

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Université A.Mira-BEJAIA  
Faculté des Sciences Exactes  
Département de Mathématiques

# THÈSE

Présentée par

**Abdelkader ELMOUMEN**

Pour l'obtention du grade de

**DOCTEUR EN SCIENCES**

Filière : Mathématiques

Option : Analyse et Probabilités

**Thème**

**Algorithmes itératifs pour la résolution de problèmes  
de calibration non linéaire**

Soutenue le : 29/11/2022

Devant le Jury composé de :

**Nom et Prénom**

**Grade**

M. F. BOUHMILA

M.C.A.

Université A. Mira – Bejaïa

Président

M. A. DAHMANI

Professeur

Université de Tamenghasset

Rapporteur

Mme M. OURBIH

Professeur

C. Universitaire de Tipaza

Examinatrice

Mme H. ZEROUATI

Professeur

Université A. Mira – Bejaïa

Examinatrice

Année Universitaire : 2021/2022

---

# *Remerciements*

Tout d'abord, je tiens à remercier particulièrement le professeur Abdelnasser DAHMANI pour avoir accepté de diriger ma thèse, pour avoir toujours été disponible pendant tout mon travail et pour n'avoir ménagé aucun effort pour sa réalisation.

J'adresse mes plus sincères remerciements aux membres du jury de cette thèse : pour avoir accepté de la lire et m'avoir fait bénéficier de leurs pertinentes remarques et suggestions.

Je voudrais remercier tous les enseignants du département Mathématiques/Informatique de l'Université de Tamanrasset. Ceux qui sont encore sur place : A. Abdelli, T. Azizi, O. Bahi, M.E. Bencheikh Le Hocine, K. Dahmani, A.C. Guidoum, M. Mebarki, N. Rahali, auxquels j'ajoute Noredine OUAHAB qui a fait sa mutation à l'université d'ADRAR.

Je remercie particulièrement Oussama BAHY qui m'a aidé dans toute la partie numérique de cette thèse.

Je tiens à remercier tous les membres du laboratoire de Mathématiques Appliquées de l'Université de Béjaia dont je fais partie comme thésard et avec qui je souhaite tisser des relations de coopération.

Enfin, je tiens à remercier mes amis enseignants non matheux et ATS de l'Université de Tamanrasset pour les temps de détente que nous avons pu partager à l'intérieur et à l'extérieur du campus.

---

## *Dédicaces*

Je dédie ce travail à mes parents pour tout ce qu'ils représentent pour moi.

A ma femme et à mes enfants.

---

# Table des matières

<b>Introduction</b>	<b>5</b>
<b>1 Approximation stochastique dans un problème de calibration non linéaire</b>	<b>6</b>
1.1 Introduction . . . . .	6
1.2 Position du problème . . . . .	7
1.3 Résultats . . . . .	10
<b>2 Algorithme de Robbins-Monro avec des erreurs aléatoires <math>\psi</math>-mélangeantes</b>	<b>17</b>
2.1 Introduction . . . . .	17
2.2 Aspect mathématique de l'approximation stochastique . . . . .	20
2.2.1 Forme générale des algorithmes . . . . .	20
2.3 Quelques résultats sur le $\psi$ -mélange . . . . .	22
2.4 Inégalités exponentielles . . . . .	27
2.4.1 Résultats préliminaires . . . . .	27
2.5 L'inégalité qui en résulte . . . . .	31
2.5.1 Intervalle de confiance . . . . .	36
2.5.2 Convergence presque complète . . . . .	38
2.5.3 Rate of convergence . . . . .	39
2.6 Etude numérique . . . . .	40
2.6.1 Introduction . . . . .	40
2.6.2 Motivation . . . . .	41
2.6.3 Application . . . . .	42
2.6.4 Présentation de la simulation . . . . .	42
2.6.5 Exemples et résultats . . . . .	43
2.6.6 Conclusion . . . . .	44

<b>3</b>	<b>Ellipses de confiance pour les paramètres d'une régression simple à erreurs fortement mélangées</b>	<b>45</b>
3.1	Introduction . . . . .	45
3.2	PRELIMINAIRES . . . . .	48
3.3	Résultats . . . . .	50
3.4	Ellipses de confiance . . . . .	58
3.5	Simulations . . . . .	60
3.5.1	Simulation avec un processus $\alpha$ -mélangeant . . . . .	62
3.5.2	Simulation avec un processus gaussien . . . . .	64
3.5.3	Étude de simulation par la méthode Monte-Carlo . . . . .	66
3.5.4	Simulation Monte-Carlo cas : gaussien . . . . .	72
3.6	Annexe . . . . .	74

---

# Introduction

De nombreux problèmes théoriques et pratiques dans divers domaines peuvent être réduits à la recherche des zéros (racines) d'une fonction. C'est le cas par exemple dans l'identification de systèmes, les coefficients inconnus du système sont estimés sur la base des données d'entrée-sortie du système de contrôle ; dans les systèmes de contrôle adaptatifs, le gain de contrôle adaptatif doit être défini sur la base des données d'observation de telle sorte que le gain tende asymptotiquement vers l'optimal ; dans l'identification aléatoire de canaux, les coefficients du canal sont estimés en utilisant les données de sortie obtenues au niveau du récepteur ; dans le traitement du signal, la matrice de pondération optimale est estimée sur la base des observations ; dans la classification de formes, les paramètres spécifiant l'hyperplan de partition sont recherchés par apprentissage, et d'autres exemples peuvent être ajoutés à cette liste.

Pour s'en convaincre, il suffit de remarquer que la résolution de nombreux problèmes consiste finalement à optimiser une certaine fonction  $L$ , c'est-à-dire à trouver son minimum (ou son maximum). Si la fonction  $L$  est différentiable, alors le problème d'optimisation se réduit à trouver les racines de  $f$  où

$$f(x) = \frac{dL(x)}{dx}$$

est la dérivée de  $L$ .

Dans le cas où la fonction ou ses dérivées peuvent être observées sans erreur, il existe de nombreuses méthodes numériques pour résoudre ce type de problèmes. Dans les applications, on préfère généralement les algorithmes récursifs, en raison de leur relative simplicité de calcul et la possibilité d'utilisation de machines de calcul. Après chaque nouvelle observation, il n'est pas nécessaire de recalculer l'estimateur à partir de toutes les données recueillies. Chaque estimation successive est obtenue comme une simple fonction de la dernière estimation et de l'observation courante.

Parmi ces méthodes itératives, nous citons la méthode de Newton-Raphson. La méthode de Newton-Raphson est un procédé très efficace dans la recherche du zéro d'une fonction réelle, sa convergence est généralement bien plus rapide que celle de la méthode dichotomique. Elle repose sur la méthode du point fixe avec une fonction particulière  $h$  qui dépend de la dérivée de  $f$ .

$$h(x) = x - \frac{f(x)}{f'(x)}$$

On voit clairement que rechercher un point fixe de l'application  $h$  revient à chercher une solution de l'équation

$$f(x) = 0.$$

Le schéma numérique de la méthode de Newton-Raphson est donné par

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Ce type de problèmes fait partie de la théorie de l'optimisation. L'optimisation est omniprésente dans divers domaines de recherche et d'application. Il arrive très souvent qu'un problème d'optimisation puisse être réduit à la recherche des zéros (racines) d'une fonction inconnue qui peut être observée, mais l'observation peut être entachée par des erreurs.

Dans une telle situation, il est impossible, de manière générale, de construire des procédures convergeant vers la solution plus rapidement que  $\frac{1}{\sqrt{n}}$  où  $n$  est le nombre d'observations.

En 1951, Robbins et Monro ont proposé une méthode pour résoudre ce type de problème qu'ils ont appelé "Méthode d'Approximation Stochastique". Supposons, par exemple, que  $f$  est croissante et que l'équation

$$f(x) = 0$$

admet une solution unique  $x^*$ .

Soit  $y_n$  le résultat de mesure de la fonction  $f$  au point  $x_n$  à l'instant  $n$

$$y_n = f(x_n) + \varepsilon_n \tag{0.0.1}$$

où  $\varepsilon_n$  est une suite de variables aléatoires définies sur le même espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ .

Le paradigme de base de la procédure d'approximation stochastique est une équation de différence stochastique

$$x_{n+1} = x_n - a_n y_n. \quad (0.0.2)$$

où  $(a_n)_n$  est une suite de nombres réels positifs vérifiant

$$\sum_{n=1}^{\infty} a_n = +\infty \text{ et } \sum_{n=1}^{\infty} a_n^2 < \infty. \quad (0.0.3)$$

L'idée de la procédure (0.0.2) est claire. Si  $x_n < x^*$  alors la différence  $x_{n+1} - x_n$  sera positive en moyenne, puisque son espérance conditionnelle sachant  $x_n$  est donnée par

$$\mathbb{E}(x_{n+1} - x_n \mid x_n) = -a_n f(x_n)$$

Dans le cas contraire, cette différence est négative en moyenne. Ainsi, la procédure (0.0.2) "force" la séquence  $x_n$  à se s'approcher de  $x^*$ .

Cependant, il faut d'abord veiller à ce que les "sauts"  $x_{n+1} - x_n$  décroissent, car sinon la suite  $(x_n)_n$  ne peut pas converger vers  $x^*$ . Ceci est assuré par la première condition de (0.0.3). Par la suite, il faut veiller à ce que l'amplitude des sauts ne doit pas diminuer trop rapidement, car sinon la séquence  $(x_n)_n$  pourrait "ne pas avoir le temps" d'atteindre le point  $x^*$ . Cette propriété de la séquence  $(x_n)_n$  est garantie par la deuxième condition de (0.0.3).

Bien que l'hypothèse d'indépendance soit parfois raisonnable, il est difficile de vérifier l'indépendance des échantillons. De plus, dans de nombreux problèmes pratiques, les échantillons ne sont pas des observations indépendantes. Par exemple, s'il s'agit d'étudier l'évolution d'un système, l'hypothèse de dépendance est plus ajustée à la réalité. Cependant, il existe de nombreuses notions de dépendance. Nous nous sommes intéressés à celles qui s'expriment en termes de coefficients de mélange entre des tribus engendrées par le passé et le futur du processus. Voici les plus importants :

Soient  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé et  $\mathcal{A}$  et  $\mathfrak{B}$  deux sous-tribus de  $\mathcal{F}$ .

- Le coefficient de mélange fort introduit par Rosenblatt (1956)

$$\alpha(\mathcal{A}, \mathfrak{B}) = \sup_{A \in \mathcal{A}, B \in \mathfrak{B}} (|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|).$$

- Le coefficient d'absolue régularité (Rozañov et Volkonskii (1959))

$$\beta(\mathcal{A}, \mathfrak{B}) = \frac{1}{2} \sup \left( \sum_{i=1}^I \sum_{j=1}^J |\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)| \right)$$



où le supremum est considéré sur toutes les partitions mesurables  $(A_i)_{1 \leq i \leq I}$  et  $(B_j)_{1 \leq j \leq J}$  de  $\Omega$ .

- le coefficient de corrélation maximale  $\rho(\mathcal{A}, \mathfrak{B})$  de Kolmogorov et Rozanov (1960)

$$\rho(\mathcal{A}, \mathfrak{B}) = \sup_{X \in L_2(\mathcal{A}), Y \in L_2(\mathfrak{B})} (|\text{corr}(X, Y)|)$$

- Le  $\varphi$ -mélange [?] :

$$\varphi(\mathcal{A}, \mathfrak{B}) = \sup_{A \in \mathcal{A}, B \in \mathfrak{B}} (|\mathbb{P}(B | A) - \mathbb{P}(B)|)$$

- Le  $\psi$ -mélange introduit par Blum, Hanson, and Koopmans.

$$\psi(\mathcal{A}, \mathfrak{B}) = \sup_{A \in \mathcal{A}, B \in \mathfrak{B}} \left( \left| \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)\mathbb{P}(B)} - 1 \right| \right).$$

Sans conditions supplémentaires, voici les seules implications qui existent entre les différents type de mélangeance.

$$\psi - \text{mélange} \implies \varphi - \text{mélange} \implies \begin{cases} \beta - \text{mélange} \\ \Downarrow \quad \Uparrow \\ \rho - \text{mélange} \end{cases} \implies \alpha - \text{mélange}.$$

Il est à noter que pour les processus stochastiques la mélangeance signifie l'indépendance asymptotique et que la mélangeance implique l'ergodicité. Ainsi, tout ce que nous avons établi à propos des processus ergodiques, s'applique donc aux processus mélangeants.

Cette thèse comprend essentiellement trois parties. Dans la première partie on traite un problème de calibration non linéaire, appelé aussi la régression inverse, en utilisant la procédure d'approximations stochastique de Robbins-Monro. Il s'agit de résoudre une équation opératorielle où l'opérateur est non linéaire, borné et défini sur un espace de Hilbert séparable à valeurs dans lui même. Nous nous situons dans le cas où l'opérateur n'est connu qu'en certains points.

Dans la deuxième partie de cette thèse, nous construisons des inégalités exponentielles pour l'algorithme de Robins-Monro en considérant les erreurs de mesure  $\psi$ -mélangeantes. Ces inégalités permettent d'obtenir la convergence presque complète ainsi que de construire un domaine de confiance pour la racine de la fonction considérée.

Dans la troisième partie, nous établissons des inégalités exponentielles pour les paramètres d'un modèle de régression linéaire simple lorsque les erreurs de mesure sont  $\alpha$ -mélangeantes.

Ces inégalités nous permettent de construire des régions de confiance, caractérisées par des ellipses, pour les paramètres dudit modèle. Pour vérifier la validité des résultats théoriques obtenus, certains résultats numériques sont considérés.

Il est important de noter que tous les calculs numériques et toutes les simulations sont réalisés sous  $\mathcal{R}$

A la fin de la thèse, on trouve une conclusion et des perspectives.

---

# Approximation

## stochastique dans un problème de calibration non linéaire

### 1.1 Introduction

Durant la dernière décennie, la littérature à propos de la calibration s'est fort développée et les techniques ont été fortement affinées et étendues. Au début des années 1990 quelques travaux importants ont ouvert la voie. Model Assisted Survey Sampling de C.-E. Sarndal, B. Swensson et J.Wretman, paru en 1992, est sans aucun doute l'ouvrage de référence le plus important pour le statisticien d'enquête [64]. La même année, l'article Calibration Estimators in Survey Sampling de J.-C. Deville et C.-E. Sarndal a été publié dans JASA [24]. Cet article important a soudainement ouvert une classe plus large de techniques de calibration. De nombreux chercheurs, théoriciens et praticiens, partout dans le monde, dans les instituts de statistique publics et les universités se sont depuis lors inspirés de cet article ce qui a donné lieu à une longue série d'études sur la calibration.

Dans ce travail, nous proposons la procédure d'approximation stochastique de Robbins-Monro [60] pour résoudre un problème de calibration non linéaire. Le cas linéaire est étudié dans [18, 20].

Les algorithmes stochastiques ont prouvé leur capacité à résoudre des problèmes numériques de divers domaines de l'ingénierie : traitement signal [25, 27], problèmes inverses [19, ?]. La communication, l'identification du système [29, 30]. Par exemple, Ding [26] a proposé un algorithme d'identification itérative basé sur les moindres carrés hiérarchiques et un algorithme d'identification hiérarchique des moindres carrés généralisé. Wang et Tang dans [69] ont utilisé l'idée d'identification de modèle auxiliaire pour étudier l'algorithme d'estimation des moindres carrés récursifs pour une classe de systèmes de moyenne mobile d'erreur de sortie linéaire en paramètres et étudié dans [70] le problème d'identification itérative d'une classe de systèmes linéaires en paramètres en utilisant la technique de filtrage des données et la technique de décomposition. Récemment, des algorithmes d'estimation des paramètres des moindres carrés récursifs ont été développés pour les systèmes à bruit coloré en utilisant la technique de filtrage [28].

## 1.2 Position du problème

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé,  $\mathbb{H}$  un espace de Hilbert séparable et  $F : \mathbb{H} \rightarrow \mathbb{H}$  un opérateur non linéaire et borné. Nous supposons que nous disposons d'un échantillon

$\{(X_i, Y_i), i = 1, \dots, n\}$  de couples indépendants et ayant chacun même loi que  $(X, Y)$  qui est un couple aléatoire à valeurs dans  $(\mathbb{H} \times \mathbb{H})$ .

On s'intéresse à résoudre l'équation

$$F(x) = y. \quad (1.2.1)$$

Notons par  $y$  la valeur exacte inconnue du deuxième membre de l'équation (1.2.1) et par  $x^*$  l'élément de  $\mathbb{H}$  pour lequel l'existence et l'unicité sont connues et qui satisfait  $F(x^*) = y$  et  $\|x^*\| \leq N_1 < +\infty$ .

Cependant, dans certaines catégories de problèmes appliqués, nous trouvons une classe pour laquelle  $F$  n'est connue que pour certains points, de plus le deuxième membre n'est connu qu'approximativement. Une façon naturelle de gérer cette situation est de considérer le cadre probabiliste. Afin de résoudre ce type de problème, nous procédons en deux étapes. La première consiste à estimer l'opérateur  $F$ . Les estimations non paramétriques de l'opérateur  $F$  sont construites par des idées de pondération locale, comme par exemple l'estimation de noyau doublement fonctionnelle suivante :

$$\widehat{F}(x) = \frac{\sum_{i=1}^n K(h^{-1} \|X_i - x\|) Y_i}{\sum_{i=1}^n K(h^{-1} \|X_i - x\|)}, \forall x \in \mathbb{H},$$

où  $K$  est une fonction noyau et  $h = h_n$  est une suite de nombres réels positifs décroissant vers zéro [36]. Le noyau  $K$  est choisi de telle sorte que

$$\widehat{F}(x_1) - \widehat{F}(x_2) = c(x_1 - x_2). \quad (1.2.2)$$

Ainsi, le problème est réduit à résoudre l'équation suivante

$$\widehat{F}(x) = y. \quad (1.2.3)$$

Nous supposons que  $\widehat{x}$  est une solution de cette équation et que  $\|\widehat{x}\| \leq N_2 < +\infty$ .

**Définition 1.2.1** [45] *La séquence de bruit  $\xi_n : \Omega \rightarrow \mathbb{H}$  est dite satisfaire la condition de*

*Kushner-Clark (KC) si  $\forall \alpha > 0, \beta > 0$  et tous ensembles infinis d'intervalles disjoints  $I_k$ ,*

*on a*

$$\left\| \sum_{n \in I_k} \frac{a \xi_n}{n} \right\| < \alpha \sum_{n \in I_k} \frac{a}{n} + \beta \quad (1.2.4)$$

*sauf pour un nombre fini de  $I_k$ .*[45]

**Remarque 1.2.1** *Vérifier que certains bruits stochastiques satisfont à la condition (1.2.4)*

*peut être accompli par une simple application de l'inégalité de Markov et du lemme de*

*Borel-Cantelli [46].*

Supposons que  $F$  a un zéro unique en  $x^*$  et satisfait une condition de régularité appropriée.

$$\forall \delta > 0, \exists h_\delta > 0 : \|x - x^*\| \geq \delta \Rightarrow \langle F(x), x - x^* \rangle \geq (r + h_\delta) \|x - x^*\| \quad (1.2.5)$$

$$\limsup_{\varepsilon \rightarrow 0} \sup_{\|u\| < \varepsilon} \|F(x^* + u)\| \leq r$$

Pour résoudre (1.2.1), nous proposons la procédure de Robbins-Monro [60]

$$x_{n+1}^* = x_n^* - \frac{a}{n} [F(x_n^*) - y + \xi_n] \quad (1.2.6)$$

où  $x_n^*$  est une estimation de la solution  $x^*$ ,  $a$  est une constante positive et  $\xi_n$  représente le bruit de mesure de moyenne nulle qui satisfait la condition de Kushner-Clark (1.2.4) et la condition de Cramer

$$\forall m \in \mathbb{N}, m \geq 2, \mathbb{E} \|\xi_i\|^m \leq \frac{m!}{2} \mathbb{E} \|\xi_1\|^2 L_\xi^{m-2}, L_\xi > 0. \quad (1.2.7)$$

Pour résoudre (1.2.3), nous utilisons encore une fois la procédure de Robbins-Monro

$$\hat{x}_{n+1} = \hat{x}_n - \frac{a}{n} [\hat{F}(\hat{x}_n) - y + \eta_n] \quad (1.2.8)$$

où  $\hat{x}_n$  est une estimation de la solution  $\hat{x}$ ,  $a$  est une constante positive et  $\eta_n$  représente le bruit de mesure de moyenne nulle qui satisfait la condition de Kushner-Clark (1.2.4) et la condition de Cramer

$$\forall m \in \mathbb{N}, m \geq 2, \mathbb{E} \|\eta_i\|^m \leq \frac{m!}{2} \mathbb{E} \|\eta_1\|^2 L_\eta^{m-2}, L_\eta > 0. \quad (1.2.9)$$

## 1.3 Résultats

**Lemme 1.3.1** *Pour  $1 \leq i, j \leq n$ , Nous avons l'égalité suivante*

$$\begin{aligned} \hat{x}_{n+1} - x_{n+1}^* &= (\hat{x}_1 - x_1^*) \prod_{i=1}^n \left(1 - \frac{ac}{i}\right) - \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \hat{F}(x_i^*) - F(x_i^*) - \\ &\quad - \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) (\xi_i - \eta_i) \end{aligned}$$

$$\text{avec } \prod_{j=n+1}^n \left(1 - \frac{ac}{j}\right) = 1.$$

**Démonstration.** A partir de (1.2.6) et (1.2.8), il est facile de voir que

$$\begin{aligned} \widehat{x}_{n+1} - x_{n+1}^* &= \widehat{x}_n - x_n^* - \frac{a}{n} \left[ \widehat{F}(\widehat{x}_n) - F(x_n^*) \right] - \frac{a}{n} (\xi_n - \eta_n) \\ &= \widehat{x}_n - x_n^* - \frac{a}{n} \left[ \widehat{F}(\widehat{x}_n) - \widehat{F}(x_n^*) + \widehat{F}(x_n^*) - F(x_n^*) \right] - \frac{a}{n} (\xi_n - \eta_n) \\ &= (\widehat{x}_n - x_n^*) - \frac{ac}{n} (\widehat{x}_n - x_n^*) - \frac{a}{n} \left( \widehat{F}(x_n^*) - F(x_n^*) \right) - \frac{a}{n} (\xi_n - \eta_n). \end{aligned} \quad (1.3.1)$$

En itérant l'équation (1.3.1), on obtient

$$\begin{aligned} \widehat{x}_{n+1} - x_{n+1}^* &= (\widehat{x}_1 - x_1^*) \prod_{i=1}^n \left(1 - \frac{ac}{i}\right) - \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \left( \widehat{F}(x_i^*) - F(x_i^*) \right) \\ &\quad - \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) (\xi_i - \eta_i). \end{aligned} \quad (1.3.2)$$

C'est ce qu'il fallait démontrer. ■

**Lemme 1.3.2** Pour tous  $a > 0, c > 0$ , on a

$$\prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \leq \left(\frac{i+1}{n+1}\right)^{ac}. \quad (1.3.3)$$

**Démonstration.** On a

$$\ln \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) = \sum_{j=i+1}^n \ln \left(1 - \frac{ac}{j}\right).$$

A partir de l'inégalité  $\ln(1-x) \leq -x$ , on obtient

$$\ln \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \leq - \sum_{j=i+1}^n \frac{ac}{j} \leq -ac \int_{i+1}^{n+1} \frac{dx}{x} = \ln \left(\frac{i+1}{n+1}\right)^{ac}.$$

Le résultat découle immédiatement de l'injectivité de la fonction  $x \mapsto \ln x$ . ■



**Lemme 1.3.3** *Pour tous  $a > 0, c > 0$ , nous avons*

$$\sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \leq \frac{2}{c}.$$

**Démonstration.** En utilisant le Lemme 1.3.2, nous obtenons

$$\begin{aligned} \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) &\leq \sum_{i=1}^n \frac{a}{i} \left(\frac{i+1}{n+1}\right)^{ac} \leq \frac{2a}{(n+1)^{ac}} \sum_{i=2}^{n+1} \frac{1}{i^{1-ac}} \\ &\leq \frac{2a}{(n+1)^{ac}} \int_1^{n+1} \frac{dx}{x^{1-ac}} \\ &\leq \frac{2}{c}. \end{aligned}$$

d'où le résultat. ■

**Théorème 1.3.1** *Si  $\mathbb{H}$  est un espace de Hilbert et  $\mathbb{E}X_i = 0$  pour tous  $i$ , alors*

$$\mathbb{E} \cosh \left( t \left\| \sum_{i=1}^n X_i \right\| \right) \leq \prod_{i=1}^n \mathbb{E} \left( e^{t\|X_i\|} - t\|X_i\| \right).$$

**Démonstration.** Voir [57]. L'énoncé de ce théorème se trouve à la page 144 et sa

démonstration en page 147. ■

**Théorème 1.3.2** *Pour tout  $\varepsilon > 0$ , on a*

$$\mathbb{P} \left\{ \left\| \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) (\xi_i - \eta_i) \right\| > \varepsilon \right\} \leq \exp \left( \frac{-\varepsilon^2}{4(\mathbb{E}\|\xi_1\|^2 + \mathbb{E}\|\eta_1\|^2) \varphi(n)} \right). \quad (1.3.4)$$

**Démonstration.** Posons

$$\zeta_i = \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) (\xi_i - \eta_i).$$

Pour l'inégalité

$$\mathbb{P} \left\{ \left\| \sum_{i=1}^n \zeta_i \right\| > \varepsilon \right\} \leq \exp(-t\varepsilon) \cdot \mathbb{E} \exp \left( t \left\| \sum_{i=1}^n \zeta_i \right\| \right) \leq 2 \exp(-t\varepsilon) \cdot \mathbb{E} \cosh \left( t \left\| \sum_{i=1}^n \zeta_i \right\| \right).$$

En vertu du théorème 1.3.1, on

$$\mathbb{P} \left\{ \left\| \sum_{i=1}^n \zeta_i \right\| > \varepsilon \right\} \leq \exp(-t\varepsilon) \cdot \prod_{i=1}^n \mathbb{E} (e^{t\|\zeta_i\|} - t\|\zeta_i\|).$$

Le développement de la fonction exponentielle en zéro donne

$$\prod_{i=1}^n \mathbb{E} (e^{t\|\zeta_i\|} - t\|\zeta_i\|) = \prod_{i=1}^n \left( 1 + \sum_{m=2}^{+\infty} \frac{t^m \mathbb{E} \|\zeta_i\|^m}{m!} \right). \quad (1.3.5)$$

D'après le Lemme 1.3.2 et les inégalités (1.2.7) et (1.2.9), on a

$$\mathbb{E} \|\zeta_i\|^m \leq \frac{a^m}{i^m} \left( \frac{i+1}{n+1} \right)^{acm} \frac{m!}{2} (\mathbb{E} \|\xi_1\|^2 + \mathbb{E} \|\eta_1\|^2) L^{m-2} \quad (1.3.6)$$

où  $L$  est tel que

$$L_\xi^{m-2} + L_\eta^{m-2} \leq L^{m-2}.$$

En combinant les inégalités (1.3.5) et (1.3.6), on obtient

$$\prod_{i=1}^n \mathbb{E} (e^{t\|\zeta_i\|} - t\|\zeta_i\|) \leq \prod_{i=1}^n \left( 1 + \frac{t^2 a^2}{2 i^2} \left( \frac{i+1}{n+1} \right)^{2ac} (\mathbb{E} \|\xi_1\|^2 + \mathbb{E} \|\eta_1\|^2) \sum_{m=2}^{+\infty} \left( \frac{ta}{i} \left( \frac{i+1}{n+1} \right)^{ac} L \right)^{m-2} \right)$$

Pour  $t \leq \frac{1}{2aL}$ , on a

$$\prod_{i=1}^n \mathbb{E} (e^{t\|\zeta_i\|} - t\|\zeta_i\|) \leq \exp \left( t^2 (\mathbb{E} \|\xi_1\|^2 + \mathbb{E} \|\eta_1\|^2) \sum_{i=1}^n \frac{a^2}{i^2} \left( \frac{i+1}{n+1} \right)^{2ac} \right).$$

Notons

$$\varphi(n) = \sum_{i=1}^n \frac{a^2}{i^2} \left( \frac{i+1}{n+1} \right)^{2ac}$$

D'après le lemme de Kronecker, nous savons que  $\varphi(n)$  tend vers 0 lorsque  $n$  tend vers l'infini, par conséquent

$$\mathbb{P} \left\{ \left\| \sum_{i=1}^n \zeta_i \right\| > \varepsilon \right\} \leq \exp(-t\varepsilon + t^2 (\mathbb{E} \|\xi_1\|^2 + \mathbb{E} \|\eta_1\|^2) \varphi(n)). \quad (1.3.7)$$

Le deuxième terme de l'inégalité (1.3.7) prend sa valeur minimale pour

$$t_0 = \frac{\varepsilon}{2 (\mathbb{E} \|\xi_1\|^2 + \mathbb{E} \|\eta_1\|^2) \varphi(n)}.$$

En remplaçant  $t$  par  $t_0$  dans (1.3.7), on obtient

$$\mathbb{P} \left\{ \left\| \sum_{i=1}^n \zeta_i \right\| > \varepsilon \right\} \leq \exp \left( \frac{-\varepsilon^2}{4 (\mathbb{E} \|\xi_1\|^2 + \mathbb{E} \|\eta_1\|^2) \varphi(n)} \right).$$

d'où le résultat. ■

**Théorème 1.3.3** *Pour tout  $\varepsilon > 0$ , nous avons*

$$\mathbb{P} \left\{ \left\| \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) (\widehat{F}(x_i^*) - F(x_i^*)) \right\| > \varepsilon \right\} \leq \exp \left( \frac{-\varepsilon^2}{4C\varphi(n)} \right). \quad (1.3.8)$$

**Démonstration.** La démonstration de ce théorème est une répétition textuelle de la démonstration de Théorème 1.3.2. Il suffit d'appliquer l'inégalité

$$\left\| \widehat{F}(x_i^*) - F(x_i^*) \right\|^m \leq \frac{m!}{2} C^{m-2},$$

et nous obtenons le résultat souhaité. ■

**Théorème 1.3.4** *Les racines des équations (1.2.1) et (1.2.3) sont presque sûrement égales, i.e.*

$$\forall \varepsilon > 0, \mathbb{P}(\|\widehat{x} - x^*\| > \varepsilon) = 0. \quad (1.3.9)$$

**Démonstration.** Pour montrer cette égalité, nous remarquons d'abord que

$$\begin{aligned} \|\widehat{x} - x^*\| &= \|\widehat{x} - \widehat{x}_{n+1} + \widehat{x}_{n+1} - x_{n+1}^* + x_{n+1}^* - x^*\| \\ &\leq \|\widehat{x} - \widehat{x}_{n+1}\| + \|\widehat{x}_{n+1} - x_{n+1}^*\| + \|x_{n+1}^* - x^*\|. \end{aligned}$$

En vertu des propriétés de probabilité, on a pour tout  $\varepsilon$  positif

$$\begin{aligned} \mathbb{P}\{\|\widehat{x} - x^*\| > \varepsilon\} &\leq \mathbb{P}\{\|\widehat{x} - \widehat{x}_{n+1}\| + \|\widehat{x}_{n+1} - x_{n+1}^*\| + \|x_{n+1}^* - x^*\| > \varepsilon\} \\ &\leq \mathbb{P}\left\{\|\widehat{x} - \widehat{x}_{n+1}\| > \frac{\varepsilon}{3}\right\} + \mathbb{P}\left\{\|\widehat{x}_{n+1} - x_{n+1}^*\| > \frac{\varepsilon}{3}\right\} + \mathbb{P}\left\{\|x_{n+1}^* - x^*\| > \frac{\varepsilon}{3}\right\}. \end{aligned}$$

Le premier et le dernier terme du second membre de l'inégalité précédente sont inférieurs

à  $\frac{\varepsilon}{3}$  en vertu de la convergence des suites  $(x_n^*)$  et  $(\widehat{x}_n)$  respectivement, vers  $x^*$  et  $\widehat{x}$ .

En utilisant l'inégalité triangulaire, on obtient de (1.3.2) que

$$\begin{aligned} \|\widehat{x}_{n+1} - x_{n+1}^*\| &\leq \|\widehat{x}_1 - x_1^*\| \prod_{i=1}^n \left(1 - \frac{ac}{i}\right) + \left\| \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) (\widehat{F}(x_i^*) - F(x_i^*)) \right\| + \\ &\quad + \left\| \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) (\xi_i - \eta_i) \right\|. \end{aligned} \quad (1.3.10)$$

De (1.3.3) et pour  $n$  un assez grand, on a

$$\left(\|\widehat{x}_1 - x_1^*\| \prod_{j=1}^n \left(1 - \frac{ac}{j}\right)\right) \leq \left(\frac{1}{n+1}\right)^{ac} \leq \frac{\varepsilon}{6}. \quad (1.3.11)$$

En utilisant encore les propriétés des probabilités et l'inégalité (1.3.11), nous avons

$$\begin{aligned} \mathbb{P}\left\{\|\widehat{x}_{n+1} - x_{n+1}^*\| > \frac{\varepsilon}{3}\right\} &\leq \mathbb{P}\left\{\left\| \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) (\xi_i - \eta_i) \right\| > \frac{\varepsilon}{12}\right\} + \\ &+ \mathbb{P}\left\{\left\| \sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) (\widehat{F}(x_i^*) - F(x_i^*)) \right\| > \frac{\varepsilon}{12}\right\}. \end{aligned}$$

En combinant (1.3.4) et (1.3.8), on obtient

$$\mathbb{P} \left\{ \|\widehat{x}_{n+1} - x_{n+1}^*\| > \frac{\varepsilon}{3} \right\} \leq \exp \left( \frac{-\varepsilon^2}{576 (\mathbb{E} \|\xi_1\|^2 + \mathbb{E} \|\eta_1\|^2) \varphi(n)} \right) + \exp \left( \frac{-\varepsilon^2}{576M\varphi(n)} \right).$$

Puisque  $\varphi(n) \rightarrow 0$  comme  $n \rightarrow +\infty$ , on a

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left\{ \|\widehat{x}_{n+1} - x_{n+1}^*\| > \frac{\varepsilon}{3} \right\} = 0. \quad (1.3.12)$$

Enfin l'inégalité (1.3.9) est démontré. ■

---

# Algorithme de Robbins-Monro avec des erreurs aléatoires $\psi$ -mélangeantes

## 2.1 Introduction

Les algorithmes d'approximation stochastique sont des méthodes récursives, de mise à jour, qui peuvent être utilisés entre autres pour résoudre des problèmes d'optimisation ou des équations du point fixe, (y compris les systèmes linéaires standards), lorsque les données collectées sont soumises à un bruit. En ingénierie, les problèmes d'optimisation sont souvent de ce type, lorsqu'on ne dispose pas d'un modèle mathématique du système (ou qu'il est trop complexe) mais que l'on souhaite tout de même optimiser son comportement en ajustant certains paramètres.

A cet effet, on peut faire des expériences ou des simulations pour évaluer la performance du système à des valeurs données des paramètres. Les algorithmes d'approximation stochastique ont également été utilisés dans les sciences sociales pour décrire une dynamique

collective : jeu fictif dans la théorie de l'apprentissage et des algorithmes de consensus qui peuvent être étudiés en utilisant leur théorie.

Apparue au milieu du vingtième siècle, l'approximation stochastique s'est avérée être un beau mélange entre la théorie des systèmes dynamiques et celle des probabilités, essentiellement quand elle est jumelée aux techniques d'approximation des équations différentielles ordinaires. Cette théorie a en plus apporté des réponses importantes nées de l'approximation numérique, dont la plus importante est : "comment approximer la solution d'une équation liée à une fonction qui est le résultat d'une expérience mais dont on ne connaît pas l'expression exacte mais plutôt sa valeur "mesurée" en chaque point?".

Les problèmes rencontrés en sciences de l'ingénieur sont en général le fruit d'expériences réalisées dans des conditions non totalement maîtrisées, c'est à dire que les informations sont toujours bruitées et par conséquent non exactes. La modélisation de tels problèmes se fait alors avec la présence inévitable d'erreurs dans les observations de la fonction objectif en chaque instant. Ces erreurs étant aléatoires, il nous est impossible d'approximer la solution du problème, qui est en général une racine de la fonction objectif, avec les techniques d'approximations déterministes. On utilise pour cela les algorithmes d'approximation stochastique. La consistance de ces algorithmes dépend grandement de la nature des erreurs rencontrées, qui sont rarement indépendantes. Bondarev et Dahmani [10] ont

construit de telles inégalités lorsque les variables aléatoires sont  $\varphi$ -mélangeantes. En [1], de telles inégalités sont construites lorsque les variables aléatoires sont  $\alpha$ -mélangeantes. En [52], ces inégalités sont construites dans le cas indépendant. Le cas où les variables sont associées a été présenté par Arab [3]. D'autres inégalités exponentielles ont été construites selon le type des variables aléatoires considérées dans [11, 21, 22, 58]. Dans ce travail, nous construisons des inégalités exponentielles pour l'algorithme de Robins-Monro en considérant des variables aléatoires de la procédure d'approximation stochastique sont  $\psi$ -mélangeantes.

Les algorithmes de l'approximation stochastique se présentent en général sous cette forme.

Soit  $f(x)$  le résultat attendu au point  $x$ , d'une expérience donnée.  $f$  étant une fonction en  $x$  supposée monotone mais non connue, et on désire trouver la solution

$$x = \theta$$

de l'équation

$$f(x) = \alpha$$

où  $\alpha$  est une constante donnée.



En 1951, Herbert Robbins et Sutton Monro ont donné une méthode pour faire des expériences successives à des points  $x_1, x_2, \dots$  de telle sorte que la suite  $x_n$  tende vers  $\theta$  presque sûrement quand  $n$  tend vers l'infini. Ainsi naquit l'approximation stochastique.

## 2.2 Aspect mathématique de l'approximation stochastique

### 2.2.1 Forme générale des algorithmes

Les algorithmes de l'approximation stochastique se présentent en général sous la forme suivante :

$$x_{n+1} = x_n + a_n Y_n$$

où

$$Y_n = f(x_n) + \xi_n,$$

$(a_n)_{n \in \mathbb{N}}$  une suite telle que

$$\sum_{n=1}^{\infty} a_n = \infty \text{ et } \sum_{n=1}^{\infty} a_n^2 < \infty.$$

et  $\xi_n$  représente l'erreur de mesure commise en évaluant la fonction inconnue à travers l'expérience  $f$  au point  $x_n$ .

La grande majorité des résultats qu'on connaît en théorie des probabilités et qui sont relatifs aux variables aléatoires, ne sont généralement valables que pour des variables aléatoires indépendantes, et il est souvent difficile voire impossible de généraliser ces résultats

aux variables aléatoires dépendantes. Pour remédier à cette anomalie, des mesures de dépendance ont été créées pour savoir à quel point ces variables étaient dépendantes.

Parmi ces mesures précédemment citées, il y a les mélanges. On cite: le  $\alpha$ -mélange, le  $\beta$ -mélange, le  $\phi$ -mélange, le  $\psi$ -mélange, le  $\rho$ -mélange, le  $\psi^*$ -mélange, ...

**Définition 2.2.1** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé.

Soit  $(X_t)_{t \in \mathbb{N}} \subset (\Omega, \mathcal{F}, \mathbb{P})$  une suite de variables aléatoires et soit

$$F_j^l = \sigma(X_k, j \leq k \leq l) \text{ avec } (k \in \mathbb{Z}).$$

où

$$\sigma(X_k, j \leq k \leq l)$$

est la tribu engendrée  $X_k, j \leq k \leq l$ .

Le  $\psi$ -mélange

$$\psi(\mathcal{A}, \mathfrak{B}) = \sup_{\substack{A \in \mathcal{A}, B \in \mathfrak{B} \\ \mathbb{P}(A)\mathbb{P}(B) \neq 0}} \left( \left| 1 - \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)\mathbb{P}(B)} \right| \right),$$

appelé aussi  $*$ -mélange, a été introduit par Blum, Hanson et Koppmans (1963).

$(X_t)_{t \in \mathbb{N}^*}$  est dite  $\psi$ -mélangeante si

$$\lim_{n \rightarrow \infty} \psi(n) = \lim_{n \rightarrow \infty} \sup_{j \in \mathbb{Z}} \psi(F_{-\infty}^j, F_{j+n}^{+\infty}) = 0.$$

## 2.3 Quelques résultats sur le $\psi$ -mélange

**Lemme 2.3.1** Soient  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé et  $X, Y$  deux variables aléatoires indépendantes définies sur  $(\Omega, \mathcal{F}, \mathbb{P})$  à valeurs dans  $\mathbb{R}$ . Alors

$$\lim_{h \rightarrow 0} \sum_{k=-\infty}^{+\infty} \mathbb{P}(\{X > z - kh\} \cap \{kh \leq Y < (k+1)h\}) = \mathbb{P}\{X + Y > z\}.$$

**Démonstration.** Il est évident que

$$\bigcup_{k=-\infty}^{k=+\infty} (\{X > z - kh\} \cap \{kh \leq Y < (k+1)h\}) \subset \{X + Y > z\}.$$

Pour l'autre inclusion, il suffit de prendre

$$k_1 = \left\lfloor \frac{Y(\omega)}{h} \right\rfloor, h \in \mathbb{R}_+^*$$

où  $\lfloor x \rfloor$  est la partie entière de  $x$ .

Comme les événements

$$\{X > z - kh\} \cap \{kh \leq Y < (k+1)h\}, k \in \mathbb{Z}$$

sont deux à deux disjoints, alors

$$\begin{aligned} & \mathbb{P}\left(\bigcup_{k=-\infty}^{k=+\infty} (\{X > z - kh\} \cap \{kh \leq Y < (k+1)h\})\right) = \\ & = \sum_{k=-\infty}^{+\infty} \mathbb{P}(\{X > z - kh\} \cap \{kh \leq Y < (k+1)h\}) = \mathbb{P}\{X + Y > z\}. \end{aligned}$$

Par passage à la limite, on a

$$\lim_{h \rightarrow 0} \sum_{k=-\infty}^{+\infty} \mathbb{P}(\{X > z - kh\} \cap \{kh \leq Y < (k+1)h\}) = \mathbb{P}\{X + Y > z\}.$$

■

**Proposition 2.3.1** Soient  $n > 1$  et  $X_1, X_2, \dots, X_n$ ,  $n$  variables aléatoires  $\psi$ -mélangeantes.

Alors il existe  $n$  variables aléatoires indépendantes  $Y_1, Y_2, \dots, Y_n$  telles que

$$(1 - \psi)^{n-1} \mathbb{P} \left\{ \sum_{i=1}^n Y_i > x \right\} \leq \mathbb{P} \left\{ \sum_{i=1}^n X_i > x \right\} \leq (1 + \psi)^{n-1} \mathbb{P} \left\{ \sum_{i=1}^n Y_i > x \right\}. \quad (2.3.1)$$

**Démonstration.** Nous allons procéder par récurrence sur  $n$ . Par hypothèse de  $\psi$ -mélange, on a pour  $n = 2$

$$\begin{aligned} (1 - \psi) \mathbb{P} \{X_1 > x_1\} \mathbb{P} \{X_2 > x_2\} &\leq \mathbb{P} \{X_1 > x_1, X_2 > x_2\} \\ &\leq (1 + \psi) \mathbb{P} \{X_1 > x_1\} \mathbb{P} \{X_2 > x_2\}. \end{aligned} \quad (2.3.2)$$

Par ailleurs, pour tout  $h > 0$ , on a

$$\{X_1 + X_2 > x\} = \bigcup_{k=-\infty}^{k=+\infty} \{X_1 + X_2 > x\} \cap \{kh \leq X_2 < (k+1)h\}.$$

Aussi, pour tout  $k$ , on a

$$\begin{aligned} &\{X_1 > x - kh\} \cap \{kh \leq X_2 < (k+1)h\} \subset \\ &\subset \{X_1 + X_2 > x\} \cap \{kh \leq X_2 < (k+1)h\} \\ &\subset \{X_1 > x - (k+1)h\} \cap \{kh \leq X_2 < (k+1)h\}. \end{aligned}$$

Ce qui donne

$$\begin{aligned} &\bigcup_{k=-\infty}^{k=+\infty} \{X_1 > x - kh\} \cap \{kh \leq X_2 < (k+1)h\} \subset \{X_1 + X_2 > x\} \\ &\subset \bigcup_{k=-\infty}^{k=+\infty} \{X_1 > x - (k+1)h\} \cap \{kh \leq X_2 < (k+1)h\}. \end{aligned}$$

En vertu des propriétés de la probabilité, on obtient

$$\begin{aligned} & \sum_k \mathbb{P}(\{X_1 > x - kh\} \cap \{kh \leq X_2 < (k+1)h\}) \leq \mathbb{P}\{X_1 + X_2 > x\} \\ & \leq \sum_k \mathbb{P}(\{X_1 > x - (k+1)h\} \cap \{kh \leq X_2 < (k+1)h\}). \end{aligned}$$

De (2.3.2), on obtient

$$\begin{aligned} & \mathbb{P}(\{X_1 > x - (k+1)h\} \cap \{kh \leq X_2 < (k+1)h\}) \leq \\ & \leq (1 + \psi(\tau)) \mathbb{P}\{X_1 > x - (k+1)h\} \mathbb{P}\{kh \leq X_2 < (k+1)h\} \end{aligned}$$

Si on note

$$\mathbb{P}\{X_1 > x_1\} = \mathbb{P}\{Y_1 > x_1\} \text{ et } \mathbb{P}\{X_2 > x_2\} = \mathbb{P}\{Y_2 > x_2\}$$

avec  $Y_1$  et  $Y_2$  indépendantes, alors

$$\begin{aligned} & \sum_k \mathbb{P}(\{Y_1 > x - kh\} \cap \{kh \leq Y_2 < (k+1)h\}) \leq \mathbb{P}\{X_1 + X_2 > x\} \\ & \leq \sum_k \mathbb{P}(\{Y_1 > x - (k+1)h\} \cap \{kh \leq Y_2 < (k+1)h\}). \end{aligned}$$

Ce qui donne

$$\begin{aligned} & \sum_k (1 - \psi) \mathbb{P}\{Y_1 > x - kh\} \mathbb{P}\{kh \leq Y_2 < (k+1)h\} \leq \mathbb{P}\{X_1 + X_2 > x\} \\ & \leq \sum_k (1 + \psi) \mathbb{P}\{Y_1 > x - kh\} \mathbb{P}\{kh \leq Y_2 < (k+1)h\}. \end{aligned}$$

Quand  $h$  tend vers 0, on obtient

$$(1 - \psi) \mathbb{P}\{Y_1 + Y_2 > x\} \leq \mathbb{P}\{X_1 + X_2 > x\} \leq (1 + \psi) \mathbb{P}\{Y_1 + Y_2 > x\}. \quad (2.3.3)$$

Ainsi le résultat est démontré pour  $n = 2$ .

Supposons que la proposition est vraie pour  $(n - 1)$  et montrons qu'elle est vraie la pour  $n$ .

Pour tout  $h > 0$ , on a

$$\left\{ \sum_{i=1}^n X_i > x \right\} = \bigcup_{k=-\infty}^{k=+\infty} \left( \left\{ \sum_{i=1}^{n-1} X_i > x \right\} \cap \{kh \leq X_n < (k+1)h\} \right), h > 0.$$

Par ailleurs, pour tout  $k$  on a

$$\begin{aligned} & \left\{ \sum_{i=1}^{n-1} X_i > x - kh \right\} \cap \{kh \leq X_n < (k+1)h\} \\ \subset & \left\{ \sum_{i=1}^n X_i > x \right\} \cap \{kh \leq X_n < (k+1)h\} \\ \subset & \left\{ \sum_{i=1}^{n-1} X_i > x - (k+1)h \right\} \cap \{kh \leq X_n < (k+1)h\}. \end{aligned}$$

En sommant sur  $\mathbb{Z}$ , on obtient

$$\begin{aligned} & \bigcup_{k=-\infty}^{k=+\infty} \left( \left\{ \sum_{i=1}^{n-1} X_i > x - kh \right\} \cap \{kh \leq X_n < (k+1)h\} \right) \subset \left\{ \sum_{i=1}^n X_i > x \right\} \\ \subset & \bigcup_{k=-\infty}^{k=+\infty} \left( \left\{ \sum_{i=1}^{n-1} X_i > x - (k+1)h \right\} \cap \{kh \leq X_n < (k+1)h\} \right). \end{aligned}$$

En vertu des propriétés de la probabilité, on a

$$\begin{aligned} & \sum_{k=-\infty}^{\infty} \mathbb{P} \left( \left\{ \sum_{i=1}^{n-1} X_i > x - kh \right\} \cap \{kh \leq X_n < (k+1)h\} \right) \leq \mathbb{P} \left\{ \sum_{i=1}^n X_i > x \right\} \\ \leq & \sum_{k=-\infty}^{\infty} \mathbb{P} \left( \left\{ \sum_{i=1}^{n-1} X_i > x - (k+1)h \right\} \cap \{kh \leq X_n < (k+1)h\} \right). \end{aligned}$$

Ce qui donne grâce à (2.3.2)

$$\begin{aligned} & \sum_{k=-\infty}^{\infty} (1 - \psi(\tau)) \mathbb{P} \left\{ \sum_{i=1}^{n-1} X_i > x - kh \right\} \mathbb{P} \{ kh \leq X_n < (k+1)h \} \leq \mathbb{P} \left\{ \sum_{i=1}^n X_i > x \right\} \\ & \leq \sum_{k=-\infty}^{\infty} (1 + \psi(\tau)) \mathbb{P} \left\{ \sum_{i=1}^{n-1} X_i > x - (k+1)h \right\} \mathbb{P} \{ kh \leq X_n < (k+1)h \}. \end{aligned}$$

La formule étant vérifiée pour  $n - 1$ , donc

$$\begin{aligned} & (1 - \psi(\tau)) \sum_{k=-\infty}^{\infty} (1 - \psi(\tau))^{n-2} \mathbb{P} \left\{ \sum_{i=1}^{n-1} Y_i > x - kh \right\} \mathbb{P} \{ kh \leq X_n < (k+1)h \} \leq \mathbb{P} \left\{ \sum_{i=1}^n X_i > x \right\} \\ & \leq (1 + \psi(\tau)) \sum_k (1 + \psi(\tau))^{n-2} \mathbb{P} \left\{ \sum_{i=1}^{n-1} Y_i > x - (k+1)h \right\} \mathbb{P} \{ kh \leq X_n < (k+1)h \}. \end{aligned}$$

En notant

$$\mathbb{P} \{ X_n > x_n \} = \mathbb{P} \{ Y_n > x_n \},$$

on obtient

$$\begin{aligned} & (1 - \psi(\tau))^{n-1} \sum_{k=-\infty}^{\infty} \mathbb{P} \left\{ \sum_{i=1}^n Y_i > x - kh \right\} \mathbb{P} \{ kh \leq Y_n < (k+1)h \} \leq \mathbb{P} \left\{ \sum_{i=1}^n Y_i > x \right\} \\ & \leq (1 + \psi(\tau))^{n-1} \sum_{k=-\infty}^{\infty} \mathbb{P} \left\{ \sum_{i=1}^{n-1} Y_i > x - (k+1)h \right\} \mathbb{P} \{ kh \leq Y_n < (k+1)h \}. \end{aligned}$$

En faisant tendre  $h$  tend vers 0, on obtient

$$(1 - \psi)^{n-1} \mathbb{P} \left\{ \sum_{i=1}^n Y_i > x \right\} \leq \mathbb{P} \left\{ \sum_{i=1}^n X_i > x \right\} \leq (1 + \psi)^{n-1} \mathbb{P} \left\{ \sum_{i=1}^n Y_i > x \right\}.$$

Ainsi, la proposition démontrée. ■

**Proposition 2.3.2** Soient  $(X_i)_i$  une suite de variables aléatoires  $\psi$ -mélangeantes définies

sur un espace probabilisé  $(\mathcal{X}, \mathcal{F}, \mathbb{P})$  et  $\varphi$  une fonction borélienne. Alors  $(\varphi(X_i))_i$  est une

suite de variables aléatoires  $\psi$ -mélangeante.

**Démonstration.** Soient  $B_1$  et  $B_2$  deux boréliens

$$\begin{aligned} \mathbb{P}(\varphi(X_i) \in B_1, \varphi(X_j) \in B_2) &= \mathbb{P}(X_i \in \varphi^{-1}(B_1), X_j \in \varphi^{-1}(B_2)) \\ &\leq (1 + \psi(\tau)) \mathbb{P}(X_i \in \varphi^{-1}(B_1)) \mathbb{P}(X_j \in \varphi^{-1}(B_2)) \\ &\leq (1 + \psi(\tau)) \mathbb{P}(\varphi(X_i) \in B) \mathbb{P}(\varphi(X_j) \in B_2) \end{aligned}$$

De même

$$(1 - \psi(\tau)) \mathbb{P}(\varphi(X_i) \in B) \mathbb{P}(\varphi(X_j) \in B_2) \leq \mathbb{P}(\varphi(X_i) \in B_1, \varphi(X_j) \in B_2)$$

ce qui montre que les variables aléatoires  $\varphi(X_i)$  sont  $\psi$ -mélangeantes. ■

## 2.4 Inégalités exponentielles

### 2.4.1 Résultats préliminaires

**Lemme 2.4.1** *Pour tous  $a > 0, c > 0, i + 1 > ac$  et  $i \in \mathbb{N}$ , on a*

$$\prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \leq \left(\frac{i+1}{n+1}\right)^{ac}.$$

**Démonstration.** On a

$$\log \left( \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \right) = \sum_{j=i+1}^n \log \left(1 - \frac{ac}{j}\right)$$

et

$$\forall x < 1, \ln(1 - x) \leq -x.$$



Donc

$$\log \left( \prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) \right) \leq \sum_{j=i+1}^n -\frac{ac}{j},$$

mais

$$\left( -\frac{ac}{j} \right)_{j \in \mathbb{N}^*} \text{ est strictement croissante,}$$

donc grâce au théorème de comparaison série-intégrale, on obtient

$$\sum_{j=i+1}^n -\frac{ac}{j} \leq -\int_{i+1}^{n+1} \frac{ac}{x} dx.$$

Par conséquent

$$\log \left( \prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) \right) \leq ac (\log(i+1) - \log(n+1)).$$

En vertu des propriétés du logarithme népérien, on a

$$\log \left( \prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) \right) \leq ac \log \left( \frac{i+1}{n+1} \right).$$

Comme l'exponentielle est une fonction croissante, alors

$$\exp \left( \log \left( \prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) \right) \right) \leq \exp \left( ac \log \left( \frac{i+1}{n+1} \right) \right).$$

Autrement dit

$$\prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) \leq \left( \frac{i+1}{n+1} \right)^{ac}.$$

Ce qui achève la démonstration. ■

**Lemme 2.4.2** Pour tous  $a > 0, c > 0, ac \leq 1$  on a

$$\sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \leq \frac{2}{c}.$$

**Démonstration.** Vue que

$$i + 1 < ac \text{ car } ac \leq 1,$$

alors par le Lemme 2.4.1, on a

$$\prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \leq \left(\frac{i+1}{n+1}\right)^{ac},$$

et par la suite

$$\frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \leq \frac{a}{i} \left(\frac{i+1}{n+1}\right)^{ac}, \forall i \in \mathbb{N}^*.$$

En sommant sur  $\mathbb{N}^*$ , on obtient

$$\sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \leq \sum_{i=1}^n \frac{a}{i} \left(\frac{i+1}{n+1}\right)^{ac}.$$

Comme

$$i + 1 \leq 2i, \forall i \in \mathbb{N}^*,$$

alors

$$\sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \leq \sum_{i=1}^n \frac{a}{i} \left(\frac{2i}{n+1}\right)^{ac}.$$

Comme  $a, 2^{ac}$  et  $(n+1)^{ac}$  sont des constantes par rapport à  $i$ , alors

$$\sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left(1 - \frac{ac}{j}\right) \leq \frac{a(2)^{ac}}{(n+1)^{ac}} \sum_{i=1}^n (i)^{ac-1}.$$

Puisque  $ac \leq 1$  alors  $(i)^{ac-1}$  est une suite décroissante. Donc grâce au théorème de comparaison série-intégrale, on a

$$\sum_{i=1}^n (i)^{ac-1} \leq \left( \int_1^n (x)^{ac-1} dx + 1 \right).$$

Ce qui fait que

$$\sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) \leq \frac{a(2)^{ac}}{(n+1)^{ac}} \left( \int_1^n (x)^{ac-1} dx + 1 \right).$$

Comme

$$\int_1^n (x)^{ac-1} dx + 1 = \frac{1}{ac} ((n)^{ac} - 1) + 1,$$

alors

$$\sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) \leq \frac{a(2)^{ac}}{(n+1)^{ac}} \left( \frac{1}{ac} ((n)^{ac} - 1) + 1 \right),$$

qui est équivalent à

$$\sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) \leq \frac{a}{ac} \frac{(2)^{ac} (n)^{ac}}{(n+1)^{ac}}.$$

Mais puisque

$$\frac{n^{ac}}{(n+1)^{ac}} \leq 1,$$

alors

$$\sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) \leq \frac{1}{c} (2)^{ac}.$$

Comme  $ac < 1$ , alors

$$\sum_{i=1}^n \frac{a}{i} \prod_{j=i+1}^n \left( 1 - \frac{ac}{j} \right) \leq \frac{2}{c}.$$

■

**Proposition 2.4.1** Soit  $(X_n)_{n \in \mathbb{N}^*}$  une suite de variables aléatoires indépendantes centrées

et  $t > 0$ , alors

$$\mathbb{E} \left( \cosh t \left| \sum_{i=1}^n X_i \right| \right) \leq \prod_{i=1}^n \mathbb{E} (e^{t|X_i|} - t|X_i|),$$

où  $\mathbb{E}$  représente l'espérance mathématique.

**Proposition 2.4.2** Soit  $f : \mathbb{R} \rightarrow \mathbb{R}$  une fonction telle que

$$\exists C > 0 : \forall x \in \mathbb{R}, \frac{f(x)}{x - x^*} > C.$$

Soit la relation de récurrence suivante

$$x_{n+1} = x_n - a_n (f(x_n) + \xi_n), \tag{2.4.1}$$

telle que  $a_n$ , est le terme général d'une série divergente et de manière que la série de terme

général  $a_n^2$  soit convergente. Alors, en posant  $x^*$  l'unique racine de la fonction  $f$ , on aura

$$x_{n+1} - x^* = (x_1 - x^*) \prod_{i=1}^n (1 - a_i f'(c_i)) - \sum_{i=1}^n a_i \prod_{j=i+1}^n (1 - a_j f'(c_j)) \xi_i.$$

**Démonstration.** La démonstration se fait par récurrence sur  $n$ . ■

## 2.5 L'inégalité qui en résulte

Dans ce qui suit, on considérera les  $\xi_i$  comme une suite de variables aléatoires  $\psi$ -mélangeantes,

centrées et vérifiant la condition de Cramer, c'est-à-dire

$$E\xi_i = 0, \forall i \in \mathbb{N}^*$$

et

$$E |\xi_i|^m \leq \frac{m!}{2} \sigma^2 H^{m-2} \text{ avec } \sigma = E |\xi_i|^2 \text{ et } H > 0,$$

où  $E$  représente l'espérance mathématique.

**Théorème 2.5.1** *Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires vérifiant la relation de récurrence (2.4.1), alors*

$$\mathbb{P} \{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi(\tau))^{n-1} \exp\left(-\frac{3}{16} \frac{\varepsilon^2}{b_n}\right). \quad (2.5.1)$$

où

$$b_n = \sigma^2 \sum_{i=1}^n a_i^2 \left(\frac{i+1}{n+1}\right)^{2ac}. \quad (2.5.2)$$

**Démonstration.** Pour tout  $\varepsilon > 0$ , on a

$$\mathbb{P} \{|x_{n+1} - x^*| > \varepsilon\} \leq \mathbb{P} \left( \left| (x_1 - x^*) \prod_{i=1}^n (1 - a_i f'(c_i)) \right| + \left| \sum_{i=1}^n a_i \prod_{j=i+1}^n (1 - a_j f'(c_j)) \xi_i \right| > \varepsilon \right).$$

et comme

$$0 < (1 - a_i f'(c_i)) < 1, \forall i \in \mathbb{N}^*,$$

alors

$$\prod_{i=1}^n (1 - a_i f'(c_i)) \rightarrow 0, \text{ quand } n \rightarrow \infty.$$

Par conséquent, il existe  $n_0 \in \mathbb{N}$  tel que  $\forall n > n_0$

$$\left| (x_1 - x^*) \prod_{i=1}^n (1 - a_i f'(c_i)) \right| \leq \frac{\varepsilon}{2}. \quad (2.5.3)$$

Ce qui nous donne, pour  $n > n_0$

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq \mathbb{P}\left(\left|\sum_{i=1}^n a_i \prod_{j=i+1}^n (1 - a_j f'(c_j)) \xi_i\right| > \frac{\varepsilon}{2}\right).$$

Posons

$$X_i = a_i \prod_{j=i+1}^n (1 - a_j f'(c_j)) \xi_i,$$

Alors

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq \mathbb{P}\left(\left|\sum_{i=1}^n X_i\right| > \frac{\varepsilon}{2}\right).$$

Puisque les  $\xi_i$  sont  $\psi$ -mélangeantes, alors les  $X_i$  le sont aussi.

Il existe donc grâce à l'inégalité (2.3.1),  $n$  variables aléatoires indépendantes  $Y_1, Y_2, \dots, Y_n$ ,

telles que

$$(1 - \psi)^{n-1} \mathbb{P}\left\{\sum_{i=1}^n Y_i > \frac{\varepsilon}{2}\right\} \leq \mathbb{P}\left\{\sum_{i=1}^n X_i > \frac{\varepsilon}{2}\right\} \leq (1 + \psi)^{n-1} \mathbb{P}\left\{\sum_{i=1}^n Y_i > \frac{\varepsilon}{2}\right\},$$

ce qui implique donc que

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi)^{n-1} \mathbb{P}\left\{\left|\sum_{i=1}^n Y_i\right| > \frac{\varepsilon}{2}\right\}.$$

Ainsi, grâce à l'inégalité de Markov et à l'inégalité triangulaire, on a pour tout  $t > 0$

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi)^{n-1} \exp\left(\frac{-t\varepsilon}{2}\right) \mathbb{E}\left(\exp\left(t \sum_{i=1}^n |Y_i|\right)\right).$$

Sachant que

$$\exp x \leq 2 \cosh x,$$

alors

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi)^{n-1} \exp\left(\frac{-t\varepsilon}{2}\right) \mathbb{E}\left(2 \cosh\left(t \sum_{i=1}^n |Y_i|\right)\right).$$

En vertu du Théorème 3 page 144 (voir [57]), on a

$$\mathbb{E}\left(2 \cosh\left(t \sum_{i=1}^n |Y_i|\right)\right) \leq \mathbb{E}\prod_{i=1}^n (e^{t|Y_i|} - t|Y_i|).$$

La fonction  $x \mapsto e^x - x$  est continue, donc borélienne. Par conséquent les variables aléatoires  $e^{t|Y_i|} - t|Y_i|$  sont indépendantes grâce à l'indépendance des variables aléatoires

$Y_i$ . Ce qui donne

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi)^{n-1} \exp\left(\frac{-t\varepsilon}{2}\right) \prod_{i=1}^n \mathbb{E}(e^{t|Y_i|} - |tY_i|).$$

En utilisant le développement en série entière de la fonction exponentielle, on obtient

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi)^{n-1} \exp\left(\frac{-t\varepsilon}{2}\right) \prod_{i=1}^n \mathbb{E}\left(\sum_{m=0}^{\infty} \frac{|tY_i|^m}{m!} - |tY_i|\right).$$

Ce qui nous donnera

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi)^{n-1} \exp\left(\frac{-t\varepsilon}{2}\right) \left(\prod_{i=1}^n \left(1 + \sum_{m=2}^{\infty} \frac{\mathbb{E}|tY_i|^m}{m!}\right)\right).$$

Les  $|tY_i|$  vérifient la condition de Cramer, c'est-à-dire

$$E|tY_i|^m \leq t^m \frac{m!}{2} a_i^m \left| \prod_{j=i+1}^n (1 - a_j f'(c_j)) \right|^m \sigma^2 H^{m-2}, \text{ avec } H > 0.$$

Ainsi, en utilisant le Lemme 2.4.1, on obtient

$$\prod_{i=1}^n \left(1 + \sum_{j=2}^{\infty} \frac{\mathbb{E}|tY_i|^j}{j!}\right) \leq \prod_{i=1}^n \left(1 + \sum_{m=2}^{\infty} \frac{t^m}{2} a_i^m \left(\frac{i+1}{n+1}\right)^{mac} \sigma^2 H^{m-2}\right).$$

Ce qui nous donne finalement

$$\begin{aligned} & \mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq \\ & \leq (1 + \psi(\tau))^{n-1} \exp\left(\frac{-t\varepsilon}{2}\right) \prod_{i=1}^n \left(1 + \frac{t^2\sigma^2 a_i^2}{2} \left(\frac{i+1}{n+1}\right)^{2ac} \sum_{m=2}^{\infty} \left(t a_i H \left(\frac{i+1}{n+1}\right)^{ac}\right)^{m-2}\right). \end{aligned}$$

Ceci étant valable pour tout  $t$  positif, on peut fixer  $t < \frac{1}{2a_i H}$ , de manière à avoir

$$\sum_{m=2}^{\infty} \left(t a_i H \left(\frac{i+1}{n+1}\right)^{ac}\right)^{m-2} \leq 1.$$

Avec ce même  $t$  on aura donc

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi(\tau))^{n-1} \exp\left(\frac{-t\varepsilon}{2}\right) \prod_{i=1}^n \left(1 + \frac{t^2\sigma^2 a_i^2}{2} \left(\frac{i+1}{n+1}\right)^{2ac}\right).$$

Puisque

$$\forall x > 0, 1 + x \leq \exp x,$$

alors

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi)^{n-1} \exp\left(\frac{-t\varepsilon}{2}\right) \prod_{i=1}^n \exp\left(\frac{t^2\sigma^2}{2} \left(\frac{i+1}{n+1}\right)^{2ac}\right).$$

Et de ce fait

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi)^{n-1} \exp\left(t^2\sigma^2 \sum_{i=1}^n a_i^2 \left(\frac{i+1}{n+1}\right)^{2ac} - \frac{t\varepsilon}{2}\right). \quad (2.5.4)$$

Notons

$$b_n = \sigma^2 \sum_{i=1}^n a_i^2 \left(\frac{i+1}{n+1}\right)^{2ac}.$$



Le  $t$  qui minimise l'expression (2.5.4) est

$$t = \frac{\varepsilon}{4b_n}.$$

En rapportant cette valeur dans (2.5.4), on obtient

$$\mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} \leq (1 + \psi(\tau))^{n-1} \exp\left(-\frac{3}{16} \frac{\varepsilon^2}{b_n}\right).$$

C'est ce qu'il fallait démontrer. ■

**Remarque 2.5.1** Par le lemme de Kronecker [?], on a

$$\lim_{n \rightarrow \infty} b_n = 0.$$

La fonction  $t \mapsto \frac{1}{(n+1)^t} \sum_{i=1}^n \frac{(1+i)^t}{i^2}$  est décroissante, on prendra  $2ac = 1.99$ .

### 2.5.1 Intervalle de confiance

**Corollaire 2.5.1** Sous les hypothèses du Théorème ?? et pour un niveau  $\alpha \in ]0, 1[$  donné,

Il existe un entier naturel  $n_\alpha$  pour lequel la solution exacte  $x^*$  de l'Equation 1.2.1 appar-

tient à l'intervalle fermé  $[x_{n_\alpha} - \varepsilon, x_{n_\alpha} + \varepsilon]$  avec une probabilité supérieure ou égale à  $1 - \alpha$ .

$$\forall \varepsilon > 0, \forall \alpha \in ]0, 1[, \exists n_\alpha \in \mathbb{N} : \mathbb{P}\{|x_{n+1} - x^*| > \varepsilon\} > 1 - \alpha. \quad (2.5.5)$$

**Démonstration.** En effet, en utilisant le lemme de Kronecker, on obtient  $\lim_{n \rightarrow +\infty} b_n =$

0 ce qui implique

$$\lim_{n \rightarrow +\infty} (1 + \psi(\tau))^{n-1} \exp\left(-\frac{3}{16} \frac{\varepsilon^2}{b_n}\right) = 0.$$

donc il existe un entier naturel  $n_\alpha$  tel que

$$\forall n \in \mathbb{N}, n > n_\alpha \implies (1 + \psi(\tau))^{n-1} \exp\left(-\frac{3}{16} \frac{\varepsilon^2}{b_n}\right) \leq \alpha \quad (2.5.6)$$

ainsi, (2.5.5) découle de (32) et (2.5.6). ■

Taille de l'échantillon

Le plus petit entier naturel  $n$  tel que

$$b_n \leq -\frac{3}{16} \frac{\varepsilon^2}{\ln \alpha}$$

$\varepsilon \backslash \alpha$	0.01	0.05	0.1
0.1	479	315	1256
0.01	46540	30283	23280

Table : Tailles de l'échantillon

$$-\frac{3}{16} \frac{(0.001)^2}{\ln(0.01)} = 4.0715 \times 10^{-8}$$

$$-\frac{3}{16} \frac{(0.001)^2}{\ln(0.05)} = 6.2589 \times 10^{-8}$$

$$-\frac{3}{16} \frac{(0.001)^2}{\ln(0.1)} = 8.143 \times 10^{-8}$$

Longueur de l'intervalle de confiance

$$\varepsilon = \sqrt{-\frac{16}{3} b_n \ln \alpha}$$

$n \setminus \alpha$	0.01	0.05	0.1
10	1.88791	1.52268	1.33495
$10^2$	0.522191	0.421171	0.369245
$10^3$	0.158682	0.127984	0.112205
$10^4$	0.0498583	0.040213	0.0352551
$10^5$	0.0157528	0.0127054	0.0111389
$10^6$	0.00498094	0.00401735	0.00352206

Niveau de confiance

$$100 \left( 1 - (1 + \psi(n))^{n-1} \exp\left(-\frac{3}{16} \frac{\varepsilon^2}{b_n}\right) \right)$$

$n \setminus \varepsilon$	0.1	0.01	0.001
10	1.28%	0.01%	0.0001%
$10^2$	15.53%	0.16%	0.0016%
$10^3$	83.94%	1.81%	0.018%
$10^4$	100%	16.91%	0.18%
$10^5$	100%	84.36%	1.83%
$10^6$	100%	100%	16.94%
$10^7$	100%	100%	
$10^8$	100%	100%	
$10^{10}$	100%	100%	
$10^{12}$	100%	100%	
$10^{14}$	100%	100%	100%

## 2.5.2 Convergence presque complète

**Corollaire 2.5.2** *Sous les hypothèses du Théorème ??, la suite  $(x_n)_n$  converge presque complètement (p.co) vers la solution de l'Equation (1.2.1).*

C'est-à-dire

$$\forall \varepsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}\{|x_n - x^*| > \varepsilon\} < +\infty.$$

**Démonstration.** La série de terme général

$$u_n = \exp\left(-\frac{3}{16} \frac{\varepsilon^2}{b_n}\right)$$

est convergente. En effet, si  $ac < \frac{1}{2}$

$$\sum_{i=1}^n \frac{(i+1)^{2ac}}{i^2} \leq \sum_{i=1}^{\infty} \frac{(i+1)^{2ac}}{i^2} = K < +\infty$$

$$\exp\left(-\frac{3}{16} \frac{\varepsilon^2}{b_n}\right) \leq \exp\left(-\frac{3k}{16} \varepsilon^2 (n+1)^{2ac}\right)$$

qui est convergente.

Si  $\frac{1}{2} \leq ac < 1$ , la fonction  $t \mapsto \frac{1}{(n+1)^t} \sum_{i=1}^n \frac{(1+i)^t}{i^2}$  est décroissante,

$$\frac{1}{(n+1)^{2ac}} \sum_{i=1}^n \frac{(1+i)^{2ac}}{i^2} \leq \frac{1}{(n+1)} \sum_{i=1}^n \frac{(1+i)}{i^2}$$

$$\sum_{i=1}^n \frac{(1+i)}{i^2} \leq \frac{\pi^2}{6} + 1 + \ln n$$

$$\exp\left(-\frac{3}{16} \frac{\varepsilon^2}{b_n}\right) \leq \exp\left(-\frac{3(n+1)}{16\left(\frac{\pi^2}{6} + 1 + \ln n\right)} \varepsilon^2\right)$$

ce qui assure la convergence presque complète ■

### 2.5.3 Rate of convergence

Dans cette sous-section, nous étudions le taux de convergence de l'algorithme stochastique

de Robbins-Monro. ???. Rappelons que  $x_n - x^* = \mathcal{O}(u_n)$ , presque complètement (*a.co.*),

où  $(u_n)$  est une séquence de nombres positifs réels, s'il existe  $\epsilon_0 > 0$  tel que

$$\sum_{n=1}^{\infty} \mathbb{P} \{ |x_n - x^*| > \epsilon_0 u_n \} < +\infty.$$

**Corollaire 2.5.3** *Sous les hypothèses du Théorème ??, on a*

$$x_n - x^* = \mathcal{O}(b_n \ln n) \text{ p.co.} \quad (2.5.7)$$

avec  $\lim_{n \rightarrow +\infty} b_n \ln n = 0$ .

**Démonstration.** Pour  $\epsilon_0$  bien choisi, par exemple,  $\epsilon_0 = \frac{16}{3}(1+d)$ ,  $d > 0$

$$\mathbb{P} \left\{ |x_n - x^*| > \frac{16}{3}(1+d)b_n \ln n \right\} \leq K_1 \frac{1}{n^{1+d}} \quad (2.5.8)$$

le second membre l'inégalité (2.5.8) est un terme général d'une série convergente. Par conséquent, le résultat souhaité (2.5.7) est démontré. ■

## 2.6 Etude numérique

### 2.6.1 Introduction

Historiquement, on s'est intéressé depuis le début du vingtième siècle à ce qu'on appelle la génération de nombres aléatoires, par l'utilisation de certaines tables formées chacune de manière différente. Ainsi, la plus ancienne table de nombres aléatoires connue est la table de Tippett, créée en 1927, elle est constituée de 10400 nombres de 4 chiffres. Puis en 1939 il y eut la table de Kandall et B.B. Smith qui contient cent mille chiffres obtenus à partir

de disques tournants. Avec le temps, les techniques de génération de ces nombres ont considérablement évolué et on peut facilement aujourd'hui obtenir des nombres aléatoires grâce aux logiciels de mathématiques présents dans le marché. Nous utiliserons pour notre application numérique le logiciel MATLAB version R2009b.

### 2.6.2 Motivation

L'approximation stochastique a fait ses preuves depuis maintenant plus d'un demi siècle, et trouve des applications dans divers domaines. Ainsi, depuis son apparition, elle a été utilisée dans différents domaines de l'ingénierie et de l'économie. Robbins et Monro ont mis au point le premier algorithme d'approximation stochastique en voulant trouver les zéros d'une fonction  $f$ , dont on ne connaît pas l'expression analytique mais dont on peut obtenir une valeur mesurée  $f(x)$  en chaque point  $x$  donné. Leur travail a été réalisé sur un phénomène chimique où  $f(x)$  représentait la valeur moyenne de l'effet qu'a produit une dose  $x$  d'une substance utilisée pour l'expérience. Depuis, des algorithmes d'approximation stochastique ont été développés pour résoudre des problèmes de ce genre, c'est-à-dire des problèmes dont on ne peut obtenir l'expression analytique de leur fonction objectif, mais dont on peut obtenir des valeurs bruitées en chaque point donné. On pourrait voir un exemple d'utilisation de ces algorithmes en économie dans [51].

### 2.6.3 Application

Considérons une suite de variables aléatoires indépendantes  $(e_t)_{t \in \mathbb{Z}}$  telle que pour tout  $t \in \mathbb{Z}$ , la variable aléatoire  $e_t$  suit une loi de Rademacher. Autrement dit,

$$\forall t \in \mathbb{Z}, \mathbb{P}\{e_t = 1\} = \mathbb{P}\{e_t = -1\} = \frac{1}{2}.$$

Considérons une suite de variables aléatoires gaussiennes  $(X_t)_{t \in \mathbb{N}}$  indépendantes de  $(e_t)_{t \in \mathbb{Z}}$ .

La suite de variables aléatoires  $\xi_t := e_t \cdot X_t$  obtenue, est une suite de variables aléatoires  $\psi$ -mélangeantes.

### 2.6.4 Présentation de la simulation

Nous avons construit notre application sur Matlab ( Matrix laboratory), qui est un langage de programmation utilisé à des fins de calcul numérique.

La fonction "rand" sur Matlab permet de retourner aléatoirement un nombre  $a \in ]0, 1[$ , elle simule donc des variables aléatoires indépendantes suivant une loi uniforme.

Pour simuler des variables aléatoires suivant une loi de Rademacher qui nous donne le résultat 1 ou  $-1$  avec la même probabilité, nous avons procédé de la manière suivante :

Nous avons remarqué que

$$P(rand < 0.5) = \frac{1}{2} \text{ et } P(rand \geq 0.5) = \frac{1}{2}.$$

Pour simuler une suite de variables indépendantes  $(e_t)_{t \in \mathbb{Z}}$  suivant une loi de Rademacher on utilise la condition "si" et on affecte à  $e_t$  la valeur 1 ou  $-1$  selon la valeur retourné par "rand", c'est-à-dire

$$\text{Si } \text{rand} < 0.5 \text{ alors } e_t = 1 \text{ sinon } e_t = -1.$$

Pour obtenir des variables aléatoires gaussiennes nous pouvons procéder comme suit,

$$X = \sqrt{(-2 * \log(\text{rand}))} \cdot \cos(2\pi \text{rand}), \text{ comme nous pouvons le vérifier à la page du cours}$$

[65]. Ainsi programmée,  $X \rightsquigarrow \mathcal{N}(0, 1)$ .

### 2.6.5 Exemples et résultats

Notre application peut approximer le zéro de toute fonction monotone à dérivée bornée.

Le tableau suivant résume les résultats que nous avons obtenus pour quelques fonctions.

	N <sup>bre</sup> d'itérations	10	500	5000	50000	200000
<i>Fonctions</i>	Points de départ					
$3x + 5$	$\frac{11}{2}$	-1.5579	-1.6731	-1.6711	-1.6673	-1.66628
$3x + 5$	2500	28.8051	-0.0344	-1.3756	-1.6492	-1.66566
$-2x - 6$	$\frac{1}{2}$	-2.8541	-2.9615	-2.9957	-3.0007	-3.00028
$-2x - 6$	-3000	-82.5274	-8.9023	-4.2725	-3.1743	-3.10739
$-7x + 2$	1	0.1795	0.2853	0.2866	0.2863	0.28545
$-7x + 2$	7500	16.8154	0.8179	0.3582	0.2955	0.28848
$6x - 4$	-6000	-17.3388	0.0389	0.5777	0.6542	0.66292
$6x - 4$	4	0.7437	0.6691	0.6660	0.6664	0.66671



### 2.6.6 Conclusion

La première chose qu'on remarque en analysant le tableau des résultats obtenus est que l'algorithme est à convergence lente, c'est-à-dire qu'il faut plusieurs, (des milliers), d'itérations pour approcher efficacement la solution théorique de la fonction.

La deuxième chose est la sensibilité du programme à la distance entre le point de départ et la solution de la fonction choisie.

On remarque par exemple que pour la fonction  $3x + 5$ , avec le point de départ  $\frac{11}{2}$ , on obtient après 50.000 itérations un résultat proche au millième près de la solution théorique. Alors que pour le point de départ 2500 il faut 200000 itérations pour obtenir un résultat comparable.

Finalement, on remarque qu'on peut quand même se faire une idée de la région où se trouve la solution après 5000 itérations, ce qui est assez bon.

---

# Ellipses de confiance pour les paramètres d'une régression simple à erreurs fortement mélangeantes

## 3.1 Introduction

Le concept de régression vient de la génétique. La théorie de la régression linéaire simple a été établie et popularisée par Galton [39] à la fin du 19<sup>ème</sup> siècle avec la publication intitulée "Regression towards mediocrity in hereditary stature". Son travail a ensuite été étendu par (Yule [72]; Pearson [55] et Fisher [37]) à un contexte statistique plus général. Depuis, de nombreux travaux se sont développés à partir de cette théorie, dans plusieurs domaines (Volkonskii et Rozanov [68]; Babu et Bai [7]; Tingley [66] et Zvara [73]). Dans la régression linéaire simple, on suppose que deux variables  $x$  et  $y$  sont liées linéairement avec des paramètres d'ordonnée à l'origine et de pente inconnus. En particulier, la variable

exogène  $x$  est supposée être précisément mesurable et la variable endogène  $y$  est supposée être une variable aléatoire dépendant de  $x$  via une fonction linéaire. De nos jours, dans de nombreux problèmes de la vie quotidienne, il est plus courant de mesurer plusieurs variables  $Y$  au lieu d'une seule, pour obtenir le modèle dit multivarié

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

où  $Y$  est un vecteur  $n \times 1$  de variables de réponse,  $X$  est une matrice  $n \times p$  de variables explicatives, variables exogènes ou encore prédicteurs,  $\beta$  est un vecteur  $p \times 1$  de coefficients inconnus et  $\varepsilon$  est un vecteur  $n \times 1$  d'erreurs ou perturbations inconnues.

La littérature sur la régression linéaire multiple est considérable (voir [53] et références citées). Plusieurs auteurs ont étudié le domaine de confiance des paramètres d'une régression multiple lorsque les erreurs sont indépendantes [14, 17, 33, 47, 50, 62]. Dans [13], des procédures statistiques de régression linéaire sont établies, lorsque le processus d'erreur est supposé corrélé.

Il existe différentes façons d'estimer la matrice de covariance asymptotique des estimateurs des moindres carrés. En utilisant l'estimation de la matrice de covariance, les intervalles de confiance et les tests habituels sur les paramètres sont modifiés. Lorsque nous parlons de mesure, que nous le voulions ou non, des erreurs seront induites en raison de la nature stochastique de toute mesure. Dans la plupart des cas, les erreurs de mesure

sont considérées comme indépendantes. Notons que les erreurs indépendantes ont échoué pour la modélisation de certains phénomènes. En pratique, cette hypothèse n'est souvent pas adaptée. En effet, les erreurs dépendantes sont plus adaptées à la réalité et il existe une classe importante de dépendance qui trouve de nombreuses applications.

Pour illustrer avec précision cette situation, nous pouvons citer le modèle d'Ising. Il décrit le comportement de deux électrons proches dans un atome. Ces électrons sont plus susceptibles d'être orientés dans la même direction que dans des directions opposées (Fortuin et al. [38]). Cet attrait s'exprime à travers la dépendance positive. Il est à noter que ce type de dépendance qui a été introduit par (Esary et al. [34]), est basé sur les travaux de Lehmann (Lehmann [49]). Leur objectif était de trouver des applications dans la fiabilité et les statistiques. Cependant, il existe de nombreuses notions de dépendance autres que la dépendance positive. Nous renvoyons ceux exprimés en termes de coefficients de mélange (Blum et al. [8]; Ibragimov et Linnik [43]; Tingley [66]). De nombreux types de conditions de mélange ont été proposés dans la littérature (Doukhan [31]; Rio [59]). Le principal inconvénient du mélange d'hypothèses est la difficulté de les vérifier. Récemment, ces conditions de mélange ont été étudiées non seulement d'un point de vue probabiliste, mais aussi pour leurs possibilités d'application en termes de prévision non paramétrique (Chai [15]).

Dans ce travail, nous considérons une condition de mélange non restrictive pour caractériser la dépendance entre les erreurs aléatoires des données. On suppose qu'il s'agit d'un mélange fort ou  $\alpha$ -mélange (Aiane [1], Arroudj [5]). Nous établissons des inégalités exponentielles, qui nous permettent de construire des régions de confiance, caractérisées par des ellipses, pour les estimations par les moindres carrés des paramètres (Dorogovtsev [32]), dans le cas d'un modèle de régression linéaire avec des erreurs  $\alpha$ -mélangeantes. Pour vérifier la validité des résultats théoriques obtenus, certains résultats numériques sont considérés.

La construction d'inégalités exponentielles pour les paramètres d'une régression linéaire multiple avec erreurs de mélange reste un problème ouvert. Entre autres, ces inégalités seront utilisées pour construire des hyperellipsoïdes pour les paramètres mentionnés.

## 3.2 PRELIMINAIRES

Considérons le modèle de régression linéaire simple

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

où  $a$  et  $b$  sont les paramètres du modèle à estimer à l'aide des observations  $(y_1, y_2, \dots, y_n)$

et  $(\varepsilon_i)_{i \in \mathbb{N}^*}$  est un processus centré, de variances non nulles  $\mathbb{E}(\varepsilon_i^2) = \sigma_i^2$  et admettant des

moments de quatrième ordre bornés.

Notons par

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

Les estimateurs des moindres carrés de  $a$  et  $b$  sont donnés par

$$\hat{a}_n = \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2}, \quad \hat{b}_n = \frac{1}{n} \sum_{i=1}^n y_i$$

et moyennant l'hypothèse  $\sum_{i=1}^n x_i = 0$ , on a

$$\hat{a}_n - a = \frac{\sum_{i=1}^n x_i \varepsilon_i}{\sum_{i=1}^n x_i^2} \text{ et } \hat{b}_n - b = \frac{\sum_{i=1}^n \varepsilon_i}{n}.$$

Il est évident que  $\hat{a}_n$  et  $\hat{b}_n$  sont des estimateurs sans biais de  $a$  et  $b$ .

La notion de  $\alpha$ -mélangeant est définie de la manière suivante (Rosenblatt [61] 1956 )

**Définition 3.2.1** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé. Considérons  $\mathcal{F}_1^i \subset \mathcal{F}$  (respectivement  $\mathcal{F}_{i+p}^\infty \subset \mathcal{F}$ ) la  $\sigma$ -algèbre engendrée par  $\{\varepsilon_j, 1 \leq j \leq i\}$  (respectivement par  $\{\varepsilon_j, j \geq i+p\}$ ).

La processus  $(\varepsilon_i)_{i \in \mathbb{N}}$  est dit  $\alpha$ -mélangeant si :

$$\alpha_\tau = \sup_{A \in \mathcal{F}_1^i, B \in \mathcal{F}_{i+\tau}^{+\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \xrightarrow{\tau \rightarrow +\infty} 0$$

Les hypothèses de mélangeance ou d'independance asymptotique est un domaine très vaste. A ce propos, nous signalons: les théorèmes centraux limites (Rosenblatt [61] ); la statistique d'ordre (Welsch [71] ); processus empiriques (Deo [23] ); logarithme itéré (Philipp [56] ); Loi forte des grands nombres (Chen [16] ); estimations robustes (Boente et Fraiman [9] ) et statistiques non paramétriques (Ferraty et Vieu [35]).

Alors que la condition de mélange fort ( $\alpha$ -mélange) a été largement adoptée dans la littérature, le nombre de procédés bien connus qui satisfont la condition de mélange fort est encore quelque peu limité. Les processus  $m$ -dépendant sont  $\alpha$ -mélangeants, puisque  $\alpha(s) = 0$ ; pour tout  $s > m$  : Kolmogorov et Rozanov (Kolmogorov et Rozanov [44]) ont démontré que les processus gaussiens avec des densités spectrales continues et positives sont  $\alpha$ -mélangeants. Ibragimov et Linnik (Ibragimov et Linnik [43]) ont montré que les processus de Markov stationnaires sont aussi  $\alpha$ -mélangeants avec certaines conditions sur les probabilités de transition. Lee (Lee [48]) montre que  $AR(1)$  satisfait est  $\alpha$ -mélangeant avec un ordre de mélange décroissant exponentiellement vers zéro. Athreya et Pantula (Athreya et Pantula [6]) établissent que certains processus de moyenne mobile autorégressifs stationnaires sont  $\alpha$ -mélangeants.

Cependant, de nombreux processus linéaires strictement stationnaires ne sont pas  $\alpha$ -mélangeants. Un exemple classique bien connu est le processus  $AR(1)$  strictement stationnaire généré par la variable aléatoire de Bernoulli  $B(p)$  et un paramètre autorégressif  $\phi \in ]0; \frac{1}{2}]$  (Andrews [2]).

### 3.3 Résultats

**Proposition 3.3.1** *Pour tout  $\varepsilon$  positif, on a*

$$\begin{aligned} \mathbb{P} \left\{ \frac{\left( \sum_{i=1}^n x_i^2 |\widehat{a}_n - a|^2 + n |\widehat{b}_n - b|^2 \right)^{\frac{1}{2}}}{\left( \frac{1}{n} \sum_{i=1}^n (y_i - \widehat{a}_n x_i - \widehat{b}_n)^2 \right)^{\frac{1}{2}}} > R \right\} &\leq \mathbb{P} \left\{ \frac{\left| \sum_{i=1}^n x_i \varepsilon_i \right|}{\left( \sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}} > r \right\} + \\ \mathbb{P} \left\{ \frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \varepsilon_i \right| > r \right\} &+ \mathbb{P} \left\{ \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 \leq \varepsilon \right\} \end{aligned} \quad (3.3.1)$$

où

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

**Démonstration.** Il suffit de remarquer l'égalité

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

et d'utiliser les deux propriétés de probabilités suivantes

$$\mathbb{P} \{X > tY\} \leq \mathbb{P} \{X > t\varepsilon\} + \mathbb{P} \{Y \leq \varepsilon\} \quad \text{et} \quad \mathbb{P} \{X + Y > t\} \leq \mathbb{P} \left\{ X > \frac{t}{2} \right\} + \mathbb{P} \left\{ Y > \frac{t}{2} \right\}$$

pour en déduire l'inégalité (3.3.1).

Sans nuire à la généralité, on supposera dans ce qui suit que la famille  $(\alpha_n)_n$  est décroissante au sens large et que  $\alpha_0 = 1$ . ■

**Théorème 3.3.1** *Soit  $(\varepsilon_n, n \geq 1)$  une suite de variables aléatoires  $\alpha$ -mélangeantes, centrées et bornées par  $L$ . Si*

$$\sum_{q=1}^{+\infty} \alpha_q < +\infty \quad \text{et} \quad 2 \frac{n\sqrt{e}}{k} (\alpha_k)^{\frac{2k}{3n}} < A \quad (3.3.2)$$



alors, pour tous  $\varepsilon > 0, n \geq 4, k \in \{1, 2, \dots, [\frac{n}{2} - 1]\}$ , ( $[\frac{n}{2} - 1]$  désigne la partie entière

de  $(\frac{n}{2} - 1)$ ), on a

$$\mathbb{P} \left\{ \frac{\left( \sum_{i=1}^n x_i^2 |\widehat{a}_n - a|^2 + n |\widehat{b}_n - b|^2 \right)^{\frac{1}{2}}}{\left( \frac{1}{n} \sum_{i=1}^n (y_i - \widehat{a}_n x_i - \widehat{b}_n)^2 \right)^{\frac{1}{2}}} > R \right\} \leq A_1 \left( e^{-A_2 r^2} + e^{-B_2 r^2} + e^{-C_2 d^2 n} \right) \quad (3.3.3)$$

où

$$d = \left( \min_{1 \leq i \leq n} \mathbb{E} \varepsilon_i^2 - \varepsilon \right), A_1 = \exp A, L_1 = L \max_i |x_i|, D_1 = \bar{\sigma} \max_i |x_i^2|, D_2 = \sup_i (\text{Var } \varepsilon_i^2)$$

$$A_2 = \frac{M}{16e \left( D_1 + 8L_1^2 \sum_{q=1}^k \alpha_q \right)}, B_2 = \frac{1}{16e \left( \bar{\sigma} + 8L^2 \sum_{q=1}^k \alpha_q \right)}, C_2 = \frac{1}{16e \left( D_2 + 32L^4 \sum_{q=1}^k \alpha_q \right)}. \quad (3.3.4)$$

En particulier, si  $R = \sqrt{\rho n}$ , alors

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

avec

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

**Démonstration.** Commençons par majorer le premier terme du second membre de l'inégalité (3.3.1). Soit  $(\zeta_i)_i$  la suite de variables aléatoires définie par  $\zeta_i = \varepsilon_i x_i$ . Cette séquence est  $\alpha$ -mélangeante, centrée et satisfait

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

L'analogie de l'inégalité de Bernstein-Fréchet (Bosq and Lecoutre [12]) appliquée à la

suite  $\alpha$ -mélangeante  $(\zeta_i)_i$  donne : Pour tous  $z > 0$ ,  $n \geq 4$ ,  $k \in \{1, 2, \dots, \lfloor \frac{n}{2} - 1 \rfloor\}$ ,

$$\mathbb{P} \left\{ \left| \sum_{i=1}^n \zeta_i \right| > r \left( \sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \right\} \leq 2 \exp \left( -zr \left( \sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} + 4z^2 e \left( D_1 + 8L_1^2 \sum_{q=1}^k \alpha_q \right) n + 2\sqrt{e} (\alpha_k)^{\frac{2k}{3n}} \frac{n}{k} \right). \quad (3.3.5)$$

En remplaçant  $z$  par la valeur qui minimise le second membre de l'inégalité (3.3.5), on

obtient

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

De (3.3.2), on en déduit que

$$\mathbb{P} \left\{ \frac{\left| \sum_{i=1}^n \varepsilon_i x_i \right|}{\left( \sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}} > r \right\} \leq A_1 e^{-A_2 r^2}. \quad (3.3.6)$$

L'inégalité

$$\mathbb{P} \left\{ \frac{1}{\sqrt{n}} \left| \sum_{i=1}^n \varepsilon_i \right| > r \right\} \leq A_1 e^{-B_2 r^2} \quad (3.3.7)$$

se déduit de la même manière.

Pour majorer le troisième terme du second membre de l'inégalité (3.3.1), on utilise le

lemme suivant ■

**Lemme 3.3.1** (Ibragimov, Linnik [43], Théorème 17.2.1 page 306). Soit  $(\varepsilon_i)_{i \in \mathbb{N}}$  une

séquence de variables aléatoires  $\alpha$ -mélangeantes. Si  $\varepsilon_i$  est mesurable par rapport à la

$\sigma$ -algèbre  $\mathcal{F}_1^i$ , et  $\varepsilon_j$   $i \neq j$ , est mesurable par rapport à la  $\sigma$ -algèbre  $\mathcal{F}_j^\infty$ , et de plus, si

$|\varepsilon_i| \leq L < +\infty$ ,  $|\varepsilon_j| \leq L < +\infty$ , alors

$$|\mathbb{E}(\varepsilon_i \varepsilon_j) - \mathbb{E}(\varepsilon_i) \mathbb{E}(\varepsilon_j)| \leq 4L^2 \alpha_{|i-j|}. \quad (3.3.8)$$

En particulier, lorsque la séquence est centrée, alors,  $|\mathbb{E}(\varepsilon_i \varepsilon_j)| \leq 4L^2 \alpha_{|i-j|}$ .

Une inégalité de la forme (3.3.8), mais avec une constante égale à 16, est démontrée dans [68].

**Démonstration.** la séquence  $(\eta_i)_i$  de variables aléatoires définies par  $\eta_i = \mathbb{E}\varepsilon_i^2 - \varepsilon_i^2$  est  $\alpha$ -mélangeante et satisfait

$$|\eta_i| \leq 2L^2 \text{ et } \mathbb{E}\eta_i^2 \leq D_2.$$

En outre, d'après l'inégalité de Markov on trouve que pour tous  $z > 0$  et  $0 < \varepsilon < \min \mathbb{E}\varepsilon_i^2$

$$\begin{aligned} & \mathbb{P} \left\{ \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 \leq \varepsilon \right\} = \mathbb{P} \left\{ \sum_{i=1}^n \varepsilon_i^2 \leq n\varepsilon \right\} = \mathbb{P} \left\{ \sum_{i=1}^n \varepsilon_i^2 - \sum_{i=1}^n \mathbb{E}\varepsilon_i^2 \leq n\varepsilon - \sum_{i=1}^n \mathbb{E}\varepsilon_i^2 \right\} \\ & \leq \mathbb{P} \left\{ \sum_{i=1}^n (\varepsilon_i^2 - \mathbb{E}\varepsilon_i^2) \leq n \left( \varepsilon - \min_{1 \leq i \leq n} \mathbb{E}\varepsilon_i^2 \right) \right\} = \mathbb{P} \left\{ \sum_{i=1}^n (\mathbb{E}\varepsilon_i^2 - \varepsilon_i^2) \geq n \left( \min_{1 \leq i \leq n} \mathbb{E}\varepsilon_i^2 - \varepsilon \right) \right\} \\ & \leq \exp(-znd) \mathbb{E} \exp(zH_n) \end{aligned} \quad (3.3.9)$$

avec  $d = (\min_{1 \leq i \leq n} \mathbb{E}\varepsilon_i^2 - \varepsilon)$  et  $H_n = \sum_{i=1}^n \eta_i$ .

Majorons l'expression  $\mathbb{E} \exp(zH_n)$ . Pour ça, posons

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

où  $N$  est tel que  $2k(N-1) < n \leq 2kN$ ,  $m_1 = \inf \{(2q+j-3)k+1, n\}$  et  $m_2 = \inf \{m_1 + k - 1, n\}$ .

Ainsi, on a

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

D'une part, la convexité de la fonction exponentielle permet d'écrire

$$\mathbb{E} \exp (z H_n) \leq \frac{1}{2} (\mathbb{E} \exp (2 z Z_{1,N}) + \mathbb{E} \exp (2 z Z_{2,N})). \quad (3.3.10)$$

D'autre part, la suite  $(\eta_m)$  étant  $\alpha$ -mélangeante, il en résulte (Hipp [42] 1979, 60 )

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

où  $h, p \in [1, +\infty]$ ;  $\frac{1}{h} + \frac{1}{p} = 1$ ;  $j = 1, 2$ ;  $q = 1, \dots, N$  avec la convention  $Z_{j,0} = 0$ .

Pour  $z$  bien choisi, par exemple

$$|z Y_{j,q}| \leq \frac{1}{4},$$

on a

$$\mathbb{E} \exp (2 z Y_{j,q}) \leq 1 + 4 z^2 \mathbb{E} Y_{j,q}^2 \leq \exp (4 z^2 \mathbb{E} Y_{j,q}^2). \quad (3.3.11)$$

Il découle alors du Lemme 3.3.1 que

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

En posant

$$C = \operatorname{Var} (\varepsilon_i^2) + 32 L^4 \sum_{q=1}^k \alpha_q \quad (3.3.12)$$

on obtient

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

Par suite, on déduit de (3.3.11) que

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

et

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

En appliquant l'inégalité de Hölder, on obtient

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

De la même manière, on obtient pour  $q \geq 2$

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

et ainsi de suite jusqu'à

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

En multipliant membre à membre toutes ces inégalités on trouve

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

Si on choisit  $h = q + 1$  et  $p = 1 + \frac{1}{q}$ ; alors

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

et

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

Ceci assure que

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

Et comme

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

alors

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

En choisissant  $q = N$  et en remarquant que  $2(N-1) < \frac{n}{k}$ ,  $N < \frac{n}{k}$ ,  $\frac{1}{h} = \frac{1}{N+1} > \frac{2k}{3n}$ , on

obtient

$$\mathbb{E} \exp \left( 2zZ_{j,q} \right) \leq \exp \left( 4z^2eCn + 2\sqrt{e}(\alpha_k)^{\frac{2k}{3n}} \frac{n}{k} \right), j = 1, 2. \quad (3.3.13)$$

L'inégalité

$$\mathbb{P} \left\{ \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 \leq \varepsilon \right\} \leq \exp \left( -znd + 4z^2eCn + 2\sqrt{e}(\alpha_k)^{\frac{2k}{3n}} \frac{n}{k} \right) \quad (3.3.14)$$

se déduit à partir du Lemme 3.3.1 et aussi de (3.3.9), (3.3.10), (3.3.12) et (3.3.13).

Le réel  $z$  qui minimise le deuxième terme de (3.3.14) est

$$z = \frac{d}{8eC}. \quad (3.3.15)$$

L'inégalité

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

découle de (3.3.12), (3.3.14) et (3.3.15).

Les hypothèses du Théorème 3.3.1 montrent qu'il existe deux constantes  $A_1$  et  $C_2$

telles que :

$$\mathbb{P} \left\{ \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 \leq \varepsilon \right\} \leq A_1 \exp(-C_2 d^2 n). \quad (3.3.16)$$

Le résultat du Théorème 3.3.1 s'obtient en regroupant (3.3.6), (3.3.7) et (3.3.16). ■

## 3.4 Ellipses de confiance

On a

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

Ainsi, pour un niveau  $\alpha$  et  $\zeta$  donné, on détermine le plus petit  $n$  qui réalise l'inégalité

suivante

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

Notons ce  $n$  par  $n_\alpha$ . Ainsi

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

ou encore

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

En d'autres termes, avec une probabilité supérieure ou égale à  $1 - \alpha$ , les paramètres  $a$  et

$b$  appartiennent à l'ellipse définie par l'équation

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

**Remarque 3.4.1** Si  $\xi$  ( $\xi = \frac{\varepsilon\rho}{2(1+\rho)}$ ) et la taille de l'échantillon sont donnés et satisfont

$$A_1 (\exp(-A_2\xi n) + \exp(-B_2\xi n) + \exp(-C_2d^2n)) < 1$$

alors, en prenant

$$\alpha = A_1 (\exp(-A_2\xi n) + \exp(-B_2\xi n) + \exp(-C_2d^2n))$$

on obtient que  $a$  et  $b$  appartiennent à l'ellipse de centre  $(\hat{a}_n, \hat{b}_n)$  et aux coordonnées des

foyers

$$F_1 = \left( \hat{a}_n - \sqrt{\rho S_n^2 \left( \frac{1}{M} - 1 \right)}, \hat{b}_n \right), F_2 = \left( \hat{a}_n + \sqrt{\rho S_n^2 \left( \frac{1}{M} - 1 \right)}, \hat{b}_n \right) \text{ si } M < 1$$

$$F_1 = \left( \hat{a}_n, \hat{b}_n - \sqrt{\rho S_n^2 \left( 1 - \frac{1}{M} \right)} \right), F_2 = \left( \hat{a}_n, \hat{b}_n + \sqrt{\rho S_n^2 \left( 1 - \frac{1}{M} \right)} \right) \text{ si } M > 1$$

avec une probabilité supérieure ou égale à  $1 - \alpha$ .

Si le niveau de confiance et la taille de l'échantillon sont donnés, nous déterminons

les foyers  $F_1$  et  $F_2$  à partir de l'équation

$$\alpha = A_1 (\exp(-A_2\xi n) + \exp(-B_2\xi n) + \exp(-C_2d^2n)).$$



## 3.5 Simulations

Dans cette section, nous simulerons la fonction de consommation keynésienne  $C_t$  qui représente la relation fonctionnelle entre la consommation totale et le revenu brut. Il est représenté par un modèle de régression linéaire simple

$$C_t = b + aR_t + \varepsilon_t \quad (3.5.1)$$

où le paramètre  $b$  est interprété comme la consommation incompressible (autonome), le coefficient de régression  $a$  de la variable de revenu  $R_t$  est interprété comme la propension marginale à consommer (PMC) et les erreurs  $\varepsilon_t$  représentent la partie inexpliquée du modèle.

Le but est de construire un domaine de confiance (ellipse de confiance) pour les paramètres  $a$  et  $b$  lorsque les erreurs  $\varepsilon_t$  sont  $\alpha$ -mélangeantes. Pour s'assurer que  $\varepsilon_t$  sont  $\alpha$ -mélangeantes, nous proposons un processus autorégressif d'ordre 1. (Il est bien connu que ce processus est  $\alpha$ -mélangeant [48]). En utilisant la méthode des moindres carrés, nous estimons  $a$  et  $b$  et à partir des résultats du Théorème 3.3.1, nous construisons des ellipses de confiance.

Compte tenu de la fonction de consommation keynésienne (3.5.1), nous construirons un domaine de confiance (ellipse de confiance) pour les paramètres  $a$  et  $b$  lorsque les erreurs

$\varepsilon_t$  sont  $\alpha$ -mélangeantes en utilisant les données du revenu mensuel moyen par habitant en

Afrique pendant 16 ans, exprimées en dollars, qui figurent dans le tableau 1

Année	Revenu disponible	Revenu centré $R_t$
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

Tableau 1: Évolution du revenu par habitant en Afrique

Source: Banque mondiale, 2019

Le PMC  $a = \frac{\Delta C_t}{\Delta R_t}$  est plus faible dans les pays développés, par exemple aux États-Unis ( $\simeq 0.04$ ), Le Royaume-Uni ( $\simeq 0.02$ ) et Canada ( $0,05 \simeq$ ) ; et plus élevé en Afrique comme au Nigeria ( $\simeq 0,64$ ), en Afrique du Sud ( $\simeq 0,7$ ) et en Zambie ( $\simeq 0,64$ ). Puisqu'une grande proportion d'Africains sont pauvres, nous avons supposé que  $a = 0,8$ . (Les méthodes de mesure du PMC peuvent être trouvées dans [41]). La consommation incompressible

(autonome)  $b$  est indépendante du revenu disponible. Il se situe généralement entre 30% et 40% de la moyenne. Dans notre contexte, il est supposé égal à 50.

Ainsi, la consommation théorique est égale à

$$C_t = 50 + 0.8R_t$$

et la consommation observée (générée)  $C_t$  est donc égale à

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

### 3.5.1 Simulation avec un processus $\alpha$ -mélangeant

Afin de caractériser les erreurs aléatoires  $\alpha$ -mélangeantes, nous considérons le modèle autorégressif du premier ordre

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

et  $(v_t)_t$  est une séquence i.i.d de variables aléatoires uniformes  $\mathcal{U}(-1, 1)$ .

(On génère le premier ordre autorégressif  $\varepsilon_t$  avec  $R$ ).

**Résultats de la simulation: cas de  $\alpha$ -mélange ( $n = 16$ )**

Année	Revenu disponible	Revenu centré $R_t$	Consommation théorique	Génération de $\varepsilon_t$	Consommation observée
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

Tableau 2 : Calcul de la consommation observée avec  $AR(1)$

Ainsi, nous obtenons un processus  $(\varepsilon_n)_n$   $\alpha$ -mélangeant avec l'ordre de mélange  $(0.04)^n$

[48].

Dans tout ce qui suit, nous considérons  $\varepsilon = 0.1$ . Il vérifie l'hypothèse  $0 < \varepsilon < \min_i \mathbb{E}\varepsilon_i^2$

( $d = \min_i \mathbb{E}\varepsilon_i^2 - \varepsilon > 0$ ). Aussi, nous prenons  $k = \lceil \frac{n}{2} - 1 \rceil$ .

D'après le Tableau 2, on obtient :  $\widehat{a}_n = 0.8041493$ ,  $\widehat{b}_n = 50.27538599$ ,  $S_n^2 = 0.507133094$

et

$$\begin{aligned} & \mathbb{P} \left\{ \frac{\left| \sum_{i=1}^{16} \varepsilon_i x_i \right|}{\left( \sum_{i=1}^{16} x_i^2 \right)^{\frac{1}{2}}} > r \right\} + \mathbb{P} \left\{ \frac{1}{\sqrt{16}} \left| \sum_{i=1}^{16} \varepsilon_i \right| > r \right\} + \mathbb{P} \left\{ \frac{1}{16} \sum_{i=1}^{16} \varepsilon_i^2 \leq \varepsilon \right\} \\ &= 1.0106 (0.98921 + 0.95741 + 0.96995) = 2.9475. \end{aligned}$$

Avec ces probabilités, nous ne pouvons pas conclure sur le niveau de confiance avec

lequel les paramètres  $a$  et  $b$  appartiennent à l'ellipse définie par l'équation :

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

### 3.5.2 Simulation avec un processus gaussien

**Remarque 3.5.1** Quand les erreurs sont gaussiennes *i.i.d*  $\mathcal{N}(0, 1)$ , on a

$$\frac{\sum_{i=1}^n \varepsilon_i}{\sqrt{n}} \rightsquigarrow \mathcal{N}(0, 1), \frac{\left| \sum_{i=1}^n x_i \varepsilon_i \right|}{\left( \sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}} \rightsquigarrow \mathcal{N}(0, 1) \text{ and } \sum_{i=1}^n \varepsilon_i^2 \rightsquigarrow \chi_n^2 \quad (3.5.2)$$

Ainsi, si on note  $\Phi$  la fonction de répartition de la variable aléatoire normale standard et

$F$  celle de  $\chi_{16}^2$ , alors de l'inégalité (3.3.1), la remarque (3.5.2) et de l'identité

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

on obtient

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

si  $\rho = 1$  (i.e.  $\xi = 0.025$ ) et  $n = 16$

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

Dans ce cas, nous ne parvenons pas non plus à construire des ellipses de confiance.

Mais d'autres choix de  $\rho$  donneront des résultats (voir (3.5.3)).

### Résultats de la simulation: cas gaussien ( $n = 16$ )

Année	Revenu disponible	Revenu centré $R_t$	Consommation théorique	Génération de $\mathcal{N}(0, 1)$	Consommation observée
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

Tableau 3: Calcul de la consommation observée avec  $\mathcal{N}(0,1)$

À partir du Tableau 3, nous obtenons :  $\hat{a}_n = 0,79$ ,  $\hat{b}_n = 49.72$ ,  $S_n^2 = 0.555$ . Ainsi, avec

une probabilité supérieure ou égale à  $1 - 4\Phi(-4\sqrt{\xi})$ , les paramètres  $a$  et  $b$  appartiennent

aux ellipses d'équations :

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

Voici quelques résultats numériques

$\rho$	Ellipse	niveau confiance
4 (i.e. $\xi = 0.04$ )	$\frac{(0.79-a)^2}{2.8774 \times 10^{-3}} + \frac{(49.72-b)^2}{2.22}$	15.26%
3 (i.e. $\xi = 0.0375$ )	$\frac{(0.79-a)^2}{2.1581 \times 10^{-3}} + \frac{(49.72-b)^2}{1.665}$	12.28%
2 (i.e. $\xi = \frac{1}{30}$ )	$\frac{(0.79-a)^2}{1.4387 \times 10^{-3}} + \frac{(49.72-b)^2}{1.11}$	6.96%

Tableau 4: Ellipses et niveau confiance ( $n = 16$ )

**Remarque 3.5.2** Pour  $\alpha$  et  $\rho$  donnés, par exemple  $\rho = 1$  (i.e.  $\xi = 0.025$ ), nous déterminons la taille minimale de l'échantillon qui nous permet d'atteindre le niveau de confiance donné en résolvant l'inégalité

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

Les différents résultats sont donnés dans le tableau ci-dessous

$\alpha$	0.01	0.05	0.1
$n$	316	201	19

### 3.5.3 Étude de simulation par la méthode Monte-Carlo

La taille de l'échantillon ( $n = 16$ ) est insuffisante pour construire des ellipses de confiance même dans le cas de variables aléatoires gaussiennes i.i.d  $\mathcal{N}(0, 1)$ . Pour surmonter cette lacune, nous utilisons la méthode de Monte-Carlo pour générer des échantillons de grande

taille ( $n = 2 \times 10^3, n = 5 \times 10^3, n = 7 \times 10^3, n = 9 \times 10^3, n = 10^4$  et  $n = 5 \times 10^4$ ). Pour ce faire, nous utilisons les données de la première colonne du Tableau 1.

Une analyse préliminaire des données, en utilisant la méthode de Cullen et Frey (voir graphique 1), nous permet de supposer que la distribution de probabilité de  $R_t$  est de Weibull avec deux paramètres. Ensuite, nous utilisons un tracé cdf (fonction `cdfcomp`), un tracé de densité (fonction `denscomp`), un tracé Q-Q de densité (fonction `qqcomp`), et un tracé P-P (fonction `ppcomp`) (voir graphique 2) et selon ce graphique, il est clair que nos données sont presque Weibull.

Enfin, pour confirmer notre affirmation, nous utiliserons le test d'ajustement de Kolmogorov-Smirnov. Après le calcul, nous avons trouvé une valeur  $p = 0,9427$  supérieure à  $0,05$ , ce qui signifie qu'elle n'est pas statistiquement significative et indique une preuve solide pour l'hypothèse nulle. Cela signifie que nous retenons l'hypothèse nulle et rejetons l'hypothèse alternative

1. Si l'on note la fonction de distribution de Weibull par

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

puis en utilisant le test de Kolmogorov-Smirnov, nous déterminons la distribution de l'échantillon donné dans la colonne 1, Tableau 1

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



2. Nous générons par la méthode Monte-Carlo  $n$  variables ( $n = 2 \times 10^3, n = 5 \times 10^3, n = 7 \times 10^3, n = 9 \times 10^3, n = 10^4$  et  $n = 5 \times 10^4$ ) de distribution Weibull  $\mathcal{W}(\sigma, \beta)$  en utilisant de la fonction cumulative inverse

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

où  $U$  a une distribution uniforme sur l'intervalle  $(0, 1)$ .

3. Nous déterminons

$$M = \frac{1}{n} \sum_{t=1}^n (R_t - \bar{R})^2$$

$$N = \max_t (R_t - \bar{R})^2$$

$$D = \max_r |R_t - \bar{R}|$$

4. Nous déterminons  $n$  variables  $\varepsilon_t$

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

5. Nous déterminons  $n$  variables  $C_t$

$$C_t = 50 + 0.8 (R_t - \bar{R}) + \varepsilon_t$$

et  $n$  variables  $y_t$

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

6. Nous déterminons  $\widehat{a}_n, \widehat{b}_n, S_n^2$  dans le cas  $\alpha$ -mélangeant et  $\widehat{\widehat{a}}_n, \widehat{\widehat{b}}_n, \widehat{\widehat{S}}_n^2$  dans le cas i.i.d

$\mathcal{N}(0, 1)$ .

### Simulation Monte-Carlo : cas $\alpha$ -mélangeant

A partir de (??) et de divers résultats donnés en annexe, on obtient

$n$	$2 \times 10^3$	$5 \times 10^3$	$7 \times 10^3$	$9 \times 10^3$	$10^4$	$5 \times 10^4$
$A_1$	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

Tableau 5 : Résultats de la simulation : cas  $\alpha$ -mélangeant

**Notation 3.5.1** Afin de simplifier l'écriture, notons par

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

1. Détermination du niveau de confiance  $(1 - \alpha)$  en pourcentage

Pour  $n$  et  $\xi$  donnés, ce qui est équivalent à  $n$  et  $\rho$  donnés par la relation

$$\rho = \frac{2\xi}{0.1 - 2\xi} \text{ avec } 0 < \xi < 0.05, \quad (3.5.4)$$

nous déterminons les différents niveaux de confiance correspondants en utilisant l'égalité

$$1 - \alpha = \left(1 - \left(\mathbb{P}_1^{(n)} + \mathbb{P}_2^{(n)} + \mathbb{P}_2^{(n)}\right)\right) * 100, n = 2 \times 10^3, \dots, n = 5 \times 10^4$$

$\xi \backslash n$	$2 \times 10^3$	$5 \times 10^3$	$7 \times 10^3$	$9 \times 10^3$	$10^4$	$5 \times 10^4$
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%

Tableau 6 : Niveau de confiance : cas  $\alpha$ -mélangeant

X représente le cas  $\mathbb{P}_1^{(n)} + \mathbb{P}_2^{(n)} + \mathbb{P}_2^{(n)} > 1$ .

**Remarque 3.5.3** Il est évident que pour différentes valeurs de  $\phi$ , on associe différentes

valeurs de probabilités  $\mathbb{P}_1^{(n)}, \mathbb{P}_2^{(n)}$  et  $\mathbb{P}_3^{(n)}$  et différents niveaux de confiance. Certaines

valeurs sont données dans le tableau suivant lorsque  $n = 2 \times 10^3$  et  $\xi = 0.049$

$\phi$	$\mathbb{P}_1^{(n=2000)}$	$\mathbb{P}_2^{(n=2000)}$	$\mathbb{P}_3^{(n=2000)}$	$\mathbb{P}_1^{(n=2000)} + \mathbb{P}_2^{(n=2000)} + \mathbb{P}_3^{(n=2000)}$ Niveau de confiance en %
0.05	0.66821	0.050444	0.941 01	1.6597 (sans intérêt)
<b>0.04</b>	<b>0.66116</b>	<b>0.03405</b>	<b>0.222 65</b>	<b>0.917 86 ; (8.214%)</b>
0.03	0.15072	0.020653	0.653 96	0.825 33 ; (17.467%)
0.02	0.07901	0.01070	0.646 62	0.736 33 ; (26.367%)
0.01	0.02208	0.004336	0.639 11	0.665 53 ; (33.447%)

Tableau 7 : Incidence de  $\phi$

## 2. Détermination de la taille minimale de l'échantillon

La taille des échantillons est déterminée en résolvant les équations

$$\mathbb{P}_1^{(n)} + \mathbb{P}_2^{(n)} + \mathbb{P}_2^{(n)} = \alpha$$

Pour  $\alpha$  et  $\xi$ . donnés. Cela donne les résultats suivants

$\xi \backslash \alpha$	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

Tableau 8 : Tailles des échantillons dans le cas  $\alpha$ -mélangeant

### 3. Détermination des foyers d'ellipse

Puisque  $M > 1$  (voir annexe), les foyers  $F_1$  et  $F_2$  des ellipses de confiance sont donnés

par

$$F_1 = \left( \hat{a}_n, \hat{b}_n - \sqrt{\rho S_n^2 \left(1 - \frac{1}{M}\right)} \right), F_2 = \left( \hat{a}_n, \hat{b}_n + \sqrt{\rho S_n^2 \left(1 - \frac{1}{M}\right)} \right).$$

$\hat{a}_n, \hat{b}_n$  et  $S_n^2$  sont regroupés dans le tableau 5 et sont déterminés à partir des équations

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

Pour  $\alpha$  et  $n$ , donnés, on trouve les foyers suivants

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

Tableau 9 : Foyers des ellipses dans le cas  $\alpha$ -demêlangeant

$X$  représente le cas  $\mathbb{P}_1^{(n)} + \mathbb{P}_2^{(n)} + \mathbb{P}_2^{(n)} > 1$ .

Notez que toutes ces ellipses sont verticales avec le centre  $(\widehat{a}_n, \widehat{b}_n)$ .

### 3.5.4 Simulation Monte-Carlo cas : gaussien

Détermination des foyers d'ellipse dans le cas gaussien

Depuis l'annexe, on peut lire les valeurs  $\widehat{a}_n, \widehat{b}_n$  et  $\widehat{S}_n^2$ . Elles sont notées dans l'annexe

comme aa\_n, bb\_n et SS\_n.

$n$	$2 \times 10^3$	$5 \times 10^3$	$7 \times 10^3$	$9 \times 10^3$	$10^4$	$5 \times 10^4$
$\widehat{a}_n$	0.7978058	0.7997622	0.7994124	0.800094	0.7996825	0.7998411
$\widehat{b}_n$	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

Tableau 10 : Résultats de simulation cas gaussien

Pour  $\alpha$  et  $n$ , donnés, on détermine  $\xi$  à partir de l'équation

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

Enfin, nous déterminons  $\rho$  à partir de la relation (3.5.4)

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

Tableau 11 : Différentes valeurs de  $\rho$  dans le cas gaussien

Pour  $\alpha$  et  $n$ , donnés, on trouve les foyers suivants

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

Tableau 12 : Foyers d'ellipse dans le cas gaussien

Notez que toutes ces ellipses sont verticales avec le centre  $(\widehat{a}_n, \widehat{b}_n)$ .

## 3.6 Annexe

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 <- fitdist(df$data, "weibull")

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 distribution for n = 2000

sigma <- 145.981173

beta <- 6.288061

n_2000 <- 2000

u_2000 <- runif(n_2000)

R_t_2000 <- sigma*(log(1/(1-u_2000)))^(1/beta)

fit.weibull_2000 <- fitdist(R_t_2000, "weibull")

xi_t_2000 <- rnorm(n_2000)

fit.weibull_2000$estimate

## shape scale

## 6.491915 147.065585

## Calculate of M, N, D

M <- 1/n_2000*(sum(round(R_t_2000-mean(R_t_2000), digits = 3)^2))
```

---

M

```
## [1] 593.6162
```

```
N <- max((R_t_2000-mean(R_t_2000))^2)
```

N

```
## [1] 8728.413
```

```
D <- max(abs(R_t_2000-mean(R_t_2000)))
```

D

```
## [1] 93.42597
```

```
## Simulation of v_t
```

```
v_t_2000 <- runif(n_2000+1, min = -1,
```

```
max = 1)
```

```
## epsilon_t
```

```
epsilon_t_2000 <- c()
```

```
epsilon_t_2000[1] <- 0
```

```
for (i in 2:1301) { epsilon_t_2000[i] <- 0.04*epsilon_t_2000[i-1]+v_t_2000[i]}
```

```
y_t_2000 <- 50 + 0.8*(R_t_2000 - mean(R_t_2000))+ xi_t_2000
```

```
## Determine C_t
```

```
C_t_2000 <- c()
```

---

```

for (i in 1:2000) { C_t_2000[i] <- 50 + 0.8*R_t_2000[i] + 0.04*epsilon_t_2000[i]+v_t_2000[i]}

## Detremine a_n, b_n, s2_n

a_n_2000 <- sum((R_t_2000 - mean(R_t_2000))*(C_t_2000 - mean(C_t_2000)))/sum((R_t_2000 -
mean(R_t_2000))^2)

a_n_2000

## [1] 0.7991316

b_n_2000 <- mean(C_t_2000) - a_n_2000*mean(R_t_2000)

b_n_2000

## [1] 50.13501

S2_n_2000 <- (1/n_2000)*sum(C_t_2000 - a_n_2000*(R_t_2000 - mean(R_t_2000))
- b_n_2000)^2

S2_n_2000

## [1] 15593677

aa_n_2000 <- sum(y_t_2000*(R_t_2000 - mean(R_t_2000)))/sum((R_t_2000 -
mean(R_t_2000))^2)

aa_n_2000

## [1] 0.7978058

bb_n_2000 <- mean(y_t_2000)

```

---

```
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)^2)

SS2_n_2000

## [1] 1.010926

# Monte Carlo simulation for weibull distribution for n = 5000

sigma <- 145.981173

beta <- 6.288061

n_5000 <- 5000

u_5000 <- runif(n_5000)

R_t_5000 <- sigma*(log(1/(1-u_5000)))^(1/beta)

fit.weibull_5000 <- fitdist(R_t_5000, "weibull")

xi_t_5000 <- rnorm(n_5000)

fit.weibull_5000$estimate

## shape scale

## 6.450028 146.362032

## Calculate of M, N, D
```

---

```
M <- 1/n_5000*(sum(round(R_t_5000-mean(R_t_5000), digits = 3)^2))

M

## [1] 611.2254

N <- max((R_t_5000-mean(R_t_5000))^2)

N

## [1] 9785.942

D <- max(abs(R_t_5000-mean(R_t_5000)))

D

## [1] 98.92392

## Simulation of v_t

v_t_5000 <- runif(n_5000+1, min = -1, max = 1)

## epsilon_t

epsilon_t_5000 <- c()

epsilon_t_5000[1] <- 0

for (i in 2:5001) {epsilon_t_5000[i] <- 0.04*epsilon_t_5000[i-1]+v_t_5000[i]}

y_t_5000 <- 50 + 0.8*(R_t_5000 - mean(R_t_5000))+ xi_t_5000

## Determine C_t

C_t_5000 <- c()
```

---

```

for (i in 1:5000) {C_t_5000[i] <- 50 + 0.8*R_t_5000[i] + 0.04*epsilon_t_5000[i]+v_t_5000[i]}

## Detremine a_n, b_n, s2_n

a_n_5000 <- sum((R_t_5000 - mean(R_t_5000))*(C_t_5000 - mean(C_t_5000)))/sum((R_t_5000 -
mean(R_t_5000))^2)

a_n_5000

## [1] 0.7999291

b_n_5000 <- mean(C_t_5000) - a_n_5000*mean(R_t_5000)

b_n_5000

## [1] 50.01532

S2_n_5000 <- (1/n_5000)*sum(C_t_5000 - a_n_5000*(R_t_5000 - mean(R_t_5000))
- b_n_5000)^2

S2_n_5000

## [1] 59455378

aa_n_5000 <- sum(y_t_5000*(R_t_5000 - mean(R_t_5000)))/sum((R_t_5000 -
mean(R_t_5000))^2)

aa_n_5000

## [1] 0.7997622

bb_n_5000 <- mean(y_t_5000)

```

---

```
bb_n_5000

## [1] 49.99761

SS2_n_5000 <- mean((y_t_5000 - aa_n_5000*(R_t_5000 - mean(R_t_5000))-
bb_n_5000)^2)

SS2_n_5000

## [1] 1.01104

# Monte Carlo simulation for weibull distribution for n = 7000

sigma <- 145.981173

beta <- 6.288061

n_7000 <- 7000

u_7000 <- runif(n_7000)

R_t_7000 <- sigma*(log(1/(1-u_7000)))^(1/beta)

fit.weibull_7000 <- fitdist(R_t_7000, "weibull")

xi_t_7000 <- rnorm(n_7000)

fit.weibull_7000$estimate

## shape scale

## 6.263327 146.161660

## Calculate of M, N, D
```

```
M <- 1/n_7000*(sum(round(R_t_7000-mean(R_t_7000), digits = 3)^2))
```

```
M
```

```
## [1] 639.5926
```

```
N <- max((R_t_7000-mean(R_t_7000))^2)
```

```
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 <- runif(n_7000+1, min = -1,
```

```
max = 1)
```

```
## epsilon_t
```

```
epsilon_t_7000 <- c()
```

```
epsilon_t_7000[1] <- 0
```

```
for (i in 2:7001) {
```

```
epsilon_t_7000[i] <- 0.04*epsilon_t_7000[i-1]+v_t_7000[i]
```

```
}
```



---

```

y_t_7000 <- 50 + 0.8*(R_t_7000 - mean(R_t_7000))+ xi_t_7000

## Determine C_t

C_t_7000 <- c()

for (i in 1:7000) {

C_t_7000[i] <- 50 + 0.8*R_t_7000[i] + 0.04*epsilon_t_7000[i]+v_t_7000[i]

}

## Determine a_n, b_n, s2_n

a_n_7000 <- sum((R_t_7000 - mean(R_t_7000))*(C_t_7000 - mean(C_t_7000)))/sum((R_t_7000 -
mean(R_t_7000))^2)

a_n_7000

## [1] 0.7996407

b_n_7000 <- mean(C_t_7000) - a_n_7000*mean(R_t_7000)

b_n_7000

## [1] 50.04262

S2_n_7000 <- (1/n_7000)*sum(C_t_7000 - a_n_7000*(R_t_7000 - mean(R_t_7000))
- b_n_7000)^2

S2_n_7000

## [1] 82701782

```

---

```
aa_n_7000 <- sum(y_t_7000*(R_t_7000 - mean(R_t_7000)))/sum((R_t_7000 -
mean(R_t_7000))^2)

aa_n_7000

## [1] 0.7994124

bb_n_7000 <- mean(y_t_7000)

bb_n_7000

## [1] 49.98879

SS2_n_7000 <- mean((y_t_7000 - aa_n_7000*(R_t_7000 - mean(R_t_7000))-
bb_n_7000)^2)

SS2_n_7000

## [1] 0.9847255

# Monte Carlo simulation for weibull distribution for n = 9000

sigma <- 145.981173

beta <- 6.288061

n_9000 <- 9000

u_9000 <- runif(n_9000)

R_t_9000 <- sigma*(log(1/(1-u_9000)))^(1/beta)

fit.weibull_9000 <- fitdist(R_t_9000, "weibull")
```

```
xi_t_9000 <- rnorm(n_9000)

fit.weibull_9000$estimate

## shape scale

## 6.350446 146.224440

## Calculate of M, N, D

M <- 1/n_9000*(sum(round(R_t_9000-mean(R_t_9000), digits = 3)^2))

M

## [1] 628.1908

N <- max((R_t_9000-mean(R_t_9000))^2)

N

## [1] 10261

D <- max(abs(R_t_9000-mean(R_t_9000)))

D

## [1] 101.2966

## Simulation of v_t

v_t_9000 <- runif(n_9000+1, min = -1,

max = 1)

## epsilon_t
```

---

```

epsilon_t_9000 <- c()

epsilon_t_9000[1] <- 0

for (i in 2:9001) {

epsilon_t_9000[i] <- 0.04*epsilon_t_9000[i-1]+v_t_9000[i]

}

y_t_9000 <- 50 + 0.8*(R_t_9000 - mean(R_t_9000))+ xi_t_9000

## Determine C_t

C_t_9000 <- c()

for (i in 1:9000) {

C_t_9000[i] <- 50 + 0.8*R_t_9000[i] + 0.04*epsilon_t_9000[i]+v_t_9000[i]

}

## Detremine a_n, b_n, s2_n

a_n_9000 <- sum((R_t_9000 - mean(R_t_9000))*(C_t_9000 - mean(C_t_9000)))/sum((R_t_9000 -
mean(R_t_9000))^2)

a_n_9000

## [1] 0.800236

b_n_9000 <- mean(C_t_9000) - a_n_9000*mean(R_t_9000)

b_n_9000

```

```
## [1] 49.97124

S2_n_9000 <- (1/n_9000)*sum(C_t_9000 - a_n_9000*(R_t_9000 - mean(R_t_9000))
- b_n_9000)^2

S2_n_9000

## [1] 106701955

aa_n_9000 <- sum(y_t_9000*(R_t_9000 - mean(R_t_9000)))/sum((R_t_9000 -
mean(R_t_9000))^2)

aa_n_9000

## [1] 0.800094

bb_n_9000 <- mean(y_t_9000)

bb_n_9000

## [1] 49.99492

SS2_n_9000 <- mean((y_t_9000 - aa_n_9000*(R_t_9000 - mean(R_t_9000))-
bb_n_9000)^2)

SS2_n_9000

## [1] 0.979984

# Monte Carlo simulation for weibull distribution for n = 10000

sigma <- 145.981173
```

---

```
beta <- 6.288061

n_10000 <- 10000

u_10000 <- runif(n_10000)

R_t_10000 <- sigma*(log(1/(1-u_10000)))^(1/beta)

fit.weibull_10000 <- fitdist(R_t_10000, "weibull")

xi_t_10000 <- rnorm(n_10000)

fit.weibull_10000$estimate

## shape scale

## 6.251053 145.916038

## Calculate of M, N, D

M <- 1/n_10000*(sum(round(R_t_10000-mean(R_t_10000), digits = 3)^2))

M

## [1] 643.3027

N <- max((R_t_10000-mean(R_t_10000))^2)

N

## [1] 10733.47

D <- max(abs(R_t_10000-mean(R_t_10000)))

D
```

```
## [1] 103.6025

## Simulation of v_t

v_t_10000 <- runif(n_10000+1, min = -1,
max = 1)

## epsilon_t

epsilon_t_10000 <- c()

epsilon_t_10000[1] <- 0

for (i in 2:10001) {

epsilon_t_10000[i] <- 0.04*epsilon_t_10000[i-1]+v_t_10000[i]

}

y_t_10000 <- 50 + 0.8*(R_t_10000 - mean(R_t_10000))+ xi_t_10000

## Determine C_t

C_t_10000 <- c()

for (i in 1:10000) {

C_t_10000[i] <- 50 + 0.8*R_t_10000[i] + 0.04*epsilon_t_10000[i]+v_t_10000[i]

}

## Determine a_n, b_n, s2_n
```

---

```

a_n_10000 <- sum((R_t_10000 - mean(R_t_10000))*(C_t_10000 - mean(C_t_10000)))/sum((R_t_
mean(R_t_10000))^2)

a_n_10000

## [1] 0.799813

b_n_10000 <- mean(C_t_10000) - a_n_10000*mean(R_t_10000)

b_n_10000

## [1] 50.02336

S2_n_10000 <- (1/n_10000)*sum(C_t_10000 - a_n_10000*(R_t_10000 - mean(R_t_10000))
- b_n_10000)^2

S2_n_10000

## [1] 117746349

aa_n_10000 <- sum(y_t_10000*(R_t_10000 - mean(R_t_10000)))/sum((R_t_10000
- mean(R_t_10000))^2)

aa_n_10000

## [1] 0.7996825

bb_n_10000 <- mean(y_t_10000)

bb_n_10000

## [1] 49.99433

```



---

```
SS2_n_10000 <- mean((y_t_10000 - aa_n_10000*(R_t_10000 - mean(R_t_10000))-
bb_n_10000)^2)

SS2_n_10000

## [1] 1.005974

# Monte Carlo simulation for weibull distribution for n = 50000

sigma <- 145.981173

beta <- 6.288061

n_50000 <- 50000

u_50000 <- runif(n_50000)

R_t_50000 <- sigma*(log(1/(1-u_50000)))^(1/beta)

fit.weibull_50000 <- fitdist(R_t_50000, "weibull")

xi_t_50000 <- rnorm(n_50000)

fit.weibull_50000$estimate

## shape scale

## 6.295943 146.041486

## Calculate of M, N, D

M <- 1/n_50000*(sum(round(R_t_50000-mean(R_t_50000), digits = 3)^2))

M
```

```
## [1] 634.5233

N <- max((R_t_50000-mean(R_t_50000))^2)

N

## [1] 11658.19

D <- max(abs(R_t_50000-mean(R_t_50000)))

D

## [1] 107.9731

## Simulation of v_t

v_t_50000 <- runif(n_50000+1, min = -1,
max = 1)

## epsilon_t

epsilon_t_50000 <- c()

epsilon_t_50000[1] <- 0

for (i in 2:50001) {

epsilon_t_50000[i] <- 0.04*epsilon_t_50000[i-1]+v_t_50000[i]

}

y_t_50000 <- 50 + 0.8*(R_t_50000 - mean(R_t_50000))+ xi_t_50000

## Determine C_t
```

---

```

C_t_50000 <- c()

for (i in 1:50000) {

C_t_50000[i] <- 50 + 0.8*R_t_50000[i] + 0.04*epsilon_t_50000[i]+v_t_50000[i]

}

## Detremine a_n, b_n, s2_n

a_n_50000 <- sum((R_t_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 <- mean(C_t_50000) - a_n_50000*mean(R_t_50000)

b_n_50000

## [1] 49.97342

S2_n_50000 <- (1/n_50000)*sum(C_t_50000 - a_n_50000*(R_t_50000 - mean(R_t_50000)) - b_n_50000)^2

S2_n_50000

## [1] 590781059

aa_n_50000 <- sum(y_t_50000*(R_t_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
```

# Conclusion-perspectives

Cette thèse s'inscrit dans le cadre de l'étude de problèmes inverses particuliers : Il s'agit de la calibration ou encore de la régression inverse. Elle est structurée en trois chapitres. Le premier porte sur la calibration non linéaire fonctionnelle. Le deuxième est consacré à la calibration non linéaire réelle et le dernier traite de la calibration affine.

Dans le chapitre un, nous avons considéré des variables aléatoires fonctionnelles (éléments aléatoires) indépendantes. Dans le second chapitre nous avons considéré le  $\psi$ -mélange et dans le dernier chapitre nous avons considéré la mélangeance forte.

Comme perspectives, il serait intéressant :

1. de construire des inégalités exponentielles dans le cas fonctionnel et de considérer les différentes formes d'indépendance asymptotique ;
2. Etudier d'autres formes de convergence ;
3. Généraliser le cas de la régression linéaire simple au cas d'une régression multivariée avec les différentes formes de mélangeance ;
4. Considérer les approximations stochastiques modifiées ;
5. Application de la procédure d'approximation stochastique à des cas réels.

---

# Bibliographie

- [1] Aiane N, Dahmani A. On the Rate of Convergence of the Robbins-Monro's Algorithm in a Linear Stochastic Ill-posed Problems with alpha-mixing Data. Communications in Statistics-Theory and Methods. Volume 46, issue 13, (2017), p. 6694-6703.
  
- [2] Andrews D.W.K. Non-strong mixing autoregressive processes, J. Appl. Probab, Volume 21, (1984), p. 930-934.
  
- [3] Arab I, Dahmani A . Consistency of stochastic approximation algorithm with quasi-associated random errors. Communications in Statistics-Theory and Methods. Volume 45, issue 23, (2016), p. 6883-6890.
  
- [4] Arab I, Dahmani A . Complete Convergence of Stochastic Approximation Algorithm in  $\mathbb{R}^d$  under Random Errors. Sequential Analysis: Design Methods and Applications. Volume 35, N°2, (2016), p. 216-225.
  
- [5] Arroudj H, Arab I, Dahmani A. Strongly mixed random errors in Mann's iteration algorithm for contractive real function, Neural, Parallel, and Scientific Computations,

- 
- Volume 25, (2017), p. 251-262.
- [6] Athreya K.B, Pantula S.G. A note on strong mixing of ARMA processes, *Statistics & Probability Letters*, Volum 4, issue 4, (1986), p. 187-190:
- [7] Babug J, Baiz. D. Edgeworth expansions for errors-in-variables models. *J. Multivariate Anal*, Volume 42, issue 2, (1992), p. 226-244
- [8] Blumj R, Hansond L, Koopmansl H. On the strong law of large numbers for a class of stochastic process. *Z. Wahrsch. Verw. Geb*, B.2, N°1, (1963), p. 1-11
- [9] Boente G .G, Fraiman .R , Robust nonparametric regression estimation for de pendent observations, *The annals of statistics*, Volume17, N°3, (1989), p. 1242-1256:
- [10] Bondarev B.V, Dahmani A . Exponential Bound in stochastic approximation procedures. 1990, p. 741-745 translated from *Ukainskii Matematicheskii Zhurnal* Volume 41, N°7, (1989), p. 867-872.
- [11] Bondarev B.V, Dahmani A. Estimates for the unknown parameters in recurrent stochastic procedures. (Russian) *Theory of random processes and its applications* (Russian), "Naukova Dumka", Kiev, (1990), p. 25-34,
- [12] Bosq D, Lecoutre J.P. *Théorie de l'estimation fonctionnelle*. *Economica*, (1987).

- [13] Caron E, Dedecker J, Michel B, Linear regression with stationary errors: the R package slm. 2019. hal-02157155v3
- [14] Carter R.A.L, Srivastava M.S, V.K. Srivastava, A. Ullah, Unbiased estimation of the MSE matrix of Stein-rule estimators, confidence ellipsoids, and hypothesis testing, *Econometric Theory*, N°6 (1990) p. 63-74.
- [15] CHaig X, Liz Y, Tian H . Consistent nonparametric estimation of error distributions linear model. *Acta Math. Appl. Sinica*, Volum 7, N°3, (1991), p. 245-256.
- [16] Chen X .R, Wu Y.H , Strong law for mixing sequence, *Acta Mathematicae Applicatae Sinica*, N°5, (1989), p. 367-371:
- [17] Chaturvedia A, Guptaa S, Bhattib M.I. Confidence ellipsoids based on a general family of shrinkage estimators for a linear model with non-spherical disturbances, *Journal of Multivariate Analysis*, Volum 104,N°1, (2012), p. 140-158.
- [18] Dahmani A, Belaide K. Exponential Inequalities in Calibration Problems with Gaussians Errors. *Communications in Statistics-Theory and Methods*, Volume 42,issue 19, (2013), p. 3596-3607.
- [19] Dahmani A, Belaide K. Exponential Inequalities in Stochastic Inverse Problems Using an Iterative Method, *Communications in Statistics - Theory and Methods*,Volume



- 44, issue 11, (2015), p. 2385-2397,
- [20] Dahmani A, Zerouti H. Exponential inequalities in linear calibration problem. *Communications in Statistics-Theory and Methods*, Volume 45, Issue 18, (2014), p. 5251-5262 .
- [21] Dahmani A, Bengrina M.H, AIT SAIDI A. Bernstein-Frechet inequalities for the Robbins Monro's Algorithm. *International Journal of Applied Mathematics and Statistics..* Volume 3, issue 5, (2005), p 17-24.
- [22] Dahmani A, Ait Saidi A . Consistency of Robbins Monro's algorithm within a mixing framework. *Bulletin de la Société Royale des Sciences de Liège*, Volume 79, (2010), p. 131-140.
- [23] Deo C.M. A note on empirical processes of strong-mixing sequences, *Ann.Probab*, Volum 1, issue 5, (1973), p.870-875.
- [24] Deville J.C, Särndal C.E. Calibration Estimators in Survey Sampling, *Journal of the American Statistical Association*, Volume 87, N°418, (1992), p. 376-382.
- [25] Ding F . Decomposition based fast least squares algorithm for output error systems. *Signal Processing*, Volum 93, issue 5, (2013), p. 1235-1242.

- [26] Ding F. Hierarchical estimation algorithms for multivariable systems using measurement information. *Information Sciences*, Volum 277, (2014), p.396-405.
- [27] Ding F. State filtering and parameter estimation for state space systems with scarce measurements. *Signal Processing*, Volum 104, (2014), p. 369-380.
- [28] Ding F, Wang Y. J, Ding J. Recursive least squares parameter estimation algorithms for systems with colored noise using the filtering technique. *Digit. Signal Process*, Volum 37, (2015), p. 100-108.
- [29] Ding J, Fan C, Lin J. Auxiliary model based parameter estimation for dual-rate output error systems with colored noise. *Applied Mathematical Modelling*, Volum 37, issue 6, (2013), p. 4051-4058.
- [30] Ding J, Lin J. Modified subspace identification for periodically non-uniformly sampled systems by using the lifting technique. *Circuits, Systems, and Signal Processing*, Volum 33, issue 5, (2014), p. 1439-1449.
- [31] Doukhan P. *Mixing : Properties and Examples*. Lecture Notes in Statist. N°85, Springer Verlag, 1994.
- [32] Dorogovtsev A.Ya. Consistency of the least squares estimator of an infinite-dimensional parameter. *Sibirsk. Mat. Zh*, Volum 33, N°4, (1992), p. 65-69. Transla-

- tion in *Siberian Math J.* , N°4, (1992), p. 603-607.
- [33] Durand D. Joint Confidence Regions for Multiple Regression Coefficients, *Journal of the American Statistical Association*, Volum 49, N°265, (1954), p. 130-146
- [34] Esary J, Proschan F. Walkup D. Association of random variables with application. *Ann. Math. Statist*, N°38, (1967), p. 1466-1476.
- [35] Ferraty F, Vieu P. *Nonparametric Functional Data Analysis*, Springer, 2006:
- [36] Ferraty F, Laksaci A, Tadj A, Vieu P. Kernel regression with functional response. *Electronic Journal of Statistics*. Volum 5, (2011), p. 159-171.
- [37] Fisher R.A, On the mathematical foundations of theoretical Statistics, *Phil. Trans. R. Soc. Lond. A* Volum 222, (1922), p. 309-368.
- [38] Fortuin C, Kastelyn P, Ginibre J. Correlation inequalities on some ordered sets. *Comm. Math. Phys*, N°22,(1971) p. 89-103.
- [39] Galton F, Regression towards mediocrity in hereditary stature. *Journal of Anrthro-pological Institute*, Volum 15, (1886) , p. 246-263.
- [40] Guo H, Liu Y , Regression estimation under strong mixing data, *Annals of the Institute of Statistical Mathematics*, Springer, volum 71, N°3, (2019), p. 553-576

- [41] Haavelmo T, Methods of Measuring the Marginal Propensity to Consume, Journal of the American Statistical Association, Volum 42, N°237, (1947), p. 105-122:
- [42] Hipp.C. Convergence rates of the strong law for stationary mixing sequences. Z. Wahrsch. Verw. Geb, N°49, (1979), p. 49-62.
- [43] Ibraguimov A, Linnik Yu V. Variables aléatoires indépendantes et stationnairement liées. Moscou, Naouka, 1965. (en russe).
- [44] Kolmogorov A .N, Yu A. Rozanov, On Strong Mixing Conditions for Stationary Gaussian Processes. Theory Probab. Appl, Volum 5, issue2, (1960), p. 204-208:
- [45] Kulkarni S.R, Horn C.S. An alternative proof for Convergence of stochastic approximation algorithms. IEEE Transactions on Automatic Control, Volum 41, N°3, (1996), p. 419-424.
- [46] Kulkarni S.R, Horn C.S. Necessary and sufficient conditions for convergence of stochastic approximation algorithms under arbitrary disturbances. Proceedings of the 34th Conference on Decision & Control. New Orleans, LA-December, Volum 4, (1995), p. 3843-3848.
- [47] Lee C, Adkins C, Carter Hill, R. An improved confidence ellipsoid for the linear regression model, Journal of Statistical Computation and Simulation, Volum 36,

- issue 1, (1990), p. 9-18.
- [48] Lee S , A note on the strong mixing property for random coefficient autoregressive process, Journal of Korean Statistical Society, volum 24, N°1, (1995), p. 243-248:
- [49] Lehüanne L . Some concepts of dependence. Ann. Math. Statist, N°37, (1966); p. 1137-1153.
- [50] Liland KH, Smilde. A, Marini. F, Næs. T, Confidence ellipsoids for ASCA models based on multivariate regression theory, Journal of Chemometrics, Volum 32, issue 5, (2018) ;e2990. <https://doi.org/10.1002/cem.2990>
- [51] Max P. L'approximation et l'optimisation stochastiques en économie : deux exemples. Revue économique, volume 24, N° 1, (1973) p.176-182.
- [52] MaoucheA F, Dahmani A, Rahmania N. stochastic procedure to solve linear ill-posed problems. Communications in Statistics - Theory and Methods, Volume 46, issue 3, (2017), p. 1519-1531.
- [53] Olive D.J, Linear Regression, Springer, (2017)
- [54] Ouvrard J.Y. Probabilités 2. Cassini.
- [55] Pearson K, Lee A. On the laws of inheritance in man. Inheritance of physical characters, Biometrika, Volum 2, issue 4, (1903) , p. 357-462.

- [56] Philipp W. A functional law of the iterated logarithm for empirical distribution functions of weakly dependent random variables, *Ann. Probab*, Volum 5, (1977), p. 319-350.
- [57] Pinelis I.F, Sakhanenko A.I. Remarks on Inequalities for Large Deviation Probabilities. *Theory of Probability & Its Applications*, Volum 30, N°1, (1986) , p. 143-14
- [58] Rahmani S, Dahmani A. Exponential Inequalities for the Robbins-Monro's Algorithm Under Mixing Condition. *Communications in Statistics - Theory and Methods*, Volume 45, issue 2, (2016) p. 520-528.
- [59] RIO E. Théorie asymptotique des processus faiblement dépendants. *Mathématiques et applications*. 1999, 31, Springer, SMAI.
- [60] Robbins H, Monro S. A stochastic approximation method. *Ann. Math. Statist.* Volum 22, (1951), p. 400-407.
- [61] Rosenblatt M. A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. USA*, Volum 42, (1956), p. 43-47.
- [62] Roy S.N, Bose R .C, Simultaneous Confidence Interval Estimation, *Annals of Mathematical Statistics*, Volum 24, (1953), p. 513-536.

- [63] Salov G.I. On a stochastic approximation theorem in a Hilbert space and its applications. *Theory of Probability and its Applications*, Volum 24, N° 2, (1980), p. 413-419.
- [64] Sarndal C.E, Swensson B, Wretman J. *Model Assisted Survey Sampling*. Springer Series in Statistics.(1992)
- [65] Suquet C . *Simulation*, Université des Sciences et Technologies de Lille, Agrégation Externe 2007-2008.
- [66] Tingley M.A. Small-sample intervals for regression. *Canad. J. Statist*, Volum 20, N°3, (1992), p. 271-280.
- [67] Vasan M.T. *Stochastic Approximation*, Cambridge University Press; Revised ed. edition (June 3, 2004).
- [68] Volkonskii V.A, Roznov Yu. A. Quelques théorèmes limites pour les fonctions aléatoires. *Théorie des Probabilités et son Application*. . T. IV, N° 2, (1959), p. 186-207.
- [69] Wang C, Tang T. Recursive least squares estimation algorithm applied to a class of linear-in-parameters output error moving average systems. *Applied Mathematics Letters*, Volum 29, (2014), p. 36-41.

- [70] Wang C, Tang T. Several gradient-based iterative estimation algorithms for a class of nonlinear systems using the filtering technique. *Nonlinear Dynamics*, Volum 77, (2014), p. 769-780.
- [71] . Welsch R . E. Limit Laws for Extreme Order Statistics Form Strong-Mixing Processes, *Ann. Math. Statist*, Volum 43, N° 2, (1972), p. 439-446:
- [72] Yule G.U. On the theory of correlation. *Journal of the Royal Statistical Society*, Volum 60, (1897), p. 812-854.
- [73] Zvara bK. Consistency of an estimate in linear regression with nonnegative errorsKybernetika (Prague), Volum 28, N°2, (1992), p. 129-139



# **Iterative algorithms for solving non-linear calibration problems**

**Abdelkader ELMOUMEN**

The main theme of this thesis is the study of regression taken from its global definition, that is, a type of statistical model. It is a matter of representing a reality of a random nature and, through the confrontation of observation and the model, of deducing the laws, or certain elements of the laws that govern this reality.

The analysis of the results of an experiment involves the use of a regression model when the observations resulting from it can be represented, each, as the sum of a systematic term, depending on the value taken by one or more other variables, and the creation of a random variable.

This thesis consists essentially of three parts. In the first part a non-linear calibration problem, also called inverse regression, was dealt with using the Robbins-Monro stochastic approximation procedure. This involves solving an operative equation where the operator is non-linear, bounded and defined on a separable Hilbert space with values in itself. We are in the case where the operator is known only at certain points.

In the second part of this thesis, we constructed exponential inequalities for the Robbins-Monro algorithm by considering  $\psi$ -mixing measurement errors. These inequalities have allowed us to establish the almost complete convergence as well as the construction of a domain of trust for the root of the function considered.

In the third part, we established exponential inequalities for the parameters of a simple linear regression model when measurement errors are  $\alpha$ -mixing. These inequalities have allowed us to build regions of confidence, characterized by ellipses, for the parameters of this model.

To validate our theoretical results, we presented a numerical study whose calculations and simulations are performed under R.

At the end of the thesis, we drew up some perspectives and a bibliographic list.