

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

UNIVERSITÉ MOHAMED KHIDER, BISKRA  
Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie  
DÉPARTEMENT DE MATHÉMATIQUES



Mémoire présenté en vue de l'obtention du Diplôme :

MASTER en Mathématiques

Option : Statistique

Par

**MEBARKI ASMA**

Titre :

**Modélisation de processus stochastique via des fonctions  
trigonométriques**

Membres du Comité d'Examen :

Pr.	BENATIA Fateh	UMKB	Président
Pr.	CHERFAOUI MOULOUD	UMKB	Encadreur
Dr.	SOLTANE Louiza	UMKB	Examinatrice

Juin 2024



## Dédicace

*Je dédie ce humble travail*

À ma mère et mon père **MEBARKI BELKACEM**, source d'amour et du courage, qui ont sacrifié pour que je puisse arriver jusque-là et qui m'ont beaucoup soutenu durant mon cursus, que DIEU les garde.

À mes chers frères **Mokhtar** et **Radhwane, Houssam Adin**

À mes chères sœurs **Imane** et **Serine, Salsabil**

À tous mes amis,

Et spéciale dédicace à M.A et Djinane

Je dédie également ce manuscrit à mon cher défunt grand-père.

Je vous dédie ce travail en témoignage de mon profond amour.

À tous ceux que j'aime.

♡Asma♡



## REMERCIEMENTS

**T**out d'abord, je remercie ALLAH Le Tout Puissant qui m'a accordée la volonté et le courage pour réaliser ce mémoire.

Je remercie le **Pr. Cherfaoui Mouloud**, avec gratitude pour son soutien sans faille et continu tout au long de mon voyage de rédaction de la thèse. Sans sa présence orientée et ses précieuses directives, je n'aurais pas pu construire ce travail correctement, et grâce à ses sages conseils et son utilité, j'ai surmonté les défis difficiles auxquels j'ai été confronté dans ma thèse et ma compréhension du sujet de la thèse. Je suis reconnaissant d'avoir été sous sa supervision et je lui souhaite une bonne santé, du succès et de la chance dans tous ses futurs efforts.

Je remercie les membres du jury qui ont accepté de juger mon travail.

Pr BENATIA Fateh, qui me fait l'honneur de présider ce jury.

Pr SOLTANE Louiza, pour avoir acceptée d'examiner ce travail.

Je tiens à remercier tous ceux qui m'ont aidé et qu'ont été à mes côtés. Merci à  
vous.

J'adresse mes remerciements les plus chaleureux envers la famille de notre département, enseignants et étudiants, dont je suis honorée de faire partie de  
celle-ci.

Enfin, je ne saurais terminer cette partie sans exprimer ma gratitude à mes parents, mes frères.

# Table des matières

Dédicace	i
Remerciements	ii
Table des matières	iv
Table des figures	v
Liste des tables	vi
Introduction	1
<b>1 Notions sur les Processus Aléatoires</b>	<b>3</b>
Introduction . . . . .	3
1.1 Quelques définitions usuelles . . . . .	3
1.2 Modélisation statistique d'un PA . . . . .	5
1.3 Régression linéaire simple . . . . .	6
1.3.1 Description du modèle de régression linéaire simple . . . . .	6
1.3.2 Modèle linéaire simple sous sa forme matricielle . . . . .	7
1.3.3 Estimation d'un modèle linéaire simple par MCO . . . . .	8
1.3.4 Propriétés des Estimateurs . . . . .	10

1.3.5	Exemple d'application . . . . .	12
1.4	Régression Linéaire Multiple . . . . .	14
1.4.1	Modèle de régression linéaire multiple . . . . .	14
1.4.2	Le modèle sous forme matricielle . . . . .	15
1.4.3	Estimation des Paramètres via MCO . . . . .	15
1.4.4	Propriétés des Estimateurs . . . . .	16
<b>2</b>	<b>Estimation des paramètre d'un PA à saisonnalité régulière</b>	<b>18</b>
	Introduction . . . . .	18
2.1	Description du modèle . . . . .	18
2.2	Estimation des paramètres : Cas sans tendance . . . . .	20
2.2.1	Méthode MCO par approche directe . . . . .	20
2.2.2	Estimation des paramètres via un modèle linéaire multiple . . . . .	23
2.3	Estimation des paramètres : cas avec tendance . . . . .	25
<b>3</b>	<b>Application numérique</b>	<b>28</b>
	Introduction . . . . .	28
3.1	Première application . . . . .	29
3.2	Deuxième application . . . . .	35
	Conclusion et remarques . . . . .	39
	<b>Conclusion</b>	<b>40</b>
	<b>Bibliographie</b>	<b>42</b>

# Table des figures

1.1	Illustration d'un Modèle additive et ses composantes . . . . .	5
1.2	Illustration d'un Modèle multiplicatif et ses composantes . . . . .	6
1.3	Nuages des points observés et la droite de regression . . . . .	14
2.1	Exemple d'illustration du Modèle retenu dans l'étude . . . . .	23
3.1	Illustration graphique du Modèle de l'exemple 3 et ses composantes : la série, la tendance et la saisonnalité . . . . .	30
3.2	Illustration graphique du Modèle de l'exemple 4 et ses composantes : la série, la tendance et la saisonnalité . . . . .	30
3.3	Variation de l' <i>EQM</i> en fonction de la taille de l'échantillon et de l'écart-type des résidus (1/2). . . . .	31
3.4	Variation de l' <i>EQM</i> en fonction de la taille de l'échantillon et de l'écart-type des résidus (2/2). . . . .	32
3.5	Allure graphiques des quatre modèles de la deuxième application. . .	36
3.6	Allure graphiques du cinquième modèle de la deuxième application. .	36
3.7	Variation de l' <i>EQM</i> en fonction de la taille de l'échantillon et de l'écart-type des résidus : cas du modèle $M_5$ . . . . .	37
3.8	Variation de l' <i>EQM</i> en fonction de la taille de l'échantillon et de l'écart-type des résidus : cas des quatre premiers modèles. . . . .	37

# Liste des tableaux

3.1	Estimateurs des paramètres du modèle linéaire multiple : Cas de l'exemple 1 et de l'exemple 2 . . . . .	31
3.2	Estimateurs des paramètres du modèle linéaire multiple : Cas de l'exemple 3 et de l'exemple 4 . . . . .	32
3.3	Estimateurs des paramètres du modèle original : Cas d'exemple 1 . . .	33
3.4	Estimateurs des paramètres du modèle original : Cas d'exemple 2 . . .	33
3.5	Estimateurs des paramètres du modèle original : Cas d'exemple 3 . . .	34
3.6	Estimateurs des paramètres du modèle original : Cas d'exemple 4 . . .	34

# Introduction

La modélisation des processus stochastiques à l'aide de fonctions trigonométriques est une méthode puissante pour analyser et prédire des phénomènes qui présentent des motifs cycliques ou périodiques. Cette approche consiste à utiliser des fonctions trigonométriques telles que le sinus et le cosinus pour représenter et approximer le comportement d'un processus aléatoire au fil du temps. En combinaison avec la méthode des Moindres Carrés Ordinaire (MCO), cette approche permet de trouver les paramètres optimaux des fonctions trigonométriques afin de minimiser l'erreur entre les observations réelles et les prédictions du modèle. La régression linéaire multiple est une technique statistique utilisée pour comprendre la relation entre plusieurs variables indépendantes,  $X$  ( $X \in \mathbb{R}^p$ ,  $p \geq 1$ ) et une variable dépendante,  $Y$  ( $Y \in \mathbb{R}$ ), en trouvant le meilleur plan (meilleure ligne droite entre  $Y$  et  $X_i$ ,  $i = \overline{1 : p}$ ) qui les représente au sens d'un critère bien déterminé (erreur quadratique, vraisemblance, ...). Cela permet de prédire la valeur de la variable dépendante en fonction des variables indépendantes.

Dans ce mémoire nous sommes intéressés à l'estimation, par la méthode MCO, les paramètres du modèle d'un processus stochastique dont l'expression est :

$$y_t = at + b + c \sin(\alpha t + \beta) + \varepsilon_t. \quad (1)$$

Dans un premier lieu, nous avons illustré que l'estimation des paramètres en question



ne peut être réalisée qu'après la transformation du (1) vers un modèle de régression linéaire multiple. Par ailleurs, à l'aide d'une étude de simulation nous avons analysé la qualité des estimations en fonction de la taille de l'échantillon et de la variance des résidus.

En plus de la présente Introduction, le mémoire est composé de trois chapitre où :

**Chapitre 1** : nous avons introduit quelques notions nécessaires pour la compréhension de la thématique abordée.

**Chapitre 2** : nous avons exposé la méthodologie à suivre pour l'estimation des paramètres du modèle considéré à l'aide de la méthode MCO.

**Chapitre 3** : nous avons présenté l'étude de simulation réalisée.

Le mémoire se termine par une conclusion et une liste bibliographique.

# Chapitre 1

## Notions sur les Processus Aléatoires

### Introduction

Dans le présent chapitre nous allons introduire, dans un premier lieu, la notion du processus aléatoire et les principales définitions liées à cette thématique. Par la suite, un cas particuliers d'un processus aléatoire à savoir le modèle linéaire simple sera exposé où nous concentrons sur l'estimation de ses paramètre par la méthode des moindres carrés ordinaire. Enfin, le modèle de régression linéaire multiple sera également présenté.

### 1.1 Quelques définitions usuelles

Un processus aléatoire (PA) (ou processus stochastique) est un processus qui décrit l'évolution d'une grandeur aléatoire en fonction du temps (ou de l'espace). Les applications principales des processus aléatoires sont : l'administration des réseaux et bien entendu dans les télécommunications.

L'étude des caractéristiques de processus aléatoire fait appel aux théories de probabilités et de statistiques dont elle constitue l'un des objectifs les plus profonds. Leur étude soulève des problèmes mathématiques intéressants et souvent très difficiles et ceci à cause de la possibilité d'existence de relations complexes entre les différentes composantes du processus.

**Définition 1.1.1** *Un processus stochastique noté  $X = (X_t)_{t \in T}$  est une famille de variables aléatoires  $X_t$  indexée par un ensemble  $T$ . En général  $T = \mathbb{R}_+$  ou  $\mathbb{R}$ , et on considère que le processus est indexé par le temps  $t \in T$ .*

- Si  $T$  est un ensemble fini, le processus est alors un vecteur aléatoire.
- Si  $T = \mathbb{N}$  ou  $T = \mathbb{Z}$  alors le processus est une suite de v.a.
- Plus généralement, quand  $T \subseteq \mathbb{Z}$ , le processus est dit discret.

**Définition 1.1.2** *Une série temporelle ou (série chronologique) notée  $X_t$  est une suite de variable aléatoire (v.a) qui dépend du temps et qui représente l'évaluation d'un phénomène ou cours du temps :*

$$X_t = (X_t)_{t \in \mathbb{N}} = \{X_0, X_1, \dots, X_t, \dots\}$$

**Définition 1.1.3** *La tendance, notée  $T_t$ , représente l'évolution à long terme de la série étudiée. Elle traduit le comportement "moyen" de la série.*

**Définition 1.1.4** *On appelle saisonnalité une fonction périodique de période  $p$ , notée  $S_{t \pm p} = S_t$ , pour tout  $t = 0, \dots, T$ . Elle est liée au rythme imposé par les saisons. Elle est de la forme :*

$$S_t = \gamma_1 S^1 + \gamma_2 S^2 + \dots + \gamma_P S^P,$$

avec

$$\gamma_j \in \mathbb{R}, \quad \sum_{j=1}^P \gamma_j = 0 \quad \text{et} \quad S^j = \begin{cases} 1, & \text{si } t = j; \\ 0, & \text{si } t \neq j. \end{cases}$$

**Définition 1.1.5** La composante résiduelle (ou bruit ou résidu), noté  $\varepsilon_t$ , correspond à des fluctuations irrégulières, en général de faible intensité mais de nature aléatoire.

## 1.2 Modélisation statistique d'un PA

Un modèle est une image simplifiée de la réalité qui vise à traduire les mécanismes de fonctionnement du phénomène étudié et permet de mieux les comprendre. En statistique on peut distinguer principalement deux types de modèles

- Modèle additif :

$$X_t = T_t + S_t + \varepsilon_t \quad (1.1)$$

- Modèle multiplicatif :

$$X_t = T_t \cdot S_t + \varepsilon_t \quad (1.2)$$

**Remarque 1.2.1** Le Modèle additif c'est le "modèle classique de décomposition" dans le traitement des modèles d'ajustement. La variable  $X_t$  s'écrit comme la somme de trois termes :  $X_t = T_t + S_t + \varepsilon_t$  où  $T_t$  représente la tendance,  $S_t$  la saisonnalité et  $\varepsilon_t$  la composante ("erreurs au modèle") aléatoire.

Les figures 1.1 et 1.2 représente un exemple d'illustration d'un modèle additif et d'un modèle multiplication respectivement.

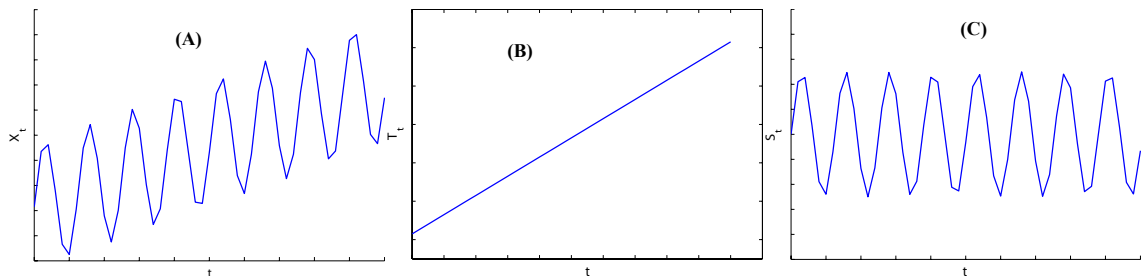


FIGURE 1.1 – Illustration d'un Modèle additive et ses composantes : (A) la série, (B) la tendance et (C) la saisonnalité

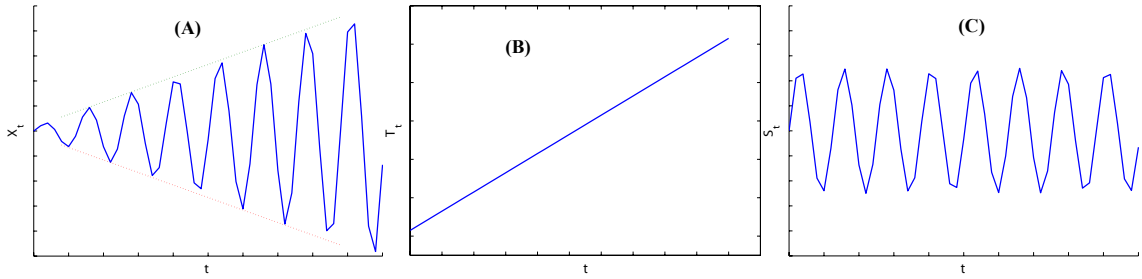


FIGURE 1.2 – Illustration d’un Modèle multiplicatif et ses composantes : (A) la série, (B) la tendance et (C) la saisonnalité

## 1.3 Régression linéaire simple

### 1.3.1 Description du modèle de régression linéaire simple

Dans ce passage, nous allons introduire un cas très particulier d’un modèle additif, à savoir le modèle linéaire simple. Il est à noter qu’un modèle linéaire simple n’est rien d’autre qu’un modèle additif lorsque  $S_t = 0, \forall t \in T$  et la tendance  $T_t$  est décrite par une droite ie  $T_t = at + b$ . Ainsi le modèle de régression linéaire simple est défini par une équation de la forme :

$$x_i = b + at_i + \varepsilon_i, \quad i = 1, \dots, n,$$

où :

$n$  : est la taille de l’échantillon.

$x_i$  : est la variable expliquée ou variable dépendante dont la valeur observée est aléatoire.

$t_i$  : est la variable explicative et indépendante (valeur fixée).

$a$  et  $b$  : sont des constantes inconnues appelées paramètres du modèle.

$\varepsilon_i$  : est le terme d’erreur aléatoire du modèle.

Dans l’étude de tels modèle généralement on impose les hypothèses suivantes :

(i) Les erreurs  $\varepsilon_i$  sont centrées, ont la même variance, et non corrélées entre elles :

$$\begin{aligned}E(\varepsilon_i) &= 0, \\E(\varepsilon_i^2) &= \sigma_\varepsilon^2, \quad i = 1, \dots, n \\cov(\varepsilon_i, \varepsilon_j) &= 0, \quad \forall(i, j) \text{ si } i \neq j,\end{aligned}$$

(ii) L'erreur est indépendante de  $t$  :

$$cov(\varepsilon, t) = 0$$

**Remarque 1.3.1** *Le modèle est linéaire en  $X$  par rapport aux deux paramètres  $a$  et  $b$ .*

### 1.3.2 Modèle linéaire simple sous sa forme matricielle

Notons que le modèle de régression linéaire simple peut encore s'écrire sous sa forme matricielle de la manière suivante :

$$X = b + aT + \varepsilon,$$

ou encore

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

avec

- $X$  désigne le vecteur à expliquer de taille  $n \times 1$ ,

- $T$  la matrice explicative de taille  $n \times 2$ ,
- $\varepsilon$  le vecteur d'erreurs de taille  $n \times 1$ ,

### 1.3.3 Estimation d'un modèle linéaire simple par MCO

Rappelons que l'estimation est une notion mathématique permettant d'apporter à un nombre d'éléments susceptibles de comporter des erreurs un ajustement afin d'obtenir des données proches de la réalité.

Pour estimer  $a$  et  $b$ , on peut utiliser la méthode des moindres carrés ordinaire (MCO) qui consiste à déterminer des valeurs pour  $a$  et  $b$  minimisant la somme des erreurs quadratiques. c'est-à-dire les estimateurs peuvent s'écrire sous la forme suivante :

$$(\hat{b}, \hat{a}) = \arg \min_{a, b \in \mathbb{R}} \Psi(b, a),$$

où :

$$\Psi(b, a) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (x_i - b - at_i)^2.$$

Pour déterminer la valeur qui minimise  $\Psi$ , annulant les dérivées partielles par rapport à  $a$  et  $b$ , c'est-à-dire

$$\left( \frac{\partial \Psi(b, a)}{\partial b} = 0, \frac{\partial \Psi(b, a)}{\partial a} = 0 \right);$$

ainsi on obtient les équations suivantes :

$$\begin{cases} \frac{\partial \Psi(b, a)}{\partial b} = -2 \sum_{i=1}^n (x_i - b - at_i) = 0 \\ \frac{\partial \Psi(b, a)}{\partial a} = -2 \sum_{i=1}^n t_i (x_i - b - at_i) = 0 \end{cases}$$

Alors,

$$\begin{cases} \sum_{i=1}^n (x_i - b - at_i) = 0 \\ \sum_{i=1}^n (t_i x_i - bt_i - at_i^2) = 0 \end{cases} \Rightarrow \begin{cases} \sum_{i=1}^n x_i - b \sum_{i=1}^n 1 - a \sum_{i=1}^n t_i = 0 \\ \sum_{i=1}^n t_i x_i - b \sum_{i=1}^n t_i - a \sum_{i=1}^n t_i^2 = 0 \end{cases}$$

La première équation donne lieu à :

$$\sum_{i=1}^n x_i = a \sum_{i=1}^n t_i + nb;$$

d'où l'on déduit immédiatement que :

$$\hat{b} = \bar{X} - \hat{a}\bar{T}, \tag{1.3}$$

avec :

$\bar{T} = \frac{1}{n} \sum_{i=1}^n t_i$  et  $\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$  sont les moyennes empiriques des  $t_i$  et des  $x_i$  respectivement.

La seconde équation donne :

$$\sum_{i=1}^n t_i x_i = \hat{a} \sum_{i=1}^n t_i^2 + \hat{b} \sum_{i=1}^n t_i$$

et en remplaçant  $\hat{b}$  par son expression (1.3), nous aurons :

$$\hat{a} = \frac{\frac{1}{n} \sum_{i=1}^n t_i x_i - \bar{T} \bar{X}}{\frac{1}{n} \sum_{i=1}^n t_i^2 - \bar{T}^2} = \frac{\text{cov}(t, x)}{\sigma_t^2} = \frac{S_{tx}}{S_t^2},$$

où



$$S_t^2 = \frac{1}{n} \sum_{i=1}^n (t_i - \bar{T})^2$$

$$S_{tx} = \frac{1}{n} \sum_{i=1}^n t_i x_i - \bar{T} \bar{X}$$

Les estimateurs de moindres carrés ordinaires sont alors :

$$\begin{cases} \hat{b} = \bar{X} - \hat{a}\bar{T}, \\ \hat{a} = \frac{S_{tx}}{S_t^2}, \end{cases}$$

### 1.3.4 Propriétés des Estimateurs

**Théorème 1.3.1** (*Biais et variance des estimateurs*) Les quantités  $\hat{a}$  et  $\hat{b}$  sont des estimateurs sans biais de  $a$  et  $b$ , c'est-à-dire

$$E(\hat{a}) = a \quad \text{et} \quad E(\hat{b}) = b.$$

De plus, on a les propriétés suivantes :

$$Var(\hat{a}) = \frac{\sigma_\varepsilon^2}{n \sum_{i=1}^n (t_i - \bar{T})^2},$$

$$Var(\hat{b}) = \frac{\sigma_\varepsilon^2 \sum_{i=1}^n t_i^2}{n \sum_{i=1}^n (t_i - \bar{T})^2},$$

**Preuve.** On a

$$\begin{cases} x_i = b + at_i + \varepsilon_i \\ \bar{X} = b + a\bar{T} + \bar{\varepsilon}. \end{cases}$$

L'écart entre  $x_i$  et la moyenne  $\bar{X}$  est donc

$$x_i - \bar{X} = a(t_i - \bar{T}) + (\varepsilon_i - \bar{\varepsilon}).$$

**Prouvons d'abord que  $E(\hat{a}) = a$  :** on a

$$\begin{aligned} \hat{a} &= \frac{S_{tx}}{S_t^2} = \frac{\sum_{i=1}^n (t_i - \bar{T})(x_i - \bar{X})}{\sum_{i=1}^n (t_i - \bar{T})^2} = \frac{a \sum_{i=1}^n (t_i - \bar{T})^2 + \sum_{i=1}^n (t_i - \bar{T})(\varepsilon_i - \bar{\varepsilon})}{\sum_{i=1}^n (t_i - \bar{T})^2} \\ \hat{a} &= a + \frac{\sum_{i=1}^n (t_i - \bar{T})(\varepsilon_i - \bar{\varepsilon})}{\sum_{i=1}^n (t_i - \bar{T})^2} = a + \frac{\sum_{i=1}^n (t_i - \bar{T})\varepsilon_i}{\sum_{i=1}^n (t_i - \bar{T})^2} \end{aligned}$$

alors,

$$E(\hat{a}) = E \left( a + \frac{\sum_{i=1}^n (t_i - \bar{T})\varepsilon_i}{\sum_{i=1}^n (t_i - \bar{T})^2} \right) = a + \frac{\sum_{i=1}^n (t_i - \bar{T})E(\varepsilon_i)}{\sum_{i=1}^n (t_i - \bar{T})^2},$$

et le fait que  $E(\varepsilon_i) = 0$  alors on en déduit que  $E(\hat{a}) = a$ .

**Prouvons maintenant que  $E(\hat{b}) = b$  :** On a d'après la méthode MCO

$$\hat{b} = \bar{X} - \hat{a}\bar{T}$$

alors,

$$E(\hat{b}) = E(\bar{X} - \hat{a}\bar{T}) = E(b + a\bar{T} + \bar{\varepsilon} - \hat{a}\bar{T}) = b + \bar{T}E(a) + E(\bar{\varepsilon}) - \bar{T}E(\hat{a})$$

et du fait que  $E(\varepsilon) = 0$  donc,

$$E(\hat{b}) = b + a\bar{T} - a\bar{T} = b.$$

$Var(\hat{a})$  et  $Var(\hat{b})$  : Pour dégager les expressions des variances des deux estimateurs il suffit de procéder avec le même raisonnement que dans les deux démonstrations précédentes (sans biais). Pour plus de détails sur la démonstration le lecteur peut se référer à [9, 2].

■

### 1.3.5 Exemple d'application

Nous avons un jeu de données qui représente le nombre d'heures étudiées en fonction des notes obtenues par un groupe d'étudiants.

Voici à quoi pourrait ressembler ce jeu de données :

Heures étudiées ( $t_i$ )	2	3	4	5	6
Note obtenue sur ./100 ( $x_i$ )	67	70	72	86	84

Le nuage de points correspondant est représenté sur la figure 1.3.

La droite de régression linéaire peut être représentée par l'équation :

$$x = \hat{a}t + \hat{b},$$

avec

$$\begin{aligned}\hat{a} &= \frac{cov(t,x)}{S_t^2} \\ \hat{b} &= \bar{X} - \hat{a}\bar{T}.\end{aligned}$$

Le tableau suivant range quelques calculs intermédiaire :

						Total
$t_i$	2	3	4	5	6	20
$x_i$	67	70	72	86	84	379
$t_i x_i$	134	210	288	430	504	1566
$t_i^2$	4	9	16	25	36	90

Ainsi, on aura

$$\bar{T} = \frac{1}{n} \sum_{i=1}^n t_i = \frac{1}{5} \times 20 = 4.$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{5} \times 379 = 75.8$$

$$cov(t, x) = \frac{1}{n} \sum_{i=1}^n t_i x_i - \bar{T} \bar{X} = \frac{1}{5} \times 1566 - 4 \times 75.8 = 10$$

$$S_t^2 = \frac{1}{n} \sum_{i=1}^n t_i^2 - (\bar{T})^2 = \frac{1}{5} \times 90 - (4)^2 = 2$$

On en déduit :

$$\hat{a} = \frac{cov(t, x)}{S_t^2} = \frac{10}{2} = 5$$

$$\hat{b} = \bar{X} - \hat{a} \bar{T} = 75.8 - 5 \times 4 = 55.8$$

Ce qui donne lieu à l'équation de régression linéaire suivante :

$$x = 5t + 55.8.$$

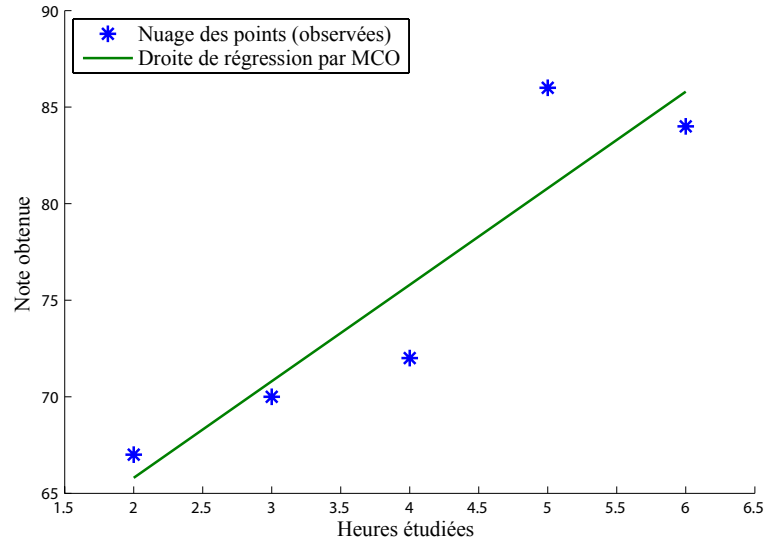


FIGURE 1.3 – Présentations graphique du nuages des points observés et la droite de régression linéaire.

## 1.4 Régression Linéaire Multiple

### 1.4.1 Modèle de régression linéaire multiple

Le modèle de régression linéaire multiple se base sur l'écriture de chaque observation  $y_i$ ,  $i = 1, \dots, n$  ( $n > p$ ), par une équation de la forme :

$$y_i = b_0 + \sum_{j=1}^p b_j x_{i,j} + \varepsilon_i,$$

où :

$x_{i,j}$  : Sont des variables déterministes,

$b_j$  : Les paramètres du modèle  $b_j$  ( $j = \overline{0 : p}$ ) sont des constantes inconnus,

$\varepsilon_i$  : Sont des termes d'erreur d'une variable aléatoire  $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^t$ .

### 1.4.2 Le modèle sous forme matricielle

Le modèle de régression linéaire multiple s'écrit sous la forme matricielle suivante :

$$Y = XB + \varepsilon,$$

avec,

$Y$  : est un vecteur aléatoire de dimension  $n$ .

$X$  : est une matrice de dimension  $(n, p + 1) \in \mathbb{R}^n \times \mathbb{R}^{p+1}$ ,

$B$  : est le vecteur des paramètres du modèle de dimension  $p + 1$ ,

$\varepsilon$  : est le vecteur des erreurs de dimension  $n$ .

Telle que :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{1,1} & \cdots & x_{1,p} \\ 1 & x_{2,1} & \cdots & x_{2,p} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n,1} & \cdots & x_{n,p} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

### 1.4.3 Estimation des Paramètres via MCO

On appelle estimateur des moindres carrés  $\hat{B}$  de  $B$  la valeur suivante :

$$\hat{B} = \arg \min_B \sum_{i=1}^n \varepsilon_i^2 = \arg \min \sum_{i=1}^n \left( y_i - \sum_{j=1}^p b_j x_{i,j} \right)^2 = \arg \min (Y - XB)'(Y - XB).$$

Pour calculer  $\hat{B}$ , nous proposons l'écriture matricielle suivante :

$$\begin{aligned} S(B) &= \varepsilon^t \varepsilon = \|Y - XB\|^2 \\ &= (Y - XB)^t (Y - XB) = (Y^t - B^t X^t)(Y - XB) \\ &= Y^t Y - 2B^t X^t Y + B^t X^t X B. \end{aligned}$$

En dérivant  $S(B)$  par rapport à  $B$  et on égalisant le résultat à zéro on obtient :

$$\begin{aligned} -2X^t Y + 2X^t X \hat{B} &= 0 \\ \Rightarrow X^t Y - X^t X \hat{B} &= 0 \\ \Rightarrow X^t Y &= X^t X \hat{B} \\ \Rightarrow (X^t X)^{-1} X^t Y &= (X^t X)^{-1} (X^t X) \hat{B} \end{aligned}$$

Donc,

$$\hat{B} = (X^t X)^{-1} X^t Y.$$

Où  $X^t$  désigne la matrice transposée de  $X$ .

#### 1.4.4 Propriétés des Estimateurs

La quantité  $\hat{B}$  forme un estimateur sans biais du paramètre  $B$ . En effet, on a :

$$\begin{aligned} E(\hat{B}) &= (X^t X)^{-1} X^t E(Y) \\ &= (X^t X)^{-1} X^t E(XB + \varepsilon) \end{aligned}$$

$$\text{Avec : } E(XB + \varepsilon) = XB + E(\varepsilon) = XB$$

$$\text{Alors : } E(\hat{B}) = B$$

Pour la variance, on procède de même

$$\begin{aligned} \text{Var}(\hat{B}) &= E[(\hat{B} - B)(\hat{B} - B)^t] \\ &= E[((X^t X)^{-1} X^t \varepsilon) ((X^t X)^{-1} X^t \varepsilon)^t] \\ &= E[(X^t X)^{-1} X^t \varepsilon^t \varepsilon X (X^t X)^{-1}] \\ &= (X^t X)^{-1} X^t E(\varepsilon^t \varepsilon) X (X^t X)^{-1} \\ &= (X^t X)^{-1} X^t \sigma_\varepsilon^2 I_n X (X^t X)^{-1} \\ &= \sigma_\varepsilon^2 (X^t X)^{-1} X^t X (X^t X)^{-1} \end{aligned}$$

$$\text{Donc : } \text{Var}(\hat{B}) = \sigma_\varepsilon^2 (X^t X)^{-1}$$



## Chapitre 2

# Estimation des paramètre d'un PA à saisonnalité régulière

### Introduction

Dans ce chapitre, nous allons considérer un processus stochastique décrit par un modèle additif sous l'hypothèse que la tendance est linéaire et la saisonnalité est régulière c'est-à-dire nous considérons ce qui suit :

$$y_i = at_i + b + c \sin(\alpha t_i + \beta) + \varepsilon_i, \quad i = 1, \dots, n. \quad (2.1)$$

Notre objectif est de discuter la méthodologie d'estimation des paramètres de tel modèle.

### 2.1 Description du modèle

Reprenant le modèle décrit par (2.1). Ce modèle est bien un modèle additif où

**La tendance  $T_t$**  : La tendance est d'écrite par une droite

$$T_t = at + b,$$

avec

- $a$  et  $b$  sont des paramètres réels et qui doivent être estimés.
- $t$  est la variable (non aléatoire) explicative et indépendante définie sur  $\mathbb{R}_+$ .

**La saisonnalité  $S_t$**  : La saisonnalité est régulière, de période  $p = a/2\pi$ , et qui donnée par :

$$S_t = c \sin(\alpha t + \beta),$$

avec

- $c$ ,  $\alpha$  et  $\beta$  sont des paramètres réels et qui doivent être estimés.
- $t$  est la variable (non aléatoire) explicative et indépendante définie sur  $\mathbb{R}_+$ .

Notre objectif est d'illustrer la méthodologie à suivre pour pouvoir estimer les paramètres de tels modèles. Pour atteindre notre objectif, nous allons considérer d'abord le cas sans tendance ( $a = 0$ ) et ceci afin d'expliquer le problème d'application de la méthode MCO d'une manière direct pour l'estimation des paramètres en question particulièrement ceux de  $S_t$ .

Par la suite, nous allons illustrer que l'estimation des paramètres de  $S_t$  par la méthode MCO est possible à condition d'apporté une transformation au modèle originale et de le ramené à un modèle linéaire multiple. Enfin, nous allons exposé l'estimation des paramètres du modèle (2.1) par la méthode de MCO à travers d'un modèle linéaire multiple.

## 2.2 Estimation des paramètres : Cas sans tendance

### 2.2.1 Méthode MCO par approche directe

Concéderons le modèle (2.1) dans le cas sans tendance ce que signifie que le modèle peut être réécrit comme suit :

$$y_i = b + c \sin(\alpha t_i + \beta) + \varepsilon_i, \quad i = 1, \dots, n. \quad (2.2)$$

Ainsi, si on opte pour l'application directe de la méthode MCO pour estimer les paramètres de ce dernier, alors la fonction à minimiser par rapport aux paramètres  $b$ ,  $c$ ,  $\alpha$  et  $\beta$  sera définie par :

$$\Psi(b, c, \alpha, \beta) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - b - c \sin(\alpha t_i + \beta))^2.$$

Par conséquent, l'estimation des paramètres précédents consiste à résoudre le système d'équations suivant :

$$\left( \frac{\partial \Psi}{\partial b} = 0, \frac{\partial \Psi}{\partial c} = 0, \frac{\partial \Psi}{\partial \alpha} = 0, \frac{\partial \Psi}{\partial \beta} = 0 \right).$$

Après le calcul des différentes dérivées on aura les détails suivants :

$$\left\{ \begin{array}{l} \frac{\partial \Psi}{\partial b} = -2 \sum_{i=1}^n (y_i - b - c \sin(\alpha t_i + \beta)) = 0 \\ \frac{\partial \Psi}{\partial c} = -2 \sum_{i=1}^n \sin(\alpha t_i + \beta) (y_i - b - c \sin(\alpha t_i + \beta)) = 0 \\ \frac{\partial \Psi}{\partial \alpha} = -2c \sum_{i=1}^n t_i \cos(\alpha t_i + \beta) (y_i - b - c \sin(\alpha t_i + \beta)) = 0 \\ \frac{\partial \Psi}{\partial \beta} = -2c \sum_{i=1}^n \cos(\alpha t_i + \beta) (y_i - b - c \sin(\alpha t_i + \beta)) = 0 \end{array} \right.$$

$$\Rightarrow \begin{cases} \sum_{i=1}^n (y_i - b - c \sin(\alpha t_i + \beta)) = 0 \\ \sum_{i=1}^n \sin(\alpha t_i + \beta) (y_i - b - c \sin(\alpha t_i + \beta)) = 0 \\ c \sum_{i=1}^n t_i \cos(\alpha t_i + \beta) (y_i - b - c \sin(\alpha t_i + \beta)) = 0 \\ c \sum_{i=1}^n \cos(\alpha t_i + \beta) (y_i - b - c \sin(\alpha t_i + \beta)) = 0 \end{cases}$$

La résolution de ce système à priori est difficile. Mais si considère l'hypothèse de la saisonnalité régulière et parfaite, c'est-à-dire  $E(\sin(\alpha t_i + \beta)) = 0$ , alors

$$E(Y) = E(b + c \sin(\alpha t + \beta)) = b + cE(\sin(\alpha t + \beta)) \Rightarrow E(Y) = b,$$

ce qui nous laisse à proposé l'estimateur suivant :

$$\hat{b} = \bar{Y},$$

et pour le paramètre  $c$  on peut envisagé les estimation suivante :

$$\hat{c} = \max_i \{|y_i - \hat{b}|\} \quad \text{ou encore} \quad \hat{c} = \frac{\max_i \{y_i - \hat{b}\} + |\min_i \{y_i - \hat{b}\}|}{2}.$$

Ainsi,

$$y_i = \hat{b} + \hat{c} \sin(\alpha t_i + \beta)$$

Il nous reste à estimer les deux paramètre  $\alpha$  et  $\beta$  pour cela soit la transformation suivante :

$$y_i = \hat{b} + \hat{c} \sin(\alpha t_i + \beta) \Rightarrow \frac{y_i - \hat{b}}{\hat{c}} = \sin(\alpha t_i + \beta) \Rightarrow \alpha t_i + \beta = \arcsin \left( \frac{y_i - \hat{b}}{\hat{c}} \right).$$

Posons  $w_i = \arcsin\left(\frac{y_i - \hat{b}}{\hat{c}}\right)$  alors on aura ce qui suit :

$$w_i = \alpha t_i + \beta.$$

Il est claire que ce dernier n'est rien d'autre qu'un modèle linéaire simple, ainsi d'après la méthode MCO, on aura :

$$\begin{cases} \hat{\alpha} = \frac{\text{cov}(w_i, t_i)}{\sigma_t^2}, \\ \hat{\beta} = \bar{w} - \hat{\alpha} \bar{t}, \end{cases}$$

Mais malheureusement, ces dernières estimations ne sont pas efficace car, quel que soit l'échantillon, elle nous fournies des estimations proche de zéro ( $\hat{\alpha} = 0$  et  $\hat{\beta} = 0$ ). Cette constatation peut être justifier et confirmé par le fait que le nuage des points  $((t_i, w_i))$  est cerné par deux droite parallèles à l'axe des abscisses (voir figure 2.1.(D)). La figure 2.1 présente une illustration des différentes variables considérées précédemment dans le cas où

$$y_t = 5 + 3 * \sin(0.5 * t + 10) + \epsilon_t.$$

avec  $\epsilon_t \rightarrow N(0, 3^2)$  et  $t = \overline{0 : 50}$  (c'st-à-dire  $n = 51$ ).

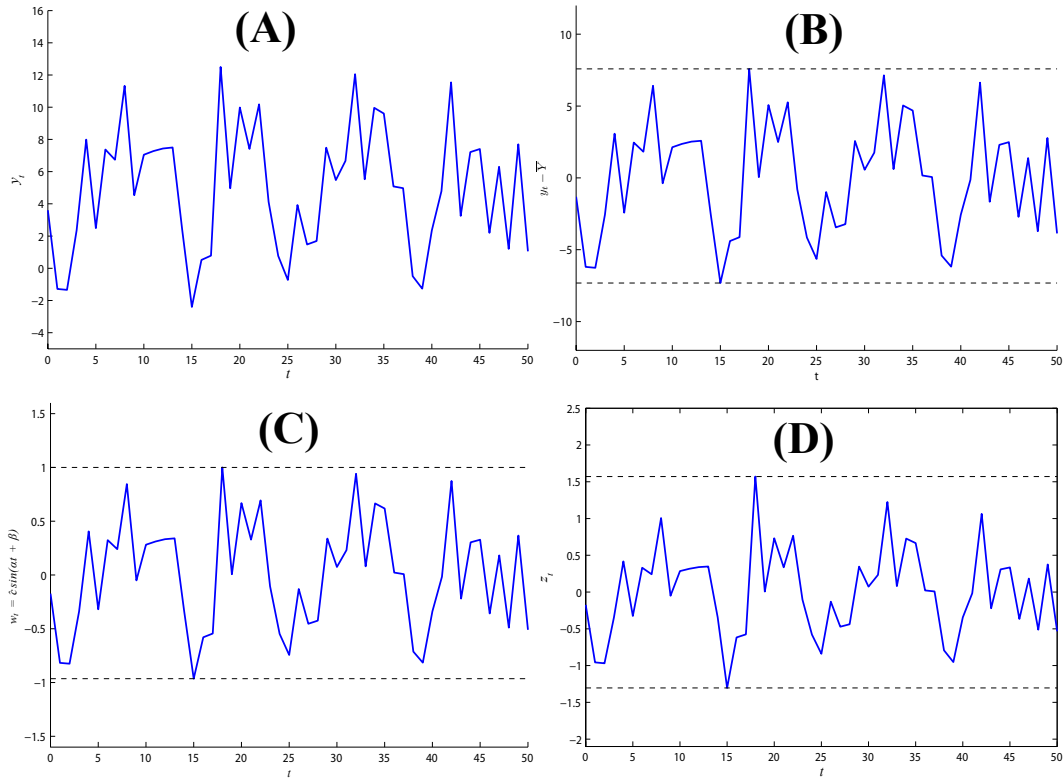


FIGURE 2.1 – Exemple d’illustration du Modèle retenu : (A) la série, (B) La série centrée et (C) Série centrée et réduite, (D) la nouvelle série

### 2.2.2 Estimation des paramètres via un modèle linéaire multiple

La tendance générale est parfois liée à l’évidence à un cycle (annuel, lunaire, etc.). La régression sinusoidale permet de tester statistiquement, et de définir un modèle de variation rigoureusement périodique. Naturellement, ce modèle doit être utilisé avec discernement. En effet, une inspection de la littérature sur les séries chronologique nous permet de se rendre compte et de noter que même si la variabilité saisonnière existe pour des abondances d’espèces, elle ne s’ajuste pas forcément à une sinusoïde si les amplitudes maximums sont décalées d’une année à l’autre, d’un mois ou plus. Et si alors on étudie les écarts entre les valeurs observés et le modèle, ceux-ci ne correspondent plus à une variabilité indépendante de la variation saisonnière. Ils ne

représentent que l'inadéquation du modèle choisi, qui lui, est strictement périodique. Grâce à une astuce, nous pouvons ajuster un modèle sinusoïdal, décrit dans (2.2), à l'aide de modèle linéaire où il suffit de rendre l'équation linéaire par un changement de variables. Rappelons que

$$\sin(\alpha + \beta) = \cos(\alpha) \sin(\beta) + \cos(\beta) \sin(\alpha)$$

Donc, notre modèle peut s'écrire selon l'équation suivante

$$y_t = b + c \sin(\alpha t + \beta) \tag{2.3}$$

$$= b + c \cos(\alpha t) \sin(\beta) + c \cos(\beta) \sin(\alpha t) \tag{2.4}$$

Si on pose

$$T_1 = \cos(\alpha t), \quad T_2 = \sin(\alpha t), \quad c_1 = c \sin(\beta), \quad c_2 = c \cos(\beta).$$

Alors, l'écriture précédente devienne

$$y_t = b + c_1 \cdot T_1 + c_2 \cdot T_2$$

Cette dernière forme est assimilable à une régression multiple sur deux variables indépendantes  $T_1$  et  $T_2$  dont l'ordonnée à l'origine vaut  $b$ , et que nous pouvons ajuster facilement à l'aide de la régression linéaire multiple. Ainsi, on peut estimer facilement les trois paramètres  $b$ ,  $c_1$  et  $c_2$  via la méthode MCO (voir chapitre 1, section 1.4.3) où

$$\hat{B} = (T^t T)^{-1} T^t Y.$$

avec

$$\hat{B} = \begin{pmatrix} \hat{b} \\ \hat{c}_1 \\ \hat{c}_2 \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad T = \begin{pmatrix} 1 & t_{1,1} & t_{1,2} \\ 1 & t_{2,1} & t_{2,2} \\ \vdots & \vdots & \vdots \\ 1 & t_{n,1} & t_{n,2} \end{pmatrix}$$

où  $t_{i,1} = \sin(\alpha t_i)$  et  $t_{i,2} = \cos(\alpha t_i)$ . Par conséquent, les estimations des paramètres du modèle original seront quantifiées comme suit :

$$\begin{cases} \hat{b} = \hat{b}, \\ \hat{\beta} = \arctan\left(\frac{\hat{c}_1}{\hat{c}_2}\right), \\ \hat{c} = \frac{\hat{c}_1}{\sin(\hat{\beta})} \text{ ou } \hat{c} = \frac{\hat{c}_2}{\cos(\hat{\beta})}. \end{cases}$$

## 2.3 Estimation des paramètres : cas avec tendance

Reprenant le modèle introduit tout au début du présent chapitre.

$$y_i = at_i + b + c \sin(\alpha t_i + \beta) + \varepsilon_i, \quad i = 1, \dots, n.$$

Si on retient à nouveau l'idée exposée dans la section 2.2.2 alors le présent modèle peut être réécrit sous sa forme linéaire multiple comme suit :

$$y_i = at_i + b + c_1 \cdot T_{i,1} + c_2 \cdot T_{i,2} + \varepsilon_i, \quad i = 1, \dots, n.$$

avec

$$T_{i,1} = \cos(\alpha t_i), \quad T_{i,2} = \sin(\alpha t_i), \quad c_1 = c \sin(\beta), \quad c_2 = c \cos(\beta).$$

Ainsi, d'après la méthode MCO l'estimateur du vecteur  $(a, b, c_1, c_2)$  est donnée par :

$$\hat{B} = (T^t T)^{-1} T^t Y.$$



avec

$$\hat{B} = \begin{pmatrix} \hat{b} \\ \hat{a} \\ \hat{c}_1 \\ \hat{c}_2 \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad T = \begin{pmatrix} 1 & t_1 & t_{1,1} & t_{1,2} \\ 1 & t_2 & t_{2,1} & t_{2,2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & t_{n,1} & t_{n,2} \end{pmatrix}$$

où  $t_{i,1} = \sin(\alpha t_i)$  et  $t_{i,2} = \cos(\alpha t_i)$ . Par conséquent, les estimations des paramètres du modèle original seront quantifiées comme suit :

$$\begin{cases} \hat{a} = \hat{a}, \\ \hat{b} = \hat{b}, \\ \hat{\beta} = \arctan\left(\frac{\hat{c}_1}{\hat{c}_2}\right), \\ \hat{c} = \frac{\hat{c}_1}{\sin(\hat{\beta})} \text{ ou } \hat{c} = \frac{\hat{c}_2}{\cos(\hat{\beta})}. \end{cases}$$

**Remarque 2.3.1** (*Estimation de  $\alpha$* ) Pour mettre en évidence une éventuelle périodicité dans les résidus, on peut également tracer le périodogramme, qui est le graphe de la projection du vecteur  $X = t(X_1, \dots, X_T)$  sur le sous- vectoriel de  $\mathbb{R}^T$  engendré par les deux vecteurs

$$C = \begin{pmatrix} \cos(\alpha t_1) \\ \vdots \\ \cos(\alpha t_T) \end{pmatrix} \quad \text{et} \quad S = \begin{pmatrix} \sin(\alpha t_1) \\ \vdots \\ \sin(\alpha t_T) \end{pmatrix},$$

où  $\alpha$  est la fréquence. On cherche ainsi à ajuster à la série  $X_t$  le  $t^{\text{ème}}$  terme d'une série trigonométrique, appelée une harmonique, et s'écrivant

$$s(t) = a \cos(\alpha \times t) + b \sin(\alpha \times t).$$

Dans le cas où la série  $X_t$  est périodique de période  $P$ , on a  $P = 2\pi/\alpha$  par conséquent

$\alpha = 2\pi/P$ . Une approximation de la norme de cette projection est

$$Q(\alpha) = \frac{T}{2}(\hat{a}^2 + \hat{b}^2),$$

où

$$\hat{a}^2 = \frac{2}{T} \sum_i X_{t_i} \cos(\alpha t_i) \quad \text{et} \quad \hat{b}^2 = \frac{2}{T} \sum_i X_{t_i} \sin(\alpha t_i).$$

Le périodogramme est donc le graphe de la fonction  $(\alpha, Q(\alpha))$ . On cherche alors la valeur  $\alpha^*$  qui maximise  $Q(\alpha)$ . Si le périodogramme présente un pic en  $\alpha^*$ , on lit cette valeur sur son graphe et on retient une période  $T^* = 2\pi/\alpha^*$ , ou plus exactement la valeur entière la plus proche de cette expression.

# Chapitre 3

## Application numérique

### Introduction

Dans le présent chapitre, nous avons réalisés une étude de simulation qui vise à analyser les propriétés statistique des estimateurs des paramètres d'un modèle décrivant une série chronologique gouvernée par une tendance linéaire et une saisonnalité régulière dont la forme est :

$$y_t = at + b + c \sin(\alpha t + \beta) + \epsilon, \quad t \in \mathbb{R}_+. \quad (3.1)$$

Plus précisément, nous allons intéresser dans un premier lieu à l'effet de la taille de l'échantillon et de la variance des résidus sur la qualité des estimations ainsi que la qualité du régresseur conçu. Par la suite, nous allons vérifier la possibilité de modéliser d'autres séries chronologiques décrites par autre modèle dont l'allure de ce dernier est similaire au modèle (3.1) mais pas forcément identique.

### 3.1 Première application

L'objectif de la présente application est de vérifier l'effet de la taille de l'échantillon et de la variance des résidus sur la qualité des estimateurs des paramètres du modèle et la qualité du régresser obtenus. Pour mesurer la qualité de ce dernier nous sommes optés pour l'Erreur quadratique moyenne défini par :

$$EQM = \frac{1}{T} \int (y_t - \hat{y}_t)^2 dt. \quad (3.2)$$

Afin de répondre à notre objectif nous avons conçu un programme MATLAB dont les principales étapes sont :

**Étape 1 :** Générer un  $n$ -échantillon  $(x_i, y_i)$  selon un modèle cible.

**Étape 2 :** Calculer  $T_1$  et  $T_2$ , matrice  $X$  (pour effectuer la transformation du modèle (3.1) vers un modèle linéaire multiples).

**Étape 3 :** Estimer les paramètres du modèle linéaire multiples.

**Étape 4 :** Calculer les estimations des paramètres du modèle original

**Étape 5 :** Calculer l'EQM associé au régresser obtenus.

Pour les calculs numériques nous avons considéré les quatre exemples suivants :

**Exemple 1 :**  $a = 0, b = 2, c = 50, \alpha = 0.2, \beta = \pi/6$ . (Sans tendance)

**Exemple 2 :**  $a = 0, b = 2, c = 50, \alpha = 0.4, \beta = \pi/6$ .

**Exemple 3 :**  $a = 3, b = 2, c = 50, \alpha = 0.2, \beta = \pi/6$ .

**Exemple 4 :**  $a = 3, b = 2, c = 50, \alpha = 0.4, \beta = \pi/6$ .

L'allure graphique des deux modèles décrits dans l'exemple 3 et l'exemple 4 sont présentés dans les figures 3.1 et 3.2, respectivement.

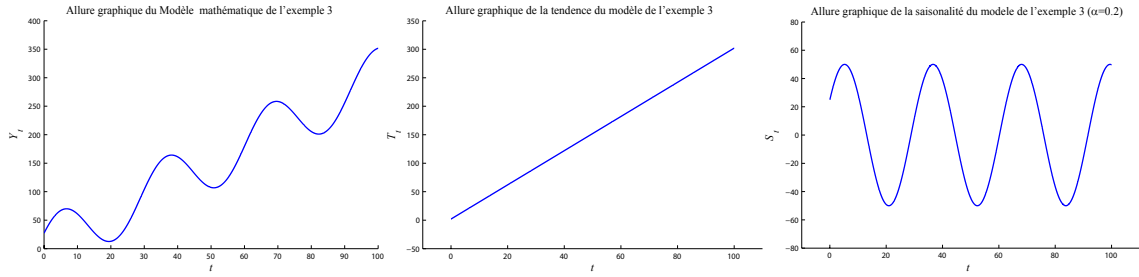


FIGURE 3.1 – Illustration graphique du Modèle de l'exemple 3 et ses composantes : la série, la tendance et la saisonnalité

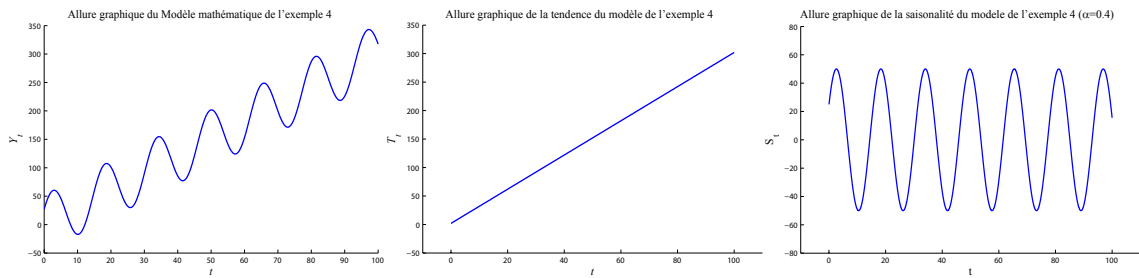


FIGURE 3.2 – Illustration graphique du Modèle de l'exemple 4 et ses composantes : la série, la tendance et la saisonnalité

L'exécution du programme conçu pour les différents paramètres ci-dessus et pour des tailles d'échantillon  $n \in \{10, 20, 30, 40, 50, 100\}$ , pour  $t \in [0 : 100/(n-1) : 100]$  et pour différentes valeurs de  $\sigma \in \{10, 1, 0.5, 0.4, 0.3\}$  nous a fournis les résultats rangés dans les Tables 3.1–3.6. Dans les colonnes 3 à 5 des Tables 3.4–3.6, la première valeur de chaque colonne représente la borne inférieure de l'intervalle de confiance à un seuil de 95% du paramètre considéré dans l'estimation, la deuxième valeur représente l'estimateur ponctuelle et la troisième valeur représente la borne supérieure de l'intervalle de confiance à un seuil de 95% du paramètre considéré. Les figures 3.2–3.4 représentent la variation de l'EQM en fonction de la taille de l'échantillon et de l'écart-type des  $\epsilon_t$ .

$\sigma_\epsilon$	n	Cas d'exemple 1				Cas d'exemple 2			
		$\hat{a}$	$\hat{C}_1$	$\hat{C}_2$	$\hat{b}$	$\hat{a}$	$\hat{C}_1$	$\hat{C}_2$	$\hat{b}$
5.00	5	0.0083	42.9614	24.8788	1.4918	0.0051	43.1769	24.7533	1.7489
	10	0.0007	43.1630	24.9202	2.1069	-0.0026	43.3090	24.9614	1.9786
	15	0.0018	43.1949	25.0094	1.7269	-0.0016	43.4226	25.2254	2.1661
	20	0.0032	43.3215	25.0042	1.7613	-0.0008	43.4121	25.1192	2.0171
	25	-0.0020	43.2445	24.9747	2.1742	0.0026	43.1762	25.1153	1.9379
	50	0.0014	43.3028	25.0505	1.7690	0.0008	43.2850	25.0048	2.0188
2.00	5	0.0010	43.1869	25.0212	1.9262	0.0008	43.4285	25.0897	2.0213
	10	-0.0009	43.2485	24.9476	2.0716	0.0027	43.2904	25.0699	1.7955
	15	0.0001	43.3046	25.0079	1.9715	-0.0004	43.2707	25.0614	2.0563
	20	-0.0002	43.3073	24.9816	1.9870	-0.0001	43.2577	24.9992	2.0269
	25	-0.0026	43.3134	24.9998	2.0699	-0.0001	43.3077	25.0137	1.9650
	50	-0.0009	43.2722	25.0347	2.0382	0.0002	43.3084	24.9542	1.9838
1.00	5	0.0002	43.3195	25.0737	1.9384	-0.0013	43.3535	24.9961	2.0472
	10	0.0008	43.2552	24.9815	1.9572	0.0001	43.2559	25.0450	1.9861
	15	0.0004	43.2621	24.9926	1.9500	-0.0002	43.3011	24.9982	2.0342
	20	0.0005	43.2771	24.9906	1.9742	0.0004	43.2982	24.9913	1.9804
	25	-0.0007	43.2968	25.0058	2.0477	0.0000	43.3118	25.0203	1.9984
	50	0.0002	43.2953	24.9977	1.9975	-0.0007	43.2910	25.0082	2.0283
0.50	5	-0.0004	43.3278	25.0131	2.0076	-0.0006	43.3156	24.9597	2.0347
	10	0.0005	43.2917	24.9896	1.9823	0.0004	43.2996	25.0154	1.9735
	15	-0.0001	43.3032	24.9992	2.0177	0.0000	43.2894	24.9813	1.9913
	20	0.0002	43.3100	24.9838	1.9901	-0.0002	43.2984	25.0029	2.0190
	25	-0.0002	43.3250	25.0018	2.0160	0.0003	43.3052	25.0159	1.9756
	50	-0.0001	43.2893	24.9995	1.9996	0.0002	43.3005	24.9972	1.9916
0.25	5	0.0000	43.2982	25.0076	1.9964	-0.0003	43.3025	25.0157	2.0169
	10	-0.0002	43.2996	25.0059	2.0125	-0.0001	43.3113	25.0091	1.9923
	15	0.0001	43.3079	24.9982	1.9984	0.0000	43.2998	25.0064	2.0040
	20	-0.0002	43.3026	24.9936	2.0084	0.0000	43.3059	24.9982	2.0004
	25	0.0003	43.2977	25.0122	1.9816	0.0003	43.3047	24.9989	1.9947
	50	0.0001	43.3031	25.0002	1.9908	-0.0001	43.3060	24.9980	2.0034

TABLE 3.1: Estimateurs des paramètres du modèle linéaire multiple : Cas de l'exemple 1 et de l'exemple 2

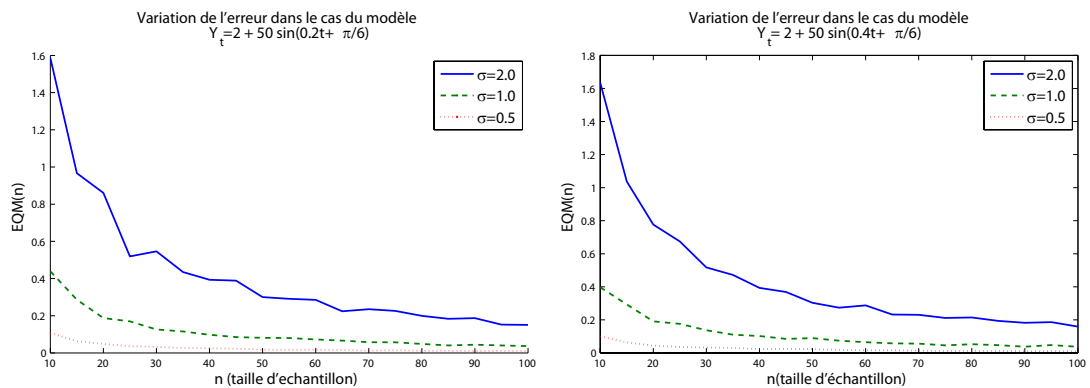


FIGURE 3.3 – Variation de l'EQM en fonction de la taille de l'échantillon et de l'écart-type des résidus (1/2).

$\sigma_\epsilon$	$n$	Cas d'exemple 3				Cas d'exemple 4			
		$\hat{a}$	$\hat{C}_1$	$\hat{C}_2$	$\hat{b}$	$\hat{a}$	$\hat{C}_1$	$\hat{C}_2$	$\hat{b}$
5.00	5	3.0077	43.6112	24.7581	1.8977	3.0114	43.0295	25.5957	1.4266
	10	2.9936	43.1727	24.9381	2.3925	2.9975	43.2722	25.1450	2.0784
	15	3.0028	43.2301	25.1205	1.7930	2.9956	43.5730	25.0487	2.3050
	20	3.0015	43.2140	25.0404	2.0304	3.0024	43.1565	25.1678	1.9287
	25	2.9986	43.3267	25.0385	2.0483	2.9997	43.2288	24.9814	1.9896
	50	3.0040	43.2917	25.0032	1.7232	2.9984	43.3625	25.0513	2.1526
2.00	5	2.9965	43.3338	25.0548	2.1906	3.0030	43.1718	24.7993	1.8145
	10	3.0020	43.2740	24.9807	1.9742	2.9996	43.2609	25.0203	1.9939
	15	2.9993	43.3023	24.9685	2.0861	2.9988	43.2985	24.9948	1.9826
	20	2.9996	43.2700	25.0278	2.0883	2.9998	43.3506	25.0180	1.9746
	25	2.9989	43.3381	25.0046	2.0631	3.0021	43.2457	25.0282	1.8943
	50	3.0005	43.3594	25.0046	1.9800	3.0001	43.2967	25.0013	2.0316
1.00	5	3.0010	43.3225	24.9868	1.9687	3.0002	43.3053	24.9962	1.9500
	10	3.0004	43.3490	24.9700	1.9565	2.9996	43.3525	24.9338	2.0235
	15	3.0012	43.3541	24.9931	1.9342	2.9995	43.2771	25.0185	2.0364
	20	3.0005	43.3618	25.0095	1.9741	2.9998	43.3203	25.0068	1.9883
	25	3.0004	43.2970	24.9829	1.9732	2.9995	43.2926	24.9741	2.0435
	50	3.0002	43.3019	24.9512	1.9971	3.0000	43.2896	24.9885	2.0135
0.50	5	2.9995	43.2997	24.9955	2.0389	2.9996	43.2757	24.9928	2.0185
	10	2.9999	43.2906	24.9896	1.9971	3.0002	43.3189	24.9993	1.9916
	15	3.0007	43.3117	24.9876	1.9667	2.9999	43.2922	24.9937	1.9831
	20	2.9997	43.2911	24.9960	2.0172	2.9997	43.3088	24.9994	2.0039
	25	2.9998	43.3027	24.9966	2.0037	3.0001	43.2905	25.0053	1.9927
	50	2.9997	43.2932	25.0164	2.0180	2.9999	43.3053	25.0017	2.0015
0.25	5	2.9996	43.3134	25.0026	2.0270	2.9999	43.3273	24.9868	2.0044
	10	2.9995	43.3070	25.0010	2.0226	3.0002	43.2922	25.0093	2.0006
	15	3.0001	43.3085	24.9945	1.9978	3.0001	43.3021	25.0017	1.9983
	20	3.0001	43.3074	24.9921	1.9918	3.0001	43.2921	24.9887	1.9952
	25	2.9998	43.3046	24.9955	2.0140	2.9998	43.3049	25.0061	2.0103
	50	3.0001	43.2999	25.0003	1.9957	3.0001	43.3002	25.0002	2.0015

TABLE 3.2: Estimateurs des paramètres du modèle linéaire multiple : Cas de l'exemple 3 et de l'exemple 4

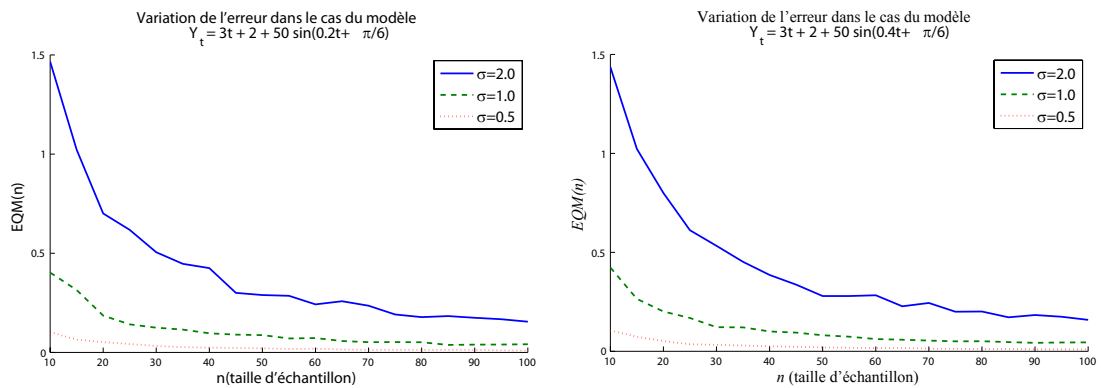


FIGURE 3.4 – Variation de l'EQM en fonction de la taille de l'échantillon et de l'écart-type des résidus (2/2).

### Chapitre 3. Application numérique

$\sigma_\epsilon$	$n$	$\hat{a}$			$\hat{b}$			$\hat{c}$			$\hat{\beta}$			EQM
5.00	5	-0.0020	0.0083	0.0185	0.917	1.492	2.067	49.266	49.692	50.119	0.516	0.525	0.534	8.983
	10	-0.0065	0.0007	0.0079	1.712	2.107	2.502	49.577	49.873	50.170	0.516	0.524	0.531	4.989
	15	-0.0032	0.0018	0.0067	1.427	1.727	2.027	49.696	49.932	50.169	0.519	0.525	0.531	3.165
	20	-0.0021	0.0032	0.0085	1.426	1.761	2.096	49.814	50.033	50.252	0.519	0.523	0.528	2.519
	25	-0.0062	-0.0020	0.0021	1.953	2.174	2.395	49.731	49.948	50.165	0.520	0.524	0.528	2.002
	50	-0.0025	0.0014	0.0052	1.552	1.769	1.986	49.886	50.031	50.175	0.522	0.524	0.527	1.058
1.00	5	-0.0016	0.0002	0.0021	1.828	1.938	2.048	49.958	50.054	50.150	0.523	0.525	0.526	0.374
	10	-0.0006	0.0008	0.0021	1.874	1.957	2.040	49.887	49.952	50.016	0.522	0.524	0.525	0.188
	15	-0.0008	0.0004	0.0016	1.873	1.950	2.027	49.914	49.963	50.012	0.523	0.524	0.525	0.132
	20	-0.0005	0.0005	0.0016	1.913	1.974	2.036	49.930	49.975	50.020	0.523	0.524	0.525	0.104
	25	-0.0017	-0.0007	0.0002	1.988	2.048	2.108	49.960	49.999	50.039	0.523	0.524	0.525	0.085
	50	-0.0005	0.0002	0.0009	1.956	1.998	2.039	49.967	49.994	50.020	0.523	0.524	0.524	0.040
0.50	5	-0.0014	-0.0004	0.0005	1.952	2.008	2.064	49.986	50.030	50.074	0.523	0.524	0.525	0.108
	10	-0.0003	0.0005	0.0012	1.937	1.982	2.028	49.957	49.987	50.016	0.523	0.524	0.524	0.045
	15	-0.0008	-0.0001	0.0006	1.973	2.018	2.062	49.977	50.001	50.026	0.523	0.524	0.524	0.032
	20	-0.0003	0.0002	0.0007	1.959	1.990	2.021	49.976	50.000	50.023	0.523	0.523	0.524	0.025
	25	-0.0007	-0.0002	0.0003	1.988	2.016	2.044	50.001	50.022	50.042	0.523	0.523	0.524	0.022
	50	-0.0004	-0.0001	0.0003	1.978	2.000	2.021	49.977	49.989	50.002	0.523	0.524	0.524	0.011
0.25	5	-0.0005	0.0000	0.0004	1.969	1.996	2.024	49.978	50.001	50.024	0.523	0.524	0.524	0.023
	10	-0.0006	-0.0002	0.0002	1.987	2.012	2.038	49.985	50.002	50.019	0.523	0.524	0.524	0.014
	15	-0.0002	0.0001	0.0004	1.980	1.998	2.017	49.993	50.005	50.017	0.523	0.524	0.524	0.008
	20	-0.0005	-0.0002	0.0000	1.993	2.008	2.024	49.987	49.998	50.009	0.523	0.523	0.524	0.007
	25	0.0001	0.0003	0.0006	1.968	1.982	1.996	49.994	50.003	50.012	0.524	0.524	0.524	0.005
	50	-0.0001	0.0001	0.0003	1.980	1.991	2.002	49.995	50.002	50.008	0.523	0.524	0.524	0.002

TABLE 3.3: Estimateurs des paramètres du modèle original : Cas d'exemple 1

$\sigma_\epsilon$	$n$	$\hat{a}$			$\hat{b}$			$\hat{c}$			$\hat{\beta}$			EQM
5.00	5	-0.0052	0.0051	0.0154	1.136	1.749	2.361	49.310	49.813	50.316	0.512	0.520	0.529	10.280
	10	-0.0114	-0.0026	0.0062	1.485	1.979	2.473	49.717	50.007	50.297	0.517	0.523	0.528	4.890
	15	-0.0072	-0.0016	0.0040	1.824	2.166	2.508	49.989	50.234	50.478	0.521	0.526	0.531	3.068
	20	-0.0068	-0.0008	0.0052	1.672	2.017	2.363	49.922	50.168	50.415	0.520	0.525	0.529	2.785
	25	-0.0018	0.0026	0.0069	1.676	1.938	2.199	49.749	49.961	50.172	0.523	0.527	0.531	2.075
	50	-0.0025	0.0008	0.0042	1.816	2.019	2.222	49.848	49.993	50.138	0.521	0.524	0.527	0.999
1.00	5	-0.0034	-0.0013	0.0008	1.924	2.047	2.170	49.946	50.045	50.145	0.521	0.523	0.525	0.432
	10	-0.0013	0.0001	0.0016	1.900	1.986	2.072	49.917	49.984	50.052	0.524	0.525	0.526	0.204
	15	-0.0015	-0.0002	0.0010	1.959	2.034	2.109	49.955	50.000	50.044	0.523	0.524	0.525	0.128
	20	-0.0007	0.0004	0.0015	1.919	1.980	2.042	49.945	49.993	50.042	0.523	0.523	0.524	0.093
	25	-0.0010	0.0000	0.0010	1.941	1.998	2.056	49.981	50.020	50.058	0.523	0.524	0.525	0.085
	50	-0.0014	-0.0007	-0.0001	1.992	2.028	2.064	49.969	49.995	50.022	0.523	0.524	0.524	0.034
0.50	5	-0.0017	-0.0006	0.0005	1.973	2.035	2.097	49.943	49.993	50.043	0.522	0.523	0.524	0.110
	10	-0.0004	0.0004	0.0011	1.932	1.974	2.015	49.974	50.006	50.039	0.523	0.524	0.524	0.047
	15	-0.0005	0.0000	0.0006	1.958	1.991	2.025	49.955	49.981	50.006	0.523	0.523	0.524	0.031
	20	-0.0007	-0.0002	0.0003	1.988	2.019	2.049	49.978	49.999	50.020	0.523	0.524	0.524	0.022
	25	-0.0001	0.0003	0.0007	1.949	1.976	2.003	49.994	50.012	50.029	0.523	0.524	0.524	0.019
	50	-0.0001	0.0002	0.0005	1.973	1.992	2.010	49.983	49.998	50.013	0.523	0.524	0.524	0.010
0.25	5	-0.0008	-0.0003	0.0002	1.988	2.017	2.046	49.989	50.009	50.029	0.523	0.524	0.524	0.023
	10	-0.0005	-0.0001	0.0003	1.969	1.992	2.015	49.997	50.013	50.030	0.523	0.524	0.524	0.014
	15	-0.0003	0.0000	0.0002	1.988	2.004	2.020	49.989	50.002	50.015	0.523	0.524	0.524	0.008
	20	-0.0002	0.0000	0.0003	1.987	2.000	2.014	49.992	50.003	50.014	0.523	0.524	0.524	0.006
	25	0.0001	0.0003	0.0005	1.981	1.995	2.008	49.992	50.002	50.012	0.523	0.524	0.524	0.005
	50	-0.0002	-0.0001	0.0001	1.993	2.003	2.013	49.996	50.003	50.010	0.523	0.524	0.524	0.003

TABLE 3.4: Estimateurs des paramètres du modèle original : Cas d'exemple 2



### Chapitre 3. Application numérique

$\sigma_\epsilon$	$n$	$\hat{a}$			$\hat{b}$			$\hat{c}$			$\hat{\beta}$			EQM
5.00	5	2.999	3.008	3.016	1.353	1.898	2.442	49.666	50.200	50.734	0.507	0.516	0.525	10.438
	10	2.986	2.994	3.001	1.948	2.392	2.837	49.589	49.882	50.175	0.518	0.524	0.530	4.577
	15	2.997	3.003	3.009	1.437	1.793	2.149	49.767	50.018	50.270	0.521	0.526	0.532	3.253
	20	2.996	3.001	3.007	1.702	2.030	2.359	49.731	49.958	50.184	0.521	0.525	0.530	2.467
	25	2.994	2.999	3.003	1.786	2.048	2.311	49.854	50.050	50.246	0.520	0.524	0.528	1.738
	50	3.000	3.004	3.008	1.520	1.723	1.927	49.849	49.998	50.148	0.521	0.524	0.527	1.037
1.00	5	2.999	3.001	3.003	1.850	1.969	2.087	49.926	50.014	50.102	0.521	0.523	0.525	0.382
	10	2.999	3.000	3.002	1.864	1.957	2.049	49.961	50.027	50.094	0.521	0.523	0.524	0.203
	15	3.000	3.001	3.002	1.863	1.934	2.005	49.993	50.043	50.093	0.522	0.523	0.524	0.125
	20	3.000	3.000	3.001	1.918	1.974	2.030	50.007	50.058	50.109	0.522	0.523	0.524	0.105
	25	3.000	3.000	3.001	1.920	1.973	2.027	49.946	49.988	50.030	0.523	0.523	0.524	0.081
	50	2.999	3.000	3.001	1.952	1.997	2.042	49.946	49.976	50.007	0.522	0.523	0.523	0.043
0.50	5	2.999	2.999	3.000	1.984	2.039	2.093	49.950	49.997	50.044	0.523	0.524	0.524	0.099
	10	2.999	3.000	3.001	1.958	1.997	2.036	49.956	49.986	50.016	0.523	0.524	0.524	0.044
	15	3.000	3.001	3.001	1.932	1.967	2.002	49.974	50.003	50.032	0.523	0.523	0.524	0.034
	20	2.999	3.000	3.000	1.982	2.017	2.052	49.968	49.989	50.010	0.523	0.524	0.524	0.025
	25	2.999	3.000	3.000	1.972	2.004	2.035	49.981	50.000	50.018	0.523	0.524	0.524	0.022
	50	2.999	3.000	3.000	1.997	2.018	2.039	49.988	50.001	50.014	0.524	0.524	0.524	0.009
0.25	5	2.999	3.000	3.000	2.000	2.027	2.054	49.988	50.012	50.036	0.523	0.524	0.524	0.027
	10	2.999	3.000	3.000	2.001	2.023	2.044	49.992	50.005	50.019	0.523	0.524	0.524	0.011
	15	3.000	3.000	3.000	1.980	1.998	2.016	49.991	50.004	50.017	0.523	0.523	0.524	0.009
	20	3.000	3.000	3.000	1.974	1.992	2.010	49.990	50.001	50.013	0.523	0.523	0.524	0.007
	25	3.000	3.000	3.000	2.000	2.014	2.028	49.990	50.001	50.012	0.523	0.523	0.524	0.006
	50	3.000	3.000	3.000	1.986	1.996	2.006	49.993	49.999	50.005	0.523	0.524	0.524	0.002

TABLE 3.5: Estimateurs des paramètres du modèle original : Cas d'exemple 3

$\sigma_\epsilon$	$n$	$\hat{a}$			$\hat{b}$			$\hat{c}$			$\hat{\beta}$			EQM
5.00	5	3.001	3.011	3.022	0.736	1.427	2.117	49.635	50.116	50.597	0.528	0.536	0.545	11.183
	10	2.990	2.998	3.006	1.638	2.078	2.519	49.740	50.073	50.407	0.520	0.526	0.533	5.305
	15	2.990	2.996	3.002	1.940	2.305	2.670	50.038	50.279	50.520	0.516	0.522	0.527	3.526
	20	2.997	3.002	3.008	1.606	1.929	2.252	49.752	49.973	50.194	0.523	0.528	0.533	2.637
	25	2.995	3.000	3.005	1.691	1.990	2.288	49.725	49.935	50.146	0.521	0.524	0.527	1.904
	50	2.995	2.998	3.002	1.941	2.153	2.364	49.953	50.085	50.216	0.521	0.524	0.527	1.031
1.00	5	2.998	3.000	3.002	1.827	1.950	2.073	49.917	50.004	50.091	0.522	0.524	0.525	0.407
	10	2.998	3.000	3.001	1.937	2.024	2.110	49.939	50.012	50.086	0.521	0.522	0.523	0.201
	15	2.998	3.000	3.001	1.967	2.036	2.106	49.935	49.989	50.044	0.523	0.524	0.525	0.137
	20	2.999	3.000	3.001	1.928	1.988	2.049	49.981	50.020	50.060	0.523	0.524	0.524	0.094
	25	2.999	3.000	3.000	1.986	2.044	2.101	49.934	49.980	50.026	0.523	0.523	0.524	0.087
	50	2.999	3.000	3.001	1.971	2.014	2.056	49.957	49.984	50.011	0.523	0.524	0.524	0.036
0.50	5	2.998	3.000	3.001	1.953	2.019	2.084	49.928	49.975	50.022	0.523	0.524	0.525	0.109
	10	2.999	3.000	3.001	1.948	1.992	2.035	49.980	50.015	50.050	0.523	0.523	0.524	0.057
	15	2.999	3.000	3.000	1.951	1.983	2.016	49.964	49.989	50.015	0.523	0.524	0.524	0.029
	20	2.999	3.000	3.000	1.977	2.004	2.031	49.980	50.006	50.033	0.523	0.524	0.524	0.026
	25	3.000	3.000	3.001	1.967	1.993	2.018	49.975	49.993	50.012	0.523	0.524	0.524	0.018
	50	3.000	3.000	3.000	1.983	2.002	2.020	49.992	50.004	50.017	0.523	0.524	0.524	0.008
0.25	5	2.999	3.000	3.000	1.976	2.004	2.033	49.992	50.016	50.041	0.523	0.523	0.524	0.027
	10	3.000	3.000	3.001	1.981	2.001	2.021	49.981	49.997	50.013	0.524	0.524	0.524	0.013
	15	3.000	3.000	3.001	1.981	1.998	2.016	49.988	50.002	50.015	0.523	0.524	0.524	0.009
	20	3.000	3.000	3.000	1.981	1.995	2.010	49.975	49.986	49.998	0.523	0.524	0.524	0.006
	25	3.000	3.000	3.000	1.997	2.010	2.023	49.996	50.006	50.017	0.524	0.524	0.524	0.005
	50	3.000	3.000	3.000	1.991	2.002	2.012	49.993	49.999	50.005	0.524	0.524	0.524	0.003

TABLE 3.6: Estimateurs des paramètres du modèle original : Cas d'exemple 4

À partir des différents résultats obtenus (numérique et graphique) on constate que :

- La taille de l'intervalle de confiance de la totalité des paramètres réduit en fait et à mesure que la taille de l'échantillon augmente ce qui se traduit par la convergence des estimations considérées en biais et en variance par conséquent en erreur quadratique moyenne.
- La qualité des estimations est inversement proportionnelle à la valeur de la variance des résidus.
- La qualité du régresseur au sens de l'EQM est proportionnelle à la taille de l'échantillon.
- La qualité du régresseur est fortement liée à la variance des résidus. Où l'EQM associé au régresseur augmente en fait et à mesure que la variance des résidus augmente.

## 3.2 Deuxième application

L'objectif de cette deuxième application est de vérifier la possibilité de modéliser une série chronologique gouvernée par une tendance linéaire et une saisonnalité régulière mais qui n'est pas décrite par une fonction sinus.

Ainsi, pour atteindre notre objectif nous avons considéré les quatre modèles suivants :

**Modèle  $M_1$  :**  $y_t = 2 + 30 \cos(0.2t + \frac{\pi}{6})$ .

**Modèle  $M_2$  :**  $y_t = 2 + 10 \sin(0.2t + \frac{\pi}{6}) + 20 \cos(0.2t + \frac{\pi}{4})$ .

**Modèle  $M_3$  :**  $y_t = 3t + 2 + 30 \cos(0.2t + \frac{\pi}{6})$ .

**Modèle  $M_4$  :**  $y_t = 3t + 2 + 10 \sin(0.2t + \frac{\pi}{6}) + 20 \cos(0.2t + \frac{\pi}{4})$ .

**Modèle  $M_5$  :**  $y_t = 3t + 2 + 10 \sin(0.4t + \frac{\pi}{6}) + 20 \cos(0.2t + \frac{\pi}{4})$ .

La figure 3.5 représente l'allure graphique des quatre premiers modèles de la présente application. Tandis que la figure 3.6 représente l'allure graphique du modèle  $M_5$  ainsi

que sa saisonnalité.

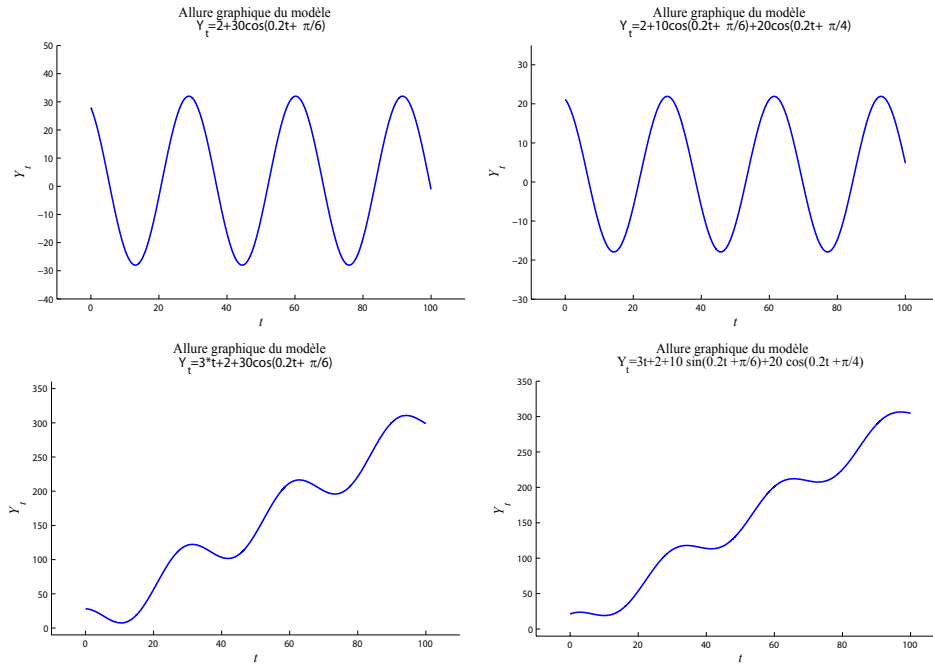


FIGURE 3.5 – Allure graphiques des quatre modèles de la deuxième application.

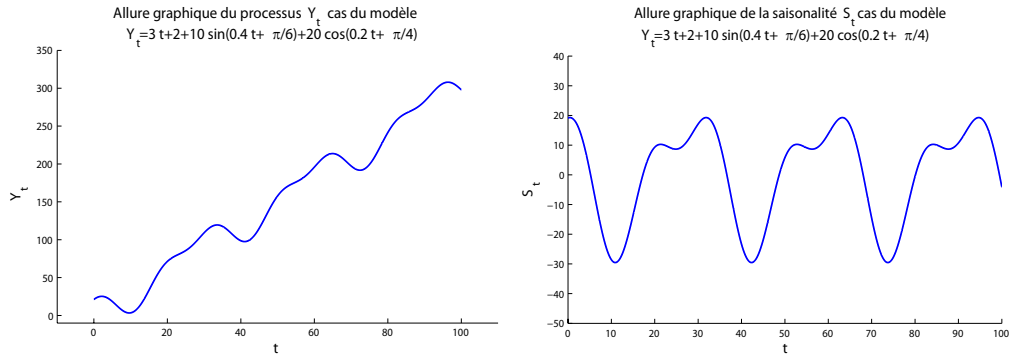


FIGURE 3.6 – Allure graphiques du cinquième modèle de la deuxième application.

Le reste des paramètres sont fixés comme suit :

- La taille d'échantillon  $n \in \{10, 20, 30, 40, 50, 100\}$ ,
- Les valeurs de la variable endogène  $t \in [0 : 100/(n - 1) : 100]$ ,
- L'écart-type des résidus  $\sigma \in \{10, 1, 0.5, 0.4, 0.3\}$ .

Pour les différents calculs nous avons exploité le programme MATLAB de la première application.

### Chapitre 3. Application numérique

Les résultats obtenus, sur 100 échantillons de chacune des situations ci-haut, sont présentés dans les figures 3.8 et 3.7.

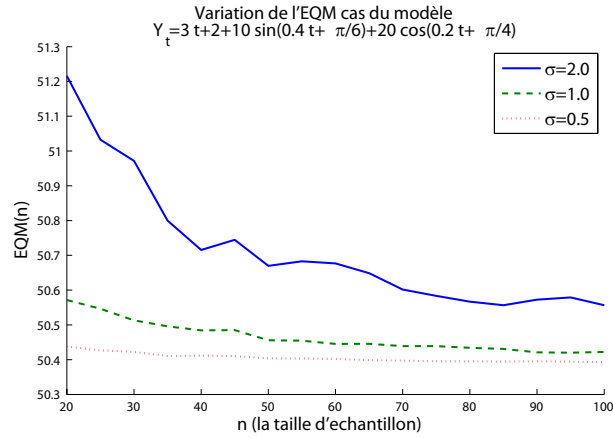


FIGURE 3.7 – Variation de l'EQM en fonction de la taille de l'échantillon et de l'écart-type des résidus : cas du modèle  $M_5$ .

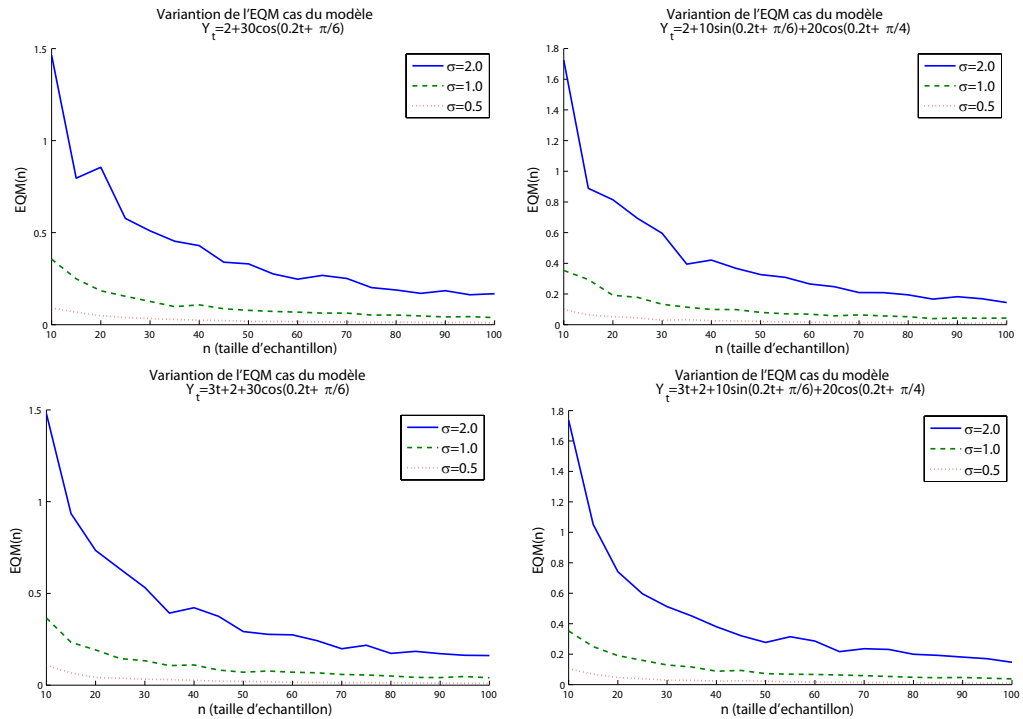


FIGURE 3.8 – Variation de l'EQM en fonction de la taille de l'échantillon et de l'écart-type des résidus : cas des quatre premiers modèles.

A partir des résultats graphiques obtenus dans cette deuxième application on constate que :

- La totalité des constats faits dans la première application sont valables pour les 4 premiers exemples considérés.
- Les 4 premiers exemples peuvent être modélisés fidèlement par le modèle (3.1), et ceci peut être justifié par la décroissance de l'EQM associé aux exemples considérés en fonction de la taille de l'échantillon ce qui se traduit par la convergence du modèle réel considéré vers le modèle (3.1). Cette convergence peut être expliquée par
  - La relation mathématique existante entre la fonction sinus et Cosinus à savoir :

$$\cos(\alpha x + \beta) = \cos(\alpha x + \beta + \pi/2)$$

dans le cas des exemples  $M_1$  et  $M_3$ .

- Le fait que la partie Sinus et la partie Cosinus ont la même période par conséquent on aura une saisonnalité régulière et qui peut être décrite par la fonction Sinus uniquement.
- Les données décrites originalement par un modèle similaire à celui de l'exemple 5, ne peuvent être modélisés par le modèle (3.1), ceci peut être justifié par le comportement de l'EQM en fonction de la taille de l'échantillon (voir figure). Ceci s'explique par le fait que la combinaison linéaire d'une fonction Sinus et la fonction Cosinus lorsqu'elles n'ont pas la même période nous fournit en générale une saisonnalité non régulière. Par conséquent, cette dernière ne peut être présentée par une fonction sinus ou une fonction Cosinus seule.
- L'EQM dans le cas de l'exemple  $M_5$ , tend vers une constante différente de zéro (voir figure 3.7), qui signifie que le comportement des données tend vers un modèle fixe mais bien évidemment différent du modèle (3.1).

### Conclusion et remarques

Dans ce chapitre, nous avons analysé la qualité des estimations des paramètres d'un modèle décrivant une série chronologique gouvernée par une tendance linéaire et une saisonnalité régulière.

Les résultats obtenus soulignent que dans le cas de saisonnalité régulière, la qualité des estimations dépend de la taille de l'échantillon et de la variance des résidus. De plus, quel que soit le modèle originale des données si ce dernier est gouverné par une saisonnalité régulière alors on peut modéliser les données issues de un tel modèle par la fonction Sinus seule c'est-à-dire on peut poser  $S_t = \sin(\alpha t + \beta)$ .

Il est à noter que dans le présent chapitre et le chapitre précédent, nous avons proposé de modéliser la saisonnalité via la fonction sinus, mais ça reste que la même démarche peut être adopté si nous désirons d'écrire cette saisonnalité via la fonction Cosinus de plus la qualité des estimations dans ce cas restera similaire aux résultats obtenus dans nos deux applications.

# Conclusion

Il est à souligné que l'analyse des séries chronologiques et l'estimation de leurs paramètres jouent un rôle crucial dans la compréhension et la prédiction des phénomènes temporels dans divers domaines.

Ce mémoire représente une tentative de comprendre en profondeur les processus temporels modélisés via des fonctions trigonométriques dans l'analyse des processus aléatoires. Plus précisément, les deux objectifs principaux étaient : 1) d'identifier les méthodes efficaces pour estimer les paramètres des séries chronologiques caractérisées par une tendance linéaire et une saisonnalité régulière et 2) comprendre comment les différents facteurs, tels que la taille de l'échantillon et la variance des données, affectent la qualité des estimations.

Lors de la réalisation de ce travail, nous avons constaté que la transformation du modèle mathématique, décrivant une série chronologique caractérisée par une saisonnalité régulière, en un modèle linéaire multiple a permis d'obtenir des estimations précises, même pour des échantillons de petite taille. En outre, l'utilisation de simulation sur différents modèles, différentes tailles d'échantillon et différentes variance des résidus a confirmé l'efficacité de cette approche et a souligné l'importance de la taille de l'échantillon et de la variance des résidus dans la qualité des estimations. Ainsi, ce travail mémoire une contribution significative à la modélisation et à l'analyse des séries chronologiques, ouvrant la voie à de futures études et applications dans divers domaines :

- L'estimation de la fréquence de la saisonnalité.
- Réaliser une étude de simulation extensive sur de modèle plus complexes (tendance non linéaire, cas multivarié, . . .)
- Appliquer la méthodologie exposé dans le présent mémoire sur des données réelles (tels : les températures mensuelles, précipitations, . . .)



# Bibliographie

- [1] Baccini, A., Besse, P., et Dejean, S. (2008). Analyse statistique de données d'expression. Université toulouse.
- [2] Benameur, S. (2021/2022). Notes de Cours : Analyse de données. Cours Master 1 option Mathématiques Appliquées, Département Mathématiques, Université de Biskra.
- [3] Besse, P. (2003). Pratique de la modélisation statistique. Publications du laboratoire de statistique et Probabilité. Université toulouse III-Paul Sabatier.
- [4] Breton, J. C. (2013). Processus stochastique. Université de Rennes 1.
- [5] Chesneau, C. (2017). Modèles de régression. Université de Caen.
- [6] Lagnoux, A. (2015). Renforcement Statistique Séries chronologiques.
- [7] Racotomalala R. Régression linéaire multiple, Université Lumière Lyon 2.
- [8] Saporta, G. (2006). Probabilités, analyse des données et statistique. Université de Toulouse Le Mirail.
- [9] Guyader A.(2012/2013) Régression linéaire. Polycopiés du cours pour Master de Statistique, Université Rennes 2.

## Résumé

Dans le présent mémoire, nous avons considéré une série chronologique gouvernée par une tendance linéaire et une saisonnalité régulière. Nous nous sommes intéressés à l'estimation des paramètres d'une telle série. Nous avons constaté que pour estimer les paramètres en question par la méthode des moindres carrés il est nécessaire de transformer le modèle mathématique décrivant la série en un modèle linéaire multiple. Par ailleurs, l'étude de simulation menée nous permet de conclure d'une part que la qualité des estimations dépend essentiellement de la taille de l'échantillon dont on dispose et de la variance des résidus. D'autre part, l'utilisation du modèle linéaire multiple nous fournit des bonnes estimations au sens de l'erreur quadratique moyenne même pour des échantillons de petites tailles.

**Mots-clés :** Série chronologique ; Tendance linéaire ; Saisonnalité régulière ; Estimation paramétrique ; Méthode des moindres carrés

## Abstract

In this work, we considered a time series governed by a linear trend and regular seasonality. We were interested in estimating the parameters of such a series. We noted that to estimate the parameters in question by the least squares method it is necessary to transform the mathematical model describing the series into a multiple linear model. Furthermore, the simulation study carried out allows us to conclude on the one hand that the quality of the estimates essentially depends on the size of the sample available and the variance of the residuals. On the other hand, the use of the multiple linear models provides us with good estimates in the sense of the mean square error even for small sample sizes.

**Key-words:** Time series; Linear trend; Regular seasonality; Parametric estimation; Least squares method.

## ملخص

في هذا البحث، قمنا بدراسة سلسلة زمنية تخضع لاتجاه خطي وموسمية منتظمة. كان اهتمامنا في تقدير المعلمات لمثل هذه السلسلة. لاحظنا أنه من الضروري تحويل النموذج الرياضي الذي يصف السلسلة إلى نموذج خطي متعدد الأبعاد لتقدير المعلمات المطلوبة بواسطة طريقة أصغر مربعات. بالإضافة إلى ذلك، أظهرت دراسة المحاكاة التي أجريناها أن جودة التقديرات تعود بشكل أساسي على حجم العينة المتاحة وعلى انحراف البيانات المتبقية (الأخطاء المتبقية). ومن ناحية أخرى، فإن استخدام نموذج الخطي متعدد الأبعاد يوفر تقديرات جيدة بمعنى الخطأ المربعي المتوسط حتى لعينات صغيرة الحجم.

## كلمات مفتاحية:

السلسلة الزمنية، الاتجاه الخطي، الموسمية المنتظمة، تقدير المعلمات، طريقة أصغر مربعات.