Article

Measuring income disparity: A Kernel estimator of the Atkinson inequality index

Komi Agbokou

Department of Mathematics/Fa.S.T./University of Kara - Togo; ffomestein@gmail.com/kagbokou@univ-lome.tg

Abstract: World Bank macrodata for every country on our planet indicate that national incomes per capita account for a significant portion of population disparity, and these incomes follow well-known distributions documented in the literature across almost all continents. Measuring and comparing disparity is a substantial task that requires assembling the relative nature of both small and large national incomes without distinctions. This is the primary reason we consider the Atkinson inequality index (in the continuous case) in this paper, which was developed towards the end of the 20th century to measure this disparity. Since then, a nonparametric estimator for the Atkinson index has not been developed; instead, a well-known classical discrete form has been utilized. This reliance on the classical form makes the estimation or measurement of economic inequalities relatively straightforward. In this paper, we construct a kernel estimator of the Atkinson inequality index and, by extension, that of its associated welfare function. We then establish their almost sure asymptotic convergence. Finally, we explore the performance of our estimators through a simulation study and draw conclusions about national incomes per capita on each continent, as well as globally, by making comparisons with the classical form based on World Bank staff estimates derived from sources and methods outlined in "The Changing Wealth of Nations". The results obtained highlight the advantages of kernel-based measures and the sensitivity of the index concerning the aversion parameter.

Keywords: probability density, Kenel estimation, statistical inference, likehood, welfare

MSC: 60E05, 60F05, 62F10, 62G05, 62G20, 62G30.

1. Introduction

Measuring inequality is essential for understanding the distribution of wealth in a society and changes in the social structure. It helps guide public policies, particularly those related to redistribution, and measures their impact. To measure economic inequality, there is a wide variety of tools that represent different perspectives on the subject studied. After reflecting on the relevant level of analysis of inequality, this article presents the various indicators that allow us to assess the extent of economic inequality, its evolution, and its persistence. It then discusses the normative implications of the choice of indicators in relation to considerations of social justice. After being ignored for several decades, the issue of inequality is back in the spotlight. The recent resurgence of inequality raises questions and evokes serious concerns. Inequalities refer to differences that generate phenomena of social hierarchy. These differences can relate to the allocation of resources that are unequally distributed or refer to unequal access to certain goods or services; this is how we speak of income inequality. Economic inequalities traditionally refer to disparities in income and wealth. These inequalities are the subject of particular attention due to the importance of the economic dimension in the social valuation of individuals in our societies. While not all inequalities are unfair, the assessment of whether a situation is fair or unfair is always made in light of a standard of equality against which the situation is evaluated. Thus, depending on whether we value equality of situations or equality of rights, a different perspective will be taken on the same situation. It is therefore impossible to think of inequalities without reference to a conception of social justice against which a judgement will be made about reality. This renewed interest in the study of inequalities partly stems from publications that have highlighted their increase. Notably, we can cite the work of Piketty [1] on this issue. One of the main contributions of this author (along with others) is to have facilitated the dissemination of statistics, particularly on high incomes, allowing for a precise measurement

of the evolution of inequalities. Documenting the existence of inequalities and having precise information on their variation is a necessary prerequisite for any scientific debate. Far from being merely technical, the discussion surrounding the choice of a relevant indicator to measure inequalities is, in reality, a scientific challenge.

It should also be noted that the issue of the well-being of the population is of major importance in the political, social, and economic spheres. Well-being means that the population has sufficient means to meet its needs, organize its life independently, use and develop its abilities, and pursue its goals. This cannot happen without appropriate framework conditions. The term well-being is used here as a synonym for quality of life. Well-being is considered not only in its material and financial dimensions but also in a broader perspective that encompasses the immaterial situation of the population. Material resources include income and wealth, which enable individuals to meet their needs. However, other material dimensions, such as housing and work, are also taken into account when measuring well-being. Education, health, and social relations are part of the immaterial dimensions of well-being, which also encompass the legal and institutional framework that allows citizens to participate in political life and ensures the physical security of individuals. Finally, the concept of well-being includes environmental aspects such as water quality, air quality, and noise pollution. In an approach to well-being that aims to be as broad as possible, it is important to consider not only objective living conditions but also their subjective perception by the population, namely what the population thinks, for example, about its housing conditions and the state of the environment, its feeling of security, and its degree of satisfaction with the world of work and with life in general. Inequality indices aim to provide information on the situation of the population. To this end, it is necessary to grasp a large number of elements that constitute well-being and to describe its different facets. Inequality indices provide statistical information on the state and evolution of well-being in a broad context, which can serve as a basis for forming public opinion and making political decisions (see also Harper [2]).

Inequality indices are classified into categories. In this paper, we will mention the most widespread ones. Generally speaking, inequality measures fall into two broad categories depending on the approach used to calculate them: descriptive measures and normative measures (see Sen [3]). Descriptive inequality measures are usually mathematical or statistical formulas, for example, the Gini index (see Agbokou et al. [4] and Banerjee [5]). Therefore, the characteristics of these indices depend on their mathematical or statistical properties, respectively. Most inequality indices are descriptive in nature. Normative inequality indices are derived from a social welfare function based on a prior value judgement about the effects of inequality on social welfare. These measures combine the inequality index with a social evaluation and specify whether inequality is harmful or not, as well as the degree of welfare that a society loses or gains because of this inequality. Atkinson inequality indices are among the most frequently cited normative measures. Note that the inequality indices examined here do not necessarily satisfy all axioms (see Sen et al. [6]). For example, the Atkinson index satisfies almost all of the axioms, but it is not additively decomposable. It is also worth noting that there are inequality measures derived from entropy that we will not discuss here.

We denote by X a continuous random variable representing income. For simplicity, let X be strictly positive (no zero or negative income), $f(x)$ is its probability density, $F(x)$ is its distribution function representing the cumulative distribution of income, and μ is the average income. The consideration of a welfare function to define an inequality index is illustrated first and foremost by the Atkinson index. We start with the following welfare function:

$$W(\theta) = \int_{\mathbb{R}} \frac{x^{1-\theta}}{1-\theta} f(x) dx. \tag{1}$$

The inequality measure that we deduce from this simple welfare function is written as:

$$A(\theta) = 1 - \left[\int_{\mathbb{R}} \left(\frac{x}{\mu} \right)^{1-\theta} f(x) dx. \right]^{\frac{1}{1-\theta}}, \tag{2}$$

where $\theta \geqslant 0$ is the inequality aversion parameter. We generally use values between 0 and 1 for θ. For $\theta = 1$, the form is indeterminate and we remove the indeterminacy by taking as a welfare function:

$$W(1) = \int_{\mathbb{R}} \log(x) f(x) dx, \tag{3}$$

and the corresponding form of the index in this case is therefore

$$A(1) = 1 - \frac{1}{\mu} \exp\left[\int_{\mathbb{R}} \log(x) f(x) dx \right]. \tag{4}$$

For $\theta \longrightarrow 0$, we come across a Rawlsian measure, Rawls [7], where only the fate of the poorest matters to society. The Atkinson index has a value in $[0, 1]$. It is worth 1 when an individual has everything and the others nothing. The welfare function associated with this index is the one that weights the observations by their rank. It is the poorest who will receive the greatest weight. We therefore deduce it, according to the formulas (1) and (2), in the form:

$$W(\theta) = \begin{cases} [\mu(1 - A(\theta))]^{1-\theta} & \text{if} \quad \theta \neq 1 , \\ \log[\mu(1 - A(\theta))] & \text{if} \quad \theta = 1. \end{cases} \tag{5}$$

Several authors have worked on the parametric estimation of the Atkinson index. We can start by citing Guerrero [8] who was one of the first to estimate the aversion parameter to determine the degree of inequality for a given income distribution, with a construction of confidence intervals. We can also cite Biewen et al. [9] and Tchamyou [10] to name only these. The aim of this paper is to provide a nonparametric estimation of the Atkinson inequality index (2), based on the kernel method. Then a simulation study will be done on simulated or real data. Finally we will make a comparative study of this estimator with the one that already exists in the literature.

2. Description, materials and methods

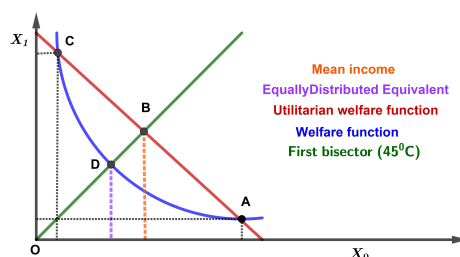### 2.1. Interpretation of the Atkinson index: graphical approach and approximation

Figure 1. Atkinson inequality index graphic description

The basis of the measurement of Atkinson inequalities is based on the concept of equitably distributed equivalent income that we denote X^*. Income X^* is the income that, if all individuals had this amount, would give the same level of social utility as the existing one ($\overline{W}$). Taking the case of two individuals, we can graphically represent the construction of the Atkinson index. Figure 1 illustrates this concept of equitably distributed equivalent. This graph shows the welfare function constructed on the space of individual incomes. The (Ox) axis shows the income of individual 1, while the (Oy) axis shows that of individual 2. If X_0 is the income of the first individual and X_1 that of the second, the average income is $\overline{X}$. Suppose the income distribution is such that point A prevails i.e. the point where we have $X_0 > X_1$. In the absence of inequality aversion ($\theta = 0$), utilitarian welfare would prevail i.e. the straight line. With this welfare function, the only way to have equal incomes at the same level of welfare is therefore to give an average income to both individuals i.e. $\overline{X}$. Since inequality aversion is zero, we are not willing to reduce the size of the butter to have more equal shares. With inequality aversion, the convex welfare function would prevail. Now, starting from A, we can find a point where incomes are equally distributed at the same level of welfare. Since the welfare function is

convex, income X^* must be less than average income $\overline{X}$. Income X^* is the abscissa of point B of the 45 degree line that has the same social welfare as A and C. Even though total income (the sum of the two individual incomes) is less than X_A, it is compensated by the gain in equality of the distribution. The reason being that, since inequality aversion is positive, we are now willing to pay the lower price of butter in order to have more equal shares. Equality is measured by the ratio $X^*/\overline{X}$. If this ratio is equal to 1 then each individual has the same level of income or if the welfare function is utilitarian (there is no perceived inequality). The approximatif Atkinson inequality index can therefore be expressed as follows (see Bell et al. [11]):

$$\widetilde{A} = 1 - \frac{X^*}{\overline{X}}. \tag{6}$$

Instinctively, this index tells us how much income we are willing to give up to have equal income. We therefore deduce that if the welfare is of the discrete form

$$\widetilde{W}(\theta) = \begin{cases} \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \dfrac{X_i^{1-\theta}}{1-\theta} & \text{if} \quad \theta \neq 1 \,, \\[4mm] \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \log(X_i) & \text{if} \quad \theta = 1 \,, \end{cases} \tag{7}$$

then the discrete Atkinson index is given by:

$$\widetilde{A}(\theta) = \begin{cases} 1 - \left[\dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \left(\dfrac{X_i}{\overline{X}} \right)^{1-\theta} \right]^{\frac{1}{1-\theta}} \,, & \text{if} \quad \theta \neq 1 \,, \\[5mm] 1 - \dfrac{1}{\overline{X}} \exp\left[\dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \log(X_i) \right] & \text{if} \quad \theta = 1 \,. \end{cases} \tag{8}$$

The formulas (or estimators if we can say so) (7) and (8) are the most used in the literature so far and would reveal inadequacies following the variation of the data since most of the data follow a distribution or at worst approximately a distribution. Hence the need to construct a non-parametric estimator or kernel estimator which often presents satisfactory results.

2.2. Construction of the kernel estimator

Kernel estimation (or Parzen-Rosenblatt method or KDE) is a non-parametric method for estimating the probability density of a random variable. It is based on a sample of a statistical population and makes it possible to estimate the density at any point of the support. In this sense, this method cleverly generalizes the histogram estimation method. Kernel density estimation is used to obtain a smooth estimate of the data distribution without making assumptions about the shape of the distribution and it is famous for its surprising smoothing character.

Let $(X_i)_{1 \leqslant i \leqslant n}$ be a random sample of size n from a population X with density function f, which represents the distribution of income. X_i, for $i \in [|1, n|]$, designating the respective income of the n individuals, are independent and identically distributed (i.i.d.) observations. The main of nonparametric density estimation is to estimate f with as few assumptions about f as possible. Among the existing kernel density estimators, we choose the classic and most famous case well known in the literature, because of its simplicity for independent observations. This estimator depends on a parameter called smoothing parameter or bandwidth parameter. It is defined by:

$$\widehat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h} \right), \quad x \in \mathbb{R} \,, \tag{9}$$

where $h = h(n) = h_n$ is the bandwidth parameter depending on the sample size n and K a probability density function called kernel function verifying certain particular properties called regularity properties.

From the estimator (9) of the distribution function f, it is obvious that we obtain an estimator of the Atkinson index (2), which is defined by:

$$\widehat{A}_n(\theta) = 1 - \left[\int_{\mathbb{R}} \left(\frac{x}{\widehat{\mu}_n} \right)^{1-\theta} \widehat{f}_n(x)dx. \right]^{\frac{1}{1-\theta}} = 1 - \left[\frac{1}{nh} \sum_{i=1}^{n} \int_{\mathbb{R}} \left(\frac{x}{\widehat{\mu}_n} \right)^{1-\theta} K\left(\frac{x-X_i}{h} \right) dx. \right]^{\frac{1}{1-\theta}} \text{ if } \theta \neq 1, \tag{10}$$

and

$$\widehat{A}_n(1) = 1 - \frac{1}{\widehat{\mu}_n} \left[\int_{\mathbb{R}} \log(x) \widehat{f}_n(x)dx \right] = 1 - \frac{1}{\widehat{\mu}_n} \exp\left[\frac{1}{nh} \sum_{i=1}^{n} \int_{\mathbb{R}} \log(x) K\left(\frac{x-X_i}{h} \right) dx. \right], \tag{11}$$

where $\widehat{\mu}_n$ is the estimator average income μ. It is easy to verify that $\widehat{\mu}_n$ satisfies:

$$\widehat{\mu}_n = \int_{\mathbb{R}} x \widehat{f}_n(x)dx = \frac{1}{n} \sum_{i=1}^{n} X_i, \tag{12}$$

and consequently, we can obtain the nonparametric estimator of (5)

$$\widehat{W}_n(\theta) = \begin{cases} [\widehat{\mu}_n(1 - \widehat{A}_n(\theta))]^{1-\theta} & \text{if} \quad \theta \neq 1, \\\\ \log[\widehat{\mu}_n(1 - \widehat{A}_n(\theta))] & \text{if} \quad \theta = 1. \end{cases} \tag{13}$$

The following paragraph provides a study of the near convergence or strong consistency of our estimators and also those of some estimators which will result from them under certain regularity hypotheses.

3. Study of the convergence of the Atkinson inequality index

In all that follows, besides the classical regularity conditions (see in Agbokou [12]), well known in the literature, on the kernel function K and on the smoothing parameter h, we consider the following hypotheses.

3.1. Assumptions

3.1.1. The model assumptions

A_1. The random variable X takes values in $\mathbb{R}_+$ or a compact subset of $\mathbb{R}_+$.

A_2. The density function f of X is uniformly continuous on its support and satisfies

$$\int_{\mathbb{R}} x^\tau f(x) = \mathbb{E}(X^\tau) > 0, \ \forall \ \tau > 0.$$

A_4. There exists a strictly positive constant η such that $\mu > \eta$ and θ takes values in $[0; 1[$.

A_5. The function $\phi : y \longmapsto y^\alpha$ on $\mathbb{R}_+$ verifies the following conditions:

i. there exist $\kappa > 1$ strictly positive constant such: $|\phi(u) - \phi(v)| \leqslant \alpha^\kappa |u - v|$, $\forall \ \alpha > 1$ and $\forall \ u, v \in \mathbb{R}_+$,

ii. there exist $C_\alpha \geqslant 1$ strictly positive constant such: $\phi(|u - v|) \leqslant C_\alpha |u - v|$, $\forall \alpha \in [0; 1]$ and $\forall \ u, v \in [0, 1]$,

iii. there exist $Q_\alpha \geqslant 1$ strictly positive constant such: $\phi(u + v) \leqslant Q_\alpha(u^\alpha + v^\alpha)$, $\forall \alpha \in [0; 1]$ and $\forall \ u, v \in \mathbb{R}_+$.

3.1.2. The kernel assumptions

K is a symmetric Kernel of bounded variation on $\mathbb{R}$ vanishing outside the interval $[a, +a]$ for some $a > 0$ and satisfying:

K_1. $\displaystyle\int_{\mathbb{R}} K(u) = 1$;

K_2. $\displaystyle\int_{\mathbb{R}} uK(u) = 0$;

K_3. $\displaystyle\int_{\mathbb{R}} u^\tau K(u) = \nu(K) > 0$, $\forall \ \tau > 0$, $\tau \neq 1$.

3.1.3. The bandwidth parameter hypothesis

The bandwidth parameter $h = (h_n)_{n\in\mathbb{N}}$ is a sequence of positive nonincreasing real numbers satisfying:

H_1. $h_n^\tau \longrightarrow 0$, $\forall \ \tau > 0$, $n \longrightarrow +\infty$;

H_2. $nh_n \longrightarrow +\infty$, $n \longrightarrow +\infty$;

Remark 1. The assumptions $A_1 - A_2$, $K_1 - K_2$ and H_2 are quite standard. $A_1 - A_4$, $A_5.i$ and $iii.$, $K_1 - K_3$ and $H_1 - H_2$ insure the strong uniform convergence of the estimators (10) and (11) to (2) and (4) respectively. Finally all the assumptions ensure the strong uniform convergence of (13) to (5).

Note 1. To simplify the writings and make them easier to handle, we adopt the following notations:

$\circ$ $\Delta(\theta) = \int_{\mathbb{R}} x^{1-\theta} f(x) dx$ and $A(\theta)$ then becomes: $A(\theta) = 1 - \left[\dfrac{\Delta(\theta)}{\mu^{1-\theta}} \right]^{\frac{1}{1-\theta}}$,

$\circ$ $\Lambda = \exp \left[\int_{\mathbb{R}} \log(x) f(x) dx \right]$ and $A(1)$ is written in the form: $A(1) = 1 - \dfrac{\Lambda}{\mu}$.

Consequently, it is clear that their kernel estimators are respectively $\widehat{\Delta}_n(\theta)$ and $\widehat{\Lambda}_n$.

3.2. Strong consistency

In this subsection, we prove the consistency of our estimator and give a rate of convergence. Our first result is the almost sure uniform convergence with an appropriate rate of the estimators $\widehat{\Delta}_n$ and $\widehat{\Lambda}_n$ stated in Proposition 1, which are the key for investigating the strong consistency of $\widehat{A}_n(\theta)$ and $\widehat{A}_n(1)$ given by Theorem 1. The last result deals with the strong consistency of the welfare function estimator $\widehat{W}_n(\theta)$ given by Corollary.

Proposition 1. *Under assumptions A_1, $A_5 - iii.$, K_1, K_3 and H_1 we have:*

(i)

$$|\widehat{\Delta}_n(\theta) - \Delta(\theta)| \longrightarrow 0 \ a.s., \quad n \to +\infty \quad \forall \, \theta \in [0,1[. \tag{14}$$

(ii) *In particular for $\theta = 1$, we have,*

$$|\widehat{\Lambda}_n - \Lambda| \longrightarrow 0 \ a.s., \quad n \to +\infty \,. \tag{15}$$

The Proposition 1 leads to the convergence of the inequality index function which is established by the following theorem:

Theorem 1. *In addition to the assumptions of the Proposition 1, if hypotheses A_2, $A_4 - A_5 - i.$, K_2 and H_2 are satisfied, then we obtain:*

(i)

$$|\widehat{A}_n(\theta) - A(\theta)| \longrightarrow 0 \ a.s., \quad n \to +\infty \quad \forall \, \theta \in [0,1[. \tag{16}$$

(ii) *In particular, if $\theta = 1$, then we get,*

$$|\widehat{A}_n(1) - A(1)| \longrightarrow 0 \ a.s., \quad n \to +\infty \,. \tag{17}$$

As a consequence of the Theorem 1, we arrive at the strong constancy of the estimator of the welfare function (13).

Corollary 1. *Under the same hypotheses as Theorem 1, assume that hypothesis $A_5 - ii.$ is verified, then we obtain:*

$$|\widehat{W}_n(\theta) - W(\theta)| \longrightarrow 0 \ a.s., \quad n \to +\infty \quad \forall \, \theta \in [0,1]. \tag{18}$$

Corollary 2. *Under the same assumptions as Theorem 1, we assume that the sequence $(X_i)_{1 \leqslant i \leqslant n}$ of i.i.d. random variables is such that $\mathbb{E}(X_i) = \mu$ and $\mathbb{V}ar(X_i) = \sigma^2 \quad \forall \, i \in \{1, \cdots, n\}$. Then for n large enough, we get:*

$$\widehat{A}_n(\theta) - A(\theta) \longrightarrow \mathcal{N}(0, s^2) \quad a.s. \quad \forall \quad \theta \in [0,1], \tag{19}$$

where

$$s^2 = (1 + \mu^2) \left(1 + \frac{\sigma^2}{\mu^2} \right).$$

3.3. Appendix of proofs

As for the proofs of the previous proposition, theorem and corollary, we need some very important or primordial lemmas whose demonstrations will be made or will be given here as well.

Let us first recall that if $\{X_i\}_1^n$ are i.i.d random variables with a distribution F. Consider a parametric function Π for which there exists an unbiased estimator. The parametric function Π can be expressed in the following form:

$$\Pi(F) = \mathbb{E}[\psi(X_1, \cdots, X_l)] = \int_{\mathbb{R}^l} \psi(X_1, \cdots, X_l) dF(x_1) \times \cdots \times dF(x_l),$$

where Ψ is a function of l i.i.d random variables from $\{X_i\}_1^n$ with $l \leqslant n$. For any function Ψ, the corresponding U-statistic for the estimation of Π based on a random sample of size n is obtained by averaging Ψ symmetrically over the observations

$$U_n = U(X_1, \cdots, X_n) = \frac{1}{\binom{n}{l}} \sum_s \psi(X_1, \cdots, X_l),$$

where $\sum_s$ represents the summation over the $\binom{n}{l}$ combinations of l distinct elements i_j, $j \in \{1, \cdots, l\}$ from $\{1, \cdots, n\}$. In particular, we have:

Lemma 1. *For $\psi(x) = x^m$ $(m > 0)$, the corresponding U-statistic is*

$$U_m = U(X_1, \cdots, X_n) = \frac{1}{n} \sum_1^n X_i^m,$$

and for n large enough, we get

$$U_m \longrightarrow \mu_m = \mathbb{E}[X^m] = \int_{\mathbb{R}} x^m dF(x) .$$

In particular for $m = 1$, we set $\widehat{\mu}_n = U_1 \longrightarrow \mu = \mu_1$.

Proof. The proof of this lemma is similar to those found in Serfling [13] and Lehmann [14]. Thus we make an exception to this one. $\square$

Lemma 2. *For every $p \in]0; 1]$, the function defined on $I = [0; +\infty)$ by $\Phi : z \longmapsto z^p$ satisfies the following condition:*

$$\forall\ u, v \in I : |\Phi(u) - \Phi(v)| \leqslant p|u - v|^p , \tag{20}$$

and the number p is the maximum of all values that verify this property.

Proof. For u, v let us suppose that $u > v$. Aware that $p - 1 \leqslant 0$, we have

$$
\begin{aligned}
|\Phi(u) - \Phi(v)| &= u^p - v^p = \int_v^u p z^{p-1} dz \\
&\leqslant \int_v^u p(z - v)^{p-1} dz = (u - v)^p \\
&\leqslant |u - v|^p \quad \text{because} \quad u > v.
\end{aligned}
$$

In this case, the corresponding p of the inequality (20) which satisfies rhe condition is equal to 1.

On the other hand, from the relation (20), setting $u = 1$ and $v = 0$. It follows $p \geqslant 1$.

Now assume that $p > 1$. If there exists another real number q such that $q > p$ and verifies the inequality (20), then we obtain

$$\forall\ u, v \in I : |\Phi(u) - \Phi(v)| \leqslant q|u - v|^q . \tag{21}$$

By setting $v = 0$ in relation (21), we have

$$|\Phi(u)| = u^p \leqslant x^q \ \forall \ u > 0.$$

Taking into account that $p - q < 0$, we get $u^{p-q} \leqslant q$. This is absurd because the left-hand side diverges towards $+\infty$ when u tends towards 0. This brings the proof of this lemma to term. □

The lemmas having been established, let's move on to the proofs of the main results obtained.

Proof of Proposition 1. *(i)* From the sequence of functions $\widehat{\Delta}_n$ in θ, we obtain

$$\widehat{\Delta}_n(\theta) = \int_{\mathbb{R}} x^{1-\theta} \widehat{f}_n(x) dx = \frac{1}{nh} \sum_{i=1}^{n} \int_{\mathbb{R}} x^{1-\theta} K\left(\frac{x - X_i}{h}\right) dx.$$

A change of variable, the use of hypotheses and $A_5 - ii$, $K_1 - K_3$, allows us to write successively

$$
\begin{aligned}
\widehat{\Delta}_n(\theta) &= \frac{1}{n} \sum_{i=1}^{n} \int_{\mathbb{R}} (X_i + wh)^{1-\theta} K(w) dw \\
&\leqslant \frac{Q_{1-\theta}}{n} \sum_{i=1}^{n} \int_{\mathbb{R}} (X_i^{1-\theta} + w^{1-\theta} h^{1-\theta}) K(w) dw \\
&\leqslant Q_{1-\theta} \left[\frac{1}{n} \sum_{i=1}^{n} X_i^{1-\theta} + h^{1-\theta} \int_{\mathbb{R}} w^{1-\theta} K(w) dw \right] \\
&\leqslant Q_{1-\theta} \left[U_{1-\theta} + h^{1-\theta} \nu(K) \right].
\end{aligned}
$$

Then we get

$$|\widehat{\Delta}_n(\theta) - U_{1-\theta}| \leqslant Mh^{1-\theta} \text{ where } M > 0; \text{ is a constant.}$$

In other words

$$|\widehat{\Delta}_n(\theta) - U_{1-\theta}| = \mathcal{O}\left(h^{1-\theta}\right) \quad a.s.$$

Furthermore, we have

$$|\widehat{\Delta}_n(\theta) - \Delta(\theta)| \leqslant |\widehat{\Delta}_n(\theta) - U_{1-\theta}| + |U_{1-\theta} - \Delta(\theta)|. \tag{22}$$

For n large enough and under hypothesis H_1, Lemma 1 applied to inequality (22) completes the first part (i) of the proof of this proposition.

(ii) The second part is inspired by the first, with a few exceptions since it is a special case. Since function $z \longmapsto exp(z)$ is a convex function, Jensen's inequality allows us to write

$$\widehat{\Lambda}_n = \exp\left[\int_{\mathbb{R}} \log(x) \widehat{f}_n(x) dx \right] \leqslant \frac{1}{n} \sum_{i=1}^{n} \int_{\mathbb{R}} x K\left(\frac{x - X_i}{h}\right).$$

The same change of variable and the use of the same hypotheses lead to

$$\widehat{\Lambda}_n \leqslant \frac{1}{n} \sum_{i=1}^{n} X_i + h \int_{\mathbb{R}} w K(w) dw \leqslant U_1 + h\nu(K) \text{ beacause here } w > 0,$$

now we get

$$|\widehat{\Lambda}_n - U_1| \leqslant Mh \text{ where } M > 0 \text{ is a constant.}$$

By using a similar inequality like that of (22) with $\widehat{\Lambda}_n$, Λ and U_1, Lemma 1 and hypothesis H_1 complete the last part of this proposition for n large enough. □

Proof of Theorem 1. (i) Thanks to hypothesis A_2, $A_5 - i$, we can write

$$
\begin{aligned}
|\widehat{A}_n(\theta) - A(\theta)| &= \left| \left[\frac{\widehat{\Delta}_n}{\widehat{\mu}_n^{1-\theta}} \right]^{\frac{1}{1-\theta}} - \left[\frac{\Delta}{\mu^{1-\theta}} \right]^{\frac{1}{1-\theta}} \right| \\
&\leqslant \left(\frac{1}{1-\theta} \right)^\kappa \left| \frac{\widehat{\Delta}_n}{\widehat{\mu}_n^{1-\theta}} - \frac{\Delta}{\mu^{1-\theta}} \right| \\
&\leqslant \left(\frac{1}{1-\theta} \right)^\kappa \left\{ \mu_{1-\theta} \left| \frac{1}{\widehat{\mu}_n^{1-\theta}} - \frac{1}{\mu^{1-\theta}} \right| + \frac{1}{\liminf_{n \longleftarrow +\infty} \widehat{\mu}_n^{1-\theta}} \left| \widehat{\Delta}_n - \Delta \right| \right\} \\
&\leqslant M \left[\left| \frac{1}{\widehat{\mu}_n^{1-\theta}} - \frac{1}{\mu^{1-\theta}} \right| + |\widehat{\Delta}_n - \Delta| \right],
\end{aligned}
$$

where

$$
0 < M = \max \left\{ \mu_{1-\theta}; \frac{1}{\liminf_{n \longleftarrow +\infty} \widehat{\mu}_n^{1-\theta}} \right\} \times \left(\frac{1}{1-\theta} \right)^\kappa.
$$

Lemma 1 associated with the Mapping Theorem allows us to obtain

$$
\left| \frac{1}{\widehat{\mu}_n^{1-\theta}} - \frac{1}{\mu^{1-\theta}} \right| \longrightarrow 0, \ a.s. \quad n \longleftarrow +\infty. \tag{23}
$$

Thus (23) and Proposition 1 – (i) completes the first part of this theorem.

(ii) As previously in (i), the same qssumptions give us:

$$
\begin{aligned}
|\widehat{A}(1) - A(1)| &= \left| \frac{\widehat{\Lambda}_n}{\widehat{\mu}} - \frac{\Lambda}{\mu} \right|, \\
&\leqslant \mu \left| \frac{1}{\widehat{\mu}} - \frac{1}{\mu} \right| + \frac{1}{\liminf_{n \longleftarrow +\infty} \widehat{\mu}_n} |\widehat{\Lambda}_n - \Lambda| \\
&\leqslant M \left\{ \left| \frac{1}{\widehat{\mu}} - \frac{1}{\mu} \right| + |\widehat{\Lambda}_n - \Lambda| \right\}
\end{aligned}
$$

where

$$
0 < M = \left\{ \mu; \frac{1}{\liminf_{n \longleftarrow +\infty} \widehat{\mu}_n} \right\}.
$$

As previously in proof (i), Lemma 1 associated with the mapping theorem and Proposition 1-(ii) complete the last part of this theorem. $\square$

Proof of Corollary 1. (i) This first part will concern the case where $\theta \in [0; 1[$. Lemma 2 and the Triangle Inequality lead us to

$$
\begin{aligned}
|\widehat{W}_n(\theta) - W(\theta)| &= \left| \left[\widehat{\mu}_n(1 - \widehat{A}_n(\theta)) \right]^{1-\theta} - \left[\mu(1 - A(\theta)) \right]^{1-\theta} \right| \\
&\leqslant (1-\theta) \left| \widehat{\mu}(1 - \widehat{A}(\theta)) - \mu(1 - A(\theta)) \right|^{1-\theta} \\
&\leqslant (1-\theta) \left\{ (1 + A(\theta)) |\widehat{\mu}_n - \mu| + \widehat{\mu}_n |\widehat{A}_n(\theta) - A(\theta)| \right\} \\
&\leqslant M_\theta \left\{ |\widehat{\mu}_n - \mu| + |\widehat{A}_n(\theta) - A(\theta)| \right\},
\end{aligned}
$$

where

$$
0 < M_\theta = \max \left\{ 1 + A(\theta); \limsup_{n \longleftarrow +\infty} \widehat{\mu}_n \right\} \times (1-\theta), \quad \forall \, \theta \in [0; 1[.
$$

Thus Lemma (1) and Theorem 1 – (i) complete the first part of this corollary.

(ii) This second and last part concerns the case where $\theta = 1$. For this case especially, we assume that $A(1)$ and its estimator $\widehat{A}_n(1)$ are all different from 1 or all do not approach 1. Therefore $\mu(1 - A(1))$ and $\widehat{\mu}_n(1 - \widehat{A}_n(1))$ are strictly positive. Let us denote $\delta = \min \left\{ \mu(1 - A(1)); \widehat{\mu}_n(1 - \widehat{A}_n(1)) \right\}$. We therefore have δ

strictly positive. Thus the function $z \longmapsto \log(z)$ is $\dfrac{1}{\delta}$–Lipschitzian on $[\delta; +\infty[$ because its derivative in absolute value is bounded above by $\dfrac{1}{\delta}$. This leads us to write:

$$
\begin{aligned}
\left| \widehat{W}_n(1) - W(1) \right| &= \left| \log[\widehat{\mu}_n(1 - \widehat{A}_n(1))] - \log[\mu(1 - A(1))] \right| \\
&\leqslant \frac{1}{\delta} \left| \widehat{\mu}_n(1 - \widehat{A}_n(1)) - \mu(1 - A(1)) \right| \\
&\leqslant M_\delta \left\{ \left| \widehat{\mu}_n - \mu \right| + \left| \widehat{A}_n(1) - A(1) \right| \right\},
\end{aligned}
$$

where

$$
0 < M_\delta = \frac{\max \left\{ 1 + A(1); \limsup_{n \longleftarrow +\infty} \widehat{\mu}_n \right\}}{\delta}.
$$

Thus Lemma (1) and Theorem 1 -(ii) complete the last part of this corollary.

$\square$

Proof of Corollary 2. ∘ For $\theta \in [0, 1[$, we directly and successively draw these inequalities from the following relation (see proof of the Proposition 1):

$$
\widehat{\Delta}_n(\theta) \leqslant Q_{1-\theta} \left[\frac{1}{n} \sum_{i=1}^{n} X_i^{1-\theta} + h^{1-\theta} \int_{\mathbb{R}} w^{1-\theta} K(w) dw \right].
$$

Using the property of monotonic increasing and the linearity of the mathematical expectation and also $\|a\| - \|b\| \leqslant |a - b|$, then we have

$$
\mathbb{E}\left[\widehat{\Delta}_n(\theta) \right] \leqslant Q_{1-\theta} \mathbb{E}\left[X^{1-\theta} \right] + h^{1-\theta} Q_{1-\theta} \nu(K) \Rightarrow \mathbb{E}\left| \widehat{\Delta}_n(\theta) - \Delta(\theta) \right| = \mathcal{O}\left(h^{1-\theta} \right). \tag{24}
$$

Moreover, we have

$$
\begin{aligned}
\left| \widehat{A}_n(\theta) - A(\theta) \right| &\leqslant M \left[\left| \frac{1}{\widehat{\mu}_n} - \frac{1}{\mu} \right| + \left| \widehat{\Delta}_n(\theta) - \Delta(\theta) \right| \right] \\
\Rightarrow \mathbb{E}\left| \widehat{A}_n(\theta) - A(\theta) \right| &\leqslant M \left[\mathbb{E}\left| \frac{1}{\widehat{\mu}_n} - \frac{1}{\mu} \right| + \mathbb{E}\left| \widehat{\Delta}_n(\theta) - \Delta(\theta) \right| \right].
\end{aligned}
$$

Thus for n large enough, we have $\mathbb{E}\left| \widehat{A}_n(\theta) - A(\theta) \right| \longrightarrow 0$. In particular we obtain

$$
\mathbb{E}\left[\widehat{A}_n(\theta) - A(\theta) \right] \longrightarrow 0 \Longleftrightarrow \mathbb{E}\left[\widehat{A}_n(\theta) \right] \longrightarrow \mathbb{E}\left[A(\theta) \right] = A(\theta). \tag{25}
$$

We have just shown that the estimator is asymptotically unbiased. We can notice that We can notice that $(a - b)^2 \leqslant a^2 + b^2$ (where a and b have the same sign) and also the function $z \longmapsto z^p$, $p \geqslant 2$ is convex (this allows us to apply Jensen's inequality). All this allows us to write

$$
\begin{aligned}
\left(\widehat{A}_n(\theta) - A(\theta) \right)^2 &\leqslant \left[\int_{\mathbb{R}} \left(\frac{x}{\widehat{\mu}_n} \right)^{1-\theta} \widehat{f}_n(x) dx \right]^{\frac{2}{1-\theta}} + \left[\int_{\mathbb{R}} \left(\frac{x}{\mu} \right)^{1-\theta} f(x) dx \right]^{\frac{2}{1-\theta}} \\
&\leqslant \int_{\mathbb{R}} \left(\frac{x}{\widehat{\mu}_n} \right)^2 \widehat{f}_n(x) dx + \int_{\mathbb{R}} \left(\frac{x}{\mu} \right)^2 f(x) dx \\
&\leqslant \frac{1}{\widehat{\mu}_n^2} \left[\frac{1}{n} \sum_{i=1}^{n} X_i^2 \int_{\mathbb{R}} K(w) dw + 2h \frac{1}{n} \sum_{i=1}^{n} X_i^2 \int_{\mathbb{R}} w K(w) dw + h^2 \int_{\mathbb{R}} w^2 K(w) dw \right] + \frac{\mu_2}{\mu^2} \\
\Rightarrow \mathbb{E}\left[\left(\widehat{A}_n(\theta) - A(\theta) \right)^2 \right] &= \mu_2 + \frac{\mu_2}{\mu^2} + \mathcal{O}\left(h^2 \right) = (\mu^2 + 1) \left(1 + \frac{\sigma^2}{\mu^2} \right) + \mathcal{O}\left(h^2 \right).
\end{aligned}
$$

Furthermore we know that

$$\mathbb{V}ar\left[\widehat{A}_n(\theta) - A(\theta)\right] = \mathbb{E}\left[\left(\widehat{A}_n(\theta) - A(\theta)\right)^2\right] - \mathbb{E}^2\left[\widehat{A}_n(\theta) - A(\theta)\right].$$

So, from the relation (24) we have

$$\mathbb{V}ar\left[\widehat{A}_n(\theta) - A(\theta)\right] = (\mu^2 + 1)\left(1 + \frac{\sigma^2}{\mu^2}\right) + \mathcal{O}\left(h^{2-2\theta}\right) \ a.s. \tag{26}$$

For n large enough, we have $\mathbb{V}ar\left[\widehat{A}_n(\theta) - A(\theta)\right] \longrightarrow s^2 = (\mu^2 + 1)\left(1 + \frac{\sigma^2}{\mu^2}\right)$.

∘ For $\theta = 1$, we also have (see proof of Proposition 1):

$$\widehat{\Lambda}_n \leqslant \frac{1}{n}\sum_{i=1}^{n} X_i + h\int_{\mathbb{R}} wK(w)dw \ \Rightarrow \mathbb{E}[\widehat{\Lambda}_n] \leqslant \mathbb{E}[X_i] + h\nu(K)\mathbb{E}|\widehat{\Lambda}_n - \mu| = \mathcal{O}(h) \ a.s.$$

On the one hand the convexity of the exponential function, the triangular inequality and $\|a| - |b\| \leqslant |a - b|$ allows us to have

$$\mathbb{E}|\widehat{\Lambda}_n - \Lambda| \leqslant \mathbb{E}|\widehat{\Lambda}_n - \mu| + \mathbb{E}|\Lambda - \mu| \implies \mathbb{E}|\widehat{\Lambda}_n - \Lambda| = \mathcal{O}(h) \ a.s.$$

On the other hand

$$|\widehat{A}_n(1) - A(1)| \quad \leqslant \quad M\left\{\left|\frac{1}{\widehat{\mu}} - \frac{1}{\mu}\right| + |\widehat{\Lambda}_n - \Lambda|\right\}.$$

Thus, for n large enough

$$\mathbb{E}|\widehat{A}_n(1) - A(1)| = \mathcal{O}(h) \ a.s. \tag{27}$$

We have also

$$\begin{aligned}
\left(\widehat{A}_n(\theta) - A(\theta)\right)^2 &\leqslant \left\{\frac{1}{\widehat{\mu}_n}\exp\left[\int_{\mathbb{R}}\log(x)\widehat{f}_n(x)dx\right]\right\}^2 + \left\{\frac{1}{\mu}\exp\left[\int_{\mathbb{R}}\log(x)f(x)dx\right]\right\}^2 \\
&\leqslant \frac{1}{\widehat{\mu}_n^2}\exp\left[2\int_{\mathbb{R}}\log(x)\widehat{f}_n(x)dx\right] + \frac{1}{\mu^2}\exp\left[2\int_{\mathbb{R}}\log(x)f(x)dx\right] \\
&\leqslant \frac{1}{\widehat{\mu}_n^2}\int_{\mathbb{R}}x^2\widehat{f}_n(x)dx + \frac{\mu_2}{\mu^2} \\
\Rightarrow \mathbb{E}\left[\left(\widehat{A}_n(\theta) - A(\theta)\right)^2\right] &\leqslant M\left[\mu_2 + h^2\nu(K)\right] + \frac{\mu_2}{\mu^2}.
\end{aligned}$$

For n large enough, we get

$$\Rightarrow \mathbb{E}\left[\left(\widehat{A}_n(\theta) - A(\theta)\right)^2\right] = \mu_2 + \frac{\mu_2}{\mu^2} + \mathcal{O}(h^2) \ a.s. \tag{28}$$

Formulas (27) and (28) give the variance as previously $\mathbb{V}ar\left[\widehat{A}_n(\theta) - A(\theta)\right] = (1 + \mu^2)\left(1 + \frac{\sigma^2}{\mu^2}\right) + \mathcal{O}(h^2) \ a.s.$ This which completes the proof of this corollary. □

4. Simulation

Before moving on to numerics, we were able to calculate the Atkinson inequality index and welfare for certain probability distributions that will be useful for the rest.

4.1. Atkinson index for some usual probability distributions

However, theoretical calculations of the Atkinson index with probability distributions are not easy or easy to determine. With other distributions, this seems very complex or even impossible. We will spare you these

details and the results obtained are grouped in the Table 1. The theoretical calculation of the Atkinson index involves certain so-called special functions which are defined by:

- Gamma function: $\Gamma(a) = \int_0^\infty z^{a-1} \exp(-z) dz$, $a \in \mathbb{R}_+$
- Digamma function: $\psi(a) = \int_0^\infty z^{a-1} \log(z) \exp(z) dz = \int_0^\infty \left(\dfrac{e^{-z}}{z} - \dfrac{e^{-az}}{1-e^{-z}} \right) dz = \dfrac{d}{da} \log[\Gamma(a)]$, $a \in \mathbb{R}_+$
- Euler - Mascheroni constant: $\gamma = -\psi(1) = -\int_0^\infty \log(z) e^{-z} dz = \lim_{n \longrightarrow +\infty} \left(\sum_{k=1}^n \dfrac{1}{k} - \log(n) \right) \simeq$ 0.577216664901532860
- Beta function: $B(a,b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, $a,\ b \in \mathbb{R}_+$

Table 1. Theoretical expressions of the Atkinson index and the welfare function

Name	p.d.f. $f(x)$	Atkinson index $A(\theta)$	Welfare function $W(\theta)$
Beta $\mathcal{B}(\alpha;\beta)$	$\dfrac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}\mathbf{1}_{[0,1]}(x)$	$\begin{cases} 1 - \dfrac{\alpha+\beta}{\alpha}\left[\dfrac{\Gamma(\alpha+1-\theta)\Gamma(\beta)}{B(\alpha,\beta)\Gamma(\alpha+\beta+1-\theta)} \right]^{\frac{1}{1-\theta}} \\ 1 - \dfrac{\alpha+\beta}{\alpha}\exp\left[\psi(\alpha) - \psi(\alpha+\beta) \right] \end{cases}$	$\begin{cases} \dfrac{\Gamma(\alpha+1-\theta)\Gamma(\beta)}{B(\alpha,\beta)\Gamma(\alpha+\beta+1-\theta)} \\ \psi(\alpha) - \psi(\alpha+\beta) \end{cases}$
Exponential $\mathcal{E}(\lambda)$	$\lambda e^{-\lambda x}\mathbf{1}_{[0,+\infty[}(x)$	$\begin{cases} 1 - \Gamma^{\frac{1}{1-\theta}}(2-\theta) \\ 1 - \lambda \exp\left[-(\gamma + \log\lambda) \right] \end{cases}$	$\begin{cases} \dfrac{\Gamma(2-\theta)}{\lambda^{1-\theta}} \\ -(\gamma + \log\lambda) \end{cases}$
Gamma $\Gamma(\alpha,\beta)$	$\dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1}e^{-\beta x}\mathbf{1}_{[0,+\infty[}(x)$	$\begin{cases} 1 - \dfrac{\Gamma^{\frac{1}{1-\theta}}(1-\theta+\alpha)}{\alpha\,\Gamma^{\frac{1}{1-\theta}}(\alpha)} \\ 1 - \dfrac{\beta}{\alpha}\exp\left[\psi(\alpha) - \log\beta \right] \end{cases}$	$\begin{cases} \dfrac{\Gamma(1-\theta+\alpha)}{\beta^{1-\theta}\Gamma(\alpha)} \\ \psi(\alpha) - \log\beta \end{cases}$
Pareto $\mathcal{P}(\alpha,\beta)$	$\dfrac{\alpha\beta^\alpha}{x^{\alpha+1}}\mathbf{1}_{[\beta,+\infty[}$	$\begin{cases} 1 - \dfrac{\alpha-1}{(\alpha+\theta-1)^{\frac{1}{1-\theta}}} \\ 1 - \dfrac{\alpha-1}{\alpha\beta}\exp\left[\dfrac{\log(\beta)+1}{\alpha} \right] \end{cases}$	$\begin{cases} \dfrac{(\alpha\beta)^{1-\theta}}{\alpha+\theta-1} \\ \dfrac{\alpha\log(\beta)+1}{\alpha} \end{cases}$
Uniform $\mathcal{U}([\alpha;\beta])$	$\dfrac{1}{\beta-\alpha}\mathbf{1}_{[\alpha;\beta]}(x)$	$\begin{cases} 1 - \dfrac{2}{\alpha+\beta}\left[\dfrac{\beta^{2-\theta}-\alpha^{2-\theta}}{(\beta-\alpha)(2-\theta)} \right]^{\frac{1}{1-\theta}} \\ 1 - \dfrac{2}{\alpha+\beta}\exp\left[\dfrac{\beta\log(\beta)-\alpha\log(\alpha)}{\beta-\alpha} - 1 \right] \end{cases}$	$\begin{cases} \dfrac{\beta^{2-\theta}-\alpha^{2-\theta}}{(\beta-\alpha)(2-\theta)} \\ \dfrac{\beta\log(\beta)-\alpha\log(\alpha)}{\beta-\alpha} - 1 \end{cases}$
Weibull $\mathcal{W}(\alpha;\beta)$	$\dfrac{\alpha}{\beta}\left(\dfrac{x}{\beta}\right)^{\alpha-1}\exp\left[-\left(\dfrac{x}{\beta}\right)^\alpha \right]\mathbf{1}_{[0,+\infty[}(x)$	$\begin{cases} 1 - \dfrac{\Gamma^{\frac{1}{1-\theta}}\left(1+\frac{1-\theta}{\alpha}\right)}{\Gamma^{\frac{1}{\alpha}}\left(1+\frac{1}{\alpha}\right)} \\ 1 - \dfrac{1}{\beta^{1/\alpha}}\exp\left[\dfrac{\log\beta+\gamma}{\alpha} \right] \end{cases}$	$\begin{cases} \dfrac{\beta^{1-\theta}}{\alpha}\Gamma\left(1+\frac{1-\theta}{\alpha}\right) \\ \dfrac{\log(\beta)+\gamma}{\alpha} \end{cases}$

Remark 2. Recall that the first lines of the expressions that are in the braces represent the general case for $\theta \neq 1$ and those of the second lines represent the particular case for $\theta = 1$.

The table being obtained, we choose one of these models to study the cobergence of our nonparametric estimator with respect to that. Our choice is the Weibull model because of its particularity in terms of application and its multitasking. The Weibull distribution intervenes in almost all areas like that of Pareto in the estimations of income inequality measures (see Agbokou [4]). To do this, with the data we have, we will determine these parameters by the Maximum Likelihood (M.L.) method to obtain the theoretical Atkinson index.

4.2. Determination of the parameters of the Weibull distribution by the M.L. method

Due to its adaptability in fitting distributions or data particularly in data science, the Weibull distribution has taken in recent years a major position in the field of parametric, semi or nonparametric estimations. One of the main barriers to a more important or wider use of the Weibull distribution is the complexity of estimating its parameters. Regrettably or sadly, the calculations that this estimation involves are not always

simple enough. This subsection deals with the estimation of maximum likelihood in samples of the Weibull probability density. The likelihood function of this sample $(X_1, \cdots, X_n)$ is given by:

$$L(X_1, \cdots, X_n, (\alpha, \beta)) = \Pi_{i=1}^n \frac{\alpha}{\beta} X_i^{\alpha-1} \exp\left(-\frac{X_i^\alpha}{\beta}\right).$$

We let $\mathcal{L}(\alpha, \beta) = \mathcal{L}(X_1, \cdots, X_n, (\alpha, \beta)) = \log L(X_1, \cdots, X_n, (\alpha, \beta))$ represent the log-likelihood, then we derive $\mathcal{L}(\alpha, \beta)$ respectively with respect to α and with respect to β then by setting the two partial derivatives equal to zero each, we obtain:

$$\begin{cases} \dfrac{\partial \mathcal{L}(\alpha, \beta)}{\partial \alpha} = \dfrac{n}{\alpha} + \displaystyle\sum_{i=1}^n \log(X_i) - \dfrac{1}{\beta} \sum_{i=1}^n X_i^\alpha \log(X_i) = 0, \\[3mm] \dfrac{\partial \mathcal{L}(\alpha, \beta)}{\partial \beta} = -\dfrac{n}{\beta} + \dfrac{1}{\beta^2} \displaystyle\sum_{i=1}^n X_i^\alpha = 0. \end{cases} \tag{29}$$

On eliminating β between these two equations of the system (29) and simpljfying the term, we get:

$$\frac{\displaystyle\sum_{i=1}^n X_i^\alpha \log(X_i)}{\displaystyle\sum_{i=1}^n X_i^\alpha} - \frac{1}{\alpha} - \frac{1}{n} \sum_{i=1}^n \log(X_i) = 0. \tag{30}$$

We note that equation (30) is difficult or even impossible to solve analytically. We will therefore use a numerical resolution or iterative methods such as the Newton method or the fixed point method or the secant and regula falsi method in order to find the estimator $\widehat{\alpha}$ of the parameter α. If the estimator $\widehat{\alpha}$ is obtained, we therefore deduce from (29), the estimator $\widehat{\beta}$ of β by the equality:

$$\widehat{\beta} = \frac{1}{n} \sum_{i=1}^n X_i^{\widehat{\alpha}}. \tag{31}$$

4.3. Real data description and simulation results

For our study, we focused on the World Bank data that could be easily found on their website DataBank [1]. We were interested in the "adjusted net national income per capita (current US dollars)" of 2021 for each country in the world (at least for those who have it because we could not find the updated data for all countries). Since the data is allocated by region according to the World Bank's breakdown, we organized this data by continent as follows:

- Europe Zone (1) = European Union
- America Zone (2) = America and the Caribbean + North America
- Asia Zone (3) = East Asia and the Pacific + South Asia + Middle East + Central Asia
- Africa Zone (4) = Sub-Saharan Africa + North Africa
- Eurasia Zone (5) = Europe Zone + Asia Zone
- World (6) = the sum of all the above zones

Remark 3. The last two groupings aim not only to study the inequalities of countries between two countries or between countries of the world but also to see the performance of our estimators when the sample size increases and also the impact of the inversion parameter on them depending on these sizes.

Table 2 provides a statistical description of some position and dispersion characteristics of these data taken to the thousandth in order to facilitate the calculations of the programming in the rest of this subsection using Matlab software.

In view of the statistical summary of Table 2, we notice that the data, whatever the area, are very heterogeneous since the coefficient of variation of each of these groupings is strictly greater than 0.30. This

---

[1] https://data.worldbank.org/indicator/NY.ADJ.NNTY.PC.KD_gl=1*1plgah*_gcl_au*NDUzNDg5MzA1LjE3MjQ4ODM3MzY.

heterogeneity character makes this study very exciting since with simulated data (which we do not present here), we always had homogeneous data which contribute to the robustness of our estimators from the size n = 30. Thus the other objective is to see the behavior of our estimators facing heterogeneous data because we know that heterogeneous data in statistical study such as estimates and forecasts negatively affect or slow down the robustness of the estimators and by ricochet impact the forecasts.

Table 2. Statistical description of data from the World Bank website

Statistic summary	Size	Minimum	Mean	Median	Mode	Maximum	Coefficient of variation
European area	27	7.6500	25.3004	19.6460	7.6500	65.3640	0.5857
America area	33	1.2510	10.6735	6.6080	1.2510	52.6190	0.9870
Asia area	51	0.3400	9.7364	3.5040	0.3400	44.5130	0.7573
Africa area	54	0.1560	1.9479	1.1455	0.1560	12.2520	0.7102
Eurasian area	78	0.3400	15.1239	9.0055	0.3400	65.3640	0.9823
World	165	0.1560	9.9217	4.0190	0.4860	65.3640	0.7618

Thus, for the simulation of our estimators, we chose as a statistical kernel K, the Epanechnikov kernel also called the "parabolic" kernel. It is named after the author Epanechnikov [15], who used and studied it for the first time in 1969. On the one hand, it is well known in the literature that the kernel has little influence on the performance of nonparametric estimators. What motivates this choice is that this kernel allows to have the most efficient estimator for the density, which concerns us in this document. On the other hand, the literature reveals that the choice of the smoothing parameter $h = h_n$ has a major influence on the robustness of the kernel estimator (see Agbokou et al. [16,17]). The process of this choice in the presence of real data is not as easy as in the presence of simulated data. We therefore opted for the numerical cross-validation method which consists in minimizing the integrated squared error defined from a series of observations or data $(x_i)_{1\leqslant i\leqslant n}$ of size n. This error is defined by:

$$\Phi(h_n) = \int_{x_{(1)}}^{x_{(n)}} \left(\widehat{A}_n(\theta) - A(\theta)\right)^2 dx, \tag{32}$$

where $x_{(1)} = \min(x_i)$ and $x_{(n)} = \max(x_i)$ for all $i \in \{1, \cdots, n\}$.

The bandwith parameter selection rule results in the minimization of this criterion:

$$\widehat{h}_n = arg \min_{h_n} \Phi(h_n). \tag{33}$$

The smoothing parameter of (33) thus obtained is asymptotically optimal. To determine the parameter h, our Matlab code is programmed in such a way that we obtain at the same time the RMSE (Residual Mean Square Error). This is the root mean square deviation, which is the standard deviation of the residuals (prediction errors). The residuals are the measure of the deviation between the data points and the regression line. The RMSE metric is the measure of the distribution of these residuals. In other words, it indicates the concentration of the data around the line of best fit. In our work we have two parameters h to determine and therefore two RMSEs. We denote them respectively h_{1n} for the estimator of the Atkinson index and h_{2n} for its associated welfare function. Similarly, their errors or residuals are respectively noted $RMSE_1$ and $RMSE_2$. To get an idea of the impact of the aversion parameter, we chose the parameter θ such that $\theta \in \{0.01, 0.05, 0.1, 0.5, 0.9, 1\}$ for each sample size (or area or region). We also calculated for each θ the classical estimators (usual classical forms) of the Atkinson index (8) and its welfare function (7) in order to see if they best estimate the theoretical Atkinson index and its theoretical welfare function.

Let us recall that the two parameters of the theoretical Atkinson index and its theoretical welfare function cannot be taken arbitrarily, we opted for the maximum likelihood (M.L.) method with the intervention of Newton's numerical method in order to be able to assign a value to each of these parameters.

The results obtained are summarized in Table 3. A visualization of all these results obtained is summarized in Figure 2.

Table 3. Comparison of theoretical values and estimates of the Atkinson index and the welfare

θ	$\hat{h}_{1n}$	$\hat{h}_{2n}$	$\hat{\alpha}$	$\hat{\beta}$	$A(\theta)$	$\hat{A}_n(\theta)$	$\tilde{A}(\theta)$	$W(\theta)$	$\hat{W}_n(\theta)$	$\tilde{W}(\theta)$	$RMSE_1$	$RMSE_2$
						European area						
0.01	0.4609	0.4268	1.8719	535.6642	0.0016	0.0015	0.0027	24.5977	24.0057	22.1368	0.0092	0.0560
0.05	0.4701	0.4193	1.8719	535.6642	0.0075	0.0077	0.1035	21.3725	20.4040	22.4932	0.0081	3.8750
0.1	0.4900	0.4191	1.8719	535.6642	0.0155	0.0154	0.2129	19.9669	19.4668	20.0677	0.0094	2.1319
0.5	0.8231	0.5110	1.8719	535.6642	0.0815	0.0816	0.0706	4.3684	4.5351	9.1670	3.08e-08	0.0278
0.9	0.2676	0.3029	1.8719	535.6642	0.1534	0.1534	0.1337	1.0515	0.3670	3.6170	1.15e-11	0.0782
1	0.7500	0.5036	1.8719	535.6642	0.9991	0.5698	0.5432	3.6651	3.0713	3.0708	0.2078	0.3532
						America area						
0.01	0.9000	0.0224	1.1726	17.2937	0.0033	0.0036	0.0037	10.4673	10.4673	10.4908	8.59e-08	2.13e-18
0.05	0.9000	0.1859	1.1726	17.2937	0.0166	0.0182	0.0185	9.3710	9.3174	9.8079	3.57e-06	0.0029
0.1	0.5255	0.2401	1.1726	17.2937	0.0332	0.0362	0.0364	8.1578	8.1482	9.0539	1.01e-05	9.21e-05
0.5	0.5290	0.4931	1.1726	17.2937	0.1703	0.1704	0.1682	2.6477	2.8142	9.9593	7.36e-10	0.0277
0.9	0.9116	0.0824	1.1726	17.2937	0.3160	0.3160	0.2799	0.8255	0.8261	2.2623	8.27e-12	3.33e-07
1	0.6641	0.1854	1.1726	17.2937	0.9953	0.4203	0.4205	2.9232	2.0131	2.0042	0.3304	0.8464
						Asia area						
0.01	0.6464	0.3758	0.7907	5.3641	0.0000	0.0000	0.0006	9.3052	9.4544	9.5485	4.36e-10	0.0223
0.05	0.2968	0.3760	0.7907	5.3641	0.0300	0.0300	0.0300	8.3408	8.4081	8.8504	1.53e-10	0.0045
0.1	0.3824	0.4680	0.7907	5.3641	0.0390	0.0390	0.0406	7.2691	7.2782	8.0833	2.84e-09	0.0065
0.5	0.3809	0.2296	0.7907	5.3641	0.2950	0.3200	0.3242	2.3305	2.4987	5.1461	8.78e-04	0.0283
0.9	0.3839	0.2992	0.7907	5.3641	0.5226	0.5206	0.5396	0.6472	0.8069	1.6730	0.0014	0.0175
1	0.0009	0.3833	0.7907	5.3641	0.9931	0.9837	0.3153	2.8545	2.8535	1.4622	8.95e-05	1.04e-06
						Africa area						
0.01	0.7214	0.6738	1.0477	22.9593	0.0039	0.0034	0.0045	18.8885	18.8885	19.0143	2.91e-07	7.28e-18
0.05	0.6010	0.4701	1.0477	22.9593	0.0246	0.0242	0.0346	16.4931	17.2964	14.6103	0.0131	3.4100
0.1	0.5973	0.5712	1.0477	22.9593	0.0395	0.0477	0.1114	13.9155	12.4864	13.4340	0.0125	0.2423
0.5	0.5802	0.3833	1.0477	22.9593	0.3082	0.3095	0.3267	3.3051	7.8615	2.0464	9.64e-04	0.0047
0.9	0.5804	0.9853	1.0477	22.9593	0.3692	0.3395	0.4800	0.8351	0.8641	0.9505	0.0195	0.0123
1	0.7422	0.5658	1.0477	22.9593	0.9985	0.5211	0.3809	3.5419	2.1503	2.4167	0.0123	8.80e-04
						Eurasian area						
0.01	0.5468	0.9741	0.8894	10.6895	0.0050	0.0049	0.0047	12.7454	12.7584	13.5694	1.01e-07	2.59e-10
0.05	0.8526	0.3224	0.8894	10.6895	0.0250	0.0252	0.0242	12.9685	12.9685	13.6272	1.92e-06	0.0195
0.1	0.8565	0.9523	0.8894	10.6895	0.0499	0.0496	0.0492	11.0575	12.2792	12.7792	7.14e-07	1.89e-09
0.5	0.8705	0.0528	0.8894	10.6895	0.2254	0.2254	0.2615	2.9975	3.2815	3.2485	0.0025	0.0032
0.9	0.5163	0.5703	0.8894	10.6895	0.4554	0.4598	0.5058	2.3421	3.2487	2.3127	0.0028	0.0083
1	0.0001	0.8003	0.8894	10.6895	0.9995	0.9998	0.5044	3.1328	3.1328	3.1045	0.0080	0.0093
						World						
0.01	0.5662	0.4996	0.7660	5.0871	0.0063	0.0062	0.0068	9.5207	9.6375	9.7281	2.35e-07	0.0137
0.05	0.8813	0.5002	0.7660	5.0871	0.0314	0.0314	0.0304	8.5243	8.5641	9.0722	7.23e-09	0.0016
0.1	0.5095	0.4999	0.7660	5.0871	0.0627	0.0627	0.0712	8.0066	8.2228	8.6013	1.36e-12	0.0086
0.5	0.5095	0.5512	0.7660	5.0871	0.3072	0.3072	0.3426	2.3449	2.5392	2.6174	7.58e-04	0.0379
0.9	0.4223	0.5093	0.7660	5.0871	0.5411	0.5411	0.6412	0.6402	0.8083	0.6392	0.0040	0.0196
1	0.0029	0.5095	0.7660	5.0871	0.9933	0.9841	0.1288	2.8771	2.8771	1.4180	8.49e-05	1.02e-10

4.4. Discussion

First of all, we note that kernel estimators provide very good adjustments for large samples, i.e. from size n = 70 and these provide better adjustments compared to those that are classical. Apparently these estimators are less sensitive (evolve linearly) for values less than 1 of the aversion parameters compared to those that are classical. In addition, we note that, the more the parameter increases, the more the indices increase sharply and their welfare functions decrease slightly. For θ = 1 all the estimators, even those qualified as classical, are very sensitive and their values sometimes differ in terms of the direction of variations (some increase when the others decrease for the Atkinson indices). Generally speaking, for Theta close to 0, the values of the estimators are very small and in the neighborhood of 1 or at 1, they are very large. This character shows the major impact of the aversion parameter for the study of inequalities in a population and it poses a great debate because a bad choice of the aversion parameter can lead us to make a hasty conclusion or one that may be far from reality. Thus for a global view, we calculated the average of all the values of the Atkinson index, which is provided by the following Figure 3:
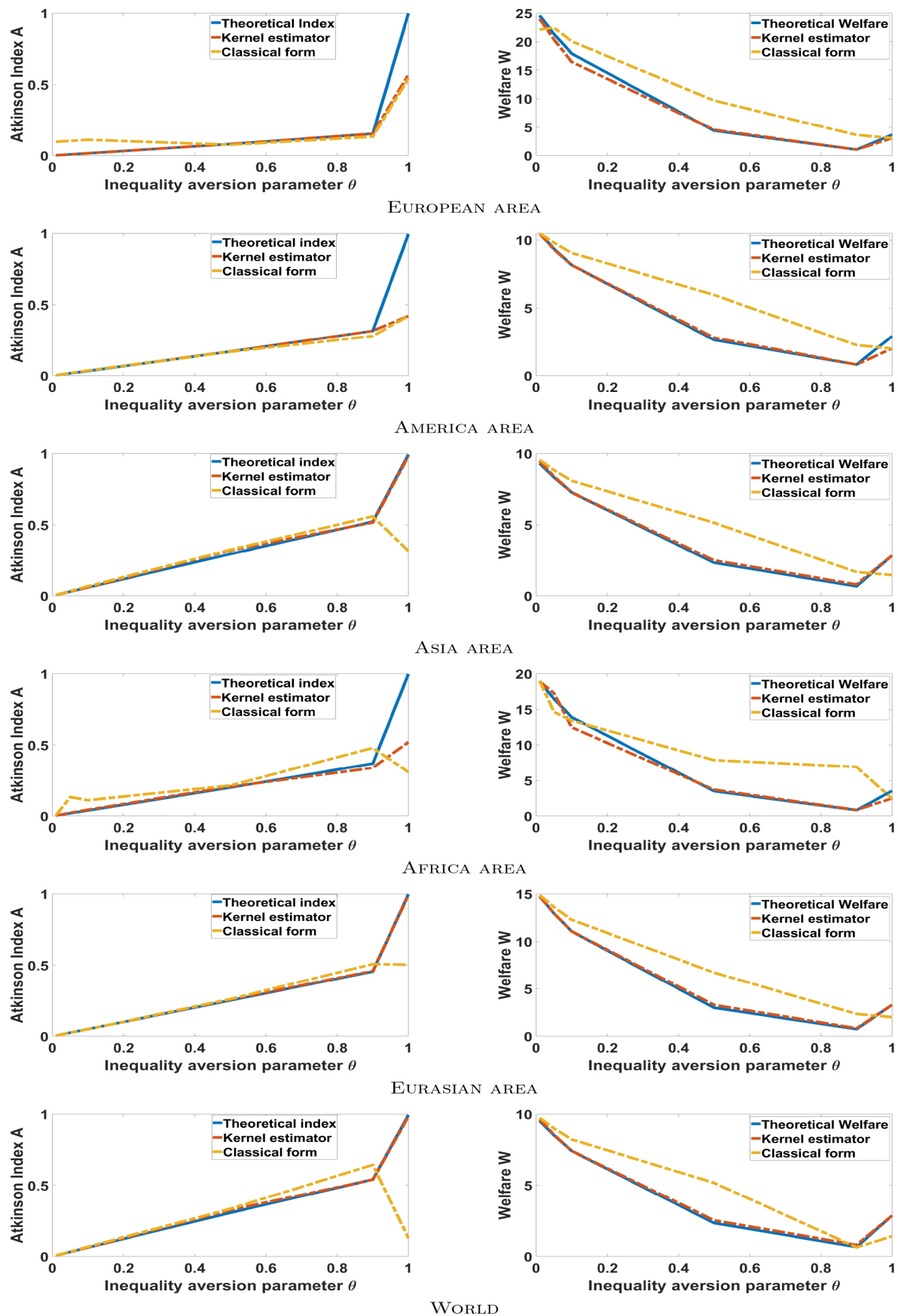
EUROPEAN AREA

ASIA AREA

AMERICA AREA

AFRICA AREA

EURASIAN AREA

WORLD

Figure 2. Comparison of theoretical and estimates curves of the Atkinson index and the welfare
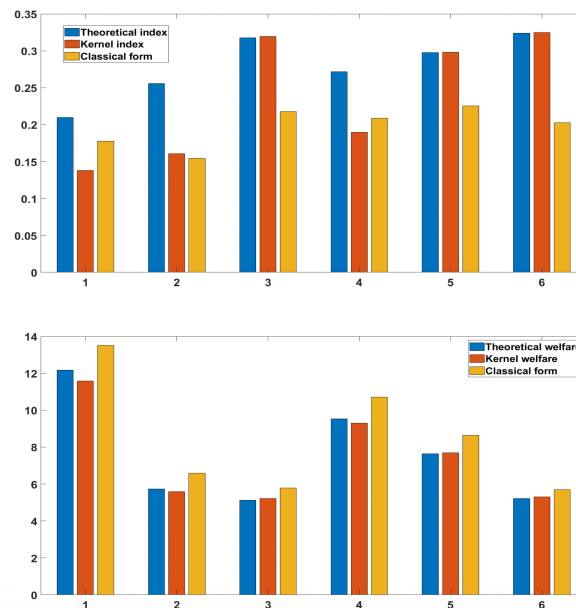
Figure 3. Diagram of the average Atkinson indices and average welfare values

This last figure shows that the Atkinson index whatever the type of estimator does not exceed 0.35 in each zone and therefore on each continent. Although these averages do not allow us to conclude effectively or to draw an efficient objective conclusion, nevertheless they allow us to have more or less or approximately an overall view of the trend of the Atkinson indices and the values of the associated welfare in each region or zone. On the one hand, we therefore conclude that Figure 2 and Table 3 have provided us with the behavior of our estimators in relation to the progressive choice of the aversion parameter θ. On the other hand, to estimate the Atkinson indices and the values of social welfare, it is therefore essential to find an estimator of the parameter θ or therefore find the corresponding value of θ for each data or observation. Thus we therefore propose here to find the value of θ for each of our observations. To do this, the function Φ (32) therefore becomes a bivariate function whose variables are the bandwidth and aversion parameters. The optimization problem (33) is therefore a bivariate optimization problem on the unit square surface and becomes:

$$\left(\widehat{h}_n, \widehat{\theta}_n\right) = \arg \min_{h_n, \theta \in [0,1]} \Phi(h_n, \theta). \tag{34}$$

Let $\widehat{\theta}_{1n}$ and $\widehat{\theta}_{2n}$ be the respective aversion parameters of the Atkinson index and welfare. The only parameters that do not change compared to the previous results are the parameters $\widehat{\alpha}$ and $\widehat{\beta}$ from the M.L method. The results obtained are grouped in Table 4, giving the values of the Atkinson indices and those of welfare after estimating the aversion parameters. Since the classical Atkinson indices and associated welfare function do not provide a very good fit, we no longer present it in this table.

Table 4. Atkinson indices and welfare values based on the estimated aversion parameter

Area or Zone	$\hat{h}_{1n}$	$\hat{\theta}_{1n}$	$RMSE_1$	$\hat{A}_n(\theta)$	$A(\theta)$	$\hat{h}_{2n}$	$\hat{\theta}_{2n}$	$RMSE_2$	$\hat{w}_n(\theta)$	$W(\theta)$
Europe	0.4495	1.0000	0.0127	0.2849	0.2981	0.9604	0.0851	0.0759	2.8102	2.8233
America	0.6547	0.5329	1.79e-10	0.1820	0.1820	0.8718	0.1152	0.0002	7.8195	7.8195
Asia	0.5710	0.7049	7.42e-08	0.4125	0.4125	0.0458	0.1140	2.13e-09	0.9093	0.9093
Africa	0.3894	0.6457	0.0018	0.2606	0.2612	0.5649	0.4643	0.0033	3.9650	3.9653
Eurasia	0.5571	0.3962	2.33e-11	0.2034	0.2034	0.4721	0.3316	9.07e-07	5.2194	5.2194
World	0.5306	0.4209	1.05e-07	0.2604	0.2604	0.7792	0.0939	6.56e-10	7.5482	7.5482

Table 4 allows us to draw a relatively effective conclusion as it reflects the true trends of the Atkinson index. However, we observe that in each region, the aversion parameter averages around 0.5 for the Atkinson index, except for the European region. This could be explained by the small sample size, which is less than 50. Conversely, for larger samples, the aversion parameter is lower. Regarding welfare, the aversion parameter

averages 0.3. This diversity in the values of the aversion parameter once again highlights the importance of its estimation, as it has a significant influence on the estimators. Thus, compared to previous results (where θ was taken arbitrarily), we see a slight significant difference, as the indices in this case are around 0.35 or even exceed it. However, it is noteworthy that inequality is more pronounced in the Americas than in other continents. This may be due to the per capita income in North America, which is more than ten times higher than that of countries in the South or the Caribbean. In Africa and Europe, this inequality is almost identical to that of the world. Inequality is less reduced in Asia, which could be explained by the nearly identical per capita income of several countries in the Gulf and the Pacific. Overall, the disparity is significant across all continents, and world leaders must make greater efforts to raise the level of net national income per capita so that the inequality index in each continent or even in each country is at least higher than 0.7, by making appropriate decisions for the development of each country.

5. Conclusion

In this paper, we have conducted a more in-depth study of the Atkinson index and its welfare function than we initially anticipated. First, we focused on the construction of the two estimators and then examined their almost pure asymptotic convergence as well as their asymptotic normality. Next, we explored a simulation study that confirms the robustness of the two estimators for large sample sizes by working with real data that is highly heterogeneous. These studies revealed that the aversion parameter has a significant influence on the estimation of the Atkinson index and cannot be chosen randomly or arbitrarily for an effective analysis. Therefore, we proposed a numerical estimation of this parameter using the cross-validation method to address the issue of selecting the aversion parameter. Finally, our study has shown that global disparity is a pressing issue that cannot be resolved in the short term; it is a collective problem that requires long-term solutions. As a consequence or open problem arising from this study, it is important to find a way to analytically estimate the aversion parameter using existing methods such as maximum likelihood, method of moments, or Bayesian methods (as discussed in Agbokou and Mensah [18]), or to explore new methods (as in Dabana et al. [19,20]) that could better assist us. This will be the focus of our future work.

References

[1] Piketty, T. (2001). *High Incomes in France in the 20th Century*. Grasset, Paris.

[2] Harper, G., & Price, R. (2011). A frame work for understanding the sociali mpacts of policy and their effects on wellbeing. A paper for the Social Impacts Task force. Defra Evidence and Analysis Series Paper 3. Defra, London. *Food and Rural Affairs.*

[3] Sen, A. (1973). *On Economic Inequality*. New York, Norton.

[4] Agbokou, K., & Mensah, Y. (2023). Varying bandwidth parameter method on Kernel Gini index estimation. *Annals of the University of Craiova-Mathematics and Computer Science Series, 50*(1), 29-41.

[5] Banerjee, A. K. (2010). A multidimensional Gini index. *Mathematical Social Sciences, 60*(2), 87-93.

[6] Sen A. (1997). *On Economic Inequality*. Oxford University Press, 2nd edition. Oxford, UK.

[7] Rawls, J. (1972). *A Theory of Justice*. Oxford University Press, London.

[8] Guerrero V., M. (1987). A note on the estimation of Atkinson's index of inequality. *Economics Letters 25*(1987) 379-384, North-Holland

[9] Biewen, M., & Jenkins, S. P. (2003). *Estimation of generalized entropy and Atkinson inequality indices from complex survey data (No. 763)*. IZA Discussion Papers.

[10] Tchamyou, V. S. (2020). Education, lifelong learning, inequality and financial access: Evidence from African countries. *Contemporary Social Science, 15*(1), 7-25.

[11] Bellù, L. G., & Liberati, P. (2006). *Policy Impacts on Inequality: Welfare Based Measures of Inequality-the Atkinson Index', Food and Agriculture Organization of the United Nations.*

[12] Agbokou, K. (2023). On the Nature of Elasticity Function: An Investigation and a Kernel Estimation. *Journal of Applied Mathematics, 2023*(1), 1346602.

[13] Serfling R. J. (1998). *Approximation Theorems of Mathematical Statistics*. The Johns Hopkins University, John Wiley & Sons.

[14] Lehmann E. L. (1998). *Elements of Large-Sample Theory*. Springer Texts in Statistics.

[15] Epanechnikov, V. A. (1969). Non-parametric estimation of a multivariate probability density. *Theory of Probability & Its Applications, 14*(1), 153-158.

[16] Agbokou, K., & Gneyou, K., (2017). On the strong convergence of the hazard rate and its maximum risk point estimators in presence of censorship and functional explanatory covariate. *Afrika Statistika, 12*(3), 1397-1416.

[17] Agbokou, K., Gneyou, K. E., & Deme, E. (2018). Almost sure representations of the conditional hazard function and its maximum estimation under right-censoring and left-truncation. *Far East Journal of Theoretical Statistics, 54*(2), 141-173.

[18] Agbokou, K., & Mensah, Y. (2022). Inference on the reproducing kernel Hilbert spaces. *Universal Journal of Mathematics and Mathematical Sciences, 15*, 11-29.

[19] Dabana, H., Agbokou, K., & Gneyou, K. E. (2024). Kernel estimation of the conditional probability density function and Mode in presence of functional covariate and censorship. *Afrika Statistika, 19*(1), 3771-3796.

[20] Dabana, H., Agbokou, K., & Gneyou, K. (2025). Local linear estimation of conditional probability density and mode under right censoring and left truncation: dependent data case. *Gulf Journal of Mathematics, 20*, 338-359.