



**UNIVERSITE KASDI MERBAH
OUARGLA**
Faculté des Mathématiques et des Sciences de la
Matiere

N° d'ordre :
N° de série :

DÉPARTEMENT DE MATHÉMATIQUES

MASTER

Spécialité : Mathématiques

Option : probabilité et statistique

Par : BERGOUG Messaouda

Thème

**Analyse de la série Chronologique
Nombre de Diabétique " Hopitale
Mohamed Boudiaff Ouargla"**

Soutenu publiquement le :31/05/2018

Devant le jury composé de :

Mr.AMARA Abdelkader	Université de KASDI Merbah-Ouargla	President
Melle.SAIDANE Hadda.	Université de KASDI Merbah-Ouargla	Examinateur
Melle. ARBIA Hanane	Université de KASDI Merbah-Ouargla	Rapporteur

Table des matières

1	THÉORIE DE SÉRIE CHRONOLOGIQUE	4
1.1	Analyse des séries temporelles : définition et objectifs	4
1.2	Domaines d'application	5
1.3	Les composantes d'une série temporelle	6
1.4	Opérateurs définis sur une série chronologique	8
1.4.1	Opérateur de retard	8
1.4.2	Opérateur de différence d'ordre d	9
1.4.3	L'opérateur Δ_s	9
1.5	Inversibilité des polynômes P(L)	10
1.6	Méthode de Bays-Ballot	11
1.7	Estimation de la tendance linéaire par la méthode de moindres carrés	13
1.8	Estimation de la tendance par lissage moyenne mobile	14
1.9	Processus bruits blancs	15
1.10	Stationnarité	15
1.10.1	Une série non stationnaire : processus TS et DS	16
1.11	Théorème de Wold	17
1.12	La fonction d'autocovariance et d'autocorrélation	17
1.12.1	La fonction d'autocovariance	17
1.12.2	La fonction d'autocorrélation	18
1.13	La fonction d'autocorrélation partielle	18
2	Les modèles des séries temporelles	20
2.1	Modèle Autorégressifs AR(p)	20
2.1.1	Processus AR(1)	21
2.1.2	Représentation stationnaire	23
2.1.3	Représentation inversible	23
2.1.4	Fonction d'autocorrélation	23
2.1.5	Estimation des paramètres d'un AR(p)	25
2.1.6	Fonction d'autocorrélation partielle PACF(h)	27
2.2	Modèle moyenne mobile MA(q)	28
2.2.1	Stationnarité	29

2.2.2	Inversibilité	29
2.2.3	Fonction d'autocorrélation	30
2.2.4	Modèle $MA(1)$:	31
2.3	Les modèles ARMA(p,q)	32
2.3.1	Auto-covariance et auto-corrélation :	33
2.3.2	Processus ARMA(1,1)	33
2.4	Les modèles ARIMA(p,q)	34
2.5	Méthode de Box-Jenkins	35
2.5.1	Les étapes de la méthodologie	35
2.5.2	La familiarisation avec les données	36
2.5.3	Identification du modèle	38
2.5.4	Estimation des paramètres	38
2.5.5	Validation du modèle	40
2.5.6	La prévision	42
3	Application de la méthode de Box-Jenkins	44
3.1	Analyse exploratoire des données :	45
3.2	Etude de la stationnarité	46

47		
	3.2.2	Test de la racine unitaire 47
	3.2.3	Stationnarisation de serie ldiabet : 48
3.3		Modélisation ARMA par la methode de BOX-JENKINS 51
	3.3.1	Identification des order p et q du modèle ARMA 51
	3.3.2	Estimation de processus AR / MA / ARMA : 52
3.4		Validation du modèle : 53
	3.4.1	Test de Student des paramètres (signiffcativité statistique) : 53
	3.4.2	Test de normalité pour les résidus du modèle MA(1) . . . 53
3.5		Prévision : 54

Chapitre 1

THÉORIE DE SÉRIE CHRONOLOGIQUE

Une série chronologique est un ensemble de valeurs d'une variable indicée dans le temps. L'analyse des séries chronologiques permet d'étudier les situations passées et présentes et peut extrapoler l'évènement dans un futur relativement proche (la prévision). La prévision se fonde donc sur la connaissance du passé et de présent. Dans ce chapitre plusieurs concepts importants liés à l'analyse de séries chronologique seront abordés parmi ceux ci, on retrouve les notions d'autocorrélation, de stationnarité et de bruit blanc. On énoncera également un résultat de *Wold (1948)* concernant les séries chronologiques stationnaires.

1.1 Analyse des séries temporelles : définition et objectifs

Les séries chronologiques, c'est-à-dire les collections de mesures ordonnées dans le temps, constituent une branche de l'économétrie dont l'objet est l'étude des variables au cours du temps. L'étude de telles séries de mesures s'est développée depuis quelques années. En conséquence, elles intéressent beaucoup de gens différents ; différents par la nature des phénomènes qu'ils étudient et par les buts qu'ils se fixent dans leur étude.

Définition 1.1.1 *Soit un espace probabilisé (Ω, A, p) , $T \subset \mathbb{R}$. Un processus stochastique est une famille $X = \{X_t, t \in T\}$ de variables aléatoires définies sur un même (Ω, A, p) . Une série chronologique est une fonction à valeurs réelles $X : T \times \Omega \rightarrow \mathbb{R}$, telle que pour tout $t \in T$ donné, $X_t(\cdot)$ soit une variable aléatoire.*

Définition 1.1.2 Une série temporelle (chronologique) est une réalisation d'un processus stochastique $\{X_t, t \in I\}$ où I est un ensemble dénombrable et totalement ordonné. C'est une suite d'observations d'une grandeur X_t ou d'un vecteur effectuées au cours du temps.

Nous allons considérer dans toute la suite que la période écoulée entre deux observations consécutives est la même, on parle donc de série annuelle, trimestrielle, mensuelle ou quotidienne.

1.2 Domaines d'application

On trouve des exemples de séries chronologiques univariées dans de très nombreux domaines. La liste suivante n'est qu'un échantillon :

1) **Finance et économétrie** : évolution des indices boursiers, des prix, des données économiques des entreprises, des ventes et achats de biens, des productions agricoles ou industrielles.

2) **Assurance** : analyse des sinistres.

3) **Médecine/Biologie** : suivi des évolutions des pathologies, analyse d'électroencéphalogrammes et d'électrocardiogrammes.

4) **Sciences de la Terre et de l'Espace** : indices de marées, variations des phénomènes physiques (Météorologie), évolution des taches solaires, phénomènes d'avalanches.

5) **Traitement du signal** : signaux de communications, de radars, de sonars, analyse de la parole.

6) **Traitement des données** : mesures successives de position ou de direction d'un objet mobile (trajectographie).

7) **Métrologie** : variation de phase ou de fréquence des oscillateurs.

8) **Démographie** : évolution de la population.

1.3 Les composantes d'une série temporelle

Les composantes d'une série temporelles sont :

1) **La tendance** ($f_i, 1 \leq i \leq n$) : représente l'évolution à long terme de la grandeur étudiée, et traduit l'aspect général de la série. C'est une fonction monotone, souvent polynomiale. La tendance prend plusieurs forme :

a- *Tendance linéaire* :

$$T_t = a + bt.$$

b- *Tendance quadratique* :

$$T_t = a + bt + ct^2.$$

c- *Tendance polynomiale d'ordre q* :

$$T_t = b_0 + b_1t + b_2t^2 + \dots + b_qt^q.$$

d- *Tendance logistique* :

$$T_t = \frac{c}{1 + be^{-at}} \text{ ou } a, b, c \in \mathbb{R}.$$

2) **Les variations saisonnières** ($s_i, 1 \leq i \leq n$) : sont liées au rythme imposé par les saisons météorologiques (production agricole, consommation de gaz,...), ou encore par des activités économiques et sociales (fêtes, vacances, solde, etc). Mathématiquement, ce sont des fonctions périodiques, c'est-à-dire qu'il existe un entier p , appelé période, tel que

$$s_i = s_{i+p}$$

pour tout $i \geq 1$. Au premier abord, cette composante est entièrement déterminée par p ses premières valeurs $s_1, s_2, s_3, \dots, s_p$.

3) **Cycles** ($c_i, 1 \leq i \leq n$) : regroupent des variations à période moins précise autour de la tendance.

4) **Les fluctuations irrégulières/résidus/bruit** ($e_i, 1 \leq i \leq n$) : sont des variations de faible intensité et de courte durée, et de nature aléatoire. En effet, elles ne sont pas clairement apercevables dans les graphiques, à cause de leur faible intensité par rapport aux autres composantes.

En résumé, nous considérerons une série chronologique comme issue de la composition de 3 composantes :

1. $(f_i, 1 \leq i \leq n)$ la tendance (intégrant éventuellement un cycle),
2. $(s_i, 1 \leq i \leq n)$ les coefficients saisonniers,
3. $(e_i, 1 \leq i \leq n)$ les fluctuations irrégulières (intégrant éventuellement des accidents).

La décomposition d'une série chronologique possédant un mouvement saisonnier peut s'effectuer selon trois types de modèles :

• **Modèle additif**

$$X_t = f_t + s_t + e_t \quad t = 1, \dots, T$$

• **Modèle multiplicatif**

$$X_t = f_t \cdot s_t \cdot e_t \quad t = 1, \dots, T$$

• **Modèle mixte**

$$X_t = f_t \cdot (1 + s_t) + e_t \quad t = 1, \dots, T$$

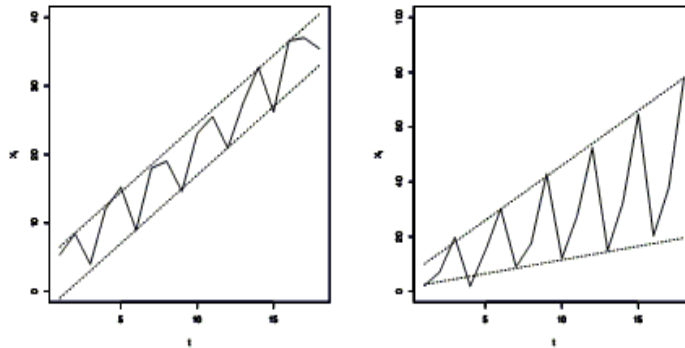


FIG. 1.1 – Exemple d'une modèle additif (gauche) et multiplicatif (droite).

1.4 Opérateurs définis sur une série chronologique

On donne dans cette partie les opérateurs de retard, de différence et Δ_s .

1.4.1 Opérateur de retard

Définition 1.4.1 On appelle opérateur retard l'opérateur L qui à tout processus $(X_t), t \in \mathbb{Z}$ associe le processus $(Y_t), t \in \mathbb{Z}$ défini par

$$\forall t \in \mathbb{Z}, Y_t = LX_t = X_{t-1}$$

Remarque 1.4.1 L'opérateur L est linéaire et inversible. Son inverse $L^{-1} = F$ est défini par

$$\forall t \in \mathbb{Z}, FX_t = X_{t+1}$$

L'opérateur F est appelé opérateur avance.

(i) Il est possible de composer les opérateurs : $L^2 = L \circ L$, et plus généralement

$$L^p = \underbrace{L \circ L \circ L \dots \circ L}_{p \text{ fois}}$$

où $p \in \mathbb{N}$ avec la convention $L^0 = I$, On notera que $L^p(X_t) = X_{t-p}$.

(ii) Soit A le polynôme, $A(z) = a_0 + a_1z + a_2z^2 + \dots + a_pz^p$, On notera $A(L)$ l'opérateur

$$A(L) = a_0I + a_1L + a_2L^2 + \dots + a_pL^p = \sum_{k=0}^p a_kL^k$$

Lemme 1.4.1 Soit une série chronologique $(X_t)_{t \in \mathbb{Z}}$, L'opérateur retard possède les propriétés suivantes :

1. $L^j X_t = X_{t-j}$, en particulier $L^0 X_t = X_t$.
2. Si $X_t = c \in \mathbb{R}$ pour tout $t \in \mathbb{Z}$, alors $L^j X_t = L^j c = c$ pour tout $c \in \mathbb{Z}$.
3. $L^j(L^k X_t) = L^{j+k} X_t = X_{t-j-k}$.
4. $(L^j + L^k)X_t = L^j X_t + L^k X_t = X_{t-j} + X_{t-k}$.

5. Si $|a| < 1$, alors

$$(1 - aL)^{-1}X_t = \sum_{j=0}^{\infty} a^j X_{t-j}.$$

A l'aide de la notation fournie par l'opérateur retard, la représentation d'un processus stationnaire d'ordre deux $(X_t)_{t \in \mathbb{Z}}$ énoncée dans le théorème de *Wold* (1948) peut s'écrire sous la forme

$$X_t = \Psi(L)\varepsilon_t + k_t,$$

où $\Psi(L)$ est un polynôme de degré infini défini par

$$\Psi(L) = \sum_{j=0}^{\infty} \Psi_j L^j.$$

1.4.2 Opérateur de différence d'ordre d

Définition 1.4.2 On définit l'opérateur Δ_d de différence d'ordre d comme l'opérateur linéaire tel que

$$\Delta_d X_t = X_t - X_{t-d} = (1 - L)^d X_t.$$

On peut aussi prendre l'opérateur d'ordre un et l'appliquer plusieurs fois :

$$\Delta^2(X_t) = \Delta((\Delta X_t)) = \Delta(X_t - X_{t-1}) = (1 - L)(X_t - X_{t-1}) = X_t - 2X_{t-1} + X_{t-2}.$$

Ces opérateurs peuvent être utilisés afin de transformer un processus de moyenne non nulle en un processus de moyenne nulle. On peut aussi s'en servir pour enlever la composante saisonnière de la série. Dans la prochaine section, on les utilise pour mieux représenter les modèles.

1.4.3 L'opérateur Δ_s

L'opérateur Δ_s permet d'éliminer la saisonnalité de périodes et on peut également l'appliquer plusieurs fois.

$$\Delta_s^2 X_t = \Delta_s(X_t - X_{t-s}).$$

Le nombre de fois où on applique est appelé ordre de désaisonnalisation.

1.5 Inversibilité des polynômes $P(L)$

Soit $A(\cdot)$ un polynôme, on cherche $B(\cdot)$ tel que

$$A(\cdot) \circ B(\cdot) = 1$$

(a) Si $|\lambda| \leq 1$ alors $(1 - |\lambda|)$ est inversible, et de plus

$$(1 - |\lambda|)^{-1} = \sum_{k=1}^{\infty} \lambda^k L^k.$$

(b) Si $|\lambda| \geq 1$ alors $(1 - |\lambda|)$ est inversible, et de plus

$$(1 - |\lambda|)^{-1} = \sum_{k=1}^{\infty} \frac{1}{\lambda^k} F^k.$$

(c) Si $|\lambda| = 1$, alors $1 - |\lambda|$ n'est pas inversible.

(d) Si $|\lambda| \leq 1$, alors

$$(1 - |\lambda|)^{-1} = \sum_{k=1}^{\infty} \lambda^k = \frac{1}{1 - \lambda} \leq +\infty.$$

donc

$$A(L) = \sum_{k=1}^{\infty} \lambda^k L^k$$

est bien défini, de plus :

$$\begin{aligned} (1 - \lambda L)A(L) &= \lim_{k \rightarrow +\infty} (1 - \lambda L) \sum_{j=0}^k \lambda^j L^j \\ &= \lim_{k \rightarrow +\infty} 1 + \lambda^{k+1} L^{k+1} = 1 \end{aligned}$$

Ce qui signifie que A est le polynôme inverse associé à $(1 - \lambda L)$.

1.6 Méthode de Bays-Ballot

Considérons le modèle

$$\begin{aligned} X_t &= \beta_1 + \beta_2 t + S_t^1 \gamma_1 + S_t^2 \gamma_2 + S_t^3 \gamma_3 + \dots + S_t^m \gamma_m + \varepsilon_t \\ \mathbf{S}_t &= S_t^1 \gamma_1 + S_t^2 \gamma_2 + S_t^3 \gamma_3 + \dots + S_t^m \gamma_m \\ \mathbf{Z}_t &= \beta_1 + \beta_2 t \end{aligned}$$

En notant N le nombre d'années entières et m la périodicité, on pose :

\bar{x}_n : moyenne des X_t relatives à l'année n .

\bar{x}_j : moyenne des X_t relatives à trimestre j .

\bar{x} : moyenne des X_t relatives à les observations X_t .

Le principe de la méthode de Bays-Ballot est de résoudre l'équation

$$\min_{\beta, \gamma} \sum_{t=1}^T \left[X_t - \beta_1 + \beta_2 t - \sum_{j=1}^m S_t^j \gamma_j \right]^2 ,$$

sous contrainte $\gamma_1 + \gamma_2 + \gamma_3 + \dots + \gamma_m = 0$.

L'équation admet alors pour solution

$$\hat{\beta}_2 = \frac{12 \sum_{n=1}^N n \bar{x}_n - \frac{N(N+1)}{2} \bar{x}}{m N(N^2 - 1)}$$

$$\hat{\beta}_1 = \bar{x} - \hat{\beta}_2 \frac{Nm + 1}{2}$$

$$\hat{\gamma}_j = \bar{x}_j - \bar{x} - \hat{\beta}_2 \left[j - \frac{m+1}{2} \right] .$$

donc

$$\hat{X}_t = \hat{\beta}_1 + \hat{\beta}_2 t + \hat{S}_t^1 \gamma_1 + \hat{S}_t^2 \gamma_2 + \hat{S}_t^3 \gamma_3 + \dots + \hat{S}_t^m \gamma_m$$

pour effectuer la prévision à l'instant $t = j + m(j - 1)$, on remplace dans le méthode théorique les paramètres par leur estimation.

Année		T_1	T_2	T_3	T_4	\bar{X}_k	$k\bar{X}_k$
2015	1	25	14	09	22	17.5	17.5
2016	2	36	18	16	31	25.25	50.5
2017	3	42	16	07	32	24.25	72.5
	\bar{X}_j	34.33	16	10.66	28.33	40	140.75
	$\bar{\gamma}_j$	13.26	-5.91	-12.09	4.73		

TAB. 1.1 – Exemple sur des données trimestriel.

N : le nombre des années 3.

m : le nombre de mois (saison) 4.

D'après la méthode de Bys-Ballot :

$$\hat{\beta}_1 = 0.84$$

$$\hat{\beta}_2 = 16.82$$

$$\hat{X}_t = \hat{\beta}_1 + \hat{\beta}_2 t + \hat{S}_t^1 \gamma_1 + \hat{S}_t^2 \gamma_2 + \hat{S}_t^3 \gamma_3 + \hat{S}_t^4 \gamma_4$$

$$\hat{X}_t = 16.82 + 0.84t + 13.26S_t^1 - 5.9S_t^2 - 12.09S_t^3 + 4.73S_t^4$$

donc

$$\hat{X}_t = 16.82 + 0.84(j + m(j - 1)) + 13.26S_t^1 - 5.9S_t^2 - 12.09S_t^3 + 4.73S_t^4$$

1.7 Estimation de la tendance linéaire par la méthode de moindres carrés

Si le phénomène augmente alors l'équation de la tendance est :

$$\hat{Z} = \hat{a} - \hat{b}x$$

La méthode de moindres carrés permet d'estimer \hat{a}, \hat{b} :

$$\hat{b} = \frac{\sum x_i z_i - n \bar{x} \bar{z}}{\sum x_i^2 - n \bar{x}^2}$$

$$\hat{a} = \bar{z} - \hat{b} \bar{x}$$

t	z	t_iz_i	t_i²
1	499.5	499.5	1
2	593.9	1187.8	4
3	659.4	1978.2	9
4	790.8	3163.2	16
5	895.8	4479	25
$\bar{t} = 3$	$\bar{z} = 687.88$	$\sum t_i z_i$	$\sum t_i^2 = 55$

TAB. 1.2 – Exemple de l'estimation de la tendance.

D'après la méthode M.C

$$n = 5$$

$$\hat{b} = \frac{11307.7 - 5(3 * 687.88)}{55 - 5 * 9}$$

$$= 98.95$$

$$\hat{a} = 687.88 - 98.95 * 3$$

$$= 391.03$$

donc

$$\hat{Z} = 391.03 + 98.95x$$

pour $t = 7$

$$\begin{aligned}\hat{Z} &= 391.09 + 998.95 * 7 \\ &= 1083.74\end{aligned}$$

1.8 Estimation de la tendance par lissage moyenne mobile

la méthode moyenne mobile est un outil intéressant pour estimer la tendance.

1^{ère} cas : $k = 2p + 1$ "impair"

$$M_k(X_t) = \frac{1}{k}(X_{t-p} + \dots + X_{t-1} + X_t + X_{t+1} + \dots + X_{t+p})$$

2^{ème} cas : $k = 2p$ "pair"

$$M_k(X_t) = \frac{1}{k}\left(\frac{1}{2}X_{t-p} + \dots + X_{t-1} + X_t + X_{t+1} + \dots + \frac{1}{2}X_{t+p}\right)$$

Le principe de cette méthode est représentée dans le tableau suivant :

t	y	m.m d'ordre 3	m.m d'ordre 4
1	y_1		
2	y_2	$z_2 = \frac{y_1+y_2+y_3}{3}$	
3	y_3	$z_3 = \frac{y_2+y_3+y_4}{3}$	$z_3 = \frac{\frac{1}{2}y_1+y_2+y_3+y_4+\frac{1}{2}y_5}{3}$
4	y_4	$z_4 = \frac{y_3+y_4+y_5}{3}$	$z_4 = \frac{\frac{1}{2}y_2+y_3+y_4+y_5+\frac{1}{2}y_6}{3}$
5	y_5	$z_5 = \frac{y_4+y_5+y_6}{3}$	
6	y_6		

TAB. 1.3 – Lissage par moyenne mobile

1.9 Processus bruits blancs

Un processus $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc si

$$\begin{cases} E(\varepsilon_t) = 0, \text{ pour tout } t \\ \text{var}(\varepsilon_t) = E(\varepsilon_t^2) = \sigma^2, \text{ pour tout } t \\ \text{cov}(\varepsilon_t, \varepsilon_{t-h}) = E(\varepsilon_t \varepsilon_{t-h}) = 0, \text{ pour tout } t \text{ et } h \neq 0 \end{cases}$$

On parle de bruit blanc gaussien lorsque $\varepsilon_1, \varepsilon_2, \dots$ sont *i.i.d.* $N(0, \sigma_\varepsilon^2)$.

1.10 Stationnarité

Une propriété importante des séries chronologiques est la stationnarité. La stationnarité est une caractéristique d'une série chronologique qui implique que le comportement de la série ne dépend pas du temps. En particulier, on dit qu'une série $(X_t)_{t \in \mathbb{Z}}$ est stable si elle ne comporte pas de tendance saisonnière, ni de tendance à la hausse ou à la baisse. Plus formellement, on distingue deux types de stationnarité, à savoir forte et faible.

Définition 1.10.1 Une suite $\{X_t : t \geq 0\}$ de variables aléatoires est dite stationnaire du second ordre si elle vérifie les propriétés suivantes :

1. Pour tout t , $E(X_t^2) < +\infty$.
2. Pour tout t , $E(X_t) = \mu$.
3. Pour tout t et pour tout h , $\gamma(h) = \text{cov}(X_t, X_{t+h})$, indépendante de t .

Définition 1.10.2 Une série $(X_t)_{t \in \mathbb{Z}}$ est dite fortement stationnaire (ou stationnaire d'ordre 1) si pour tout $l \in \mathbb{Z}$

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{L}{=} (X_{t_1+h}, \dots, X_{t_n+h}).$$

La notion de stationnarité d'ordre 1 est généralement beaucoup trop forte pour les applications. Dans la pratique, il est habituellement suffisant de considérer des modèles faiblement stationnaires.

1.10.1 Un série non stationnaire : processus TS et DS

On présente deux classes des processus non stationnaires les processus *TS* et *DS*.

a/ Processus TS :

Le processus *TS* (Trend Stationary) s'écrit :

$$y_t = \alpha + \beta t + \varepsilon_t$$

ou ε_t représente l'erreur du modèle à la date t .

Il présente une non stationnarité de nature déterministe. Le processus *TS* est non stationnaire car $E(y_t) = \alpha + \beta t$ dépend du temps t .

b/Processus DS :

Le processus *DS* (Differncy Stationnary) avec dérive ($\beta \neq 0$) s'exprime comme suit :

$$y_t = y_{t-1} + \beta + \varepsilon_t.$$

Par récurrence, on obtient (dans le cas avec dérive) :

$$y_1 = y_0 + \beta + \varepsilon_1$$

$$y_2 = y_1 + \beta + \varepsilon_2$$

$$= y_0 + \beta + \varepsilon_1 + \beta + \varepsilon_2 =$$

$$= y_0 + 2\beta + \varepsilon_1 + \varepsilon_2$$

.

.

.

$$y_t = y_0 + \beta t + \sum_{i=1}^t \varepsilon_i \text{ ou } \varepsilon_i \sim iid(0, \sigma_\varepsilon^2),$$

1.11 Théorème de Wold

Le théorème de *Wold (1948)* est fondamental pour l'analyse de séries temporelles stationnaires. Ce résultat se veut, d'une certaine façon, un préalable aux processus ARMA qui seront abordés au chapitre 2.

Théorème 1.11.1 *Tout processus $(X_t)_{t \in \mathbb{Z}}$ faiblement stationnaire peut s'écrire sous la forme*

$$X_t = \sum_{j=0}^{+\infty} \Psi_j \varepsilon_{t-j} + k_t,$$

où les paramètres ψ_1, ψ_2, \dots sont des réels tels que $\psi_0 = 1$, et $\sum_{j=0}^{+\infty} \Psi_j^2 \leq \infty$. $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien et k_t est une composante linéaire telle que $\text{cov}(k_t, \varepsilon_{t-j}) = 0$ pour tout $j \in \mathbb{Z}$.

Selon le théorème de *Wold (1948)*, tout processus stationnaire d'ordre 2 peut être représenté comme une somme pondérée infinie de chocs passés, caractérisés par un bruit blanc.

1.12 La fonction d'autocovariance et d'autocorrélation

1.12.1 La fonction d'autocovariance

La fonction d'autocovariance $\gamma_X(h)_{h \in \mathbb{Z}}$ mesure la covarianance entre une variable et cette même variable à des dates différentes pour un délai h :

$$\begin{aligned} \gamma_X(h) &= \text{cov}(X_t, X_{t-h}) \\ &= E[(X_t - E(X_t))(X_{t-h} - E(X_{t-h}))] \end{aligned}$$

ainsi

$$\begin{aligned} \gamma_X(0) &= \text{var}(X_t) \\ &= E[(X_t - E(X_t))^2] \end{aligned}$$

Elle donne une information sur la variabilité de la série et sur les liaisons temporelles qui existent entre les diverses composantes de la série X_t .

La fonction d'autocovariance d'une processus stationnaire est une fonction :

a) **paire** : $\gamma_X(h) = \gamma_X(-h) \cdot \forall h$.

b) **semi-définie positive** :

$$\sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma_X(t_j - t_k) \geq 0, \forall n \in \mathbb{N}, \forall a_j \in \mathbb{R}, \forall t_j \in \mathbb{Z}.$$

1.12.2 La fonction d'autocorrélation

La fonction d'autocorrélation $\rho_X(h)$ est une fonction mesure la corrélation de la série avec elle-même décalée de h périodes, elle est définie par :

$$\begin{aligned} \rho_X(h) &= \frac{\text{cov}(X_t, X_{t-h})}{\sqrt{\text{var}(X_t)} \sqrt{\text{var}(X_{t-h})}} \\ &= \frac{\gamma_X(h)}{\gamma_X(0)}. \end{aligned}$$

$\rho_X(\cdot)$ est à valeurs dans $[-1, 1]$, et $\rho_X(0) = 1$.

1.13 La fonction d'autocorrélation partielle

Elle mesure la liaison linéaire entre X_t , X_{t-h} par les variables intermédiaires $X_{t-1}, \dots, X_{t-h+1}$.

Le coefficient d'autocorrélation partielle d'ordre h , noté $\pi(h)$ est définie par

$$\pi(h) = \text{cov}(X_t, X_{t-h} / X_{t-1}, \dots, X_{t-h+1})$$

Le coefficient d'autocorrélation partielle d'ordre h d'un processus stationnaire est alors se calcule de la manière suivante :

$$\pi(h) = \frac{|R(h)^*|}{|R(h)|}$$

avec

$$R(h) = \begin{pmatrix} 1 & \rho_1 & \cdot & \cdot & \rho_{h-1} \\ \rho_1 & 1 & \cdot & \cdot & \rho_{h-2} \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{h-1} & \rho_{h-2} & \cdot & \cdot & 1 \end{pmatrix}$$

$R(h)^*$ est la matrice $R(h)$ dans la quelle on a remplacé la colonne h par $\begin{pmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \rho_2 \end{pmatrix}$,

$$R(h)^* = \begin{pmatrix} 1 & \rho_1 & \cdot & \cdot & \rho_{h-1} \\ \rho_1 & 1 & \cdot & \cdot & \rho_{h-2} \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{h-1} & \rho_{h-2} & \cdot & \cdot & 1 \end{pmatrix}$$

Ainsi

$$\pi(1) = \rho(1), \quad \pi(2) = \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2}, \dots$$

Chapitre 2

Les modèles des séries temporelles

Les modèles **ARIMA** permettent de représenter un grand nombre de processus aléatoires stationnaires. On s'intéresse à la mise en oeuvre d'une modélisation linéaire de série chronologique par la méthodologie de Box et Jenkins. Dans ce chapitre, nous effectuons quelques rappels sur les processus de trois type : les modèles autorégressifs (*AR*), les modèles moyenne mobile (*MA*) et les modèles intégrés (*I*).

2.1 Modèle Autorégressifs AR(*p*)

Dans l'étude d'une série chronologique, il est naturel de penser que la valeur de la série à la date t peut dépendre des valeurs prises aux dates précédentes :

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-p}).$$

Il n'est généralement pas nécessaire de prendre en compte tout le passé de la série et on peut le plus souvent se limiter à p valeurs :

$$X_t = f(X_{t-1}, X_{t-2}, \dots, X_{t-p}) + \varepsilon_t$$

où $\{\varepsilon_t\}$ est un bruit blanc.

Définition 2.1.1 On appelle processus Autorégressif (X_t) vérifiant une relation du type :

$$\begin{aligned} X_t &= \mu + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \alpha_3 X_{t-3} + \dots + \alpha_p X_{t-p} + \varepsilon_t \\ &\Leftrightarrow X_t - \alpha_1 X_{t-1} - \alpha_2 X_{t-2} - \alpha_3 X_{t-3} - \dots - \alpha_p X_{t-p} = \varepsilon_t + \mu \\ &\Leftrightarrow (1 - \alpha_1 L - \dots - \alpha_p L^p) X_t = \varepsilon_t + \mu \\ &\Leftrightarrow \Phi(L) X_t = \mu + \varepsilon_t, \quad \Phi(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p, \forall t \in \mathbb{Z} \end{aligned}$$

où les α_i sont des réels et (ε_t) est bruit blanc de variance σ^2 , avec $p \in \mathbb{N}^*$, $\alpha_p \neq 0$.

2.1.1 Processus AR(1)

On dit que la série X_t suit un processus Autorégressif d'ordre 1 (AR(1)) si on peut écrire :

$$\begin{aligned} X_t &= \mu + \alpha_1 X_{t-1} + \varepsilon_t \\ &\Leftrightarrow X_t - \alpha_1 X_{t-1} = \mu + \varepsilon_t \\ &\Leftrightarrow (1 - \alpha_1 L) X_t = \mu + \varepsilon_t \end{aligned}$$

où la série ε_t est un bruit blanc.

On peut remarquer qu'on fait une regression de la série décalée de 1 sur la série elle même et les résidus forment un bruit blanc.

Fonction transfert :

On peut écrire

$$\begin{aligned} X_t &= \mu + \alpha_1 X_{t-1} + \varepsilon_t \\ X_t &= \mu + \alpha_1^2 X_{t-2} + \alpha_1 \varepsilon_{t-1} + \varepsilon_t \\ X_t &= \mu + \alpha_1^k X_{t-k} + \dots + \alpha_1^2 \varepsilon_{t-2} + \alpha_1 \varepsilon_{t-1} + \varepsilon_t \\ X_t &= \mu + \sum_{i=0}^{+\infty} \alpha_1^i \varepsilon_{t-i} \\ X_t &= \mu + \sum_{i=0}^{+\infty} \alpha_1^i L^i \varepsilon_t \\ X_t &= \mu + \Psi(L) \varepsilon_t. \end{aligned}$$

La fonction de transfert a donc une infinité de termes. Si on revient à notre écriture formelle sous forme de polynôme, on remarque qu'on a calculé l'inverse du polynôme $(1 - \alpha_1 L)$ et que :

$$(1 - \alpha_1 L)^{-1} = \sum_{i=0}^{+\infty} \alpha_1^i L^i$$

Proposition 2.1.1 *Si $X_t \sim AR(p)$ tel que $\Phi(L)X_t = \mu + \varepsilon_t$ alors*

$$\begin{aligned} E(X_t) &= \frac{\mu}{\Phi(1)} \\ &= \frac{\mu}{1 - (\alpha_1 + \alpha_2 + \dots + \alpha_p)} \end{aligned}$$

Démonstration

$$\begin{aligned} X_t &= \mu + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \alpha_3 X_{t-3} + \dots + \alpha_p X_{t-p} + \varepsilon_t \\ E(X_t) &= \mu + \alpha_1 E(X_{t-1}) + \alpha_2 E(X_{t-2}) + \alpha_3 E(X_{t-3}) + \dots + \alpha_p E(X_{t-p}) + E(\varepsilon_t) \\ m &= \mu + \alpha_1 m + \alpha_2 m + \alpha_3 m + \dots + \alpha_p m \\ m &= \frac{\mu}{1 - (\alpha_1 + \alpha_2 + \dots + \alpha_p)} \\ &= \frac{\mu}{\Phi(1)} \end{aligned}$$

Proposition 2.1.2 *Si $X_t \sim AR(p)$ tel que $\Phi(L)X_t = \mu + \varepsilon_t$ et si l'on pose $Y_t = X_t - m$ où $m = E(X_t)$, on a alors*

$$\Phi(L)Y_t = \varepsilon_t \text{ et } E(Y_t) = 0$$

Démonstration

$$\begin{aligned} E(X_t) &= m \\ &= \frac{\mu}{\Phi(L)} \\ \text{et } \Phi(L)(X_t - m) &= \Phi(L)X_t - \Phi(L)m \end{aligned}$$

Or $Lm = m$ et par conséquent :

$$\begin{aligned}\Phi(L)m &= (1 - (\alpha_1 L + \alpha_2 L^2 + \dots + \alpha_p L^p))m \\ &= \Phi(1)m\end{aligned}$$

Finalement

$$\begin{aligned}\Phi(L)(X_t - m) &= \Phi(L)X_t - \mu \\ &= \varepsilon_t\end{aligned}$$

2.1.2 Représentation stationnaire

Ce processus est pour l'instant défini sous forme implicite et en particulier il n'est pas certain que cette dernière équation admette toujours une solution stationnaire.

Si le polynôme Φ a toutes ses racines de module différent de 1, on peut inverser l'opérateur $\Phi(L)$. On en déduit que l'équation admet une solution unique, avec une écriture $MA(\infty)$:

$$Y_t = \Phi(L)^{-1}\varepsilon_t = \sum_{i=-\infty}^{+\infty} h_i \varepsilon_{t-i}$$

On peut alors montrer que l'on a $\sum_{i=-\infty}^{+\infty} |h_i| < \infty$ et donc que la représentation est stationnaire.

2.1.3 Représentation inversible

La représentation $AR(p)$ est inversible par définition.

2.1.4 Fonction d'autocorrélation

-L'auto-covariance

$$\gamma(h) = \text{cov}(X_t, X_{t-h}) = E(X_t X_{t-h}) \text{ pour } h \geq 0$$

$$\begin{aligned}
X_t &= \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \alpha_3 X_{t-3} + \dots + \alpha_p X_{t-p} + \varepsilon_t \\
X_t X_{t-h} &= \alpha_1 X_{t-1} X_{t-h} + \alpha_2 X_{t-2} X_{t-h} + \alpha_3 X_{t-3} X_{t-h} + \dots + \alpha_p X_{t-p} X_{t-h} + \varepsilon_t X_{t-h} \\
E(X_t X_{t-h}) &= \alpha_1 E(X_{t-1} X_{t-h}) + \alpha_2 E(X_{t-2} X_{t-h}) + \dots + \alpha_p E(X_{t-p} X_{t-h}) + E(\varepsilon_t X_{t-h}) \\
\gamma(h) &= \alpha_1 \gamma(h) + \alpha_2 \gamma(h-1) + \alpha_3 \gamma(h-2) + \dots + \alpha_p \gamma(h-p) + \underbrace{E(\varepsilon_t X_{t-h})}_{=0}, h \geq 0
\end{aligned}$$

Donc on obtient

$$\gamma(h) = \sum_{j=1}^p \alpha_j \gamma(h-j), \quad h \geq 0$$

où

$$\begin{aligned}
\sigma_X^2 &= \gamma(0) \\
&= \text{cov}(X_t, X_t) \\
&= E(X_t^2)
\end{aligned}$$

donc

$$\begin{aligned}
X_t^2 &= \alpha_1 X_{t-1} X_t + \alpha_2 X_{t-2} X_t + \alpha_3 X_{t-3} X_t + \dots + \alpha_p X_{t-p} X_t + \varepsilon_t X_t \\
E(X_t^2) &= \alpha_1 E(X_{t-1} X_t) + \alpha_2 E(X_{t-2} X_t) + \alpha_3 E(X_{t-3} X_t) + \dots + \alpha_p E(X_{t-p} X_t) + E(\varepsilon_t X_t) \\
&= \alpha_1 \gamma(1) + \alpha_2 \gamma(2) + \alpha_3 \gamma(3) + \dots + \alpha_p \gamma(p) + E(\varepsilon_t X_t)
\end{aligned}$$

Or

$$\begin{aligned}
E(\varepsilon_t X_t) &= E((\alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \alpha_3 X_{t-3} + \dots + \alpha_p X_{t-p} + \varepsilon_t) \varepsilon_t) \\
&= \underbrace{E((\alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \alpha_3 X_{t-3} + \dots + \alpha_p X_{t-p}) \varepsilon_t)}_{=0} + E(\varepsilon_t^2)
\end{aligned}$$

$$E(\varepsilon_t \varepsilon_t) = E(\varepsilon_t^2) = \sigma_\varepsilon^2$$

Alors

$$\gamma(0) = \alpha_1 \gamma(1) + \alpha_2 \gamma(2) + \alpha_3 \gamma(3) + \dots + \alpha_p \gamma(p) + \sigma_\varepsilon^2$$

-L'auto-corrélation ACF(h) :

A partir de la relation de récurrence de $\gamma(h)$ on déduit que

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \quad h \geq 0.$$

Donc

$$\begin{aligned} \rho(h) &= \alpha_1 \frac{\gamma(h-1)}{\gamma(0)} + \alpha_2 \frac{\gamma(h-2)}{\gamma(0)} + \alpha_3 \frac{\gamma(h-3)}{\gamma(0)} + \dots + \alpha_p \frac{\gamma(h-p)}{\gamma(0)} \\ &= \alpha_1 \rho(h-1) + \alpha_2 \rho(h-2) + \alpha_3 \rho(h-3) + \dots + \alpha_p \rho(h-p) \\ &= \sum_{j=1}^p \alpha_j \rho(h-j), \quad h \geq 0 \end{aligned}$$

Déduire de $\gamma(0)$:

$$\gamma(0) = \frac{\sigma_\varepsilon^2}{1 - (\alpha_1 \rho(1) + \alpha_2 \rho(2) + \alpha_3 \rho(3) + \dots + \alpha_p \rho(p))}$$

2.1.5 Estimation des paramètres d'un AR(p)

Définition 2.1.2 On désigne par \mathfrak{R}_p la matrice d'autocorrélation :

$$\mathfrak{R}_p = \begin{pmatrix} 1 & \rho(1) & \cdot & \cdot & \rho(h-1) \\ \rho(1) & 1 & \cdot & \cdot & \rho(h-2) \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho(h-1) & \rho(h-2) & \cdot & \cdot & 1 \end{pmatrix}$$

Les équations de Yule-Walker pour $h = 1, \dots, p$ peuvent s'écrire :

$$\begin{pmatrix} 1 & \rho(1) & \cdot & \cdot & \rho(h-1) \\ \rho(1) & 1 & \cdot & \cdot & \rho(h-2) \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho(h-1) & \rho(h-2) & \cdot & \cdot & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \alpha_p \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \rho(2) \\ \cdot \\ \cdot \\ \rho(p) \end{pmatrix}$$

Les solutions de l'équation de récurrence sont complètement déterminées par la donnée de conditions initiales $\rho(1), \dots, \rho(p)$: elles permettent d'obtenir $\alpha_1, \alpha_2, \dots, \alpha_p$.

En particulier : elles donneront une estimation préliminaire de $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ en fonction de $\hat{\rho}(1), \dots, \hat{\rho}(p)$:

$$\left\{ \begin{array}{l} \rho(1) = \alpha_1 + \alpha_2\rho(1) + \dots + \alpha_p\rho(p-1) \\ \vdots \\ \rho(p) = \alpha_1\rho(p-1) + \dots + \alpha_{p-1}\rho(1) + \alpha_p \end{array} \right.$$

Alors

$$\left\{ \begin{array}{l} \alpha_1 = (1 - \alpha_1)\rho(1) - \dots - \alpha_p\rho(p-1) \\ \vdots \\ \alpha_p = \rho(p) - \alpha_1\rho(p-1) + \dots + \alpha_{p-1}\rho(1) \end{array} \right.$$

On peut donc aussi obtenir $\rho(1), \dots, \rho(p)$ en fonction de $\alpha_1, \dots, \alpha_p$.

Le système de Yule Waller prend la forme matricielle :

$$\mathfrak{R}_p^{-1}\hat{\alpha} = \hat{\rho}.$$

$$\hat{\rho} = (\hat{\rho}(1), \dots, \hat{\rho}(p)) \text{ et } \hat{\alpha} = (\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p).$$

d'où

$$\hat{\alpha} = \mathfrak{R}_p^{-1}\hat{\rho}.$$

Ce qui nous aidera à faire les estimations $(\alpha_j)_{j=1, \dots, p}$ à l'aide des $\hat{\rho}_j$:

$$\hat{\alpha} = \mathfrak{R}_p^{-1}\hat{\rho}$$

Exemple 2.1.1 Soit $X_t \sim AR(2)$.

$$\begin{aligned} X_t &= \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t \\ \varepsilon_t &\sim \beta\beta(0, \sigma^2) \end{aligned}$$

Pour ce modèle

$$\begin{aligned} \hat{\alpha} &= \mathfrak{R}_p^{-1}\hat{\rho} \\ &= \mathfrak{R}_2^{-1}\hat{\rho} \end{aligned}$$

et

$$\mathfrak{R}_2 = \begin{pmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{pmatrix}, \quad \hat{\mathfrak{R}}_2^{-1} = \frac{1}{1 - \rho^2(1)} \begin{pmatrix} 1 & \hat{\rho}(1) \\ \hat{\rho}(1) & 1 \end{pmatrix}$$

donc

$$\begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{pmatrix} = \frac{1}{1 - \hat{\rho}^2(1)} \begin{pmatrix} 1 & \hat{\rho}(1) \\ -\hat{\rho}(1) & 1 \end{pmatrix} \begin{pmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \end{pmatrix}$$

déduire

$$\hat{\alpha}_1 = \frac{\hat{\rho}(1) + \hat{\rho}(1)\hat{\rho}(2)}{1 - \hat{\rho}^2(1)}$$

$$\hat{\alpha}_2 = \frac{-\hat{\rho}^2(1) + \hat{\rho}(2)}{1 - \hat{\rho}^2(1)}$$

Si $X_t \sim AR(p)$ alors les $|\rho(h)|$ et $\gamma(h)$ décroissent vers 0 exponentiellement avec h .

2.1.6 Fonction d'autocorrélation partielle PACF(h)

Si $X_t \sim AR(p)$ alors :

$$\begin{cases} \pi(h) = 0 & \text{si } h \geq p \\ \pi(h) \neq 0 & \text{sinon} \end{cases}$$

Dans un tel processus X_t et X_{t-p+1} sont indépendants conditionnellement aux valeurs intermédiaires $\{X_{t-1}, X_{t-2}, \dots, X_{t-p}\}$ et donc $h \geq p \implies \pi(h) = 0$.

La valeur à la date t dépend des p dates précédent et pas des autres. Cette propriété sert à l'indentification des modèles et à déterminer l'ordre p d'un processus $AR(p)$ au vu du corrélogramme partiel.

Généralement $(X_t)_{t \in \mathbb{Z}} \sim AR(p)$ alors :

$$\begin{cases} \gamma(h) = 0 & \text{si } |h| \geq p \\ \gamma(h) \neq 0 & \text{si } |h| = p \end{cases}$$

$$\begin{cases} \rho(h) = 0 & \text{si } |h| \geq p \\ \rho(h) \neq 0 & \text{si } |h| = p \end{cases}$$

$$\begin{cases} \pi(h) = 0 \text{ si } h \geq p \\ \pi(h) \neq 0 \text{ si non} \end{cases}$$

Écriture $MA(\infty)$: On suppose que $\Phi(L)X_t = \varepsilon_t$ et les racines de Φ sont de module strictement supérieur à 1.

$$\Phi(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p.$$

On suppose que

$$\Phi(Z) = \prod_{i=1}^p (1 - \lambda_i Z)$$

où

$$|\lambda_i| = \frac{1}{|Z_i|} \leq 1$$

Alors $\Phi(L)$ est inversible et

$$\Phi(L)^{-1} = \sum_{k=0}^{+\infty} \beta_k L^k$$

tel que

$$\sum_{k=1}^{+\infty} |\beta_k| \leq \infty, \beta_0 = 1$$

Alors

$$\begin{aligned} X_t &= \sum_{k=0}^{+\infty} \beta_k L^k \varepsilon_t \\ &= \sum_{k=0}^{+\infty} \beta_k \varepsilon_{t-k} \end{aligned}$$

2.2 Modèle moyenne mobile $MA(q)$

Définition 2.2.1 On appelle processus moyenne Mobile (moving average) d'ordre q , noté $MA(q)$ un processus (X_t) vérifiant une relation du type

$$X_t = \mu + \varepsilon_t + \sum_{j=1}^q \beta_j \varepsilon_{t-j}, \text{ pour tout } j \in \mathbb{Z}.$$

où les β_j sont des réels et $\varepsilon_t \sim \beta\beta(0, \sigma^2)$.

Equivalent l'écriture

$$X_t = \mu + \Theta(L)\varepsilon_t \text{ où } \Theta(L) = 1 - \beta_1 L + \dots + \beta_q L^q.$$

Ecriture $AR(\infty)$ quand les racines de β sont de module ≥ 1 :

sous ces hypothèses $\Theta(L)$ est inversible et $\Theta(L)^{-1} = \sum_{k=0}^{+\infty} a_k L^k$ avec $a_0 = 1$ et

$\sum_{k=0}^{+\infty} |a_k| \leq +\infty$, Il s'en suit que

$$\begin{aligned} X_t - m &= \Theta(L)\varepsilon_t \Leftrightarrow \Theta(L)^{-1}(X_t - m) = \varepsilon_t \\ &\Leftrightarrow \Theta(L)^{-1}X_t = \frac{m}{\Theta(1)} + \varepsilon_t \end{aligned}$$

Soit encore

$$\sum_{k=0}^{\infty} a_k X_{t-k} - \mu = \varepsilon_t \text{ où } \mu = \frac{m}{\Theta(1)}$$

D'où la représentation canonique $AR(\infty)$:

$$X_t = - \sum_{k=1}^{\infty} a_k X_{t-k} + \frac{m}{\Theta(1)} + \varepsilon_t.$$

2.2.1 Stationnarité

Les modèles $MA(q)$ sont toujours stationnaire.

2.2.2 Inversibilité

La modèle $MA(q)$ ne sont inversibles que si les racines du polynôme caractéristique sont à l'extérieur du disque unité.

2.2.3 Fonction d'autocorrélation

La fonction d'autocovariance est donnée par :

$$\gamma(h) = \begin{cases} 0 & \text{si } |h| \geq q \\ -\beta_q \sigma_\varepsilon^2 \neq 0 & \text{si } |h| = q \\ \sigma_\varepsilon^2 \left(-\beta_h + \sum_{j=h+1}^q \beta_j \beta_{j-h} \right) & \text{si } 1 \leq |h| \leq q. \\ \sigma_\varepsilon^2 \left(1 + \sum_{j=h+1}^q \beta_j^2 \right) & \text{si } h = 0 \end{cases}$$

Preuve

$$\begin{aligned} \gamma(h) &= E(X_t X_{t-h}) \\ &= E([\varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q}] [\varepsilon_{t-h} + \beta_1 \varepsilon_{t-1-h} + \dots + \beta_q \varepsilon_{t-q-h}]) \\ &= \begin{cases} [\beta_h + \beta_1 \beta_{h+1} + \dots + \beta_q \beta_{q-h}] \sigma^2 & \text{si } 1 \leq h \leq q \\ 0 & \text{si } h \geq q \end{cases} \end{aligned}$$

Avec pour $h = 0$, la relation :

$$\gamma(0) = [1 + \beta_1^2 + \dots + \beta_q^2] \sigma^2$$

Cette dernière relation peut se réécrire :

$$\gamma(k) = \sigma^2 \sum_{j=0}^q \beta_j \beta_{j+k}$$

avec la convention $\beta_0 = 1$.

D'où la fonction d'autocorrélation :

$$\rho(h) = \frac{\beta_h + \beta_1 \beta_{h+1} + \dots + \beta_q \beta_{q-h}}{1 + \beta_1^2 + \dots + \beta_q^2} \quad \text{si } 1 \leq h \leq q.$$

On peut noter un particulier que $\gamma(q) = \sigma^2 \beta_q \neq 0$, alors que $\gamma(q+1) = 0$.

Cette propriété sera relativement pratique pour faire l'estimation de l'ordre de processus *MA*.

2.2.4 Modèle $MA(1)$:

La forme générale des processus de type $MA(1)$ est :

$$X_t = \varepsilon_t + \beta\varepsilon_{t-1} \text{ pour tout } t \in \mathbb{Z}$$

où $\varepsilon_t \sim \beta\beta(0, \sigma^2)$, Les autocorrélations sont données par

$$\rho(1) = \frac{-\beta}{1 + \beta^2}, \text{ et } \rho(h) = 0, \text{ pour } h \geq 2.$$

On peut noter que $-1/2 \leq \rho(1) \leq 1/2$: les modèles $MA(1)$ ne peuvent avoir de fortes autocorrélations à l'ordre 1.

Pour un modèle $MA(1)$:

$$E(X_t) = 0.$$

$$\gamma(0) = \text{var}(X_t) = \sigma^2(1 + \beta^2),$$

$$\gamma(1) = -\beta\sigma_\varepsilon^2,$$

$$\rho(1) = \frac{-\beta}{1 + \beta^2},$$

$$\gamma(k) = \rho(k) = 0 \text{ pour } k \geq 2.$$

-Tableaux récapitulatifs des différents situations :

Processus	Autocorrélation $\rho(h)$	Autocorrélation partielle $\gamma(h)$
AR(p)	<i>amortie</i>	nulle pour $h \geq p$
MA(q)	pour $h \geq q$	<i>amortie</i>

TAB. 2.1 – Table récapitulatifs des différents situations

2.3 Les modèles ARMA(p,q)

Nous allons maintenant introduire un modèle de processus stationnaire comportant une partie *AR* et une partie *MA*. C'est pourquoi il porte le nom de processus *ARMA* (AutoRegressive Moving Average). Les processus *ARMA* sont très importants en pratique car on peut montrer que tout processus stationnaire peut être approché par un processus *ARMA*.

Définition 2.3.1 *On dit qu'un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ admet une représentation ARMA(p,q) s'il vérifie l'équation :*

$$X_t = \mu + \varepsilon_t + \sum_{j=1}^q \beta_j \varepsilon_{t-j} + \sum_{i=1}^p \alpha_i X_{t-i} \quad \text{pour } t \in \mathbb{Z}$$

où ε_t est un bruit blanc $\beta \beta(0, \sigma^2)$ et $\alpha_1, \alpha_2, \dots, \alpha_p, \beta_1, \beta_2, \dots, \beta_q$ sont des coefficients réels, avec $\alpha_p \neq 0$ et $\beta_q \neq 0$.

Autrement dit, en utilisant les polynômes en L .

$$\Phi(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p$$

et

$$\Theta(L) = 1 - \beta_1 L + \dots + \beta_q L^q$$

Un processus *ARMA*(p,q) vérifie l'équation :

$$\Phi(L)X_t = \Theta(L)\varepsilon_t$$

pour tout t dans \mathbb{Z} .

Naturellement un processus *AR*(p) est un *ARMA*(p, 0) et un processus *MA*(q) est *ARMA*(0, q).

La représentation d'un processus *ARMA*(p,q).

$$\Phi(L)X_t = \Theta(L)\varepsilon_t$$

est dite

-**minimale** : si les polynômes $\Phi(z)$ et $\Theta(z)$ n'ont pas de racine commune.

-**causale** : si le polynôme $\Phi(z)$ a toutes ses racines à l'extérieur du disque unité.

-**inversible** : si le polynôme $\Theta(z)$ a toutes ses racines à l'extérieur du disque unité.

-**canonique** : si elle est causale et inversible.

2.3.1 Auto-covariance et auto-corrélation :

i) Pour $h \geq q$, les $\gamma(h)$ et les $\rho(h)$ vérifiant les équations de récurrence d'ordre p :

$$\gamma(h) - \alpha_1\gamma(h-1) - \dots - \alpha_p\gamma(h-p) = 0$$

$$\rho(h) - \alpha_1\rho(h-1) - \dots - \alpha_p\rho(h-p) = 0$$

ii) Elles décroissent donc vers 0 exponentiellement pour $h \geq q$.

ii) Un processus $ARMA(p, q)$ vérifie alors *équations de Yule-Walker*.

2.3.2 Processus ARMA(1,1)

En prenant $p = 1$ et $q = 1$, la série X_t suit un processus $ARMA(1, 1)$ si on peut écrire :

$$(1 - \alpha L)X_t = (1 - \beta L)\varepsilon_t$$

Les propriétés d'un $ARMA(1, 1)$ sont :

- Pour que la série X_t soit stationnaire et inversible si $|\Phi| \leq 1$ et $|\Theta| \leq 1$.
- L'ACF d'un $ARMA(1, 1)$ décroît exponentiellement à partir de $h = 1$ alors que l'ACF d'un $AR(1)$ décroît exponentiellement à partir de $h = 0$.
- Le PACF d'un $ARMA(1, 1)$ ressemble à celle d'un $MA(1)$ à partir de $h = 2$.

La sortie ci-dessus montre l'évolution de l'autocorrélogramme dans le cas $ARMA(1, 1)$

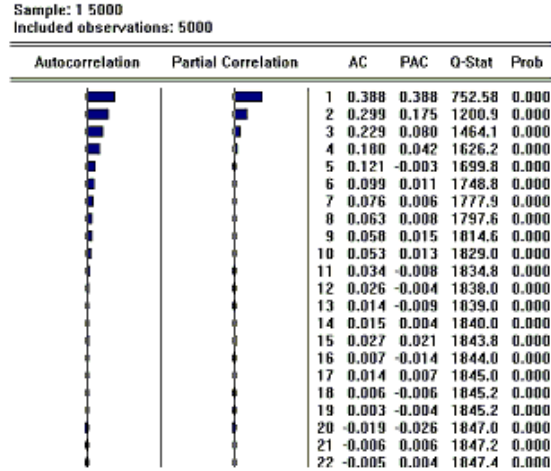


FIG. 2.1 – Autocorrélogramme d'un $ARMA(1, 1)$.

2.4 Les modèles $ARIMA(p, q)$

Les processus étant dans la pratique rarement stationnaires, on a introduit une généralisation des processus $ARMA$ vus précédemment de manière à les étendre à des processus non stationnaires. L'idée générale, essentiellement conçue pour les processus non stationnaires a tendance polynômiale, est de différencier suffisamment le processus initial afin d'obtenir un processus sans tendance et sur la partie différenciée appliquer un modèle $ARMA$. Cette classe de modèle est connue sous le nom de processus $ARIMA$ (AutoRegressive Integrated Moving Average).

Définition 2.4.1 On dit qu'un processus $(X_t)_{t \in \mathbb{Z}}$ admet une représentation $ARIMA(p, d, q)$ s'il vérifie l'équation :

$$\Phi(L)\nabla^d X_t = \Theta(L)\varepsilon_t, \text{ pour tout } t \text{ dans } \mathbb{Z}.$$

où

$$\nabla^d = (I - L)^d$$

Et ceci avec

$$\varepsilon_t \sim \beta\beta(0, \sigma^2)$$

$$\Phi(L) = 1 - \alpha_1 L - \dots - \alpha_p L^p \quad \text{où } \alpha_p \neq 0$$

$$\Theta(L) = 1 - \beta_1 L + \dots + \beta_q L^q \quad \text{où } \beta_q \neq 0$$

$\Theta(L)$ et $\Phi(L)$ ont leurs racines de module ≥ 1 et n'ont pas de racines communes.

On peut montrer que, dans le cas d'un processus $ARIMA(p, d, q)$, le processus $\nabla^d X_t$ est asymptotiquement, au sens quand $t \rightarrow +\infty$, un processus $ARMA(p, q)$. On peut aussi établir des représentations $AR(\infty)$ et $MA(\infty)$ pour les processus $ARIMA$.

2.5 Méthode de Box-Jenkins

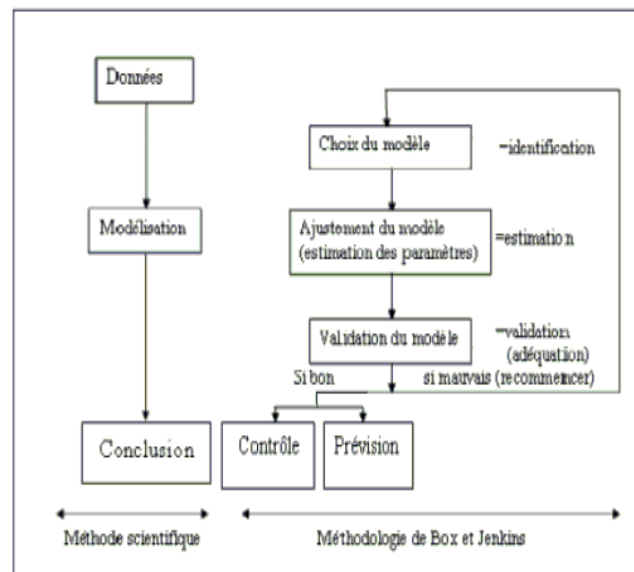
Box & Jenkins (1970) ont élaboré une méthodologie pour identifier un modèle adéquat pour une série chronologique. Leur méthode est fondée sur les modèles $ARIMA$. Dans ce chapitre, les principales étapes de cette technique sont présentées. Pour les méthodes d'inférence présentées dans la suite, on supposera que T réalisations d'une série chronologique univariée, notées X_1, X_2, \dots, X_n ont été observées.

2.5.1 Les étapes de la méthodologie

La méthodologie de Box et Jenkins comporte essentiellement les étapes suivantes :

- Transformation des données afin de stabiliser la variance (log, sqrt,...) et différenciation des données pour les stationariser.
- Visualisation des ACF et des $PACF$ empiriques pour identifier les paramètres p et q appropriés.
- Estimation des paramètres du(des) modèle(s) sélectionné(s).
- Diagnostique et tests d'adéquation du modèle.
- Prévision : La dernière étape consiste à la prévision des valeurs futures à travers le modèle retenu.

jenkins



3.pdf

FIG. 2.2 – Organigramme de la méthodologie de Box et Jenkins.

2.5.2 La familiarisation avec les données

Avant toute étude de séries il est indispensable de s’informer sur le domaine dont relèvent les données, leur nature (mensuelle, annuelle,...), les objectifs poursuivis (prévision ponctuelle ou par intervalle,...), puis de la représenter graphiquement. En effet, quelques fois une simple visualisation du graphe permet de détecter ou soupçonner l’existence de plusieurs composantes (tendance, saisonnalité,...).

Analyse préliminaire

Cette étape met à notre disposition des options utiles pour : corriger les données aberrantes, abandonner une partie des données manquantes, transformer les données (transformation logarithmiques,...). En effet, l’examen visuel du graphique ou du tableau ne permet pas toujours de déterminer avec certitude l’existence d’une tendance ou d’une saisonnalité, de surcroît il interdit l’automatisme de traitement qui peut s’avérer nécessaire dans le cas d’un nombre important de séries à examiner. Les tests permettent de pallier ces deux inconvénients.

1-Analyse de la tendance

Le test de Dickey-Fuller Augmenté permettent de déterminer si la série est stationnaire et dans le cas d’une non-stationnarité de quel type il s’agit : *TS* et

DS.

- Si la série étudiée est de type *TS*, il convient de la stationnariser par régression sur le temps et le résidu d'estimation est alors étudié selon la méthodologie de Box-Jenkins.

- Si la série étudiée est de type *DS*, il convient de la stationnariser par passage aux différences selon l'ordre d'intégration $I = d$ (c'est-à-dire le nombre de fois qu'il faut différencier la série pour la rendre stationnaire). La série différenciée est alors étudiée selon la méthodologie de Box-Jenkins.

Stratégie de tests de Dickey et Fuller :

Les différents modèles retenus pour construire les statistiques des tests les suivants

Modèle[1] :

$$\Delta y_t = \rho y_{t-1} + \alpha + \beta t + \sum_{i=1}^p \phi_i \Delta y_{t-i} + \varepsilon_t$$

Modèle[2] :

$$\Delta y_t = \rho y_{t-1} + \alpha + \sum_{i=1}^p \phi_i \Delta y_{t-i} + \varepsilon_t$$

Modèle[3] :

$$\Delta y_t = \rho y_{t-1} + \sum_{i=1}^p \phi_i \Delta y_{t-i} + \varepsilon_t$$

2-Analyse de la saisonnalité

Le test de Fisher suppose la chronique sans tendance ou encore sans saisonnalité. Dans le cas d'une série affectée d'un mouvement saisonnier, on procède à la dessaisonnalisation de la série.

Transformation

La transformation log permet de stabiliser la variance d'une série si la variance de la série originale croît avec les valeurs de celle-ci. Cette transformation permet aussi parfois de se rapprocher de la normalité. Une série non stationnaire est une série qui en général croît avec le temps. On peut donc être tenté de retirer par régression un trend temporel de cette série. L'autre solution consiste à utiliser un filtre du type $(1 - L)$. On va alors différencier la série. Si l'on a tout d'abord pris le logarithme, la série transformée :

$$(1 - L) \log(y_t) = \log(y_t) - \log(y_{t-1})$$

2.5.3 Identification du modèle

La méthode d'identification est essentiellement fondée sur l'analyse conjointe des auto-corrélations et des auto-corrélations partielles. Nous supposons le processus stationnaire $ARMA(p, q)$. Elle est fondée sur l'étude des corrélogrammes simple et partiel :

- Si le corrélogramme simple n'a que ses q premiers termes ($q = 3$ maximum) différents de 0 et que les termes du corrélogramme partiel diminuent lentement, nous pouvons pronostiquer un $MA(q)$.

- Si le corrélogramme partiel n'a que ses p premiers termes ($p = 3$ maximum) différents de 0 et que les termes du corrélogramme simple diminuent lentement, cela caractérise un $AR(p)$.

- Si les fonctions d'autocorrélation simple et partiel ne paraissent pas tronquées, il s'agit alors d'un processus de type ARMA, dont les paramètres dépendent de la forme particulière des corrélogrammes.

2.5.4 Estimation des paramètres

Une fois que l'on a décidé de chercher pour la série temporelle X_t un modèle $ARMA(p, q)$, reste à estimer les coefficients qui le caractérisent :

Si X_t est un processus $ARMA(p, q)$, vérifiant $\phi(L)X = \theta(L)\varepsilon$, alors sa fonction d'autocovariance vérifie la relation de récurrence suivante :

$$\forall t \geq q + 1, \gamma(t) + \phi_1\gamma(t - 1) + \dots + \phi_p\gamma(t - p) = 0$$

Les estimateurs de Yule-Walker de $\phi_1, \phi_2, \dots, \phi_p$ sont naturellement définis par :

$$\begin{pmatrix} \hat{\rho}(q) & \cdot & \cdot & \cdot & \hat{\rho}(q - p + 1) \\ \cdot & \cdot & \cdot & \cdot & \hat{\rho}(q - p + 2) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \hat{\rho}(q + p - 1) & \cdot & \cdot & \cdot & \hat{\rho}(q) \end{pmatrix} \begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \cdot \\ \cdot \\ \hat{\phi}_p \end{pmatrix} = - \begin{pmatrix} \hat{\rho}(q + 1) \\ \hat{\rho}(q + 2) \\ \cdot \\ \cdot \\ \hat{\rho}(q + p) \end{pmatrix}$$

Critères de choix des modèles

Souvent il n'est pas facile de déterminer un modèle unique. Le modèle qui est finalement choisi est celui qui minimise l'un des critères à partir de T observations.

1-Critères d'information :

critères de choix de modèles les plus courants sont :

– Le critère *AIC* (Akaike Information Criterion), qui sera généralement préféré si l'objectif de l'étude est de faire de la prévision, et qui est défini par :

$$AIC(p, q) = -2 \log L + 2(p + q)$$

– Le critère *BIC* (Bayesian Information Criterion) sera quant à lui généralement préféré si l'objectif de l'étude est de s'ajuster à la série observée, et est défini par :

$$BIC(p, q) = -2 \log L + (p + q) \log T$$

où $\log L$ est la log-vraisemblance du modèle $ARMA(p, q)$ estimé et T est le nombre d'observations. Les modèles ayant la plus petite valeur du critère devront être choisis.

2-Critères d'erreur :

Dans l'échantillon initial, (X_1, \dots, X_T) on considère seulement une partie d'observations. Le reste sera à prévoir par le modèle. On peut alors considérer les erreurs :

– Racine quadratique moyenne :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{X}_i - X_i)^2}$$

– Quadratique moyenne :

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{X}_i - X_i)^2$$

– Absolue :

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{X}_i - X_i|$$

2.5.5 Validation du modèle

A l'étape de l'identification, les incertitudes liées aux méthodes employées font que plusieurs modèles en général sont estimés et c'est l'ensemble de ces modèles qui subissent alors l'épreuve des tests. Il en existe de très nombreux permettant d'une part de valider le modèle retenu.

Tests concernant les paramètres

Tous les coefficients du modèle retenu doivent être significativement différents de zéro. Il convient donc d'utiliser le test de Student classique.

1-Test de student sur les paramètres : Afin que le modèle soit valide, il faut que tous les coefficients soient significativement différents de zéro (donc que leurs probabilités critiques soient inférieurs à 0,05).

2-Les tests de bruit blanc normal

Le processus estimé est bien évidemment de bonne qualité si la chronique calculée suit les évolutions de la chronique empirique. Les résidus entre les valeurs observées et les valeurs calculées par le modèle, doivent donc se comporter comme un bruit blanc normal.

a-Tests de recherche d'autocorrélation :

Si les résidus obéissent à un bruit blanc, il ne doit pas exister d'autocorrélation dans la série. Le test suivant peut être utilisé.

b-Tests de normalité :

Le test de Jarque et Bera (1984) peut s'appliquer pour tester la normalité des résidus, il est fondé sur la notion de Skewness et de Kurtosis.

Les tests du Skewness et de Kurtosis : Soit

$$\mu_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k$$

le moment centré d'ordre k .

le coefficient de Skewness (coefficient d'asymétrie de la distribution) est égal à :

$$s = \frac{\mu_3}{\mu_2^{\frac{3}{2}}}$$

et le coefficient de Kurtosis (aplatissement-épaisseur des queues)

$$k = \frac{\mu_4}{\mu_2^2}$$

Si la distribution est normale et le nombre d'observations grand :

$$s \rightarrow N\left(0, \sqrt{\frac{6}{n}}\right)$$

et

$$k \rightarrow N\left(3, \sqrt{\frac{6}{n}}\right)$$

On construit alors les statistiques :

$$v_1 = \frac{s - 0}{\sqrt{\frac{6}{n}}}$$

et

$$v_2 = \frac{k - 3}{\sqrt{\frac{24}{n}}}$$

que l'on compare à 1,96 au seuil de 5%. Soient les hypothèses suivantes :

$$\begin{cases} H_0 : v_1 = 0, v_2 = 0 \\ H_1 : v_1 \neq 0 \text{ ou } v_2 \neq 0 \end{cases}$$

On acceptera avec un risque de 5% l'hypothèse H_0 si $v_1 < 1.96$ et $v_2 < 1.96$, dans le cas contraire, l'hypothèse de normalité est rejetée.

Le test de Jarque et Bera :

il s'agit d'un test qui regroupe les résultats précédents, si s et k obéissent à des lois normales alors la quantité :

$$JB = \frac{n}{6}s^2 + \frac{n}{24}(k - 3)^2 \sim \chi^2(2).$$

Donc si $JB \geq \chi_{1-\alpha}^2$, on rejette l'hypothèse H_0 de normalité des résidus au seuil α .

2.5.6 La prévision

On suppose (p, d, q) connus. On dispose d'observations X_1, \dots, X_T et on veut faire une prévision à l'horizon h , c'est-à-dire prévoir X_{T+1}, \dots, X_{T+h} .

1-Prévision d'un modèle AR(p) :

Le modèle s'écrit :

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t.$$

La prévision optimale à la date $T + 1$, faite à la date T est

$$\hat{X}_{T+1} = E(X_{T+1}/X_T, X_{T-1}, \dots, X_1).$$

Donc

$$\hat{X}_{T+1} = \phi_1 X_T + \dots + \phi_p X_{T+1-p}$$

De façons analogue

$$X_{T+h} = \phi_1 X_{T+h-1} + \dots + \phi_p X_{T+h-p} + \varepsilon_{T+h}.$$

Donc

$$\hat{X}_{T+h} = E(X_{T+h}/X_T, X_{T-1}, \dots, X_1).$$

De façons réccursive

$$\hat{X}_{T+h} = \begin{cases} \phi_1 \hat{X}_{T+h-1} + \dots + \phi_{h-1} \hat{X}_{T+1} + \phi_h X_T + \dots + \phi_h X_{T+h-p} & \text{pour } h \leq p, \\ \phi_1 \hat{X}_{T+h-1} + \dots + \phi_p \hat{X}_{T+h-p} & \text{pour } h > p. \end{cases}$$

2-Prévision d'un modèle MA(q) :

Soit le modèle $MA(q)$

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} = \Theta(L)\varepsilon_t.$$

La prévision optimale à la date $T + 1$ est

$$\hat{X}_{T+1} = \theta_1 \varepsilon_T + \dots + \theta_q \varepsilon_{T+1-q}.$$

De facons analogue

$$\hat{X}_{T+h} = \begin{cases} \theta_h \hat{X}_T + \dots + \theta_q \hat{X}_{T+h-q} & \text{pour } h \leq q. \\ 0 & \text{pour } h > q. \end{cases}$$

3-Prévision d'un modèle ARMA(p,q) :

(X_t) est un ARMA(p, q), alors :

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j}.$$

Donc

$$X_{T+h} = \sum_{i=1}^p \phi_i X_{T+h-i} + \varepsilon_{T+h} - \sum_{j=1}^q \theta_j \varepsilon_{T+h-j}.$$

Ona alors

$$\hat{X}_{T+h} = \sum_{i=1}^p \phi_i \hat{X}_{T+h-i} - \sum_{j=1}^q \theta_j \hat{\varepsilon}_{T+h-j}.$$

Où

$$\begin{cases} \hat{\varepsilon}_{T+h-j} = 0 & \text{si } h > j \\ \hat{\varepsilon}_{T+h-j} = \varepsilon_{T+h-j} & \text{si } h \leq j \end{cases}$$

Chapitre 3

Application de la méthode de Box-Jenkins

Les données considérées à cet effet sont mensuelles et la période retenue pour l'étude va de *Janvier 2013* à *décembre 2017* l'évolution du nombre des malades diabètes de l'hôpital de Ouargla Mohamed Boudiaff, 60 observations, comme indiqué dans le tableau ci-dessous. Pour construire le modèle de prévision, nous utilisons logiciels Eviews et gretl (prédiction), mais avant tout ,nous commençons par l'analyse exploratoire (plot) de la série chronique puis l'analyse de la stationnarité (test de la racine unitaire).

Mois/Années	2013	2014	2015	2016	2017
Janvier	557	495	501	671	860
Fevrier	359	428	600	556	1083
Mars	402	525	200	754	777
Avril	548	592	302	761	800
Mai	547	609	700	814	939
Juin	546	701	560	917	461
Juillet	414	278	200	480	822
Août	243	330	600	659	922
Septembre	424	496	411	450	717
Octobre	466	459	800	928	1008
Novembre	444	429	788	860	893
Décembre	534	581	700	874	400

TAB. 3.1 – Nombre des malades diabètes de l'hôpital de Ouargla Mohamed Boudiaff 2013-2017.

3.1 Analyse exploratoire des données :

La série DIABET est représentée dans le graphe suivant :

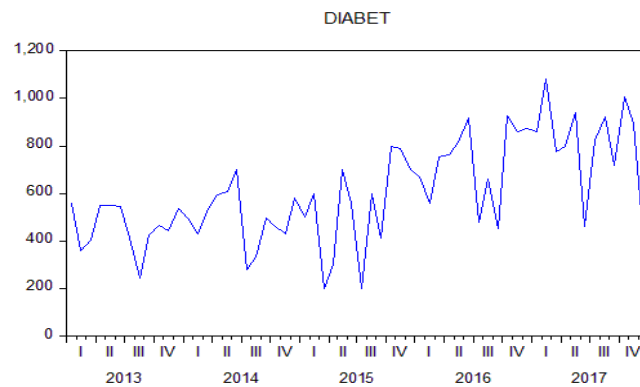


FIG. 3.1 – La série diabet.

On peut également travailler avec de données logarithmiques. Dans ce cas, le plot de la série linéarisée se présente comme suit :

$$\text{genre ldiabet} = \text{Log}(\text{diabet})$$

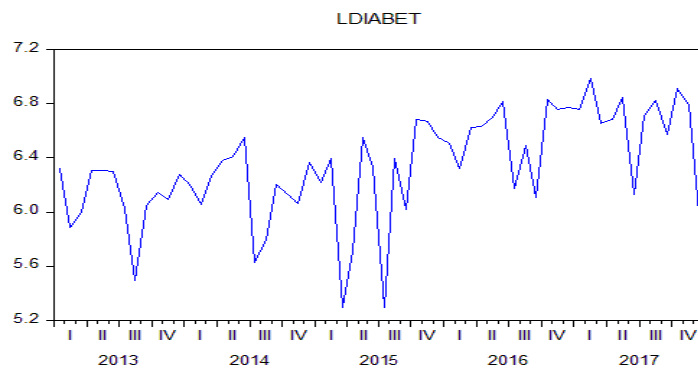


FIG. 3.2 – La série ldiabet.

3.2 Etude de la stationnarité

3.2.1 Analyse des Corrélogrammes de la séries ldiabet

Date: 05/03/18 Time: 19:12
Sample: 2013M01 2017M12
Included observations: 59

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
█	█	1 -0.369	-0.369	8.4656	0.004
█	█	2 -0.138	-0.317	9.6596	0.008
█	█	3 0.038	-0.188	9.7541	0.021
█	█	4 -0.024	-0.172	9.7927	0.044
█	█	5 -0.046	-0.199	9.9367	0.077
█	█	6 0.170	0.043	11.891	0.064
█	█	7 -0.152	-0.115	13.492	0.061
█	█	8 0.126	0.094	14.619	0.067
█	█	9 -0.017	0.068	14.639	0.101
█	█	10 -0.177	-0.138	16.932	0.076
█	█	11 0.073	-0.079	17.332	0.098
█	█	12 0.224	0.177	21.184	0.048
█	█	13 -0.190	0.028	23.998	0.031
█	█	14 0.097	0.127	24.746	0.037
█	█	15 -0.213	-0.196	28.468	0.019
█	█	16 0.051	-0.112	28.681	0.026
█	█	17 0.111	-0.040	29.728	0.028
█	█	18 0.036	0.047	29.845	0.039
█	█	19 -0.078	0.043	30.389	0.047
█	█	20 0.057	-0.030	30.690	0.059

6.pdf

Les coefficients d'autocorrélation simples décroissent lentement, de manière linéaire, qui montre la non stationnarité de notre série.

3.2.2 Test de la racine unitaire

Le test de racine unitaire (test ADF) est confirmé la non stationnarité.

1) Modèle avec tendance et constante

Null Hypothesis: LDABET has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-6.501157	0.0000
Test critical values:		
1% level	-4.121303	
5% level	-3.487845	
10% level	-3.172314	

*Mackinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LDABET)
Method: Least Squares
Date: 05/02/18 Time: 23:14
Sample (adjusted): 2013M02 2017M12
Included observations: 59 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LDABET(-1)	-0.893933	0.137504	-6.501157	0.0000
C	5.318763	0.820593	6.480823	0.0000
@TREND("2013M01")	0.011421	0.003115	3.565247	0.0005

R-squared 0.430457 Mean dependent var -0.005612
Adjusted R-squared 0.410126 S.D. dependent var 0.426594
S.E. of regression 0.327638 Akaike info criterion 0.855695
Sum squared resid 6.011415 Schwarz criterion 0.761332
Log likelihood -16.34299 Hannan-Quinn criter. 0.696931
F-statistic 21.16308 Durbin-Watson stat 1.881382
Prob(F-statistic) 0.000000

La série non stationnaire avec trend et intercept significatif car la p-value < 0.05.

2) Modèle sans tendance et avec constante

Null Hypothesis: LDIABET has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.869259	0.0002
Test critical values:		
1% level	-3.54099	
5% level	-2.911730	
10% level	-2.593551	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LDIABET)
Method: Least Squares
Date: 05/02/18 Time: 23:31
Sample (adjusted): 2013M02 2017M12
Included observations: 59 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LDIABET(-1)	-0.594225	0.122036	-4.869259	0.0000
C	3.761414	0.775066	4.853027	0.0000
R-squared	0.293765	Mean dependent var		-0.005612
Adjusted R-squared	0.281375	S.D. dependent var		0.426594
S.E. of regression	0.361631	Akaike info criterion		0.836928
Sum squared resid	7.454302	Schwarz criterion		0.907353
Log likelihood	-22.68936	Hannan-Quinn criter.		0.864419
F-statistic	23.70969	Durbin-Watson stat		2.034257
Prob(F-statistic)	0.000009			

La série non stationnaire avec intercept significatif (p-value < 0.05).

3) Modèle sans tendance et sans constante

Null Hypothesis: LDIABET has a unit root
Exogenous: None
Lag Length: 2 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	0.134417	0.7219
Test critical values:		
1% level	-2.606163	
5% level	-1.946654	
10% level	-1.613122	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LDIABET)
Method: Least Squares
Date: 05/02/18 Time: 23:34
Sample (adjusted): 2013M04 2017M12
Included observations: 57 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LDIABET(-1)	0.001062	0.007981	0.134417	0.8936
D(LDIABET(-1))	-0.650049	0.133217	-4.128973	0.0001
D(LDIABET(-2))	-0.370057	0.131747	-2.808850	0.0069
R-squared	0.256338	Mean dependent var		-8.75E-05
Adjusted R-squared	0.228795	S.D. dependent var		0.429932
S.E. of regression	0.377559	Akaike info criterion		0.941015
Sum squared resid	7.697731	Schwarz criterion		1.048544
Log likelihood	-23.81891	Hannan-Quinn criter.		0.962804
Durbin-Watson stat	2.079101			

La série non stationnaire avec une racine unitaire (p-value>0.05).

Alors la série ldiabet est non stationnaire.

3.2.3 Stationnarisation de serie ldiabet :

Il faut appliquer le filtre de différence pour stationnariser la série :

$$\text{genre dldiabet} = d(\text{ldiabet})$$

Test de la racine sur dldiabet :

On applique le test de ADF :

1) Modèle avec tendance et constante

```

Null Hypothesis: DLDIABET has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on SIC, maxlag=10)

```

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-8.564284	0.0000
Test critical values:		
1% level	-4.127338	
5% level	-3.490662	
10% level	-3.173943	

*MacKinnon (1996) one-sided p-values.

```

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DLDIABET)
Method: Least Squares
Date: 05/03/18 Time: 18:11
Sample (adjusted): 2013M04 2017M12
Included observations: 57 after adjustments

```

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLDIABET(-1)	-1.917923	0.223944	-8.564284	0.0000
D(DLDIABET(-1))	0.368489	0.132795	2.774938	0.0076
C	0.040736	0.107586	0.378831	0.7065
@TREND("2013M01")	-0.000910	0.003068	-0.296576	0.7679

R-squared	0.723426	Mean dependent var	-0.016075
Adjusted R-squared	0.707771	S.D. dependent var	0.704111
S.E. of regression	0.330630	Akaike info criterion	0.973614
Sum squared resid	7.578599	Schwarz criterion	1.116986
Log likelihood	-23.74799	Hannan-Quinn criter.	1.029333
F-statistic	46.21015	Durbin-Watson stat	2.082413
Prob(F-statistic)	0.000000		

Il se dégage que la tendance n'est pas significativement différente de zéro car la probabilité critique associée au trend est supérieure à 5 pour cent.

2) Modèle sans tendance et avec constante

Null Hypothesis: DLDIABET has a unit root
Exogenous: None
Lag Length: 1 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-8.730733	0.0000
Test critical values:		
1% level	-2.606163	
5% level	-1.946654	
10% level	-1.613122	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DLDIABET)
Method: Least Squares
Date: 05/03/18 Time: 18:23
Sample (adjusted): 2013M04 2017M12
Included observations: 57 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLDIABET(-1)	-1.917262	0.219599	-8.730733	0.0000
D(DLDIABET(-1))	0.368866	0.130270	2.831556	0.0061
R-squared	0.722644	Mean dependent var		-0.016071
Adjusted R-squared	0.717601	S.D. dependent var		0.704111
S.E. of regression	0.374173	Akaike info criterion		0.906286
Sum squared resid	7.700306	Schwarz criterion		0.977941
Log likelihood	-23.82845	Hannan-Quinn criter.		0.934121
Durbin-Watson stat	2.078912			

Il se dégage que la constante n'est pas significativement différent de zéro .

3) Modèle sans tendance et sans constante

Null Hypothesis: DLDIABET has a unit root
Exogenous: None
Lag Length: 1 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-8.730733	0.0000
Test critical values:		
1% level	-2.606163	
5% level	-1.946654	
10% level	-1.613122	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DLDIABET)
Method: Least Squares
Date: 05/03/18 Time: 18:23
Sample (adjusted): 2013M04 2017M12
Included observations: 57 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLDIABET(-1)	-1.917262	0.219599	-8.730733	0.0000
D(DLDIABET(-1))	0.368866	0.130270	2.831556	0.0061
R-squared	0.722644	Mean dependent var		-0.016071
Adjusted R-squared	0.717601	S.D. dependent var		0.704111
S.E. of regression	0.374173	Akaike info criterion		0.906286
Sum squared resid	7.700306	Schwarz criterion		0.977941
Log likelihood	-23.82845	Hannan-Quinn criter.		0.934121
Durbin-Watson stat	2.078912			

La série n'a pas une racine unitaire (p-value<0.05).

On accepte donc l'hypothèse alternative selon laquelle la série dldiabet est stationnaire. $dldiabet \rightarrow I(0)$: la série dldiabet suit un processus ARMA et ldiabet $\rightarrow I(1)$: la série ldiabet suit processus ARIMA .

Graphique de la série stationnaire :

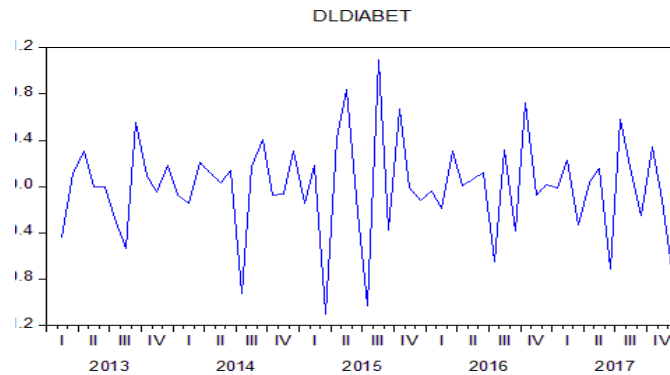


FIG. 3.3 – La série dldiabet.

Il ressort de ce graphique (après différence première) que la série fluctue autour d'une moyenne constante.

3.3 Modélisation ARMA par la méthode de BOX-JENKINS

La série dldiabet est stationnaire .On va lui chercher un modèle de type $ARMA(p, q)$.

3.3.1 Identification des order p et q du modèle ARMA

Pour connaître les ordres du modèle $ARMA(p, q)$, nous allons nous servir de corrélogrammes de la série stationnaire dldiabet. En effet, le corrélogramme simple permet d'identifier un modèle $MA(q)$, alors que le corrélogramme partiel permet d'identifier un modèle $AR(p)$.

corrélogramme de la série stationnaire

Date: 05/03/18 Time: 19:12
Sample: 2013M01 2017M12
Included observations: 59

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
█	█	1	-0.369	-0.369	8.4656	0.004
█	█	2	-0.138	-0.317	9.6596	0.008
█	█	3	0.038	-0.188	9.7541	0.021
█	█	4	-0.024	-0.172	9.7927	0.044
█	█	5	-0.046	-0.199	9.9367	0.077
█	█	6	0.170	0.043	11.891	0.064
█	█	7	-0.152	-0.115	13.492	0.061
█	█	8	0.126	0.094	14.619	0.067
█	█	9	-0.017	0.068	14.639	0.101
█	█	10	-0.177	-0.138	16.932	0.076
█	█	11	0.073	-0.079	17.332	0.098
█	█	12	0.224	0.177	21.184	0.048
█	█	13	-0.190	0.028	23.998	0.031
█	█	14	0.097	0.127	24.746	0.037
█	█	15	-0.213	-0.196	28.468	0.019
█	█	16	0.051	-0.112	28.681	0.026
█	█	17	0.111	-0.040	29.728	0.028
█	█	18	0.036	0.047	29.845	0.039
█	█	19	-0.078	0.043	30.389	0.047
█	█	20	0.057	-0.030	30.690	0.059

Il ressort de ces corrélogrammes que la première autocorrélation (Simple et partielle) et deuxième autocorrélation (partielle) de la série *dldiabet* est significativement différente de zéro.

Comme il s'agit de processus $ARIMA(1, 1, 1)$ et $ARIMA(2, 1, 0)$ nous trouvons les modèles pour les estimations : $AR(1)$, $MA(1)$, $AR(1)MA(1)$, $AR(2)$, $AR(2)MA(1)$.

3.3.2 Estimation de processus AR / MA / ARMA :

Après l'estimation des différents modèles, l'analyse de la significativité des coefficients nous conduisons à conserver trois modèles. Seuls les modèles dont les résidus sont de bruit blancs seront validés.

$$MA(1) : dldiabet_t = \varepsilon_t + 0.817429\varepsilon_{t-1}.$$

$$AR(1) : dldiabet_t = -0.394338dldiabet_{t-1} + \varepsilon_t.$$

$$AR(2) : dldiabet_t = -0.544337dldiabet_{t-1} - 0.362682dldiabet_{t-2} + \varepsilon_t.$$

Et pour pouvoir choisir un bon modèle parmi ceux présents, on construit un tableau de comparaison de leurs critères d'information et nous retenons le modèle qui minimise ces critères.

Tableau de comparaison :

Modèle	AKAIKE	SCHWARZ
MA(1)	0.761664	0.832089
AR(1)	1.027049	1.097474
AR(2)	0.929617	1.035255

TAB. 3.2 – Critères d’information pour les trois modèles

D’après le tableau nous pouvons retenir le modèle $MA(1)$ parce que c’est le modèle qui minimise les critères d’information de AKAIKE et SCHWARZ par rapport aux autres modèles. Alors notre modèle est :

$$dldiabet_t = \varepsilon_t + 0.817429\varepsilon_{t-1}.$$

3.4 Validation du modèle :

Après avoir estimé le modèle $MA(1)$, il convient à présent de valider ce modèle en servant d’une part des tests suivants :

3.4.1 Test de Student des paramètres (significativité statistique) :

Ce test consiste à vérifier que les paramètres du modèle qui ont été estimés sont statistiquement différents de 0.

Les hypothèses du test sont :

$$\begin{cases} H_0 : \alpha_1 = 0, \text{ le coefficient est non significatif.} \\ H_1 : \alpha_1 \neq 0, \text{ le coefficient est significatif.} \end{cases}$$

Ainsi, au risque de 5% le paramètre d’un modèle est statistiquement différent à Zéro car $|t_stat| = 8.735454 > 2$.

3.4.2 Test de normalité pour les résidus du modèle MA(1)

Pour tester la normalité des résidus du modèle $MA(1)$ on utilise la statistique Jarque-Bera, on remarque que $BJ < \chi_{1-\alpha}^2(2)$ ($1.026646 < 5.991$) avec $\alpha = 0.05$ donc accepte l’hypothèse \mathbf{H}_0 de normalité des résidus au seuil α du modèle $MA(1)$.

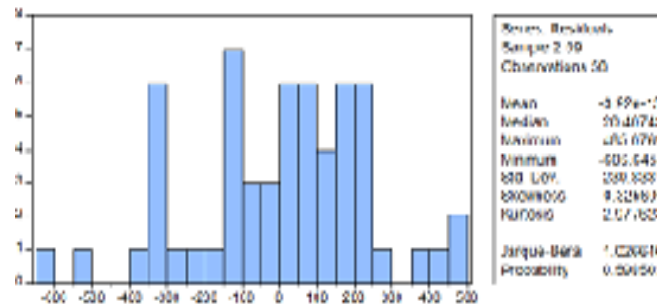


FIG. 3.4 – Normalité des résidus du modèle MA(1).

3.5 Prédiction :

sous gretl , nous obtenons les prévisions suivantes sur 12 mois .

Mois(année 2018)	Prévisions	IC(0.95)2
Janvier	842	[518; 1166]
Février	849	[525; 1173]
Mars	857	[532; 1173]
Avril	864	[540; 1189]
Mai	872	[548; 1196]
Juin	880	[556; 1204]
Juillet	887	[563; 1211]
Aout	895	[571; 1219]
Septembre	902	[578; 12227]
Octobre	910	[586; 1234]
Novembre	917	[593; 1242]
Décembre	925	[601; 1249]

TAB. 3.3 – Prévision

Résumé

L'analyse des séries chronologiques est basée sur la méthodologie Box Jenkins qui été utilisée dans la recherche scientifique ces deriers temps, Cette méthode est très utilisée dans la recherche scientifique ces derniers temps. De ce fait, l'objet de cette méthode est de modéliser une série temporelle en fonction de ses valeurs passées, en utilisant des modèles ARMA et ARIMA.

Mots clés: ARIMA, ARMA, série temporelle, méthodologie de Box Jenkins.

Abstract

Time series analysis is based on the Box & Jenkins methodology that has been used in scientific research in recent times . This methodology has been widely used in scientific research in recent times. Therefore, the purpose of this method is to model a time series based on its past values, using ARMA and ARIMA models.

Keywords:ARIMA, ARMA, time series, Box & Jenkins methodology.

المخلص

تحليل السلاسل الزمنية يعتمد على منهجية بوكس جنكينز في هذه المذكرة ، وقد استخدمت هذه الطريقة في البحث العلمي في الآونة الأخيرة . وبالتالي فإن الغرض من هذه الطريقة هو نمذجة سلسلة زمنية استنادًا إلى قيمها السابقة باستخدام نماذج ARMA و ARIMA.

الكلمات المفتاحية: السلاسل الزمنية, التنبؤ, منهجية بوكس و جنكينز .