A COLLECTION OF PAPERS IN MATHEMATICS AND RELATED SCIENCES,
A FESTSCHRIFT IN HONOUR OF THE LATE GALAYE DIA
EDITORS : PROFS HAMET SEYDI, GANE SAMB LO, ABOUBAKARY DIAKHABY
SPAS-EDS (SAINT-LOUIS, CALGARY, 2018), WWW.STATPAS.ORG/SPASEDS/
IN EUCLID (WWW.PROJECTEUCLID.ORG)
DOI : 10.16929/SBS/2018.100

CHAPTER 11

Asymptotic Theory and Statistical Decomposability gap Estimation for Takayama's Index, *by M.C Haidara, T.A. Kpanzou, P.D. Mergane and G.S. Lo*

Mohamed Cheikh Haidara [1], chheikhh@yahoo.fr
Tchilabalo Abozou Kpanzou [2], kpanzout@gmail.com
Pape Djiby Mergane[1], merganedjiby@gmail.com
Gane Samb Lo [1,3], gane-samb.lo@ugb.edu.sn, gslo@aust.edu.ng
[1] LERSTAD, Gaston Berger University, Saint-Louis, Senegal [2] Université de Kara, Togo. [3] African University of Sciences and Technology, AUST, Abuja, Nigeria

Abstract. The asymptotic statistical representation of the non-decomposable Takayama's index and its statistical decomposability gap estimations are addressed.

Keywords. Welfare index; Asymptotic Representation; Asymptotic Laws; Statistical estimation of decomposability; Welfare Axiomatic; Functional Gaussian Process; Gaussian Fields .
AMS 2010 Mathematics Subject Classification. 62G30; 62P20; 91B82

Cite the chapter as : :
Haidara M.C, Kpanzou T.A, Mergane P.D and Lo G.S (2018). Asymptotic Theory and Statistical Decomposability gap Estimation for Takayama's Index. In *A Collection of Papers in Mathematics and Related Sciences, a festschrift in honour of the late Galaye Dia* (Editors : Seydi H., Lo G.S. and Diakhaby A.). Spas Editions, Euclid Series Book, pp. 195 –220.
Doi : 10.16929/sbs/2018.100-02-08

A Collection of Papers in Mathematics and Related Sciences, a festschrift in honour of the late Galaye Dia. **Haidara M.C, Kpanzou T.A, Mergane P.D, and Lo G.S (2018). Asymptotic Theory and Statistical Decomposability gap Estimation for Takayama's Index**. Pages 195 — 220.

Full Abstract. In the spirit of recent asymptotic works on the General Poverty Index (GPI) in the field of Welfare Analysis, the asymptotic statistical representation of the non-decomposable Takayama's index, which has failed to be incorporated in the unified GPI approach, is addressed and established here. This representation also allows to extend to it, recent results of statistical decomposability gaps estimations. The theoretical results are applied to real databases. The conclusions of the undertaken applications recommend to use Takayama's index as a practically decomposable one, in virtue of the low decomposability gaps with respect to the large values of the index.

1. Introduction

1.1. General introduction, motivations and objectives. In this paper, we are concerned with the asymptotic theory of the (Takayama (1979)) welfare statistic and the estimation of its decomposability gap. Let us begin to define it. Let X_1, X_2, etc., be independent observations of a non-negative random variable X with cumulative distribution function (cdf) $F_{(1)}$, all of them defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and for each $n \geq 1$, let μ_n be the sample mean.

In Welfare Analysis, a poverty line is set as the lowest income under which an individual is declared poor. The number of poor individuals in the sample is denoted by Q_n. Now, we consider for each $n \geq 1$, the Takayama statistic as defined by

$$T_n(X) = 1 + \frac{1}{n} - \frac{2}{\mu_n n^2} \sum_{j=1}^{q} (n - j + 1)\, X_{j,n}. \tag{1.1}$$

We will see in Theorem 29 that, under suitable conditions, Takayama's index (1.1) converges in probability to

$$T = 1 - \frac{2}{\mu} \int y(1 - F_{(1)}(x)) \mathbb{1}_{(x \leq Z)} dF_{(1)}(x). \tag{1.2}$$

as $n \to +\infty$. Accordingly, the number T is defined as the Takayama parameter of the *cdf* $F_{(1)}$.

This statistic has been extensively studied by many authors from the axiomatic point of view. Indeed, Welfare Analysis researchers investigate the quality of an index with respect to a number of desired axioms. In that sense, the review paper of Zheng (1997) is a useful reference. We are going to quote Zheng (1997), just to highlight its importance and, for this reason, we will not enter into the details of the meanings of these axioms. According to Zheng (1997), Takayama's measure satisfies the following axioms : Focus, symmetry, replication invariance, continuity, minimal transfer, restricted continuity, nonpoverty growth, normalization. And it fails to fulfill the others : weak transfer, progressive transfer, decomposability, regressive transfer, weak transfer sensitivity, subgroup consistency, weak monotonicity, strong monotonicity.

With respect to its relation with the Gini inequality, the Takayama measure is a smoother translation of the Gini coefficient; but such advantage is obtained at a substantial cost. (See Zheng (1997)). As Takayama himself admitted, his measure may violate the monotonicity axioms, which is a serious drawback. It violates every axiom that the Sen measure fails to satisfy except continuity and replication invariance. The claim by Takayama that his estimator is superior to that of Sen has been challenged..".

One of the most desired axiom of a welfare measure is the decomposability one. Let us explain this property at a statistical level.

Suppose that we are monitoring some index I over a given population of size N. When I is applied to the whole population, we may use the notation $I = I_N$. In a large population subjected to a number of inequalities between areas and in which there are groups with specific features at the exclusion of the others, public policy efficiency usually requires to target disadvantaged areas or groups and to implement therein strong strategies aimed at improving the status of this group in relation to a given pattern (for example poverty, health covering, education level, etc.), monitored by the index I. In such a case, the population is divided into sensitive K subgroups of interest $S_1, ..., S_K$ of respective sizes N_i, $i \in \{1, ..., K\}$, and the studied behavior is followed up by an index, say I, taking the values $I_{N_i}(i)$ in each subgroup S_i, $i \in \{1, ..., K\}$.

The index I is said to be decomposable if we may express the *global* index on the whole population with respect to the partial indices at the subgroup level as follows, that is

$$I_N = \sum_{1 \leq i \leq K} \frac{N_i}{N} I_{N_i}(i). \tag{1.3}$$

Formula (1.4) offers the practical and comfortable latitude to work at the local level with the possibility to recompose the global index at the global level. This explains why decomposable indices are so preferred, in particular the (Foster *et al.* (1984)) index of index $\alpha \geq 0$

$$FGT_n(\alpha) = \frac{1}{n} \sum_{1 \leq i \leq n} \max\left(\frac{Z - X_i}{Z},\ 0\right)^{\alpha}, \quad \alpha \geq 0.$$

The problem is that some the most interesting measures are not decomposable, in particular the weighted ones. Indeed, successful policies require to target disadvantaged or vulnerable groups. For example, suppose that we are dealing with poverty. A measure that counts all poor individuals with the same weight is less interesting than another that puts bigger weights to poorer individuals. A variation of such an index in the good direction tends to be negligible if the less poor individual behave better, and to be noticeable if the poorer individuals among the poor become better off.

Our problematic is to keep using weighted measures like Sen (1976), Kakwani (1980), Shorrocks (1995) and Takayama measures, to cite a few, and yet, to have a quick approach to report the global situation. The solution resides certainly in the estimation of the decomposability gap.

$$g_N = I_N - \sum_{1 \leq i \leq K} \frac{N_i}{N} I_{N_i}. \tag{1.4}$$

We will see later that we will be able to estimate this gap. Then we will be able to work at a local level and to report the global index in accurate confidence interval.

Recently, Haidara *et al.* (2012) motivated the estimation of decomposability gap of non-decomposable measures in the sense described above. But the results in Haidara *et al.* (2012), although including almost all the known measures, ignored the Takayama for the main reason that this latter is not based on the poverty deficits $Z - X_{j,n}$ nor on the relative poverty deficits $(Z - X_{j,n})/Z$, $1 \leq j \leq Q_n$, where Q_n is the number of poor individuals, numbered from 1 to j, and Z is the poverty line .

This is the main motivation of studying the Takayama's index with respect to two directions :

(a) Provide a full asymptotic theory of Takayama's index, in parallel with that of General Poverty Index (GPI) in which its fails to a part of it. This asymptotic theory is based on the use of the functional empirical process and provides the results in the form of asymptotic representation s.

(b) Based on the results of Point (a), the same functional empirical process is used again to handle the decomposability gap.

Our best achievements are the complete description of the asymptotic distribution of the generalized Takayama measure using the functional empirical process and an auxiliary empirical process we name *residual one* in Lo *et al.* (2010), and the statistical estimation of decomposability gap.

The rest of paper is organized as follows. In the rest of this Section 1, we describe the probability space on which the proofs will take place. In Section 2, an asymptotic representation theorem for the generalized Takayama index is stated and proved. In Section 3, the statistical estimation of the decomposability gap will be fully developed. In Section 4 data-driven applications using real data are provided. A conclusion section will end the paper.

1.2. Notations and Probability Space.

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 $\mathcal{F}_{(1)}$ 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 $\mathcal{F}_{(1)}$, which is endowed with the L_2-norm

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

we defined the Gaussian process $\{\mathbb{G}_{(1)}(f), f \in \mathcal{F}_{(1)}\}$, which is characterized by its variance-covariance function

$$\Gamma_{(1)}(f,g) = \int^2 (f - \mathbb{E}_X(f))(g - \mathbb{E}_X(g))d\mathbb{P}_X, (f,g) \in \mathcal{F}_{(1)}^2.$$

This Gaussian process is the asymptotic weak limit of the sequence of functional empirical processes (fep) defined as follows. Let $X_1, X_2, ...$ be a sequence of independent copies of X. 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 $\mathbb{G}_{n,(1)}$ weakly converges to $\mathbb{G}_{(1)}$ in $\ell^\infty(\mathbb{R})$, as stochastic processes indexed by $\mathcal{F}_{(1)}$, whenever it is a Donsker class. The details of this highly elaborated theory may be found in Billingsley (1968), Pollard (1984), Vaart (1996) and similar sources.

For each $n \geq 1$, we define the functional empirical process associated with X by

$$\mathbb{G}_{n,(1)}(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (f(X_i) - \mathbb{E}f(X_i)), f \in \mathcal{F}_{(1)},$$

and denote the integration with respect to the empirical measure by

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

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 (2016).

We also have to use the Renyi's representation of the random variable X of interest by means of the cumulative distribution function (*cdf*) $F_{(1)}$ as follows

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

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

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

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

$$X_1 = F_{(1)}^{-1}(U_1),\quad X_2 = F_{(1)}^{-1}(U_2),\quad etc. \tag{1.5}$$

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

To the sequences of $(U_n)_{n\geq 1}$, we also associate the sequence of real empirical functions

$$\mathbb{U}_{n,(1)}(s) = \frac{1}{n}\#\{i, 1 \leq i \leq n,\ U_i \leq s\},\ s \in (0,1)\ n \geq 1 \tag{1.6}$$

and the sequence of real uniform quantile functions

$$\mathbb{V}_{n,(1)}(s) = U_{1,n}1_{(0\leq s\leq 1/n)} + \sum_{j=1}^{n} U_{j,n}1_{((j-1)/n\leq s\leq (j/n))},\ s \in (0,1),\ n \geq 1 \tag{1.7}$$

and next, the sequence of real uniform empirical processes

$$\alpha_{n,(1)}(s) = \sqrt{n}(\mathbb{U}_{n,(1)} - s),\ s \in (0,1)\ n \geq 1 \tag{1.8}$$

and the sequence of real uniform quantile processes

$$\gamma_{n,(1)}(s) = \sqrt{n}(s - \mathbb{V}_{n,(1)}),\ s \in (0,1)\ n \geq 1. \tag{1.9}$$

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

$$\mathbb{G}_{n,r,(1)}(x) = \sqrt{n}\left(\mathbb{F}_{n,(1)}(x) - F_{(1)}(x)\right), \ x \in \mathbb{R}, \ n \geq 1, \tag{1.10}$$

where

$$\mathbb{F}_{n,(1)}(s) = \frac{1}{n}\#\{i, 1 \leq i \leq n, \ X_i \leq s\}, \ x \in \mathbb{R} \ n \geq 1, \tag{1.11}$$

is the associated sequence of empirical functions, and the associated sequence of quantile processes

$$\mathbb{Q}_{n,(1)}(x) = \sqrt{n}\left(\mathbb{F}^{-1}_{(n),(1)}(s) - F^{-1}(s)\right), \ s \in (0,1), \ n \geq 1 \tag{1.12}$$

where

$$\mathbb{F}^{-1}_{(n),(1)}(s) = X_{1,n}1_{(0 \leq s \leq 1/n)} + \sum_{j=1}^{n} X_{j,n}1_{((j-1)/n \leq s \leq (j/n))}, \ s \in (0,1), \ n \geq 1, \tag{1.13}$$

is the associated sequence of quantile processes.

By passing, we recall that $\mathbb{F}^{-1}_{(n),(1)}$ is actually the generalized inverse of $\mathbb{F}_{(n),(1)}$. In virtue of the representation (1.5), we have the following remarkable relations :

$$\mathbb{G}_{n,r,(1)}(x) = \alpha_{n,(1)}(F_{(1)}(x)), \ x \in \mathbb{R} \tag{1.14}$$

and

$$\mathbb{Q}_{n,(1)}(x) = \sqrt{n}\left(F^{-1}_{(1)}(\mathbb{V}_{n,(1)}(s)) - F^{-1}_{(1)}(s)\right) \ s \in (0,1), \ n \geq 1, \tag{1.15}$$

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

$$\mathbb{F}_{n,(1)}(x) = \mathbb{U}_{n,(1)}(F_{(1)}(x)), \ x \in \mathbb{R} \tag{1.16}$$

and

$$\mathbb{F}_{n,(1)}^{-1}(s) = F_{(1)}^{-1}(\mathbb{V}_{(n),(1)}(s)), \ s \in (0,1), \ n \geq 1. \tag{1.17}$$

As well, the real and functional empirical processes are related as follows

$$\mathbb{G}_{n,r,(1)}(x) = \mathbb{G}_{n,(1)}(f_x^*), \ \alpha_{n,(1)}(s) = \mathbb{G}_{n,(1)}(f_s), \ s \in (0,1) \ x \in \mathbb{R}, \ n \geq 1, \tag{1.18}$$

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

To finish the description, a result of (See Bahadur (1966)) that says that the addition of the sequences of uniform empirical processes and uniform quantile processes (1.8) and (1.9) is asymptotically, and uniformly on $[0,1]$, zero in probability, that is

$$\sup_{s\in[0,1]} \left|\alpha_{n,(1)}(s) + \gamma_{n,(1)}(s)\right| = o_{\mathbb{P}}(1) \text{ as } n \rightarrow +\infty. \tag{1.19}$$

This result is a powerful tool to handle the rank statistics when our studied statistics are L-statistics.

All the needed notation are now complete and will allow the expression of the asymptotic theory we undertake here.

2. The asymptotic behavior of Takayama's statistic

Let us introduce the following notation. The mean value of X is finite and is denoted by

$$\mu = \mathbb{E}(X)$$

For a measurable numerical function f, we set

$$\mathbb{P}_X(f) = \int f(x) dF_{(1)}(x)$$

and

$$\mathbb{P}_n(f) = n^{-1} \sum_{j=1}^{n} f(X_j).$$

Let us define

$$\mu_n = \mathbb{P}_n(I_d),$$

where I_d is the identity application on $\mathbb{R}$. Fix also, for $y \in \mathbb{R}_+ \setminus \{0\}$,

$$\ell(x) = x\, \mathbb{1}_{(x \leq Z)},$$

$$h(x) = x(1 - G(x))\mathbb{1}_{(x \leq Z)},$$

$$g(x) = 2\,\left(\mathbb{P}_X(h)\, \mathbb{E}^{-2}\,(X)\; x - \mathbb{E}(X)^{-1}\, h(x)\right),$$

$$q(x) = -2\mathbb{E}(X)\ell(y)^{-1}.$$

and for all $s \in [0,1]$

$$\nu(s) = q\left(F_{(1)}^{-1}(s)\right)\; \mathrm{I}_(F_{(1)}^{-1}(s) \leq Z).$$

and

$$f_s^{**}(x) = f^{*}_{F_{(1)}^{-1}(s)}(x) = \mathrm{I}_{(x \leq F_{(1)}^{-1}(s))},\; x \in \mathbb{R}.$$

Finally, we suppose that the *cdf* $F_{(1)}$ is increasing so that we have

(2.1) $\quad F_{(1)}^{-1}(F_{(1)}(x)) = x$ and $F_{(1)}(F_{(1)}^{-1}(s)) = s,$ for $\; x \in \mathbb{R},\; s \in (0,1).$

We have the following results for the asymptotic behavior of Takayama's statistic.

THEOREM 29. *Let* $0 < \mathbb{E}(X^2) < \infty$. *Suppose that the regularity condition 2.1 holds. Then we have as* $n \to \infty$

$$\sqrt{n}(T_n - T) = \mathbb{G}_{n,(1)}(g) + \beta_n(\nu) + o_{\mathbb{P}}(1), \tag{2.2}$$

with

$$\beta_n(\nu) = -\int_0^1 \mathbb{G}_{n,(1)}(f_s^{**})\nu(s)ds.$$

We also have

$$\sqrt{n}(T_n - T) = \mathbb{G}_{(1)}(g) + \beta(\nu) + o_{\mathbb{P}}(1), \tag{2.3}$$

with

$$\beta(\nu) = -\int_0^1 \mathbb{G}_{(1)}(f_s^{**})\nu(s)ds.$$

In particular, we have

$$\sqrt{n}(T_n - T) \to \mathcal{N}(0, \sigma^2), \tag{2.4}$$

where

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 + 2\sigma_{1,2} \tag{2.5}$$

and

$$\sigma_1^2 = \int_0^\infty (g(x) - \mathbb{P}_X(g))^2 \, dF_{(1)}(x),$$

$$\sigma_2^2 = \int_{[0,1]^2} \nu(s)\nu(t)(\min(s,t) - st)\, ds\, dt \tag{2.6}$$

and

$$\sigma_{1,2} = \int_0^1 \nu(s) \left(\int_{(x \leq F_{(1)}^{-1}(s))} g(x)\, dF_{(1)}(x) - s\,\mathbb{P}_X(g) \right) ds. \tag{2.7}$$

Before we begin the proof of this main theorem, the following lemma will allow us to make straightforward computations on formulas based on the functional empirical processes. (See Lo (2016) , Chapter 5, more details on this lemma and other manipulations on $o_{\mathbb{P}}(c_n)$ with positive sequences c_n, $n \geq 1$).

Lemma 17. *Let (A_n) and (B_n) be two sequences of real valued random variables defined on the same probability space holding the sequence X_1, X_2, etc.*

Let A and B be two real numbers and Let $L(x)$ and $H(x)$ be two real-valued functions of $x \in \mathbb{R}$, with $(L, H) \in \mathcal{F}_{(1)}^2$.

Suppose that

$$A_n = A + n^{-1/2}\mathbb{G}_{n,(1)}(L) + o_{\mathbb{P}}(n^{-1/2})$$

and

$$A_n = B + n^{-1/2}\mathbb{G}_{n,(1)}(H) + o_{\mathbb{P}}(n^{-1/2}).$$

Then, we have

$$A_n + B_n = A + B + n^{-1/2}\mathbb{G}_{n,(1)}(L+H) + o_{\mathbb{P}}(n^{-1/2}),$$

and

$$A_n B_n = AB + n^{-1/2}\mathbb{G}_{n,(1)}(BL+AH) + o_{\mathbb{P}}(n^{-1/2})$$

and if $B \neq 0$, *we also have*

$$\frac{A_n}{B_n} = \frac{A}{B} + n^{-1/2}\mathbb{G}_{n,(1)}(\frac{1}{B}L - \frac{A}{B^2}H) + o_{\mathbb{P}}(n^{-1/2})$$

Proof of Theorem 29. Let us begin by recalling that

$$T_n = 1 + \frac{1}{n} - \frac{2}{\mu_n n^2}\sum_{j=1}^{q}(n-j+1)\,X_{j,n}. \tag{2.8}$$

Let us denote

$$A_n = \frac{1}{n^2}\sum_{j=1}^{q}(n-j+1)\,X_{j,n}.$$

Also, let $R_n = (R_{1,n}, ..., R_{n,n})$ be the rank statistic based on $X_1, ..., X_n$. We have

$$\begin{aligned} A_n &= \frac{1}{n}\sum_{j=1}^{n}\left(1-\frac{j}{n}\right)\,X_{j,n}\,\mathbb{1}_{(X_{j,n}\leq Z)} + \frac{1}{n^2}\sum_{j=1}^{n}X_{j,n}\ \mathbb{1}_{(X_{j,n}\leq Z)} \\ &= \frac{1}{n}\sum_{j=1}^{n}\left(1-\frac{R_{j,n}}{n}\right)\,X_{j,n}\,\mathbb{1}_{(X_{j,n}\leq Z)} + \frac{1}{n}\mathbb{P}_n(\ell), \end{aligned}$$

Let us define

$$\beta_n^*(q) = -\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\mathbb{G}_{n,r,(1)}(X_j) - F_{(1)}(X_j)\right)q(X_j). \tag{2.9}$$

This process has been introduced by Lo *et al.* (2010). It may be directly related to the functional empirical process by using the Bahadur theorem (Bahadur (1966)) as explained below. Let us use the representations given Subsection 1.2 of Section 1.

Now since $\mathbb{P}_X(\ell)$ is finite, we have $n^{-1}\mathbb{P}_n(\ell) = o_{\mathbb{P}}(n^{-1})$ and then,

$$\begin{aligned}
A_n &= \frac{1}{n}\sum_{j=1}^{n}\left(1-\mathbb{G}_{n,r,(1)}(X_j)\right)\ell(X_j)+o_{\mathbb{P}}(n^{-1})\\
&= \frac{1}{n}\sum_{j=1}^{n}\left(1-F_{(1)}(X_j)\right)\ell(X_j)+\frac{1}{n}\sum_{j=1}^{n}\left(F_{(1)}(X_j)-\mathbb{G}_{n,r,(1)}(X_j)\right)\ell(X_j)+o_{\mathbb{P}}(n^{-1})\\
&= \mathbb{P}_n\left(h\right)+\frac{1}{\sqrt{n}}\beta_n^{*}\left(\ell\right)+o_{\mathbb{P}}(n^{-1}).
\end{aligned}$$

Next, we obtain

$$\sqrt{n}\left(A_n-\mathbb{P}_X(h)\right)=\mathbb{G}_{n,(1)}(h)+\beta_n^{*}\left(\ell\right)+o_{\mathbb{P}}(n^{-1/2}). \tag{2.10}$$

Finally the Takayama index can be written as

$$T_n=1+\frac{1}{n}-\frac{2}{\mathbb{P}_n\left(I\right)}\left(\mathbb{P}_n\left(h\right)+\frac{1}{\sqrt{n}}\beta_n^{*}\left(\ell\right)+o_{\mathbb{P}}(n^{-1})\right).$$

Let us go further. We recall that $T=1-\frac{2}{\mu}\mathbb{P}_X\left(h\right)$. We have

$$\sqrt{n}\left(T_n-T\right)=-2\left(\frac{\sqrt{n}\left(A_n-\mathbb{P}_X\left(h\right)\right)}{\mu_n}-\frac{\mathbb{P}_X\left(h\right)}{\mu\,\mu_n}\sqrt{n}\left(\mu_n-\mu\right)\right)+\frac{1}{n}.$$

But we also have $\sqrt{n}\left(\mu_n-\mu\right)=\mathbb{G}_{n,(1)}\left(I_d\right)$. From Equation (2.10), we get

$$\sqrt{n}\left(T_n-T\right)=-\frac{2}{\mu_n}\left(\mathbb{G}_{n,(1)}(h)+\beta_n^{*}\left(\ell\right)+o_{\mathbb{P}}(n^{-1/2})-\frac{\mathbb{P}_X\left(h\right)}{\mu}\mathbb{G}_{n,(1)}\left(I_d\right)\right)+o_p(1)$$

By applying the last conclusion in Lemma 17, we arrive at

$$\sqrt{n}\left(T_n-T\right)=-2\left\{\mathbb{G}_{n,(1)}\left(\mu^{-1}\,h-\mathbb{P}_X\left(h\right)\,\mu^{-2}\,I_d\right)+\mu^{-1}\,\beta_n^{*}\left(\ell\right)\right\}+o_p(1).$$

By using the definitions given above, in particular the definition of the function q, we may write

$$\sqrt{n}\left(T_n-T\right)=\mathbb{G}_{n,(1)}\left(g\right)+\beta_n^{*}\left(q\right)+o_p(1). \tag{2.11}$$

By (1.13), we have

$$\begin{aligned}
\beta_n^*(q) &= -\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\mathbb{G}_{n,r,(1)}(X_{j,n}) - F_{(1)}(X_{j,n})\right) \\
&= -\frac{1}{\sqrt{n}}\sum_{j=1}^{n} n\int_{\frac{j-1}{n}}^{\frac{j}{n}}\left(\mathbb{G}_{n,r,(1)}(X_{j,n}) - F_{(1)}(X_{j,n})\right) q(X_{j,n})ds \\
&= -\sqrt{n}\int_0^1\left(\mathbb{G}_{n,r,(1)}(F_{n,(1)}^{-1}(s)) - F_{(1)}(F_{n,(1)}^{-1}(s))\right) q(F_{n,(1)}^{-1}(s))ds.
\end{aligned}$$

Now, by (1.17), and next by (1.16) and by Assumption 2.1, we have

$$\begin{aligned}
\beta_n^* &= -\sqrt{n}\int_0^1\left(\mathbb{G}_{n,r,(1)}(F_{(1)}^{-1}(\mathbb{V}_{n,(1)}(s))) - F_{(1)}(F_{(1)}^{-1}(\mathbb{V}_{(n),(1)}(s)))\right) q(F_{(1)}^{-1}(\mathbb{V}_{(n),(1)}(s)))ds \\
&= -\sqrt{n}\int_0^1\left(\mathbb{U}_{n,(1)}(F_{(1)}(F_{(1)}^{-1}(\mathbb{V}_{n,(1)}(s)))) - F_{(1)}(F_{(1)}^{-1}(\mathbb{V}_{n,(1)}(s)))\right) q(F_{n,(1)}^{-1}(s)) \\
&= -\sqrt{n}\int_0^1\left(\mathbb{U}_{n,(1)}(\mathbb{V}_{n,(1)}(s)) - \mathbb{V}_{n,(1)}(s)\right) q(F_{n,(1)}^{-1}(s))ds \\
&= -\int_0^1 \sqrt{n}\left(s - \mathbb{V}_{n,(1)}(s)\right) q\left(F_{(1)}^{-1}\left(\mathbb{V}_{n,(1)}(s)\right)\right) ds \\
&- \int_0^1 \sqrt{n}\left(\mathbb{U}_{n,(1)}\left(\mathbb{V}_{n,(1)}(s)\right) - s\right) q\left(F_{(1)}^{-1}\left(\mathbb{V}_{n,(1)}(s)\right)\right) ds.
\end{aligned}$$

From Shorack *et al.* (1986), page 511, we have for any $n \geq 1$,

$$\sup_{0\leq s\leq 1}\left|\mathbb{U}_{n,(1)}\left(\mathbb{V}_{n,(1)}(s)\right) - s\right| \leq \frac{1}{n}.$$

Thus, for $n \geq 1$,

$$\begin{aligned}
\beta_n(q) = -\int_0^1 \sqrt{n}\left(s - \mathbb{V}_{n,(1)}(s)\right) q(F_{(1)}^{-1}(\mathbb{V}_{n,(1)}(s))), ds + o_{\mathbb{P}}(1) \\
= -\int_0^1 \gamma_{n,(1)}(s) q(F_{(1)}^{-1}(\mathbb{V}_{n,(1)}(s))) + o_{\mathbb{P}}(1).
\end{aligned}$$

Here, we may use the Bahadur property (See (1.19) in Subsection 1.2, Section 1). We recall that by (2.1) and (1.18) in 1.2, Section 1, we have $\alpha_{n,(1)}(s) = \mathbb{G}_{n,r,(1)}\left(F_{(1)}^{-1}(s)\right)$ and that $\nu(s) = q(F_{(1)}^{-1}(s))$. Next, we have

$$\mathbb{G}_{n,r,(1)}\left(F_{(1)}^{-1}(s)\right) = \mathbb{G}_{n,(1)}(f^{*}_{F_{(1)}^{-1}(s)}) = \mathbb{G}_{n,(1)}(f^{**}_s),$$

where, accordingly to the notation before Theorem 29, we simplified and wrote

$$f^{*}_{F_{(1)}^{-1}(s)} = f^{**}_s.$$

We get

$$\beta_n(\nu) = \int_0^1 \mathbb{G}_{n,(1)}(f^{**}_s)\, \nu(\mathbb{V}_{n,(1)}(s))\, ds + o_{\mathbb{P}}(1). \tag{2.12}$$

Now, we have

$$\begin{aligned}\beta_n(\nu) &= \int_0^1 \mathbb{G}_{n,(1)}(f_s)\, \nu(s)\, ds \\ &+ \int_0^1 \mathbb{G}_{n,(1)}(f^{**}_s)\left(\nu\left(\mathbb{V}_{n,(1)}(s)\right) - \nu(s)\right) ds + o_{\mathbb{P}}(1)\end{aligned}$$

with

$$\begin{aligned}&\left|\int_0^1 \mathbb{G}_{n,(1)}(f^{**}_s)\left(\nu\left(\mathbb{V}_{n,(1)}(s)\right) - \nu(s)\right) ds\right| \\ \leq &\left(\sum_{s\in[0,1]} \left|\mathbb{G}_{n,(1)}(f_s)\right|\right) \int_0^1 \left|\nu(\mathbb{V}_{n,(1)}(s) - \nu(s)\right| ds.\end{aligned}$$

Since $C_n = \sup_{s\in[0,1]} \left|\mathbb{G}_{n,(1)}(f_s)\right|$ weakly converges to $\sup_{s\in[0,1]} \left|\mathbb{G}_{(1)}(f_s)\right|$, which is an *a.s.* finite random variable (the supremum of the Brownian bridge on $[0,1]$ is bounded in probability), we have that the sequence C_n is bounded in probability (See lemma 8, Chapter 5, Lo (2016), page 120). Next $D_n(s) = \left|\nu\left(\mathbb{V}_{n,(1)}(s)\right) - \nu(s)\right|$ almost-surely converges to zero form the *a.s.* convergence of $\sup_{s\in[0,1]} \left|\mathbb{V}_{n,(1)}(s) - s\right|$ to zero. Since

$$D_n(s) \leq 2\left|\frac{F_{(1)}^{-1}(Z)}{\mathbb{P}_X(I_d)}\right|,$$

we may apply the Lebesgue Convergence Theorem to have, as $n \to +\infty$,

$$\left(\sup_{s\in[0,1]} \left|\mathbb{G}_{n,(1)}(f_s)\right|\right) \int_0^1 \left|\nu\left(\mathbb{V}_{n,(1)}(s)\right) - \nu(s)\right| \, ds \rightarrow_{\mathbb{P}} 0.$$

Finally, by putting together the previous facts, we have

$$\sqrt{n}(T_n - T) = \mathbb{G}_{n,(1)}(g) + \int_0^1 \mathbb{G}_{n,(1)}(f_s^{**}) \, \nu(s) \, ds + o_{\mathbb{P}}(1).$$

This is the representation (2.2).

A simple argument based on Riemann sums along with weak law criteria using characteristic functions yields

$$\sqrt{n}(T_n - T) = \mathbb{G}_{(1)}(g) + \int_0^2 \mathbb{G}_{(1)}(f_s^{**}) \, \nu(s) \, ds + o_{\mathbb{P}}(1).$$

It is clear that $\sqrt{n}(T_n - T)$ is asymptotically Gaussian $\mathcal{N}(0, \sigma^1)$ since the couple $(\mathbb{G}_{(1)}(g), \int_0^1 \mathbb{G}_{(1)}(f_s^{**}) \, \nu(s) \, ds)$ is normal. The computation of the variance-covariance of this vector is straightforward and is given below. We have

$$\begin{aligned} \sigma^2 &= \mathbb{E}\left(\mathbb{G}_{(1)}(g)^2\right) + \mathbb{E}\left(\beta(\nu)^2\right) + 2\,\mathbb{E}\left(\mathbb{G}_{(1)}(g)\,\beta(\nu)\right) \\ &=: \sigma_1^2 + \sigma_2^2 + 2\,\sigma_{1,2}. \end{aligned}$$

where $=:$ stands for the definition of the three terms of the left-hand member in the latter equality as σ_1^2, σ_2^2 and $2\sigma_{1,2}$. We easily find that

$$\sigma_1^2 = \int_0^{\infty} \left(g(x) - \mathbb{P}_X(g)\right)^2 \, dF_{(1)}(x).$$

Next, by a well-known formula, we have

$$\sigma_2^2 = \mathbb{E}\left(\beta(\nu)\beta(\nu')\right) = \int_{(0,1)^2} \nu(s)\nu(t)(\min(s,t) - st) \, ds \, dt. \tag{2.13}$$

Concerning $\sigma_{1,2}$, we have

$$\mathbb{E}\left(\mathbb{G}_{(1)}(g)\,\beta(\nu)\right) = \mathbb{E}\left(\mathbb{G}_{(1)}(g) \int_{[0,1]} \mathbb{G}_{(1)}(f_s^{**})\nu(s) \, ds\right).$$

Fubini's theorem implies that

$$\mathbb{E}\left(\mathbb{G}_{(1)}(g)\,\beta(\nu)\right) = \int_{[0,1]} \nu(s)\mathbb{E}\left(\mathbb{G}_{(1)}(g)\mathbb{G}_{(1)}(f_s^{**})\right)\,ds.$$

But, we remark that

$$\mathbb{E}\left(\mathbb{G}_{(1)}(g)\mathbb{G}_{(1)}(f_s^{**})\right)\,ds = \mathbb{P}_X\left(g\,f_s^{**}\right) - \mathbb{P}_X(g)\,\mathbb{P}_X(f_s^{**}),$$

$$\mathbb{P}_X\left(g\,f_s^{**}\right) = \int_{\mathbb{R}_+^*} 1\!\!1_{(x\leq F_{(1)}^{-1}(s))} g(x)\,dF_{(1)}(x)$$

We finally get

$$\sigma_{1,2} = \int_0^1 \nu(s)\left(\int_{(x\leq F_{(1)}^{-1}(s))} g(x)\,dF_{(1)}(x) - s\,\mathbb{P}_X(g)\right)\,ds.$$

This completes the proof of the Theorem 29. ■.

3. Statistical estimation of the default of decomposability

3.1. Introduction. In this section, we are concerned with the statistical estimation of the decomposability gap of the Takayama statistic. This statistic is surely non decomposable in the classical definition of Welfare analysts. This study comes as a continuation of the works of Haidara and Lo who first considered such an estimation. The reader is then referred to Haidara *et al.* (2012) for a general introduction on this topic.

It is remarkable that the results of Haidara and Lo extend to Takayama's statistic although they used indices based on the relative poverty gaps. The reason is that they derived their estimation from the representation (2.2). This means that such results hold whenever that representation holds.

In that sense, the coming theorem is a consequence of Formula (2.2) and Theorem 1 in Haidara *et al.* (2012).

As a result, we will focus on the data driven applications and on comparison results with the Sen measure. In this context, we will rephrase the statistical decomposability gap problem. We did this in the introduction with the deterministic index. We are going to describe it with the random index.

Now, suppose that the population is divided into K subgroups $S_1, ..., S_K$ and for each $i \in \{1, ..., K\}$, let us denote the subset of the random sample $\{X_1, ..., X_n\}$ coming from S_i by $\mathcal{E}_i = \{X_{i,1}, ..., X_{i,n_i}\}$ and then put $T_{n_i}(i) = T(X_{i,1}, ..., X_{i,n_i})$ the Takayama statistics on the i^{th} subgroup. The decomposability gap is defined by

$$gd_n = T_n - \frac{1}{n}\sum_{i=1}^{K} n_i T_{n_i}(i).$$

At this step, we have to precise our random drawing. We are going to use a probability space in the form $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2, \mathbb{P}_1 \otimes \mathbb{P}_2)$ with $\mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$. We draw the observations in the following way. In each trial, we draw a subgroup, the ith subgroup $(\mathcal{E}_i)$ having the occurring probability p_i. We define

$$\pi_{i,j}(\omega_1) = 1_{(\text{the } i^{th} \text{ subgroup is drawn at the } j^{th} \text{ trial})}(\omega_1),$$

where, $1 \leq i \leq K, 1 \leq j \leq n$. Now, given that the i^{th} subgroup is drawn at the j^{th} trial, we pick one individual in this subgroup and observe its income $X_j(\omega_1, \omega_2)$. We then have the observations

$$\{X_j(\omega_1, \omega_2),\ 1 \leq j \leq n\}.$$

We have these simple facts. First, for $1 \leq i \leq K$,

$$n_i^* = \sum_{j=1}^{n} \pi_{i,j}. \tag{3.1}$$

Let us denote the distribution of X_j given $(\pi_{i,j} = 1)$, by $F_{i,(1)}$ that is

$$\mathbb{P}(X_j \leq y\ /\pi_{i,j} = 1) = F_{i,(1}(x).$$

We simply put $F_{i,(1}(x) = F_i$, $y \in \mathbb{R}$, to keep the notation simple. Then we have

$$\begin{aligned}\forall(x \in \mathbb{R}), \mathbb{P}(X_j \leq y\ ) &= \sum_{i=1}^{K} \mathbb{P}(\pi_{i,j} = 1)\mathbb{P}(X_j \leq y\ /\pi_{i,j} = 1)\\ &= \sum_{i=1}^{K} p_i F_i(x).\end{aligned}$$

We conclude that $\{X_1, ..., X_n\}$ is an independent sample drawn from $F_{(1)}(x) = \sum_{i=1}^{K} p_i F_i(x)$, which is the mixture of the distribution functions of the subgroups incomes.

Finally, we readily see that conditionally on $n^* \equiv (n_1^*, n_2^*, ..., n_K^*) = (n_1, n_2, ..., n_K) \equiv \overline{n}$ with $n_1 + n_2 + ... + n_K = n$, $\{X_{i,j},\ 1 \leq j \leq n_i^*\}$ are independent random variables with distribution function F_i.

3.2. Notation. Given all the previous preliminaries, we are able to state similar results of Haidara *et al.* (2012). Denote for each subgroup i $(1 \leq i \leq K)$

$$g_i(x) = 2\mathbb{E}^{-2}(X^i)I_i x - 2\mathbb{E}^{-1}(X^i)\,(1 - F_i(x))\, x\mathbb{1}_{(x<Z)}$$

and

$$\nu_i(x) = -2\mathbb{E}^{-1}(X^i)x\mathbb{1}_{(x<Z)}$$

Finally introduce as in Haidara *et al.* (2012),

$$A_1 = \sum_{i=1}^{K} p_i \left\{ \int_0^1 (g - g_i)^2 (F_i^{-1}(t))dt - \left(\int_0^1 (g - g_i)(F_i^{-1}(t))dt \right)^2 \right\},$$

$$A_2 = \sum_{i=1}^{K} p_i \int_0^1 \int_0^1 (s \wedge t - st)(p_i\nu - \nu_i)(F_i^{-1}(s))(p_i\nu - \nu_i)(F_i^{-1}(t))dsdt,$$

$$A_{31} = \sum_{i=1}^{K} p_i^2 \sum_{h \neq i}^{K} p_h \int_0^1 \int_0^1 \left[F_h(F_i^{-1}(s)) \wedge F_h(F_i^{-1}(t)) \right.$$
$$\left. - F_h(F_i^{-1}(s))F_h(F_i^{-1}(t)) \right] \nu(F_i^{-1}(s))\nu(F_i^{-1}(t))dsdt,$$

$$A_{32} = \sum_{i=1}^{K} p_i \sum_{j \neq i}^{K} p_j \sum_{h \notin \{i,j\}}^{K} p_h \int_0^1 \int_0^1 \left[F_h(F_i^{-1}(s)) \wedge F_h(F_j^{-1}(t)) \right.$$
$$\left. - F_h(F_i^{-1}(s))F_h(F_j^{-1}(t)) \right] \nu(F_i^{-1}(s))\nu(F_j^{-1}(t))dsdt,$$

$$B_1 = \sum_{i=1}^{K} p_i \int_0^1 \left\{ \int_0^s (g - g_i)(F_i^{-1}(t))dt \right.$$
$$\left. - s\int_0^1 (g - g_i)(F_i^{-1}(t))dt \right\} (p_i\nu - \nu_i)(F_i^{-1}(s))ds,$$

$$B_2 = \sum_{j=1}^{K} p_j \sum_{i \neq j}^{K} p_i \int_0^1 \int_0^1 [s \wedge F_i(F_j^{-1}(t)) - sF_i(F_j^{-1}(t))],$$

$$\times(p_i\nu - \nu_i)(F_i^{-1}(s))\nu(F_j^{-1}(t))dsdt,$$

$$B_3 = \sum_{j=1}^{K} p_j \sum_{i\neq j}^{K} p_i \int_0^1 \left\{ \int_0^{F_i(F_j^{-1}(s))} (g-g_i)(F_i^{-1}(t))dt \right.$$

$$\left. -F_i(F_j^{-1}(s)) \times \int_0^1 (g-g_i)(F_i^{-1}(t))dt \right\} \nu(F_j^{-1}(s))ds,$$

$$gd = T(F_{(1)}) - \sum_{i=1}^{K} p_i T(F_i);\ gd_{0,n} = T(F_{(1)}) - \sum_{i=1}^{K} (n_i/n)T(F_i)$$

3.3. The theoretical result. We have the following result.

THEOREM 30. *Let $\mathbb{E}X^2 < \infty$ and for each $i \in 1,...,K$,*

$$0 < \int x\ dF_{(1)}(x)\ dx < +\infty$$

and, $F_{(1)}$ and each F_i, $1 \leq i \leq K$ are increasing so that they are invertible.

Then we have

$$gd_{n,0}^{*} = \sqrt{n}(gd_n - gd_0) \rightsquigarrow \mathcal{N}(0, \vartheta_1^2 + \vartheta_3^2)$$

and

$$gd_n^{*} = \sqrt{n}(gd_n - gd) \rightsquigarrow \mathcal{N}(0, \vartheta_1^2 + \vartheta_2^2)$$

with

$$\vartheta_1^2 = A_1 + A_2 + A_3 + 2(B_1 + B_2 + B_3)$$

and

$$\vartheta_2^2 = \sum_{h=1}^{K} {F_h}^2 p_h - \left(\sum_{h=1}^{K} F_h p_h\right)^2$$

for $F_h = \mathbb{E}g(X^h) - J(F_h) + \sum_{i=1}^{K} p_i \mathbb{E}F_h(X^i)\nu(X^i)$, and

$$\vartheta_3^2 = \sum_{h=1}^{K} {M_h}^2 p_h - \left(\sum_{h=1}^{K} M_h p_h\right)^2$$

for $M_h = \mathbb{E}g(X^h) + \sum_{i=1}^{K} p_i \mathbb{E}F_h(X^i)\nu(X^i)$.

Proof. Based on Formula (2.2), the proof of Theorem 1 in Haidara *et al.* (2012) applies line by line. ■.

4. Data-driven applications

A - ESAM 1 Database, 1996.

We consider the Senegalese database ESAM 1 of 1996 which includes 3278 households. We first consider the geographical decomposition into the areas (Dakar is the Capital). We have the Takayama measure values for the whole Senegal and for its ten sub-areas. The FAO scale has been used to obtain the equivalence-adult income for the households and poverty line has been taken equal to 143080 local monetary units (*CFA franc CFA*).

Area	Senegal	Kolda	Dakar	Diourbel	St-Louis	Louga
Takayama (%)	93.14	78.57	96.51	86.65	93.92	88.59
Size	3278	198	1122	231	314	174

Area	Tambacounda	Kaolack	Thies	Fatick	Ziguinchor
Takayama (%)	80.81	89.10	88.24	87.37	94.60
Size	126	316	401	180	216

Let us compute the different variances $\vartheta_1^2, \vartheta_2^2$ and ϑ_3^2 of Theorem 30 with the empirical estimations $p_i \approx n_i/n,$. We obtain for the geographical decomposability in Senegal : $\vartheta_1^2+\vartheta_2^2 = 0.0891$; $\vartheta_1^2+\vartheta_3^2 = 0.0891$ and $gd_n = 0.0203$. This gives the 95% -confidence interval :

$$dg \in [0.0101; 0.0305], \quad \sum_{i=1}^{k} \frac{n_i}{n} T_{n_i}(G_i) = 0.9111,$$

that is

$$T(F_{(1)}) = 0.9314 \in [0.9009; 0.9416]. \tag{4.1}$$

Now for a decomposition with respect to the household chief gender, we get the Takayama measure values.

Gender	Senegal	Male	female
Takayama Index	93.14%	93.66%	90.73%
size	3278	2559	719

We get here $\vartheta_1^2+\vartheta_2^2 = 2,4147.10^{-4}$ $\vartheta_1^2+\vartheta_3^2 = 2,4147.10^{-4}$, $gd_n = 0.0124$ This gives the 95% -confidence interval :

$$dg \in [-0.0008031,\ 0.0033020], \sum_{i=1}^{k} \frac{n_i}{n} T_{n_i}(G_i) = 0.9302$$

and

$$T(F_{(1)}) = 0.9314 \in [0.9310; 0.9335]. \tag{4.2}$$

We get the conclusion that, in this case, the gap of decomposability is not that low. Rather, it is statistically significant.

B - EPVC Database, 2004.

We consider the Mauritanian database EPCV of 2004 which includes 9360 households. We first consider the geographical decomposition into the areas, Nouakchott (Nktt) is the Capital. We have the Takayama measure values for the whole Mauritania and for its thirteen sub-areas. The Oxford scale has been used to obtain the equivalence-adult income for the households and poverty line has been taken equal to 94600 local monetary units (*ougiya*)

Area	Mauritania	Hod el Charghy	Hod el Gharby	Assaba
Takayama (%)	87.49	85.40	85.83	91.78
Size	9360	1211	469	514

Area	Gorgol	Brakna	Trarza	Adrar	Dakhlet Nouadhibou
Takayama (%)	75.76	77.93	84.01	88.82	98.79
Size	796	1190	1217	568	585

Area	Tagant	Guidimagha	Tiris Zemmour	Inchiri	Nktt
Takayama (%)	71.89	77.91	93.00	83.37	95.16
Size	490	234	284	205	1597

We obtain for the geographical decomposability in Mauritania: $\vartheta_1^2+\vartheta_2^2 = 0.0357908$ $\vartheta_1^2+\vartheta_3^2 = 0.0357908$ and $gd_n = 0.0167$. This gives the 95% -confidence interval :

$$dg \in [0.0128; 0.0205], \ \ \sum_{i=1}^{k} \frac{n_i}{n} T_{n_i}(G_i) = 0.8582$$

and

$$T(F_{(1)}) = 0.8749 \in [0.8453, \ 0.8787]. \tag{4.3}$$

We get the conclusion that, in thees cases, the decomposability gap is not significantly low.

C - Analysis and comparisons.

(a) In Haidara *et al.* (2008), we have seen that decomposability gaps for the Sen index, have been estimated with confidence intervals with extreme lower and upper points not more far from zero that 1 to 9 per thousand, both for the gender and for the areas decompositions. We seize this opportunity to correct that paper and to say that there is no percentage in any confidence interval concerning the decomposability gap (dg). Instead, we have absolute numbers therein.

This conclusion is backed the empirical research in Haidara *et al.* (2008), where the Sen index index has been observed as decomposable on the

ESAM data.

(b) Compared to these indices, the Takayama index seems much less decomposable, from the statistical point of view. The gap of decomposability is statistically significant and are estimated at least at 0.7%. But compared to the values of the Takayama's index, which turn around 80%, the gaps are still relatively low.

And, in this case of significant lack of decomposition, our results may be used to recompose the global index in (4.1), (4.2) and (4.3). And we see that decomposability gaps are low with respect to the values of the Takayama index values. We conclude that, based on the Senegal and Mauritania date, we may recommend to use the Takayama index as a decomposable one, at a statistical level.

5. conclusion

As in Haidara *et al.* (2012), the Takayama's index which is theoretically non-decomposable has been observed as practically a decomposable one, based on the available data. But more importantly, the asymptotic law decomposability gap has been entirely described and the decomposability gap has been confined in 95%-confidence intervals. This was possible because of the asymptotic representation of the Takayama's with respect to the functional empirical process and the Lo and Sall residual empirical process. The conclusions obtained in this paper are recommendable to other databases studies.

Bibliography

Bahadur R.R.(1966). A Note on Quantiles in Large Samples.Ann. *Math. Statist.*, 37 (3), 577-580.

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

Haidara, M. C and G. S. Lo. (2008). Sur la décomposabilité empirique des indicateurs de pauvreté. ArXiv:1701.02649

Haidara M.C and LO G.S.(2012) Statistical Estimation of Gap of Decomposability of the General Poverty Index. *International Journal of Statistics and Probability*, 1 (2). doi:10.5539/ijsp.v1n2p211

Lo, G. S., Sall, S.T. and Seck, C.T. (2009). Une théorie asymptotique des indicateurs de pauvreté . *C. R. Math. Rep. Acad. Sci. Canada*, 31 (2), 45-52.

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

Lo, G.S. and Sall. S.T.(2010). Asymptotic representation theorems for poverty indices. *Afr. Stat.* 5 (2010). 238–244. MR2920300.

Sen A. K.(1976). Poverty: An Ordinal Approach to Measurement. *Econometrica*, **44**, 219-231. http://dx.doi.org/10.2307/1912718

Foster, J., Greer, J. and Shorrocks, A.(1984). A class of Decomposable Poverty Measures. *Econometrica* **52**, 761-766. Doi : 10.2307/1913475

Kakwani, N.(1980). On a Class of Poverty Measures. *Econometrica*, **48**, 437-446. (MR0560520). Doi : 10.2307/1911106.

Mergane P. D. and Lo G.S. (2013). On the Functional Empirical Process and Its Application to the Mutual Influence of the Theil-Like Inequality Measure and the Growth. *Applied Mathematics*, 4 (7). Doi : 10.4236/am.2013.47136

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

Takayama, N.(1979). Poverty, Income Inequality, and Their Measures: Professor Sen's Axiomatic Approach Reconsidered, Econometrica **47**, 747-759.

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

Zheng, B.(1997). Aggregate Poverty Measures. *Journal of Economic Surveys*, **11** (2), 123-162. Doi : 10.1111/1467-6419.00028

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

Billingsley, P.(1968). *Convergence of Probability measures*. John Wiley, New-York.

Lo, G.S.(2016). Weak Convergence (IA). Sequences of random vectors. SPAS Books Series.(2016). *Arxiv* : 1610.05415. Doi : 10.16929/sbs/2016.0001.