by (37), $(2L_n)/\sqrt{n}$ is bounded in probability, and then for any $\eta > 0$, there exist $\lambda_0 > 0$ and $n_0 \geq 1$ such that for any $n \geq n_0$,

$$A_n(2) = \mathbb{P}(\{2L_n > \lambda_0\sqrt{n}\}) \leq \eta. \quad (F1)$$

Next

$$A_n(1) \leq \mathbb{P}\biggl(\sup_{x\in[0,\ 1+\lambda_0/\sqrt{n}]}\biggl|(x+2\lambda_0\sqrt{n})^{\nu-2}-x^{\nu-2}\biggr|>\varepsilon\biggr).$$

For n big enough with $n \geq n_0$, $\lambda_0/\sqrt{n}$ is less than 2 and for $x \in [0,1]$, $1+\lambda_0/\sqrt{n} \leq 2$, we get

$$\begin{aligned} A_n(1) &\leq \mathbb{P}\biggl(\sup_{(x,y)\in[0,\ 2]^2,\ |x-y|\leq\lambda_0/\sqrt{n}}\biggl|x^{\nu-2}-y^{\nu-2}\biggr|>\varepsilon\biggr) \\ &= 0, \end{aligned}$$

for n big enough, since by uniform continuity of the function $x \mapsto x^{\nu-2}$ on $[0,\ 2]$, as $n \to +\infty$,

$$\sup_{(x,y)\in[0,\ 2]^2,\ |x-y|\leq\lambda/\sqrt{n}}\biggl|x^{\nu-2}-y^{\nu-2}\biggr| \to 0.$$

Then, for any $\eta > 0$, for n big enough,

$$\mathbb{P}(V_n > \varepsilon) \leq \eta,$$

so that $V_n = o_{\mathbb{P}}(1)$. Finally

$$\biggl|R_n(1)\biggr| \leq o_{\mathbb{P}}(n^{-1/2})\hat{R}e_n(q_2),$$

for $q_2(x) = x$. From Lo et al. (2023), $\hat{R}e_n(q_2) = O_{\mathbb{P}}(1)$ for $\mathbb{E}(X^2)$ finite. Finally, under (H2),

$$\biggl|R_n(1)\biggr| = o_{\mathbb{P}}(n^{-1/2}).$$

So, we get

$$J_n(\nu) = \frac{1}{\mathbb{E}(X)+n^{-1/2}\mathbb{G}_n(h_1)}\biggl(\mathbb{E}H(X)) + n^{-1/2}\mathbb{G}_n(H) + \hat{R}e_n(q_1) + o_{\mathbb{P}}(n^{-1/2}\biggr)$$

which gives

$$J_n(\nu) = \frac{1}{\mathbb{E}(X)+n^{-1/2}\mathbb{G}_n(h_1)}\biggl(\mathbb{E}H(X))+n^{-1/2}\mathbb{G}_n(H)+n^{-1/2}\int_0^1 \mathbb{G}_n(\tilde{f}_s)\ell(s)\,ds+o_{\mathbb{P}}(n^{-1/2}\biggr).$$

Finally, we get the *GAR* in (35), ending the proof of Theorem 2. ■

5. Conclusion

In this paper we have established a general asymptotic representation (*GAR*) of the S-Gini index of inequality $J(\nu)$, $\nu \geq 2$, using the functional empirical process (*fep*) and the Lo's residual functional empirical process introduced by Lo and Sall (2010a) and Lo and Sall (2010b), abbreviated *lrfep*. This serves as a basis to studying asymptotic normality of the nonparametric estimator of $J(\nu)$.

Acknowledgment. The authors are indebted to Professor Gane Samb Lo, who made significant contributions to this paper, despite not being listed as a co-author. The same applied to Da Gandasor who partly shared his earlier results on the subject. We also extend our appreciation to the scientific team of the project for their insightful remarks, editing, and critiques, which have notably enhanced the quality of this paper. Finally, we would like to thank the editorial team of the journal for their excellent work in producing the final version of this paper.

References

Bahadur, R.R. (1966). A note on quantiles in large samples, *Ann. Math. Stat.* **37**, pp. 577–580. MR :32:6522. ZL: 0147.18805. doi:10.1214/aoms/1177699450. euclid.aoms/1177699450.

Billingsley, P. (1968). *Convergence of probability measures*. Wiley, new-York.

David H.A. (1968). Gini's mean difference rediscovered. *Biometrika*, Vol 55, pp. 573-575.

Donalson D. and Weymark J. (1983). Ethical flexible Gini indices for income distributions in the continuum. *Journal of Economic Theory*, Vol. 22 (9), April, pp 353-358.

Foster, J., Greer J. and Shorrocks A. (1984). A class of Decomposable Poverty Measures.

Gisbert F.J.G, de la Vega M.C.L. and Urritia A.M. (2010). Generalized S-Gini family: some properties. *Working paper Series ECINEQ*, ECINEQ WP 2010 - 170, Society for the Study of Economic Welfare (www.ecineq.org, visted suceesfully on 2024 August 6).

Kakwani N. (1980). On a Class of Poverty Measures. *Econometrica*, 48, 437-446.

Lo G.S. (2013). On the General Poverty Index. *Far East Journal of Theoretical Statistics*. Volume 42. (1), 1-22.

Lo G.S., Kpanzou T.A. and Da B.O.G. (2023). General asymptotic representations of indexes based on the functional empirical process and the residual functional empirical process and applications. *Afrika Statistika*, Vol 18, (3), pp. 3521-5330. Doi:http://dx.doi.org/10.16929/as/3521.2023.317.ssFEPLREP.

Lo G.S., Ngom M. and Kpanzou T.A. (2016). Weak Convergence (IA). Sequences of random vectors. SPAS Books Series. (2016). Doi : 10.16929/sbs/2016.0001. Arxiv : 1610.05415.

Lo G.S. and Sall S.T. (2010a). Asymptotic Representation Theorems for Poverty Indices. *Afrika Statistika, Special Volume (5) : Proceedings of the International Work-*

Special issue of the statistical applications of the fep and the *lrfep* to nonparametric estimation. E. Kpatcha, T.A. Kpanzou and D. Ba, Afrika Statistika, Vol. 18 (3), 2023, pages 3603 - 3621.

shop on Multiple Risks and Copula, Biskra 2010, pp. 238-244. Ed. Abdelhakim Necir.

Lo G.S. and Sall S.T. (2010b). Uniform weak convergence of the time-dependent poverty measures for continuous longitudinal data, *Brazilian Journal of Probability and Statistics*, Vol. 24, Issue 3, pp 457-467.

Mergane P.D., Kpanzou T.A., Ba D. and Lo G.S. (2018a). A Theil-like class of inequality measures, its asymptotic normality theory and applications. *Afrika Statistika* Vol. 13 (3), pp 1699-1715.

Mergane P.D., Lo G.S. and Kpanzou T.A. (2018b). On the joint distribution of variations of the Gini index and Welfare indices. *In A Collection of Papers in Mathematics and Related Sciences, a festschrift in honour of the late Galaye Dia.* Spas Editions, Euclid Series Book, pp. 163– 194. Doi : 10.16929/sbs/2018.100-02-07.

Pollard D. (1984). *Convergence of Stochastic Processes*. Springer-Verlag, Berlin.

Sen A.K. (1976). Poverty: An Ordinal Approach to Measurement. *Econometrica*, 44, 219-231.

van der Vaart A.W. and Wellner J.A. (1996). *Weak Convergence and Empirical Processes With Applications to Statistics*. Springer, New-York.

Xu K. (2000). Inference for generalized Gini indices using iterated-boostrap method. *J. Bus. Econom. Statist.* Vol 18, pp 223-227.

Yitzhaki S. (1983). On an extension of the Gni inequality index. Internat. Econom. Rev., Vol. 24, pp. 617-628.

Zheng B. (1997). Aggregate poverty measures. *Journal of Economic Surveys*, 11 123–162.

Zitikis R. and Gastwirth J.L. (2002). The Asymptotic Distribution of The S-gini Index. *Aust. N. Z. J. Stat.* 44 (4), pp. 439–446.