African Diaspora Journal of Mathematics
Volume 8, Number 1, pp. 54–67 (2009)
ISSN 1539-854X
www.african-j-math.org

THE ASYMPTOTIC THEORY OF THE KAKWANI CLASS OF POVERTY MEASURES

GANE SAMB LO*
LERSTAD, Université Gaston Berger, Saint-Louis, Sénégal,
and LSTA, Université Pierre et Marie Curie, Paris, France.

SERIGNE TOUBA SALL†
LERSTAD and FASTEF, Université Cheikh Anta Diop, Dakar.

Abstract

Let $Y_1, Y_2, ...$ be independent observations of the income variable of some population, with underlying distribution G. Given a poverty line Z, for each n$\geq$ 1, Q_n is the number of poor individuals in the population. The Kakwani poverty measure

$$P_{kak,n}(k) = \frac{Q_n}{n\Phi_k(Q_n)} \sum_{j=1}^{Q_n} (Q_n - j + 1)^k \left(\frac{Z - Y_{j,n}}{Z} \right),$$

where $\Phi_k(Q_n) = \sum_{j=1}^{j=Q_n} j^k$, and $Y_{1,n} \leq Y_{2,n} \leq ... \leq Y_{n,n}$ are the ordered incomes of the individuals of the population, is one of the most important tools for monitoring poverty in Economics. Here, we complete our asymptotic normality theory for poverty measures by a special study for the Kakwani index and for the Sen measure which is $P_{kak,n}(1)$. The results are positively simulated and data driven examples are also given.

Keywords: Poverty measures, extreme value theory, asymptotic normality, empirical processes, hungarian approximation.

1 Introduction

In this paper, we are concerned with the statistical analysis of poverty indices and its application to the data driven examples. The poverty indices are defined as follows. We consider a population of individuals, each of which having a random income or expenditure Y with distribution function $G(y) = P(Y \leq y)$. An individual is classified as poor whenever his income or expenditure Y fulfills $Y < Z$, where Z is a specified treshold level (the poverty line).

Consider now a random sample $Y_1, Y_2, ...Y_n$ of size n of incomes, with empirical distribution function $G_n(y) = n^{-1}\#\{Y_i \leq y : 1 \leq i \leq n\}$. The number of poor individuals within the sub-population is then equal to $Q_n = nG_n(Z)$.

*E-mail address : ganesamblo@ufrsat.org
†E-mail address: stsall@ufrsat.org

Given these preliminaries, we introduce measurable functions $a(p)$, $b(p)$, $w(p,t,q)$, and $d(t)$ of $p,q \in \mathbb{N}$, and $t \in \mathbb{R}$. The meaning of these functions will be discussed later on.

Let now $Y_{1,n} \leq Y_{2,n} \leq ... \leq Y_{n,n}$ be the order statistics of the sample $Y_1, Y_2, ...Y_n$ of Y. We consider the general poverty indices of the form

$$J_n = \frac{1}{a(Q_n)b(n)} \sum_{j=1}^{Q_n} w(n, Q_n, j)\; d\left(\frac{Z - Y_{j,n}}{Z}\right), \tag{1.1}$$

In the sequel, (1.1) will be called a poverty index (indices in the plural) or simply a poverty measure according to the economists terminology.

A number of such indices have been introduced in the literature since the pioneering work of Sen(1976) who first derived poverty measures (see [18]) from an axiomatic point of view. Many others followed after. Each author introduced *its* own measure in a particular form with specific functions $a(.), b(.)$ and $w(.)$, as we will illustrate it, in coming paragraphs. We put them all in the general form (1.1) and next in

$$J_n = \frac{1}{n} \sum_{j=1}^{Q_n} c(n, Q_n, j)\; d\left(\frac{Z - Y_{j,n}}{Z}\right), \tag{1.2}$$

with $[c(n, Q_n, j) = \{nw(n, Q_n, j)\} / \{a(Q_n)b(n)\}]$ for a global statistical study through a unified approach. Indeed, a survey of these indices may to be found in Zheng [21], who also discussed their characteristics and properties, from an axiomatic point of view. In the form (1.2), $c(n, Q_n, j)$ represents the weight of the poverty index.

Poverty indices are used by economists to monitor and follow the poverty comparison (between specific geographical areas) and the poverty evolution (in the time). To put the reader in his ease and to enable him to have ideas on the functions $a()$, $b()$ and $c()$, we recall here the most used poverty indices in the literature. Among the non-weighted ones, those for which $c(n, Q_n, j) \equiv 1$, the most popular of them is surely the Foster-Greer-Thorbecke (1984) ([7]) class (denoted FGT class) given, for $\alpha \geq 0$,

$$FGT_n(\alpha) = \frac{1}{n} \sum_{j=1}^{Q_n} \left(\frac{Z - Y_{j,n}}{Z}\right)^{\alpha}. \tag{1.3}$$

In Economics terminology, $FGT_n(0) = Q_n/n$ is called the headcount of poor individuals or households, $FGT_n(1)$ the depth in poverty and $FGT_n(2)$ the severity of poverty. In our general frame, the non-weighted poverty index is simply put in the form

$$J_n = \frac{1}{n} \sum_{j=1}^{n} d\left(\frac{Z - Y_{j,n}}{Z}\right). \tag{1.4}$$

In the welfare research fields, this class is very preferred to the others because of the decomposability property. This means that if the population is subdivided into K groups of proportion ω_i $(i = 1, ..., K)$, it suffices to determine the poverty index $J_n(i)$ in each subgroup i in order to get the global index J_n on the whole population by

$$J_n = \sum_{j=1}^{K} \omega_i\, J_n(i).$$

But beside this class, we have several weighted indices, which also have important properties in connection with the Gini or Theil inequality coefficients (see [1], for instance). We mention two of such indices. First, Sen(1979), (see [18]) introduced

$$J_{sen,n} = \frac{2}{n(Q_n+1)} \sum_{j=1}^{Q_n} (Q_n - j + 1) \left(\frac{Z - Y_{j,n}}{Z} \right). \tag{1.5}$$

This has been generalized later by Kakwani(1980) to

$$J_{kak,n}(k) = \frac{Q_n}{n\Phi_k(Q_n)} \sum_{j=1}^{Q_n} (Q_n - j + 1)^k \left(\frac{Z - Y_{j,n}}{Z} \right), \tag{1.6}$$

for $k \geq 1$ and $\Phi_k(Q_n) = \sum_{j=1}^{j=Q_n} j^k$ (see [8]). The generalization, here, means that Sen measure is obtained from (1.6) by putting $k = 1$, that is $J_{sen,n} = J_{kak,n}(1)$. Next Shorrocks(1995), (see [16]) proposed

$$J_{sh,n} = \frac{1}{n^2} \sum_{j=1}^{Q_n} (2n - 2j + 1) \left(\frac{Z - Y_{j,n}}{Z} \right). \tag{1.7}$$

The weighted indices evaluate the poverty intensity by putting a more important weight on the poorest individuals. This means that a small decrease of the income of one household will correspond to a significant increase of the index. The contrary is also valid. This is one of their most crucial characteristics. As mentioned above, a thorough survey of the axiomatics of poverty measures can be found in Zheng [21]. We are not directly concerned with these questions in the present paper nor in those related to the poverty line determination.

In earlier papers, we have been concerned with the consistency of such statistics and further with their general asymptotic theory. Indeed, we proved that (1.2), under some very mild conditions of the distribution function, converges in probability to

$$J = \int_0^Z L(u, G) \, d(\frac{Z-u)}{u})dG(u), \tag{1.8}$$

where L is a weight function, taking a specific form for each particular index. This exact parameter, that we may call as the Exact Poverty Index (EPI) of an infinite population, is new in the literature in its generality. We also remark that the weight is $L(s, G) \equiv 1$ for the not weighted indices while it is equal to $2(1-s)$ (resp. $2(1 - s/G(Z))$) for the Shorrocks (resp. Sen) index.

It then becomes natural to find out a normal asymptotic law of $\sqrt{n}(J_n - J)$ for an accurate estimation of the poverty level of the studied population. Precisely, we proved that $\sqrt{n}(J_n - J)$ converges in distribution to a centered Gaussian random variable $\mathcal{N}(0, \vartheta^2)$ of variance ϑ^2. This allows to use the confidence interval

$$[J_n - \frac{\vartheta}{\sqrt{n}} u_{1-\alpha/2}, J_n + \frac{\vartheta}{\sqrt{n}} u_{1-\alpha/2}] =: [J(1, 1-\alpha), J(2, 1-\alpha)] \tag{1.9}$$

of probability $1 - \alpha$, where $u_{1-\alpha/2}$ is the (1-$\alpha/2$)-quantile of the standard normal distribution function verifying $\mathbb{P}(N(0,1) \leq u_{1-\alpha/2}) = 1 - \alpha/2$.

Let us explain how this kind of results may help to th monitoring of poverty. Given a specific zone and two periods A and B, the poverty comparison relies on their two EPI's, J(A) and J(B) through, for example, their two 95%-confidence intervals

$$[J(1,0.95,A),J(2,0.95,A)] \text{ and } [J(1,0.95,B),J(2,0.95,B)].$$

For $J(A) < J(B)$, we will report positive poverty change if and only if $\Delta(A,B,0.95) = J(1,0.95,B) - J(1,0.95,A) > 0$ and $\Delta(A,B,0.95)$ is the measure of the significant poverty alleviation from period A to period B. The same method is to be used for geographical poverty comparison by the use of the EPI.

This may be achieved through asymptotic normality results. Related works are available in Kakwani (1993) ([9]), Bishop, Formby and Zheng(1997) ([2]), Barrett and Donald (2000) ([1]), Dia (2005) ([6]), Lo (2008) ([11]) and Lo and Sall (2008) ([10]), etc. But as mentioned above, these results are given for particular indices and do not cover the large variety of measures such as the Clark-Hemming(1981) ([3]), the Ray(1989) ([15]), the Kakwani (1980) ([8]), the Thon (1979) ([19]), the Watts(1968) ([20]) measures and so on. We do not want to make this paper any longer and we direct the reader to ([21]) for the definition of theses statistics covered by our general results as showed in ([11])

In [12], we have set a general asymptotic normality of (1.2) for a large class of poverty measures, especially those for which the weight $c(n, j)$ does not include the random poor number Q_n. Unfortunately, the very important case of the Sen measure, and the Kakwani one, which both have weights of the form $c(n,Q_n, j)$, are not covered by our earlier results. This motivates us to devote a special study of the Kakwani class of poverty measures, including the Sen measure as a particular case.

Here, we will give our results in the general form (1.2) before we specialize them for the Kakwani and Sen indices. Since we are directly concerned with poor countries, we shall require this condition

$$G_n(Z) = Q_n/n \to \xi = G(Z) \in]0,1[, \ as\ n \to +\infty.$$

Now, we must remark that Y is here an income or expenditure variable. Then its lower endpoint y_0 is not negative. This allows to study (1.2) via the transform $X = 1/(Y - y_0)$. Throughout this paper, the distribution function of X is F, whose upper endpoint is then $+\infty$. Hence (1.2) is transformed as

$$J_n = \frac{1}{n}\sum_{j=1}^{Q_n} c(n,Q_n, j)\ d\left(\frac{Z - y_0 - X_{n-j+1,n}^{-1}}{Z}\right), \tag{1.10}$$

We organize the paper as follows. We state and prove the asymptotic normality of (1.2) in the next section, and further specialize it for the Sen and Kakwani indices. In a last section, we give simulations of the obtained results and give some data driven applications on the Senegalese data.

2 Results and Proofs

We need conditions on the function $d(.)$ and on the weight $c(n,Q_n, j)$. First assume that

(D1) d($\cdot$) admits a continuous derivative on]0,1).

(D2) d'($\frac{z-y_0}{z}$) and d($\frac{z-y_0}{z}$) are finite.

For $A(u) = 1/F^{-1}(1-u)$, we assume :

(C1) A($\cdot$) admits a derivative on (0,1) denoted A'(u)=a(u).

(C2) a($\cdot$) is continuous on an interval $[a', a'']$ with $0 < a' < a'' < 1$.

(C3) $\exists\, u_0 > 0, \exists\, \eta > -3/2, \forall\, u \in (0, u_0), |a(u)| < C_0\, u^{\eta} \exp(\int_u^1 b(t)t^{-1}dt)$, with b(t)$\rightarrow$ 0 as t$\rightarrow$ 0.

The condition (C3) means that h is bounded by a regularly varying function S(u)=C_0 $u^{\eta} \exp(\int_u^1 b(t)t^{-1}dt)$ of exponent $\eta > -3/2$.

We also need some conditions on the weight $c(\cdot)$. To this purpose, we introduce further notation. In fact, we use in this paper the representations of the studied random variables X_i, $i \geq 1$, by $F^{-1}(1-U_i)$, $i \geq 1$, where $U_1, U_2, ...$ is a sequence of independent random variables uniformly distributed on $(0,1)$. Now let $U_n(\cdot)$ and $V_n(\cdot)$ the uniform empirical distribution and the quantile function based on $U_i, 1 \leq i \leq n$. We have

$$j \geq 1,\ \ \frac{j-1}{n} < s \leq \frac{j}{n} \Longrightarrow \frac{j}{n} = U_n(V_n(s)) \tag{2.1}$$

so that

$$j \geq 1,\ \ \frac{j-1}{n} < s \leq \frac{j}{n} \Longrightarrow c(n,q,j) = c(n, Q_n, nU_n(V_n(s)) \equiv L_n(s). \tag{2.2}$$

Since $U_n(V_n(s)) \rightarrow s$, as n$\rightarrow \infty$, our condition on the weight c($\cdot$) is that the function $L_n(\cdot)$ is uniformly bounded by some constant $D > 0$ and

$$L_n(s) \rightarrow L(s), as\ n \rightarrow \infty, \tag{2.3}$$

where L($\cdot$) is a non-negative C^1-function on (0,1). We focus in this paper on the poverty measures for which

$$\sup_{0 \leq s \leq 1} \left| \sqrt{n}(L_n(s) - L(s)) - \gamma(s)\sqrt{n}(G_n(Z) - G(Z)) \right| = o_p(1) \rightarrow_P 0,\ as\ n \rightarrow \infty, \tag{2.4}$$

for some function $\gamma(\cdot)$. In [12], we delt with the case $\gamma(s) = 0$ for 0$\leq s \leq 1$. We will check later that the Kakwani measure fulfills this condition for all k$\geq$ 1.

Finally, let

$$m(s) = L(s)\ d\left(\frac{Z - y_0 - 1/F^{-1}(1-s)}{Z}\right) ds.$$

We are now able to give our general theorem

Theorem 2.1. *Suppose that (C1-2-3), (D1-2) and (2.4) hold, and let*

$$\mu = \int_0^{G(Z)} \gamma(s) d\left(\frac{Z - y_0 - 1/F^{-1}(1-s)}{Z}\right) ds.$$

Then

$$\sqrt{n}(J_n - J) \to N(0, \vartheta^2)$$

where

$$J = \int_0^{G(Z)} L(s)\, d\left(\frac{Z - y_0 - 1/F^{-1}(1-s)}{Z}\right) ds,$$

$$\vartheta^2 = \theta^2 + (m(G(Z)) + \mu)^2 G(Z)(1 - G(Z)) + (1 - G(Z))\frac{2(m(G(Z)) + \mu)}{Z}\int_0^{G(Z)} sL(s)h(s)ds$$

and

$$\theta^2 = Z^{-2}\int_0^{G(Z)}\int_0^{G(Z)} L(u)\, L(v)\, h(u)h(v)(u \wedge v - uv)\, du\, dv$$

with

$$h(s) = a(s)\, d'(\frac{Z - y_0 - 1/F^{-1}(1-s)}{Z}).$$

Now, we specialize this result for the Sen measure.

Theorem 2.2. *Let Suppose that (C1-2-3) hold and denote by*

$$J_{sen,n} = \frac{2}{n(Q_n + 1)}\sum_{j=1}^{Q_n}(Q_n - j + 1)\left(\frac{Z - y_0 - 1/X_{n-j+1,n}}{Z}\right),$$

the Sen Statistic. Then

$$L(s) = 2(1 - s/G(Z)), \quad \mu = 2G(Z)^{-2}\int_0^{G(Z)} s\, d\left(\frac{Z - y_0 - 1/F^{-1}(1-s)}{Z}\right) ds$$

and

$$\sqrt{n}(J_{sen,n} - J) \to N(0, \vartheta^2)$$

with

$$\vartheta^2 = \theta^2 + \mu^2 G(Z)(1 - G(Z)) + \frac{2\mu}{Z}(1 - G(Z))\int_0^{G(Z)} sL(s)h(s)ds$$

and

$$\theta^2 = Z^{-2}\int_0^{G(Z)}\int_0^{G(Z)} L(u)\, L(v)\, h(u)h(v)(u \wedge v - uv)\, du\, dv.$$

Next, we apply it to the Kakwani measure.

Theorem 2.3. *Suppose that (C1-2-3) hold and let denote by*

$$J_{kak,n} = \frac{Q_n}{n\Phi_k(Q_n)} \sum_{j=1}^{Q_n} (Q_n - j + 1)^k \left(\frac{Z - y_0 - 1/X_{n-j+1,n}}{Z} \right),$$

be the Kakwani statistic. Then for k≥ 1

$$L(s) = (k+1)(1 - s/G(Z))^k, \quad \mu = \int_0^{G(Z)} \gamma(s)\, d\left(\frac{Z - y_0 - 1/F^{-1}(1-s)}{Z} \right) ds$$

and

$$\gamma(s) = k(k+1)(1 - s/G(Z))^{k-1}(s/G(Z)^2)$$

Then

$$\sqrt{n}(J_{kak,n}(k) - J) \to N(0, \vartheta^2)$$

with

$$\vartheta^2 = \theta^2 + \mu^2 G(Z)(1 - G(Z)) + \frac{2\mu}{Z}(1 - G(Z)) \int_0^{G(Z)} sL(s)h(s)ds \tag{2.5}$$

and

$$\theta^2 = Z^{-2} \int_0^{G(Z)} \int_0^{G(Z)} L(u)\, L(v)\, h(u)h(v)(u \wedge v - uv)\, du\, dv.$$

3 Proofs of the Results

All our results will be derived from the lemma below. But, first we place ourselves on a probability space where the so-called Hungarian constructions hold. Namely, Csőrgö and al. (see [5]) have constructed a probability space holding a sequence of independent uniform random variables U_1, U_2, ... and a sequence of Brownian bridges B_1,B_2,... such that for each $0 < \nu < 1/2$, as n$\to\infty$,

$$\sup_{1/n \le s \le 1-1/n} \frac{|\beta_n(s) - B_n(s)|}{(s(1-s))^{1/2-\nu}} = O_p(n^{-\nu}) \tag{3.1}$$

and for $0 < \nu < 1/4$,

$$\sup_{1/n \le s \le 1-1/n} \frac{|B_n(s) - \alpha_n(s)|}{(s(1-s))^{1/2-\nu}} = O_p(n^{-\nu}), \tag{3.2}$$

where

$$\left\{ \alpha_n(s) = \sqrt{n}\,(U_n(s) - s)\,, 0 \leqslant s \leqslant 1 \right\}$$

is the uniform empirical process and

$$\left\{ \beta_n(s) = \sqrt{n}\,(s - V_n(s))\,, 0 \leqslant s \leqslant 1 \right\}$$

is the uniform quantile process. (see ([4]) for a more direct and shorter proof, and ([14]) for a dual version in the sense that in [5], (3.1) is established for pour $0 < \nu < 1/2$ and (3.2) for $0 < \nu < 1/4$, while in [14], (3.1) is proved for $0 < \nu < 1/4$ and (3.2) for $0 < \nu < 1/2$). Now, throughout ν will be fixed and $0 < \nu < 1/4$. Now we are able to give the lemma.

Lemma 3.1. *Suppose that (C1-2-3) and (D1-2) hold and*

$$\sup_{0\leq s\leq 1} \sqrt{n}\,|L_n(s) - L(s)| = O_P(1) \text{ as } n \to \infty. \tag{3.3}$$

Let

$$J = \int_0^{G(Z)} L(s)\, d\left(\frac{Z - y_0 - 1/F^{-1}(1-s)}{Z}\right) ds.$$

Then we have the expansion

$$\sqrt{n}(J_n - J) = N_n(1) + N_n(2) + \int_{1/n}^{G(Z)} \sqrt{n}(L_n(s) - L(s)) d\left(\frac{Z - y_0 - 1/F^{-1}(1-s))}{Z}\right) ds + o_P(1)$$

with

$$N_n(1) = \frac{1}{Z}\int_{1/n}^{G(Z)} L(s) B_n(s) h(s) ds \tag{3.4}$$

and

$$N_n(2) = m(G(Z)) B_n(G(Z)) \tag{3.5}$$

for

$$m(s) = L(s)\, d\left(\frac{Z - y_0 - 1/F^{-1}(1-s)}{Z}\right) ds.$$

Proof. This expansion is Formula (4.14) in [12] (see [10] as well). The condition (3.3) leads to the result. □

We are now proving the first and general Theorem.

Proof. (Theorem 1). Let (Ω, Σ, P) be the probability space on which (3.1) and (3.2) hold. We shall use in this space the representations of the studied random variables X_i, $i \geq 1$, by $F^{-1}(1 - U_i)$, $i \geq 1$,where $U_1, U_2, \ldots$ is a sequence of independent random variables uniformly distributed on $(0,1)$. It follows that :

$$\{\, X_{n-i+1,n}, 1 \leq i \leq n \,;\, n \geq 1\} \stackrel{d}{=} \left\{F^{-1}(1 - U_{i,n}), 1 \leq i \leq n \,;\, n \geq 1\right\} \tag{3.6}$$

where $U_{1,n} < \ldots < U_{n,n}$ are the order statistics based on $\mathrm{U}_1, \ldots, U_n$.

We use at this step the lemma together with (3.3), (2.4) and (3.5), we arrive at

$$\sqrt{n}(J_n - D) = N_n(1) + N_n(3) + o_P(1),$$

where $\mathrm{N}_n(1)$ is defined in (3.4) and

$$N_n(3) = (m(G(Z) + \mu)\alpha_n(G(Z)) + o_P(1) = (m(G(Z) + \mu) B_n(G(Z)) + o_P(1). \tag{3.7}$$

The vector $(\mathrm{N}_n(1), N_n(3))$ is Gaussian and

$$Cov(N_n(1), N_n(3)) = \frac{m(G(Z)) + \mu}{Z} \mathbb{E}\int_{1/n}^{G(Z)} L(s) h(s) B_n(G(Z)) B_n(s) ds \tag{3.8}$$

$$= \frac{m(G(Z)) + \mu}{Z}(1 - G(Z)) \int_{1/n}^{G(Z)} s\, L(s)\, h(s)\, ds.$$

Then $\sqrt{n}(J_n - J)$ is a linear transform $\mathrm{N}_n(1) + N_n(3)$ of the Gaussian vector $(N_n(1), N_n(3))$ plus an $o_P(1)$ term. The variance of this Gaussian term is easily computed through (3.8) and the conclusion follows, that is $\sqrt{n}(J_n - J)$ is asymptotically a centered Gaussian random variable with variance (3.8). □

The main part of this paper is to check that (3.3) holds for the Sen and Kakwani measures.

Proof. (Theorem 2). We have here

$$L_n(s) = \frac{2Q_n}{Q_n+1} - 2\left(\frac{n}{Q_n+1}\right)\frac{j}{n} + \frac{2}{Q_n+1} \quad for \quad \frac{j-1}{n} < s \leq \frac{j}{n}$$

$$= \frac{2Q_n}{Q_n+1} - \frac{2Q_n}{(Q_n+1)G_n(Z)}U_n(V_n(s)) + \frac{2}{Q_n+1} \quad for \quad \frac{j-1}{n} < s \leq \frac{j}{n}.$$

Let $L(s) = 2(1 - s/G(Z))$. We get by straight manipulations that

$$\sup_{0\leq s\leq 1} |L_n(s) - L(s)| \leq G(Z)^{-1}\left[4 \times Q_n^{-1} + \sup_{0\leq s\leq 1} |U_n(V_n(s)) - s|\right] + \left|G(Z)^{-1} - G_n(Z)^{-1}\right|,$$

$$\to 0,\ a.s.,\ as\ n \to +\infty,$$

by [17], p.151. Hence (3.3) holds and the lemma is valid. We must check (2.4). We have for $\frac{j-1}{n} \leq s \leq \frac{j}{n}$,

$$\sqrt{n}(L_n(s) - L(s)) = 2\frac{\sqrt{n}(U_n(V_n(s)) - s)}{G_n(Z)G(Z)} + 2\frac{s\sqrt{n}(G_n(Z) - G(Z))}{G_n(Z)G(Z)} + o_P(1)$$

$$= 2G(Z)^{-2}\, s\sqrt{n}(G_n(Z) - G(Z)) + o_P(1)$$

$$= 2sG(Z)^{-2}\, \alpha_n(G(Z)) + o_P(1).$$

uniformly in s$\in (0,1)$. Put $\gamma(s) = 2s/G(Z)^2$. Finally remark that the conditions (D1-2) hold and that $m(G(Z)) = 0$. This concludes the proof. □

Proof. (Theorem 3). First remark that conditions (D1-2) of Theorem 2.1 hold since $d(x) = x$. The weight of the Kawkani measure is

$$\frac{Q_n}{\Phi_k(Q_n)}(Q_n - j + 1)^k = \frac{Q_n}{\sum_{i=1}^{i=Q_n}(i/n)^k}(Q_n/n - j/n + 1/n)^k.$$

But

$$\sum_{i=1}^{i=Q_n}(i/n)^k = n\sum_{i=1}^{i=Q_n}\int_{(i-1)/n}^{i/n} U_n(V_n(s))^k\, ds = n\int_0^{Q_n/n} U_n(V_n(s))^k\, ds$$

Since $\mathrm{U}_n(V_n(\cdot))$ converges uniformly to $I(s) = s$ in (0,1) and $Q_n/n \to G(Z)$, $\int_0^{Q_n/n} U_n(V_n(s))^k$ $ds \to G(Z)^{k+1}/(k+1)$. Now, for $\frac{j-1}{n} < s \leq \frac{j}{n}$,

$$L_n(s) = \frac{Q_n}{\Phi_k(Q_n)}(Q_n - j + 1)^k = \left\{\frac{1}{n}\sum_{i=1}^{i=Q_n}(i/n)^k\right\}^{-1}(Q_n/n)\ \{Q_n/n - U_n(V_n(s)) + 1/n\}^k$$

$$\equiv I_n \times G_n(Z) \times H_n(s).$$

Then for

$$L(s) = (k+1)(1 - s/G(Z))^k,$$

we have

$$\sup_{0 \le s \le G(Z)} |L_n(s) - L(s)| \to 0.$$

We first handle

$$I_n - \frac{G(Z)^{k+1}}{k+1} = \int_0^{G_n(Z)} U_n(V_n(s))^k\, ds - \int_0^{G(Z)} U_n(V_n(s))^k\, ds + \int_0^{G(Z)} U_n(V_n(s))^k\, ds - \int_0^{G(Z)} s^k\, ds$$

$$= \int_{G(Z)}^{G_n(Z)} U_n(V_n(s))^k\, ds + \int_0^{G(Z)} (U_n(V_n(s))^k - s^k)\, ds$$

$$= (G_n(Z) - G(Z)) U_n(V_n(a_n(s)))^k\, ds + \int_0^{G(Z)} k(U_n(V_n(s)) - s) b_n(s)^{k-1}\, ds,$$

with $\mathrm{a}_n(s) \in [\min(G_n(Z), G(Z)), \max(G_n(Z), G(Z))]$ and $\mathrm{b}_n \in [\min(U_n(V_n(s)), s), \max(U_n(V_n(s)), s)]$ by the mean value theorem. We get by uniform continuity on (0,1), and by [17], p.151, that, for $I = G(Z)^{k+1}/(k+1)$,

$$\sqrt{n}(I_n - I) = \sqrt{n}(G_n(Z) - G(Z))G(Z)^k + o_P(1) \tag{3.9}$$

$$= G(Z)^k \alpha_n(G(Z)) + o_P(1) = G(Z)^k B_n(G(Z)) + o_P(1),$$

by (3.2). We now handle

$$H_n(s) - (G(Z) - s)^k = k\{(G_n(Z) - G(Z) + (s - U_n(V_n(s))\}\, a_n(s)^{k-1},$$

with $\mathrm{a}_n(s) \in [\min(G(Z) - s, G_n(Z) - U_n(V_n(s)) + 1/n), \max(G(Z) - s, G_n(Z) - U_n(V_n(s)) + 1/n)]$. By the same arguments used above, we also have for $H(s) = (G(Z) - s)^k$,

$$\sqrt{n}\{H_n(s) - H(s)\} = k(G(Z) - s)^{k-1} B_n(G(Z)) + o_P(1), \tag{3.10}$$

where we have used

$$\sqrt{n}(G_n(Z) - G(Z)) = B_n(G(Z)) + o_P(1). \tag{3.11}$$

We put together (3.9), (3.10) and (3.11) in this way :

$$\frac{G_n(Z)H_n(s)}{I_n} - \frac{G(Z)H(s)}{I}$$

$$= \frac{G_n(Z)H_n(s)}{I_n} - \frac{G_n(Z)H(s)}{I_n} + \frac{G_n(Z)H(s)}{I_n} - \frac{G(Z)H(s)}{I_n} + \frac{G(Z)H(s)}{I_n} - \frac{G(Z)H(s)}{I}$$

and have

$$L_n(s) - L(s) = \frac{G_n(Z)}{I_n}(H_n(s) - H(s) + \frac{H(s)}{I_n}(G_n(Z) - G(Z)) - \frac{G(Z)H(s)}{I \times I_n}(I_n - I).$$

And then, uniformly in s,

$$\sqrt{n}(L_n(s)-L(s)) = (k\frac{k+1}{G(Z)^k}(G(Z)-s)^{k-1} + \frac{k+1}{G(Z)^{k+1}}(G(Z)-s)^k$$

$$-\frac{G(Z)(k+1)^2}{G(Z)^{2k+2}}(G(Z)-s)^k)B_n(G(Z)) + o_P(1)$$

$$= k(k+1)(1-s/G(Z))^{k-1}(s/G(Z)^2)\ B_n(G(Z)) + o_p(1).$$

And hence

$$\sqrt{n}(L_n(s)-L(s)) = \gamma(s)B_n(G(Z)) + o_P(1),$$

with

$$\gamma(s) = k(k+1)(1-s/G(Z))^{k-1}(s/G(Z)^2).$$

Put $\gamma(s) = 2s/G(Z)^2$. Finally remark that the conditions (D1-2) hold and $m(G(Z)) = 0$. This concludes the proof □

Remark 3.2. We should notice that for k=1, we obtain the behavior of the Sen measure.

4 Simulation study and data driven applications

4.1 Simulations

Since, we want to base poverty comparison analysis on our results for the Kakwani measure, it is important to be convinced that they perform well in the environment of the study. In the Senegalese data, it is striking to remark that the variable $X = 1/(Y-y_0)$, where Y is the expenditure, is well fitted to the log-normal law (see [10]), in all the geographical areas and for the databases of 1996 and 2001. This is why we conducted the simulation studies with the special choice of the log-normal law for the distribution function F. We begin to fix values for the parameters m and σ^2 of the log-normal law, a value k and a given poverty line Z corresponding to a poor proportion $G(Z)$ and next evaluate, by the Newton-Raphson algorithm, the values of $J(k)$ et ϑ^2 (denoted ET^2). We generate $n = 150$ independent random variables from the log-normal law of parameters m and σ^2, and compute the sample index $J_{kak,n}(k)$, the individual error $|J_{kak,n}(k)-J(k)|$ and the P-value $\mathbb{P}(\mathcal{N}(0,\vartheta^2) \geq |\sqrt{n}(J_{kak,n}(j)-J(k))|)$. Now after $B = 1000$ replications, we finally report the mean value of the poverty index (always denoted by $J_{kak,n}(k)$), the mean quadratic error from the individual errors (denoted E) and the mean value of the P-value (denoted PV). For different values of $G(Z)$, and for values of m and σ^2 near those of our data, the simulations are very conclusive and support very well to the theoretical results. We may remark the small values for ET and great values PV (greater than 5%) in the table below. This shows how conclusive are the simulation studies.

Tab. 1

Q/n(%)	35	45	55	65
k=1 (Sen)	$J_n(\%) = 19,77$	$26,48$	$34,99$	$44,75$
	$J(\%) = 21,05$	$26,37$	36.95	$43,94$
	$E = 0.028$	0.035	0.0296	0.036
	$ET = 0.312$	0.319	0.364	0.364
	$PV(\%) = 24$	23	$25,80$	$27,47$
k=3	$J_n = 19,19$	$25,63$	$27,49$	$44,16$
	$J = 19,42$	$25,40$	$26,57$	$45,10$
	$E = 0.027$	0.032	0.0311	0.343
	$ET = 0.464$	0.490	0.505	0.527
	$PV = 31,12$	$33,55$	$30,98$	$29,66$

4.2 data driven applications

We just want here to illustrate how good are the application results of our methods to poverty databases. A large scale and thorough study on Senegalese data is reported in [10]. We shall apply Theorems 2.2 and 2.3 to the data of the poorest area in Senegal (Kolda) and the richest area (Dakar). In both cases, we give the sample size, compute the empirical poverty index value J_n and the Exact Poverty value (EPI) which is J and and report the 95%-confidence interval for $k = 1$ (Sen measure) and $k = 3$. We should remark that the computations are based on a log-normal fitting of the the data $X_1, X_2, \ldots$ for a per day poverty line of approximately of half a US Dollar.

$$\left[J_n - 1.96\frac{\vartheta}{\sqrt{n}}, J_n + 1.96\frac{\vartheta}{\sqrt{n}}\right] = [J(1), J(2)]$$

Area	size	k	$J_n\%$	D%	J(1)%	J(2)%
Kolda	198	1	24.86	24.47	21.25	28.47
Kolda	198	3	29.83	29.25	25.75	33.9
Dakar	1122	1	1.48	1.85	1.05	1.92
Dakar	1122	3	1.48	1.95	1.42	2.48

This shows how good are the applications of the results to the Senegalese data. We may also compare the two areas with the help of the confidence intervals not by relying of the empirical values. We report a significant difference of the poverty level of 24.47 $\% - 1.85\% = 22.62\%$ with Sen's index and of $25.75\% - 1.95\% = 23.8\%$ with Kakwani's index of parameter k=3.

Conclusion

In this paper we finished the settling of a general asymptotic theory for the general poverty measure, using a unified approach. We devoted ([10]) to a large study of poverty analysis in Senegal with comparison results as well in space as in time with the help of the asymptotic results given here. In next papers, we will investigate the time-dependent poverty measure $\{J_n(t), 0 \leq t \leq T\}$ for $T > 0$, for longitidunal data study.

References

[1] G. Barrett and S. Donald, Statistical Inference with Generalized Gini Indices of Inequality and Poverty, (2000). *http://www.eco.utexas.edu/˜donald/research/genginir.pdf.*

[2] J. A. Bishop, J. P. Formby and B. Zheng, Statistical Inference and the Sen Index of Poverty. *Int. Econom. Rev.*, **38**, (1997), (2), pp.381-387.

[3] S. Clark, R. Hemming and D. Ulph, On Indices for the Measurement of Poverty. *Econom. J.*, **91**, (1981), pp.525-526

[4] M. Csörgő, L. Horvàth, Approximations of weighted empirical and quantile processes. *Statist. Probab. lett.*, **4**, (1986), pp.275-280.

[5] M. Csőrgö, S. Csőrgö, L. Horvàth and D. M. Mason, Weighted empirical and quantile processes. *Ann. Probab.*, **14**, (1986), pp.31-85.

[6] G. Dia, Répartition Ponctuelle Aléatoire des Revenus et Estimation de l'Indice de Pauvreté. *Afrik. Statist.*, **1**, (2005), (1), pp.47-66

[7] J. Foster, J. Greer and A. Shorrocks, A class of Decomposable Poverty Measures. *Econometrica*, **52**, (1984), pp.761-766.

[8] N. Kakwani, On a class of Poverty measures. *Econometrica*, **48**, (1980), pp.437-446

[9] N. Kakwani, N.(1993). Statistical inference in the measurement of poverty. *Rev. Econom. Stat.*, **75**, (1983), (4), pp. 632-639.

[10] G. S. Lo, Estimation asymptotique des indices de pauvreté : modélisation continue et analyse spatio-temporelle de la pauvreté au Sénégal. *J. Afric. Commun. Sci. Techn.*, **4**, (2009), pp.341-377

[11] G. S. Lo, The General Poverty Index Measure. To appear in *Far East J. Theor. Stat.*, (2009). Avalaible at : http://www.ufrsat.org/lerstad/pub/GPI.pdf.

[12] G. S. Lo, S. T. Sall and C. T. Seck. Une Thorie asymptotique des indicateurs de pauvret. To appear in *C. R. Math. Acad. Sci. Soc. R. Can.*, (2009).

[13] K. S. Amartya, Poverty: An Ordinal Approach to Measurement. *Econometrica*, **44**, (1976), pp.219-231.

[14] D. M. Mason and W. R. van Zwet, A refinement of the KMT inequality for the uniform empirical process. *Ann. Probab.*, **15**, (1987), (3), pp.871-884.

[15] R. Ray, A new class of decomposable poverty measures. *Indian Econom. J.*, **36**, pp.30-38.

[16] A. Shorrocks, Revisiting the Sen Poverty Index. *Econometrica*, **63**, (1995), pp.1225-1230.

[17] G. R Shorack and J. A. Wellner, *Empirical Processes with Applications to Statistics*, (1986), wiley-Interscience, New-York.

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

[19] D. Thon, On Measuring Poverty. *Rev. Inc. Weal.*, **25**, (1979), pp.429-440.

[20] H. Watts, An Economics Definition of Poverty. *In D. P. Moyniha (ed), On Understaining Poverty*, (1968), New-York : Basic Books.

[21] B. Zheng, Aggregate Poverty Measures. *J. Econom. Surv.*, **11**, (1997), (2), pp.123-162.