

LA MÉTHODE BOX JENKINS

SOMMAIRE

INTRODUCTION

PARTIE I : LA THEORIE.

PARTIE II : LE PROCESSUS (AR) ET LE
PROCESSUS
(MA).

PARTIE III : ARMA ET ARIMA.

PARTIE IV : APPLICATION

INTRODUCTION

La méthodologie de BOX JENKINS pour analyser une série chronologique représente la réponse statistique à ce problème. Il s'agit de choisir dans la vaste classe des modèles ARIMA le modèle reproduisant au mieux la série étudiée. La panoplie statistique habituelle peut s'appliquer : estimation des paramètres du modèle, tests d'hypothèse, analyse des résidus, identification d'observations atypiques, prévision.

Lorsque les données ont une structure probabiliste suffisamment stable au cours du temps et sont assez nombreuses pour permettre une estimation de cette structure, l'approche BOX JENKINS permet d'obtenir les prévisions les plus précises.

La théorie sous-jacente à la méthodologie de BOX JENKINS est complexe, elle est cependant indispensable à une bonne utilisation des logiciels. Il faut avoir en mémoire les propriétés des modèles ARIMA pour pouvoir choisir le modèle adapté aux données. Nous présentons donc les résultats théoriques essentiels dans une première partie.

L'objectif de cet exposé est de permettre une bonne utilisation des logiciels d'analyse d'une série chronologique à l'aide de la méthode de BOX JENKINS.

PARTIE : LA THEORIE

Nous allons exposer dans cette partie les principaux résultats théorique indispensables à l'utilisation de la méthode de BOX JENKINS. La plupart des résultats seront donnés sans démonstration, et nous renvoyons implicitement aux références citées pour démonstrations et approfondissements.

Chapitre I : Les concepts de base

Les modèles ARIMA reposent sur un certain nombre de concepts de base que nous allons présenter.

1.1 Processus aléatoire stationnaire

Un processus aléatoire est une suite de variables aléatoires (z_t) indicées par le temps. Nous limitons dans cet exposé au cas d'un temps discret $t = \dots, -j, \dots, -1, 0, 1, \dots, j, \dots$. Le processus (z_t) est stationnaire lorsque sa structure probabiliste est stable au cours du temps : La loi de probabilité de tout n-uple (z_{t1}, \dots, z_{tn}) est la même que celle du n-uple $(z_{t1+K}, \dots, z_{tn+K})$ pour tout décalage K . Ainsi, pour un processus stationnaire (z_t) , tous ont même moyenne, même variance, et les autocorrélations sont indépendantes de l'instant t pour tout décalage K .

1.2 Processus gaussien stationnaire

Un processus est gaussien lorsque la loi de probabilité de tout n-uple (z_{t1}, \dots, z_{tn}) est une loi multinormale. Pour un processus gaussien la stationnarité au niveau des moyennes, les variances et des autocorrélations suffit à assurer la stationnarité du processus.

1.3 Bruit blanc

Un bruit (a_t) est un processus stationnaire tel que :

$$E(a_t) = 0, V(a_t) = \delta^2, \rho_k = \text{cor}(a_t, a_{t-k}) = 0 \text{ pour tout } K > 0$$



Cette notion de bruit correspond aux hypothèses faites sur les résidus en régression multiple. Les variables aléatoires at sont aussi appelées chocs aléatoires. On suppose implicitement que les chocs at suivent une loi normale $N(0, \delta)$.

1.4 Les données

On considère que la série étudiée, notée aussi (z_t, \dots, z_n) , est une réalisation particulière d'une portion d'un processus aléatoire (z_t) . Les séries économiques sont rarement stationnaires. Elles peuvent présenter une tendance et des variations saisonnières. L'amplitude des variations de la série peut aussi dépendre de son niveau. Il est cependant souvent possible de stationnariser la série à l'aide de transformations permettant d'éliminer tendance et saisonnalité, et de stabiliser la variance.

- **Une série chronologique assez longue ($n \geq 50$).**
- **Exemple : Ventes d'anti-inflammatoires en France de janvier 1978 à juillet 1982.**
- **Objectif : Prévoir les ventes d'août à décembre 1982.**

Marché total des anti-inflammatoires

date	ventes	date	ventes	date	ventes
JAN 1978	3 741	JAN 1980	4 687	JAN 1982	4 764
FEB 1978	3 608	FEB 1980	4 704	FEB 1982	4 726
MAR 1978	3 735	MAR 1980	4 579	MAR 1982	5 080
APR 1978	3 695	APR 1980	4 800	APR 1982	4 952
MAY 1978	3 810	MAY 1980	4 485	MAY 1982	4 633
JUN 1978	3 819	JUN 1980	4 617	JUN 1982	4 830
JUL 1978	3 291	JUL 1980	4 491	JUL 1982	4 460
AUG 1978	3 053	AUG 1980	3 832		
SEP 1978	3 908	SEP 1980	4 669		
OCT 1978	4 035	OCT 1980	5 193		
NOV 1978	3 933	NOV 1980	4 544		
DEC 1978	4 004	DEC 1980	4 676		
JAN 1979	3 961	JAN 1981	4 709		
FEB 1979	4 025	FEB 1981	4 705		
MAR 1979	4 336	MAR 1981	4 677		
APR 1979	4 335	APR 1981	4 627		
MAY 1979	4 412	MAY 1981	4 555		
JUN 1979	4 268	JUN 1981	4 570		
JUL 1979	3 968	JUL 1981	4 457		
AUG 1979	3 505	AUG 1981	3 589		
SEP 1979	4 434	SEP 1981	4 636		
OCT 1979	4 854	OCT 1981	5 077		
NOV 1979	4 592	NOV 1981	4 623		
DEC 1979	4 264	DEC 1981	4 591		

Marché total des anti-inflammatoires



1.5 Les autocorrélations

Pour estimer les autocorrélations ρ_k d'un processus stationnaire (z_t) à partir des données observées (z_1, \dots, z_n) , on utilise la formule :

$$\rho_k = \text{Cor}(w_t, w_{t-k})$$

$$r_k = \frac{\sum_{t=k+1}^N (w_t - \bar{w})(w_{t-k} - \bar{w})}{\sum_{t=1}^N (w_t - \bar{w})^2} = \text{estimation de } \rho_k$$

1.6 Autocorrélations partielles



Régression de w_t sur w_{t-1}, \dots, w_{t-k} :

$$w_t = \varphi_{0k} + \varphi_{1k} w_{t-1} + \dots + \varphi_{kk} w_{t-k} + \varepsilon_t$$

Autocorrélations partielle d'ordre k : φ_{kk}

C'est une corrélation partielle :

$$\varphi_{kk} = \text{Cor}(w_t, w_{t-k} \mid w_{t-1}, \dots, w_{t-k+1})$$

Calcul pratique de $\hat{\varphi}_{kk}$ estimation de φ_{kk}

$$\varphi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \text{L} & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \text{L} & \rho_{k-3} & \rho_2 \\ & & \text{M} & & \\ \rho_{k-1} & \rho_{k-2} & \text{L} & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \text{L} & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \text{L} & \rho_{k-3} & \rho_{k-2} \\ & & \text{M} & & \\ \rho_{k-1} & \rho_{k-2} & \text{L} & \rho_1 & 1 \end{vmatrix}}$$

Soit :

$$\varphi_{11} = \frac{|\rho_1|}{|1|} = \rho_1$$

$$\varphi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

ETC.....

On obtient les estimations $\hat{\varphi}_{kk}$ des φ_{kk} en remplaçant les ρ_k par r_k

PARTIE II : LE PROCESSUS (AR) ET LE

Chapitre I : Le processus AR

Nous allons présenter une famille de processus aléatoires qui sont censés recouvrir une gamme très large d'évolution possible de séries chronologiques :

Il s'agit dans ce chapitre du processus autorégressif AR et le processus de moyenne mobile MA, et dans le 3ème chapitre on verra les deux autres méthodes ARMA et ARIMA.

1-1 Modèle autorégressif (AR) :

Parmi les procédures courantes qui permettent de modéliser des séries, dont on ne connaît pas les variables explicatives mais dont on peut penser qu'elles suivent des lois temporelles, est celle des modèles autorégressifs.

Un processus autorégressif d'ordre p est celui où l'observation présente y_t est générée par une moyenne pondérée des observations passées jusqu'à la p -ième période sous la forme suivante :

$$AR(1) : y_t = \theta_1 y_{t-1} + \varepsilon_t$$

$$AR(2) : y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + \varepsilon_t$$

...

$$AR(p) : y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + \dots + \theta_p y_{t-p} + \varepsilon_t \quad [E]$$

Où $\theta_1, \theta_2, \dots, \theta_p$ sont des paramètres à estimer positifs ou négatifs, ε_t est un aléa gaussien. En introduisant l'opérateur décalage D , l'équation [E] peut aussi s'écrire comme :

$$(1 - \theta_1 D - \theta_2 D^2 - \dots - \theta_p D^p) y_t = \varepsilon_t$$

1-2 Modèle MA (Moving Average : Moyenne Mobile) :

Dans le processus de moyenne mobile d'ordre q à chaque fois que l'observation y_t est générée par une moyenne pondérée d'aléas jusqu'à la q -ième période.



$$\begin{aligned}
MA(1) : y_t &= \varepsilon_t - \alpha_1 \varepsilon_{t-1} \\
MA(2) : y_t &= \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2} \\
&\dots \\
MA(q) : y_t &= \varepsilon_t - \alpha_1 \varepsilon_{t-1} - \alpha_2 \varepsilon_{t-2} - \dots - \alpha_q \varepsilon_{t-q} \quad [E']
\end{aligned}$$

Où $\alpha_1, \alpha_2, \dots, \alpha_q$ sont des paramètres pouvant être positifs ou négatifs et ε_t est un aléa gaussien.

L'équation [E'] peut aussi s'écrire, tout en prenant en considération l'opérateur décalage D :

$$(1 - \alpha_1 D - \alpha_2 D^2 - \dots - \alpha_q D^q) \varepsilon_t = y_t$$

Dans ce processus, comme dans le modèle autorégressif AR, les aléas sont supposés être engendrés par un processus de type bruit blanc. Le modèle MA peut être interpréter comme étant représentatif d'une série chronologique fluctuant autour de sa moyenne de manière aléatoire, ce qui justifie le terme de moyenne mobile dans la mesure où celle-ci "gomme" le bruit créé par l'aléa.

Il est à noter qu'il y équivalence entre un processus MA(1) et un processus AR d'ordre p infini :

$$MA(1) = AR(\infty)$$

Caractéristiques des corrélogramme

Le corrélogramme simple d'un processus MA(q) est de la forme générale :

$$\rho_k = \frac{\sum_{i=0}^{i=q-k} \alpha_i \alpha_{i+k}}{\sum_{i=0}^{i=q} \alpha_i^2} \text{ Pour } k=0,1,\dots,q \text{ et } \rho_k = 0 \text{ pour } k > q$$

C'est à dire que seuls les q premiers termes du corrélogramme simple sont significativement différents de zéro.

Le corrélogramme partiel est caractérisé par une décroissance géométrique des retards

Exemple : Marché Total Différence régulière/saisonnière : $d = 1, D = 1$
Autocorrélations partielles calculées

Partial Autocorrelations

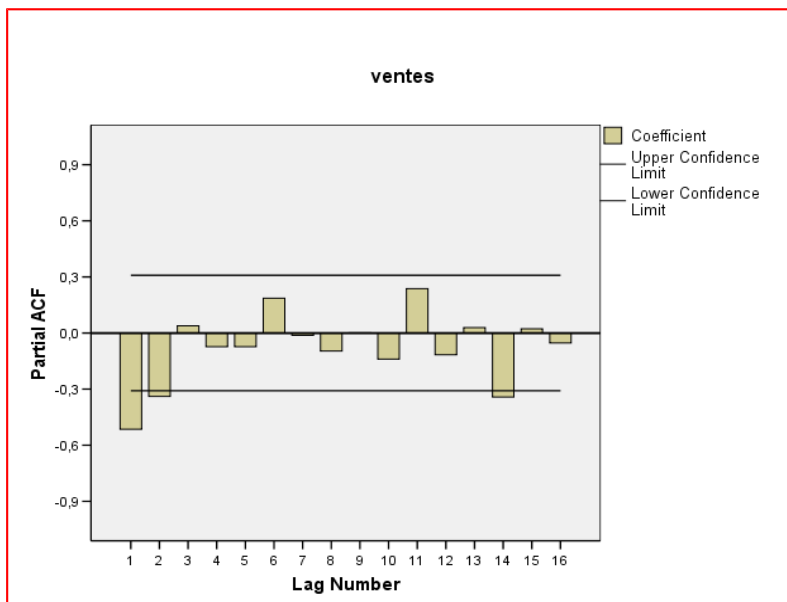
Series: ventes

Lag	Partial Autocorrelation	Std. Error
1	-.515	.154
2	-.339	.154
3	.039	.154
4	-.073	.154
5	-.073	.154
6	.186	.154
7	-.012	.154
8	-.097	.154
9	.001	.154
10	-.139	.154
11	.238	.154
12	-.116	.154
13	.029	.154
14	-.343	.154
15	.022	.154
16	-.053	.154

Rejet de
 $H_0 : \phi_{kk} = 0$
 si:

$$|\hat{\phi}_{kk}| > 2 / \sqrt{N}$$

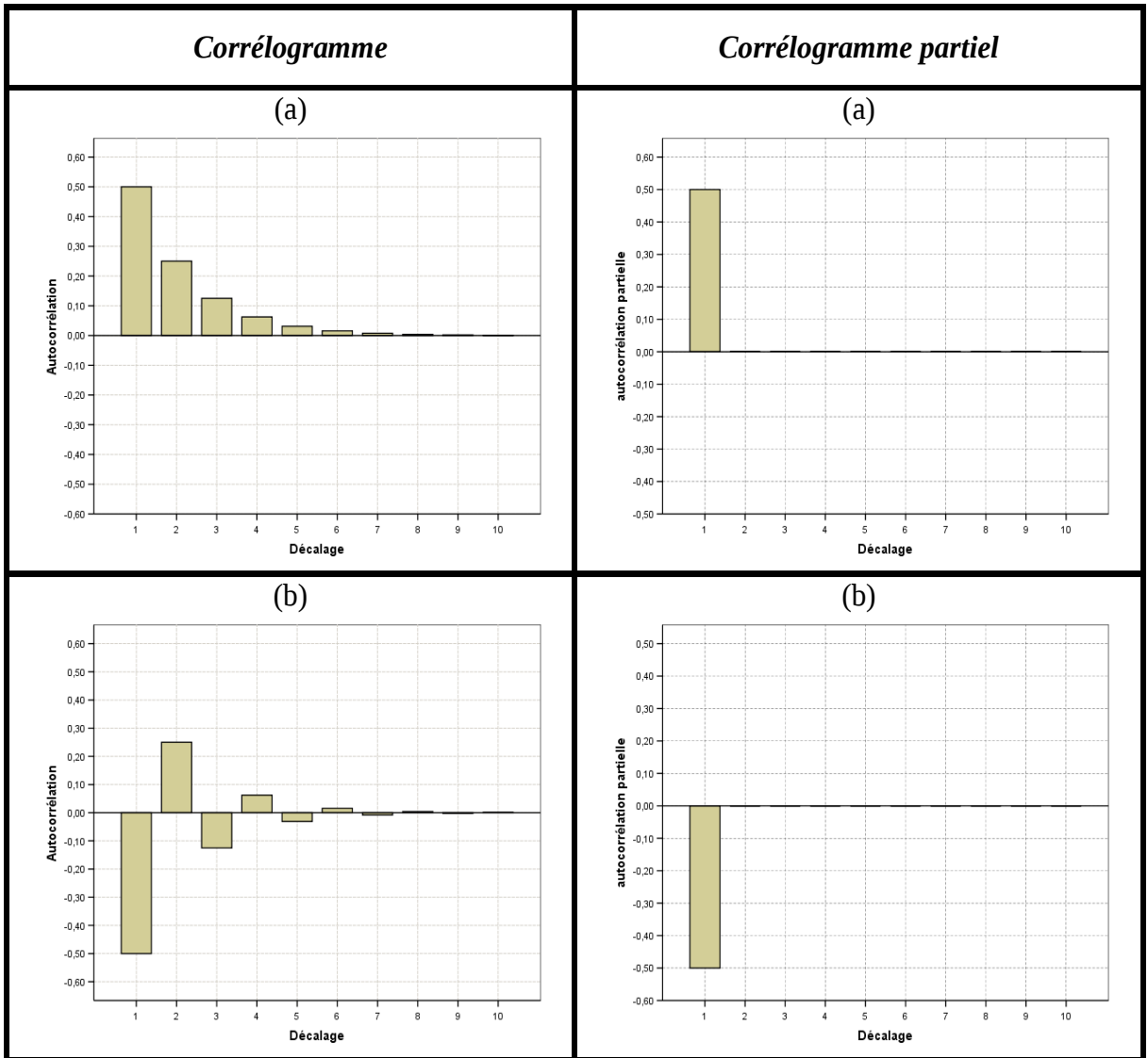
Corrélogramme partiel observé



Autocorrélations et autocorrélations partielles des modèles AR(p) et MA(q)

AR(1)

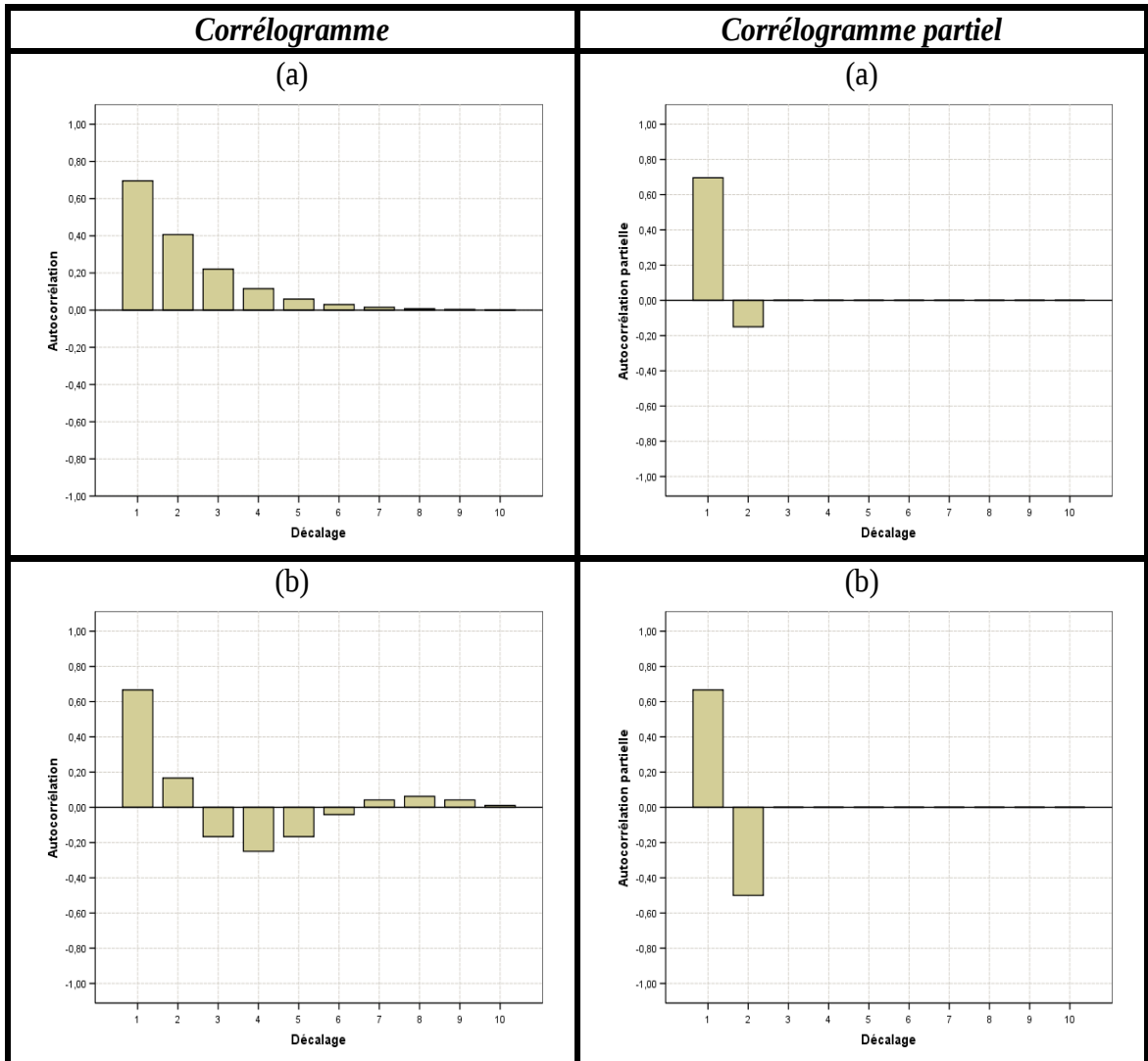
(a): $w_t = 0.5w_{t-1} + a_t$, (b): $w_t = -0.5w_{t-1} + a_t$



AR(2)

(a) : $w_t = .8w_{t-1} - .15w_{t-2} + a_t$

(b) : $w_t = w_{t-1} - .5w_{t-2} + a_t$

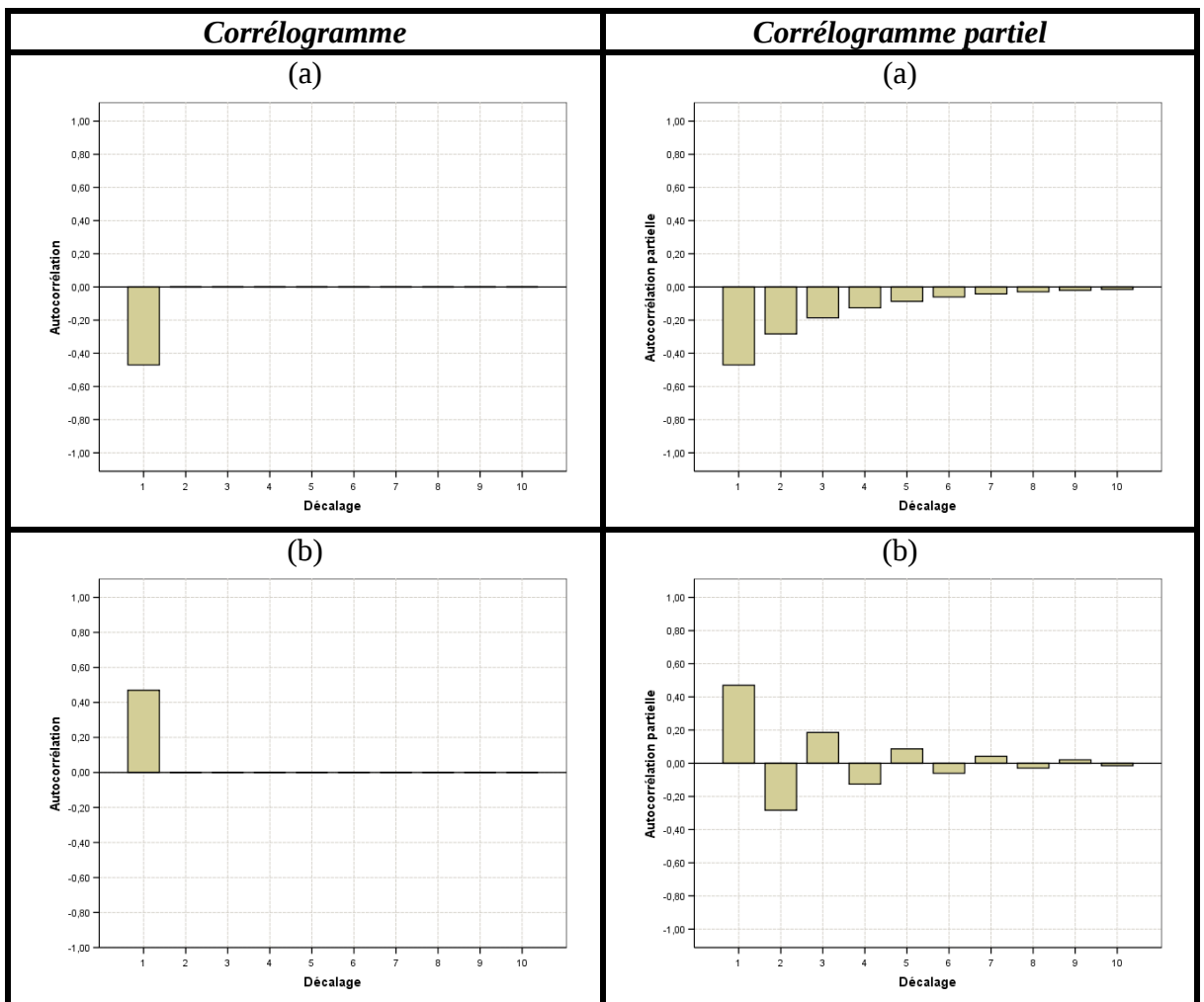


Le dernier pic significatif du corrélogramme partiel donne l'ordre p du modèle $AR(p)$.

MA(1)

(a) : $w_t = a_t - .7a_{t-1}$

(b) : $w_t = a_t + .7a_{t-1}$



MA(q)

(a) : $q = 2$

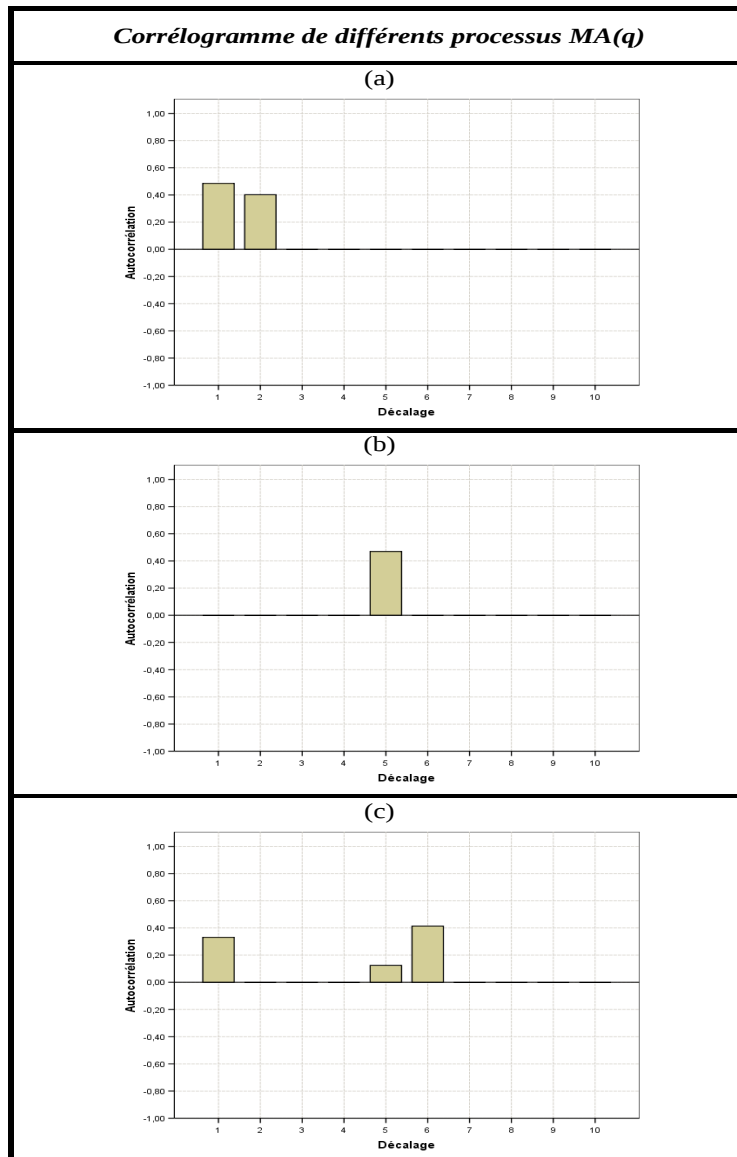
$$w_t = a_t + .5a_{t-1} + .3a_{t-2}$$

(b) : $q = 5$

$$w_t = a_t + .7a_{t-5}$$

(c) : $q = 6$

$$w_t = a_t + .3a_{t-1} + .6a_{t-6}$$



Le dernier pic significatif du corrélogramme donne l'ordre q du modèle MA(q).

PARTIE III : ARMA ET ARIMA.

Maintenant que l'on dispose des tests permettant de diagnostiquer le caractère stationnaire d'une série, on s'interroge sur la façon de représenter les séries stationnaires, c'est-à-dire sur la nature du processus aléatoire sous-jacent. Deux processus sont couramment utilisés : les processus ARMA proposés par Box et Jenkins et les processus à hétéroscédasticité conditionnelle de Engle.

Chapitre I : Les Processus ARMA

Il existe deux types de processus aléatoires permettant de décrire l'évolution des séries stationnaires : les processus autorégressifs d'ordre p notés $AR(p)$ et les processus moyennes mobiles d'ordre q notés $MA(q)$. Ceux-ci peuvent être combinés pour former des processus ARMA (p,q) « Autorégressive Moving Average Process ».

■ Définition

	Processus AR	Processus MA	Processus ARMA
Définition	$X_t = \sum_p \theta_i X_{t-i} + \varepsilon_t$	$X_t = -\sum_q \delta_i \varepsilon_{t-i} + \varepsilon_t$	$X_t = \sum_p \theta_i X_{t-i} + \varepsilon_t - \sum_q \delta_i \varepsilon_{t-i} + \varepsilon_t$
Définition avec opérateur de retard	$\varepsilon_t = (1 - \sum_p \theta_i L^i) x_t$	$x_t = (1 - \sum_q \delta_i L^i) \varepsilon_i$	$(1 - \sum_p \theta_i L^i) x_t = (1 - \sum_q \delta_i L^i) \varepsilon_i$

Dans un processus $AR(p)$ la variable dépend de ses valeurs passées jusqu'au retard p :

$$X_t = \sum_p \theta_i X_{t-i} + \varepsilon_t \quad (1)$$

L'équation (1) peut être réécrite en utilisant l'opérateur de retard L soit L^{i-1} :

$$\varepsilon_t = (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_i L^i - \dots - \theta_p L^p) x_t$$

Dans un processus $MA(q)$ à variable dépend d'une moyenne pondérée des écarts aléatoires passés :

$$X_t = -\sum_q \delta_i \varepsilon_{t-i} + \varepsilon_t$$

En utilisant la notation de l'opérateur de retard, on obtient :

$$x_t = (1 - \delta_1 L - \delta_2 L^2 - \dots - \delta_i L^i - \dots - \delta_q L^q) \varepsilon_t$$

Dans le cas d'une série stationnaire, on doit noter qu'un processus $MA(q)$ peut être inversé et s'écrire sous la forme d'un processus $AR(\infty)$. De la même façon un processus $AR(p)$ peut s'inverser et s'écrire sous d'un $MA(\infty)$.



Dans le processus ARMA, a série stationnaire dépend de ses valeurs passées et d'une moyenne mobile des innovations passées :

$$X_t = \sum_{i=1}^p \theta_i X_{t-i} - \sum_{i=1}^q \varepsilon_{t-i} \delta_i \varepsilon_{t-i} + \varepsilon_t$$

$$(1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_p L^p) x_t = (1 - \delta_1 L - \delta_2 L^2 - \dots - \delta_q L^q) \varepsilon_t$$

▣ Le théorème de la décomposition WOLD

Le théorème permet de décomposer un processus stationnaire en une composante déterministe parfaitement prévisible et une composante stochastique. Ainsi ce théorème permet d'écrire n'importe quel processus stationnaire sous la forme d'un ARMA. Plus précisément pour une valeur de p donnée, une série stationnaire x_t peut s'écrire sous la forme :

$$X_t = \sum_{i=1}^p \theta_i X_{t-i} - \sum_{i=1}^q \varepsilon_{t-i} \delta_i \varepsilon_{t-i} + \varepsilon_t \quad (2)$$

La composante déterministe peut être interprétée comme une espérance mathématique c'est-à-dire la valeur prédite de x_t étant ses valeurs passées.

Il est important de comprendre que les processus AR.MA.ARMA sont des représentations alternatives de la même série. Ainsi un processus ARMA (p,q) peut s'écrire sous la forme AR(∞) ou d'un MA (∞). Il permet donc une représentation plus parcimonieuse des données qu'un AR ou qu'un MA.

▣ Problématique :

Pratiquement on ne peut pas estimer l'équation (2) dans la mesure où le nombre q de retard est infini. BOX et Jenkins proposent une méthode permettant de déterminer p et q et d'estimer le processus.

▣ La méthode BOX et Jenkins :

Cette méthode comporte 4 étapes :

A- La phase de l'identification :

On doit s'assurer que la série ne contient ni racine unitaire, ni composante saisonnière le choix de p et de q repose sur l'observation des corrélogramme, c'est-à-dire de l'ensemble des fonctions d'autocorrélations simple et des fonctions partielle.

▣ Fonction d'autocorrélations simple



On définit Y_t la fonction d'autocorrélations simple à l'ordre k de la manière suivante :

$$Y_t = \frac{\sum_{i=k+1}^T (x_t - \text{xmoy})(X_{t-k} - \text{xmoy})}{\sum_{i=k+1}^T (x_t - \text{xmoy})^2}$$

Dans le cas d'un AR (p), il est possible à partir des coefficients du modèle on peut montrer que les fonctions d'autocorrélations décroissent régulièrement avec k . On montre que les autocorrélations dans un processus MA(q) deviennent nulles lorsque k devient plus grand que q et donc que le corrélogramme des autocorrélations simples est troqué.

▣ Fonction d'autocorrélations partielle

On entend par Fonction d'autocorrélations partielle à 'ordre k , la corrélation entre X_t et X_{t-k} sachant l'influence de $X_{t-1}, \dots, X_{t-k+1}$.

Dans le cas d'un processus AR(p), la fonction d'autocorrélations partielle est troquée à partir de p . dans le cas d'un processus MA (q), les autocorrélations partielles décroissent régulièrement. on en déduit que les fonctions d'autocorrélation simple et partielle ne sont pas tronquées, la série doit être modélisée comme un processus ARMA.

B - La phase d'estimation :

Un processus AR (p) peut être estimé pas MCO alors que l'estimation des processus A (q) ou ARMA (p, q) repose généralement sur des techniques d'estimation fondées sur la maximisation d'une fonction de vraisemblance. Cette méthode exige cependant une hypothèse supplémentaire sur la distribution des écarts aléatoires.

C - La phase des tests :

Les tests portent d'abord sur les coefficients du modèle. On applique le ratio de Student pour tester séquentiellement le caractère significatif des retards. Ce test porte sur le retard le plus élevé. Autrement dit, dans l'hypothèse nulle, le processus est ARMA ($p-1, q$) et dans l'hypothèse alternative il s'agit d'un ARMA (p, q). un test identique peut être réalisé sur l'ordre du processus MA.

Es tests portent ensuite sur les écarts aléatoire. Celui-ci doit être un bruit blanc. on s'en assurera par la batterie usuelle su tests : Normalité, homoscedacité et indépendance sérielle. En cas d'échec, il convient de reformuler le modèle.

D - La phase de prévision :

Sachant que la même série peut avoir plusieurs représentations ARMA, il convient alors de retenir le modèle dont le pouvoir prédictif est le plus élevé. Pour cela, on dispose des critères d'information. Les plus utilisés sont ceux d'Akaike (AIC) et de Schwartz (SC).

$AIC = \ln \left(\frac{SCR_e}{T} \right) + \frac{2(p+q)}{T}$ avec SCR_e : somme des carrés des résidus de l'ARMA (p,q).

$SC = \ln \left(\frac{SCR_e}{T} \right) + \frac{(p+q) \ln(T)}{T}$

Chapitre II : Identification, estimation et tests d'ARIMA :

2-1 : Identification :

Le résultat final après le processus de ARMA est le modèle de d'identification ou modèle autorégressif intégré de moyenne mobile ARIMA (p, d, q) :

P : Ordre de la composante AR

D : Nombre de différences nécessaires à la stationnarité

q : Ordre de la composante MA

L'estimation du modèle ARIMA suppose qu'on travaille sur une série stationnaire.

Cela signifie que la variance et la moyenne de la série sont constantes dans le temps. La meilleure méthode pour éliminer toute tendance est de remplacer la série originale par la série des différences adjacentes.

Après la stationnarisation, nous pouvons identifier les valeurs des paramètres p et q du modèle ARMA.

- Si le corrélogramme simple n'a que ses q premiers termes (q=3maximum) 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).

2-2 : Evaluation des modèles :

La modélisation sert à déterminer combien de paramètres autorégressifs « p » et de moyennes mobiles « q » pour avoir un modèle effectif et qui possède le moins de paramètres et le plus grand nombre de degrés de liberté pour ajuster les données. Le nombre de paramètres p et q ne dépasse pas souvent 2.



Estimation des paramètres : les méthodes d'estimation diffèrent selon le type de processus diagnostiqué.

Dans le cas d'un modèle AR, nous pouvons appliquer une méthode des moindres carrés ou bien nous pouvons utiliser les relations existantes entre les autos corrélations et les coefficients des modèles (équations YULLE-WALKER).

L'estimation des paramètres d'un modèle MA s'avère plus complexe.

Box et Jenkins suggèrent d'utiliser une procédure itérative de type balayage qui peut être illustré de la manière suivante :

Supposons le processus :

$$(1 - \theta_1 D - \theta_2 D^2) y_t = (1 - \alpha_1 D - \alpha_2 D^2) \varepsilon_t$$

Que nous pouvons écrire :

$$y_t = \frac{1}{1 - \theta_1 D - \theta_2 D^2} (1 - \alpha_1 D - \alpha_2 D^2) \varepsilon_t$$

On a que :

$$v_t = \frac{1}{1 - \theta_1 D - \theta_2 D^2} \varepsilon_t$$

Et que :

$$v_t - \theta_1 v_{t-1} - \theta_2 v_{t-2} = \varepsilon_t$$

Ce qui nous donne :

$$y_t = v_t - \alpha_1 v_{t-1} - \alpha_2 v_{t-2}$$

$$D'où \quad v_t = y_t + \alpha_1 v_{t-1} + \alpha_2 v_{t-2}$$

Nous pouvons donc initialiser la procédure de balayage en partant de deux intervalles de valeurs plausibles pour $(\hat{\alpha}_1; \hat{\alpha}_2)$.

Puis pour chaque couple de valeur $(\hat{\alpha}_1; \hat{\alpha}_2)$, nous posons : $\hat{v}_0 = 0$ et $\hat{v}_1 = 0$ et nous calculons les valeurs estimées de \hat{v}_t à partir de la relation $v_t = y_t + \alpha_1 v_{t-1} + \alpha_2 v_{t-2}$

$$\hat{v}_2 = y_2$$

$$\hat{v}_3 = y_3 + \hat{\alpha}_1 \hat{v}_2$$

$$\hat{v}_4 = y_4 + \hat{\alpha}_1 \hat{v}_3 + \hat{\alpha}_2 \hat{v}_2$$

Après le calcul des valeurs de \hat{v}_t , nous estimons les paramètres θ_1 et θ_2 par la méthode des moindres carrés appliqué à l'équation $v_t - \theta_1 v_{t-1} - \theta_2 v_{t-2} = \varepsilon_t$

$$v_t = \theta_1 v_{t-1} + \theta_2 v_{t-2} + \varepsilon_t$$

Nous aurons les valeurs $\hat{\theta}_1$ et $\hat{\theta}_2$, $\hat{\alpha}_1$ et $\hat{\alpha}_2$, ce qui va rendre en minimum la somme des carrés des résidus issus de la régression de l'équation $v_t = \theta_1 v_{t-1} + \theta_2 v_{t-2} + \varepsilon_t$.

On ne peut valider cette méthode d'estimation que si le nombre de paramètre à estimer n'est pas trop important.



2-3 : Tests de validation du modèle Box-Jenkins :

Quand on finit d'estimer les paramètres du modèle, on examine les résultats de cette estimation.

En premier lieu, les coefficients du modèle doivent être significativement différents de 0. Si un coefficient n'est pas significativement différent de 0, il convient d'envisager une nouvelle spécification éliminant l'ordre du modèle AR ou MA non valide.

En deuxième lieu, le coefficient de détermination doit être traité de la même manière que les autres modèles ordinaires.

En troisième lieu, l'analyse des résidus s'effectue à partir de deux critères à respecter :

- Dans le cas où la moyenne n'est pas nulle il convient d'ajouter une constante au modèle :

Soit T le nombre de données disponibles (après avoir enlevé les retards correspondant aux termes AR et MA): Si le processus $(\varepsilon_t, t \in \mathbb{Z})$ est i.i.d. $(0; \sigma_\varepsilon^2)$; on doit avoir:

$$\bar{\varepsilon}_t = \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t \xrightarrow{T \rightarrow \infty} 0$$

- Le résidu est un bruit blanc, les statistiques **Q** et **Q'** de Box-Pierce et de Ljung-Box permettent de tester cette hypothèse.

Si le résidu n'est pas un bruit blanc, la spécification du modèle est incomplète et il manque au moins un ordre à un processus.

La statistique Q et Q' sont donnés comme suit :

$$Q = n \sum_{k=1}^h \hat{\rho}_k^2$$

$$Q' = n(n+2) \sum_{k=1}^h \hat{\rho}_k^2 / (n-k)$$

n : Nombre d'observation

p : Autocréation d'ordre k

h : Nombre de retards

Les statistiques Q et Q' sont distribués asymptotiquement comme une khi-deux à (h-p-q) degré de liberté. Donc nous rejetons l'hypothèse de bruit blanc, au seuil α , si les statistiques Q et Q' sont supérieur à khi-deux lu dans la table $(1-\alpha)$ et (h-p-q) ddl.



Chapitre III : La prévision :

Prolonger les séries au-delà de la période d'échantillonnage c'est l'objectif principal de l'estimation de processus ARMA. De telles projections sont parfois utilisées comme une référence pour des comparaisons avec des prévisions faites sur la base de modèles à plusieurs variables plus compliqués. On trouve inévitablement deux sources d'erreurs dans le futur et la prévision à savoir:

- Les erreurs dues à la méconnaissance des innovations futures ;
- Les erreurs dues aux différences entre les valeurs estimées et les vraies valeurs des paramètres

Ce qui sera intéressant dans cette section c'est la première source d'erreurs qu'on va traiter.

3-1-Prévision du processus AR (1) :

$$y_t - \mu = \alpha(y_{t-1} - \mu) + \varepsilon_t \quad |\alpha| < 1$$

Dans lequel les ε_t sont *i.i.d* $(0, \sigma^2)$ et la forme la plus pratique s'écrit comme suit :

$$y_t = (1 - \alpha) \mu + \alpha y_{t-1} + \varepsilon_t$$

Dans tout ce qui suit, nous supposons qu'on dispose de valeurs observées de y pour la période allant de 1 à n , et que toutes les prévisions sont faites conditionnellement à l'information disponible à l'instant n . Ainsi que :

Y_{n+s} = valeur inconnu de y à la période futur $n+s$

\hat{Y}_{n+s} = prévision de la valeur y_{n+s} faite sur la base de l'information disponible à n

$e_{n+s} = y_{n+s} - \hat{y}_{n+s}$ = erreur de prévision

Considérons la prévision de y_{n+1} pour le processus AR (1), la vraie valeur est donnée par

$$y_{n+1} = (1 - \alpha) \mu + \alpha y_n + \varepsilon_{t+1}$$

Estimée par

$$\hat{Y}_{n+1} = E(y_{n+1} / y_n) = (1 - \alpha) \mu + \alpha y_n$$

D'où

$$\hat{Y}_{n+1} - \mu = \alpha(y_n - \mu)$$

De la même façon on peut trouver la prévision de la période $n+s$

$$\hat{y}_{n+s} - \mu = \alpha^s (y_n - \mu)$$

\hat{Y}_{n+s} tend vers μ quand s tend vers l'infini



3-2 : Prédiction du processus MA (1) :

Un processus MA (1) est défini par : $y_t = \mu + \varepsilon_t - \beta\varepsilon_{t-1}$

En faisant à nouveau les prévisions à partir de la période n, on trouve :

$$\hat{y}_{n+1} = \mu - \varepsilon_n$$

Car ε_{t+1} est inconnu à la période n donc on la pose égale à l'espérance en zéro de manière à débiter le processus. Il est clair que l'importance de cette approximation décroît avec la taille de l'échantillon. On a manifestement $\text{var}(\varepsilon_{n+1}) = \sigma^2$. On se projetant de deux périodes dans le

Futur, on voit que : $y_{n+2} = \mu + \varepsilon_{n+2} - \beta\varepsilon_{n+1}$

Et la prévision pour l'espérance

$$\hat{Y}_{n+2} = \mu$$

Donc, pour le processus MA (1)

$$\text{Et } \hat{y}_{n+s} = \mu \quad \text{Var}(\varepsilon_{n+s}) = (1 + \beta^2) \sigma^2 = \sigma^2 y \quad s \geq 2$$

Processus ARMA (1,1)

Un processus ARMA (1,1) s'écrit comme : $Y_n - \mu = \alpha(y_{n-1} - \mu) + \varepsilon_n - \beta\varepsilon_{n-1}$

La prévision de l'ECE minimum pour la période n+1 est alors :

$$\hat{Y}_{n+1} - \mu = \alpha(y_n - \mu) - \beta\varepsilon_n$$

La seule différence avec la prévision de modèle AR (1) réside dans le terme ($\beta\varepsilon_n$). La variance de l'erreur de prévision est $\text{var}(\varepsilon_{n+1}) = \sigma^2$

La prévision pour la période n+2 est donnée par $\hat{Y}_{n+2} - \mu = \alpha^2(y_n - \mu) - \alpha\beta\varepsilon_n = \alpha(\hat{y}_{n+1} - \mu)$

En continuant de cette manière on peut montrer que :

$$\hat{y}_{n+s} - \mu = \alpha^s (y_n - \mu) - \alpha^{s-1} \beta \varepsilon_n$$

La prévision tend donc vers l'espérance inconditionnelle quand de la prévision s'accroît. De même on peut montrer que :

$$\text{Var}(\varepsilon_{n+s}) \rightarrow \frac{(1 - 2\alpha\beta + \beta^2) \sigma^2}{(1 - \alpha^2)} \text{ quand } s \rightarrow \infty$$



Dans la pratique le modèle BOX-JENKINS est bien plus complexe. il incorpore également un trend et une composante saisonnière, et il est connu sous le nom SARIMA il peut être utilisé à des fins prévisionnelles une fois que les paramètres sont estimés.

3-3 : La Prédiction des modèles ARMA (p,q) :

Le modèle ARMA permet d'écrire $Y_t = \Phi^{-1}(B) \Theta(B)\epsilon_t$

En posant $\Phi^{-1}(B) \Theta(B) = \Psi(B)$ on aura :

$$Y_t = \Psi(B) \epsilon_t$$

$$Y_t = \sum_{j=1}^T \Psi_j \epsilon_{t-j}$$

On a transformé le modèle ARMA en une moyenne mobile donc

$$Y_{t+L} = \Psi_0 \epsilon_{t+L} + \Psi_1 \epsilon_{t+L-1} + \dots + \Psi_L \epsilon_t + \Psi_{L+1} \epsilon_{t-1} + \dots + \Psi_{L+t-1} \epsilon_1$$

De même :

$$Y_{t+L} = \Psi_0 \epsilon_{t+L} + \Psi_1 \epsilon_{t+L-1} + \dots + \Psi_{L-1} \epsilon_{t+1} + \sum_{j=1}^T \Psi_{L+j} \epsilon_{t-j}$$

La somme $\sum_{j=1}^{L+j} \Psi_{L+j} \epsilon_{t-j}$ contient l'information passée jusqu'à la date t incluse. Il va de soi que la prévision $\hat{Y}_t(L)$ ne peut s'appuyer que sur l'information disponible à la date t :

$$\hat{Y}_t(L) = \sum_{j=1}^T \Psi_{L+j} \epsilon_{t-j}$$

La façon à opérer est la suivante : on calcule la prévision à l'horizon L=1. Puis on s'appuie sur cette prévision pour effectuer des prévisions à l'horizon de deux périodes L=2 et ainsi de suite jusqu'à atteindre l'horizon finale prévision.

On a :

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} + \zeta$$

Calculons la

$$Y_{t+1} = \phi_1 Y_t + \dots + \phi_p Y_{t-p+1} + \epsilon_{t+1} - \theta_1 \epsilon_t - \dots - \theta_q \epsilon_{t-q+1} + \zeta$$

Prenons l'espérance conditionnelle pour obtenir la prévision donc,

$$\hat{Y}_t(1) = E(y_{t+1} / y_t ; y_{t-1} \dots) = \phi_1 Y_t + \dots + \phi_p Y_{t-p+1} - \theta_1 \epsilon_t - \dots - \theta_q \epsilon_{t-q+1}$$



En s'appuyant sur $\hat{Y}_t(1)$; $\hat{Y}_t(2)$...Et ainsi de suite jusqu'à atteindre l'horizon final L pour lequel la prévision s'écrit

$$\hat{Y}_t(L) = \phi_1 \hat{Y}_t(L-1) + \dots + \phi_L Y_t + \dots + \phi_p Y_{t-p+1} - \theta_L \varepsilon_t \dots - \theta_q \varepsilon_{t-q+L}$$

PARTIE IV : Application

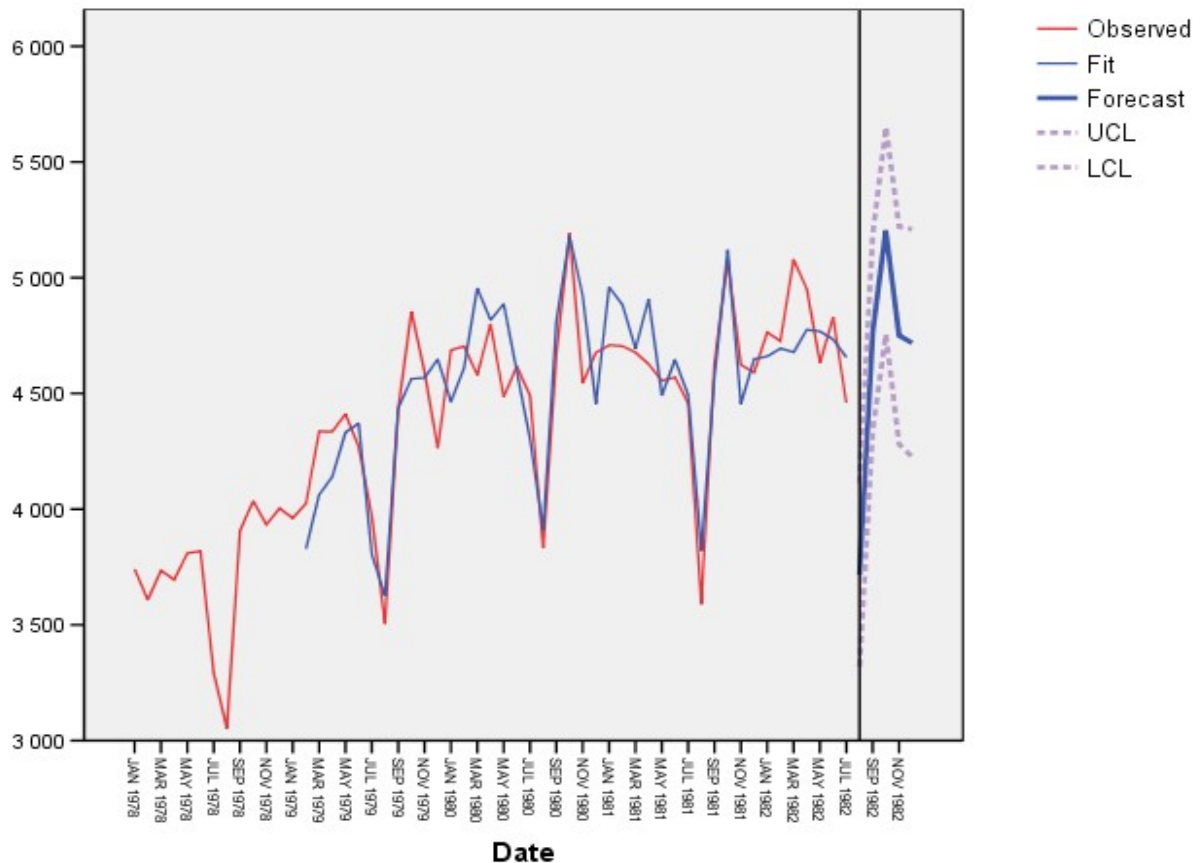
	DATE.	Fit for ventes	95% LCL	95% UCL	SE of Fit
1	AUG 1982	3716.13	3319.22	4113.04	196.53
2	SEP 1982	4763.13	4340.43	5185.82	209.30
3	OCT 1982	5204.13	4757.13	5651.12	221.34
4	NOV 1982	4750.13	4280.09	5220.17	232.75
5	DEC 1982	4718.13	4226.12	5210.14	243.62

Intervalle de prévision à 95% de z_{55+h}

Chaque modèle a sa propre formule de construction de l'intervalle de prévision.

Modèle MA(1) :

$$\hat{z}_{55+h} \pm t_{.975} (N - r) \hat{\sigma} \sqrt{1 + (h - 1)(1 - \hat{\theta})^2}$$



Amélioration du modèle MA(1)

$$r_{12}(\hat{a}_t) = -.281 \quad \text{est significatif}$$

- On suppose maintenant le modèle

$$w_t = \delta + b_t - \theta b_{t-1}$$

$$b_t = a_t - \Theta a_{t-12}, \quad \text{où } a_t \text{ bruit blanc}$$

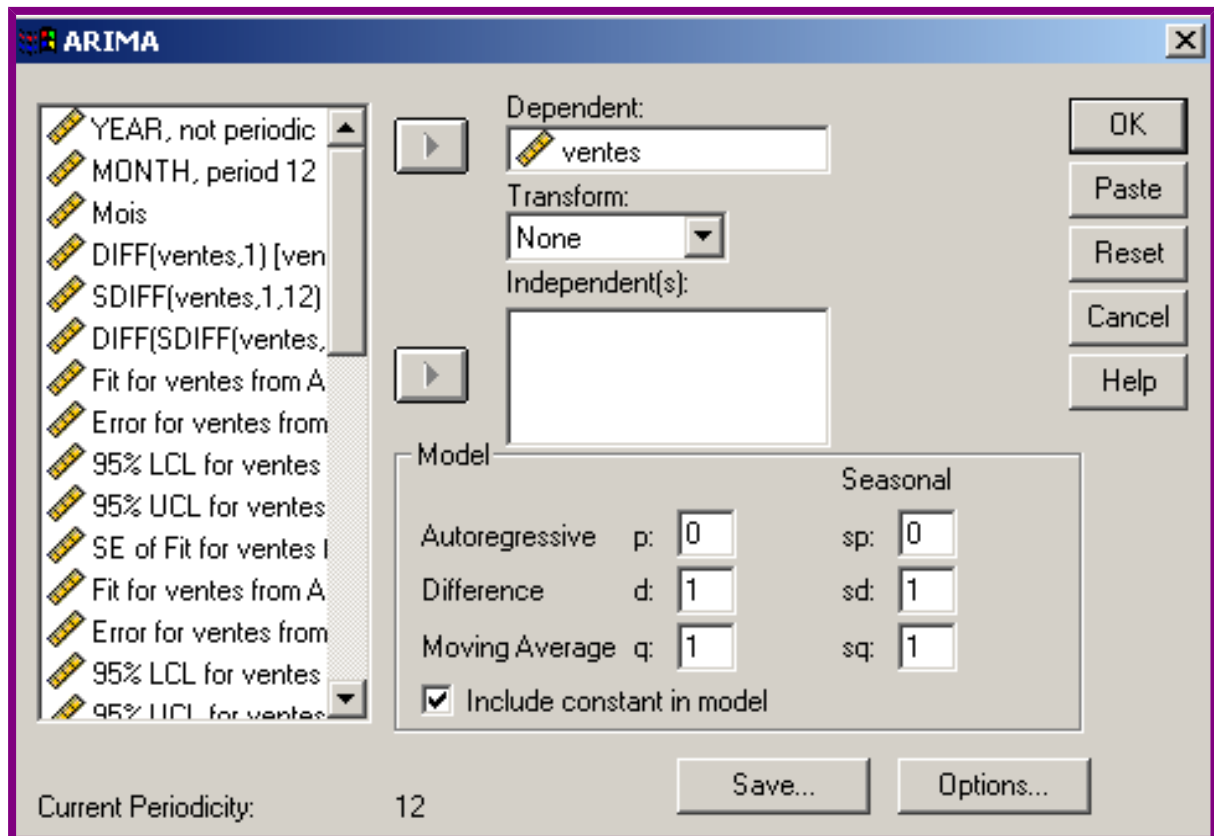
- De

$$w_t = \delta + (1 - \theta B)b_t \quad \text{et} \quad b_t = (1 - \Theta B^{12})a_t$$

On déduit :

$$w_t = \delta + (1 - \theta B)(1 - \Theta B^{12})a_t$$

Demande SPSS



Résultats

$$w_t = \delta + (1 - \theta B)(1 - \Theta B^{12})a_t$$

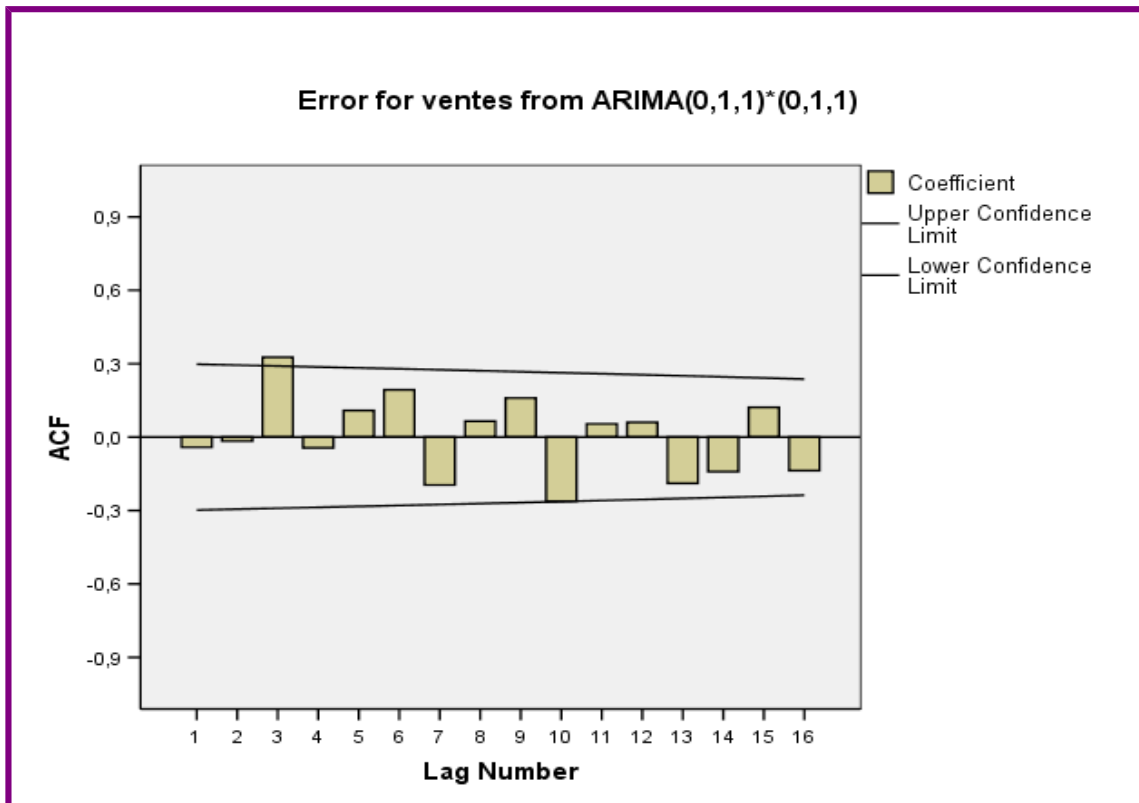
Residual Diagnostics

Number of Residuals	42
Number of Parameters	2
Residual df	39
Adjusted Residual Sum of Squares	1268226.611
Residual Sum of Squares	1336414.106
Residual Variance	25544.245
Model Std. Error	159.826
Log-Likelihood	-276.531
Akaike's Information Criterion (AIC)	559.062
Schwarz's Bayesian Criterion (BIC)	564.275

Parameter Estimates

	Non-Seasonal Lags	Seasonal Lags	Constant
	MA1	Seasonal MA1	
Estimates	.715	.765	-11.468
Std Error	.107	.399	5.219
t	6.693	1.918	-2.197
Approx Sig	.000	.062	.034

Melard's algorithm was used for estimation.



Étude de la voie autorégressive

On suppose que w_t suit un modèle AR(14) :

$$w_t = \delta + \varphi_1 w_{t-1} + \dots + \varphi_{14} w_{t-14} + a_t$$

$$\text{Var}(a_t) = \sigma^2$$

et on a $\delta = \mu \times (1 - \varphi_1 - \dots - \varphi_{14})$.

On choisit les paramètres μ , $\varphi_1, \dots, \varphi_{14}$ et σ^2 à l'aide de la méthode du maximum de vraisemblance.

Résultats

$$w_t = \delta + \varphi_1 w_{t-1} + \dots + \varphi_{14} w_{t-14} + a_t$$

Residual Diagnostics

Number of Residuals	42
Number of Parameters	14
Residual df	27
Adjusted Residual Sum of Squares	949178.0
Residual Sum of Squares	1041062
Residual Variance	28699.741
Model Std. Error	169.410
Log-Likelihood	-270.689
Akaike's Information Criterion (AIC)	571.379
Schwarz's Bayesian Criterion (BIC)	597.444

Parameter Estimates

		Estimates	Std Error	t	Approx Sig
Non-Seasonal Lags	AR1	-.680	.156	-4.367	.000
	AR2	-.441	.169	-2.614	.014
	AR3	.059	.188	.311	.758
	AR4	.034	.184	.185	.855
	AR5	.107	.191	.560	.580
	AR6	.138	.214	.644	.525
	AR7	-.051	.254	-.200	.843
	AR8	-.016	.240	-.067	.947
	AR9	-.006	.232	-.026	.980
	AR10	-.054	.237	-.228	.821
	AR11	.185	.234	.791	.436
	AR12	-.307	.227	-1.355	.187
	AR13	-.428	.208	-2.059	.049
	AR14	-.572	.156	-3.668	.001
Constant		-10.788	9.983	-1.081	.289

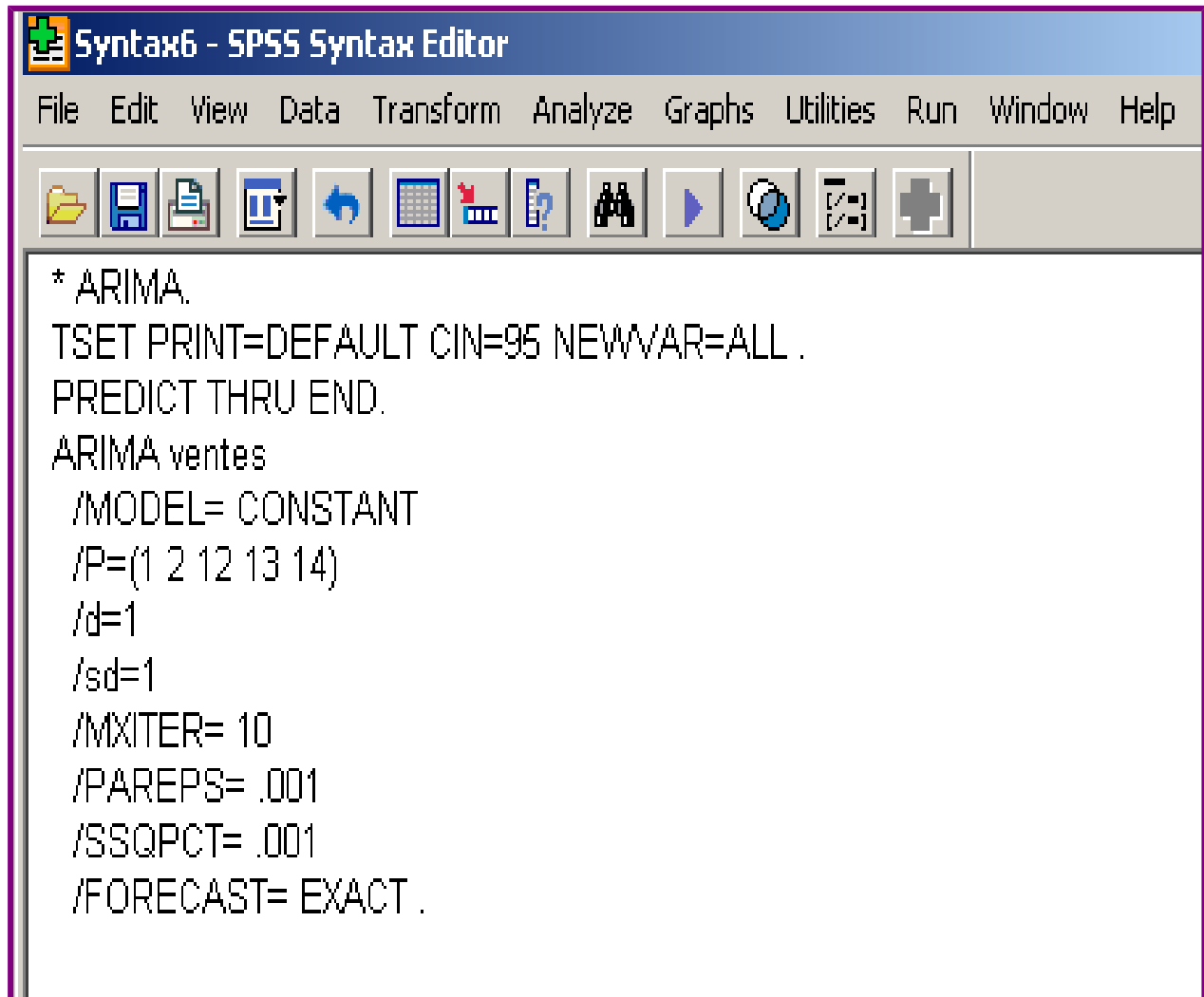
Melard's algorithm was used for estimation.

Modèle AR : p = (1,2,12,13,14) avec constante

$$w_t = \delta + \varphi_1 w_{t-1} + \varphi_2 w_{t-2} + \varphi_{12} w_{t-12} + \varphi_{13} w_{t-13} + \varphi_{14} w_{t-14} + a_t$$



Demande SPSS



Syntax6 - SPSS Syntax Editor

File Edit View Data Transform Analyze Graphs Utilities Run Window Help

```
* ARIMA.  
TSET PRINT=DEFAULT CIN=95 NEWVAR=ALL .  
PREDICT THRU END.  
ARIMA ventes  
  /MODEL= CONSTANT  
  /P=(1 2 12 13 14)  
  /d=1  
  /sd=1  
  /MXITER= 10  
  /PAREPS= .001  
  /SSQPCT= .001  
  /FORECAST= EXACT .
```

Résultats

$$w_t = \delta + \varphi_1 w_{t-1} + \varphi_2 w_{t-2} + \varphi_{12} w_{t-12} + \varphi_{13} w_{t-13} + \varphi_{14} w_{t-14} + a_t$$

Residual Diagnostics

Number of Residuals	42
Number of Parameters	5
Residual df	36
Adjusted Residual Sum of Squares	1093774.600
Residual Sum of Squares	1192109.813
Residual Variance	25711.840
Model Std. Error	160.349
Log-Likelihood	-273.114
Akaike's Information Criterion (AIC)	558.228
Schwarz's Bayesian Criterion (BIC)	568.654

Parameter Estimates

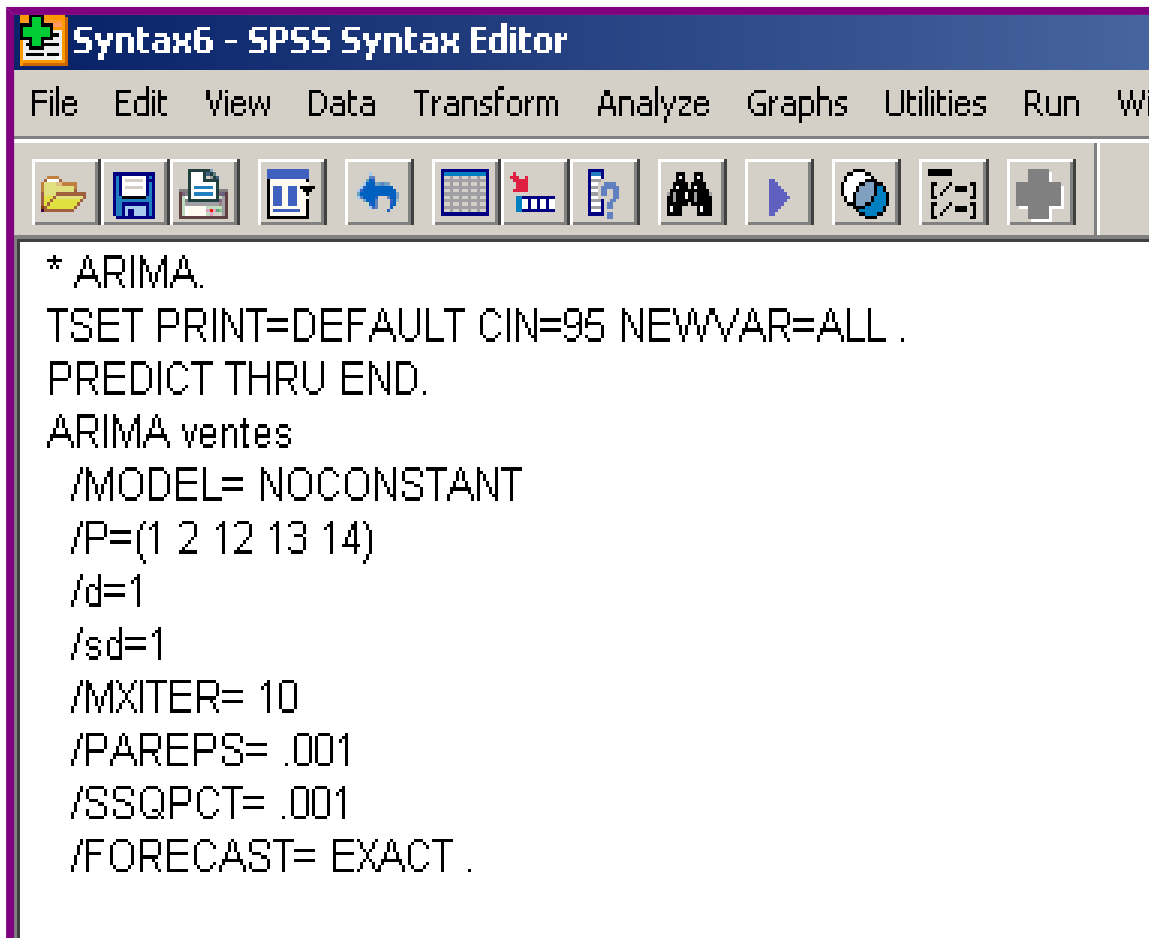
		Estimates	Std Error	t	Approx Sig
Non-Seasonal Lags	AR1	-.775	.127	-6.083	.000
	AR2	-.490	.122	-4.006	.000
	AR12	-.512	.138	-3.711	.001
	AR13	-.594	.159	-3.733	.001
	AR14	-.526	.145	-3.619	.001
Constant		-12.797	7.487	-1.709	.096

Melard's algorithm was used for estimation.

Modèle AR : $p = (1, 2, 12, 13, 14)$ sans constante

$$w_t = \varphi_1 w_{t-1} + \varphi_2 w_{t-2} + \varphi_{12} w_{t-12} + \varphi_{13} w_{t-13} + \varphi_{14} w_{t-14} + a_t$$

Demande SPSS



```
* ARIMA.  
TSET PRINT=DEFAULT CIN=95 NEWVAR=ALL .  
PREDICT THRU END.  
ARIMA ventes  
  /MODEL= NOCONSTANT  
  /P=(1 2 12 13 14)  
  /d=1  
  /sd=1  
  /MXITER= 10  
  /PAREPS= .001  
  /SSQPCT= .001  
  /FORECAST= EXACT.
```

Résultats

$$w_t = \varphi_1 w_{t-1} + \varphi_2 w_{t-2} + \varphi_{12} w_{t-12} + \varphi_{13} w_{t-13} + \varphi_{14} w_{t-14} + a_t$$

Residual Diagnostics

Number of Residuals	42
Number of Parameters	5
Residual df	37
Adjusted Residual Sum of Squares	1172013
Residual Sum of Squares	1233379
Residual Variance	27877.941
Model Std. Error	166.967
Log-Likelihood	-274.563
Akaike's Information Criterion (AIC)	559.127
Schwarz's Bayesian Criterion (BIC)	567.815

Parameter Estimates

		Estimates	Std Error	t	Approx Sig
Non-Seasonal	AR1	-.747	.134	-5.591	.000
Lags	AR2	-.460	.129	-3.568	.001
	AR12	-.454	.148	-3.066	.004
	AR13	-.508	.171	-2.975	.005
	AR14	-.467	.154	-3.041	.004

Melard's algorithm was used for estimation.

Modèle AR : $p = 2$, $P = 1$ avec constante

$$(1 - \varphi_1 B - \varphi_2 B^2)(1 - \Phi B^{12})w_t = \delta + a_t$$

Demande SPSS

ARIMA

Dependent: ventes

Transform: None

Independent(s):

Model

			Seasonal
Autoregressive	p:	2	sp: 1
Difference	d:	1	sd: 1
Moving Average	q:	0	sq: 0

Include constant in model

Current Periodicity: 12

Save... Options...

Résultats

$$(1 - \varphi_1 B - \varphi_2 B^2)(1 - \Phi B^{12})w_t = \delta + a_t$$

Residual Diagnostics

Number of Residuals	42
Number of Parameters	3
Residual df	38
Adjusted Residual Sum of Squares	1196121
Residual Sum of Squares	1286077
Residual Variance	27725.190
Model Std. Error	166.509
Log-Likelihood	-274.998
Akaike's Information Criterion (AIC)	557.997
Schwarz's Bayesian Criterion (BIC)	564.948

Parameter Estimates

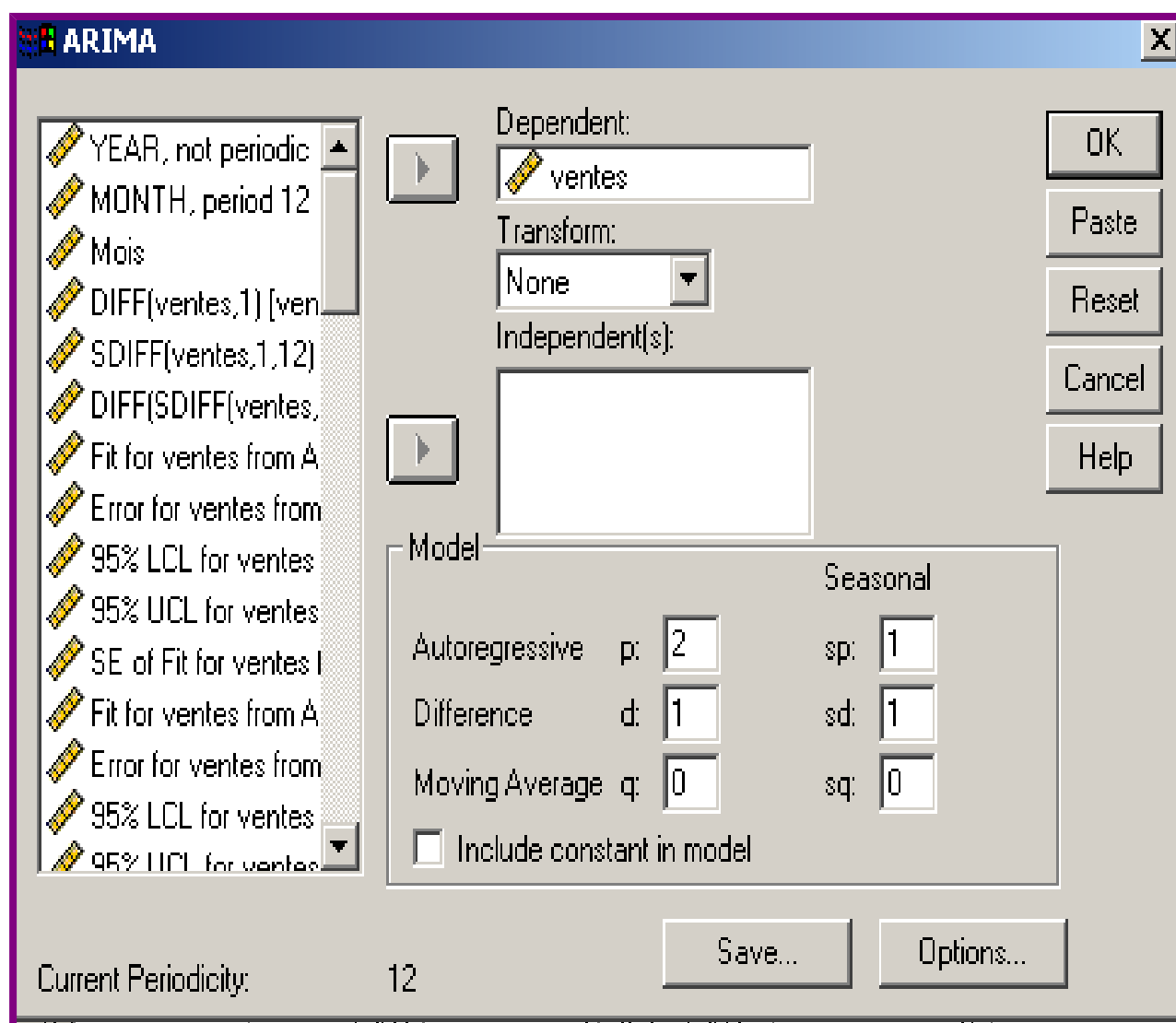
		Estimates	Std Error	t	Approx Sig
Non-Seasonal	AR1	-.759	.139	-5.445	.000
Lags	AR2	-.523	.132	-3.970	.000
Seasonal Lags	Seasonal AR1	-.557	.146	-3.812	.000
Constant		-12.289	8.308	-1.479	.147

Melard's algorithm was used for estimation.

Modèle AR : $p = 2$, $P = 1$ sans constante

$$(1 - \varphi_1 B - \varphi_2 B^2)(1 - \Phi B^{12})w_t = a_t$$

Demande SPSS



Résultats

$$(1 - \varphi_1 B - \varphi_2 B^2)(1 - \Phi B^{12})w_t = a_t$$

Residual Diagnostics

Number of Residuals	42
Number of Parameters	3
Residual df	39
Adjusted Residual Sum of Squares	1256636
Residual Sum of Squares	1315334
Residual Variance	29246.908
Model Std. Error	171.017
Log-Likelihood	-276.033
Akaike's Information Criterion (AIC)	558.066
Schwarz's Bayesian Criterion (BIC)	563.279

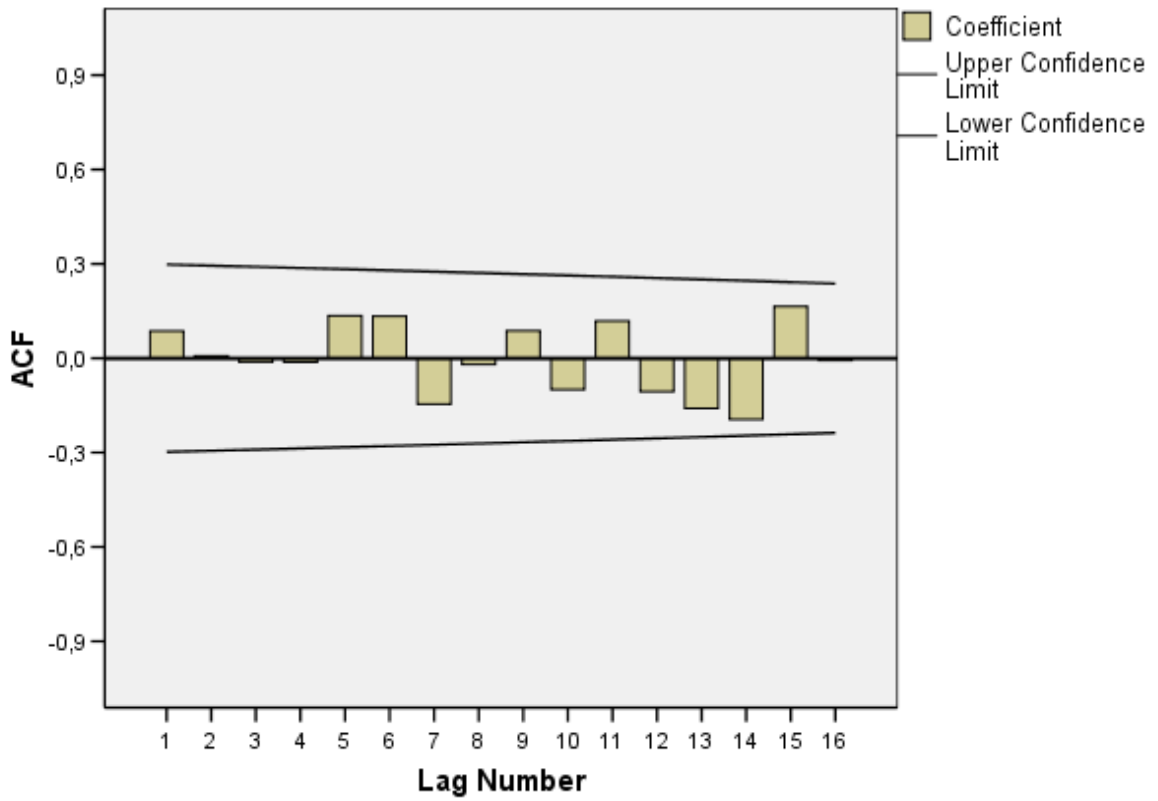
Parameter Estimates

		Estimates	Std Error	t	Approx Sig
Non-Seasonal	AR1	-.731	.143	-5.101	.000
Lags	AR2	-.481	.135	-3.562	.001
Seasonal Lags	Seasonal AR1	-.489	.154	-3.186	.003

Melard's algorithm was used for estimation.



Error for ventes from ARIMA(2,1,0)*(1,1,0) sans cste



Bibliographie

- ❖ Méthode Statistiques En gestion : Michel Tenenhaus, Professeur au Groupe HEC, édition Dunod
- ❖ Régie Bourbonnais : Econométrie manuel et exercices corrigés 4ème édition
- ❖ Econométrie : Claudio Araujo, Jean-François Brun, Jean-Louis Combès
- ❖ Quelques articles- cours d'internet