Filomat 37:2 (2023), 591–613
https://doi.org/10.2298/FIL2302591E

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Confidence ellipses for simple regression parameters with strongly mixing errors

Abdelkader Elmoumen[a], Halima Zerouati[b], Abdelnasser Dahmani[c]

[a]*Faculty of Science and Technology, University of Tamenghasset, Algeria*
[b]*Laboratoire de Mathématiques Appliquées, Faculté des Sciences Exactes, Université de Bejaia, Bejaia, Algérie*
[c]*Laboratoire de Mathématiques Appliquées, Faculté des Sciences Exactes, Université de Bejaia, Bejaia, Algérie and University of Tamenghasset, Algeria*

Abstract. In this work, we establish exponential inequalities which allow us to construct confidence regions, characterized by ellipses, for the least squares estimates of the parameters in the case of a linear regression model with α-mixing errors.

1. Introduction

The concept of regression comes from genetics. Simple linear regression theory was established and popularized by Galton [22] during the late 19th century with the publication entitled "Regression towards mediocrity in hereditary stature". His work was later extended by (Yule [41]; Pearson [33] and Fisher [20]) to a more general statistical context. Since then, many works have been developed from this theory, in several fields (Volkonskii and Rozanov [39]; Babu and Bai [5]; Tingley [38] and Zvara [42]). In simple linear regression, it is assumed that two variables x and y are linearly related with unknown intercept and slope parameters. In particular, the exogenous variable x is assumed to be precisely measurable and the endogenous variable y is assumed to be a random variable depending on x via a linear function. Nowadays, in many problems of daily life, it is more usual to measure several Y variables instead of just one, to get the so called multivariate model

$$Y = X\beta + \varepsilon$$

where Y is an $n \times 1$ vector of response variables, X is an $n \times p$ matrix of predictors, β is a $p \times 1$ vector of unknown coefficients, and ε is an $n \times 1$ vector of unknown errors.

The literature on multiple linear regression is considerable (see [32] and references cited). Several authors have studied the confidence domain of the parameters of a multiple regression when the errors are independent [10, 12, 17, 28, 31, 36]. In [9], statistical procedures for linear regression are established, when the error process are assumed to be correlated.

2020 *Mathematics Subject Classification*. 62J02.
Keywords. Regression; $\alpha-$mixing random data; Confidence ellipses.
Received: 02 February 2021; Accepted: 26 June 2021
Communicated by Aleksandar Nastić
Correspondng author: Halima Zerouati
Email addresses: elmoumenaek01ra@gmail.com (Abdelkader Elmoumen), h_zerouati@yahoo.fr (Halima Zerouati)

Different ways to estimate the asymptotic covariance matrix of the least squares estimators are available. Using the covariance matrix's estimation, the confidence intervals and the usual tests on the parameters are modified. When we speak about measurement, whether we like it or not, errors will be induced due to the stochastic nature of any measurement. In most cases, specification errors are considered independent. Note that independent errors failed for modeling some phenomena. In practice this hypothesis is frequently not adapted. Indeed, dependent errors are more adjusted to reality and there is an important class of dependence which finds many applications.

To illustrate accurately this situation, we can mention the Ising model. It describes the behavior of two close electrons in an atom. These electrons are more susceptible to be oriented in the same direction than in opposite directions (Fortuin et al. [21]). This attraction is expressed through the positive dependence. It should be noted that this type of dependence which has been introduced by (Esary et al. [18]), is based on Lehmann's works (Lehmann [30]). Their objective was to find applications in reliability and statistics. However, there are many notions of dependence other than positive dependence. We refer those expressed in terms of mixing coefficients(Blum et al. [6]; Ibragimov and Linnik [26]; Tingley [38]). Many types of mixing conditions have been proposed in the literature (Doukhan [16]; Rio [35]). The main drawbacks of mixing assumptions is the difficulty of checking them. Recently, these mixing conditions have been studied not only from a probabilistic point of view, but also for their application possibilities in terms of non parametric prevision (Chai [11]).

In this work, we consider a non restrictive mixing condition to characterize the dependence between the data random errors. We suppose them to be strong mixing or α-mixing (Aiane [1], Arroudj [3]). We establish exponential inequalities, which allow us to construct confidence regions, characterized by ellipses, for the least squares estimates of the parameters (Dorogovtsev [15]), in the case of a linear regression model with α-mixing errors. To check the validity of the obtained theoretical results, some numerical results are considered.

The construction of exponential inequalities for the parameters of a multiple linear regression with mixing errors remains an open problem. Among other things, these inequalities will be used to construct hyperellipsoids for the mentioned parameters.

2. Preliminaries

Let us consider the simple linear regression model

$$y_i = ax_i + b + \varepsilon_i$$

where a and b are the parameters to be estimated using observations $(y_1, y_2, \cdots, y_n)$ and $(\varepsilon_i)_{i\in\mathbb{N}^*}$ is a centered sequence of non-zero variances $\mathbb{E}\left(\varepsilon_i^2\right) = \sigma_i^2$ and admitting bounded fourth order moments

Denote

$$\overline{\sigma} = \max_i \sigma_i^2 \text{ and } M = \frac{1}{n}\sum_{i=1}^{n} x_i^2.$$

The least squares estimates of the parameters a and b are given by

$$\widehat{a_n} = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2} \text{ and } \widehat{b_n} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

and assuming that $\sum_{i=1}^{n} x_i = 0$, we obtain

$$\widehat{a_n} - a = \frac{\sum_{i=1}^{n} x_i\varepsilon_i}{\sum_{i=1}^{n} x_i^2} \text{ and } \widehat{b_n} - b = \frac{\sum_{i=1}^{n} \varepsilon_i}{n}.$$

It is obvious that $\widehat{a_n}$ and $\widehat{b_n}$ are unbiased estimators of a and b.

The α-mixing notion is defined in the following way (Rosenblatt [37] 1956)

Definition 2.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and we note $\mathcal{F}_1^i \subset \mathcal{F}$ (respectively $\mathcal{F}_{i+p}^{\infty} \subset \mathcal{F}$) the σ-algebra generated by $\left\{\varepsilon_j, 1 \le j \le i\right\}$ (respectively by $\left\{\varepsilon_j, j \ge i+p\right\}$). The sequence $(\varepsilon_i)_{i\in\mathbb{N}}$ is said to be α-mixing if:*

$$\alpha_\tau = \sup_{A\in F_1^i, B\in F_{i+\tau}^{+\infty}} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| \underset{\tau\to+\infty}{\to} 0.$$

The field of strong mixing conditions is a vast area. In this regard, we point out: central limit theorems (Rosenblatt [37]); order Statistics (Welsch [40]); empirical processes (Deo [14]); iterated logarithm (Philipp [34]); Strong law of large numbers (Chen [13]); robust estimations (Boente and Fraiman [7]) and nonparametric statistics (Ferraty and Vieu [19]).

While the strong mixing condition has been widely adopted in the literature, the number of well-known processes which satisfy the strong mixing condition is still somewhat limited. m-dependent processes are trivially strong mixing, since $\alpha(s) = 0$, for all $s > m$. Kolmogorov and Rozanov (Kolmogorov and Rozanov [27]) proved that Gaussian processes with continuous and positive spectral densities are strong mixing. Ibragimov and Linnik (Ibragimov and Linnik [26]) have shown that stationary Markov processes are strong mixing under some conditions on the transition probabilities. Lee (Lee [29]) shows that $AR(1)$ satisfies the strong mixing condition with mixing order decaying exponentially to zero. Athreya and Pantula (Athreya and Pantula [4]) establish that certain stationary autoregressive moving average processes are strong mixing.

However, many strictly stationary linear processes fail to be α-mixing. A well known classic example is the strictly stationary $AR(1)$ process generated by Bernoulli $\mathfrak{B}(p)$ innovation random variable and an autoregressive parameter $\phi \in]0, 1/2]$ (Andrews [2]).

3. Results

Proposition 3.1. *For any $\varepsilon > 0$, we have*

$$\mathbb{P}\left\{\frac{\left(\sum_{i=1}^n x_i^2\left|\widehat{a_n}-a\right|^2 + n\left|\widehat{b_n}-b\right|^2\right)^{\frac{1}{2}}}{\left(\frac{1}{n}\sum_{i=1}^n\left(y_i-\widehat{a_n}x_i-\widehat{b_n}\right)^2\right)^{\frac{1}{2}}} > R\right\} \le \mathbb{P}\left\{\frac{\left|\sum_{i=1}^n x_i\varepsilon_i\right|}{\left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}} > r\right\} + \mathbb{P}\left\{\frac{1}{\sqrt{n}}\left|\sum_{i=1}^n \varepsilon_i\right| > r\right\} + \mathbb{P}\left\{\frac{1}{n}\sum_{i=1}^n \varepsilon_i^2 \le \varepsilon\right\} \tag{1}$$

with

$$r = R\sqrt{\frac{\varepsilon}{2\left(1+\frac{R^2}{n}\right)}}.$$

Note that

$$\frac{1}{n}\sum_{i=1}^n\left(y_i-\widehat{a_n}x_i-\widehat{b_n}\right)^2 = \frac{1}{n}\sum_{i=1}^n\varepsilon_i^2 - \frac{1}{n}\left(\sum_{i=1}^n x_i^2\left|\widehat{a_n}-a\right|^2 + \left|\widehat{b_n}-b\right|^2\right).$$

Using the following properties

$$\mathbb{P}\{X > tY\} \le \mathbb{P}\{X > t\varepsilon\} + \mathbb{P}\{Y \le \varepsilon\} \text{ and } \mathbb{P}\{X+Y > t\} \le \mathbb{P}\left\{X > \frac{t}{2}\right\} + \mathbb{P}\left\{Y > \frac{t}{2}\right\}$$

we deduce the inequality (1).

Without loss of generality, we will suppose that the sequence $(\alpha_n)_n$ is decreasing and that $\alpha_0 = 1$.

Theorem 3.2. *Let $(\varepsilon_i, n \geq 1)$ be a centered sequence of α-mixing random elements, bounded by L. If*

$$\sum_{q=1}^{+\infty} \alpha_q < +\infty \text{ and } 2\frac{n\sqrt{e}}{k}(\alpha_k)^{\frac{2k}{3n}} < A \tag{2}$$

then, for any $0 < \varepsilon < \min_i \mathbb{E}\varepsilon_i^2, n \geq 4, k \in \left\{1, 2...., \left[\frac{n}{2} - 1\right]\right\}$, ($\left[\frac{n}{2} - 1\right]$ refers to the integer part of $\left(\frac{n}{2} - 1\right)$), we have

$$\mathbb{P}\left\{\frac{\left(\sum_{i=1}^{n} x_i^2 \left|\widehat{a}_n - a\right|^2 + n\left|\widehat{b}_n - b\right|^2\right)^{\frac{1}{2}}}{\left(\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{a}_n x_i - \widehat{b}_n\right)^2\right)^{\frac{1}{2}}} > R\right\} \leq A_1\left(e^{-A_2 r^2} + e^{-B_2 r^2} + e^{-C_2 d^2 n}\right) \tag{3}$$

where In particular if $R = \sqrt{\rho n}$, then

$$\mathbb{P}\left\{\frac{\left(\sum_{i=1}^{n} x_i^2 \left|\widehat{a}_n - a\right|^2 + n\left|\widehat{b}_n - b\right|^2\right)^{\frac{1}{2}}}{\left(\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{a}_n x_i - \widehat{b}_n\right)^2\right)^{\frac{1}{2}}} > R\right\} \leq A_1\left(e^{-A_2 \xi n} + e^{-B_2 \xi n} + e^{-C_2 d^2 n}\right).$$

with

$$\xi = \frac{\rho\varepsilon}{2(1+\rho)}.$$

We deal with the second member of the inequality (1). Let the sequence (ζ_i) be defined by $\zeta_i = \varepsilon_i x_i$. This sequence is α-mixing, centered and satisfies:

$$|\zeta_i| = |\varepsilon_i x_i| \leq L_1, \mathbb{E}\zeta_i^2 \leq D_1.$$

The similar of the Bernstein-Fréchet's inequality (Bosq and Lecoutre [8]) applied to the α-mixing sequence provides: For any $z > 0$, $n \geq 4$, $k \in \left\{1, 2, ..., \left[\frac{n}{2} - 1\right]\right\}$,

$$\mathbb{P}\left\{\left|\sum_{i=1}^{n}\zeta_i\right| > r\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}}\right\} \leq 2\exp\left(-zr\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}} + 4z^2 e\left(D_1 + 8L_1^2\sum_{q=1}^{k}\alpha_q\right)n + 2\sqrt{e}(\alpha_k)^{\frac{2k}{3n}}\frac{n}{k}\right). \tag{4}$$

Then, z is replaced by the value which minimizes the second member of the inequality (4), and we obtain

$$\mathbb{P}\left\{\left|\sum_{i=1}^{n}\zeta_i\right| > r\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}}\right\} \leq 2\exp\left(\frac{-r^2\sum_{i=1}^{n}x_i^2}{16e\left(D_1 + 8L_1^2\sum_{q=1}^{k}\alpha_q\right)n} + 2\sqrt{e}(\alpha_k)^{\frac{2k}{3n}}\frac{n}{k}\right).$$

According to (2), it follows that

$$\mathbb{P}\left\{\frac{\left|\sum_{i=1}^{n}\varepsilon_i x_i\right|}{\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}}} > r\right\} \leq A_1 e^{-A_2 r^2}. \tag{5}$$

The inequality

$$\mathbb{P}\left\{\frac{1}{\sqrt{n}}\left|\sum_{i=1}^{n}\varepsilon_i\right|>r\right\}\leq A_1e^{-B_2r^2} \tag{6}$$

is deduced as above.

The following lemma is used for the third term of the second member of (1).

Lemma 3.3. *(Ibragimov, Linnik [26], Theorem 17.2.1 page 306). Let $(\varepsilon_i)_{i\in\mathbb{N}}$ be a strong mixing sequence. If ε_i is measurable in regard to the σ−algebra $\mathcal{F}_1^i$, and ε_i $i\neq j$, is measurable with regard to the σ−algebra $\mathcal{F}_j^\infty$, and moreover, if $|\varepsilon_i|\leq L<+\infty$, $\left|\varepsilon_j\right|\leq L<+\infty$, then*

$$\left|\mathbb{E}\left(\varepsilon_i\varepsilon_j\right)-\mathbb{E}\left(\varepsilon_i\right)\mathbb{E}\left(\varepsilon_j\right)\right|\leq 4L^2\alpha_{|i-j|}. \tag{7}$$

In particular, when the sequence is centered, then,

$$\left|\mathbb{E}\left(\varepsilon_i\varepsilon_j\right)\right|\leq 4L^2\alpha_{|i-j|}.$$

An inequality of the form (7), but with a constant equal to 16, is proved in [39].

The sequence (η_i) defined by $\eta_i=\mathbb{E}\varepsilon_i^2-\varepsilon_i^2$ is α-mixing and satisfies

$$\left|\eta_i\right|\leq 2L^2 \text{ and } \mathbb{E}\eta_i^2\leq D_2.$$

Furthermore, from Markov's inequality, we obtain for any $z>0$ and $0<\varepsilon<\min\mathbb{E}\varepsilon_i^2$

$$\begin{aligned}
&\mathbb{P}\left\{\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2\leq\varepsilon\right\}=\mathbb{P}\left\{\sum_{i=1}^{n}\varepsilon_i^2\leq n\varepsilon\right\}=\mathbb{P}\left\{\sum_{i=1}^{n}\varepsilon_i^2-\sum_{i=1}^{n}\mathbb{E}\varepsilon_i^2\leq n\varepsilon-\sum_{i=1}^{n}\mathbb{E}\varepsilon_i^2\right\}\\
\leq\ &\mathbb{P}\left\{\sum_{i=1}^{n}\left(\varepsilon_i^2-\mathbb{E}\varepsilon_i^2\right)\leq n\left(\varepsilon-\min_{1\leq i\leq n}\mathbb{E}\varepsilon_i^2\right)\right\}=\mathbb{P}\left\{\sum_{i=1}^{n}\left(\mathbb{E}\varepsilon_i^2-\varepsilon_i^2\right)\geq n\left(\min_{1\leq i\leq n}\mathbb{E}\varepsilon_i^2-\varepsilon\right)\right\}\\
\leq\ &\exp\left(-znd\right)\mathbb{E}\exp\left(zH_n\right)
\end{aligned} \tag{8}$$

with $d=\left(\min_{1\leq i\leq n}\mathbb{E}\varepsilon_i^2-\varepsilon\right)$ and $H_n=\sum_{i=1}^{n}\eta_i$. Let us bound $\mathbb{E}\exp\left(zH_n\right)$.

To this end, let us set

$$Y_{j;q}=\sum_{m=m_1}^{m_2}\eta_m;\ j=1,2;\ q=1,2,...,N \text{ and } Z_{j,q}=\sum_{s=1}^{q}Y_{j;s};\ j=1,2;\ q=1,2,...,N$$

where N is defined by $2k(N-1)<n\leq 2kN$, $m_1=\inf\{(2q+j-3)k+1,n\}$, $m_2=\inf\{m_1+k-1,n\}$.

Thus, we have

$$H_n=Z_{1,N}+Z_{2,N}.$$

The convexity of the exponential function allows to get

$$\mathbb{E}\exp\left(zH_n\right)\leq\frac{1}{2}\left(\mathbb{E}\exp\left(2zZ_{1,N}\right)+\mathbb{E}\exp\left(2zZ_{2,N}\right)\right). \tag{9}$$

On the other hand, the sequence (η_m) being α-mixing, it follows that (Hipp [25] 1979, 60)

$$\mathbb{E}\exp\left(2zZ_{j,q}\right)\leq\mathbb{E}\exp\left(2zZ_{j,q-1}\right)\mathbb{E}\exp\left(2zY_{j,q}\right)+4\left(\alpha_k\right)^{\frac{1}{h}}\left\|\exp\left(2zY_{j,q}\right)\right\|_\infty\mathbb{E}^{\frac{1}{p}}\left(\exp\left(2zpZ_{j,q-1}\right)\right)$$

where $h, p \in [1, +\infty]$; $\frac{1}{h} + \frac{1}{p} = 1$; $j = 1, 2$; $q = 1, ..., N$ with the convention $Z_{j,0} = 0$.
For z well chosen, for example

$$\left|zY_{j,q}\right| \leq \frac{1}{4},$$

we get

$$\mathbb{E}\exp\left(2zY_{j,q}\right) \leq 1 + 4z^2\mathbb{E}Y_{j,q}^2 \leq \exp\left(4z^2\mathbb{E}Y_{j,q}^2\right). \tag{10}$$

Lemma 4 leads to

$$\mathbb{E}Y_{j,q}^2 = \sum_{m=m_1}^{m_2} \mathbb{E}\eta_m^2 + \sum_{\substack{m_1 \leq s, s' \leq m_2 \\ s \# s'}} E\eta_s\eta_{s'} \leq kVar\left(\varepsilon_i^2\right) + 32L^4 \sum_{q=1}^{k} \alpha_q.$$

Denote

$$C = Var\left(\varepsilon_i^2\right) + 32L^4 \sum_{q=1}^{k} \alpha_q \tag{11}$$

to obtain

$$\mathbb{E}Y_{j,q}^2 \leq kC.$$

Consequently, (10) gives

$$\mathbb{E}\exp\left(2zY_{j,q}\right) \leq \exp\left(4z^2kC\right)$$

and

$$\mathbb{E}\exp\left(2zZ_{j,q}\right) \leq \mathbb{E}\exp\left(2zZ_{j,q-1}\right)\exp\left(4z^2kC\right) + 4\left(\alpha_k\right)^{\frac{1}{h}} \sqrt{e}\mathbb{E}^{\frac{1}{p}}\left(\exp\left(2zpZ_{j,q-1}\right)\right).$$

Applying Hölder's inequality, we obtain

$$\mathbb{E}\exp\left(2zZ_{j,q}\right) \leq \left(\exp\left(4z^2kC\right) + 4\left(\alpha_k\right)^{\frac{1}{h}} \sqrt{e}\right)\mathbb{E}^{\frac{1}{p}}\left(\exp\left(2zpZ_{j,q-1}\right)\right).$$

Similarly, we obtain for $q \geq 2$

$$\mathbb{E}^{\frac{1}{p}}\exp\left(2zpZ_{j,q-1}\right) \leq \left(\exp\left(4z^2p^2kC\right) + 4\left(\alpha_k\right)^{\frac{1}{h}}\left\|\exp\left(2zpY_{j,q-1}\right)\right\|_{\infty}\right)^{\frac{1}{p}} \mathbb{E}^{\frac{1}{p^2}}\left(\exp\left(2zp^2Z_{j,q-2}\right)\right),$$

and so on until

$$\begin{aligned}\mathbb{E}^{\frac{1}{p^{q-2}}}\exp\left(2zp^{q-2}Z_{j,2}\right) \quad &\leq \quad \left(\exp\left(4z^2p^{2(q-2)}kC\right) + 4\left(\alpha_k\right)^{\frac{1}{h}}\left\|\exp\left(2zp^{q-2}Y_{j,2}\right)\right\|_{\infty}\right)^{\frac{1}{p^{q-2}}} \times \\ &\times\mathbb{E}^{\frac{1}{p^{q-1}}}\left(\exp\left(2zp^{q-1}Y_{j,1}\right)\right).\end{aligned}$$

Multiplying all these inequalities member to member, we get

$$\begin{aligned}\mathbb{E}\exp\left(2zZ_{j,q}\right) \quad &\leq \quad \prod_{i=1}^{q-1}\left(\left(\exp\left(4z^2p^{2(i-1)}kC\right) + 4\left(\alpha_k\right)^{\frac{1}{h}}\left\|\exp\left(2zp^{i-1}Y_{j,q-i+1}\right)\right\|_{\infty}\right)^{\frac{1}{p^{i-1}}}\right) \times \\ &\times \quad \mathbb{E}^{\frac{1}{p^{q-1}}}\left(\exp\left(2zp^{q-1}Y_{j,1}\right)\right).\end{aligned}$$

If we choose $h = q+1$ and $p = 1 + \frac{1}{q}$; then

$$p^{i-1} \leq \left(1+\frac{1}{q}\right)^q \leq e, i = 1, ..., q$$

and, we get

$$\left\|\exp\left(2zp^{i-1}Y_{j,q-i+1}\right)\right\|_\infty \leq \sqrt{e},$$

which ensures that

$$\left(\exp\left(4z^2p^{2(i-1)}kC\right) + 4\left(\alpha_k\right)^{\frac{1}{h}}\left\|\exp\left(2zp^{i-1}Y_{j,q-i+1}\right)\right\|_\infty\right)^{\frac{1}{p^{i-1}}} \leq \exp\left(4\sqrt{e}\alpha_k^{\frac{1}{h}} + 4z^2p^{2(i-1)}kC\right), \; i = 1, ..., q-1.$$

And since $\mathbb{E}^{\frac{1}{p^{q-1}}}\left(\exp\left(2zp^{q-1}Y_{j,1}\right)\right) \leq \exp\left(4z^2ekC\right)$, we obtain

$$\mathbb{E}\exp\left(2zZ_{j,q}\right) \leq \exp\left(4z^2keCq + 4\sqrt{e}\left(\alpha_k\right)^{\frac{1}{h}}(q-1)\right).$$

Choosing $q = N$ and noting that $2(N-1) < \frac{n}{k}$, $N < \frac{n}{k}$, $\frac{1}{h} = \frac{1}{N+1} > \frac{2k}{3n}$, we obtain

$$\mathbb{E}\exp\left(2zZ_{j,q}\right) \leq \exp\left(4z^2eCn + 2\sqrt{e}\left(\alpha_k\right)^{\frac{2k}{3n}}\frac{n}{k}\right), j = 1, 2. \tag{12}$$

The inequality

$$\mathbb{P}\left\{\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 \leq \varepsilon\right\} \leq \exp\left(-znd + 4z^2eCn + 2\sqrt{e}\left(\alpha_k\right)^{\frac{2k}{3n}}\frac{n}{k}\right) \tag{13}$$

is deduced from Lemma 3.3 and also from (8), (9), (11) and (12).

The real z which minimizes the second term of (13) is

$$z = \frac{d}{8eC}. \tag{14}$$

Hence the inequality

$$\mathbb{P}\left\{\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 \leq \varepsilon\right\} \leq \exp\left(-\frac{nd^2}{16e\left(Var\left(\varepsilon_i^2\right) + 32L^4\sum_{q=1}^{k}\alpha_q\right)} + 2\sqrt{e}\left(\alpha_k\right)^{\frac{2k}{3n}}\frac{n}{k}\right)$$

arises from (11), (13) and (14).

The assumptions of the Theorem 3 suggest that there are two constants A_1 and C_2 such that:

$$\mathbb{P}\left\{\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 \leq \varepsilon\right\} \leq A_1\exp\left(-C_2d^2n\right). \tag{15}$$

The results of the Theorem 3 are obtained from (5), (6) and (15).

4. Confidence ellipses

We have

$$\lim_{n\to+\infty} A_1\left(\exp\left(-A_2\xi n\right)+\exp\left(-B_2\xi n\right)+\exp\left(-C_2 d^2 n\right)\right)=0.$$

Thus, for a given level α and ξ, we identify the smallest n which fulfills the following inequality

$$A_1\left(\exp\left(-A_2\xi n\right)+\exp\left(-B_2\xi n\right)+\exp\left(-C_2 d^2 n\right)\right)\leq \alpha.$$

We will note this n by n_α, then, we get

$$\mathbb{P}\left\{\frac{\left(\sum_{i=1}^{n_\alpha} x_i^2\left|\widehat{a}_{n_\alpha}-a\right|^2+n_\alpha\left|\widehat{b}_{n_\alpha}-b\right|^2\right)^{\frac{1}{2}}}{\left(\frac{1}{n_\alpha}\sum_{i=1}^{n_\alpha}\left(y_i-\widehat{a}_{n\alpha}x_i-\widehat{b}_{n_\alpha}\right)^2\right)^{\frac{1}{2}}}>\sqrt{\frac{2\xi}{\varepsilon-2\xi}n_\alpha}\right\}\leq\alpha$$

or else

$$\mathbb{P}\left\{\frac{\left(\sum_{i=1}^{n_\alpha} x_i^2\left|\widehat{a}_{n_\alpha}-a\right|^2+n_\alpha\left|\widehat{b}_{n_\alpha}-b\right|^2\right)^{\frac{1}{2}}}{\left(\frac{1}{n_\alpha}\sum_{i=1}^{n_\alpha}\left(y_i-\widehat{a}_{n_\alpha}x_i-\widehat{b}_{n_\alpha}\right)^2\right)^{\frac{1}{2}}}\leq\sqrt{\frac{2\xi}{\varepsilon-2\xi}n_\alpha}\right\}\geq 1-\alpha.$$

In other words, with a probability greater than or equal to $1-\alpha$, the parameters a and b belong to the ellipse defined by the equation

$$\frac{\left|\widehat{a}_{n_\alpha}-a\right|^2}{\frac{\rho n_\alpha S_{n_\alpha}^2}{\sum_{i=1}^{n_\alpha}x_i^2}}+\frac{\left|\widehat{b}_{n_\alpha}-b\right|^2}{\rho S_{n_\alpha}^2}\leq 1 \text{ where } S_{n_\alpha}^2=\frac{1}{n_\alpha}\sum_{i=1}^{n_\alpha}\left(y_i-\widehat{a}_{n_\alpha}x_i-\widehat{b}_{n_\alpha}\right)^2,\ \rho=\frac{2\xi}{\varepsilon-2\xi}.$$

Remark 4.1. *If ξ $\left(\xi=\frac{\varepsilon\rho}{2(1+\rho)}\right)$ and the sample size are given and satisfy*

$$A_1\left(\exp\left(-A_2\xi n\right)+\exp\left(-B_2\xi n\right)+\exp\left(-C_2 d^2 n\right)\right)<1$$

then taking

$$\alpha=A_1\left(\exp\left(-A_2\xi n\right)+\exp\left(-B_2\xi n\right)+\exp\left(-C_2 d^2 n\right)\right)$$

we obtain that a and b belong to the ellipse of center $\left(\widehat{a}_n,\widehat{b}_n\right)$ and with the coordinates of the foci

$$\begin{aligned}F_1 &= \left(\widehat{a}_n-\sqrt{\rho S_n^2\left(\frac{1}{M}-1\right)},\widehat{b}_n\right), F_2=\left(\widehat{a}_n+\sqrt{\rho S_n^2\left(\frac{1}{M}-1\right)},\widehat{b}_n\right) \text{ if } M<1\\ F_1 &= \left(\widehat{a}_n,\widehat{b}_n-\sqrt{\rho S_n^2\left(1-\frac{1}{M}\right)},\right), F_2=\left(\widehat{a}_n,\widehat{b}_n+\sqrt{\rho S_n^2\left(1-\frac{1}{M}\right)}\right) \text{ if } M>1\end{aligned}$$

with a probability greater than or equal to $1-\alpha$.

If the confidence level and the sample size are given, we determine the foci F_1 and F_2 from the equation

$$\alpha=A_1\left(\exp\left(-A_2\xi n\right)+\exp\left(-B_2\xi n\right)+\exp\left(-C_2 d^2 n\right)\right).$$

5. Simulation

In this section, we will simulate a Keynesian consumption function C_t, which is an economic formula which represents the functional relationship between total consumption and gross national income. It is represented by a simple linear regression model

$$C_t = b + aR_t + \varepsilon_t \tag{16}$$

where the constant b is interpreted as the incompressible (autonomous) consumption, the regression coefficient a of the income variable R_t is interpreted as the marginal propensity to consume (MPC) and the errors ε_t represent the unexplained part of the model.

The purpose is to build a confidence domain (Confidence ellipse) for the parameters a and b when the errors ε_t are α-mixing. To ensure that ε_t are α-mixing we propose an autoregressive process of order 1. (It is well known that this process is α-mixing [29]). Using the least squares method we estimate a and b and from Theorem 3's results, we build confidence ellipses.

Given the Keynesian consumption function (16), we will build a confidence domain (Confidence ellipse) for the parameters a and b when the errors ε_t are α-mixing using data of average monthly income per capita in Africa for 16 years expressed in dollars that appear in Table 1.

Year	Available income	Centered income R_t
2004	74	−61.19
2005	88	−47.19
2006	100	−35.18
2007	112	−23.19
2008	127	−8.19
2009	133	−2.19
2010	132	−3.19
2011	139	3.81
2012	144	8.81
2013	167	31.81
2014	169	33.81
2015	167	31.81
2016	156	20.81
2017	150	14.81
2018	151	15.81
2019	154	18.81

Table 1: Evolution of per capita income in Africa
Source: World Bank, 2019

The MPC $a = \frac{\triangle C_t}{\triangle R_t}$ is lower in developed countries, for example the USA ($\simeq 0.04$), the United Kingdom ($\simeq 0.02$) and Canada ($\simeq 0.05$) ; and higher in Africa as in Nigeria ($\simeq 0.64$) , South Africa ($\simeq 0.7$) and Zambia, ($\simeq 0.64$) . Since a large proportion of Africans are poor, we assumed that $a = 0.8$. (Methods for measuring MPC can be found in [24]). The incompressible (autonomous) consumption b is independent of disposable income. It is usually between 30% and 40% of the average. In our context, it is assumed to be equal to 50.

Thus, the theoretical consumption is equal to

$$C_t = 50 + 0.8R_t$$

and the observed consumption (generated) C_t is therefore equal to

$$C_t = 50 + 0.8R_t + \varepsilon_t.$$

5.1. Simulation with α-mixing process

In order to characterize the strong mixing random errors, we consider the first order autoregressive model

$$\varepsilon_t = \phi\varepsilon_{t-1} + v_t, \varepsilon_0 = 0, \phi = 0.04$$

and $(v_t)_t$ is an i.i.d sequence of uniform random variables $\mathcal{U}(-1, 1)$.

(We generate the first order autoregressive ε_t with R).

5.1.1. Simulation results: α-mixing case (n=16)

Year	Available income	Centered income R_t	Theoretical consumption	Generation of ε_t	Observed Consumption
2004	74	−61.19	1.05	0, 39561626	1, 44561626
2005	88	−47.19	12.25	−0, 28058696	11, 96941304
2006	100	−35.18	21.85	−0, 35656872	21, 49343128
2007	112	−23.19	31.45	0, 56408866	32, 01408866
2008	127	−8.19	43.45	0, 39911624	43, 84911624
2009	133	−2.19	48.25	0, 96873086	49, 21873086
2010	132	−3.19	47.45	0, 97374001	48, 42374001
2011	139	3.81	53.05	−0, 90880095	52, 14119905
2012	144	8.81	57.05	−0, 06660959	56, 98339041
2013	167	31.81	75.45	−0, 6759005	74, 7740995
2014	169	33.81	77.05	0, 07378586	77, 12378586
2015	167	31.81	75.45	0, 38713788	75, 83713788
2016	156	20.81	66.65	0, 57094363	67, 22094363
2017	150	14.81	61.85	0, 83260872	62, 68260872
2018	151	15.81	62.65	−0, 47168471	62, 17831529
2019	154	18.81	65.05	0, 7906821	65, 8406821

Table 2: Calculation of the observed consumption with $AR(1)$

Thus, we obtain an α-mixing process (ε_t) with mixing order $(0.04)^k$ [29].

In all of what follows, we consider that in all calculations $\varepsilon = 0.1$. It verifies the hypothesis $0 < \varepsilon < \min_i \mathbb{E}\varepsilon_i^2 \left(d = \min_i \mathbb{E}\varepsilon_i^2 - \varepsilon > 0\right)$. We also take $k = \left[\frac{n}{2} - 1\right]$.

According to table 2, we get : $\widehat{a_n} = 0.8041493, \widehat{b_n} = 50.27538599, S_n^2 = 0.507133094$ and

$$\begin{aligned} &\mathbb{P}\left\{\frac{\left|\sum_{i=1}^{16}\varepsilon_i x_i\right|}{\left(\sum_{i=1}^{16}x_i^2\right)^{\frac{1}{2}}} > r\right\} + \mathbb{P}\left\{\frac{1}{\sqrt{16}}\left|\sum_{i=1}^{16}\varepsilon_i\right| > r\right\} + \mathbb{P}\left\{\frac{1}{16}\sum_{i=1}^{16}\varepsilon_i^2 \le \varepsilon\right\} \\ = &\ 1.0106\,(0.989\,21 + 0.957\,41 + 0.969\,95) = 2.\,947\,5. \end{aligned}$$

With these probabilities we cannot conclude on the confidence level with which the parameters a and b belong to the ellipse defined by the equation:

$$\frac{(0.8 - a)^2}{\frac{8.112\rho}{12344.4}} + \frac{(50.2 - b)^2}{0.507\rho} = 1.$$

5.2. Simulation with Gaussian process

Remark 5.1. *When the errors are Gaussian i.i.d $\mathcal{N}(0,1)$, we have*

$$\frac{\sum_{i=1}^{n}\varepsilon_i}{\sqrt{n}} \rightsquigarrow \mathcal{N}(0,1), \frac{\left|\sum\limits_{i=1}^{n} x_i\varepsilon_i\right|}{\left(\sum\limits_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}} \rightsquigarrow \mathcal{N}(0,1) \text{ and } \sum_{i=1}^{n}\varepsilon_i^2 \rightsquigarrow \chi_n^2 \tag{17}$$

Thus, if we denote by Φ the cumulative distribution function of the standard normal random variable and F the one of χ^2_{16}, then from the inequality (1) and according to (17) and also to the identity

$$1-\Phi(x)=\Phi(-x)$$

we obtain

$$\mathbb{P}\left\{\frac{\left(\sum\limits_{i=1}^{n} x_i^2\left|\widehat{a_n}-a\right|^2+n\left|\widehat{b_n}-b\right|^2\right)^{\frac{1}{2}}}{\left(\frac{1}{n}\sum\limits_{i=1}^{n}\left(y_i-\widehat{a_n}x_i-\widehat{b_n}\right)^2\right)^{\frac{1}{2}}}>r\right\}\le 4\Phi\left(-\sqrt{n\xi}\right)+F(n\varepsilon), \xi=\frac{\rho\varepsilon}{2(1+\rho)}$$

For $\rho = 1$ (i.e. $\xi = 0.025$) and $n = 16$

$$4\Phi\left(-\sqrt{16*0.025}\right)+F(16*0.1)=\mathbf{1.05418}>\mathbf{1}$$

In this case, we also fail to build confidence ellipses, but other choices of ρ will give results (see (18)).

5.2.1. Simulation results: Gaussian case (n=16)

Year	Available income	Centered income R_t	Theoretical consumption	Generation of $\mathcal{N}(0,1)$	Observed Consumption
2004	74	−61.19	1.05	0,768741345	1,818741345
2005	88	−47.19	12.25	−0,397436753	11,85256325
2006	100	−35.18	21.85	0,820283334	22,67028333
2007	112	−23.19	31.45	−1,530953749	29,91904625
2008	127	−8.19	43.45	0,123859096	43,5738591
2009	133	−2.19	48.25	0,168483654	48,41848365
2010	132	−3.19	47.45	−0,84723396	46,60276604
2011	139	3.81	53.05	0,266346526	53,31634653
2012	144	8.81	57.05	−0,706227055	56,34377295
2013	167	31.81	75.45	−0,961058755	74,48894125
2014	169	33.81	77.05	0,406303922	77,45630392
2015	167	31.81	75.45	−1,544547995	73,90545201
2016	156	20.81	66.65	−0,269142373	66,38085763
2017	150	14.81	61.85	−1,442049114	60,40795089
2018	151	15.81	62.65	0,17407425	62,82407425
2019	154	18.81	65.05	0,512285927	65,56228593

Table 3: Calculation of the observed consumption with $\mathcal{N}(0.1)$

From Table 3 we get : $\widehat{a_n} = 0,79, \widehat{b_n} = 49.72, S_n^2 = 0.555$. Thus, with a probability greater than or equal to $1 - 4\Phi\left(-4\sqrt{\xi}\right)$, the parameters a and b belong to the ellipses of equation:

$$\frac{(0.79 - a)^2}{\frac{8.88\rho}{12344.4}} + \frac{(49.72 - b)^2}{0.555\rho} = 1.$$

Here are some numerical results

ρ	Ellipse	Confidence level
4 (i.e. $\xi = 0.04$)	$\frac{(0.79-a)^2}{2.8774\times10^{-3}} + \frac{(49.72-b)^2}{2.22}$	15.26%
3 (i.e. $\xi = 0.0375$)	$\frac{(0.79-a)^2}{2.1581\times10^{-3}} + \frac{(49.72-b)^2}{1.665}$	12.28%
2 $\left(\text{i.e. } \xi = \frac{1}{30}\right)$	$\frac{(0.79-a)^2}{1.4387\times10^{-3}} + \frac{(49.72-b)^2}{1.11}$	6.96%

(18)

Table 4: Ellipses and confidence levels $(n = 16)$

Remark 5.2. *For α and ρ given, for example $\rho = 1$ (i.e. $\xi = 0.025$), we determine the minimum sample size that allows us to reach the given confidence level by solving the inequality*

$$4\Phi\left(-\sqrt{0.025n}\right) + F(0.1n) \leq \alpha.$$

The different results are given in the table below

α	0.01	0.05	0.1
n	316	201	19

.

5.3. Monte Carlo simulation study

The sample size ($n = 16$) is insufficient to construct confidence ellipses even in the case of Gaussian random variables i.i.d $\mathcal{N}(0,1)$. To overcome this shortcoming, we use the Monte Carlo method to generate large sample sizes ($n = 2\times10^3, n = 5\times10^3$, $n = 7\times10^3, n = 9\times10^3, n = 10^4$ and $n = 5\times10^4$). To perform this, we use the data of the first column of Table 1.

A preliminary analysis of the data using Cullen and Frey method (see Graph 1) allows us to assume that the probability distribution of R_t is of Weibull with two parameters. Then, we use a cdf plot (function cdfcomp), a density plot (function denscomp), a density Q-Q plot (function qqcomp), and a P-P plot (function ppcomp) (see Graph 2) and according to this graph, it is clear that our data are almost Weibull.

Finally, to confirm our assertion we will use the Kolmogorov-Smirnov goodness of fit test. After calculation we found p-value = 0.9427 greater than 0.05 which means it is not statistically significant and indicates strong evidence for the null hypothesis. This means we retain the null hypothesis and reject the alternative hypothesis.

1. If we note the distribution function of the Weibull distribution by

$$F(x) = \begin{cases} 1 - \exp\left(-\left(\dfrac{x}{sigma}\right)^{beta}\right) & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x \leq 0 \end{cases}$$

then using the Kolomogorov-Smirnov test we determine the distribution of the sample given in column 1, Table 1

$$R_t \rightsquigarrow \mathcal{W}(sigma, beta)$$

2. We generate by the Monte Carlo simulation n variables ($n = 2\times10^3, n = 5\times10^3, n = 7\times10^3, n = 9\times10^3$, $n = 10^4$ and $n = 5\times10^4$) from the Weibull distribution $\mathcal{W}(sigma, beta)$ using the inverting cumulative function

$$\widetilde{R}_t = sigma\left(\ln\frac{1}{1-U}\right)^{1/beta}$$

where U has uniform distribution over the interval $(0,1)$.

3. We determine

$$\begin{aligned} M &= \frac{1}{n}\sum_{t=1}^{n}\left(R_t-\overline{R}\right)^2 \\ N &= \max_t\left(R_t-\overline{R}\right)^2 \\ D &= \max_r\left|R_t-\overline{R}\right| \end{aligned}$$

4. We determine n variables ε_t

$$\varepsilon_t = 0.04\varepsilon_{t-1}+\nu_t \text{ with } \varepsilon_0 = 0, v_t \text{ i.i.d } \mathcal{U}(-1,1)$$

5. We determine n variables C_t

$$C_t = 50+0.8\left(R_t-\overline{R}\right)+\varepsilon_t$$

and n variables y_t

$$y_t = 50+0.8\left(R_t-\overline{R}\right)+\widetilde{\varepsilon}_t, \widetilde{\varepsilon}_t \text{ i.i.d } \mathcal{N}(0,1)$$

6. We determine $\widehat{a}_n, \widehat{b}_n, S_n^2$ in the α-mixing case and $\widehat{\widehat{a}}_n, \widehat{\widehat{b}}_n, \widehat{S}_n^2$ in the i.i.d $\mathcal{N}(0,1)$ case.

5.3.1. α-mixing Monte Carlo simulation

From (1) and various results given in the Appendix, we obtain

n	2×10^3	5×10^3	7×10^3	9×10^3	10^4	5×10^4
A_1	1	1	1	1	1	1
A_2	0.004222	0.0038775	0.004129	0.0038006	0.0037207	0.0033788
B_2	0.034489	0.034489	0.034489	0.034489	0.034489	0.034489
C_2	0.000751	0.000751	0.000751	0.000751	0.000751	0.000751
$\widehat{a}_n$	0.7991316	0.7999291	0.7996407	0.800236	0.799813	0.8001871
$\widehat{b}_n$	50.13501	50.01532	50.04262	49.97124	50.02336	49.97342
S_n^2	15593677	59455378	82701782	106701955	117746349	590781059

Table 5: Simulation results α-mixing case

Notation 5.3. *In order to simplify the writing, let us note by*

$$\begin{aligned} \mathbb{P}_1^{(n)} &= \exp\left(-\frac{M\xi n}{16e\left(\frac{1}{3}\max_i x_i^2+8\left(\frac{1}{3}\max_i|x_i|\right)^2\sum_{q=1}^{(n/2)-1}(0.04)^q\right)}\right) \\ \mathbb{P}_2^{(n)} &= \exp\left(-\frac{\xi n}{16e\left(\frac{1}{3}+8\sum_{q=1}^{(n/2)-1}(0.04)^q\right)}\right) \\ \mathbb{P}_3^{(n)} &= \exp\left(-\frac{\left(\frac{1}{3}-0.1\right)^2 n}{16e\left(\frac{1}{3}+32\sum_{q=1}^{(n/2)-1}(0.04)^q\right)}\right) \end{aligned}$$

1. **Determining the confidence level** $(1-\alpha)$ in %

For given n and ξ which is equivalent to n and ρ given by the relation

$$\rho = \frac{2\xi}{0.1-2\xi} \text{ with } 0<\xi<0.05. \tag{19}$$

we determine the different corresponding confidence levels by using the equality

$$1-\alpha = \left(1-\left(\mathbb{P}_1^{(n)}+\mathbb{P}_2^{(n)}+\mathbb{P}_2^{(n)}\right)\right)*100, n=2\times10^3,\cdots,n=5\times10^4$$

$\xi\backslash n$	2×10^3	5×10^3	7×10^3	9×10^3	10^4	5×10^4
0.049	8.2143%	61.277%	75.082%	81.173%	91.245%	99.975%
0.04	0.0642%	53.824%	67.86%	74.428%	91.245%	99.884%
0.03	X	43.507%	57.248%	64.037%	67.19%	99.371%
0.02	X	28.936%	42.454%	49.229%	52.33%	96.591%
0.01	X	X	15.552%	24.366%	27.837%	81.537%

Table 6: Confidence level in α-mixing case

X represents the case $\mathbb{P}_1^{(n)}+\mathbb{P}_2^{(n)}+\mathbb{P}_2^{(n)}>1$.

Remark 5.4. *It is obvious that for different values of ϕ, we associate different values of probabilities $\mathbb{P}_1^{(n)}, \mathbb{P}_2^{(n)}$ and $\mathbb{P}_3^{(n)}$ and different confidence levels. Some values are given in the following table when $n=2\times10^3$ and $\xi=0.049$*

ϕ	$\mathbb{P}_1^{(n=2000)}$	$\mathbb{P}_2^{(n=2000)}$	$\mathbb{P}_3^{(n=2000)}$	$\mathbb{P}_1^{(n=2000)}+\mathbb{P}_2^{(n=2000)}+\mathbb{P}_3^{(n=2000)}$ *Confidence level in %*
0.05	0.66821	0.050444	0.94101	1.6597 *(without interest)*
0.04	**0.66116**	**0.03405**	**0.22265**	**0.91786** ; (8.214%)
0.03	0.15072	0.020653	0.65396	0.82533 ; (17.467%)
0.02	0.07901	0.01070	0.64662	0.73633 ; (26.367%)
0.01	0.02208	0.004336	0.63911	0.66553 ; (33.447%)

Table 7: Incidence of ϕ

Determining the minimum sample size
Sample sizes are determined by solving the equations

$$\mathbb{P}_1^{(n)}+\mathbb{P}_2^{(n)}+\mathbb{P}_2^{(n)}=\alpha; n=2\times10^3,\cdots,n=5\times10^4$$

for given α and ξ. This gives the following results

$\xi\backslash\alpha$	0.01	0.05	0.1
0.049	22261	14483	11142
0.04	27270	17740	13638
0.03	36360	23653	18180
0.02	54539	35479	27270
0.01	109091	70956	54539

Table 8: Sample sizes in α-mixing case

Determination of ellipse foci
Since $M>1$ (see Appendix), the foci F_1 and F_2 of the confidence ellipses are given by

$$F_1 = \left(\widehat{a_n}, \widehat{b_n} - \sqrt{\rho S_n^2 \left(1 - \frac{1}{M}\right)},\right), F_2 = \left(\widehat{a_n}, \widehat{b_n} + \sqrt{\rho S_n^2 \left(1 - \frac{1}{M}\right)}\right).$$

$\widehat{a_n}, \widehat{b_n}$ and S_n^2 are grouped in the Table 5 and ρ is determined from the equations

$$\left(\exp\left(-4.222 \times 10^{-3} * \xi * n\right) + \exp\left(-3.8775 \times 10^{-3} * \xi * n\right) + \exp\left(-7.5109 \times 10^{-4} n\right)\right) \leq \alpha$$

For given α and n, we find the following foci

$n \backslash \alpha$	0.01	0.05	0.1
2×10^3	X	X	X
5×10^3	X	$F_1 = (0.7999, -6857)$ $F_2 = (0.7999, 6957)$	$F_1 = (0.7999, -5437)$ $F_2 = (0.7999, 5537)$
7×10^3	$F_1 = (0.7996, -7884)$ $F_2 = (0.7996, 7984)$	$F_1 = (0.7996, -5506)$ $F_2 = (0.7996, 5606)$	$F_1 = (0.7996, -4752)$ $F_2 = (0.7996, 4852)$
9×10^3	$F_1 = (0.8, -6937)$ $F_2 = (0.8, 7037)$	$F_1 = (0.8, -5350)$ $F_2 = (0.8, 5450)$	$F_1 = (0.8, -4689)$ $F_2 = (0.8, 4789)$
10^4	$F_1 = (0.7998, -6748)$ $F_2 = (0.7998, 6848)$	$F_1 = (0.7998, -5285)$ $F_2 = (0.7998, 5385)$	$F_1 = (0.7998, -4648)$ $F_2 = (0.7998, 4748)$
5×10^4	$F_1 = (0.8, -6039)$ $F_2 = (0.8, 6139)$	$F_1 = (0.8, -4973)$ $F_2 = (0.8, 5073)$	$F_1 = (0.8, -4455)$ $F_2 = (0.8, 4555)$

Table 9: Ellipse foci in α-mixing case

X represents the case $\mathbb{P}_1^{(n)} + \mathbb{P}_2^{(n)} + \mathbb{P}_2^{(n)} > 1$.
Note that all these ellipses are vertical with center $\left(\widehat{a_n}, \widehat{b_n}\right)$.

5.4. *Gaussian Monte Carlo Simulation*

Determination of ellipse foci in the Gaussian case

From the appendix, we can read the values $\widehat{\widehat{a_n}}, \widehat{\widehat{b_n}}$ and $\widehat{S_n^2}$. They are noted in the Appendix as aa_n, bb_n and SS_n.

n	2×10^3	5×10^3	7×10^3	9×10^3	10^4	5×10^4
$\widehat{\widehat{a_n}}$	0.7978058	0.7997622	0.7994124	0.800094	0.7996825	0.7998411
$\widehat{\widehat{b_n}}$	50.01131	49.99761	49.98879	49.99492	49.99433	50.00023
$\widehat{S_n^2}$	1.010926	1.01104	0.9847255	0.979984	1.005974	1.000784

Table 10: Simulation results Gaussian case

For given α and n, we determine ξ from the equation

$$\begin{aligned} \mathbb{P}^{(n)} &= \alpha \Longleftrightarrow 4\Phi\left(-\sqrt{\xi n}\right) + F(0.1n) = \alpha \\ &\Longleftrightarrow 4\Phi\left(-\sqrt{\xi n}\right) = \alpha \text{ since } F(0.1n) = 0 \text{ for } n = 2 \times 10^3, \cdots, n = 5 \times 10^4 \end{aligned}$$

Finally, we determine ρ from the relation (19)

$n\backslash\alpha$	0.01	0.05	0.1
2×10^3	0.012270964	0.007789259	0.005945072
5×10^3	0.003161741	0.002013601	0.001538948
7×10^3	0.002256348	0.001437459	0.001098766
9×10^3	0.001754058	0.001117667	0.000854387
10^4	0.001578375	0.001005788	0.000768882
5×10^3	0.000315277	0.002013601	0.000153682

Table 11: Different values of ρ in Gaussian case

For given α and n, we find the following foci

$n\backslash\alpha$	0.01	0.05	0.1
2×10^3	$F_1=(0.7978,49.886)$ $F_2=(0.7978,50.109)$	$F_1=(0.7978,49.909)$ $F_2=(0.7978,50.086)$	$F_1=(0.7978,49.958)$ $F_2=(0.7978,50.037)$
5×10^3	$F_1=(0.7994,49.941)$ $F_2=(0.7994,50.054)$	$F_1=(0.7994,49.953)$ $F_2=(0.7994,50.043)$	$F_1=(0.7994,49.941)$ $F_2=(0.7994,50.054)$
7×10^3	$F_1=(0.7996,49.996)$ $F_2=(0.7996,50.090)$	$F_1=(0.7996,49.436)$ $F_2=(0.7996,50.649)$	$F_1=(0.7996,49.519)$ $F_2=(0.7996,50.567)$
9×10^3	$F_1=(0.8,49.930)$ $F_2=(0.8,50.013)$	$F_1=(0.8,49.938)$ $F_2=(0.8,50.004)$	$F_1=(0.8,49.942)$ $F_2=(0.8,50)$
10^4	$F_1=(0.7998,49.984)$ $F_2=(0.7998,50.063)$	$F_1=(0.7998,49.992)$ $F_2=(0.7998,50.055)$	$F_1=(0.7998,49.996)$ $F_2=(0.7998,50.051)$
5×10^4	$F_1=(0.8,49.953)$ $F_2=(0.8,49.989)$	$F_1=(0.8,49.953)$ $F_2=(0.8,49.989)$	$F_1=(0.8,49.959)$ $F_2=(0.8,49.984)$

Table 12: Ellipse foci in Gaussian case

Note that all these ellipses are vertical with center $\left(\widehat{\widehat{a}_n}, \widehat{\widehat{b}_n}\right)$.

6. Appendix

Monte Carlo simulation study
```
# Import packages
library(fitdistrplus)
## Warning: package ’fitdistrplus’ was built under R version 4.0.4
## Loading required package: MASS
## Loading required package: survival
library(logspline)
# Create the dataset
df ¡- as.data.frame(data¡ -c(74,88,100,112,127,133,132,139,144,167,169
,167,156,150,151,154))
# Preliminary analysis of the data using Cullen and Frey
descdist(df$data, discrete = FALSE)
## summary statistics
## ——
## min: 74 max: 169
## median: 141.5
## mean: 135.1875
## estimated sd: 28.68732
## estimated skewness: -0.8523543
```

```
## estimated kurtosis: 2.908225
# Parametre estimation
fit.weibull$estimate
## shape scale
## 6.288061 145.981173
plot(fit.weibull)
# Kolmogorov Smirnov goodness of fit test
ks.test(df$data, ”pweibull”, shape = fit.weibull$estimate[1],
scale = fit.weibull$estimate[2])
## Warning in ks.test(df$data, ”pweibull”, shape = fit.weibull$estimate[1], : ties
## should not be present for the Kolmogorov-Smirnov test
##
## One-sample Kolmogorov-Smirnov test
##
## data: df$data
## D = 0.13213, p-value = 0.9427
## alternative hypothesis: two-sided
# Monte Carlo simulation for weibull distribition for n = 2000
sigma¡- 145.981173
beta ¡- 6.288061
n_2000
u_2000
R_t_2000
fit.weibull_2000$estimate
## shape scale
## 6.491915 147.065585
## Calculate of M, N, D
3)ˆ2))
M
## [1] 593.6162
N
## [1] 8728.413
D
D
## [1] 93.42597
## Simulation of v_t
max = 1)
## epsilon_t
S2_n_2000
## [1] 15593677
mean(R_t_2000)))/sum((R_t_2000 - mean(R_t_2000))ˆ2)
aa_n_2000
## [1] 0.7978058
bb_n_2000
## [1] 50.01131
SS2_n_2000 ¡- mean((y_t_2000 - aa_n_2000*(R_t_2000 - mean(R_t_2000))-bb_n_2000)ˆ2)
SS2_n_2000
## [1] 1.010926
# Monte Carlo simulation for weibull distribition for n = 5000
fit.weibull_5000$estimate
## shape scale
## 6.450028 146.362032
```

```
## Calculate of M, N, D
3)ˆ2))
M
## [1] 611.2254
N
## [1] 9785.942
D
## [1] 98.92392
## Simulation of v_t
## epsilon_t
## Determine C_t
0.04*epsilon_t_5000[i]+v_t_5000[i]}
## Detremine a_n, b_n, s2_n
mean(C_t_5000)))/sum((R_t_5000-mean(R_t_5000))ˆ2)
a_n_5000
## [1] 0.7999291
b_n_5000
## [1] 50.01532
a_n_5000*(R_t_5000 - mean(R_t_5000)) - b_n_5000)ˆ2
S2_n_5000
## [1] 59455378
mean(R_t_5000)))/sum((R_t_5000 - mean(R_t_5000))ˆ2)
aa_n_5000
## [1] 0.7997622
bb_n_5000
## [1] 49.99761
mean(R_t_5000))-bb_n_5000)ˆ2)
SS2_n_5000
## [1] 1.01104
# Monte Carlo simulation for weibull distribition for n = 7000
fit.weibull_7000$estimate
## shape scale
## 6.263327 146.161660
## Calculate of M, N, D
3)ˆ2))
M
## [1] 639.5926
N
## [1] 9653.994
D ¡- max(abs(R_t_7000-mean(R_t_7000)))
D
## [1] 98.25474
## Simulation of v_t
v_t_7000
max = 1)
## epsilon_t
epsilon_t_7000
epsilon_t_7000[1]
for (i in 2:7001) {
epsilon_t_7000[i]
}
y_t_7000 xi_t_7000
```

```
## Determine C_t
C_t_7000
for (i in 1:7000) {
C_t_7000[i] 0.04*epsilon_t_7000[i]+v_t_7000[i]
}
## Detremine a_n, b_n, s2_n
a_n_7000 mean(C_t_7000)))/sum((R_t_7000-mean(R_t_7000))^2)
a_n_7000
## [1] 0.7996407
b_n_7000
b_n_7000
## [1] 50.04262
S2_n_7000 a_n_7000*(R_t_7000 - mean(R_t_7000)) - b_n_7000)^2
S2_n_7000
## [1] 82701782
aa_n_7000 mean(R_t_7000)))/sum((R_t_7000 - mean(R_t_7000))^2)
aa_n_7000
## [1] 0.7994124
bb_n_7000
bb_n_7000
## [1] 49.98879
SS2_n_7000 mean(R_t_7000))-bb_n_7000)^2)
SS2_n_7000
## [1] 0.9847255
# Monte Carlo simulation for weibull distribition for n = 9000
sigma
beta
n_9000
u_9000
R_t_9000
fit.weibull_9000
xi_t_9000
fit.weibull_9000$estimate
## shape scale
## 6.350446 146.224440
## Calculate of M, N, D
M 3)^2))
M
## [1] 628.1908
N
N
## [1] 10261
D
D
## [1] 101.2966
## Simulation of v_t
v_t_9000
max = 1)
## epsilon_t
epsilon_t_9000
epsilon_t_9000[1]
for (i in 2:9001) {
```

```
epsilon_t_9000[i]
}
y_t_9000 xi_t_9000
## Determine C_t
C_t_9000
for (i in 1:9000) {
C_t_9000[i] 0.04*epsilon_t_9000[i]+v_t_9000[i]
}
## Detremine a_n, b_n, s2_n
a_n_9000 mean(C_t_9000)))/sum((R_t_9000-mean(R_t_9000))^2)
a_n_9000
## [1] 0.800236
b_n_9000
b_n_9000
## [1] 49.97124
S2_n_9000 a_n_9000*(R_t_9000 - mean(R_t_9000)) - b_n_9000)^2
S2_n_9000
## [1] 106701955
aa_n_9000 mean(R_t_9000)))/sum((R_t_9000 - mean(R_t_9000))^2)
aa_n_9000
## [1] 0.800094
bb_n_9000
bb_n_9000
## [1] 49.99492
SS2_n_9000 mean(R_t_9000))-bb_n_9000)^2)
SS2_n_9000
## [1] 0.979984
# Monte Carlo simulation for weibull distribition for n = 10000
sigma
beta
n_10000
u_10000
R_t_10000
fit.weibull_10000
xi_t_10000
fit.weibull_10000$estimate
## shape scale
## 6.251053 145.916038
## Calculate of M, N, D
M digits = 3)^2))
M
## [1] 643.3027
N
N
## [1] 10733.47
D
D
## [1] 103.6025
## Simulation of v_t
v_t_10000
max = 1)
## epsilon_t
```

```
epsilon_t_10000
epsilon_t_10000[1]
for (i in 2:10001) {
epsilon_t_10000[i] 0.04*epsilon_t_10000[i-1]+v_t_10000[i]
}
y_t_10000 xi_t_10000
## Determine C_t
C_t_10000
for (i in 1:10000) {
C_t_10000[i] 0.04*epsilon_t_10000[i]+v_t_10000[i]
}
## Detremine a_n, b_n, s2_n
a_n_10000 mean(R_t_10000))*(C_t_10000 - mean(C_t_10000)))/sum((R_t_10000-mean(R_t_10000))^2)
a_n_10000
## [1] 0.799813
b_n_10000
b_n_10000
## [1] 50.02336
S2_n_10000 a_n_10000*(R_t_10000 - mean(R_t_10000)) - b_n_10000)^2
S2_n_10000
## [1] 117746349
aa_n_10000 mean(R_t_10000)))/sum((R_t_10000 - mean(R_t_10000))^2)
aa_n_10000
## [1] 0.7996825
bb_n_10000
bb_n_10000
## [1] 49.99433
SS2_n_10000 - mean(R_t_10000))-bb_n_10000)^2)
SS2_n_10000
## [1] 1.005974
# Monte Carlo simulation for weibull distribition for n = 50000
sigma
beta
n_50000
u_50000
R_t_50000
fit.weibull_50000
xi_t_50000
fit.weibull_50000$estimate
## shape scale
## 6.295943 146.041486
## Calculate of M, N, D
M digits = 3)^2))
M
## [1] 634.5233
N
N
## [1] 11658.19
D
D
## [1] 107.9731
## Simulation of v_t
```

```
v_t_50000
max = 1)
## epsilon_t
epsilon_t_50000
epsilon_t_50000[1]
for (i in 2:50001) {
epsilon_t_50000[i] 0.04*epsilon_t_50000[i-1]+v_t_50000[i]
}
y_t_50000 xi_t_50000
## Determine C_t
C_t_50000
for (i in 1:50000) {
C_t_50000[i] 0.04*epsilon_t_50000[i]+v_t_50000[i]
}
## Detremine a_n, b_n, s2_n
a_n_50000 mean(R_t_50000))*(C_t_50000 - mean(C_t_50000)))/sum((R_t_50000-mean(R_t_50000))^2)
a_n_50000
## [1] 0.8001871
b_n_50000
b_n_50000
## [1] 49.97342
S2_n_50000 a_n_50000*(R_t_50000 - mean(R_t_50000)) - b_n_50000)^2
S2_n_50000
## [1] 590781059
aa_n_50000 mean(R_t_50000)))/sum((R_t_50000 - mean(R_t_50000))^2)
aa_n_50000
## [1] 0.7998411
bb_n_50000 ¡- mean(y_t_50000)
bb_n_50000
## [1] 50.00023
SS2_n_50000 ¡- mean((y_t_50000 - aa_n_50000*(R_t_50000 - mean(R_t_50000))-bb_n_50000)^2)
SS2_n_50000
## [1] 1.000784
```

Acknowledgments
The authors would like to thank the reviewer for his careful reading and valuable comments which improved presentation of the manuscript.

References

[1] N. Aiane, A. Dahmani, On the rate of convergence of the Robbins-Monro's algorithm in a linear stochastic ill-posed problem with $\alpha-$mixing data, Communications in Statistics-Theory and Methods, 46(13) (2017) 6694-6703.
[2] D.W.K. Andrews, Non-strong mixing autoregressive processes, J. Appl. Probab., 21 (1984) 930-934.
[3] H. Arroudj , I. Arab, A. Dahmani. Strongly mixed random errors in Mann's iteration algorithm for contractive real function, Neural, Parallel, and Scientific Computations, 25 (2017) 251-262.
[4] K.B. Athreya, S.G. Pantula, A note on strong mixing of ARMA processes, Statistics & Probability Letters, 4 (1986) 187-190.
[5] G.J. Babu, Z.D. Bai, Edgeworth expansions for errors-in-variables models, J. Multivariate Anal. 42 (2) (1992) 226-244
[6] J.R. Blum, D.L. Hanson, L.H. Koopmans, On the strong law of large numbers for a class of stochastic process, Z. Wahrsch. Verw. Geb. B 2 (1) (1963) 1-11.
[7] G.G. Boente, R. Fraiman, Robust nonparametric regression estimation for dependent observations, The annals of statistics, 17 (3) (1989) 1242-1256.
[8] D. Bosq, J.P. Lecoutre, Théorie de l'estimation fonctionnelle, Economica, 1987.
[9] E. Caron, J. Dedecker, B. Michel, Linear regression with stationary errors: the R package slm. 2019. hal-02157155v3
[10] R.A.L. Carter, M.S. Srivastava, V.K. Srivastava, A. Ullah, Unbiased estimation of the MSE matrix of Stein-rule estimators, confidence ellipsoids, and hypothesis testing, Econometric Theory, 6 (1990) 63-74.

[11] G.X. Chai, Z.Y. Li, H. Tian, Consistent nonparametric estimation of error distributions linear model. Acta Math. Apll. Sinica. 7 (3) (1991) 245-256.
[12] A. Chaturvedia, S. Guptaa, M. I. Bhattib, Confidence ellipsoids based on a general family of shrinkage estimators for a linear model with non-spherical disturbances, Journal of Multivariate Analysis, 104 (1) (2012) 140-158.
[13] X.R. Chen, Y. H. Wu, Strong law for mixing sequence, Acta Mathematicae Applicatae Sinica, 5 (1989) 367-371.
[14] C.M. Deo, A note on empirical processes of strong-mixing sequences, Ann. Probab., 1 (5) (1973) 870-875.
[15] A. Ya. Dorogovtsev, Consistency of the least squares estimator of an infinite-dimensional parameter, Sibirsk. Mat. Zh. 33 (4) (1992) 65-69. Translation in Siberian Math J. 4(1992) 603-607.
[16] P. Doukhan, Mixing : Properties and Examples, Lecture Notes in Statist. N°85, Springer Verlag, New-York, 1994.
[17] D. Durand, Joint Confidence Regions for Multiple Regression Coefficients, Journal of the American Statistical Association, 49 (265) (1954) 130-146.
[18] J. Esary, F. Proschan, D. Walkup, Association of random variables with application, Ann. Math. Statist. 38 (1967) 1466-1476.
[19] F. Ferraty, P. Vieu, Nonparametric Functional Data Analysis, Springer, 2006.
[20] R.A. Fisher, On the mathematical foundations of theoretical Statistics, Phil. Trans. R. Soc. Lond. A 222 (1922) 309-368.
[21] C. Fortuin, P. Kastelyn, J. Ginibre, Correlation inequalities on some ordered sets, Comm. Math. Phys.22 (1971) 89-103.
[22] F. Galton, Regression towards mediocrity in hereditary stature. Journal of Anrthropological Institute, 15 (1886) 246-263.
[23] H. Guo, Y. Liu, Regression estimation under strong mixing data, Annals of the Institute of Statistical Mathematics, Springer, 71(3) (2019) 553-576.
[24] T. Haavelmo, Methods of Measuring the Marginal Propensity to Consume, Journal of the American Statistical Associatio, 42 (237) (1947) 105-122.
[25] C. Hipp, Convergence rates of the strong law for stationary mixing sequences, Z. Wahrsch. Verw. Geb 49 (1) (1979) 49-62.
[26] I.A. Ibragimov, Yu. A. Linnik, Independent and stationary sequences of variables random. Groningen, The Netherlands: Wolters-Noordhoff, 1971.
[27] A.N. Kolmogorov, Yu.A. Rozanov, On Strong Mixing Conditions for Stationary Gaussian Processes. Theory Probab. Appl., 5 (2) (1960) 204-208.
[28] C. Lee, C. Adkins, R. Carter Hill, An improved confidence ellipsoid for the linear regression model, Journal of Statistical Computation and Simulation, 36 (1) (1990) 9-18.
[29] S. Lee, A note on the strong mixing property for random coefficient autoregressive process, Journal of Korean Statistical Society, 24 (1) (1995) 243-248.
[30] E.L. Lehmann, Some concepts of dependence, Ann. Math. Statist. 37 (1966) 1137-1153.
[31] KH. Liland, A. Smilde, F. Marini, T. Næs, Confidence ellipsoids for ASCA models based on multivariate regression theory, Journal of Chemometrics. 2018;e2990. https://doi.org/10.1002/cem.2990
[32] D.J. Olive, Linear Regression, Springer, 2017.
[33] K. Pearson, A. Lee, On the laws of inheritance in man. Inheritance of physical characters, Biometrika, 2 (4) (1903) 357-462.
[34] W. Philipp, A functional law of the iterated logarithm for empirical distribution functions of weakly dependent random variables, Ann. Probab., 5 (1977) 319-350.
[35] E. Rio, Théorie asymptotique des processus faiblement dépendants, Mathématiques et applications, 31, Springer, SMAI.1999.
[36] S.N. Roy, R.C. Bose, Simultaneous Confidence Interval Estimation, Annals of Mathematical Statistics, 24 (1953), 513-536.
[37] M.A. Rosenblatt, A central limit theorem and a strong mixing condition, Proc. Nat. Acad. Sci. USA. 42 (1956.) 43-47.
[38] M.A. Tingley, Small-sample intervals for regression. Canad. J. Statist. 20 (3) (1992) 271-280.
[39] V.A. Volkonskii, Yu.A. Rozanov, Some limit theorems for random functions. I, Theory of Probability and its Applications, 4 (2) (1959) 178-197.
[40] R.E. Welsch, Limit Laws for Extreme Order Statistics Form Strong-Mixing Processes, Ann. Math. Statist., 43 (2) (1972).439-446.
[41] G.U. Yule, On the theory of correlation. Journal of the Royal Statistical Society, 60 (1897) 812-854.
[42] K. Zvara, Consistency of an estimate in linear regression with nonnegative errors, Kybernetika, Prague, 28 (2) (1992) 129-139.