الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي و البحث العلمي

Université Abou Bekr Belkaid
Tlemcen Algérie
جامعة أبي بكر بلقايد
تلمسان الجزائر

FACULTÉ DES SCIENCES

DÉPARTEMENT DE MATHÉMATIQUES

MEMOIRE DE MASTER
EN MATHEMATIQUES

Option :Probabilité et statistique

Sujet :

Processus Moyennes Mobiles .Théorèmes limites et estimation

Candidat(e) : **Bouzidi Soujoud**

Date : 03/10/2019

Membres du Jury :

Président :	Mr ALLAM Abdelaziz,	maître de conférences, Université Abou Bakr Belkaid de Tlemcen
Examinateur :	Mme BOUKHIAR Souad,	maître de conférences , Université Abou Bakr Belkaid de Tlemcen
Encadreur :	Mr LABBAS Ahmed ,	maître de conférences, Université Abou Bakr Belkaid de Tlemcen

Année Universitaire 2018/2019

Dédicaces

A mes chers parents.
A mes chères sœurs Mawhouba et Ikram.
A mes chers tantes .
A mes chers amis .
Ainsi qu'à mon cher ancle Morad et à ma tante Rachida.

Remerciements

Avant toute chose, je tiens à remercier Dieu le tout puissant, qui m'a toujours guidée dans tout ce que j'ai entrepris de faire.

J'exprime mes profonds remerciements, ma vive reconnaissance et ma sinsère gratitude à **Mr. A. LABBAS**, maitre de conférences à la faculté des Sciences,de l' Universsité Abou Bekr BelKaid pour avoir accepté de m'encadrer et pour ses conseils et ses précieuses orientations qu'il n'a cessé de m'apporter tout au long de ce travail.

J'adresse mes sinsères remerciements à **Mr. A. ALLAM** professeur à la faculté des sciences, Universsité Abou Bekr BelKaid pour l'honneur qu'il me fait de présider le jury de ce mémoire.

Je remercie également **Mme.S.BOUKHIAR** maitre de conférences à la faculté des Sciences, Universsité Abou Bekr BelKaid de m'avoir fait l'honneur d'accepter d'examiner ma thèse de Master.

Mes remerciements vont aussi à tous mes professeurs qui ont contribué d'une manière ou d'une autreà l'accomplissement de ce travail.

Plus que quiconque, il me faut remercier ma chère mère qui m'a apporté son appui durant toutes mes années d'études, pour son sacrifice et son soutien qui m'ont donné confiance, courage et sécurité.

Je remercie mon cher père qui m'a appris le sens de la persévérance tout au long de mes études, ses conseils et ses encouragements.

Mes sentiments de reconnaissance et mes remerciements vont également à tous mes amis pour les sympathiques moments que nous avons passés ensemble.

Enfin je remercie gracieusement toute personne qui a contribué de près ou de loin à la réalisation de ce travail.

Tlemcen, le 9 octobre 2019.

Table des matières

Introduction

Depuis la nuit des temps,l'homme a toujours voulu prédire l'avenir pour satisfaire sa curiosité . En effet, il s'avère que les meilleurs décisions ont été prises suite aux inventions et expériences humaines de l'histoire , même s'il est parfois difficile de prédire à juste titre si telle ou telle nouvelle invention fera un succès ou un échec .
Effectivement, une méthode de prévision populaire basée sur une étude rigoureuse a pu donner une découverte de séries chronologiques $(X_t)_{t\in\mathbb{Z}}$ qui sont des suites de variables aléatoires indexées par le temps.
Cette approche permet de prédire par exemple de nombreux phénomènes naturels et financiers telle que une série chronologique qui est constituée de valeurs observées à des intervalles réguliers de temps . À titre d' exemple , les débits annuels sur un cours d'eau ou encore les valeurs mensuelles de titres boursiers sont des séries chronologiques .
À la base , l'étude formelle des séries chronologiques consiste à trouver un modèle mathématique qui explique le mieux possible les données observées .
Dans ce mémoire ,nous nous intéressons particulièrement aux processus moyennes mobiles réelles d'ordre fini q notées par $MMR(q)$.
Notre mémoire est composé de trois chapitres.

Dans le premier chapitre on étudie le processus moyenne mobile et les différents caractéristiques de celui-ci.

Dans le second chapitre , nous étudions l'approche asymptotique des processus moyennes mobiles ,en commençant par développer la loi des grands nombres pour les moyennes mobiles d'ordre infini , puis nous développons le théorème Central Limite pour les processus stationnaires m-dépendants .Enfin, nous traitons un corollaire qui représente une application de théorème central limite sur un $MMR(1)$.

Le troisième et le dernier chapitre englobera les différentes méthodes d'estimation du paramétre de processus moyenne mobile d'ordre 1 .

Chapitre introductif 1

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, sur lequel nous considérerons un processus stochastique $(X_t)_{t\in T}$ avec $T = \mathbb{N}$ ou $\mathbb{Z}$. Sachant qu'un processus stochastique est un ensemble de variables aléatoires qui sont ordonnées dans le temps. On définit deux types de processus stochastiques :

1. à temps continu : $X(t)$
2. à temps discret : $X_t : t = 0, \pm 1, \pm 2, \pm 3, \dots$

Évidemment, pour modéliser et analyser les séries temporelles on utilise les processus stochastiques à temps discret.

Dans toute la suite $\mathcal{L}(X)$ désigne la loi de la variable aléatoire X.
On note

$$\mathbb{R}^T = \{x = (x_t)_{t\in T} / x_t \in \mathbb{R}, \forall t \in T\}.$$

On appelle tribu produit sur $\mathbb{R}^T$ la plus petite tribu rendant mesurables les applications coordonnées :

lorsque t parcours T, on la note $\mathcal{B}_{\mathbb{R}^T}$.
Soit maintenant l'application :

$$\begin{array}{rccc} X: & \Omega & \to & \mathbb{R}^T \\ & \omega & \mapsto & X_t(\omega) \end{array}$$

On a alors $(X_t)_{t\in T}$ est un processus stochastique à valeurs dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ si, et seulement si, l'application X est mesurable de $(\Omega, \mathcal{F}$ dans $(\mathbb{R}^T, \mathcal{B}_{\mathbb{R}^T})$. On peut identifier donc un processus stochastique $(X_t)_{t\in T}$ et une application mesurable X. Ainsi, la loi d'un processus stochastique $(X_t)_{t\in T}$, la probabilité image de $\mathbb{P}$ par X, sont notées $\mathbb{P}_X$, de sorte que si $A = \mathbb{R} \times ...\mathbb{R} \times A_i \times A_{i+1} \times ... \times A_{i+k} \times \mathbb{R} \times \mathbb{R} \times ... \in \mathcal{B}_{\mathbb{R}^T}$, avec $i \in T$ et $k \geq 0$ alors

$$\mathbb{P}_X(A) = \mathbb{P}(X_i \in A_i, X_{i+1} \in A_{i+1}, ..., X_{i+k} \in A_{i+k}).$$

La loi $\mathbb{P}_X$ est caractérisée donc par les lois $\mathcal{L}(X_i, X_{i+1}, ..., X_{i+k})$ appelées les distributions fini-dimensionnelles du processus.

1.1 Processus linéaires et moyennes mobiles :

Nous définissons tout d'abord les premiers moments d'un processus

Définition 1.1.1 *La moyenne du processus $X = \{X_t\}_{t\in\mathbb{Z}}$ notée $(\mu_t)_{t\in\mathbb{Z}}$ est définie par $\mu_t = \mathbb{E}(X_t)$,$\forall t \in \mathbb{Z}$*
On dit que le processus est centré , si $\mu_t = \mu = 0, \forall t \in \mathbb{Z}$

Remarque 1.1.1 *Si $X = \{X_t\}_{t\in\mathbb{Z}}$ est à valeur dans $\mathbb{R}^m$ sa moyenne noteée μ_t est le vecteur de $\mathbb{R}^m$ donné par :*

$$\mathbb{E}(X_t) = \begin{pmatrix} \mathbb{E}(X_{1t}) \\ \mathbb{E}(X_{2t}) \\ \\ \\ \mathbb{E}(X_{mt}) \end{pmatrix} = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \\ \\ \\ \mu_{mt} \end{pmatrix} = \mu_t \tag{1.1}$$

1.1.1 Stationarité du processus :

Un processus peut être strictement stationnaire ou il peut être caractérisé par la stationnarité faible (du deuxième ordre).

Définition 1.1.2 *Un processus $X = \{X_t\}_{t\in\mathbb{Z}}$ est du second ordre si*

$$\mathbb{E}(X_t^2) < \infty$$

de plus, on dira que ce processus est faiblement stationnaire (ou stationnaire d'ordre 2) si les deux conditions suivantes sont vérifiées pour tout $t, h \in \mathbb{Z}$:

(i) $\mathbb{E}(X_t) = \mu$ *pour tout* $t \in \mathbb{Z}$

(ii) $\mathbb{E}(X_t X_t + h) = \gamma(h)$ *est indépendant de* t

Définition 1.1.3 *1. $\{X_t, t \in \mathbb{Z}\}$ est dit strictement stationnaire si pour toute suite $(t_1, t_2, ..., t_n) \in \mathbb{Z}$ avec $t_1 < t_2 < ... < t_n$ et $h \in \mathbb{Z}$*

$$\mathcal{L}(X_{t_1+h}, X_{t_2+h}, ..., X_{t_n+h}) = \mathcal{L}(X_{t_1}, X_{t_2}, ..., X_{t_n}).$$

1.1.2 processus bruit blanc et moyenne mobile :

Définition 1.1.4 *Un processus aléatoire $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ est dit Bruit blanc réel faible noté $BBR(0, \sigma_\varepsilon^2)$ si*

(1) $\mathbb{E}(\varepsilon_t) = 0$

(2) $\mathbb{E}(\varepsilon_{t_1}\varepsilon_{t_2}) = \sigma_\varepsilon^2 \delta_{t_1 t_2}$

Doú :

$$\delta_{t_1 t_2} = \begin{cases} 1 & si\ t_1 = t_2 \\ 0 & sinon \end{cases}$$

Si les ε_t sont iid (indépendnts identiquement distribuée) , alors il est dit bruit blanc fort et il est notée par $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$

Exemple 1.1.1 *l'exemple le plus connu d'un processus stationnaire est le processus bruit blanc.*
car

$$\begin{cases} \mathbb{E}(\varepsilon_t) = 0 & \forall t \in T \\ \gamma(h) = \mathbb{E}(\varepsilon_t \varepsilon_{t-h}) \end{cases}$$

avec

$$\gamma(h) = \begin{cases} \sigma_\varepsilon^2 & h = 0 \\ 0 & h \neq 0 \end{cases}$$

Définition 1.1.5 *Soit $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ un bruit blanc fort .*
Un processus $X = (X_t, t \in \mathbb{Z})$ est dit un processus linéaire s'il vérifie l'equation suivante :

$$X_t = \sum_{i \geq 0} \theta_i \varepsilon_{t-i}$$

avec $\sum\limits_{i \geq 0} \theta_i^2 < \infty$

Remarque 1.1.2
1. *Dans la définition (1.1.5) l'egalité est prise au sens de L^2et comme les ε_t sont indépendants alors c'est même $P - p.s$*
2. *X ainsi défini est strictement stationnaire .*

Proposition 1.1.1 *si $\{X_t\}_{t \in \mathbb{Z}}$ processus stationnaire et si $(a_i, i \in \mathbb{Z})$ une suite de nombre réels absolument sommable $(\sum \mid a_i \mid < \infty)$*
alors $Z_t = \sum a_i X_{t-i}$ $t \in \mathbb{Z}$ est un nouveau processus stationnaire.

Définition 1.1.6 *Un processus $X = (X_t, t \in \mathbb{Z})$ moyenne mobile réelle d'ordre q noté $MMR(q)$ est défini par*

$$X_t = \mu_t + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{1.2}$$

où $\varepsilon = (\varepsilon_t\ ,\ t \in \mathbb{Z})$ est un $BBR(0, \sigma_\varepsilon^2)$ et $\mu_t = \mathbb{E}(X_t)$
et $\forall j = 1, \ldots, q$ on a $\theta_j \in \mathbb{R}$ avec $\theta_q \neq 0$

Le processus défini précédemment peut s'écrire sous une autre forme .
Si B est l'opérateur retard tel que $B^j(X_n) = X_{n-j}$ pour tout $j = 1, 2, \ldots,$ alors on note por tout$t \in \mathbb{Z}$:

$$X_t = \Theta(B) \varepsilon_t$$

où $\Theta(B) = 1 + \theta_1 B^1 + \theta_2 B^2 + \cdots + \theta_q B^q$

1.2 Définitions et propriétées :

Définition 1.2.1 (Inversibilité)
Un processus MMR(q) défini pour tout $t \in \mathbb{Z}$ par l'équation $X_t = \Theta(B)\varepsilon_t$ est inversible s'il existe une suite sommable de réels$\{\alpha_j\}_{j\in\mathbb{N}}$ i.e $\sum\limits_{j=0}^{\infty} |\alpha_j| < \infty$ telle que :

$$\varepsilon_t = \sum_{j=0}^{\infty} \alpha_j X_{t-j}$$

Remarque préliminaire :"inversion d'un polynôme"
Soit le polynôme $1 - aB$
Calculons son inverse $(1 - aB)^{-1}$ on a les deux cas suivants :

1. si $|a| < 1$: donc si la racine de $1 - az = 0$ est supérieur à 1 en module $|z| = \dfrac{1}{|a|} > 1$ alors l'inverse est donnée par :

$$(1 - aB)^{-1} = \sum_{i=0}^{\infty} a^i B^i$$

$$\begin{aligned}(1 + aB)^{-1}(1 - aB) &= \left(\sum_{i=0}^{\infty} a^i B^i\right)(1 - aB)\\ &= \sum_{i=0}^{\infty} a^i B^i - a^{i+1}B^{i+1}\\ &= B^0 = 1\end{aligned}$$

   $(1 - aB)$ est donc inversible et son inverse $(1 + aB)^{-1} = \sum_{i=1}^{\infty} a^i B^i$.
2. si $|a| > 1$: donc si la racine de $1 - az = 0$ est inférieur à 1 en valeur absolue$|z| = \dfrac{1}{|a|} < 1$ alors l'inverse est donnée par :

$$(1 - aB)^{-1} = -\sum_{i=1}^{\infty} a^{-i}B^{-i}$$

3. Si $|a| = 1$, $(1 + aB)$ n'est pas inversible, si par exemple $a = 1$, tout processus constant est *annulé* par l'opérateur $1 - B$, l'application $1 - B$ n'est donc pas injective.
   Plus généralement, si l'on considère un polynôme :

$$\Theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q$$

   de racines $z_j = \frac{1}{a_j}$ de module supérieurs à 1.
   Alors on sait qu'il existe une séries entière $\sum_{i\in\mathbb{Z}} \Psi_i z^i := \Psi(z)$, telle que $\sum_{i=0}^{\infty} |\Psi_i| < \infty$ et $\Theta(z)\Psi(z) = 1$.
   En appliquant ce résultat aux séries en B, $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ est inversible et son inverse est $\Psi(B)$.

Remarque 1.2.1 *il existe autre manière de vérifier l'inversibilité d'un MMR*(q) *est de voir si les racines du polynôme* $\Theta(B)$ *sont en module supérieur à* 1

Définition 1.2.2 *(m-dépendance)*
Une suite strictement stationnaire (X_n) *de variables aléatoires de* $\mathbb{R}$ *est dite m-dépendante, avec* $m \in \mathbb{N}^*$ *si pour tout les ensembles* $\{X_j, j \leqslant t\}$ *et* $\{X_j, j \geqslant t+m+1\}$ *sont indépendants .*
La m-dépendance se traduit par l'indépendance des observations dès qu'elles sont séparées de $m+1$ *unités de temps .*

Proposition 1.2.1 *Un processus MMR*(q) *est un processus q-dépendant.*

Définition 1.2.3 *(la fonction d'autocovariance)*
La fonction d'autocovariance du processus $X = (X_t, t \in \mathbb{Z})$ *est dfinie par :*

$$\gamma(s,t) = Cov(X_s, X_t) = \mathbb{E}[(X_s - \mathbb{E}(X_s))(X_t - \mathbb{E}(X_t))] \quad \forall s,t \in \mathbb{Z}$$

Pour un processus stationnaire , la fonction d'autocovariance définie précédement s'écrit comme fonction à une seule variable comme suit :

$$\forall t,h \in \mathbb{Z} \quad \gamma(t,t+h) = \gamma(h,0) = \gamma(h)$$

Proposition 1.2.2 *la fonction d'autocovariance d'un processus stationnaire centré est une fonction :*

1. $\gamma(0) = Var(X_t)$ *et* $|\gamma(h)| < \gamma(0)$
2. *paire :*$\gamma(-h) = \gamma(h)\ \forall h$
3. *définie semi-positive : pour tout* $a = (a_1, \ldots, a_p) \in \mathbb{R}^p ; t = (t_1, \ldots, t_p) \in \mathbb{Z}^p$ *on a*

$$\sum_{i,j=1}^{p} a_i \gamma(t_i - t_j) a_j \geq 0$$

puisque cette quantité est égal à $V(\sum_{i=1}^{p} a_i X_{t_i})$

Démonstration. On suppose que $\mathbb{E}(X_t) = 0$

1. $\gamma(0) = Var(X_t) = \mathbb{E}(X_t^2) \geq 0$
   par l'inégalité de *Cauchy-Shwartz* :

$$\begin{aligned}|\gamma(h)| &= |\mathbb{E}(X_t X_{t+h})| \\ &\leq \left(\mathbb{E}(X_t^2)\right)^{\frac{1}{2}} \left(\mathbb{E}(X_{t+h}^2)\right)^{\frac{1}{2}} \\ &= \mathbb{E}(X_t^2) \\ &= \gamma(0).\end{aligned}$$

   Car on a $((X_t)_t)_{t \in \mathbb{Z}}$ est un processus stationnaire, alors

$$\mathbb{E}(X_t^2) = \mathbb{E}(X_{t+h}^2).$$

2. $\gamma(h) = Cov((X_t, X_{t+h})$, puisque $t = s - h \Rightarrow t + h = s$
   Donc $\gamma(h) = Cov((X_{s-h}, X_s) = \gamma(-h)$.
3. D'abord $V\left(\sum_{i=1}^{p} a_i X_{t_i}\right) \geq 0$, $\forall a = (a_1, a_2, \ldots, a_p) \in \mathbb{R}^p \quad t = (t_1, t_2, \ldots, t_p) \in \mathbb{R}^p$
   et

$$\begin{aligned} V\left(\sum_{i=1}^{p} a_i X_{t_i}\right) &= \mathbb{E}\left(\sum_{i=1}^{p} a_i X_{t_i} \sum_{j=1}^{p} a_j X_{t_j}\right) \\ &= \mathbb{E}\left(\sum_{i=1}^{p}\sum_{j=1}^{p} a_i a_j X_{t_i} X_{t_j}\right) \\ &= \sum_{i=1}^{p}\sum_{j=1}^{p} a_i a_j \mathbb{E}\left(X_{t_i} X_{t_j}\right) \end{aligned}$$

   Alors

$$\begin{aligned} V\left(\sum_{i=1}^{p} a_i X_{t_i}\right) &= \sum_{i=1}^{p}\sum_{j=1}^{p} a_i a_j \gamma\left(X_{t_i} X_{t_j}\right) \\ &\geq 0. \end{aligned}$$

□

Définition 1.2.4 *(la fonction d'autocorrélation)*
La fonction d'autocorrélation d'un processus stationnaire est définie par

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

1.2.1 Étude des moments d'un MMR(q)

Pour un processus $MMR(q)$ définie par (1.2)

.1 sa moyenne :

$$\begin{aligned} \mathbb{E}(X_t) &= \mu + \mathbb{E}(\varepsilon_t) + \theta_1 \mathbb{E}(\varepsilon_{t-1}) + \cdots + \theta_q \mathbb{E}(\varepsilon_{t-q}) \\ &= \mu + 0 + \theta_1 \times 0 + \cdots + \theta_q \times 0 \\ &= \mu. \end{aligned}$$

.2 sa variance : $V(X_t) = Cov(X_t, X_t) = \gamma(0)$.
   Pour $h = 0$:

$$\begin{aligned} \gamma(0) &= \mathbb{E}(X_t - \mathbb{E}(X_t))^2 \\ &= \mathbb{E}(X_t - \mu)^2 \\ &= \mathbb{E}[(\varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q})^2] \\ &= \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2 + \cdots + \theta_q^2 \sigma_\varepsilon^2 \\ &= (1 + \theta_1^2 + \cdots + \theta_q^2)\sigma_\varepsilon^2. \end{aligned}$$

Alors
$$\gamma(0) = (1 + \theta_1^2 + \cdots + \theta_q^2)\sigma_\varepsilon^2.$$

.3 sa fonction d'autocovariance :
$$Cov(X_t, X_{t-h}) = \gamma(|t-(t-h)|) = \gamma(h)$$

Pour X_t centrée on aura trois cas

i) Pour $1 < h < q$:

$$\begin{aligned}
\gamma(h) &= \mathbb{E}\left[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q})(\varepsilon_{t-h} + \theta_1\varepsilon_{t-h-1} + \cdots + \theta_q\varepsilon_{t-h-q})\right] \\
&= \mathbb{E}\left[\theta_h\varepsilon_{t-h}^2 + \theta_{h+1}\theta_1\varepsilon_{t-h-1}^2 + \cdots + \theta_q\theta_{q-h}\varepsilon_{t-q}^2\right] \\
&= \theta_h\sigma_\varepsilon^2 + \theta_{h+1}\theta_1\sigma_\varepsilon^2 + \theta_{h+2}\theta_2\sigma_\varepsilon^2 + \cdots + \theta_q\theta_{q-h}\sigma_\varepsilon^2 \\
&= (\theta_h + \theta_{h+1}\theta_1 + \theta_{h+2}\theta_2 + \cdots + \theta_q\theta_{q-h})\sigma_\varepsilon^2
\end{aligned}$$

i) Pour $h = q$:
D'un part on a
$$X_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q}$$
d'autre part on a
$$X_{t-h} = X_{t-q} = \varepsilon_{t-q} + \theta_1\varepsilon_{t-q-1} + \cdots + \theta_q\varepsilon_{t-2q}$$
Alors
$$\begin{aligned}
\gamma(h) &= \mathbb{E}\left[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q})(\varepsilon_{t-q} + \theta_1\varepsilon_{t-q-1} + \cdots + \theta_q\varepsilon_{t-2q})\right] \\
&= \mathbb{E}\left[\varepsilon_t\varepsilon_{t-q} + \theta_q\varepsilon_t\varepsilon_{t-2q} + \theta_1\varepsilon_{t-1}\varepsilon_{t-q} + \theta_1\theta_q\varepsilon_{t-1}\varepsilon_{t-2q} + .. + \theta_q\varepsilon_{t-q}^2 + \theta_q^2\varepsilon_{t-q}\varepsilon_{t-2q}\right] \\
&= 0 + \cdots + 0 + \theta_q\mathbb{E}(\varepsilon_{t-q}^2) + 0 + \cdots + 0 \\
&= \theta_q\sigma_\varepsilon^2.
\end{aligned}$$

iii) Pour $h > q$:
Il n'y a pas des ε dont les dates sont communes dans la définition de $\gamma(h)$, et donc l'espérence est nulle

Par conséquent,
$$\gamma(h) = \begin{cases} (\theta_h + \theta_{h+1}\theta_1 + \theta_{h+2}\theta_2 + \cdots + \theta_q\theta_{q-h})\sigma_\varepsilon^2 & \text{si } 1 < h < q \\ \theta_q\sigma_\varepsilon^2 & \text{si } h = q \\ 0 & \text{si} h > q \end{cases}$$

La fonction d'autocorrélation pour un processus $MMR(q)$ à partir de la fonction d'autocovariance de $MMR(q)$
$$\rho(h) = \begin{cases} \dfrac{\theta_h + \theta_{h+1}\theta_1 + \cdots + \theta_q\theta_{q-h}}{1 + \theta_1^2 + \cdots + \theta_q^2} & \text{si } 1 < h < q \\ \dfrac{\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2} & \text{si } h = q \\ 0 & \text{si } h > q \end{cases}$$

1.2.2 Étude des moments d'un MMR(1)

Le processus $MMR(1)$ définie par

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1}, \quad t \in \mathbb{Z}$$

Où $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$ et $\theta \in \mathbb{R}$ telle que $|\theta| < 1$.

1. Sa moyenne :

$$\begin{aligned}\mathbb{E}(X_t) &= \mathbb{E}(\varepsilon_t + \theta\varepsilon_{t-1}) \\ &= 0.\end{aligned}$$

2. Sa variance

$$V(X_t) = \gamma(0) = (1+\theta)\sigma_\varepsilon^2.$$

3. Sa fonction d'autocovariance :

$$\gamma(h) = \begin{cases} (1+\theta)\sigma_\varepsilon^2 & \text{si } h = 0 \\ \theta\sigma_\varepsilon^2 & \text{si } |h| = 1 \\ 0 & \text{si } |h| > 1 \end{cases} \tag{1.3}$$

4. Sa fonction d'autocorélation :

$$\rho(h) = \begin{cases} 1 & \text{si } h = 0 \\ \dfrac{\theta}{1+\theta^2} & \text{si } |h| = 1 \\ 0 & \text{si } |h| > 1 \end{cases} \tag{1.4}$$

1.3 Les modes de convergence

Définition 1.3.1 1. *Une suite de v.a.r $(X_n)_n$ converge en loi vers une v.a. réelle X si et seulement si*

$$\lim_{n\to\infty} F_{X_n}(t) = F_X(t)$$

en tout points $t \in \mathbb{R}$ où la fonction de répartion F_i est continue, on note $X_n \xrightarrow{\mathcal{L}} X$ ou $F(X_n) \xrightarrow{\mathcal{L}} F(X)$

2. *on dit que $(X_n)n$ converge en probabilité vers la v.a.Xsi pour $\varepsilon > 0$*

$$\lim_{n\to+\infty} \mathbb{P}\left(|X_n - X| > \varepsilon\right) = 0.$$

On note $X_n \xrightarrow{\mathbb{P}} X$.

3. *Soit $p \geq 1$ supposons que les v.a.(X_n) et X sont dans L^p. on dit que X_n converge dans L^p vers la v.a.X si*

$$\lim_{n\to+\infty} \mathbb{E}\left(|X_n - X|^p\right) = 0.$$

On note $X_n \xrightarrow{L^p} X$.

Proposition 1.3.1 *1. La convergence dans L^p, $p > 1$ entraîne la convergence en probabilité.*

2. La convergence en probabilité implique la convergence en loi.

3. $\forall C \in \mathbb{R}$, $X_n \xrightarrow{\mathcal{L}} C \Leftrightarrow X_n \xrightarrow{\mathbb{P}} C$

Théorème 1.3.1 *(caractérisation de la convergence en loi)*
Soit $(F_0, F_1, \ldots, F_k)$ une suite des fonctions de répartitions sur $\mathbb{R}^k$ avec les fonctions caractéristiques $\phi_n(t) = \int_{\mathbb{R}^k} \exp(it'X)dF_n(X), \quad n = 0, 1, \ldots, k$
Alors les trois assertions ci-dessous sont équivalentes

i) $F_n \to F_0$.
ii) $\int_{\mathbb{R}^k} g(X)dF_n(X) \to \int_{\mathbb{R}^k} g(X)dF_0(X) \quad \forall g$ une fonction continue borné.
iii) $\lim\limits_{n\to\infty} \phi_n(t) = \phi_0(t), \quad \forall t = (t_0, t_1, \ldots, t_k)' \in \mathbb{R}^k$.

Théorème de convergence dominée : Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires avec $|X_n| \leq Z$ et Z une variable aléatoire intégrable *i.e* $\mathbb{E}|Z| < \infty$.

$$\text{Si} \quad X_n \to X \; p.s \quad \text{Alors} \quad \mathbb{E}(X_n) \to \mathbb{E}(X).$$

Thérème de Lévy : Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires, X une variable aéatoire. ϕ_{X_n} et ϕ_X sont les fonctions caractéristiques respectives des variables aléatoires X_n et X.

$$\{\forall t \in \mathbb{R} : \phi_{X_n}(t) \to \phi_X(t)\} \Leftrightarrow \left\{X_n \xrightarrow{\mathcal{L}} X\right\}$$

Thérème de Slutsky : Soient $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ suites de variables aléatoires et $C \in \mathbb{R}$ telque

$$\begin{cases} X_n \xrightarrow{\mathcal{L}} X \\ Y_n \xrightarrow{\mathcal{L}} C \end{cases}$$

Alors

1. $X_n + Y_n \xrightarrow{\mathcal{L}} X + C$.
2. $X_n Y_n \xrightarrow{\mathcal{L}} XC$.
3. $\frac{X_n}{Y_n} \xrightarrow{\mathcal{L}} \frac{X}{C}, \quad C \neq 0$.

Théorème 1.3.2 *Soient $(m_n)_{n\in\mathbb{N}}$ une suite de $\mathbb{R}$ et $(\sigma_n)_{n\in\mathbb{N}}$ une suite de $\mathbb{R}_+^*$ et $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelle avec $\forall n \geq 1, X_n \hookrightarrow \mathcal{N}(m_n, \sigma_n^2)$. Alors*
Si les suites $(m_n)_{n\in\mathbb{N}}$ et $(\sigma_n)_{n\in\mathbb{N}}$ convergent ç-a-d

$$\begin{cases} m_n \xrightarrow[n\to\infty]{} m \\ \sigma_n \xrightarrow[n\to\infty]{} \sigma \end{cases}$$

Alors $X_n \xrightarrow{\mathcal{L}} X$ avec $X \hookrightarrow \mathcal{N}(m, \sigma^2)$.

Démonstration. On utilise la fonction caractéristique.

$$\begin{aligned}\phi_{X_n}(\lambda) &= \mathbb{E}\left(e^{i\lambda X_n}\right)\\ &= \exp\left(i\lambda m_n + \frac{\lambda^2}{2}\sigma_n^2\right)\end{aligned}$$

Comme $\begin{cases} m_n \xrightarrow[n\to\infty]{} m \\ \sigma_n \xrightarrow[n\to\infty]{} \sigma \end{cases}$
et la fonction e^x est une fonction continue, alors $\forall\lambda \in \mathbb{R}$,

$$\exp\left(i\lambda m_n + \frac{\lambda^2}{2}\sigma_n^2\right) \xrightarrow[n\to\infty]{} \exp\left(i\lambda m + \frac{\lambda^2}{2}\sigma^2\right)$$

ç-a-d $\phi_{X_n}(\lambda) \xrightarrow[n\to\infty]{} \phi_X(\lambda)$ avec $X_n \hookrightarrow \mathcal{N}(m_n, \sigma_n^2)$ et $X \hookrightarrow \mathcal{N}(m, \sigma^2)$.
Par le **Thérème de Lévy** on a $X_n \xrightarrow{\mathcal{L}} X$ avec $X \hookrightarrow \mathcal{N}(m, \sigma^2)$. $\square$

Théorie asymptotique pour les processus moyennes mobiles 2

2.1 Introduction

Pour évaluer le comportement asymptotique des séries temporelles,nous avons besoin de connaitre la distribution de statistiques (telles que la moyenne , la fonction d'autocovariance...)à partir des données dont nous disposons .Cependent .même pour un nombre fini d'observations la distribution exacte n'est pas toujours facile à déterminer .Dans ces cas-là , nous nous basons sur l'inférence statistique obtenue à partir des grands échantillons pour obtenir les résultats de [1] , [3] et [5] .

Définition :

Une suite de variables aléatoires $X = (X_n, n \in \mathbb{Z})$ est **asymptotiquement normale** de moyenne μ_n et d'écart-type σ_n,
si $\sigma_n > 0$ pour n grand et

$$\frac{X_n - \mu_n}{\sigma_n} \xrightarrow{\mathcal{L}} Z \qquad avec Z \sim \mathcal{G}(0,1)$$

où $\mathcal{G}(0,1)$ indique une loi gaussienne de moyenne nulle et d'écart-type 1 .Par la suite nous utiliserons la notation de Serfting (1980) à savoir

$$X \sim \mathcal{AN}(\mu_n, \sigma_n^2)$$

2.2 Loi des grands nombres pour les moyennes mobiles

Théorème 2.2.1 *soit $(X_t)_{t\in\mathbb{Z}}$ un processus moyenne mobile réelle d'ordre infini, défini par*

$$X_t = \sum_{j=-\infty}^{+\infty} \theta_j \varepsilon_{t-j}$$

Où $\{\varepsilon_t\}$ *est iid de moyenne* μ *avec* $\sum_{j=-\infty}^{+\infty} |\theta_j| < \infty$

Alors

$$\overline{X}_n \xrightarrow{\mathbb{P}} (\sum_{j=-\infty}^{+\infty} \theta_j)\mu$$

Pour démontrer ce théorème , on a besoin des propositions et du théorèmes suivants :

Proposition 2.2.1 *Si* $\{\varepsilon_t\}$ *un ensemble de variable aléatoire sachant que* $\sup \mathbb{E}|\varepsilon_t| < \infty$ *et si* $\sum_{j=-\infty}^{+\infty} |\theta_j| < \infty$ *alors la série* $\sum_{j=-\infty}^{+\infty} \theta_j \varepsilon_{t-j}$ *est intégrable*

Démonstration. Si on pose $g_n := \sum_{j=-n}^{n} |\theta_j||\varepsilon_{t-j}|$

alors $\{g_n\}_n$ est positive et croissante (suite des termes positifs)

Donc par le théorème de convergence monotone et $\sup \mathbb{E}|\varepsilon_t| < \infty$ on a :

$$\begin{aligned}
\mathbb{E}(\sum_{j=-\infty}^{+\infty} |\theta_j||\varepsilon_{t-j}|) &= \mathbb{E}\left(\lim_{n\to+\infty} \sum_{j=-n}^{+n} |\theta_j||\varepsilon_{t-j}\right) \\
&\leq \lim_{n\to+\infty} \mathbb{E}\left(\sum_{j=-n}^{+n} |\theta_j|\right) \sup_t(\mathbb{E}|\varepsilon_t|) \\
&< \infty
\end{aligned}$$

□

Proposition 2.2.2 *(L'inégalité de Tchebychev)*

Si $\mathbb{E}(|X|^r) < \infty, r \geqslant 0$ *et* $\lambda > 0$ *alors*

$$\mathbb{P}(|X| \geqslant \lambda) \leqslant \lambda^{-r}\mathbb{E}(|X|^r)$$

Démonstration.

$$\begin{aligned}
\mathbb{P}(|X| \geqslant \lambda) &= \mathbb{P}(|X|^r\lambda^{-r} \geqslant 1) \\
&= \mathbb{E}(\mathbb{1}_{[1,+\infty]}(|X|^r\lambda^{-r})) \\
&\leqslant \mathbb{E}(|X|^r\lambda^{-r}\mathbb{1}_{[1,+\infty]}(|X|^r\lambda^{-r})) \\
&\leqslant \lambda^{-r}\mathbb{E}|X|^r
\end{aligned}$$

□

Proposition 2.2.3 *Si* (X_n) *une suite de variables aléatoires dans* $\mathbb{R}^k$ *sachant que* $X_n \xrightarrow{\mathbb{P}} X$ *et si* $g : \mathbb{R}^k \longrightarrow \mathbb{R}^m$ *est continue alors*

$$g(X_n) \xrightarrow{\mathbb{P}} g(X)$$

Démonstration. Soit k un réel positif alors pour chaque $\varepsilon > 0$ on a

$$\mathbb{P}(|g(X_n) - g(X)| > \varepsilon) \leqslant \mathbb{P}(|g(X_n) - g(X)| > \varepsilon, |X| \leqslant K, |X_n| \leqslant k) + \mathbb{P}(\{|X| > K\} \cup \{|X_n| > k\})$$

Puisque g est uniformément continue sur $\{|X| \leq K\}$ il existe $\gamma(\varepsilon) > 0 :\ \forall n$ telle que

$$\{|g(X_n) - g(X)| > \varepsilon, |X| \leq K, |X_n| \leq k\} \subseteq \{|X_n - X| > \gamma(\varepsilon)\}$$

Par conséquent

$$\begin{aligned}\mathbb{P}(|g(X_n) - g(X)| > \varepsilon) &\leqslant \mathbb{P}(|X_n - X| > \gamma(\varepsilon)) + \mathbb{P}(|X| > k) + \mathbb{P}(|X_n| > k)\\ &\leqslant \mathbb{P}(|X_n - X| > \gamma(\varepsilon)) + \mathbb{P}(|X| > k) + \mathbb{P}(|X| > \frac{k}{2}) + \mathbb{P}(|X_n - X| > \frac{k}{2})\end{aligned}$$

Maintenant, pour tout $\delta > 0$,on peut choisir k qui rend chacun du deuxième et troisième termes inférieur à $\frac{\delta}{4}$ $\forall n$ assez grand

Par conséquent , $g(X_n) \xrightarrow{\mathbb{P}} g(X)$ □

Proposition 2.2.4 *(Loi Faible des Grands Nombres) si $\{\varepsilon_n\}$ est une suite de variables aléatoires i.i.d de moyenne μ alors*

$$\bar{\varepsilon}_n \xrightarrow{\mathbb{P}} \mu$$

Avec $\bar{\varepsilon}_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} \varepsilon_k$

Démonstration. Comme $\bar{\varepsilon}_n - \mu = \frac{1}{n}[(\varepsilon_1 - \mu) + \cdots + (\varepsilon_n - \mu)]$ il suffit de prouver le résultat pour des suites de variables de moyenne nulle

En supposant que $\mu = 0$ et en utilisant l'indépendance de $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ nous avons :

$$\begin{aligned}\Phi_{\bar{\varepsilon}_n}(t) &= \mathbb{E}(e^{it\bar{\varepsilon}_n})\\ &= \mathbb{E}(e^{it\frac{1}{n}\sum\limits_{k=1}^{n}\varepsilon_k})\\ &= \prod_{k=1}^{n} \mathbb{E}(e^{\frac{it}{n}\varepsilon_k})\\ &= (\mathbb{E}(e^{\frac{it}{n}\varepsilon_k})^{n}\\ &= (\Phi_{\varepsilon_1}(\frac{t}{n}))^{n}\end{aligned}$$

De l'inégalité : $|1 - y^n| \leqslant n|1 - y|$, $|y| \leqslant 1$ et de l'hypothèse que $\mathbb{E}(\varepsilon_1) = 0$
Il s'ensuit que

$$\begin{aligned}|1 - \Phi_{\bar{\varepsilon}_n}(t)| &\leqslant n|1 - \Phi_{\varepsilon_1}(\frac{t}{n})|\\ &= n|\mathbb{E}(1 + itn^{-1}\varepsilon_1) - e^{itn^{-1}\varepsilon_1}|\end{aligned}$$

Par l'inégalité de Jensen

$$\leqslant \mathbb{E}|n((1+itn^{-1}\varepsilon_1)-e^{itn^{-1}\varepsilon_1})|$$

On pose $z = tn^{-1}\varepsilon_1$

$$\begin{aligned}|1+iz-e^{iz}| &= |1+iz-\cos z - i\sin z|\\ &\leqslant |1-\cos z| + |z-\sin z|\\ &\leqslant \min(2|z|,|z|^2)\end{aligned}$$

Pour chaque réel z ,En remplaçant z par $tn^{-1}x$ nous voyons que $\forall x$

$$|n((1+itn^{-1}x)-e^{itn^{-1}x})| \leqslant 2|t||x| \quad , n=1,2,\ldots$$

et $|n((1+itn^{-1}x)-e^{itn^{-1}x})| \longrightarrow 0$ quand $n \longrightarrow \infty$
Comme $\mathbb{E}(\varepsilon_1) = 0$ par hypothèse , $\mathbb{E}|n((1+itn^{-1}\varepsilon_1)-e^{itn^{-1}\varepsilon_1})| \longrightarrow 0$ par le théorème de convergence dominé .
Par conséquent $\Phi_{\bar{\varepsilon}_n}(t) \longrightarrow 1 \ \forall t$ et comme la fonction caractéristique pour ε_0
En appliquant le **Thérème de Lévy** et (1.3.1)(3) on conclut que

$$\bar{\varepsilon}_n \xrightarrow{\mathbb{P}} 0$$

□

Théorème 2.2.2 *(Théorème d'Approximation Basique "TAB")*
Soient { X_n , $Y_{nk}, n = 1,2,\ldots;k = 1,2,\ldots$ } des vecteurs aléatoires de dimension k sachant que

i) $Y_{nk} \xrightarrow{\mathcal{L}} Y_k$ *qd* $n \longrightarrow \infty : \forall k = 1,2,\ldots$

ii) $Y_k \xrightarrow{\mathcal{L}} Y$ *qd* $k \longrightarrow \infty$

iiI) $\lim\limits_{k\to+\infty} \limsup\limits_{n\to+\infty} \mathbb{P}(||X_n - Y_{nk}|| > \lambda) = 0 \ \forall \lambda > 0$

Alors

$$X_n \xrightarrow{\mathcal{L}} Y$$

La toisiéme condition de TAB est impliqué de l'énagalité de ***Tchebychev*** *si*
(iii') $\mathbb{E}\,[||X_n - Y_{n_k}||] \xrightarrow[n,m\to+\infty]{} 0$
et (iii') est souvent beaucoup plus facile à établir que (iii)

Démonstration. Pour démonter ce théorème on utilise la fonction caractéristique en appliquant le **Théorème de Lévy**. Nous avons besoin de montrer que

$$|\phi_{X_n} - \phi_Y| \to 0$$

on note par $\phi \equiv \phi(\lambda)$ pour simplifier l'écriture .
D'abord

$$|\phi_{X_n} - \phi_Y| \leq |\phi_{X_n} - \phi_{Y_{n_k}}| + |\phi_{Y_{n_k}} - \phi_{Y_k}| + |\phi_{Y_k} - \phi_Y| \tag{2.1}$$

1. Par la condition (i) et le **Théorème de Lévy** le seconde terme converge vers 0.
2. Par la condition (ii') et le **Théorème de Lévy** le troisième terme converge vers 0.
3. Il reste à montrer que le 1er terme de (2.1) tendre vers 0.

En effet,

$$\begin{aligned}
|\phi_{Y_{n_k}} - \phi_{Y_k}| &= \left|\mathbb{E}\left(e^{i\lambda' X_n} - e^{i\lambda Y_{n_k}}\right)\right| \\
&\leq \mathbb{E}\left|e^{i\lambda' X_n}\left(1 - e^{i\lambda'(Y_{n_k} - X_n)}\right)\right| \\
&= \mathbb{E}\left|1 - e^{i\lambda'(Y_{n_k} - X_n)}\right| \\
&= \mathbb{E}\left[\left|1 - e^{i\lambda'(Y_{n_k} - X_n)}\right| \mathbb{1}_{\{|Y_{n_k} - X_n| < \delta\}}\right] + \mathbb{E}\left[\left|1 - e^{i\lambda'(Y_{n_k} - X_n)}\right| \mathbb{1}_{\{|Y_{n_k} - X_n| \geq \delta\}}\right]
\end{aligned}$$

Avec $\delta > 0$.
Étant donné $\varepsilon > 0$, on choisi $\delta(\varepsilon) > 0$ tel que

$$\left|1 - e^{i\lambda'(Y_{n_k} - X_n)}\right| < \varepsilon$$

Si $|Y_{n_k} - X_n| < \delta$ et le premier terme inférieure à ε, une constante arbitraire petite. Alors

$$\mathbb{E}(0) = 0$$

Pour le second terme on a $\left|1 - e^{i\lambda'(Y_{n_k} - X_n)}\right| \leq 2$. Alors on aura

$$\mathbb{E}\left[\left|1 - e^{i\lambda'(Y_{n_k} - X_n)}\right| \mathbb{1}_{\{|Y_{n_k} - X_n| \geq \delta\}}\right] \leq 2\mathbb{P}\left(|Y_{n_k} - X_n| \geq \delta\right)$$

et

$$\mathbb{P}\left(|Y_{n_k} - X_n| \geq \delta\right) \xrightarrow[n \to +\infty]{} 0.$$

Alors le second terme tendre vers 0 quand $n \to +\infty$
Par la condition (iii). □

Démonstration. On note que la série $\sum_{j=-\infty}^{+\infty} \theta_j \varepsilon_{t-j}$ converge absolument d'aprés la proposition (2.2.1)
et $\forall k$ on a par la loi faible des grands nombres

$$n^{-1} \sum_{k=1}^{n} \varepsilon_k \xrightarrow{\mathbb{P}} \mu$$

On définit les variables

$$\begin{aligned}
Y_{nk} &:= \frac{1}{n} \sum_{t=1}^{n} \sum_{|j| \leqslant k} \theta_j \varepsilon_{t-j} \\
&= \sum_{|j| \leqslant k} \theta_j \frac{1}{n} \sum_{t=1}^{n} \varepsilon_{t-j}
\end{aligned}$$

On obtient par la proposition (2.2.3)

$$Y_{nk} \xrightarrow[n\to+\infty]{\mathbb{P}} (\sum_{|j|\leqslant k} \theta_j)\mu$$

Si on définit $Y_k = (\sum_{|j|\leqslant k} \theta_j)\mu$

Alors $Y_k \longrightarrow Y := (\sum_{-\infty}^{+\infty} \theta_j)\mu$ quand $k \longrightarrow \infty$

il reste à montrer la troisième condition de **TAB** (i.e)

$$\lim_{k\to+\infty} \limsup_{n\to+\infty} \mathbb{P}(|\overline{X}_n - Y_{nk}| > \lambda) = 0 \;\forall \lambda > 0$$

En appliquant (L'inégalité de Tchebychev) pour $r = 1$ on obtient :

$$\begin{aligned}
\mathbb{P}(|\overline{X}_n - Y_{nk}| > \lambda) &= \mathbb{P}(|n^{-1} \sum_{t=1}^{n} \sum_{j=-\infty}^{+\infty} \theta_j \varepsilon_{t-j} - n^{-1} \sum_{t=1}^{n} \sum_{|j|\leqslant k} \theta_j \varepsilon_{t-j}| > \lambda) \\
&= \mathbb{P}(|n^{-1} \sum_{t=1}^{n} \sum_{|j|>k} \theta_j \varepsilon_{t-j}| > \lambda) \\
&\leqslant \frac{1}{\lambda} \mathbb{E}| \sum_{|j|>k} \theta_j \varepsilon_{t-j}| \\
&\leqslant \frac{1}{\lambda} (\sum_{|j|>k} |\theta_j|) \mathbb{E}|\varepsilon_1|
\end{aligned}$$

Comme $\sum_{j=-\infty}^{+\infty} |\theta_j| < \infty$ par hypothése alors $\sum_{|j|>k} |\theta_j| \xrightarrow[k\to+\infty]{} 0$

D'où la troxième condition du TAB est vérifiée , Alors

$$\overline{X}_n \xrightarrow{\mathcal{L}} (\sum_{j=-\infty}^{+\infty} \theta_j)\mu$$

D'aprés la proposition (1.3.1) vu dans le chapitre 1 on conclut que

$$\overline{X}_n \xrightarrow{\mathbb{P}} (\sum_{j=-\infty}^{+\infty} \theta_j)\mu \quad \text{C.Q.F.D}$$

□

2.3 Théorèmes limites centraux

Théorème 2.3.1 ***(Théorème centrale limite des variables aléatoires indépendantes)***
Soit $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires iid ç.à.d* $X_n \sim IID(\mu, \sigma^2)$
et $\overline{X}_n = n^{-1}(X_1 + X_2 + \cdots + X_n)$ *,Alors*

$$\overline{X}_n \sim \mathcal{AN}(\mu, n^{-1}\sigma^2)$$

Démonstration. on définie $(Y_t)_{t\in\mathbb{Z}}$ une suite de variables aléatoires iid de moyenne nulle et de variance égale à 1.
Par

$$Y_t = \frac{(X_t - \mu)}{\sigma}$$

et soit $\overline{Y}_n = n^{-1}\sum_{i=1}^{n} Y_i$
Par le **Théorème de Lévy**, il suffit de montrer que

$$\phi_{n^{\frac{1}{2}}\overline{Y}_n}(t) \longrightarrow e^{\frac{-t^2}{2}}$$

Par l'indépendance des variables $(Y_t)_{t\in\mathbb{Z}}$ (par construction),on a

$$\begin{aligned}
\phi_{n^{\frac{1}{2}}\overline{Y}_n}(t) &= \mathbb{E}\left[\exp\left(itn^{\frac{1}{2}}\overline{Y}_n\right)\right] \\
&= \mathbb{E}\left[\exp\left(itn^{-\frac{1}{2}}\sum_{j=1}^{n} Y_j\right)\right] \\
&= \left[\phi_{Y_1}(itn^{-\frac{1}{2}})\right]^n.
\end{aligned}$$

Tout d'abord on a besoin de l'inégalité suivante :

$$|x^n - y^n| \leq n|x-y| \qquad pour \qquad |x| \leq 1 \ et \ |y| \leq 1 \tag{2.2}$$

$$\text{si on pose} \begin{cases} x = \phi_{Y_1}(tn^{-\frac{1}{2}}) \ \ avec \ \ |\phi_{Y_1}(tn^{-\frac{1}{2}})| \leq 1 \\ y = 1 - \dfrac{t^2}{2n} \ \ avec \ \ |1 - \dfrac{t^2}{2n}| \leq 1 \ pour \ n \geq \dfrac{t^2}{4} \end{cases}$$

$|1 - \dfrac{t^2}{2n}| \leq 1 \ pour \ n \geq \frac{t^2}{4}$ car,
D'un part $-\dfrac{t^2}{2n} \leq 0$ alors $1 - \dfrac{t^2}{2n} \leq 1$.
D'autre part ,
Comme $n \geq \dfrac{t^2}{4}$ alors $2 \geq \dfrac{t^2}{2n}$
Ce qui implique $1 - \dfrac{t^2}{2n} \geq -1$.
Donc

$$|1 - \frac{t^2}{2n}| \leq 1$$

En appliquant l'inégalité (2.2) pour $n \geq \dfrac{t^2}{4}$ on a

$$\left|\left[\phi_{Y_1}(itn^{-\frac{1}{2}})\right]^n - \left[1 - \frac{t^2}{2n}\right]^n\right| \leq n\left|\phi_{Y_1}(itn^{-\frac{1}{2}}) - \left(1 - \frac{t^2}{2n}\right)\right| \tag{2.3}$$

$$= n\left|\mathbb{E}\left[e^{itn^{-\frac{1}{2}}Y_1} - (1 + itn^{-\frac{1}{2}} - \frac{t^2Y_1^2}{2n})\right]\right| \tag{2.4}$$

car :

$$\mathbb{E}\left(1 + itn^{-\frac{1}{2}} - \frac{t^2Y_1^2}{2n}\right) = 1 + itn^{-\frac{1}{2}}\mathbb{E}(Y_1) - \frac{t^2}{2n}\mathbb{E}(Y_1^2)$$
$$= 1 - \frac{t^2}{2n} \quad car \;\; \mathbb{E}(Y_1) = 0 \quad et \;\; \mathbb{E}(Y_1^2) = 1.$$

En utilisnt le développement de Taylor de la fonction $f(x) = e^{itn^{-\frac{1}{2}}x}$ au voisinage de $x = 0$
on a $f(x) = 1 + itn^{-\frac{1}{2}}x - \dfrac{t^2}{2n}x^2$
Alors

$$n\left|e^{itn^{-\frac{1}{2}}x} - (1 + itn^{-\frac{1}{2}}x - \frac{t^2}{2n}x^2)\right| \longrightarrow 0 \quad quand \quad n \longrightarrow +\infty.$$

et

$$n\left|e^{itn^{-\frac{1}{2}}x} - (1 + itn^{-\frac{1}{2}}x - \frac{t^2}{2n}x^2)\right| \leq (tx)^2 \quad pour \;\; tout \quad n \quad et \quad x.$$

Ainsi, par le théorème de convergence domonée le second membre de l'inegalitée (2.3) converge vers 0 quand $n \longrightarrow \infty$. et puisque $\left(1 - \dfrac{t^2}{2n}\right)^n \longrightarrow e^{-\frac{t^2}{2}}$ quand $n \longrightarrow \infty$ .on obtient

$$\phi_{n^{\frac{1}{2}}\overline{Y}_n}(t) \longrightarrow e^{\frac{-t^2}{2}} \quad C.Q.F.D$$

□

Ce théorème s'applique aux variables indépendantes , or nous avons vu que le modèle $MMR(1)$ est un processus 1-dépendant.Nous énonçons donc des résultas qui étendent le **Théorème de la Limite Centrale aux processus m-dépendant,dûs à Hoeffding et Robbins** (1948) [1]

Théorème 2.3.2 *Soit* $X = (X_t, t \in \mathbb{Z})$ *une suite strictement stationnaire de variables aléatoires m-dépendant centrées et de fonction d'autocovariance* $\gamma(.)$.
Si $\nu_m = \gamma(0) + \sum\limits_{j=1}^{m} 2\gamma(j) \neq 0$, *alors*

1) $\lim\limits_{n\to+\infty} n\, var(\overline{X}_n) = \nu_m$

2) $\overline{X}_n \sim \mathcal{AN}(0, \dfrac{\nu_m}{n})$

Démonstration. 1) Montrons que $\lim\limits_{t\to\infty} n.var(\overline{X}_n) = \nu_m$
avec

$$\nu_m = \gamma(0) + \sum_{j=1}^{m} 2\gamma(j)$$
$$= \sum_{j=-m}^{m} \gamma(j) \neq 0$$

On va calculer $Var(\overline{X}_n)$

$$\begin{aligned}
\mathsf{V}(\overline{X}_n) &= \mathsf{V}\left(\tfrac{1}{n}\sum_{t=1}^{n} X_t\right) \\
&= \frac{1}{n^2}Cov\left(\sum_{t=1}^{n} X_t, \sum_{t=1}^{n} X_t\right) \\
&= \frac{1}{n^2}\left[Cov(X_1 + X_2 + \cdots + X_n, X_1 + X_2 + \cdots + X_n)\right] \\
&= n^{-2}\left[Cov(X_1, X_1) + .. + Cov(X_1, X_n) + .. + Cov(X_n, X_0) + .. + Cov(X_n, X_n)\right].
\end{aligned}$$

ce qui donne :

$$\begin{aligned}
\mathsf{V}(\overline{X}_n) = n^{-2}[&n\gamma(0) + (n-1)\gamma(1) + \cdots + \gamma(n-1) + (n-1)\gamma(-1) \\
&+ (n-2)\gamma(-2) + \cdots + \gamma(1-n)]
\end{aligned}$$

Comme la fonction d'autocovariance est une fonction paire alors :

$$\begin{aligned}
\mathsf{V}(\overline{X}_n) &= n^{-2}\left[n\left(\gamma(0) + \frac{n-1}{n}\gamma(1) + \frac{n-2}{n}\gamma(2) + \dots\right)\right] \\
&= n^{-1}\left(\sum_{j=-n}^{n} \frac{n-|j|}{n}\gamma(j)\right)
\end{aligned}$$

Alors

$$\mathsf{V}(\overline{X}_n) = n^{-1}\left(\sum_{|j|<n} 1 - \frac{|j|}{n}\gamma(j)\right) \tag{2.5}$$

Donc

$$\begin{aligned}
n\mathsf{V}(\overline{X}_n) &= \sum_{|j|<n} (1 - \tfrac{|j|}{n})\gamma(j) \\
&= \sum_{|j|<m} (1 - \tfrac{|j|}{n})\gamma(j) \quad pour \quad n > m \\
&= \sum_{j=-m}^{m} (1 - \tfrac{|j|}{n})\gamma(j).
\end{aligned}$$

En appliquant le Théorème de la convergence dominé on trouve

$$\begin{aligned}
\lim_{n\to+\infty} n\mathsf{V}(\overline{X}_n) &= \lim_{n\to+\infty} \sum_{j=-m}^{m} (1 - \frac{|j|}{n})\gamma(j) \\
&= \sum_{j=-m}^{m} \lim_{n\to+\infty} (1 - \frac{|j|}{n})\gamma(j) \\
&= \sum_{j=-m}^{m} \gamma(j) \\
&= \nu_m
\end{aligned}$$

Alors

$$\lim_{t\to\infty} n.var(\overline{X}_n) = \nu_m$$

2) Montrons que si $\nu_m = \gamma(0) + \sum_{j=1}^{m} 2\gamma(j) \neq 0$ alors $\overline{X}_n \sim \mathcal{AN}(0, \frac{\nu_m}{n})$
pour démontrer cette partie, on utilise le théorème d' Approximation Basique (TAB), on doit construire une séquence de variables y_{kn} approximations de

$$n^{\frac{1}{2}}\overline{X}_n = n^{-\frac{1}{2}} \sum_{t=1}^{n} X_t$$

Dans le cas dépendant on peut simplement vérifier les trois conditions du théorème (TAB).
Tout d'abord, on considère pour $k > 2m$ l'approximation

$$\begin{aligned} y_{kn} = n^{-\frac{1}{2}}[&(X_1 + \cdots + X_{k-m}) + (X_{k+1} + \cdots + X_{2k-m}) \\ &+ (X_{2k+1} + \cdots + X_{3k-m}) + \cdots + (X_{(r-1)k+1} + \cdots + X_{rk-m})] \end{aligned}$$

Alors
$$y_{kn} = n^{-\frac{1}{2}}(Z_1 + Z_2 + \cdots + Z_r)$$
D'où $r = [\frac{n}{k}]$, avec $[\frac{n}{k}]$ le plus grand entier naturel inférieur ou égale à $\frac{n}{k}$.
Cette approximation contient qu'une partie de $n^{\frac{1}{2}}\bar{X}_n$.
Ainsi, les variables $Z_1, Z_2, .., Z_r$ sont indépendants car ils sont séparées par plus de m unités de temps ç-à-d :

$$k + 1 - (k - m) = m + 1$$

$m + 1$ unités séparent les variables Z_1 et Z_2.
Comme les variables $Z_1, Z_2, .., Z_r$ sont strictement stationnaires alors ils sont identiquement distribuées à la même loi de moyenne nulle et de variances :

$$S_{k-m} = \sum_{|u| \leqslant m} (k - m - |u|)\gamma(u)$$

En effet,

1. $\mathbb{E}(Z_1) = \mathbb{E}(X_1 + \cdots + X_{k-m}) = 0$ car $(X_t)_{t\in\mathbb{Z}}$ sont centrées
2.

$$\begin{aligned} var(Z_1) &= V(X_1 + \cdots + X_{k-m}) \\ &= \sum_{j=1}^{k-m} V(X_j) + 2 \sum_{i,j=1}^{k-m-1} cov(X_i, X_j) \\ &= (k-m)\gamma(0) + \sum_{i=1}^{k-m-j} (k-m-j)\gamma(j) \\ &= \sum_{|j| \leqslant m} (k-m-|j|)\gamma(j) \\ &:= S_{k-m} \end{aligned}$$

Maintenant on va vérifier les conditions de TAB

i) En appliquant le théorème central limite classique à la somme y_{kn} qui donne :

$$\begin{aligned} y_{kn} &= n^{-\frac{1}{2}} \sum_{i=1}^{r} Z_i \\ &= \left(\frac{n}{r}\right)^{-\frac{1}{2}} r^{-\frac{1}{2}} \sum_{i=1}^{r} Z_i \end{aligned}$$

Comme

$$\left(\frac{n}{r}\right)^{-\frac{1}{2}} \longrightarrow k^{-\frac{1}{2}}$$

et

$$r^{-\frac{1}{2}} \sum_{i=1}^{r} Z_i \xrightarrow{\mathcal{L}} \mathcal{N}(0, S_{k-m})$$

par **le théorème de Slutsky** on déduit que :

$$y_{kn} \xrightarrow{\mathcal{L}} y_k$$

Avec

$$y_k \sim \mathcal{N}(0, \frac{S_{k-m}}{k}) \tag{2.6}$$

pour un k fixé.

ii) Notons que

$$\lim_{k \to +\infty} \frac{S_{k-m}}{k} = \nu_m$$

En effet,

$$\begin{aligned} \lim_{k \to +\infty} \frac{S_{k-m}}{k} &= \lim_{k \to +\infty} \frac{1}{k} \sum_{|u| \leqslant m} (k - m - |u|)\gamma(u) \\ &= \lim_{k \to +\infty} \sum_{|u| \leqslant m} (\frac{k - m - |u|}{k})\gamma(u) \\ &= \lim_{k \to +\infty} \sum_{|u| \leqslant m} (1 - \frac{m - |u|}{k})\gamma(u) \end{aligned}$$

Si on pose

$$g_k(u) = (1 - \frac{m - |u|}{k})\gamma(u)$$

On a

$$\begin{aligned} |g_k(u)| &= |1 - \frac{m - |u|}{k}||\gamma(u)| \\ &\leqslant (1 + |m - |u||)|\gamma(u)| \\ &\leq (1 + 2m)\gamma(0) \\ &< \infty \end{aligned}$$

Car $|\gamma(u)| < \gamma(0)$ et $|u| \leq m$
Si on pose

$$g(u) = (1 + |m - |u||)|\gamma(u)|$$

Alors

$$g \in \mathbf{L^1}$$

Par le théorème de convergence dominée :

$$\begin{aligned}
\lim_{k\to+\infty} \sum_{|u|\leqslant m} (1 - \frac{m-|u|}{k})\gamma(u) &= \sum_{|u|\leqslant m} \lim_{k\to+\infty} (1 - \frac{m-|u|}{k})\gamma(u) \\
&= \sum_{|u|\leqslant m} \gamma(u) \\
&= \nu_m
\end{aligned}$$

car

$$\lim_{k\to+\infty} \frac{m-|u|}{k} = 0$$

Alors

$$\lim_{k\to+\infty} \frac{S_{k-m}}{k} = \nu_m \tag{2.7}$$

Donc la fonction caractéristique de y_k notée Φ_{y_k} converge.
Quand on applique le théorème (1.3.2) á partir des résultats (2.6) et (2.7) on a :

$$\Phi_{y_k}(\lambda) = \exp(-\frac{\lambda^2}{2}\frac{S_{k-m}}{k}) \longrightarrow \exp(-\frac{\lambda^2}{2}\nu_m)$$

Avec

$$\exp(-\frac{\lambda^2}{2}\nu_m) = \Phi_{N(0,\nu_m)}$$

Alors la suite des variables aléatoires $(y_k)_{k\geqslant 2m}$ converge en loi vers une variable aléatoire $y \sim \mathcal{N}(0, \nu_m)$ par théorème de Lévy. On déduit que :

$$y_k \xrightarrow{\mathcal{L}} y$$

iii) Pour vérifier la troisième condition du TAB, on va montrer sa condition équivalente (iii') ç-á-d il faut qu'on vérifie que :

$$\mathbb{E}[(n^{\frac{1}{2}}\overline{X}_n - y_{kn})^2] = Var(n^{\frac{1}{2}}\overline{X}_n - y_{kn}) \underset{k,n\to+\infty}{\longrightarrow} 0$$

$$
\begin{aligned}
n^{\frac{1}{2}}\overline{X}_n - y_{kn} &= \sqrt{n}\frac{1}{n}\sum_{t=1}^{n} X_t - \frac{1}{\sqrt{n}}\sum_{i=1}^{r} Z_i \\
&= \frac{1}{\sqrt{n}}[(X_1 + \cdots + X_n) - (X_1 + \cdots + X_{k+m} + \cdots + X_{rk-m})] \\
&= n^{-\frac{1}{2}}[(X_{k-m+1} + \cdots + X_k) + (X_{2k-m+1} + \cdots + X_{2k}) \\
&+ (X_{(r-1)k-m+1} + \cdots + X_{(r-1)k}) \\
&. \\
&. \\
&. \\
&+ (X_{rk-m+1} + \cdots + X_n)] \\
&= n^{-\frac{1}{2}}(W_1 + W_2 + \cdots + W_r)
\end{aligned}
$$

et

$$
\begin{aligned}
V(W_r) &= \sum_{|u|\leqslant k-m} (n - [\tfrac{n}{k}]k + m - |u|)\gamma(u) \\
&\leqslant \sum_{|u|\leqslant k-m} (k + m - |u|)\gamma(u)
\end{aligned}
$$

Donc,

$$
\begin{aligned}
Var(n^{\frac{1}{2}}\overline{X}_n - y_{kn}) &= n^{-1}[(r-1)S_m + Var(W_r)] \\
&\underset{n\to+\infty}{\longrightarrow} k^{-1}S_m \underset{k\to+\infty}{\longrightarrow} 0
\end{aligned}
$$

D'où les trois conditions de **TAB** sont vérifées .
Alors

$$\overline{X}_n \sim \mathcal{AN}(0, \frac{\nu_m}{n}) \text{ C.Q.F.D}$$

□

L'application de ce théorème au processus $MMR(1)$ donne

Corollaire 2.3.1 *une suite strictement stationnaire de variables aléatoires* 1*-dépendant définies par une* $MMR(1)$ *telle que*

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1} \qquad \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

Ces variables sont centrée et ont $\gamma(.)$ *comme fonction d'autocovariance .*
Comme $\nu_1 = \gamma(0) + 2\gamma(1) \neq 0$*, alors*

i) $\lim_{n\to+\infty} n\, var(\overline{X}_n) = (1+\theta)^2\sigma_\varepsilon^2$

ii) $\overline{X}_n \sim \mathcal{AN}(0, n^{-1}(1+\theta)^2\sigma_\varepsilon^2)$

Démonstration. Puisque le processus MMR(1) est 1-dépendant alors on peut appliquer le théoréme (2.3.2) en prenant $m = 1$ si $\nu_1 \neq 0$.
Tout d'abord on calcul ν_1 :

$$\nu_1 = \gamma(0) + 2\gamma(1)$$

on a par la relation (1.4) dans le chapitre 1

1. $\gamma(0) = (1+\theta^2)\sigma_\varepsilon^2$,
2. $\gamma(1) = \theta\sigma_\varepsilon^2$.

Alors

$$\begin{aligned} \nu_1 &= (1+\theta^2)\sigma_\varepsilon^2 + 2\theta\sigma_\varepsilon^2 \\ &= (1+\theta^2+2\theta)\sigma_\varepsilon^2. \end{aligned}$$

Donc

$$\nu_1 = (1+\theta)^2\sigma_\varepsilon^2 \neq 0$$

De plus

$$\begin{aligned} n\ Var(\overline{X}_n) &= n\ Var\left[\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right] \\ &= n^{-1} Var\left[(\varepsilon_1 + \theta\varepsilon_0) + (\varepsilon_2 + \theta\varepsilon_1) + \cdots + (\varepsilon_n + \theta\varepsilon_{n-1})\right] \\ &= n^{-1} Var\left[\theta\varepsilon_0 + (1+\theta)\varepsilon_1 + (1+\theta)\varepsilon_2 + \cdots + (1+\theta)\varepsilon_{n-1} + \varepsilon_n\right] \end{aligned}$$

Puisque $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$ donc on aura

$$\begin{aligned} n\ Var(\overline{X}_n) &= n^{-1}(1+\theta)^2 n Var(\varepsilon_1) \\ &= (1+\theta)^2\sigma_\varepsilon^2 \end{aligned}$$

Alors par le théorème (2.3.2)

i) $\lim\limits_{n\to\infty} n\ Var(\overline{X}_n) = \nu_1$,

ii) $\overline{X}_n \sim \mathcal{AN}(0, n^{-1}\nu_1)$.

En remplaçant ν_1 par $(1+\theta)^2\sigma_\varepsilon^2$,on conclut que :

i) $\lim\limits_{n\to\infty} n\ Var(\overline{X}_n) = (1+\theta)^2\sigma_\varepsilon^2$,

ii) $\overline{X}_n \sim \mathcal{AN}(0, n^{-1}(1+\theta)^2\sigma_\varepsilon^2)$.

$\square$

LES MÉTHODES D'ESTIMATION DU PARAMÈTRE D'UN PROCESSUS MMR(1) 3

Dans ce chapitre nous étudions des méthodes pour l'estimation des paramètres d'un processus moyenne mobile . Nous considérons ici uniquement le cas $MMR(1)$, centrée et inversible vérifiant

$$X_t = \varepsilon_t + \theta \varepsilon_{t-1} \tag{3.1}$$

Notre problème consiste à estimer le paramètre θ à partir des observations $X_1, X_2, \ldots, X_n$
Bien qu'il existe plusieurs méthodes d'estimation ,nous présentons les quatres méthodes les plus courants à savoir ;en développant les résultats des articles [2],[4],[7],[8] et [9]

1) la méthode des moments .
2) la méthode des moindres carrées.
3) la méthode du maximum de la fontion de vraisemblance .
4) la méthode récursive "*RIV*"

3.1 La méthode des moments

Cette méthode basée sur la fonction d'autocorrélation.consiste à substituer les moments empiriques aux moments théoriques et à résoudre les équations obtenues .
pour un $MMR(1)$ l'estimateur de θ noté $\widehat{\theta}$ doit être solution de

$$\widehat{\rho}(1) = \frac{\widehat{\gamma}(1)}{\widehat{\gamma}(0)}$$

Où $\widehat{\rho}(1)$ est la fonction d'autocorrélation empirique .
Soit à partir de l'quation (1.4) vu dans la chapitre 1

$$\widehat{\rho}(1) = \frac{\widehat{\theta}}{1 + \widehat{\theta}^2}$$

$\widehat{\theta}$ est alors la solution de l'équation

$$\widehat{\rho}(1)\widehat{\theta}^2 - \widehat{\theta} + \widehat{\rho}(1) = 0 \tag{3.2}$$

et vérifie

$$\widehat{\theta} = \frac{1 \pm \sqrt{1 - 4\widehat{\rho}(1)^2}}{2\widehat{\rho}(1)}$$

En calculant le detérminant de l'équation (3.2) on aurra trois cas se présentent :

1. Si $|\widehat{\rho}(1)| < \frac{1}{2}$ alors l'équation (3.2) à deux solutions .Nous choisissons celle qui donne un modèle inversible c'est à dire vérifiant $|\widehat{\theta}| < 1$ on obtient donc

$$\widehat{\theta} = \frac{1 - \sqrt{1 - 4\widehat{\rho}(1)^2}}{2\widehat{\rho}(1)}$$

2. Si $|\widehat{\rho}(1)| = \frac{1}{2}$ alors l'équation (3.2) admet une unique solution $|\widehat{\theta}| = 1$. Le modèle n'a pas inversible .
3. Si $|\widehat{\rho}(1)| > \frac{1}{2}$ alors les racines ne sont pas réelles et donc l'équation (3.2) n'a pas de solution réelle .

En résumé ,sous l'hypothѕe de l'existence d'un modèle $MMR(1)$ inversible défini par $X_t = \varepsilon_t + \theta\varepsilon_{t-1}$, l'estimateur des moments se limite au cas où $|\widehat{\rho}(1)| < \frac{1}{2}$.

Remarque 3.1.1 *Dans le cas général , la méthode des moments est souvent difficile à résoudre pour un $MMR(q)$, car les autocorrélations $\rho(h)$ sont des fonctions non linéaires des paramètres $\theta_1, \ldots, \theta_q$. Du fait de la sensibilité de cette méthode face aux erreurs d'arrondi , la méthode des moments est d'avantage utilisée pour calculer les valeurs initiales des paramètres avant l'utilisation d'autres méthodes d'estimation , que pour estimer le paramètre en lui même .*

3.2 La méthode des moindres carrés

Le principe des moindres carrés repose sur la recherche de la valeur θ qui minimise la somme des carrés des erreurs commises lors de l'estimateur. Soit $S(\theta)$ cette somme définie par :

$$S(\theta) = \sum_{t=-\infty}^{n} \mathbb{E}(\varepsilon_t \backslash \theta, X)^2 = \sum_{t=-\infty}^{n} \widehat{\varepsilon_t}^2 \tag{3.3}$$

Cette méthode est basée sur une relation de récurrence nécessitent des valeurs initiales de $\widehat{\varepsilon_t}$. Selon la méthode d'initialisation de $\widehat{\varepsilon_t}$, pour $t \leq 0$, on distingue deux types d'estimateurs :

- L'estimateur des moindres carrés conditionnel.
- L'estimateur des moindres carrés non-conditionnel.

3.2.1 L'estimateur des moindres carrés conditionnel :

on considère que les valeurs initiales des bruits blancs sont des valeurs fixées . Pour un $MMR(q)$ les valeurs $\widehat{\varepsilon}_{1-q}, \widehat{\varepsilon}_{2-q}, \ldots, \widehat{\varepsilon}_{-1}, \widehat{\varepsilon_0}$ sont supposées nulles .Nous présumons que $\widehat{\varepsilon_0} = 0$ et cherchons à minimiser

$$\mathrm{SC}(\theta) = \sum_{t=1}^{n} {\varepsilon_t}^2$$

La relation $X_t = \varepsilon_t + \theta\varepsilon_{t-1}$ s'écrit :

$$\varepsilon_t = X_t - \theta\varepsilon_{t-1}$$

Lemme 3.2.1 *Pour tout $t \geq 2$. on obtient*

$$\varepsilon_t = \sum_{i=0}^{t-1} (-\theta)^i X_{t-i}$$

Démonstration. On a

$$\begin{aligned} \varepsilon_0 &= 0 \\ \varepsilon_1 &= X_1 \end{aligned}$$

Donc

$$\begin{aligned} \varepsilon_2 &= X_2 - \theta\varepsilon_1 \\ &= X_2 - \theta X_1 \\ \varepsilon_3 &= X_3 - \theta\varepsilon_2 \\ &= X_3 - \theta(X_2 - \theta X_1) \\ \varepsilon_4 &= X_4 - \theta\varepsilon_3 \\ &= X_4 - \theta(X_3 - \theta X_2 + \theta^2 X_1) \\ &= X_4 - \theta X_3 - \theta^2 X_2 + \theta^3 X_1 \\ &\vdots \\ \varepsilon_t &= \sum_{i=0}^{t-1} (-\theta)^i X_{t-i} \end{aligned}$$

□

Donc d'après le lemme on a :

$$\varepsilon_t = \sum_{i=0}^{t-1} (-\theta)^i X_{t-i}$$

Ainsi , la somme des carrées conditionnellement à $\widehat{\varepsilon_0} = 0$ devient

$$\mathrm{SC}(\theta) = \sum_{t=1}^{n} {\varepsilon_t}^2 = \sum_{t=1}^{n} \left(\sum_{i=0}^{t-1} (-\theta)^i X_{t-i} \right)^2$$

L'estimateur $\widehat{\theta}_{\mathrm{SC}}$ est la valeur qui minimise la fonction $\mathrm{SC}(\theta)$. Celle-ci n'étant pas quadratique en θ , l'équation permettant d'estimer θ n'est donc pas linéaire.

3.2.2 L'estimateur des moindres carrés non-conditionnel :

Cette méthode proposée par Box et Jenkins [2] ,consiste à déterminer des valeurs initiales aux chocs aléatoires $\widehat{\varepsilon_0}, \widehat{\varepsilon}_{-1}, \dots, \widehat{\varepsilon}_{-Q}$ meilleurs que les valeurs nulles afin d'améliorer l'approximation de $\mathrm{S}(\theta)$ en enrichissant l'historique des observations . Ainsi on cherche à minimiser pour une valeur Q déterminée , la somme

$$\mathrm{SNC}(\theta) = \sum_{t=Q}^{n} \widehat{\varepsilon_t}^2 = \sum_{t=-Q}^{n} \mathbb{E}(\varepsilon_t \backslash \theta, X)^2$$

où les valeurs $\widehat{\varepsilon_t} = \mathbb{E}(\varepsilon_t \backslash \theta, X)$ sont estimées , pour $Q \leq t \leq 0$. Selon Box et Jenkins , pour un $MMR(q)$ vérifie $Q = 1 - q$, d'où $Q = 0$ pour un $MMR(1)$ et l'unique valeur à déterminer est $\widehat{\varepsilon_0}$.
le problème tout d'abord est de minimiser

$$SNC(\theta) = \sum_{i=0}^{n} \varepsilon_t^2.$$

Mais minimiser $\sum_{i=0}^{n} \varepsilon_t^2$ revient à minimiser $\varepsilon'\varepsilon$ et ε s'écrit sous la forme suivante :

$$\varepsilon = MX + T\varepsilon_0$$

Où

— ε est un paramètre

— M et T sont deux matrices telle que $dim\ M(n+1;n)$ et $dim\ T(n+1;1)$

En effet,

$$\begin{cases} \varepsilon_0 &= \varepsilon_0 \\ \varepsilon_1 &= X_1 - \theta\varepsilon_0 \\ \varepsilon_2 &= X_2 - \theta\varepsilon_1 \\ \varepsilon_3 &= X_3 - \theta\varepsilon_2 \\ &\vdots \\ \varepsilon_n &= X_n - \theta\varepsilon_{n-1} \end{cases} \iff \begin{cases} \varepsilon_0 &= \varepsilon_0 \\ \varepsilon_1 &= X_1 - \theta\varepsilon_0 \\ \varepsilon_2 &= X_2 - \theta X_1 + \theta^2\varepsilon_0 \\ \varepsilon_3 &= X_3 - \theta X_2 - \theta^2 X_1 + \theta^3\varepsilon_0 \\ &\vdots \\ \varepsilon_n &= X_n - \theta X_{n-1} + -(\theta)^2 X_{n-2} + \cdots + (-\theta)^{n-1}X_1 + (-\theta)^n \end{cases}$$

Alors

$$\varepsilon_i = X_i + \sum_{k=1}^{i-1} (-1)^k \theta^k X_{i-k} + (-1)^i \theta^i \varepsilon_0.$$

On peut écrire ce qui précède comme suit :

$$\begin{pmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -\theta & 1 & 0 & \cdots & \cdots & 0 \\ \theta^2 & -\theta & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (-1)^{n-1}\theta^{n-1} & 1 & 0 & \cdots & \cdots & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ \vdots \\ X_n \end{pmatrix} + \begin{pmatrix} 1 \\ -\theta \\ \theta^2 \\ \vdots \\ (-1)^n\theta^n \end{pmatrix} \varepsilon_0$$

avec

$$\begin{cases} MX &= \sum_{k=1}^{i-1}(-1)^k\theta^k X_{i-k} \\ T\varepsilon_0 &= (-1)^i\theta^i\varepsilon_0 \end{cases}$$

Commençons d'abord par estimer les valeurs initiales pour améliorer l'estimation .

Soit $\varepsilon = (\varepsilon_{1-q}, \ldots, \varepsilon_n)$,on prend $MMR(1)$ donc $\varepsilon = (\varepsilon_0, \ldots, \varepsilon_n)$,on a la proposition suivante

Proposition 3.2.1 *l'estimateur des moindres carrés de ε_0 noté $\hat{\varepsilon}_0$ est donné par :*

$$\hat{\varepsilon}_0 = -(T'T)^{-1}T'MX$$

Démonstration. D'après $Y = Z\beta + U$ tel que $U \hookrightarrow \mathcal{N}(0, \sigma^2)$.Alors l'estimateur de moindre carré de β est donné par $\hat{\beta} = (Z'Z)^{-1}Z'Y$.

Cepandant ,on a $\varepsilon = MX + T\varepsilon_0$ tel que $\varepsilon_t \hookrightarrow \mathcal{N}(0, \sigma_t^2)$

$$\begin{aligned} \varepsilon = MX + T\varepsilon_0 \iff -MX &= T\varepsilon_0 - \varepsilon \\ \iff \hat{\varepsilon}_0 &= (T'T)^{-1}T'(-MX) \\ &= -(T'T)^{-1}T'MX \end{aligned}$$

En prenant

$$\begin{cases} T &= Z \\ \varepsilon &= U \\ y &= MX \end{cases}$$

$$\begin{aligned} S(\theta) &= \varepsilon'\varepsilon = \sum_{t=0}^{n}\varepsilon_t^2 \\ &= (MX + T\varepsilon_0)'(MX + T\varepsilon_0) \end{aligned}$$

Alors

$$\hat{S}(\theta) = (MX + T\hat{\varepsilon}_0)'(MX + T\hat{\varepsilon}_0).$$

Avec

$$\begin{aligned} MX + T\varepsilon_0 &= MX + T\hat{\varepsilon}_0 - T\hat{\varepsilon}_0 + T\varepsilon_0 \\ &= MX + T\hat{\varepsilon}_0 + T(\varepsilon_0 - \hat{\varepsilon}_0) \end{aligned}$$

Donc

$$\begin{aligned} S(\theta) &= [MX + T\hat{\varepsilon}_0 + T(\varepsilon_0 - \hat{\varepsilon}_0)]'[MX + T\hat{\varepsilon}_0 + T(\varepsilon_0 - \hat{\varepsilon}_0)] \\ &= \left[(MX + T\hat{\varepsilon}_0)' + (\varepsilon_0 - \hat{\varepsilon}_0)T'\right][MX + T\hat{\varepsilon}_0 + T(\varepsilon_0 - \hat{\varepsilon}_0)] \\ &= (MX + T\hat{\varepsilon}_0)'(MX + T\hat{\varepsilon}_0) + (MX + T\hat{\varepsilon}_0)'T(\varepsilon_0 - \hat{\varepsilon}_0) \\ &+ (\varepsilon_0 + \hat{\varepsilon}_0)'T'(MX + T\hat{\varepsilon}_0) + (\varepsilon_0 - \hat{\varepsilon}_0)'T'T(\varepsilon_0 - \hat{\varepsilon}_0). \end{aligned}$$

De plus

$$\begin{cases} (MX+T\hat{\varepsilon}_0)'T(\varepsilon_0-\hat{\varepsilon}_0) &= 0 \\ (\varepsilon_0-\hat{\varepsilon}_0)'T'(MX+T\hat{\varepsilon}_0) &= 0 \end{cases}$$

car $\hat{\varepsilon}_0 = -(T'T)^{-1}T'MX \Longleftrightarrow T'T\hat{\varepsilon}_0 = -T'MX$.

Donc

$$\begin{aligned} S(\theta) &= (MX+T\hat{\varepsilon}_0)'(MX+T\hat{\varepsilon}_0) + (\varepsilon_0-\hat{\varepsilon}_0)'T'T(\varepsilon_0-\hat{\varepsilon}_0) && (3.4) \\ &= \hat{S}(\theta) + (\varepsilon_0-\hat{\varepsilon}_0)'T'T(\varepsilon_0-\hat{\varepsilon}_0) && (3.5) \end{aligned}$$

□

et en fin trouver l'estimateur des moindrs carrés non-conditionnel revient á minimiser le terme $\hat{S}(\theta)$

3.3 la méthode du maximum de la fontion de vraisemblance :

là encore , nous étudions deux types d'estimateurs selon le choix des valeurs initiales , à savoir

- L'estimateur du maximum de la vraisemblance conditionnelle.
- L'estimateur du maximum de la vraisemblance non-conditionnelle ou vraisemblance exacte .

3.3.1 L'estimateur du maximum de la vraisemblance conditionnelle

Ayant une relation entre X_t et ε_t , la fonction de vraisemblance des X_t est déterminée à partir de celle de ε_t.
Donc nous supposons que les ε_t sont indépendantes de loi $\mathcal{N}(0,\sigma_\varepsilon^2)$.
Partant de la densité des probabilités conditionnelles des bruits blancs , on a

$$f(\varepsilon_t/\theta,\sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{1}{2}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma_\varepsilon^2}\right) \tag{3.6}$$

En réecrivant ε_t à partir de (3.1) c'est à dire :

$$\sum_{t=1}^{n} (x_t-\theta\varepsilon_{t-1})^2 = \sum_{t=1}^{n} \varepsilon_t^2 = \varepsilon'\varepsilon$$

Donc on peut écrire la fonction de vraisemblance en fonction des paramètres $(\theta,\sigma_\varepsilon^2)$ en supposant que les valeurs initiales $X_0, \varepsilon_{-1}, \varepsilon_0$ sont fixées à 0 .
la fonction de vraisemblance conditionnelle est alors

$$\mathrm{L}_C(\theta,\sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_\varepsilon^2}\sum_{t=1}^{n}\varepsilon_t^2\right)$$

D'où la fonction de la log-vraisemblance

$$l_c(\theta, \sigma_\varepsilon^2) = \log \mathrm{L}_C(\theta, \sigma_\varepsilon^2) = -\frac{n}{2}\log(2\pi\sigma_\varepsilon^2) - \frac{1}{2\sigma_\varepsilon^2}\mathrm{SC}(\theta)$$

où $\mathrm{SC}(\theta) = \sum\limits_{t=1}^{n} \varepsilon_t^2(\theta)$
Les premiers termes de $\log \mathrm{L}_C(\theta, \sigma_\varepsilon^2)$ étant négligeables lorsque le nombre d'observations est grand , il en découle que $\log \mathrm{L}_C(\theta, \sigma_\varepsilon^2)$ est dominée par $\mathrm{SC}(\theta)$, et de ce fait rechercher le maximum du logarithme de la vraisemblance conditionnelle revient à minimiser le terme $\mathrm{SC}(\theta)$.
On en déduit que L'estimateur du maximum de vraisemblance conditionnelle coïncide avec l'estimateur conditionnel des moindres carrés étudié précédement .

Démonstration. on a d'abord $X_t = \varepsilon_t + \theta\varepsilon_{t-1}$ où $\varepsilon_t \hookrightarrow \mathcal{N}(0, \sigma_\varepsilon^2)$.
Donc $\mathcal{L}\left(X_t|\varepsilon_{t-1}\right) \hookrightarrow \mathcal{N}(\theta\varepsilon_{t-1}, \sigma_\varepsilon^2)$
Alors

$$f(X_t|\varepsilon_{t-1}, \theta, \sigma_\varepsilon^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(X_t - \theta\varepsilon_{t-1})^2}{\sigma_\varepsilon^2}\right).$$

On suppose ε_0 fixé et $\varepsilon_1 = X_1 - \theta\varepsilon_0$. Alors

$$\mathcal{L}\left(X_2|X_1, \varepsilon_0\right) = \mathcal{L}\left(X_2|\varepsilon_1\right)$$

d'où

$$\begin{aligned} f\left(x_2|x_1, \varepsilon_0, \theta, \sigma_\varepsilon^2\right) &= f\left(x_2|\varepsilon_1, \theta, \sigma_\varepsilon^2\right) \\ &= \frac{1}{\sigma_\varepsilon\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x_2 - \theta\varepsilon_1)^2}{\sigma_\varepsilon^2}\right) \end{aligned}$$

Si ε_1 , ε_0 fixés, $\varepsilon_2 = X_2 - \theta\varepsilon_1$,alors

$$\begin{aligned} \varepsilon_1 &= X_1 - \theta\varepsilon_0 \\ \varepsilon_2 &= X_2 - \theta\varepsilon_1 \\ &\vdots \\ \varepsilon_n &= X_n - \theta\varepsilon_{n-1} \end{aligned}$$

d'où

$$f\left(x_t|x_{t-1}, \ldots, x_1, \varepsilon_0, \theta, \sigma_\varepsilon^2\right) = \mathcal{L}\left(X_t, \varepsilon_{t-1}\right)$$

Ainsi la fonction de vraissemblance est :

$$\begin{aligned}
L_c(\theta,\sigma^2) &= f(x_1,\theta,\varepsilon_0,\sigma_\varepsilon^2)\cdot f(x_0|x_1,\theta,\varepsilon_0,\sigma_\varepsilon^2)\dots f(x_n|x_{n-1},\dots,x_1,\theta,\varepsilon_0,\sigma_\varepsilon^2)\\
&= \frac{1}{2\pi\sigma_\varepsilon^2}e^{-1/2\left(\frac{x_1-\varepsilon_0}{\sigma_\varepsilon}\right)^2}\prod_{t=2}^{n}e^{-\frac{1}{2}\left(\frac{x_t-\theta\varepsilon_{t-1}}{\sigma_\varepsilon}\right)}\\
&= \left(\frac{1}{2\pi\sigma_\varepsilon^2}\right)^{n/2}e^{-\frac{1}{2\sigma_\varepsilon^2}\sum_{t=1}^{n}\varepsilon_t^2}\\
&= \left(\frac{1}{2\pi\sigma_\varepsilon^2}\right)^{n/2}e^{-\frac{1}{2\sigma_\varepsilon^2}S(\theta)}
\end{aligned}$$

d'où la fonction de log-vraissemblance est :

$$\begin{aligned}
l(\theta,\sigma^2) &= \log(L_c(\theta,\sigma^2))\\
&= -\frac{n}{2}\log 2\pi-\frac{n}{2}\log\sigma_\varepsilon^2-\frac{1}{2\sigma_\varepsilon^2}S(\theta).
\end{aligned}$$

□

Ainsi l'estimateur de maximum de vraisemblance coïncide avec l'estimateur de moindre carré cas conditionnel

3.3.2 L'estimateur du maximum de la vraisemblance non-conditionnelle

Contrairement à la vraisemblance conditionnelle , le critère de vraisemblance non-conditionnelle (ou vraisemblance exacte) ne correspond pas au critère des moindres carrès non-conditionnel.
Comme on a déja vu dans la partie conditionnelle on connait la fonction de vraisemblance des ε_t

$$\mathrm{L}(\theta,\sigma_\varepsilon^2)=(2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}}\exp\left(-\frac{1}{2\sigma_\varepsilon^2}\sum_{t=1}^{n}\varepsilon_t^2\right)$$

et pour obtenir la fonction de vraisemblance, nous utilisons la densité de probabilité d'une série d'observations $X=(X_1,\dots,X_n)$, en supposant que chaque observation de cette série est générée par une MMR(1) définie par (3.1). Sous l'hypothèse de normalité des ε_i, et par conséquent $X=(X_1,\dots,X_n)$ est gaussien.
car on a

$$\varepsilon = MX+T\varepsilon_0 \Longrightarrow X=M^{-1}(\varepsilon-T\varepsilon_0)$$

D'ou $\varepsilon=\begin{pmatrix}\varepsilon_0\\ \varepsilon_1\\ \vdots\\ \vdots\\ \varepsilon_n\end{pmatrix}$ et $M=\begin{pmatrix}0 & \cdots & \cdots & \cdots & \cdots & 0\\ 1 & 0 & \cdots & \cdots & \cdots & 0\\ -\theta & 1 & 0 & \cdots & \cdots & 0\\ \theta^2 & -\theta & 1 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots\\ (-1)^{n-1}\theta^{n-1} & 1 & 0 & \cdots & \cdots & 1\end{pmatrix}$

$$\text{et } X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ \vdots \\ X_n \end{pmatrix} \text{ et } T = \begin{pmatrix} 1 \\ -\theta \\ \theta^2 \\ \vdots \\ (-1)^n\theta^n \end{pmatrix}$$

telle que $dim\ M(n+1;n)$ et $dim\ L(n+1;1)$ Comme ε et ε_0 suivent des lois normal et X est une transformation affine de ε et ε_0 alors

$$X \sim \mathcal{N}(0, \Gamma_x)$$

D'où Γ_x est la matrice de variance -covariance de X de dim$(n \times n)$ définie positive comme suit :

$$\Gamma_x = \sigma_\varepsilon^2 \begin{pmatrix} 1+\theta^2 & \theta & 0 & . & . & . & 0 \\ \theta & 1+\theta^2 & \theta & . & . & . & 0 \\ 0 & \theta & 1+\theta^2 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & 1+\theta^2 \end{pmatrix}$$

on obtient la fonction de densité jointe

$$f(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}} (\det \Gamma_X)^{-1/2} e^{-\frac{1}{2}\langle X, \Gamma_X^{-1} X\rangle}.$$

On pose $\Gamma_x = \sigma_\varepsilon^2 K \implies \det \Gamma_x = (\sigma_\varepsilon^2)^n \det K$; avec k est une matrice de dimension $(n \times n)$

Alors

$$\ln(f(x_1, \ldots, x_n)) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln \sigma^2 - \frac{1}{2}\ln\det(K) - \frac{1}{2\sigma_\varepsilon^2} X' K^{-1} X. \tag{3.7}$$

Par ailleurs

$$\begin{aligned} f(\varepsilon, \theta, \sigma_\varepsilon^2) &= \left(\frac{1}{2\pi\sigma_\varepsilon^2}\right)^{n+1} e^{-\frac{1}{2\sigma_\varepsilon^2}\sum_{t=0}^{n}\varepsilon_t^2} \\ &= \left(\frac{1}{2\pi\sigma_\varepsilon^2}\right)^{n+1} e^{-\frac{1}{2\sigma_\varepsilon^2}\varepsilon'\varepsilon} \end{aligned}$$

Or $\varepsilon = MX + T\varepsilon_0$
On pose $(X, \varepsilon_0) = h(\varepsilon) \iff \varepsilon = h^{-1}(x, \varepsilon_0) = MX + T\varepsilon_0$ avec $|J_{h^{-1}}(x, \varepsilon_0)| = 1$ et de plus

$$f(x, \varepsilon_0) = f(h^{-1}(x, \varepsilon_0))|J_{h^{-1}}(x, \varepsilon_0)|$$

$$\implies f(x, \varepsilon_0, \theta, \sigma_\varepsilon^2) = \left(2\pi\sigma_\varepsilon^2\right)^{-\frac{n+1}{2}} e^{-\frac{1}{2\sigma_\varepsilon^2}S(\theta)}.$$

Or on a par (3.4) $S(\theta) = \hat{S}(\theta) + (\varepsilon_0 - \hat{\varepsilon}_0)' T' T (\varepsilon_0 - \hat{\varepsilon}_0)$ ce qui imlique

$$\begin{aligned}
f(x, \varepsilon_0, \theta, \sigma^2) &= \left(2\pi\sigma_\varepsilon^2\right)^{-\frac{n}{2}} \left(2\pi\sigma_\varepsilon^2\right)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma_\varepsilon^2}\hat{S}(\theta)} e^{-\frac{1}{2\sigma_\varepsilon^2}(\varepsilon_0 - \hat{\varepsilon}_0)' T' T (\varepsilon_0 - \hat{\varepsilon}_0)}. \\
&= \left[\left(2\pi\sigma_\varepsilon^2\right)^{-\frac{n}{2}} |T'T|^{-\frac{1}{2}} e^{-\frac{1}{2\sigma_\varepsilon^2}\hat{S}(\theta)}\right] \left[\left(2\pi\sigma_\varepsilon^2\right)^{-\frac{n}{2}} |T'T|^{-\frac{1}{2}} e^{-\frac{1}{2\sigma_\varepsilon^2}(\varepsilon_0 - \hat{\varepsilon}_0)' T' T (\varepsilon_0 - \hat{\varepsilon}_0)}\right] \\
&= f(x, \theta, \sigma_\varepsilon^2).f(\varepsilon_0 | x, \theta)
\end{aligned}$$

Avec

$$\mathrm{L}(\theta, \sigma_\varepsilon^2) = f(x, \theta, \sigma_\varepsilon^2) = \left(2\pi\sigma_\varepsilon^2\right)^{-\frac{n}{2}} |T'T|^{-\frac{1}{2}} e^{-\frac{1}{2\sigma_\varepsilon^2}\hat{S}(\theta)} \tag{3.8}$$

En identifiant (3.7) avec la fonction de vraisemblance (3.8) on obtient

$$\begin{cases} |K| = |T'T| \\ \hat{S}(\theta) = X' K^{-1} X \end{cases}$$

D'où le résultat final

$$l(\theta, \sigma_\varepsilon^2) - \frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma_\varepsilon^2 - \frac{1}{2} \ln |T'T| - \frac{1}{2\sigma_\varepsilon^2} \hat{S}(\theta)$$

Trouver l'estimateur du maximum de la vraisemblance non conditionnelle revient à minimiser le terme $\hat{S}(\theta)$

3.4 La méthode du *RIV*

L'estimateur proposée est appellé "estimateur de valeurs initiales aléatoires " (RIVE) et la méthode est dite la méthode du RIV en utilisant Les résultats de l'article de **Anna Clara Monti** [4] .
l'estimateur est basée sur l'idée que toute l'information concernant les paramètres est transmise par l'échantillon et que le bruit blanc a un effet qui devient négligeable quand n tend vers l'infini. Soit $\{X_t, t \in \mathbb{Z}\}$ un processus moyenne mobile d'ordre un défini par

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1} \tag{3.9}$$

avec $|\theta| \leqslant 1$ et $\{\varepsilon_t,\ t \in \mathbb{Z}\}$un bruit blanc faible.
Nous remarquons que dans la définition (3.9) nous avons opté pour un signe $-$ devant le paramètre pour des comodités de calcul contrairement à la définition d'un processus moyenne mobile $MMR(1)$ dans le chapitre 1.
$(X_1, X_2, \ldots, X_n)$ des observatios d'un processus moyenne mobile $MMR(1)$ définie par (3.9).

La méthode d'estimation *RIV* du paramètre θ est basée sur le procédé itératif composé des étapes suivantes :

i) Générer un échantillon de taille n : $b_0^0, b_1^0, \ldots, b_{n-1}^0$ v.a.i.i.d de moyenne égale à zéro et variance égale à un (de loi quelconque).

ii) Prendre $i = 1$

iii) Estimer θ par

$$\widehat{\theta}^{(i)} = - \frac{\sum_{t=1}^{n} X_t b_{t-1}^{(i-1)}}{\sum_{t=1}^{n} \left(b_{t-1}^{(i-1)}\right)^2}$$

iv) On définit de nouveau $b_t^{(i)}$ par

$$b_t^{(i)} = X_t + \widehat{\theta}^{(i)} b_{t-1}^{(i-1)} \qquad t = 1, \ldots, n$$

v) Prendre $i = i + 1$.

vi) Répéter les étapes iii) et v) jusqu'à ce que $|\widehat{\theta}^{(i)} - \widehat{\theta}^{(i-1)}|$ soit "trés petit"

L'estimateur obtenu par la méthode du *RIV* est défini par

$$\widehat{\theta}_{RIV} = \lim_{i \to +\infty} \widehat{\theta}^{(i)} \tag{3.10}$$

Un estimateur *RIV* de la variance du bruit blanc est défini par

$$\widehat{\sigma}^2_{RIV} = n^{-1} \sum_{t=1}^{n} b_t^2$$

où

$$b_t = \lim_{i \to +\infty} b_t^{(i)}$$

Remarque 3.4.1 1. *Les nombres* $b_0^0, b_1^0, \ldots, b_{n-1}^0$ *peuvent être générés, par exemple en utilisant une distribution gaussienne.*

2. *Quand le nombre d'itération i augmante la variable aléaltoire* $b_t^{(i)}$ *converge vers le bruit* ε_t *en probabilité.*

3. *Soit* $\widehat{\theta}_{RIV}$ *l'estimateur de* ***RIV*** *défini par l'équation* (3.10) *,par les résultats de l'article de* ***Anna Clara Monti*** *[4] on a les deux propriétés suivantes :*

   i) $\widehat{\theta}_{RIV} \xrightarrow{\mathbb{P}} \theta$

   ii) $n^{\frac{1}{2}} \left(\widehat{\theta}_{RIV} - \theta\right)$ *est asymptotiquement distribuée à la loi normal centrée réduite* $\mathcal{N}(0, 1)$.

Bibliographie

[1] Hoeffding, W.,et Robbins, H.(1948). The central limit theorem for dependent random variables. Duke Mathematical Journal,15(3), 773˘780

[2] Box, G.E.P. and G.M. Jenkins. Time series analysis : forecasting and control (Holden Day, San Franscisco.1970. revised edition. 1976).

[3] Robert H.Shumway and David S. Stoffer ,Time Series Analysis And Its Application ,third edition ,Springer. pp $507-526$

[4] Anna Clara Monti(1996) ,A newpreliminary estimator for $MA(1)$ models. computational statistics , data analysis $21(1996)1-15$.

[5] Brockwell, P. J. and Richard A.davis (1991). Time Series : Theory and Methods pp$203-214$. (Second ed.). New York : Springer.

[6] Brockwell, P. J. et Davis, R. A. (1996). Introduction to Time Series and Forecasting. New York : Springer

[7] DeniseR.Osborn (1976).Maximum likelihood estimation of moving average processes .An- nals of Economic and Social Measurement $5/1, 1976$.

[8] A. M. Walker.Large-Sample Estimation of Parameters for Moving-Average models.Biometrika, Vol.48, $No.3/4(Dec., 1961)$, pp.$343-357$

[9] James D.Hamilton , Time Series Analysis , pp $127-131$

Résumé

L'objectif principal de ce mémoire est d'étudier les théorèmes limites des processus moyennes mobiles

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1}, \quad t \in \mathbb{Z}.$$

Puis citer quelques méthodes d'estimation du paramètre θ

Mots clés : Stationnarité, m-dṕendance, bruit blanc, processus moyenne mobile.

Abstract

The main purpose of this text is to study the limit theorems of moving average processes

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1}, \quad t \in \mathbb{Z}.$$

Then cite some estimation methods of parameter θ

Keywords : Stationarity, m dependence, white noise, moving average process.