THE ASYMPTOTIC THEORY OF THE POVERTY INTENSITY IN VIEW OF EXTREME VALUE THEORY FOR TWO SIMPLE CASES.

GANE SAMB LO AND SERIGNE TOUBA SALL

Abstract. Let Y_1, Y_2 .... be independent observations of the income variable of some given population, with underlying distribution G. Given a poverty line Z, then for each n$\geq$ 1, q =q_n is the numbers of poor in some given the population. The general form of poverty measures used by economists to monitor the welfare evolution of this population is

$$P_n = \frac{1}{a(q)b(n)}\sum_{j=1}^{q} c(n,q,j)\ d\left(\frac{Z-Y_{j,n}}{Z}\right).$$

This class includes the most popular poverty measures like the Sen, Shorrocks and Greer-Foster-Thorbecke statistics. We give a complete asymptotic normality theory in the frame of extreme value theory. In this paper, the two simple cases of the Pareto and exponential distributions are studied with the intensity poverty. Simulations and applications to the senegalese data are driven.

Résumé

Soit Y_1, Y_2 .... des observations indépendantes de la variable revenu d'une population donnée. Etant donné un seuil de pauvreté Z, alors q_n est pour tout n$\geq$ 1 le nombre de pauvres dans la population. La forme générale des indicateurs de pauvreté est

$$P_n = \frac{1}{a(q)b(n)}\sum_{j=1}^{q} c(n,q,j)\ d\left(\frac{Z-Y_{j,n}}{Z}\right).$$

Cette classe de mesures contient les indicateurs classiques tels que ceux de Sen, de Shorrocks et de Greer-Foster-Thorbecke. Dans cet article, nous préparons la théorie asymtotique complète de ces indicateurs dans le cadre des valeurs extrêmes. Nous nous concentrons ici sur l'intensité de la pauvreté pour deux cas simples : La loi de Paréto et la loi exponentielle. Des simulations et des applications sur les données du Sénégal sont présentées.

1. INTRODUCTION.

Poverty is measured in Economics by statistics of the general form

$$P_N = \frac{1}{a(Q)b(N)}\sum_{j=1}^{Q} c(N,Q,j)\ d\left(\frac{Z-Y_{j,N}}{Z}\right) \tag{1.1}$$

where a, b, c, d are given functions, Q is the number of poor in the studied population P of size N, Z the poverty line and $\mathrm{Y}_{1,N} \leq Y_{2,N} \leq ... \leq Y_{N,N}$ are

2000 *Mathematics Subject Classification.* 2E20,62E17, 62G30, 91B82, Secondary 60G15
JEL Classification: I32.
Key words and phrases. Poverty measurement and analysis, asymptotic normality, empirical processes, Gaussian approximations, estimation and simulations, extreme value theory.

Received on January 20,2006. Revised version accepted on May 10, 2007

the ordered incomes of the individuals of P. The poverty line Z is defined by economic specialists or governmental authorities so that any individual or household with income (say yearly) less than Z is considered as poor. The economic welfare of the population $\mathcal{P}$ is monitored both by the measures (1.1) and other measures of the inequality of the income distribution. A

very large class of poverty measures can be found in the literature. One may divide them into two classes. The first includes the non weighted ones, for which c(N, Q, j)$\equiv$ 1. The most popular of them is the Foster-Greer-Thorbecke class which is, for $\alpha \geq 0$,

$$FGT(\alpha) = \frac{1}{N}\sum_{j=1}^{Q}\left(\frac{Z - Y_{j,N}}{Z}\right)^{\alpha}. \tag{1.2}$$

For $\alpha = 0$, (1.2) reduces to Q/N, the headcount of the poor while for $\alpha = 1$ and $\alpha = 2$, it is respectively interpreted as the severity of poverty and the depth in poverty. The second class consists of the weighted measures. We mention here two of its famous members, the Sen(1976) measure

$$P_{SE,N} = \frac{2}{N(Q+1)}\sum_{j=1}^{Q}(Q - j + 1)\left(\frac{Z - Y_{j,N}}{Z}\right) \tag{1.3}$$

and the Shorroks(1995) one

$$P_{SH,N} = \frac{1}{N^2}\sum_{j=1}^{Q}(2N - 2j + 1)\left(\frac{Z - Y_{j,N}}{Z}\right). \tag{1.4}$$

These discrete statistics have been widely prospected in the frame of poverty reduction. The interested readers are referred to [9], [10], [14], [18].

We now want to settle an asymptotic theory for the statistics (1.1). We suppose that we have independent and identically observations of the income Y_1, Y_2 .... with underlying distribution G with lower endpoint $y_0 = \inf\{x, G(x) > 0\} \geq 0$. Given a poverty line Z, we have the (random) number of poor $q_n = q$ in the sample of size n : $Y_1, Y_2, ..., Y_n$. The sample measure of poverty becomes

$$P_n = \frac{1}{a(q)b(n)}\sum_{j=1}^{q}c(n, q, j)\ d\left(\frac{Z - Y_{j,n}}{Z}\right). \tag{1.5}$$

Clearly, this depends on the q lower extreme values and suggests to use the extreme value theory to handle such statistics. If q/n$\rightarrow$ 0, the study will only need the lower tail of G. But for many poor countries q/n is greater than 50% so that the reasonable hypothesis must be q/n$\rightarrow \xi = G(Z) \in]0, 1[$. Thus, at most the lower tail and the center of G will be used. This is in conformity with the focalization hypothesis of poverty measure which says that any poverty statistics is strictly function of the poor income. In général, the law of such statistics are usually guided by the extremal domain of attraction of the sequence of minima $Y_{1,n}$, n$\geq$ 1. However, observe that much of the theory of extreme value is set up by using the extremal domain of attraction of the maxima. We make a brief recall of our needs in extreme value theory using maxima in Section 2 . In Section 3, we state our results for the intensity poverty whose the asymptotic theory is completely described for the Pareto and the exponential cases.

2. Basics of extreme values theory using Maxima

Let $Y_1, Y_2, ...$ be a sequence of independent and indentically distributed random variable with $P(Y_1 \leq x) = G(x)$ and G(0) = 0. The estimation of the law of the maximum $Y_{n,n} = \max(Y_1, ..., Y_n)$ is one of the most important questions of the extreme value theory. Recall that $Y_{n,n}$ converges in type to some random variable Z with non degenerated distribution function $H(.)$ if and only if there exists two sequences of real numbers $(a_n > 0, \;\; n \geq 1)$ and $(b_n, \; n \geq 1)$ such that

$$\lim_{n\to\infty} \mathbb{P}(Y_{n,n} \leq a_n x + b_n) = \lim_{n\to\infty} G^n(a_n x + b_n) = H(x), \quad \forall x \in \mathbb{R}. \tag{2.1}$$

It is then said that G is in the domain of attraction of H, which is denoted by G$\in D(H)$. It is well-known that H is necessarily one of these three types :

$$H(x) = \varphi_\alpha(x) = \exp(-x^{-\alpha}) \; \mathbb{I}_{(x\geq 0)} \; ; \quad (type\ I) \tag{2.2}$$

for $\alpha > 0$,

$$H(x) = \Lambda(x) = \exp\left(-e^{-x}\right) \; ; \qquad (type\ II) \tag{2.3}$$

$$H(x) = \psi_\gamma(x) = \exp(-(-x)^\gamma) \; \mathbb{I}_{(x\leq 0)} + \mathbb{I}_{(x>0)} \; ; \; \gamma > 0 \;\; ; \;\; (type\ III) \tag{2.4}$$

where I_A stands for the indicator function of the set A. It should be noticed that the function H has to be seen as an equivalence :

$$H_1 \mathcal{R} \; H_2 \; \Leftrightarrow (\exists \; (a,b) \in \mathbb{R}_+^* \times \mathbb{R}, \forall \; x \in \mathbb{R}, \; H_1(x) = H_2(ax+b)) \tag{2.5}$$

In this case, any of the function H in (2.2), (2.3) and (2.4) may be represented as

$$H(x) = H_\gamma(x) = \exp(-(1+\gamma x)^{-1/\gamma}), \text{ for } 1{+}\gamma x > 0,$$

where $(1+\gamma x)^{-1/\gamma}$ is interpreted as e^{-x} for $\gamma = 0$. The values $\gamma > 0$, $\gamma = 0$ and $\gamma < 0$ are respectively related to the types I, II et III.

The class of the whole domain of attraction (the set of all distribution functions attracted to some H) is denoted by $\mathfrak{D} = \cup_{\alpha>0} \cup_{\gamma>0} D(\varphi_\alpha) \cup D(\psi_\gamma) \cup D(\Lambda)$.

The interested reader is referred to [5], [7] or [15] for a general introduction to extreme value theory. The characterization of these three domains received very great attention in the two past decades. The frame of extreme value theory has also been widely used to find the asymptotic normality of sums of extreme values as well as Hill type statistics (see [12], mainly and [13], [16], [6], [11]).

The following representations are then used for the generalized inverse of G defined by $G^{\leftarrow}(s) = \inf\{x, G(x) \geq s\}$, $0 \leq s \leq 1$. First, w have for G$\in D(\varphi_\gamma)$, $z_0 = \sup\{x, G(x) < 1\} = +\infty$ and

$$G^{\leftarrow}(1-u) = c(1+f(u))\ u^{1/\gamma} \exp(\int_u^1 b(t)t^{-1}dt), \quad 0 < u < 1, \tag{2.6}$$

Next, for G$\in D(\psi_\gamma)$, $z_0 = \sup\{x, G(x) < 1\} < +\infty$ and

$$z_0 - G^{\leftarrow}(1-u) = c(1+f(u))\ u^{1/\gamma} \exp(\int_u^1 b(t)t^{-1}dt), 0 < u < 1 \tag{2.7}$$

and, finally, for G$\in D(\Lambda)$, $z_0 = \sup\{x, G(x) < 1\} \leq +\infty$.

$$G^{\leftarrow}(1-u) = d - s(u) + \int_u^1 s(t)t^{-1}dt), 0 < u < 1, \tag{2.8}$$

where $s(u) = c(1+f(u)) \exp(\int_u^1 b(t)t^{-1}dt)$, $0 < u < 1$. In each of these formulae, f and b are functions such taht (f(u), b(u))→ (0,0) when u→ 0, while c> 0 and d are constants. (2.6) and (2.7) are the Karamata representations while (2.8) is the one of de Haan.

To finish, we recall the characterization on G for the different domains. First, by Theorem 2.4.3 in [7], G$\in D(\Lambda)$ iff

$$\frac{1 - G(t + xr(t))}{1 - G(t)} \to e^{-x}, \text{ as x} \to z_0 \tag{2.9}$$

for any x and for some positive function r(·). Moreover, all the functions r(·) in (2.9) are equivalent between them as x→ z_0 Further, by Theorem 2.4.1 in [6], G $\in D(\Lambda)$ iff

$$\forall x > 0, \{G^{\leftarrow}(1-ux) - G^{\leftarrow}(1-u)\}/r(u) \quad \to \ \text{-}\log x, \tag{2.10}$$

for some positive function r(·). Notice that z_0 may be finite or not. Next, G$\in D(\varphi_\gamma)$ iff $z_0 = +\infty$ and

$$\forall \lambda > 0,\ (1 - G(\lambda x))/(1 - G(x)) \to \lambda^{-\gamma}. \tag{2.11}$$

In this case, G admits the Karamata representation

$$1 - G(x) = c(1 + f_1(x))\ x^{-\gamma} \exp(\int_1^x b_1(t)t^{-1}dt), \quad y_0 < x < z_0, \tag{2.12}$$

where $(f_1(u), b_1(u)) \to (0,0)$ as u→ 0. The function

$$L(x) = c(1 + f_1(u)) \exp(\int_1^x b_1(t)t^{-1}dt)$$

is the representation of a function slowly varying at infinity that is

$$L(\lambda x)/L(x) \to 1 \ as \ x \uparrow \infty$$

for any $\lambda > 1$. Finally G$\in D(\psi_\gamma)$ iff $z_0 < \infty$ and $G(z_0 - 1/\cdot) \in D(\varphi_\gamma)$.

Now, we must remark in our frame that Y is an income variable. Then its lower endpoint y_0 is not negative. This allows us to study (1.5) via the transform $X=1/(Y-y_0)$. Throughout this paper, the distribution function of

X denoted by F, whose upper endpoint is then $+\infty$. Its extremal domain is D(Λ) or D(φ_γ). Hence (1.5) is transformed as

$$P_n = \frac{1}{a(q)b(n)}\sum_{j=1}^{q} c(n,q,j)\ d\left(\frac{Z - y_0 - X^{-1}_{n-j+1,n}}{Z}\right), \tag{2.13}$$

and studied following the extremal law of X using the general representations (2.6) and (2.8).

The aim of this paper is to prepare the general asymptotic theory of the poverty measures by giving the complete study of one represent of each of the possible extremal domain : The Pareto distribution of parameter γ (for the type I) and the exponential distribution (for the type II). Only the non weighted measures are also considered.

Throughout the paper, we use the representations of the studied random variables X_i, $i \geq 1$, by $F^{\longleftarrow}(1-U_i)$, $i \geq 1$,where $U_1, U_2, ...$ is a sequence of independent random variables uniformly distributed on $(0, 1)$. It follows that :

$$\{\ X_{n-i+1,n}, 1 \leq i \leq n\ ;\ n \geq 1\} \overset{d}{=} \{F^{\leftarrow}(1 - U_{i,n}), 1 \leq i \leq n\ ;\ n \geq 1\} \tag{2.14}$$

where $U_{1,n} < ... < U_{n,n}$ are the order statistics based on $U_1, ..., U_n$.

Our best achievement is the complete asymptotic characterization of the poverty intensity measure

$$I_n = \frac{1}{n}\sum_{j=1}^{q}\left(\frac{Z - Y_{j,n}}{Z}\right) \tag{2.15}$$

when the exponential distribution of parameter $\lambda > 0$,

$$F(x) = \begin{cases} 1 - exp(-\lambda x), & \lambda > 0, \text{ for x} > 0 \\ 0 & \text{else} \end{cases} \tag{2.16}$$

and when the Paréto distribution of parameter $\gamma > 0$, i. e.,

$$F(x) = \begin{cases} 1 - x^{-1/\gamma}, & \text{for x} > 1 \\ 0 & \text{else} \end{cases} \tag{2.17}$$

are retained for $F(\cdot) = 1 - G(y_0 + 1/\cdot)$. Now, our results are given.

3. RESULTS

Theorem 1. *Let F= 1-G($y_0 + 1/\cdot$) be the exponential distribution function of parameter $\lambda > 0$, then $\sqrt{n}(I_n - C_n) \to N(0, \sigma^2)$ where*

$$C_n = \frac{1}{nZ}q(Z - y_0) + \frac{\lambda}{Z}\int_{\frac{1}{n}}^{q/n} \frac{1}{\log(s)} ds$$

and

$$\left(\frac{Z}{\lambda}\right)^2 \sigma^2 = \frac{-2}{\log G(Z)}\int_0^{G(Z)} (\log s)^{-2} ds + 2\int_0^{G(Z)} (\log s)^{-3} ds - (\int_0^{G(Z)} \frac{ds}{(\log s)^2})^2$$

with q/n=$G_n(Z)$ and $G(Z) = \exp(-\lambda/(Z - y_0))$ for $Z > y_0$.

Theorem 2. *Let F=1-G($y_0+1/\cdot$) be the Pareto distribution function of parameter $\gamma^{-1}>0$, then $\sqrt{n}(I_n-D_n)\rightarrow N(0,\theta^2)$ with*

$$D_n=\frac{1}{nZ}q(Z-y_0)-\frac{1}{Z(\gamma+1)}((q/n)^{\gamma+1}-n^{-\gamma-1})$$

and

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

with $G(Z)=((Z-y_0)^{\frac{1}{\gamma}}\wedge 1)$ for Z> y_0.

4. PROOFS

In the beginning we use a special representation of

$$I_n=\frac{1}{n}\sum_{j=1}^{q}\left(\frac{Z-Y_{j,n}}{Z}\right) \tag{4.1}$$

By using the transform X=1/(Y-y_0), we get

$$I_n=\frac{1}{nZ}\sum_{j=1}^{q}(Z-y_0-\frac{1}{X_{n-j+1,n}}) \tag{4.2}$$

$$I_n=\frac{1}{nZ}q(Z-y_0)-\frac{1}{nZ}\sum_{j=1}^{q}\frac{1}{X_{n-j+1,n}} \tag{4.3}$$

$$=\frac{1}{nZ}q(Z-y_0)-\frac{1}{nZ}\sum_{j=1}^{q}n\int_{\frac{n-j}{n}}^{\frac{n-j+1}{n}}\frac{1}{Q_n(s)}ds, \tag{4.4}$$

where Q_n is the empirical quantile function based on $X_1,X_2,...,X_n$ defined by $Q_n(s)=X_{j,n}$ for $\frac{j-1}{n}<s\leqslant\frac{j}{n}$ for j=1,...,n and $Q_n(0)=X_{1,n}$. Furthermore, we get by (2.14) the following representation of Q_n via the empirical quantile function V_n based on $U_1,U_2,...,U_n$:

$$\{Q_n(t),0\leq t\leq 1\}=^d\{F^{\leftarrow}(V_n(t)),0\leqslant t\leqslant 1\} \tag{4.5}$$

$$=^d\{F^{\leftarrow}(1-V_n(t)),\ 0\leqslant t\leqslant 1\}, \tag{4.6}$$

where $F^{\leftarrow}$(s)=inf{x, F(x)$\geq s$}, 0$\leq s\leq 1$, is the quantile function of F. Recall that $G_n(Z)=q/n$ when G_n denotes the empirical distribution function based on $Y_1,Y_2,...,Y_n$. Thus, the final representation of I_n is

$$I_n=\frac{1}{nZ}q(Z-y_0)-\frac{1}{nZF^{\leftarrow}(1-U_{1,n})}-\frac{1}{Z}\int_{\frac{1}{n}}^{G_n(Z)}\frac{1}{F^{\leftarrow}(1-V_n(s))}ds. \tag{4.7}$$

Now we are ready to prove the two theorems.

Proof of theorem 1.
We have $F^{\leftarrow}(s) = -\log(1-s)/\lambda, 0 \leqslant s < 1$ and from (4.7),

$$I_n = \frac{1}{nZ}q(Z - x_0) + \frac{\lambda}{nZ\log(U_{1,n})} + \frac{\lambda}{Z}\int_{\frac{1}{n}}^{G_n(Z)} \frac{1}{\log(V_n(s))}ds. \tag{4.8}$$

Put

$$C_n = \frac{1}{nZ}q(Z - x_0) + \frac{\lambda}{Z}\int_{\frac{1}{n}}^{G_n(Z)} \frac{1}{\log(s)}ds \tag{4.9}$$

By the mean value theorem

$$I_n - C_n = \frac{\lambda}{nZ\log(U_{1,n})} - \frac{\lambda}{Z}\int_{\frac{1}{n}}^{G_n(Z)} \frac{V_n(s) - s}{\zeta_n(s)(\log(\zeta_n(s)))^2}ds, \tag{4.10}$$

where $\zeta_n(s) \in [\min(s, V_n(s)); \max(s, V_n(s))]$. Thus

$$\sqrt{n}\,(I_n - C_n) = \frac{\lambda}{\sqrt{n}Z\log(U_{1,n})} - \frac{\lambda}{Z}\int_{\frac{1}{n}}^{G_n(Z)} \frac{\sqrt{n}(V_n(s) - s)}{\zeta_n(s)(\log(\zeta_n(s)))^2}ds. \tag{4.11}$$

At this step, we need one version of the so-called hungarian construction. Precisely, Csörgö and al. (see [2]) have constructed a probability space holding a sequence of independent uniformly distributed random variables U_1, U_2, ... and a sequence of brownian bridges $B_1, B_2, \ldots$ such that for each $0 < \nu < 1/2$, as $n \to \infty$,

$$\sup_{1/n \leq s \leq 1-1/n} \frac{|\beta_n(s) - B_n(s)|}{(s(1-s))^{1/2-\nu}} = O_p(n^{-\nu}). \tag{4.12}$$

where $\{\beta_n(s) = \sqrt{n}\,(s - V_n(s))\,, 0 \leqslant s \leqslant 1\}$ is the uniform quantile process. From this and (4.11), we get

$$\sqrt{n}\,(I_n - C_n) = \frac{\lambda}{\sqrt{n}Z\log(U_{1,n})} + \frac{\lambda}{Z}\int_{\frac{1}{n}}^{G_n(Z)} \frac{\beta_n(s) - B_n(s)}{\zeta_n(s)(\log(\zeta_n(s)))^2}ds + \frac{\lambda}{Z}\int_{\frac{1}{n}}^{G_n(Z)} \frac{B_n(s)}{\zeta_n(s)(\log(\zeta_n(s)))^2}ds$$

$$= \frac{\lambda}{\sqrt{n}Z\log(U_{1,n})} + \frac{\lambda}{Z}\int_{\frac{1}{n}}^{G_n(Z)} \frac{\beta_n(s) - B_n(s)}{s^{\frac{1}{2}-\nu}} * \frac{s^{\frac{1}{2}-\nu}}{\zeta_n(s)(\log(\zeta_n(s)))^2}ds +$$

$$\frac{\lambda}{Z}\int_{\frac{1}{n}}^{G_n(Z)} B_n(s)\left(\frac{1}{\zeta_n(s)(\log(\zeta_n(s)))^2} - \frac{1}{s(\log s)^2}\right)ds + \frac{\lambda}{Z}\int_{G(Z)}^{G_n(Z)} B_n(s)\left(\frac{1}{s(\log s)^2}\right)ds$$

$$+\frac{\lambda}{Z}\int_{1/n}^{G(Z)} B_n(s)\left(\frac{1}{s(\log s)^2}\right)ds =: R_{0,n} + R_{1,n} + R_{2,n} + R_{3,n} + N_n. \tag{4.13}$$

We first claim that each of the $R_{j,n}$ tends to zero in probability as $n \to \infty$.

Claim 1.1. $R_{0,n} \to_p 0$. We have

$$R_{\mathbf{0},n} = \frac{\lambda}{\sqrt{n} Z \log(U_{1,n})} = \frac{\lambda}{\sqrt{n} Z (\log(nU_{1,n}) - \log n)} \tag{4.14}$$

But it is well-known that $\log(nU_{1,n}) = O_p(1)$ so that $R_{0,n} = o_p(1)$.

Claim 2.1. $R_{1,n} \to_p 0$. We have

$$R_{1,n} = \frac{\lambda}{Z} \int_{\frac{1}{n}}^{G_n(Z)} \frac{\beta_n(s) - B_n(s)}{s^{\frac{1}{2}-\nu}} * \frac{s^{\frac{1}{2}-\nu}}{\zeta_n(s)(\log(\zeta_n(s)))^2} ds \tag{4.15}$$

$$= O_p(n^{-\upsilon}) \int_{\frac{1}{n}}^{G_n(Z)} \frac{s^{-\frac{1}{2}-\nu}}{(\zeta_n(s)/s) \ \log \zeta_n(s))^2} ds =: O_p(n^{-\nu}) \times R'_{1,n}. \tag{4.16}$$

Now, by Lemma 13 in [3],

$$\lim_{\rho \to \infty} \inf \lim_{n \to \infty} \sup \mathbb{P}(\rho^{-1} \leq \inf\{V_n(s)/s, \ 1/\text{n} \leq s \leq 1\} \leq \sup\{V_n(s)/s, \ 1/\text{n} \leq s \leq 1\} \leq \rho) = 1.$$

Recall again that $G_n(Z) \to G(Z) > 0$. Hence, for any $G(Z) > \varepsilon > 0$, for large values of n and for large values of ρ, we have with probability at least greater than 1-$\varepsilon/3$ (denoted w.p.$\varepsilon/3$),

$$(\zeta_n(s)/s)^{-1} \leqslant \rho \tag{4.17}$$

and we have w.p.$\varepsilon/3$ for large n,

$$1/n \leq s \leq G_n(Z) \leq G(Z) + \varepsilon/2 \implies 0 \leq \zeta_n(s) \leq \max(V_n(s), s) \tag{4.18}$$

$$\leq \max(V_n(G(Z) + \varepsilon/2), \ G(Z) + \varepsilon/2) \tag{4.19}$$

and since $V_n(G(Z) + \varepsilon/2) \to_p G(Z) + \varepsilon/2$, we also get w.p.$\varepsilon/3$ for large n,

$$0 \leq V_n(G(Z) + \varepsilon/2) \leq G(Z) + \varepsilon, \tag{4.20}$$

and then w.p.$2\varepsilon/3$, for large n

$$1/n \leq s \leq G_n(Z) \leq G(Z) + \varepsilon \implies 0 \leq \zeta_n(s) \leq G(Z) + \varepsilon. \tag{4.21}$$

This latter yields w.p.$2\varepsilon/3$, for large n

$$1/n \leq s \leq G_n(Z) \leq G(Z) + \varepsilon \implies 0 \leq -1/\log \zeta_n(s) \leq -1/\log(G(Z) + \varepsilon) \tag{4.22}$$

By (4.17), (4.21), we have w.p.ε, for large values of n and ρ,

(4.23)
$$|R'_{1,n}| \leq \int_{\frac{1}{n}}^{G(Z)+\varepsilon} \frac{s^{-\frac{1}{2}-\nu}}{(\zeta_n(s)/s)(\log \zeta_n(s))^2} ds \leq \frac{\rho}{(\log(G(Z)+\varepsilon))^2} (G(Z)+\varepsilon)^{1/2-\nu}.$$

Finally $R'_{1,n} = O_p(1)$ and then by (4.16), $R_{1,n} \to_p 0$.

Claim 3.1. $R_{2,n} \to_p 0$. We have

$$R_{2,n} = \frac{\lambda}{Z} \int_{\frac{1}{n}}^{G_n(Z)} B_n(s) \left(\frac{1}{\zeta_n(s)(\log(\zeta_n(s)))^2} - \frac{1}{s(\log s)^2} \right) ds \tag{4.24}$$

$$= \frac{\lambda}{Z}\int_{\frac{1}{n}}^{G_n(Z)} B_n(s) * \frac{(\zeta_n(s)-s)(-2-\log(\pi_n(s))}{\pi_n(s)^2(\log \pi_n(s))^3} ds. \tag{4.25}$$

where $\pi_n(s) \in [\zeta_n(s)\wedge s, \zeta_n(s)\vee s]$. By (5) and (6) in [17], p.500, we have for any $0 < c < \infty$,

$$\sup_{c/(n+1)<s<1-c/(n+1)} \left| B_n(s)/\sqrt{s(1-s)} \right| = O_p(b_n) \tag{4.26}$$

and

$$\sup_{c/(n+1)<s<1-c/(n+1)} \left| \beta_n(s)/\sqrt{s(1-s)} \right| = O_p(b_n) \tag{4.27}$$

with $b_n = \sqrt{2\log\log n}$. From (4.25) and since $\pi_n(s) \in [V_n(s)\wedge s, V_n(s)\vee s]$

$$|R_{2,n}| \leq= \frac{\lambda}{Z}\int_{\frac{1}{n}}^{G_n(Z)} |V_n(s)-s| \ |B_n(s)| \frac{|2+\log(\pi_n(s)|}{\pi_n(s)^2 \left|\log \pi_n(s)\right|^3} ds.$$

$$|R_{2,n}| \leq \frac{\lambda}{Z\sqrt{n}} \int_{\min(\frac{1}{n},G_n(Z))}^{\max(\frac{1}{n},G_n(Z))} s(1-s) \left|\frac{B_n(s)}{\sqrt{s(1-s)}}\right| \ \left|\frac{\beta_n(s)}{\sqrt{s(1-s)}}\right| \frac{|2+\log(\pi_n(s)|}{\pi_n(s)^2 \times \left|\log(\pi_n(s)\right|^3} \ ds$$

$$= O_p(n^{-1/2}\log\log n) \int_{\min(\frac{1}{n},G_n(Z))}^{\max(\frac{1}{n},G_n(Z))} s \frac{|2+\log(\pi_n(s)|}{\pi_n(s)^2 \times \left|\log(\pi_n(s)\right|^3} \ ds \tag{4.28}$$

$$\leq O_p(n^{-1/2}\log\log n) \int_{\min(\frac{1}{n},G_n(Z))}^{\max(\frac{1}{n},G_n(Z))} s^{-1} \frac{|1+2/\log(\pi_n(s)|}{(\pi_n(s)/s)^2 \times \left|\log(\pi_n(s)\right|^2} \ ds$$

By the same methods that led to (4.17), we have, for any $\varepsilon > 0$ such that $G(Z)+\varepsilon < 1$, w.p.ε for large values of n and $\rho > 0$,

$$|R_{2,n}| \leq O_p(n^{-1/2}\log\log n)\rho^2 \frac{|1+2/\log(G(Z)+\varepsilon)|}{(\log(G(Z)+\varepsilon))^2} \int_{\frac{1}{n}}^{G(Z)+\varepsilon} s^{-1} ds$$

(4.29)

$$\leq O_p(n^{-1/2}\log\log n)\rho^2 \frac{1+|2/\log(G(Z)+\varepsilon)|}{\log(G(Z)+\varepsilon)^2} (|\log(G(Z)+\varepsilon| + \log n) \to_p 0.$$

which finishes the proof of the claim.

Claim 4.1. $R_{3,n} \to_p 0$. Since $G_n(Z) \to_p G(Z)$ as $n \to \infty$,

$$\sup\{s(logs)^2, s \in [\min(G_n(Z), G(Z)), \max(G_n(Z), G(Z))] \to_p G(Z)(\log(G(Z))^2$$

and

$$\inf\{s(logs)^2, s \in [\min(G_n(Z), G(Z)), \max(G_n(Z), G(Z))] \to_p G(Z)(\log(G(Z))^2$$

and then

$$R_{3,n} = \frac{\lambda}{(Z\log Z)^2}(1+o_p(1)) \int_{G(Z)}^{G_n(Z)} B_n(s) ds \tag{4.30}$$

But for any $\varepsilon > 0$, we have w.p.ε, for large values of n

$$\left| \int_{G(Z)}^{G_n(Z)} B_n(s) ds \right| \leq \int_{G(Z)-\varepsilon}^{G(Z)+\varepsilon} |B_n(s)| \, ds =: T_n(\epsilon) \tag{4.31}$$

Now $E(\mathrm{T}_n(\varepsilon)) = E(\int_{G(Z)-\varepsilon}^{G(Z)+\varepsilon} |B_n(s)|\, ds) = \sqrt{2/\pi}\int_{G(Z)-\varepsilon}^{G(Z)+\varepsilon} \sqrt{s(1-s)}\ ds = T(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$. We conclude that for any ε, we get for large values of n, by Markov inequality

$$\forall\eta > 0, \mathbb{P}(\left|\int_{G(Z)}^{G_n(Z)} B_n(s)ds\right| \geq \eta) \leq E(T(\varepsilon))/\eta \tag{4.32}$$

that is

$$\forall\eta > 0, \lim_{n\rightarrow\infty} \sup\ \mathbb{P}(\left|\int_{G(Z)}^{G_n(Z)} B_n(s)ds\right| \geq \eta) \leq \mathbb{E}(T(\varepsilon))/\eta \tag{4.33}$$

Now, as $\varepsilon \downarrow 0$, we conclude that $\int_{G(Z)}^{G_n(Z)} B_n(s)ds \longrightarrow_p 0$ and thus $\mathrm{R}_{3,n} \rightarrow_p 0$ by (4.30) et (4.31).

Finally we achieve our proof by remarking that N_n is a Riemannian integral of a Brownian bridge and hence, is a Gaussian and centered random variable with variance

$$\sigma_n^2 = \left(\frac{\lambda^2}{Z^2}\right)\int_{\frac{1}{n}}^{G(Z)}\int_{\frac{1}{n}}^{G(Z)} \frac{\min(s,t)-st}{s(\log s)^2 t(\log t)^2}dsdt \tag{4.34}$$

$$\frac{Z^2}{\lambda^2}\sigma_n^2 = \int_{\frac{1}{n}}^{G(Z)} \frac{1-s}{s(\log s)^2}\int_{\frac{1}{n}}^{s}\frac{dt}{(\log t)^2}ds + \int_{\frac{1}{n}}^{G(Z)}\frac{1}{(\log s)^2}\int_{s}^{G(Z)}\frac{(1-t)dt}{t(\log t)^2}ds$$

$$= \int_{\frac{1}{n}}^{G(Z)} \frac{1}{s(\log s)^2}\int_{\frac{1}{n}}^{s}\frac{dt}{(\log t)^2}ds + \int_{\frac{1}{n}}^{G(Z)}\frac{1}{(\log s)^2}\int_{s}^{G(Z)}\frac{dt}{t(\log t)^2}ds$$

$$-\int_{\frac{1}{n}}^{G(Z)}\int_{\frac{1}{n}}^{G(Z)}\frac{dt\ ds}{(\log t)^2(\log s)^2}\,, \tag{4.35}$$

Now by integrating by parts the first integral with $s^{-1}(\log s)^{-2} = (-\int_s^{G(Z)} \frac{dt}{t(\log t)^2})'$, we get

$$\frac{Z^2}{\lambda^2}\sigma_n^2 = 2\int_{\frac{1}{n}}^{G(Z)}\frac{1}{(\log s)^2}\int_{s}^{G(Z)}\frac{dt}{t(\log t)^2}ds - (\int_{\frac{1}{n}}^{G(Z)}\frac{dt}{(\log t)^2})^2\,. \tag{4.36}$$

Noting that $-1/s(\log s)^2$ is the derivative of $1/\log s$, we conclude that

(4.37)
$$\frac{Z^2}{\lambda^2}\sigma_n^2 = \frac{-2}{\log G_n(Z)}\int_{1/n}^{G(Z)}(\log s)^{-2}ds + 2\int_{1/n}^{G(Z)}(\log s)^{-3}ds - (\int_{\frac{1}{n}}^{G(Z)}\frac{ds}{(\log s)^2})^2$$

(4.38)
$$\rightarrow \frac{Z^2}{\lambda^2}\sigma^2 = \frac{-2}{\log G(Z)}\int_{0}^{G(Z)}(\log s)^{-2}ds + 2\int_{0}^{G(Z)}(\log s)^{-3}ds - (\int_{0}^{G(Z)}\frac{ds}{(\log s)^2})^2,$$

with $\mathrm{G}(Z) = exp(-\lambda/(Z-y_0))$ for $Z > y_0$.

Proof of theorem2. We have for for $\gamma > 0$ and x≥ 1, F(x)=1-x$^{-1/\gamma}$, that is $F^{\leftarrow}(s) = (1-s)^{-\gamma}$ for $0 \leqslant s < 1$. Thus by (4.7)

$$I_n = \frac{1}{nZ}q(Z-y_0) - \frac{1}{nZ}U_{1,n}^{\gamma} - \int_{\frac{1}{n}}^{G_n(Z)} V_n(s)^{\gamma}\, ds. \tag{4.39}$$

Put

$$D_n = \frac{1}{nZ}q(Z-y_0) - \frac{1}{Z}\int_{\frac{1}{n}}^{G_n(Z)} s^{\gamma} ds \tag{4.40}$$

By the mean value theorem, we get

$$I_n - D_n = \frac{1}{nZ}U_{1,n}^{\gamma} + \frac{\gamma}{Z}\int_{\frac{1}{n}}^{G_n(Z)} (V_n(s)-s)(\zeta_n(s))^{\gamma-1} ds, \tag{4.41}$$

where $\zeta_n(s) \in [\min(s, V_n(s)), \max(s, V_n(s))]$ and

$$\sqrt{n}\,(I_n - D_n) = \frac{1}{\sqrt{n}Z}U_{1,n}^{\gamma} - \frac{\gamma}{Z}\int_{\frac{1}{n}}^{G_n(Z)} \beta_n(s)(\zeta_n(s))^{\gamma-1} ds.$$

$$= \frac{1}{\sqrt{n}Z}U_{1,n}^{\gamma} - \frac{\gamma}{Z}\int_{\frac{1}{n}}^{G_n(Z)} (\beta_n(s) - B_n(s))\,(\zeta_n(s))^{\gamma-1} ds - \frac{\gamma}{Z}\int_{\frac{1}{n}}^{G(Z)} B_n(s)((\zeta_n(s))^{\gamma-1} - s^{\gamma-1}) ds$$

(4.42)

$$-\frac{\gamma}{Z}\int_{G(Z)}^{G_n(Z)} B_n(s)(s)^{\gamma-1} ds - \frac{\gamma}{Z}\int_{\frac{1}{n}}^{G(Z)} B_n(s)(s)^{\gamma-1} ds \;=: R_{0,n} + R_{1,n} + R_{2,n} + R_{3,n} + N_n$$

We first claim that each of the R$_{j,n}$ tends to zero in probability as n$\rightarrow \infty$.

Claim 1.2. R$_{0,n} \rightarrow_p 0$. This immediate since $\gamma > 0$ and U$_{1,n} \rightarrow_p 0$ as n$\rightarrow \infty$.
Claim 2.2. R$_{1,n} \rightarrow_p 0$. Put $0 < \nu < 1/2$ and $\nu + 1/2 < 1 + \gamma > 1$ so that $\delta = \frac{1}{2} + \nu - \gamma < 1$. Then by (4.12),

$$\mathbf{R}_{1,n} = O_p(n^{-\nu})\int_{\frac{1}{n}}^{G_n(Z)} s^{\frac{1}{2}-\upsilon}\; \zeta_n(s)^{-1+\gamma} ds \tag{4.43}$$

$$= O_p(n^{-\nu})\int_{\frac{1}{n}}^{G_n(Z)} s^{-(\frac{1}{2}+\upsilon-\gamma)}\; (\zeta_n(s)/s)^{-1+\gamma}\; ds = O_p(n^{-\nu})R'_{1,n}.$$

We may use the same methods used in (4.23) to have for any $\varepsilon > 0$, for large values of n and ρ,

$$0 \leq R'_{1,n} \leqslant \max(\rho^{-1+\gamma}, \rho^{1-\gamma})\int_{\frac{1}{n}}^{G(Z)+\varepsilon} s^{-(\frac{1}{2}+\upsilon-\gamma)}\; ds.$$

$$0 \leq R'_{1,n} \leq \max(\rho^{-1+\gamma}, \rho^{1-\gamma})\int^{G(Z)+\varepsilon} s^{-\delta}\; ds$$

$$= \max(\rho^{-1+\gamma}, \rho^{1-\gamma})(G(Z)+\varepsilon)^{1-\delta}. \tag{4.44}$$

The formulae (4.43) and (4.44) together achieve the proof of Claim 2.2

Claim 3.2 $R_{2,n} \to_p 0$

$$R_{2,n} = -\frac{\gamma}{Z}\int_{\frac{1}{n}}^{G_n(Z)} B_n(s)\ (\zeta_n(s)^{\gamma-1} - s^{\gamma-1})\ ds \tag{4.45}$$

$$\mathbf{R}_{2,n} = -\frac{\gamma(\gamma-1)}{Z\sqrt{n}}\int_{\frac{1}{n}}^{G_n(Z)} B_n(s)\ \sqrt{n}(\zeta_n(s)-s)\ \pi_n(s)^{\gamma-2}ds \tag{4.46}$$

where $\pi_n(\cdot)$ is defined above, so that

$$|\mathbf{R}_{2,n}| \leq \frac{\gamma\,|\gamma-1|}{Z\sqrt{n}}\int_{\min(G_n(Z),1/n)}^{\max(G_n(Z),1/n)} |B_n(s)| \times \left|\sqrt{n}(V_n(s)-s)\right|\ \pi_n(s)^{\gamma-2}ds. \tag{4.47}$$

By using (4.26) and (4.27), we get

$$|\mathbf{R}_{2,n}| \leq O_p(n^{-1/2}\log\log n)\int_{\min(G_n(Z),1/n)}^{\max(G_n(Z),1/n)} s^{\gamma-1}(1-s)\ (\pi_n(s)/s)^{\gamma-2}ds. \tag{4.48}$$

As in what preceeds, we have for any $\varepsilon > 0$, for large values of n and $\rho > 0$,

$$|\mathbf{R}_{2,n}| \leq O_p(n^{-1/2}\log\log n)\max(\rho^{-2+\gamma}, \rho^{2-\gamma})\int_0^{G(Z)+\varepsilon} s^{\gamma-1}ds. \tag{4.49}$$

Hence

$$|\mathbf{R}_{2,n}| \leq O_p(n^{-1/2}\log\log n)\ (G(Z)+\varepsilon)^{\gamma}. \tag{4.50}$$

This achieves the proof of Claim 3.2.

Claim 4.2. $R_{3,n} \to_p 0$. The proof is the same as that of Claim 4.1 with $(-1/(s(logs)^2)$ replaced by $(s^{\gamma-2})$.

It remains to remark as below that N_n is a centered Gaussian random variable with variance

$$\theta_n^2 = \left(\frac{\gamma}{Z}\right)^2\int_{\frac{1}{n}}^{G_n(Z)}\int_{\frac{1}{n}}^{G_n(Z)}(\min(s,t)-st)\ t^{\gamma-1}s^{\gamma-1}dt\ ds. \tag{4.51}$$

A direct calculation yields $\theta_n^2 \to \theta^2 = (\gamma/Z)^2 G(Z)^{2\gamma+1}(\frac{2}{(\gamma+1)(2\gamma+1)} - \frac{G(Z)}{(\gamma+1)^2})$ with $G(Z) = ((Z-y_0)^{\frac{1}{\gamma}} \wedge 1)$.

5. Simulations of the results

We will simulate our results in the conditions of the senegalease data base ESAM (1996) related to a national survey with 3278 households. The income variable is available for ten (10) districts with a minimal size of n=172. The poor headcounts in the different areas vary from 35% to 81%. As a consequence, we make our simulations for the current estimations of γ (wich varies around 1,5) and λ (wich varies around 25000), with n=50 and for several values of the headcounts (from 35% to 85%). For each case, we will use B=100 replications and report I_n the mean value of the empirical poverty intensity, D or C the centering coefficient, E the mean quadratic error, ET^2 the variance of the empirical intensity and the mean P-value for the normality test.

The two following tables show that convergence is very fast even for a small size since the mean quadratic error is very small. The variances will permit to have the confidence intervals of the exact poverty intensity. The high values of the p-values P are very conclusive for the asymptotic normality.

Here are our results for the Pareto law :

Q/n(%)	35	45	55	75	85
γ =0,8	$I(\%) = 14,989$ $C = 15,29$ $E(\%) = 1,7$ $ET = 0.266$ $P(\%) = 26,38$	$I = 19,59$ $C = 19,79$ $E = 2,22$ $ET = 0,29$ $P = 20,49$	$I = 24,48$ $C = 24,10$ $E = 2,46$ $ET = 0,3$ $P = 24,94$	$I = 31,38$ $C = 31,14$ $E = 2,11$ $ET = 0,31$ $P = 28,25$	$I = 37,82$ $C = 37,47$ $E = 2,23$ $ET = 0,3$ $P = 27$
γ =2	$I(\%) = 22,93$ $C = 22,95$ $E(\%) = 3,35$ $ET = 0,36$ $P(\%) = 23,89$	$I = 29,35$ $C = 29,79$ $E = 2,97$ $ET = 0,89$ $P = 24,62$	$I = 36,74$ $C = 36,28$ $E = 3,18$ $ET = 0,4$ $P = 23,5$	$I = 46,53$ $C = 46,92$ $E = 3,37$ $ET = 0,4$ $P = 23,88$	$I = 55,91$ $C = 55,93$ $E = 3,3$ $ET = 0,36$ $P = 20,84$

Here are our results for the exponential law :

Q/n(%)	35	45	55	75	85
λ =2	$I(\%) = 13,46$ $C = 13,6$ $E(\%) = 0,04$ $ET = 0,476$ $P(\%) = 25$	$I = 19,8$ $C = 19,73$ $E = 0,038$ $ET = 0,447$ $P = 25,18$	$I = 27,38$ $C = 27,55$ $E = 0,07$ $ET = 0,55$ $P = 24,26$	$I = 42,34$ $C = 42,4$ $E = 0,08$ $ET = 0,579$ $P = 22,88$	$I = 61,68$ $C = 0,48$ $E = 0,08$ $ET = 0,567$ $P = 24,74$
λ =25000	$I(\%) = 13,70$ $C = 13,6$ $E(\%) = 0,04$ $ET = 0,476$ $P(\%) = 25$	$I = 19,8$ $C = 19,73$ $E = 0,03$ $ET = 0,517$ $P = 25,93$	$I = 27,39$ $C = 27,4$ $E = 0,06$ $ET = 0,55$ $P = 25,51$	$I = 42,57$ $C = 42,48$ $E = 0,07$ $ET = 0,574$ $P = 23,68$	$I = 61,83$ $C = 61,18$ $E = 0,07$ $ET = 0,563$ $P = 24,73$

6. Conclusion

We have been able to get the asymptotic normality of the poverty intensity for two special representatives of the two extremal domains. The results are very well supported by the simulations. This opens the way for a general study of this poverty measure in the frame of extreme value theory first and futher for the nonweighted measures. Immediate steps after this study will consist of deep comments of the results as well as simulations from the 1995 senegalease data.

References

[1] Document Stratégique de Réduction de la Pauvreté, site www.finances.gouv.sn/dsrp02.htlm, Sénégal.
[2] Csörgö, S. Csörgö M. Horvàth, L. and Mason, M. (1986). Weighted empirical and quantile processes. Ann. Probab. 14,31-85.
[3] Csörgö, S., Deheuvels P. and Mason, D.M. (1985). Kernel estimates of the tail index of a distribution. Ann. Statist. 13, 1467-1487.
[4] Csörgö, S. and Mason, D.M. (1985). Central limit theorems for sums of extreme values. Math Proc. Cambridge Philos. Soc, 98, 547-558,
[5] de Haan, L. (1970). On regular variation and its application to the weak convergence of sample extremes. Mathematical contre Tracts, 32, Amsterdam.
[6] Dekkers, A.L. M, Einmahl, J.H.J. et De Haan, L. (1989). A moment estimator Ann. Statist.,17, 1833-1855.
[7] Galambos, J.(1985). The Asymptotic Theory of Extreme Order Statistics. Wiley, New-York.
[8] Lo, G.S.(1989). A note on the asymptotic normality of sums of extreme values. J. statistics. Plann. Inf., 22, 127-136.
[9] Lo, G.S.(2003). Comportement asymptotique des indicateurs de pauvreté. Préprint du Laboratoire de Techniques Quantitatives 03/9, Université Cheikh Anta Diop, Dakar, Editées par le Centre de Recherches Appliquées (CREA), Sénégal.
[10] Lo, G.S.(2003). Convergence des indicateurs de pauvreté pondérés de Sen et de Shorrocks. Publications de l'UFR SAT, Lerstad, 1, 2003, Université Gaston Berger de Saint-Louis.
[11] Hasofer, A.M. et Wang Z.(1987). A test for extreme value domain of attraction. J. Amer. Statist. Assoc., vol 57, 717, Theory and Method, 171-177.
[12] Hill, B.M. (1975). A simple general approach to the inference about the tail index of a distribution. Ann Statist. 3, 1163-1174.
[13] Mason, D.M. (1982). Laws of large numbers for sums of extreme values. Ann. Probab. 10, 754 - 764.
[14] Sen Amartya K.(1976). Poverty: An Ordinal Approach to Measurement. Económetrica, 44, 219-231.
[15] Resnick, S.I. (1987). Extreme Values, Regular Variation and Point Processes. Springer-Verbag, New-York.
[16] Smith, R.L.(1987). Estimating tail of probability distribution. Ann. Statist. 15, 1174-1207
[17] Shorack G.R. and Wellner J. A.(1986). Empirical Processes with Applications to Statistics, wiley-Interscience, New-York.
[18] Taleb Ely Ould Taleb Ahmed(2003). Les mesures de la pauvreté, de l'inégalité et l'impact de la croissance économique. Mémoire de DEA, Université Gaston Berger de Saint-Louis.

Lerstad and Université Gaston Berger. BP 234, Saint-Louis, Sénégal.
E-mail address: ganesamblo@ufrsat.org.sn

Current address: Lerstad and Ecole Normale Supérieure, Dakar, Sénégal.
E-mail address: toubasall@yahoo.fr