Afrika Statistika

ISSN 2316-090X

Afrika Statistika
Vol. 13 (3), 2018, pages 1699 –1715.
DOI: http://dx.doi.org/10.16929/as/1699.129

A Theil-like class of inequality measures, its asymptotic normality theory and applications

Pape Djiby Mergane *,(1) **, Tchilabalo Abozou Kpanzou** (2)**, Diam Ba** (1) **and Gane Samb LO** (1,3)

(1) LERSTAD, Gaston Berger University, Saint-Louis, SENEGAL
(2) University of Kara, Kara, TOGO
(3) Associate Researcher, LASTA, Pierre et Marie University, Paris, FRANCE
Associated Professor, African University of Sciences and Technology, Abuja, NIGERIA
Personal Address : Evanston Drive, NW, Calgary, AB, Canada, T3P 0J9

Received on July 14, 2018; Accepted on September 14, 2018

Abstract. In this paper, we consider a coherent theory about the asymptotic representations for a family of inequality indices called Theil-Like Inequality Measures (*TLIM*), within a Gaussian field. The theory uses the functional empirical process approach. We provide the finite-distribution and uniform asymptotic normality of the element of the TLIM class in a unified approach rather than in a case by case one. The results are then applied to some UEMOA countries databases.

Key words: Inequality measure; Asymptotic behaviour; Asymptotic representations; functional empirical process
AMS 2010 Mathematics Subject Classification : 62G05; 62G20; 62G07; 91B82; 62P20

---

* Corresponding author: Pape Djiby Mergane :
Tchilabalo Abozou Kpanzou (kpanzout@gmail.com)
Diam Ba (diamba79@gmail.com)
Gane Samb LO (gane-samb.lo@ugb.edu.sn, gslo@aust.edu.ng, ganesamblo@ganesamblo.net

Résumé. (French) Dans cet article, nous présentons une théorie cohérente de représentations asymptotiques d'une famille de mesures d'inégalité dénommée *TLIM* dans un champ gaussien précis. Notre méthode est fondée sur le processus empirique fonctionnel. Nous tirons de la représentation asymptotique les limites en distribution des estimateurs plug-in des membres de la famille en dimension finie. Les résultats sont ensuite appliqués à des données issues des pays de l'UEMOA.

1. Introduction

In this paper, we deal with a modern weak theory for some large class of inequality indices that, further, will allow to handle easy comparison studies with different kinds of statistics.

According to earlier economists, inequality indices are functional relations between the income and the economic welfare (see Dalton (1920)). This explains, among others, the wide variety of such indices in the literature (See, e.g., Cowell (1980a,b, 2000)).

Such statistics, of course, have been widely studied with respect to a great variety of interests, including statistical characterizations and asymptotic properties (See Davidson and Duclos (2000), Barrett and Donald (2009), for recent studies).

Recently, Greselin et al. (2009) provided a mathematical investigation of these indices in a modern setting including Vervaat processes, L-statistics and empirical processes.

Having in mind the necessity of comparing inequality measures with different kind of statistics such as growth statistics, we aim at providing a coherent asymptotic weak theory for some class of inequality measures. Indeed we propose the functional empirical process setting (see Van der Vaart and Wellner (1996)) which provide natural Gaussian field in which many statistics used in Economics may be represented in.

Our best achievement consists of the asymptotic representations for the elements of our class of inequality measures, in terms of the above mentioned Gaussian field. The results are illustrated in data driven applications, on Senegalese data for instance.

The class on which we focus here is a functional family of inequality measures which gathers various ones around the central Theil measure. This class named after the Theil-Like Inequality Measure (TLIM) will be the central point of our study. It includes the Generalized Entropy Measure, the Mean Logarithmic Deviation (Cowell (2003); Theil (1967); Cowell (1980a)), the different inequality measures of Atkinson (1970), Champernowne and Cowell (1998), Kolm (1976), and the divergence of Renyi (1961).

This means that, here, we will not discuss other inequality statistics such as the Gini, the Generalized Gini, the S-Gini, the E-Gini (See Barrett and Donald (2009)). Those statistics and similar ones will be treated in separate papers.

Now we are going to introduce our family. For that, let X denote the income (or expense) random variable related to a given population. We assume that X and its independent observations are defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and take their values in on some interval $\mathcal{V}_X \subset \mathbb{R}_+^*$ and have common cumulative distribution function *(cdf)*, $F(x)$, $x \in \mathcal{V}_X$. In this paper, we only use Lebesgue-Stieljes integrals and for any measurable function $\ell : \mathbb{R} \to \mathbb{R}$, we have, whenever it makes sense,

$$\mathbb{E}(\ell(X)) = \int \ell(x)\ dF(x) \equiv \int_{\mathcal{V}_X} \ell\ d\mathbb{P}_X,$$

where $\mathbb{P}_X = \mathbb{P}X^{-1}$ is the measure image of $\mathbb{P}$ by X, but is also Lebesgue-Stieljes probability measure characterized by: $\mathbb{P}_X(]a,b]) = F(b) - F(a)$ for any $-\infty \leq a \leq b \leq +\infty$.

Now, consider a sample of $n \geq 1$ individuals or households of that population and observe their income $X_1, X_2, \cdots, X_n$. We define the following family of inequality indices, indexed by $\phi = (\tau, h, h_1, h_2) \in \mathcal{P}_0$ as follows

$$T_n(\phi, X) = \tau\left(\frac{1}{n}\sum_{j=1}^{n}\left(\frac{h(X_j)}{h_1(\mu_n)} - h_2(\mu_n)\right)\right),\ h_1(\mu_n) \neq 0, \tag{1}$$

where $\mu_n = \frac{1}{n}\sum_{j=1}^{n} X_j$ is the empirical mean while $h(x)$, $h_1(y)$, $h_2(z)$, and $\tau(t)$ are real and measurable functions of $x, y, z \in \mathcal{V}_X$ and $t \in \mathbb{R}$. The exact form of $\mathcal{P}_0$ is not important here, in opposite to the conditions on the functions τ, h, h_1 and h_2 under which the results are valid. In a future paper on the uniform limits in ϕ, that class will be crucial.

We will see below that T_n under specific hypotheses on τ, h, h_1, h_2 and μ_n, converges to the exact inequality measure

$$T(\phi, X) = \tau\left(\frac{1}{h_1(\mu)}\int_{\mathcal{V}_X} h(x)\ dF(x) - h_2(\mu)\right),\ h_1(\mu) \neq 0, \tag{2}$$

where $\mu = \mathbb{E}(X)$ is the mathematical expectation of X that we suppose finite here. We will come back later on the function classes $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$ and $\mathcal{F}_4$ in which h, τ, h_1 and h_2 are supposed to lie.

Each measure of this Theil-like family has its own particular properties, that are derived from the combination of different concepts. One may mention the concept of welfare criteria (Atkinson (1970), Sen (1973)), that of the analogy with analysis of risks (Harsanyi (1953), Harsanyi (1955), Rothschild and Stiglitz (1973)), that of the complaints approach (Temkin (1993)) etc. The Theil inequality itself finds all its interest in the information-theoretic idea following that of main components (Kullback (1959)). It is based on the three following axioms: Zero-valuation of certainty, Diminishing-valuation of probability, Additivity of independent events. A deep review of such of individual properties for a number inequality measures can be found in Cowell (Cowell (1980a,b, 2000)) for instance.

It is worth mentioning that the TLIM presented here, is rather a mathematical form gathering a number of different measures.

The rest of the paper is organized as follows. In Section 2, we describe the TLIM family and show how the particular indices are derived from it. In Section 3, we briefly recall the functional empirical processes setting. In section 4, we deal with the asymptotic theory of the TLIM, state and describe our main results. The results are proved in Section 5. Section 6 is devoted to data driven applications. We finish by a conclusion in Section 7.

2. Description of the TLIM

The inequality measures mentioned above are derived from (1) with the particular values of the measurable functions τ, h, h_1 and h_2 as described below for all $s > 0$.

2.1. Generalized Entropy

$$GE_{n,\alpha}(X) = \frac{1}{n\alpha(\alpha-1)} \sum_{j=1}^{n} \left(\left(\frac{X_j}{\mu_n} \right)^{\alpha} - 1 \right);$$

$$\alpha \neq 0,\ \alpha \neq 1,\ \tau(s) = \frac{s-1}{\alpha(\alpha-1)},\ h(s) = h_1(s) = s^{\alpha},\ h_2(s) \equiv 0.$$

2.2. Theil's measure

$$Th_n(X) = \frac{1}{n} \sum_{j=1}^{n} \frac{X_j}{\mu_n} \log \frac{X_j}{\mu_n};$$

$$\tau(s) = s,\ h(s) = s\log(s),\ h_1(s) = s,\ h_2(s) = \log(s).$$

2.3. Mean Logarithmic Deviation

$$MLD_n(X) = \frac{1}{n} \sum_{j=1}^{n} \log \left(\frac{X_j}{\mu_n} \right)^{-1};$$

$$\tau(s) = s,\ h(s) = h_2(s) = \log(s^{-1}),\ h_1(s) \equiv 1.$$

2.4. Atkinson's measure

$$Atk_{n,\alpha}(X) = 1 - \frac{1}{\mu_n}\left(\frac{1}{n}\sum_{j=1}^{n} X_j^{\alpha}\right)^{1/\alpha};$$

$$\alpha < 1 \text{ and } \alpha \neq 0,\ \tau(s) = 1 - s^{1/\alpha},\ h(s) = h_1(s) = s^{\alpha},\ h_2(s) \equiv 0.$$

2.5. Champernowne's measure

$$Ch_n(X) = 1 - \frac{1}{\exp\left(-\frac{1}{n}\sum_{j=1}^{n}\log\frac{X_j}{\mu_n}\right)};$$

$$\tau(s) = 1 - \exp(s),\ h(s) = h_2(s) = \log(s),\ h_1(s) \equiv 1.$$

2.6. Kolm's measure

$$Ko_{n,\alpha}(X) = \log\left(\frac{1}{n}\sum_{j=1}^{n}\exp\left(-\alpha(X_j - \mu_n)\right)\right)^{1/\alpha};$$

$$\alpha > 0,\ \tau(s) = \frac{1}{\alpha}\log(s),\ h(s) = h_1(s) = \exp(-\alpha s),\ h_2(s) \equiv 0.$$

2.7. Divergence of Renyi

$$DR_{n,\alpha}(X) = \frac{1}{\alpha - 1}\log\left(\frac{1}{n}\sum_{j=1}^{n}\left(\frac{X_j}{\mu_n}\right)^{\alpha}\right);$$

$$\alpha \in \mathbb{R}_+ \setminus \{1\},\ \tau(s) = \frac{1}{\alpha - 1}\log(s),\ h(s) = h_1(s) = s^{\alpha},\ h_2(s) \equiv 0.$$

3. The functional empirical process

Let $Z_1, Z_2, \ldots, Z_n$ be a sequence of independent and identically distributed random elements defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with values in some metric space (S, d). Given a collection $\mathcal{F}$ of measurable functions $f : S \to \mathbb{R}$ satisfying

$$\sup_{f \in \mathcal{F}} |f(z) - \mathbb{P}(f)| < \infty, \text{ for every } z,$$

where $\mathbb{P}(f) = \mathbb{E}(f(Z))$ is the mathematical expectation of $f(Z)$, the functional empirical process (*FEP*) based on the $(Z_j)_{j=1,\ldots,n}$ and indexed by $\mathcal{F}$ is defined by :

$$\forall f \in \mathcal{F}, \mathbb{G}_n(f) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(f(Z_j) - \mathbb{P}(f)\right).$$

This process is widely studied in Van der Vaart and Wellner (1996) for instance. It is readily derived from the real Law of Larges Numbers (*LLN*) and the real Central Limit Theorem (*CLT*) that $\mathbb{P}_n(f) = \frac{1}{n}\sum_{j=1}^{n} f(Z_j) \to \mathbb{P}(f) \ \ a.s.$ and that $\mathbb{G}_n(f) \to \mathcal{N}\left(0, \sigma_f^2\right)$, where

$$\sigma_f^2 = \mathbb{P}\left((f - \mathbb{P}(f))^2\right) < \infty, \tag{3}$$

whenever $\mathbb{E}\left(f(Z)^2\right) < \infty$.

When using the *FEP*, we are often interested in uniform *LLN*'s and weak limits of the *FEP* considered as stochastic processes. This gives the so important results on Glivenko-Cantelli classes and Donsker ones. Let us define them here (for more details see Van der Vaart and Wellner (1996)).

Since we may deal with non measurable sequences of random elements, we generally use the outer almost sure convergence defined as follows:

a sequence U_n converges outer almost surely to zero, denoted by $U_n \to 0\, as^*$, whenever there is a measurable sequence of measurable random variables V_n such that

1. $\forall n,\ |U_n| \leq V_n,$
2. $V_n \longrightarrow 0\ a.s.$

The weak convergence generally holds in $\ell^\infty(\mathcal{F})$, the space of all bounded real functions defined on $\mathcal{F}$, equipped with the supremum norm $\|x\|_{\mathcal{F}} = \sup_{f\in\mathcal{F}} |x(f)|$.

Definition 1. $\mathcal{F} \subset L_1(\mathbb{P})$ is called a *Glivenko-Cantelli class* for $\mathbb{P}$, if

$$\lim_{n\to\infty} \left\| \frac{1}{n}\sum_{j=1}^{n} \left(f(Z_j) - \mathbb{P}(f)\right) \right\|_{\mathcal{F}} = 0\,\text{a.s.}^*.$$

Definition 2. $\mathcal{F} \subset L_2(\mathbb{P})$ is called a *Donsker class* for $\mathbb{P}$, or $\mathbb{P}$-*Donsker class* if $\{\mathbb{G}_n(f); f \in \mathcal{F}\}$ converges in $l^\infty(\mathcal{F})$ to a centered Gaussian process $\{\mathbb{G}(f); f \in \mathcal{F}\}$ with covariance function

$$\Gamma(f,g) = \int_{\mathbb{R}} (f(z) - \mathbb{P}(f))\,(g(z) - \mathbb{P}(g))\; d\mathbb{P}_Z(z)\,;\ \forall f,g \in \mathcal{F}.$$

Remark 1. When $S = \mathbb{R}$ and $\mathcal{F} = \left\{f_t = \mathbf{1}_{(-\infty,t]}, t \in \mathbb{R}\right\}$, $\mathbb{G}_n$ is called real empirical process and is denoted by α_n.

In this paper, we only use finite-dimensional forms of the *FEP*, that is $(\mathbb{G}_n(f_i), i = 1, \ldots, k)$. And then, any family $\{f_i, i = 1, \ldots, k\}$ of measurable functions satisfying (3), is a Glivenko-Cantelli and a Donsker class, and hence

$$(\mathbb{G}_n(f_i), i = 1, \ldots, k) \xrightarrow{d} (\mathbb{G}(f_1), \mathbb{G}(f_2), \ldots, \mathbb{G}(f_k)),$$

where $\mathbb{G}$ is the Gaussian process, defined in Definition 2. We will make use of the linearity property of both $\mathbb{G}_n$ and $\mathbb{G}$. Let $f_1, \ldots, f_k$ be measurable functions satisfying (3) and $a_i \in \mathbb{R}, i = 1, \ldots, k$, then

$$\sum_{j=1}^{k} a_j \mathbb{G}_n(f_j) = \mathbb{G}_n\left(\sum_{j=1}^{k} a_j f_j\right) \xrightarrow{d} \mathbb{G}\left(\sum_{j=1}^{k} a_j f_j\right).$$

The materials defined here, when used in a smart way, lead to a simple handling of the problem which is addressed here.

4. Our results

Let us introduce some notation.

$$B_{h,n} = \frac{1}{n}\sum_{j=1}^{n} h(X_j), \; B_h = \int_{\mathcal{V}_X} h(x)\, dF(x);$$

$$K_\phi = \tau'\left(\frac{B_h}{h_1(\mu)} - h_2(\mu)\right) \neq 0;$$

for all $x \in \mathcal{V}_X$, we define the following function

$$F_\phi(x) = K_\phi\left(\frac{1}{h_1(\mu)}h(x) - \left(\frac{B_h\, h_1'(\mu)}{h_1^2(\mu)} + h_2'(\mu)\right) I_d(x)\right)$$

with $I_d(x) = x$, and τ' is the derivative of the function τ.

The following general condition will be assumed in all the paper:

(C) h_1 is not null in a neighborhood of μ.

Here are our main results.

4.1. Pointwise asymptotic laws

Consider the following hypotheses based on the functions h, τ, h_1, h_2. The *A1.x* series concern the almost-sure limits and the *A2.x* the asymptotic normality.

(A1.1) $\mathbb{E}h(X) < \infty$;
(A1.2) τ is a continuous function on $\mathcal{V}_X$;
(A1.3) for $i \in \{1, 2\}$, $h_i(\mu) < \infty$ and h_i is continuous on $\mathcal{V}_X$.

(A2.1) $\mathbb{E}h^2(X) < \infty$, $\mathbb{E}\,(X\,h(X)) < \infty$;
(A2.2) τ is continuously differentiable such that $\tau' \neq 0$;
(A2.3) $\forall i \in \{1,2\}$, $h_i(\mu) < \infty$, h_i is continuously differentiable at μ.

We have :

Theorem 1. *Suppose that the conditions (**C**), (**A1.1**), (**A1.2**) and (**A1.3**) are satisfied, then T_n converges almost surely to T.*

Theorem 2. *Suppose that the conditions (**C**), (**A2.1**), (**A2.2**) and (**A2.3**) are satisfied, and K_ϕ is finite. Then*
(a) we have the following asymptotic representation in the empirical functional process

$$\sqrt{n}\,(\tau(I_n) - \tau(I)) = \mathbb{G}_n\,(F_\phi) + o_{\mathbb{P}}(1),\ as\ n \to +\infty,$$

where

$$F_\phi = \tau'\left(\frac{\mathcal{E}h(Y)}{h_1(\mu)} - h_2(\mu)\right)\left(\frac{1}{h_1(\mu)}\,h - \left(\frac{\mathcal{E}h(Y)\,h_1'(\mu)}{h_1^2(\mu)} + h_2'(\mu)\right)\,I_d\right)$$

(b) and we have the convergence in distribution, as n tends to infinity, of $\sqrt{n}(T_n(\phi,X) - T(\phi,X))$ to centered normal Gaussian law:

$$\sqrt{n}(T_n(\phi,X) - T(\phi,X)) \rightsquigarrow \mathcal{N}(0,\sigma_\phi^2),$$

where

$$\begin{aligned}\sigma_\phi^2 &= \int\left(F_\phi(x) - \int F_\phi(x)\,d\mathbb{P}_X(x)\right)^2\,d\mathbb{P}_X(x)\\ &= a_\phi^2\,\mathbb{E}\,(h(X) - \mathbb{E}h(X))^2 + b_\phi^2\,\mathbb{E}\,(X-\mu)^2\\ &- 2a_\phi b_\phi\,\mathbb{E}\,(h(X) - \mathbb{E}h(X))\,(X-\mu)\,,\end{aligned}$$

with

$$a_\phi = \frac{K_\phi}{h_1(\mu)}\ \text{ and } b_\phi = K_\phi\left(\frac{B_h\,h_1'(\mu)}{h_1^2(\mu)} + h_2'(\mu)\right).$$

Remark. The main result is the one given in Point (a). From it, Point (b) is deduced in a straightforward way.

The results above cover all the *TLIM* class. They should be particularized for the practitioner who would pick one of the elements of that class for analyzing data. Here are then the details for each case.

4.2. Particular cases for pointwise results

a. *The Theil's measure*
The empirical form of Theil measure is defined as follows

$$Th_n = \frac{1}{\mu_n}\frac{1}{n}\sum_{j=1}^{n} X_j \log X_j - \log \mu_n,$$

$$\forall s > 0,\ \tau(s) = s,\ h(s) = s\log(s),\ h_1(s) = s,\ h_2(s) = \log(s),$$

Denote by

$$Th = \frac{1}{\mu}\int_{\mathcal{V}_X} x\log x\, dF(x) - \log \mu$$

the continuous form of the Theil measure.

All these functions are continuous on $\mathcal{V}_X$, then the assumptions defined above become for the *a.s.* requires that $\mathbb{E}X\log X$ is finite and $0 < \mu < \infty$. As for the asymptotic normality, we need that

$$\mathbb{E}\left|X\right|^2,\ \mathbb{E}\left|X\log X\right|^2\ \mathbb{E}\left|X^2\log X\right|^2 \text{ are finite.}$$

And we have $K_\phi = 1$, $B_h = \mathbb{E}\left(X\log X\right)$. We conclude that

$$\sqrt{n}\left(Th_n - Th\right) \rightsquigarrow \mathcal{N}(0, \sigma^2_{Theil})$$

with

$$\sigma^2_{Theil} = \frac{\mathbb{E}\left(X\log X\right)^2}{\mu^2} + \frac{\mathbb{E}X^2}{\mu^2}\left(\frac{B_h}{\mu}+1\right)^2 - \frac{2\mathbb{E}\left(X^2\log X\right)}{\mu^2}\left(\frac{B_h}{\mu}+1\right) - 1.$$

b. *The Mean Logarithmic Deviation*

Let

$$MLD_n = \frac{1}{n}\sum_{j=1}^{n}\log X^{-1} - \log \mu_n^{-1}$$

be the empirical form of the Mean Logarithmic Deviation. Its theoretical form is given as folllows

$$MLD = \int_{\mathcal{V}_X}\log x^{-1}\, dF(x) - \log \mu^{-1}.$$

These specific functions are given by:

$$\forall s > 0,\quad \tau(s) = s,\quad h(s) = h_2(s) = \log s^{-1},\quad h_1(s) \equiv 1.$$

The consistency requires that $\mathbb{E}\log X < \infty$ and that $0 < \mu < \infty$ while the normality is got when

$$\mathbb{E}\,|X|^2\,, \mathbb{E}\,|\log X|^2 \text{ and } \mathbb{E}\,|X\log X| \text{ are finite.}$$

In that case, we find easily that $K_\phi = 1$, $B_h = \mathbb{E}\log X^{-1}$ and

$$\sqrt{n}\,(MLD_n - MLD) \rightsquigarrow \mathcal{N}(0, \sigma^2_{MLD})$$

where

$$\sigma^2_{MLD} = \frac{\mathbb{E}\left(X^2\right)}{\mu^2} + \mathbb{E}\left(\log^2 X\right) - \frac{2}{\mu}\mathbb{E}\,(X\log X) - (B_h+1)^2\,.$$

c. *The Champernowne's measure*

In this case, the specific functions are given by:

$$\tau(s) = 1 - \exp(s)\,,\ h(s) = h_2(s) = \log(s),\ h_1(s) \equiv 1.$$

And, the various forms are:

$$Ch_n = 1 - \exp\left(\frac{1}{n}\sum_{j=1}^{n}\log\frac{X_j}{\mu_n}\right);$$

$$Ch = 1 - \exp\left(\int \mathcal{V}_X \log\frac{x}{\mu}\,dF(x)\right).$$

We find that $Ch_n = \tau(-MLD_n)$ and $Ch = \tau(-MLD)$, where MLD is the Mean Logarithmic Deviation. As τ is continuous on $\mathcal{V}_X$, we consider the same hypotheses as in the case of Mean Logarithmic Deviation.
The function τ is continuously differentiable, we put $B_h = \mathbb{E}\log X^{-1}$ and $K_\phi = \frac{\exp(-B_h)}{\mu}$, then we have

$$\sqrt{n}\,(Ch_n - Ch) \rightsquigarrow \mathcal{N}(0, \sigma^2_{Ch})$$

with

$$\sigma^2_{Ch} = K^2_\phi\,\sigma^2_{MLD}.$$

d. *Cases of the Generalized Entropy* $(\alpha \neq 0,\ \alpha \neq 1)$; *the Atkinson's measure* $(\alpha < 1, \alpha \neq 0)$; *the Divergence of Renyi* $(\alpha > 0, \alpha \neq 1)$.

We may gather these indices into one subclass by giving different values to the function τ and to the parameter α, with this common expression

$$\forall s > 0, h(s) = h_1(s) = s^\alpha \text{ and } h_2 \equiv 0,$$

and then give a general description of the results. For that, Let $I_{n,\alpha} = \mathbb{P}_n(h)/h_1(\mu_n)$ and $I_\alpha = \int_{\mathcal{V}_X} \frac{h(x)}{h_1(\mu)} \, dF(x)$.

We require for consistency that $\mathbb{E}X^\alpha < \infty$ and that $\mu \neq 0$ and, for asymptotic normality that

$$\mathbb{E}\,|X|^{2\alpha} < \infty, \ \mathbb{E}\,|X|^2 < \infty \text{ and } \mathbb{E}\,|X|^{\alpha+1} < \infty.$$

Further, let $B_h = \mathbb{E}X^\alpha$ and $K_\phi = \tau'(I_\alpha)$. Then we get

$$\sqrt{n}\,(I_{n,\alpha} - I_\alpha) = \mathbb{G}_n\left(\frac{h}{\mu^\alpha} - \frac{\alpha\mathbb{E}X^\alpha}{\mu^\alpha + 1} I_d\right) + o^*_{\mathbb{P}}(1),$$

which tends towards a centered Gaussian process with variance

$$\sigma^2_{I_\alpha} = \frac{1}{\mu^{2\alpha}}\left(\mathbb{E}X^{2\alpha} + \frac{(\alpha B_h)^2}{\mu^2}\mathbb{E}X^2 - \frac{2\alpha B_h}{\mu}\mathbb{E}X^{\alpha+1}\right) - \frac{B_h^2}{\mu^{2\alpha}}(1-\alpha)^2. \quad (4)$$

Now, we may return to the individual cases.
d.1. *Generalized Entropy*
We find $K_\phi = 1/(\alpha(\alpha-1))$, from there, we get the variance

$$\sigma^2_{GE_\alpha} = K^2_\phi \sigma^2_{I_\alpha}, \text{where } \sigma^2_{I_\alpha} \text{ is given in Equation (4).}$$

d.2. *Atkinson's measure*
Put $K_\phi = (\mathbb{E}X^\alpha)^{(1/\alpha-1)}/\alpha$. We similarly get that

$$\sigma^2_{Atk_\alpha} = \frac{1}{\alpha^2}\,(\mathbb{E}X^\alpha)^{\left(\frac{1-\alpha}{\alpha}\right)^2}\,\sigma^2_{I_\alpha}.$$

d.3. *Divergence of Renyi*
By taking $K_\phi = ((\alpha-1)\mathbb{E}X^\alpha)^{-1}$, we obtain by the same way, that

$$\sigma^2_{DR_\alpha} = \frac{\sigma^2_{I_\alpha}}{((\alpha-1)\,\mathbb{E}X^\alpha)^2}$$

where $\sigma^2_{I_\alpha}$ is given in (4).

e. *Case of the Kolm's measure*

This index is defined for $\alpha > 0$, and its specific functions are:

$$\tau(s) = \frac{1}{\alpha}\log(s), \ h(s) = h_1(s) = \exp(-\alpha s), \ h_2(s) \equiv 0, \ \forall s > 0.$$

Its empirical form is given by

$$Ko_{n,\alpha} = \log\left(\frac{1}{n}\sum_{j=1}^{n}\exp\left(-\alpha(X_j-\mu_n)\right)\right)^{1/\alpha};$$

and its theoretical form is defined as follows

$$Ko_\alpha = \log\left(\int_{\mathcal{V}_X}\left(\frac{e^{-x}}{e^{-\mu}}\right)^{\alpha} dF(x)\right)^{1/\alpha}.$$

We need for consistency that $\mu < \infty$ and that $\mathbb{E}\exp(-\alpha X) < \infty$ and, for asymptotic normality that

$$\mathbb{E}\left(|X|^2\right), \mathbb{E}\left(\left|e^{-\alpha X}\right|\right), \mathbb{E}\left(\left|e^{-2\alpha X}\right|\right) \text{ and } \mathbb{E}\left(\left|Xe^{-\alpha X}\right|\right) \text{ are finite.}$$

Then we have $B_h = \mathbb{E}\left(e^{-\alpha X}\right)$ and $K_\phi = (\alpha B_h e^{\alpha\mu})^{-1}$.

Put

$$I_{n,\alpha} = \frac{1}{e^{-\alpha\mu_n}}\frac{1}{n}\sum_{j=1}^{n}e^{-\alpha X_j} \text{ and } I_\alpha = \frac{1}{e^{-\alpha\mu}}\int_{\mathcal{V}_X} e^{-\alpha x}\, dF(x).$$

Then

$$\sqrt{n}\left(I_{n,\alpha}-I_\alpha\right) = \mathbb{G}_n\left(e^{\alpha\mu}\left(h+\alpha B_h I_d\right)\right) + o^*_{\mathbb{P}}(1).$$

Since $Ko_{n,\alpha} = \tau(I_{n,\alpha})$, we deduce that

$$\sigma^2_{Ko_\alpha} = \frac{\mathbb{E}e^{-2\alpha X}}{(\alpha B_h)^2} + \mathbb{E}X^2 + \frac{2}{\alpha B_h}\mathbb{E}\left(Xe^{\alpha X}\right) - \left(\frac{1}{\alpha}+\mu\right)^2.$$

Finally, we summarize the used abbreviations in Table 1, and, for each index, the expression of the function F_ϕ and $\mathbb{P}(F_\phi)$ in Table 2 where we can find the expressions of a_ϕ and b_ϕ.

Notations	Indices
$GE(\alpha)$, $\alpha \neq 0, 1$	Generalized Entropy with parameter α
THEIL	Theil
MLD	Mean Logarithmic Deviation
ATK(α), $\alpha < 1$ and $\alpha \neq 0$	Atkinson with parameter α
CHAMP	Champernowne
KOLM(α)$\alpha > 0$	Kolm with parameter α
DR(α)$\alpha \geq 0, \alpha \neq 1$	Divergence of Renyi with parameter α

Table 1. Notations of the indices

Indices	B_h	$F_\phi(x),\ \forall x \in \mathcal{V}_X$	$\mathbb{P}(F_\phi)$
$GE(\alpha)$	$\int_{\mathcal{V}_X} x^\alpha \, dF(x)$	$\frac{1}{\alpha(\alpha-1)} \frac{1}{\mu^\alpha} \left(x^\alpha - \frac{\alpha B_h}{\mu} x \right)$	$\frac{-B_h}{\alpha \mu^\alpha}$
THEIL	$\int_{\mathcal{V}_X} x \log x \, dF(x)$	$\frac{1}{\mu} \left(x \log x - \left(\frac{B_h}{\mu} + 1 \right) x \right)$	-1
MLD	$\int_{\mathcal{V}_X} \log x^{-1} \, dF(x)$	$\frac{1}{\mu} x - \log x$	$1 + B_h$
ATK(α)	$\int_{\mathcal{V}_X} x^\alpha \, dF(x)$	$\frac{B_h^{1/\alpha}}{\mu} \left(\frac{1}{\mu} x - \frac{B_h^{-1}}{\alpha} x^\alpha \right)$	$\left(1 - \frac{1}{\alpha}\right) \frac{B_h^{1/\alpha}}{\mu}$
CHAMP	$\int_{\mathcal{V}_X} \log x \, dF(x)$	$\left(\frac{1}{\mu} x - \log x \right) \frac{\exp(B_h)}{\mu}$	$\frac{1 - B_h}{\mu} \exp(B_h)$
KOLM(α)	$\int_{\mathcal{V}_X} \exp(-\alpha x) \, dF(x)$	$x + \frac{1}{\alpha B_h} \exp(-\alpha x)$	$\mu + \frac{1}{\alpha}$
DR(α)	$\int_{\mathcal{V}_X} x^\alpha \, dF(x)$	$\frac{1}{\alpha - 1} \left(\frac{1}{B_h} x^\alpha - \frac{\alpha}{\mu} x \right)$	-1

Table 2. Summary of the functions F for each index

5. Proof of Theorems 1 and 2

Proof of Theorem 1.

On one hand, denote by

$$I_n = \frac{\mathbb{P}_n(h)}{h_1(\mu_n)} - h_2(\mu_n) \text{ and } I = \frac{\mathbb{P}(h)}{h_1(\mu)} - h_2(\mu), \tag{5}$$

by decomposing the difference of I_n and I, we get the next equality

$$I_n - I = \frac{(\mathbb{P}_n - \mathbb{P})(h)}{h_1(\mu_n)} - \frac{\mathbb{P}(h)}{h_1(\mu)h_1(\mu_n)}\left(h_1(\mu_n) - h_1(\mu)\right) - \left(h_2(\mu_n) - h_2(\mu)\right).$$

As for all $i = 1, 2$; the function h_i is continuous on $\mathcal{V}_X$ and using the fact that μ_n converges almost surely to μ, then we have when n tends to infinity

$$h_i(\mu_n) \xrightarrow{as} h_i(\mu) < \infty. \tag{6}$$

We have also

$$(\mathbb{P}_n - \mathbb{P})(h) = \frac{1}{n}\sum_{j=1}^{n}\left(h(X_j) - \mathbb{E}h(X_j)\right).$$

Or the sequence of the random variables $\{h(X_j)\}_{j=1,\cdots,n}$ is independent and identically distributed, and as $\mathbb{E}h(X) < \infty$ by the hypothesis (**A1.1**), then the Law of Large Numbers implies that

$$(\mathbb{P}_n - \mathbb{P})(h) \xrightarrow{as} 0. \tag{7}$$

Finally, using (6) and (7), we get

$$I_n \xrightarrow{as} I, \text{ when } n \to \infty.$$

On the other hand, as τ satisfies the hypothesis (**A1.2**), then we deduce that

$$T_n \xrightarrow{as} T, \text{ when } n \to \infty. \square$$

Proof of Theorem 2.

Using the equation (5), we have

$$I_n - I = \frac{(\mathbb{P}_n - \mathbb{P})(h)}{h_1(\mu_n)} - \frac{B_h}{h_1(\mu)h_1(\mu_n)}\left(h_1(\mu_n) - h_1(\mu)\right) - \left(h_2(\mu_n) - h_2(\mu)\right).$$

Since h_i is continuously differentiable at μ for $i = 1, 2$, we get

$$h_i(\mu_n) - h_i(\mu) = h_i'(\mu)\,(\mathbb{P}_n - \mathbb{P})(I_d) + o_{\mathbb{P}}(n^{-\frac{1}{2}}).$$

Then

$$I_n - I = \frac{(\mathbb{P}_n - \mathbb{P})\,(h)}{h_1(\mu_n)} - \frac{B_h}{h_1(\mu)\,h_1(\mu_n)}\left(h_1'(\mu)\;(\mathbb{P}_n - \mathbb{P})\,(I_d) + o^*_{\mathbb{P}}(n^{-\frac{1}{2}})\right)$$

$$-h_2'(\mu)\;(\mathbb{P}_n - \mathbb{P})\,(I_d) + o^*_{\mathbb{P}}(n^{-\frac{1}{2}}).$$

But

$$\frac{B_h}{h_1(\mu)\,h_1(\mu_n)} o_{\mathbb{P}}(n^{-\frac{1}{2}}) + o^*_{\mathbb{P}}(n^{-\frac{1}{2}}) = o_{\mathbb{P}}(n^{-\frac{1}{2}}),$$

then we get the next expression

$$I_n - I = \frac{(\mathbb{P}_n - \mathbb{P})\,(h)}{h_1(\mu)} - \left(\frac{B_h\,h_1'(\mu)}{h_1^2(\mu)} + h_2'(\mu)\right)(\mathbb{P}_n - \mathbb{P})\,(I_d) + o_{\mathbb{P}}(n^{-\frac{1}{2}}).$$

Then

$$\sqrt{n}(I_n - I) = \frac{1}{h_1(\mu)}\mathbb{G}_n\,(h) - \left(\frac{B_h\,h_1'(\mu)}{h_1^2(\mu)} + h_2'(\mu)\right)\mathbb{G}_n\,(I_d) + o_{\mathbb{P}}(1).$$

By the linearity property of $\mathbb{G}_n$, we get

$$\sqrt{n}\,(I_n - I) = \mathbb{G}_n\left(\frac{1}{h_1(\mu)}\,h - \left(\frac{B_h\,h_1'(\mu)}{h_1^2(\mu)} + h_2'(\mu)\right)\,I_d\right) + o_{\mathbb{P}}(1).$$

Since K_ϕ is finite by assumption, we apply a gain the delta-method to the function τ to have

$$\sqrt{n}\,(\tau(I_n) - \tau(I)) = \mathbb{G}_n\left(K_\phi\left(\frac{1}{h_1(\mu)}\,h - \left(\frac{B_h\,h_1'(\mu)}{h_1^2(\mu)} + h_2'(\mu)\right)\,I_d\right)\right) + o_{\mathbb{P}}(1).$$

Using the notations of the equation (2), therefore

$$\sqrt{n}\,(\tau(I_n) - \tau(I)) = \mathbb{G}_n(a_\phi h - b_\phi I_d) + o^*_{\mathbb{P}}(1) = \mathbb{G}_n(F_\phi) + o_{\mathbb{P}}(1)$$

and we easily obtain by (3) the variance σ_ϕ^2. This ends the proof of Theorem 2. ■

Data	Years of data collection	Number of households	Mean of the expenses
ESAM2	2001-2002	6565	995.20
ESPS	2005-2006	13568	898.70

Table 3. Descriptive Statistics for the Distribution

6. Data driven applications and variance computations

We here give data driven applications to show how our results work. We consider the ESAM2 (Enquête Sénégalaise auprès des Ménages, 2$^{\text{ème}}$ édition) and the ESPS (Enquête de Suivi de la Pauvreté au Sénégal) databases respectively collected in 2001-2002 and in 2005-2006. (See ANSD SENEGAL (2001-2006)). Sénégal is a member of UEMOA. In both databases, we consider expense variables aggregated at the level of Heads households as indicators of welfare.

We present the data in Table 3.

We proceeded to the computations of the inequality measures and the corresponding variances using R Development Core Team (2009). We obtained the results in Table 4.

ESAM2		
TLIM	T(in %)	σ^2_ϕ
$GE(.5)$	36.362	1.643
$GE(2)$	100.984	148.274
THEIL	43.102	4.371
MLD	34.286	1.024
ATK(.5)	17.355	0.339
ATK(−.5)	37.497	0.532
CHAMP	4.846	1.636
DR(.5)	19.061	0.339
DR(2)	110.515	65.043

ESPS		
TLIM	T(in %)	σ^2_ϕ
$GE(.5)$	22.684	0.238
$GE(2)$	34.206	3.709
THEIL	24.007	0.411
MLD	23.060	0.202
ATK(.5)	11.021	0.053
ATK(−.5)	29.591	0.280
CHAMP	3.334	0.519
DR(.5)	11.677	0.053
DR(2)	52.124	5.231

Table 4. Results of the variances computations

7. Conclusion

The family we introduced allows a flexible and unified approach in the asymptotic theory of a class of inequality indices. In parallel, the computer packages also may be presented in more compact forms. We illustrated both aspects (theoretical and computational) in the paper. Hence the practitioner has all he needs about these indices in one place. But we only studied the finite dimensional limits. In a future paper, we will try to present uniform asymptotic laws of the family index by the parameter $\phi = (\tau, h, h_1, h_2)$.

References

ANSD SENEGAL (Agence Nationale de la Statistique et de la Démographie), Dakar, Sénégal. *www.ansd.sn.*

Atkinson, A.B. (1970). On the Measurement of Inequality, *Journal of Economic Theory,* **2**, 244-263.

Barrett, G., & Donald, S. (2009). Statistical inference with generalized Gini indices of inequality, poverty, and welfare. *J. Bus. Econom. Statist.*, **27**(1), 1-17. http://dx.doi.org/10.1198/jbes.2009.0001

Champernowne, D.G. and Cowell, F. A. (1998). *Economic inequality and income distribution,.* Cambridge: Cambridge University Press.

Cowell, F.A. (1980a). Generalized entropy and the measurement of distributional change. *European Economic Review,* **13**, 147-159.

Cowell, F.A. (1980b). On the structure of additive in equality measures. *Review of Economic Studies,* **47**, 521-531.

Cowell, F.A. (2000). Measurement of inequality. In A.B. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Chapter 2, pp.87-166. Amsterdam: North Holland.

Cowell, Frank A. (2003). *Theil, Inequality and the Structure of Income Distribution.* London School of Economics and Political Sciences. available at: *http://eprints.lse.ac.uk/2288/*.

Dalton H. (1920). The Measurement of Inequality of Incomes, *Economic Journal*, vol. **30**, 348-361.

Davidson, R. and Duclos, J.Y. (2000). Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality. *Econometrica*, *68* (6), 1435-1464.

Greselin, F., Puri, M.L. and Zitikis, R. (2009). L-functions, processes, and statistics in measuring economic inequality and actuarial risks. *Statistics and Its Interface,* **2**, 227-245.

Harsanyi, J. C. (1953). Cardinal utility in welfare economics of concentration. *Journal of the Royal Statistical Society, Series A 123,* 423-34.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics and interpersonal comparisons of utility. *Journal of Political Economy* 63, 309-321.

Kolm S. (1976a): Unequal Inequalities I, *Journal of Economic Theory,* **12**, 416-442.

Kullback, S. (1959). *Inference Theory and Statistics.* New-York: John Wiley.

R Development Core Team (2009) *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Renyi A. (1961). On measures of entropy and information. In 4^{th} Berkeley Symposium on Mathematical Statistics and Probability, **1**, 547-561.

Rothschild, M. and Stiglitz, J. E. (1973). Some further results on the measurement of inequality. *Journal of Economic Theory,* **6**, 188-203.

Sen, A.K. (1973). *on Economic Inequality.* Oxford: Clarendon Press.

Temkin, L. (1993). *Inequality.* Oxford University Press.

Theil, H. (1967). Economics and Information Theory, Amsterdam, North Holland.

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