



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



UNIVERSIT KASDI MERBAH OUARGLA

FACULTE DES MATHMATIQUES ET SCIENCES DE LA MATIERE

DEPARTEMENT DE MATH MATIQUES

MEMOIRE PRESENTE EN VUE DE L'OTENTION DU DIPLOME :

MASTER en Mathématiques

Option : Probabilités Et Statistique

Par

BOUKHRIS Maroua

Thème :

**Prévisions par l'utilisation des modèles SARIMA
(Nombre des naissances à l'hôpital de Sidi Abd-
elkader de Ouargla)**

Membres du Comité d'Examen :

Mr Abbasi.H	Université KASDI Merbah-Ouargla	Président
Mr Agoune.R	Université KASDI Merbah- Ouargla	Examineur
Mm Khorsi.R	Université KASDI Merbah-Ouargla	Encadreur

Année universitaire 2019- 2020

Table des matières

Table des matières	i
Dédicace	i
Remerciements	ii
Notations et abreviations	iii
Liste des tableaux	v
Table des figures	vi
Introduction	1
1 Généralités sur les séries chronologique :	2
1.1 Définition d'une série chronologique	2
1.2 Objectifs de l'analyse d'une chronique	3
1.3 Les composantes d'une série chronologique	4
1.4 Schémas de décomposition d'une série chronologique	5
1.5 Coefficients saisonniers	8
1.6 Série désaisonnalisée	9
1.7 Autocorrélations simple et partielle	9
1.7.1 La fonction d'autocovariance et d'autocorrélation	10
1.7.2 La fonction d'autocorrélation partielle	11
1.8 Bruit blanc	12
1.9 stationnarité	12
1.9.1 Relation entre stationnarité faible et stricte	13
1.9.2 Test de Dickey Fuller Augmenté	14
1.10 Théorème de Wold	14
1.11 Opérateurs définis sur une série chronologique	16
1.11.1 Opérateur de retard	16

1.11.2	Opérateur de différence d'ordre d	17
1.12	Analyse de la tendance	17
1.12.1	Rappels sur la régression linéaire	17
1.12.2	La méthode des moindres carrés	18
1.12.3	Propriétés et interprétation du coefficient de corrélation linéaire	18
1.12.4	Estimation non paramétrique	19
1.13	Les moyennes mobiles	19
1.13.1	Propriétés d'un lissage par moyenne mobile	22
1.13.2	La série lissée par moyenne mobile	23
2	Modélisation <i>ARIMA</i> et <i>SARIMA</i> des séries chronologique	24
2.1	Les modèles autorégressifs (AR)	24
2.1.1	La stationnarité	25
2.1.2	L'inversibilité	26
2.1.3	La causalité	26
2.1.4	Les fonctions auto-covariance, auto-corrélations et équivalence de Yule-Walker :	27
2.1.5	Prédiction dans le modèle AR(p)	31
2.2	Les modèles moyennes mobiles, (MA)	32
2.2.1	stationnarité	32
2.2.2	causalité	33
2.2.3	inversibilité	33
2.2.4	La fonction d'autocorrélation	33
2.2.5	Prédiction dans le modèle MA(q)	34
2.3	Les modèles ARMA(p, q) (Autorégressive Moving Average)	35
2.3.1	Inversibilité et stationnarité en terme de valeurs passées	35
2.3.2	Auto-corrélation	36
2.3.3	Auto-corrélations partielles	37
2.3.4	Prédiction dans modèle ARMA(p,q)	37
2.4	Les modèles ARIMA(p, d, q)	38
2.5	Les modèles ARIMA saisonniers (SARIMA)	39
2.6	Les processus SARIMA	39
2.7	Box-Jenkins pour des modèles saisonniers	41
2.7.1	Identification et ajustement des modèles SARIMA aux données	42
2.7.2	Prévision des processus SARIMA	50

3 Application	52
3.1 Identification de la série	52
3.1.1 Les tests informels	53
3.1.2 Tests formels : Test de racine unitaire (ADF)	55
3.2 Stationnarisation de la série	57
3.3 Identification du modèle optimal provisoire	60
3.4 Estimation	60
3.5 Validation du modèle	62
3.6 La prévision	64
Conclusion générale	65
Bibliographie	66

Dédicace

D'abord, je remercie mon DIEU qui ma donnée le courage et la force pour accomplir ce modeste travail.

Je présente mes dédicaces aux deux êtres les plus chers de ma vie "mon père" et

"ma mère".

En plus de mon mari et de mon fils, je prie Dieu de les protéger pour moi

A mes frères et . A mes Sœurs .

A toutes ma famille

Je veux aussi me consacrer à tous les enseignants qui ont bâti mon niveau d'éducation dès mon plus jeune âge et à tous ceux qui me connaissent.

Remerciements

Je remercie **ALLAH** tout puissant de m'avoir donné la volonté et le courage de mener à bien ce travail.

Que Toutes les formes de Prière et de Salat soient adressées à notre Ame et Conscience Sidna Mohammed, notre lumière dans cette vie.

Je remercie Madame "*Khorsi.R*" pour l'honneur qu'elle a fait en acceptant d'être mon encadreur, pour l'intérêt scientifique qu'elle a porté à ce travail, pour ses précieux conseils tout le long de l'élaboration de ce travail.

J'adresse aussi mes sincères remerciements à "*Mr.Agoun .R*" et "*Mr.Abbasi.H*" qui ont accepté de faire partie de mon jury de prendre de leur temps pour examiner mon travail .

Je remercie tous mes enseignantes et enseignants.

Et enfin à tous nos collègues de la promotion 2019/2020 et à tous ceux qui ont contribué, de près ou de loin, à la réalisation de ce travail.

Notations et abriviations

ADF	Dickey-Fuller Augmenté
AIC (p, q)	Critère d'Akaïke
AR	Autorégressifs, (Autoregresive)
ARMA(p, q)	Autoregresive Moving Average
ARIMA(p,d, q)	Autoregresive Integrated Moving Average
B	Opérateur retard
BB	Bruit blanc
BIC	Critère de Schwarz
DS	Differncy Stationnary
F	Opérateur d'avance
FAC	Fonction d'autocorrélation

FACP	Fonction d'autocorrélation partielle
JB	Test de Jarque & Bira
MA	Moyennes Mobiles (Moving Average)
MCO	Méthode de moindre careés
MM	Moyenne mobile
SARIMA	Saisonnier Autoregressive Integrated Moving Average
TS	Tend Stationnary
VCM	Valeur Critique de MacKinnon

Liste des tableaux

1.1	Moyenne Mobile d'ordre 3	20
1.2	Moyenne Mobile d'ordre 2	21
1.3	Moyenne Mobile d'ordre 4	22
3.1	L'évolution du nombre des naissances à l'hôpital de Sidi Abd- elkader de Ouargla (Janvier 2015 - Aout 2020).	52
3.2	Prévision du nombre de naissances.	64

Table des figures

1.1	La population algérienne de 2000 à 2020.	3
1.2	Les composantes d'une série chronologique.	5
1.3	schéma additif.	6
1.4	schéma multiplicatif.	6
2.1	Schéma général de la modélisation d'une série temporelle par un modèle SARIMA.	51
3.1	Graphique de l'évolution du nombre des naissances à l'hôpital de Sidi Abd-elkader de Ouargla (Janvier 2015 - Aout 2020)	53
3.2	Histogramme et statistiques.	54
3.3	Corrélogramme de la série SARIMA	55
3.4	Test ADF sur premier série.	57
3.5	Test ADF sur la série LOGSARIMA.	58
3.6	Test ADF sur la série DLOGSARIMA.	58
3.7	Graphique de la série stationnaire.	59
3.8	Correlogramme de la série stationnaire.	59
3.9	le meilleur modèle.	60
3.10	Estimer l'équation.	61
3.11	Test de normalité.	62
3.12	Correlogramme Q-Statistics.	63
3.13	Correlogram Squared Residuals.	63

Introduction

Les séries chronologiques, ou temporelles, encore nommées chroniques sont des séries statistiques ordonnées dans le temps que prennent des valeurs numériques. Elles peuvent utiliser n'importe quelle mesure du temps, dates ou périodes, à condition que l'espacement des observations soit constant. On emploie aussi le temps de périodicité malgré son ambiguïté (il est aussi utilisé pour désigner le temps qui s'écoule entre deux périodes identiques). Une série temporelle peut aussi s'observer de façon continue, par exemple en physique. Leur domaine d'application est très vaste et s'étend de l'astronomie à l'économie et finance en passant par la biologie, psychologie, géophysique ou la théorie du signal...etc. L'étude d'un processus aléatoire à partir d'une série chronologique peut avoir généralement les objectifs suivants : Comprendre le passé : expliquer les variations observées, Prédire les valeurs futures (proches), Etudier le lien avec d'autres séries. L'idée est de prendre un échantillon de données et de construire le meilleur modèle qui ajuste ces données. Ce modèle nous permet de tirer certaines conclusions sur la série.

Ce travail est réparti comme suit :

Le premier chapitre est consacré à comprendre les notions de base sur les séries chronologiques et tout ce qui y est lié.

Le deuxième chapitre présente les principaux modèles : AR, MA, ARIMA, et les modèles saisonniers SARIMA avec les étapes méthodologiques de Box & Jenkins. L'étude des séries chronologiques randomisées est mentionnée.

Le troisième chapitre et dernier chapitre est consacré à l'application de modèle SARIMA par la méthode de Box & Jenkins avec le programme Eviews sur le nombre de naissances à l'hôpital de Sidi Abd-elkader de Ouargla (*Janvier 2015 - Août 2020*).

Chapitre 1

Généralités sur les séries chronologique :

Une série chronologique est une suite formée d'observations au cours du temps, L'analyse des série chronologique est un outil couramment utilisé de nos jours pour la prédiction de données futures, il est appliqué dans beaucoup de domaines à savoir en finance en médecine en économétrie et en météorologie et dans bien d'autre domaines.

La prévision se base sur la connaissance du passé et du présent. L'analyse des séries chronologiques, présente plusieurs méthodes entre d'autre le lissage exponentiel et les modèles basés sur les processus aléatoires linéaires (modèles ARIMA saisonnier SARIMA,). dans la suite on présentera les notions de base des séries chronologiques nécessaire à la compréhension de la suite du mémoire.

1.1 Définition d'une série chronologique

Une séries statistique est désignée comme étant chronologique quand les données qui la constituent sont les valeurs d'une variable enregistrée en fonction de la date pendant une certaine période [2] [3] .

Définition 1.1.1 *On appelle série chronologique une suite d'observations numériques d'une grandeur effectuées à intervalles réguliers au cours du temps.*

Définition 1.1.2 *Une série chronologique, ou chronique ou série temporelle, est une suite finie de données indexée par le temps.*

Si t_1, t_2, \dots, t_n sont les n instants successifs d'observation et si X_{t_i} est la valeur mesurée à l'instant t_i , on notera la série chronologique $\{X_t\}_{t \in T}$ où T est l'ensemble ordonné $T = \{t_1, t_2, \dots, t_n\}$.

Exemple 1.1.1 (*de séries chronologiques*)

1. Ecologie (pollution).
 2. Nombre annuel de grèves aux Etats-Unis de 1951 à 1980.
 3. Nombre mensuel de décès accidentelsarx USA 1973 - 1979 .
- Taille de la population algérienne de 2000 à 2020.

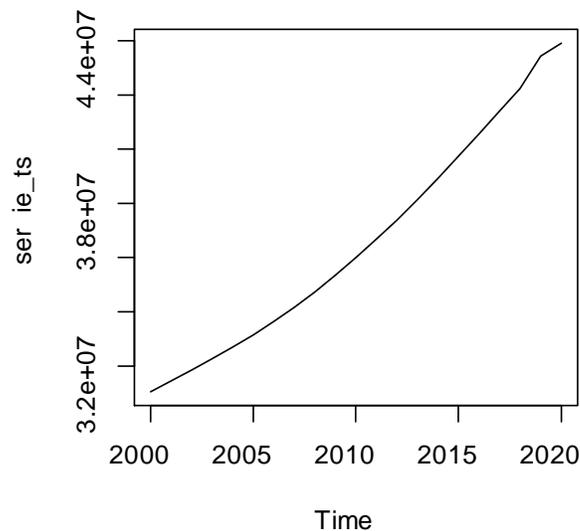


FIG. 1.1 – La population algérienne de 2000 à 2020.

1.2 Objectifs de lanalyse d'une chronique

Les objectifs de l'analyse d'une série chronique sont :

1. Modélisation

Elle consiste à :

-Développer des modèles permettant de décrire le comportement d'une ou plusieurs séries chronologiques.

-Mettre au point une méthodologie pour spécifier, estimer, valider (juger) un modèle approprié pour des données particulières.

2. Prévision

Étant données des observations X_1, \dots, X_T , la prévision consiste à évaluer un valeur non observée, X_{T+h} . La prévision peut être ponctuelle, ou prendre la forme d'un intervalle de prévision.

1.3 Les composantes d'une série chronologique

La première étape dans l'étude des séries chronologiques consiste à représenter graphiquement l'événement qui permet d'entrevoir les quatre composantes fondamentales d'une chronique [4] :

1. **La tendance** f_t : Représente l'évolution à long terme de la grandeur étudiée, et traduit l'aspect général de la série.
2. **Les variations saisonnières** S_t : 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).

Elles sont de nature périodique, c'est-à-dire qu'il existe un entier p , appelé période, tel que $S_t = S_{t+p}$ pour tout $t \geq 1$.

3. **Cycle** C_t : Regroupe les variations autour de la tendance avec des alternances de phases d'expansion et de recession. Ces phases durent généralement plusieurs années, mais n'ont pas de durée fixe.
4. **Les variations accidentelles ou résiduelles** ε_t : les variations accidentelles sont des fluctuations irrégulières et imprévisibles. Elles sont supposées en général de faible amplitude. Elles proviennent de circonstances non prévisibles : catastrophes naturelles, crise boursière, grèves,...

Le graphique suivant présente l'ensemble des composantes précitées

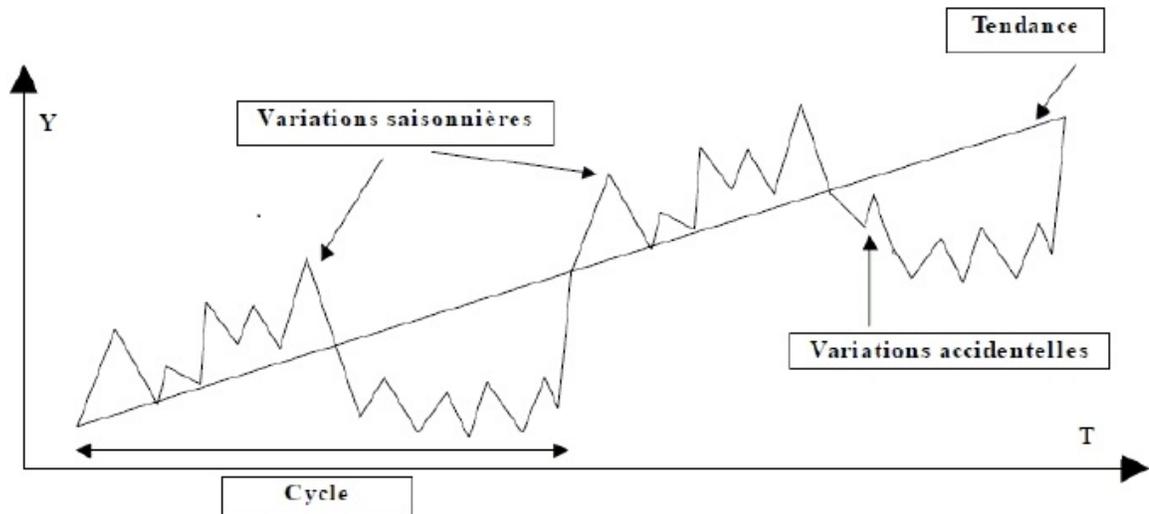


FIG. 1.2 – Les composantes d'une série chronologique.

1.4 Schémas de décomposition d'une série chronologique

La technique de décomposition d'une série chronologique, repose sur un modèle qui l'autorise. Ce modèle porte le nom de schéma de décomposition. Il en existe essentiellement deux grands types [5] :

1. Schéma additif

Dans un modèle additif, on utilise les 4 composantes : tendance, variations saisonnières, composante cyclique et composante résiduelle sont indépendantes les unes des autres. On considère que la série X_t s'écrit comme la somme de ces 4 composantes :

$$X_t = f_t + C_t + S_t + \varepsilon_t.$$

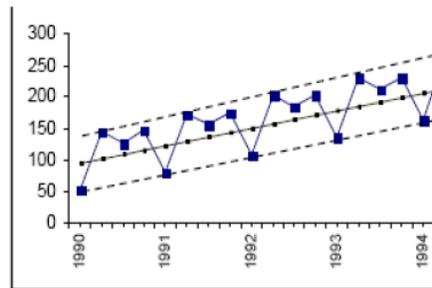


FIG. 1.3 – schéma additif.

2. Schéma multiplicatif

Dans un modèle multiplicatif on utilise aussi les 4 composantes : tendance, variations saisonnières, composante cyclique et composante résiduelle sont indépendantes les unes des autres. On considère que la série X_t s'écrit comme la produit de ces 4 composantes :

$$X_t = f_t \times C_t \times S_t \times \varepsilon_t.$$

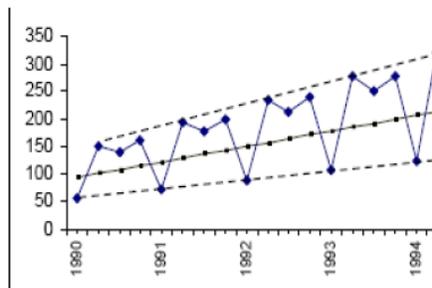


FIG. 1.4 – schéma multiplicatif.

3. Schéma mixte

Dans un modèle mixte on utilise les deux schéma multiplicatif et additif avec les 4 composantes : tendance, variations saisonnières, composante cyclique et composante résiduelle sont indépendantes les unes des autres on considère que la série X_t s'écrit comme la produit et la somme de ces 4 composant :

$$X_t = (f_t + c_t) \times s_t + \varepsilon_t$$

$$X_t = f_t + (c_t \times s_t) + \varepsilon_t$$

.

.

.

Remarque 1.4.1 *La tendance (Z_t) prend différentes forme :*

- a. **Linéaire** : $Z_t = a + bt$.
- b. **Quadratique** : $Z_t = a + bt + ct^2$.
- c. **Exponentielle** : $Z_t = ae^{bt}$.

- **Choix du modèle**

Le choix du modèle est défini par plusieurs méthodes :

Méthode de la bande :

On utilise le graphe de la série et la droite passant par les minima et celle passant par les maxima.

- Si ces 2 droites sont à peu près parallèles : le modèle est additif.
- Si ces 2 droites ne sont pas parallèles : le modèle est multiplicatif.

Méthode du profil :

On utilise le graphique des courbes superposées.

- Si les différentes courbes sont à peu près parallèles : le modèle est additif.
- Sinon (les pics et les creux s'accroissent) : le modèle est multiplicatif.

Méthode du tableau de Buys et Ballot :

On calcule pour chacune des années, la moyenne et l'écart type.

On trace les points d'abscisse la moyenne et d'ordonnée l'écart type de la même année.

On trace la droite des moindres carrés de ces points.

- Si l'écart type est indépendant de la moyenne le modèle est additif.

- La pente (a) de la droite des moindres carrés est très proche de 0.
- Sinon ;
- l'écart type est fonction de la moyenne le modèle est multiplicatif.
- La pente (a) de la droite des moindres carrés n'est pas nulle.

1.5 Coefficients saisonniers

On sait que l'influence des variations saisonnières doit être neutre sur l'année et que ces variations (S_t) se répètent théoriquement à l'identique de période en période. Dans toute série chronologique observée sur un cas réel, les variations saisonnières ne sont jamais identiques. Donc, pour satisfaire aux exigences du modèle théorique, et pour pouvoir étudier la série réelle, il faut estimer, à la place des (S_t) observées, des variations périodiques identiques chaque année (mois par mois, ou trimestre par trimestre) qu'on appelle coefficients saisonniers. On les note S_j , $j = 1$ à 12 pour des données mensuelles. $j = 1$ à 4 pour des données trimestrielles.

•Méthode de calcul des coefficients saisonniers

La série Y_t est observée sur n années par période p . ($p = 12$) mois ($j = 1, 2, \dots, 12$) ou 4 trimestres ($j = 1, 2, 3$ ou 4). Les variations saisonnières S_t sont égales, par hypothèse du modèle **additif** à :

$$S_t = Y_t - f_t,$$

nous obtenons donc $n \times j$ valeurs de S_t , que nous pouvons écrire S_{ij} . On retiendra 12 valeurs de S_j (mois) ou 4 valeurs de S_j (trimestres) comme coefficients saisonniers, en calculant, mois par mois, ou trimestre par trimestre, la moyenne arithmétique des S_t , sur l'ensemble des n années, on obtient :

$$S_j = \frac{1}{n} \sum_{i=1}^n S_{ij}.$$

La somme sur l'année de ces coefficients saisonniers S_j devrait en toute logique être égale à 0. En fait, bien souvent, les approximations des calculs conduisent à un résultat légèrement différent. Dès lors, dans le cas où la somme des S_j est différente de 0, on calcule un coefficient correcteur qui est la moyenne des S_j sur l'année

$$S_j = \frac{1}{12} \sum_{j=1}^{12} S_j.$$

Et l'on retient en définitive, comme coefficient saisonnier corrigé la valeur :

$$S_j^* = S_j - \rho.$$

Le principe théorique selon lequel la moyenne (ou la somme) des coefficients saisonniers est égale à zéro est respectée par les S_j^* (coefficients saisonniers corrigés).

1.6 Série désaisonnalisée

Nous appelons série désaisonnalisée ou série corrigée des variations saisonnières notée série *CVS*, la série chronologique Y_t à laquelle on a enlevé les variations saisonnières.

Dans le cas du modèle additif : La série désaisonnalisée est :

$$Y_t^* = Y_t - S_t.$$

Dans le cas du modèle multiplicatif : La série désaisonnalisée est :

$$Y_t^* = \frac{Y_t}{S_t}.$$

La particularité de la série *CVS* est que les données de Y_t^* sont directement comparables : on a enlevé l'effet des saisons et donc le caractère propre de chaque mois on peut par exemple comparer les données d'un mois de janvier et celle d'un mois de juillet.

1.7 Autocorrélations simple et partielle

Les principales caractéristiques temporelles sont données par l'autocorrélation (simple) et l'autocorrélation partielle .

1.7.1 La fonction d'autocovariance et d'autocorrélation

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

$$\gamma(h) = \text{cov}(X_t, X_{t-h}) = E[(X_t - E(X_t))(X_{t-h} - E(X_{t-h}))]$$

Ainsi

$$\gamma(0) = \text{Var}(X_t) = E[((X_t) - (E(X_t)))^2] = \sigma_X^2.$$

Elle fournit 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'autocorrélation est définie par :

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, h \in \mathbb{Z}.$$

Avec $\rho(0) = 1$ et $|\rho(h)| < 1$.

L'équivalent empirique de la fonction d'autocorrélation, noté $\hat{\rho}(h)$, est obtenu à partir de l'estimateur suivant pour l'autocovariance $\hat{\gamma}(h)$ à l'ordre h :

$$\hat{\gamma}(h) = \frac{1}{T-h-1} \sum_{t=h+1}^T (X_t - \bar{X})(X_{t-h} - \bar{X}).$$

On définit la matrice de corrélation (de dimension m) de la manière suivante :

$$R(m) = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(m-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(m-2) \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \rho(m-1) & \rho(m-2) & \dots & \rho(1) & 1 \end{pmatrix}$$

1.7.2 La fonction d'autocorrélation partielle

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

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

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

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

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

avec

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

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

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

Ainsi

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

De manière empirique, les autocorrélations partielles s'estiment par l'estimation des autocorrélations simples et en calculant $\hat{r}(h)$ à partir de la formule ci-dessus.

1.8 Bruit blanc

Définition 1.8.1 *On dit que la suite de variables aléatoires $\{\varepsilon_t\}$ constitue un bruit blanc faible si elle possède les propriétés suivantes [6] :*

$$\left\{ \begin{array}{l} E(\varepsilon_t) = 0, \text{ pour tout } t \in \mathbb{Z} \\ E(\varepsilon_t^2) = \sigma_\varepsilon^2 \neq 0, \\ \text{cov}(\varepsilon_s, \varepsilon_t) = 0, \text{ si } t \neq s. \end{array} \right.$$

En d'autres termes, les variables aléatoires ε_t sont de moyenne nulle, de variance constante et non corrélées.

Remarque 1.8.1 *On dit que $\{\varepsilon_t\}$ est un bruit*

Remarque 1.8.2 *blanc fort s'il est un bruit blanc faible et que les variables aléatoires ε_t sont i.i.d. (indépendantes et identiquement distribuées).*

1.9 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 est stable si elle ne comporte pas de tendance à la hausse ou à la baisse. Plus formellement, on distingue deux types de stationnarité, à savoir forte et faible [7] .

Définition 1.9.1 (Stationnarité faible) *Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit **stationnaire** ou **faiblement stationnaire** ou **stationnaire au second ordre** si*

- i $E(X_t^2) < \infty$,
- ii $E(X_t) = \mu, \forall t \in \mathbb{Z}$ (donc ne dépend pas du temps) ,
- iii $Cov(X_t, X_{t-h}) = E(X_t X_{t-h}) - E(X_t)E(X_{t-h}) = \gamma(h) \forall t \in \mathbb{Z}, \forall h \in \mathbb{Z}$ (ne dépend pas du temps) .

En résumé, un processus X_t est dit stationnaire du second ordre si sa moyenne, sa variance et sa covariance sont indépendantes du temps et si sa variance est finie.

Définition 1.9.2 (Stationnarité stricte) *Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit **strictement stationnaire** ou **stationnaire au sens strict** si les lois jointes de $(X_{t_1}, \dots, X_{t_k})$ et de $(X_{t_1+h}, \dots, X_{t_k+h})$ sont identiques pour tout entier positif k et pour tous $t_1, \dots, t_k, h \in \mathbb{Z}$.*

Intuitivement, une série chronologique strictement stationnaire doit avoir le même comportement statistique sur des intervalles de temps égaux.

1.9.1 Relation entre stationnarité faible et stricte

Propriété 1.9.1 *Un processus strictement stationnaire du second ordre est faiblement stationnaire. La réciproque n'est pas vraie en général.*

Contre-exemple : Soit (X_t) une suite de variables aléatoires indépendantes telle que :

- $X_t \sim \exp(1)$ lorsque t est pair,
- $X_t \sim N(1, 1)$ lorsque t est impair,

alors (X_t) est stationnaire avec $\gamma_X(0) = 1$ et $\gamma_X(h) = 0$ lorsque $h \neq 0$. Cependant X_1 et X_2 n'ont pas la même loi donc (X_t) n'est pas strictement stationnaire.

Difference-Stationary

Un processus X_t est dit difference-stationary (*DS*) ou stationnaire en différence s'il peut s'écrire sous l'une des formes suivante :

- (1) $X_t = X_{t-1} + \varepsilon_t$.
- (2) $X_t = X_{t-1} + \beta_t + \varepsilon_t$ (avec $\beta \neq 0$).

Trend-Stationary

Il sera dit Trend-stationary (*TS*) ou stationnaire en tendance s'il peut se mettre sous la forme :

$$X_t = g(t) + \varepsilon_t,$$

$g(t)$ est une fonction polynômiale du temps.

La bonne manière de stationnariser une série *TS* consiste à estimer, en général par les moindres carrés ordinaires (*MCO*), l'expression de la tendance et à la retirer. Tandis que la stationnarisation des séries *DS* se fait par passage aux différences.

1.9.2 Test de Dickey Fuller Augmenté

Les tests de stationnarité permettent de vérifier si une série est stationnaire ou non. Le test de Dickey-Fuller, le test augmenté de Dickey-Fuller (*ADF*), ou encore le test de Phillips-Perron (*PP*) sont des tests de racine unitaire, pour lesquels l'hypothèse nulle est que la série a été générée par un processus présentant une racine unitaire, et donc, qu'elle n'est pas stationnaire. On donne ici le test de *ADF* :

Les hypothèses du test de *ADF* se définissent de la façon suivante ; Un processus non stationnaire, il correspond à une de ces formes de non stationnarité :

$$\Delta X_t = \phi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t \text{ avec } \phi = \rho - 1 \quad (1)$$

$$\Delta X_t = \alpha + \phi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t \quad (2)$$

$$\Delta X_t = \alpha + \beta t + \phi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t \quad (3)$$

1.10 Théorème de Wold

Théorème 1.10.1 *Tout processus $(X_t)_{t \in \mathbb{Z}}$ faiblement stationnaire peut s'écrire sous la forme [8] :*

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

où les paramètres ψ_1, ψ_2, \dots sont des réels tels que $\psi_0 = 1$ et $\sum_{j=0}^{+\infty} \psi_j^2 < +\infty$,

$\varepsilon_{t \in \mathbb{Z}}$ est un bruit blanc et k_t est une composante linéaire telle que $Cov(k_t, \varepsilon_{t-j}) = 0$ pour tout $j \in \mathbb{Z}$.

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

Dans ce qui suit, deux exemples de séries chronologiques stationnaires sont présentés. On obtiendra, dans chaque cas, leur représentation selon la formule du théorème de Wold.

Exemple 1.10.2 *Soit le processus*

$$X_t = \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^j V_{t-j+1},$$

où $(V_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien de variance unitaire. Afin que la condition $\psi_0 = 1$ du théorème de Wold soit satisfaite, il s'agit de poser

$$\varepsilon_t = \frac{1}{2}V_t.$$

on a alors

$$X_t = \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^{j-1} \varepsilon_{t-j+1} = \varepsilon_t + \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^j \varepsilon_{t-j},$$

où $(\varepsilon)_{t \in \mathbb{Z}}$ est un bruit blanc de variance $Var(\varepsilon_t) = Var(V_t/2) = \frac{1}{4}$. Ainsi, $\psi_0 = 1$ et $\psi_j = (1/2)^j$. On note que la condition $\sum_{j=0}^{+\infty} \psi_j^2 < +\infty$ est également satisfaite car

$$\sum_{j=0}^{+\infty} \psi_j^2 = \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^{2j} = \sum_{j=0}^{+\infty} \left(\frac{1}{4}\right)^j = \left(\frac{4}{3}\right),$$

où le résultat

$$\sum_{j=0}^{+\infty} \alpha^j = \frac{1}{1-\alpha}, \quad |\alpha| < 1,$$

été utilisé.

Exemple 1.10.3 *Soit le processus*

$$X_t = \rho X_{t-1} + \varepsilon_t.$$

où $|\rho| < 1$ et $\varepsilon_{t \in \mathbb{Z}}$ est un bruit blanc de variance $\sigma_\varepsilon^2 = 1$.

Alors X_t a une représentation sous la forme

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

qui est précisément un cas particulier de la formule $X_t = \sum_{j=0}^{+\infty} \psi_j \varepsilon_{t-j} + k_t$ avec $\psi_j = \rho^j$. En utilisant le résultat sur la somme géométrique infinie, on note que

$$\sum_{j=0}^{+\infty} \psi_j^2 = \sum_{j=0}^{+\infty} (\rho^j)^2 = \sum_{j=0}^{+\infty} (\rho^2)^j = \frac{1}{1 - \rho^2} < +\infty.$$

1.11 Opérateurs définis sur une série chronologique

Si la série X_t n'est pas stationnaire l'opérateur suivant rend cette série stationnaire [9]

1.11.1 Opérateur de retard

Définition 1.11.1 *L'opérateur de retard B se définit de la manière suivante :*

$$B(X_t) = X_{t-1}$$

Remarque 1.11.1 *L'opérateur B est linéaire et inversible. Son inverse $B^{-1} = F$ est défini par*

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

L'opérateur F est appelé opérateur d'avance.

Si on compose B avec lui-même on obtient $B^2 = B \circ B$ tel que

$$\forall t \in \mathbb{Z}, B^2 X_t = X_{t-2}.$$

On peut itérer cette application et définir par récurrence

$$B^k X_t = X_{t-k}, k \in \mathbb{N}.$$

Par convention, B^0 est l'opérateur identité I .

1.11.2 Opérateur de différence d'ordre d

Définition 1.11.2 On définit l'opérateur Δ^d de différence d'ordre d comme l'opérateur linéaire tel que

$$\Delta^d X_t = (1 - B)^d X_t.$$

On peut aussi prendre l'opérateur d'ordre 1 et l'appliquer plusieurs fois :

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

1.12 Analyse de la tendance

Dans cette section, nous nous plaçons dans le cadre d'un modèle composé uniquement d'une tendance et de fluctuations irrégulières et donnons différentes méthodes permettant d'estimer la tendance.

1.12.1 Rappels sur la régression linéaire

Lorsqu'une liaison linéaire forte entre deux variables X et Y semble raisonnable au vu du nuage de points, on a alors une relation du type :

$$Y \simeq aX + b.$$

où les coefficients a et b sont inconnus.

Le problème est que les points du nuage sont rarement (parfaitement) alignés : ils sont proches d'une droite.

Nous cherchons maintenant la droite qui passe au plus près des points du nuage. Pour cela, il faut donc mesurer l'éloignement des points du nuage par rapport à une droite D d'équation $y = ax + b$ puis minimiser un critère d'erreur donné. On peut envisager de minimiser

- la somme des erreurs en valeur absolue : $\min_{a,b} \sum_{i=1}^n |y_i - ax_i - b|$.
- la somme des erreurs au carré : $\min_{a,b} \sum_{i=1}^n (y_i - ax_i - b)^2$.

La méthode des moindres carrés minimisant le second critère est la plus utilisée.

1.12.2 La méthode des moindres carrés

On démontre en minimisant la fonction de deux variables

$$g(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2.$$

que le couple solution (\hat{a}, \hat{b}) est donné par

$$\hat{a} = \frac{\frac{1}{n} \sum_{i=1}^n y_i x_i - (\frac{1}{n} \sum_{i=1}^n y_i)(\frac{1}{n} \sum_{i=1}^n x_i)}{\frac{1}{n} \sum_{i=1}^n x_i^2 - (\frac{1}{n} \sum_{i=1}^n x_i)^2} = \frac{Cov(X, Y)}{Var(X)},$$

et

$$\hat{b} = \frac{1}{n} \sum_{i=1}^n y_i - \hat{a} \frac{1}{n} \sum_{i=1}^n x_i = \bar{Y} - \hat{a} \bar{X}.$$

La droite d'équation $y = \hat{a}x + \hat{b}$ est appelée droite de régression de Y en X et est notée : $\Delta_{Y/X}$.

1. Cette droite passe par le point moyen $M(\bar{X}; \bar{Y})$.
2. Le coefficient directeur a de $\Delta_{Y/X}$, $Cov(X; Y)$ et $r(X; Y)$ sont de même signe :
 - Lorsqu'ils sont positifs, on parle de corrélation positive (y augmente quand x augmente).
 - Lorsqu'ils sont négatifs, on parle de corrélation négative (y diminue quand x augmente).

1.12.3 Propriétés et interprétation du coefficient de corrélation linéaire

Afin de confirmer qu'il est raisonnable d'approximer le nuage de points par une droite, on calcule le coefficient de corrélation linéaire :

$$r(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}.$$

a. Le coefficient de corrélation linéaire est symétrique :

$$r(X, Y) = r(Y, X).$$

b. L'inégalité de **Cauchy-Schwarz** donne :

$$-1 \leq r(X, Y) \leq 1.$$

En pratique, il faut commencer par tracer le nuage de points puis calculer $r(X, Y)$ et ce n'est que si la corrélation linéaire est assez forte que l'on cherchera la droite de régression de Y en X .

1.12.4 Estimation non paramétrique

Dans certaines situations, il n'est pas facile de trouver le degré du polynôme d'ajustement pour Z_t ou de changement de variable adéquat. On pourrait utiliser un polynôme avec un degré élevé mais le nombre de paramètres à estimer serait important et rendrait les calculs fastidieux. Par ailleurs, on ne sait pas non plus déterminer l'allure de cette fonction. Dans cette situation, on a recours à la théorie non paramétrique de l'estimation de la tendance qui ne suppose rien sur celle-ci a priori et on approxime la tendance par la moyenne mobile arithmétique d'ordre k .

$$Z_t^* = M_k(Z_t).$$

Alors la tendance à la date t peut être estimée par la moyenne mobile (centrée) d'ordre k à la date t .

1.13 Les moyennes mobiles

Définition 1.13.1 On appelle **moyenne mobile**, une transformation de X_t s'écrivant comme combinaison linéaire finie des valeurs de la série correspondant à des dates entourant t . La série transformée s'écrit

$$M_{m_1+m_2+1}X_t = \sum_{i=-m_1}^{m_2} \theta_i X_{t+i} = \theta_{-m_1}X_{t-m_1} + \theta_{-1}X_{t-1} + \theta_0X_t + \theta_1X_{t+1} + \dots + \theta_{m_2}X_{t+m_2}.$$

où $\theta_{-m_1}, \dots, \theta_{m_2}$ sont des réels et $m_1, m_2 \in \mathbb{N}$. On appelle **ordre** de la moyenne mobile la valeur $m_1 + m_2 + 1$.

– On dit que la moyenne mobile est **centrée** lorsque $m_1 = m_2 = m$.

- Une moyenne mobile centrée est **symétrique** si et seulement si $\theta_{-i} = \theta_i, i = 1, \dots, m$.
- Une moyenne mobile **arithmétique** est une moyenne mobile **centrée**, d'ordre (impair) $2m + 1$ et telle que

$$\theta_i = \frac{1}{2m + 1}, \forall i = -m \dots m.$$

Une moyenne mobile arithmétique est donc centrée (par définition) et symétrique. On a en particulier dans ce cas, $\sum_{i=-m}^m \theta_i = 1$ et $M_{2m+1}X_t$ apparaît comme la moyenne des observations $X_{t-m}, \dots, X_t, \dots, X_{t+m}$.

Cas particulier : Moyenne mobile arithmétique

La série des moyennes mobiles arithmétiques d'ordre k (k est impair), est la série des moyennes de k observations consécutives et elle prend ses valeurs aux dates moyennes correspondantes. Plus précisément, on calcule les moyennes de k termes consécutifs pour les dates

$$\frac{t_1 + \dots + t_k}{k} \text{ puis } \frac{t_2 + \dots + t_{k+1}}{k} \text{ jusqu'à } \frac{t_{T-k+1} + \dots + t_T}{k},$$

et pour la variable d'intérêt

$$\frac{X_1 + \dots + X_k}{k} \text{ puis } \frac{X_2 + \dots + X_{k+1}}{k} \text{ jusqu'à } \frac{X_{T-k+1} + \dots + X_T}{k}.$$

Exemple 1.13.1 Calcul d'une moyenne mobile arithmétique d'ordre 3

Date t	Série y_t	Date $M_3(t)$ de la $M M$	$M_3(y_t)$
1	5		
2	3	$(1 + 2 + 3)/3 = 2$	$(5 + 3 + 4)/3 = 4$
3	4	$(2 + 3 + 4)/3 = 3$	$(3 + 4 + 5)/3 = 4$
4	5	$(3 + 4 + 5)/3 = 4$	$(4 + 5 + 4)/3 = 4.33$
5	4	$(4 + 5 + 6)/3 = 5$	$(5 + 4 + 4)/3 = 4.33$
6	4		

TAB. 1.1 – Moyenne Mobile d'ordre 3

Lorsque k est impair, $k = 2m + 1$, la série moyenne mobile est calculée aux mêmes instants que les observations initiales. En revanche, lorsque k est pair,

$k = 2m$, la moyenne mobile est calculée entre les dates d'observations. Si l'on veut comparer la série transformée à la série initiale, on a besoin d'avoir les valeurs pour les mêmes dates d'observations. Pour pallier cet inconvénient, on prendra plutôt comme transformation

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

combinaison linéaire des moyennes mobiles arithmétiques sur $2m$ valeurs

$$M_k^1 X_t = \frac{1}{k} (X_{t-m} + \dots + X_{t+m-1}) = M_{m_1+m_2+1} X_t \text{ avec } m_1 = m, m_2 = m - 1,$$

et

$$M_k^2 X_t = \frac{1}{k} (X_{t-m-1} + \dots + X_{t+m}) = M_{m_1+m_2+1} X_t \text{ avec } m_1 = m - 1, m_2 = m.$$

Exemple 1.13.2 *Calcul des moyennes mobiles arithmétiques d'ordre 2 et 4*

Date t	Série y_t	Date $M_2(t)$ de la M	$M_2(y_t)$
1	5		
2	3	$(1/2 + 2 + 3/2)/2 = 2$	$(5/2 + 3 + 4/2)/2 = 3.75$
3	4	$(2/2 + 3 + 4/2)/2 = 3$	$(3/2 + 4 + 5/2)/2 = 4$
4	5	$(3/2 + 4 + 5/2)/2 = 4$	$(4/2 + 5 + 4/2)/2 = 4.5$
5	4	$(4/2 + 5 + 6/2)/2 = 5$	$(5/2 + 4 + 4/2)/2 = 4.25$
6	4		

TAB. 1.2 – Moyenne Mobile d'ordre 2

Date t	Série y_t	Date $M_4(t)$ de la $M M$	$M_4(y_t)$
1	5		
2	3		
3	4	$(1/2 + 3 + 4 + 5/2)/4 = 3$	$(5/2 + 3 + 4 + 5 + 4/2)/4 = 4.125$
4	5	$(2/2 + 3 + 4 + 5 + 6/2)/4 = 4$	$(3/2 + 4 + 5 + 4 + 4/2)/4 = 4.125$
5	4		
6	4		

TAB. 1.3 – Moyenne Mobile d'ordre 4

1.13.1 Propriétés d'un lissage par moyenne mobile

Effet d'une moyenne mobile sur une tendance

L'application d'une moyenne mobile arithmétique (paire ou impaire) ne modifie pas une tendance constante.

L'application d'une moyenne mobile arithmétique (paire ou impaire) conserve une tendance linéaire.

Plus précisément,

1. Une moyenne mobile conserve les constantes si et seulement si

$$\theta_{-m_1} + \dots + \theta_{m_2} = 1.$$

2. Une moyenne mobile symétrique conservant les constantes et les polynômes de degré 1.

Effet d'une moyenne mobile sur une composante saisonnière

Si la série X_t possède une composante saisonnière de période P alors l'application d'une moyenne mobile d'ordre P supprime cette saisonnalité. La série $(M_P X_t)_t$ ne possède plus de composante saisonnière de période P .

Plus précisément, cherchons les séries chronologiques qui sont arrêtées par le filtre moyenne mobile centrée d'ordre $2m + 1$. Ce sont les séries S_t telles que leur transformée S_t^* par la moyenne mobile vérifie $S_t^* = 0$.

On cherche donc à déterminer le noyau d'une moyenne mobile arithmétique d'ordre $2m + 1$ ou encore les vecteurs propres associés à la valeur propre nulle.

Pour trouver les séries arrêtées par une moyenne mobile, on considère le polynôme

$$P(x) = \theta_{-m_1} + \theta_{-m_1+1}x + \dots + \theta_{m_2}x^{m_1+m_2}.$$

dont on cherche les racines. Dans le cas centré, on considère simplement

$$P(x) = \theta_{-m_1} + \theta_{-m_1+1}x + \dots + \theta_m x^{2m}.$$

Propriété 1.13.3 *La moyenne mobile d'ordre $m_1 + m_2 + 1$ arrête les fonctions de la forme a^t ssi a est racine du polynôme $P(x)$.*

Effet d'une moyenne mobile sur les fluctuations irrégulières

Jusqu'à présent nous nous sommes intéressés juste à l'effet d'une moyenne mobile sur la partie déterministe de la série (tendance et saisonnalité). Nous allons étudier maintenant l'effet d'une moyenne mobile sur le résidu lorsque celui-ci est un bruit blanc. Par construction, une moyenne mobile consiste à faire des moyennes partielles de proche en proche. On obtient donc un lissage de la série. L'effet de la composante irrégulière est d'autant plus atténué que l'ordre de la moyenne mobile est grand.

Plus précisément,

Propriété 1.13.4 *Les moyennes mobiles arithmétiques d'ordre $2m + 1$ sont les moyennes mobiles minimisant la variance d'un bruit blanc parmi les moyennes mobiles centrées telles que $\sum_{-m}^m \theta_i = 1$.*

1.13.2 La série lissée par moyenne mobile

Tout d'abord, on applique une moyenne mobile arithmétique d'ordre $2m + 1$ dans le cas d'une saisonnalité d'ordre impair $2m + 1$ ou une moyenne mobile arithmétique modifiée d'ordre $2m + 1$ dans le cas d'une saisonnalité de période paire $2m$. Dans chaque cas, on a vu que la saisonnalité est ainsi annulée et que la variance du bruit est diminuée. Si le modèle est bon, la série transformée ne contient plus aucun mouvement saisonnier. Notons que la série ainsi obtenue est de longueur $T - 2m$.

Chapitre 2

Modélisation *ARIMA* et *SARIMA* des séries chronologique

Le modèle **ARMA** est une sorte de modèle de série chronologique aléatoire commun, fondé par **Box** et **Jenkins**, il est également appelé méthode **B-J**. Ces dernières années, la méthode est largement appliquée dans les prévisions météorologiques, en raison de ses avantages et de sa simple application selon la théorie, La moyenne mobile intégrée autorégressive ou **ARIMA**, est l'une des méthodes de prévision les plus largement utilisées pour la prévision de données de séries chronologiques univariées. Les modèles **ARIMA** visent à décrire les autocorrélations dans les données. Bien que la méthode puisse gérer des données avec une tendance, elle ne prend pas en charge les séries chronologiques avec une composante saisonnière.

Une extension d'**ARIMA** qui prend en charge la modélisation directe de la composante saisonnière de la série est appelée **SARIMA**, Le modèle **SARIMA** a été développé à partir du modèle **ARMA**.

Les modèles **ARIMA** permettent de combiner trois types de modèle temporels : les modèle autorégressifs (*AR*), les modèles moyenne mobile (*MA*) et les modèles intégrés (*I*). Nous exposons ces modèles avec le modèle **SARIMA** dans ce chapitre.

2.1 Les modèles autorégressifs (**AR**)

Définition 2.1.1 *On dit que la série X_t suit un processus autorégressif d'ordre 1 ($AR(1)$) si on peut écrire :*

$$X_t = \phi X_{t-1} + \varepsilon_t,$$

$$\begin{aligned} X_t - \phi X_{t-1} &= \varepsilon_t, \\ (1 - \phi B)X_t &= \varepsilon_t, \end{aligned}$$

où ε_t est un bruit blanc.

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

Définition 2.1.2 On appelle processus autorégressif d'ordre p , noté $AR(p)$, un processus $\{X_t\}_{t \in \mathbb{Z}}$ vérifiant une relation du type :

$$X_t - \sum_{i=1}^p \phi_i X_{t-i} = \varepsilon_t, \quad \forall t \in \mathbb{Z}.$$

$\{\varepsilon_t\}_{t \in \mathbb{Z}}$ est un bruit blanc de variance δ_ε^2 .

ϕ_i : des coefficients réels sachant que $i = 1, \dots, p$, $\phi_p \neq 0$.
cette relation équivalent à

$$(1 - \phi_1 B - \dots - \phi_p B^p) X_t = \varepsilon_t,$$

$$\iff \Phi(B)X_t = \varepsilon_t.$$

$\Phi(B)$: est une equation caractéristique de B (L'opérateur retard).

avec $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$

2.1.1 La stationnairité

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

$$X_t = \Phi(B)^{-1} \varepsilon_t = \sum_{i=-\infty}^{+\infty} h_i \varepsilon_{t-i}.$$

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

2.1.2 L'inversibilité

Définition 2.1.3 Si $(\zeta_t)_t$ peut être exprimé en fonction de $\{X_s, s \leq t\}$ seulement,

$$\zeta_t = \sum_{j=0}^{+\infty} \varphi_j X_{t-j},$$

alors $(X_t)_t$ est inversible.

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

2.1.3 La causalité

Définition 2.1.4 Si $(X_t)_t$ peut être exprimé en fonction de $\{\zeta_s, s \leq t\}$ seulement,

$$X_t = \sum_{i=0}^{+\infty} \phi_i X_{t-i},$$

alors $(X_t)_t$ est causal.

Si le polynôme Φ a toutes ses racines de module strictement supérieur à 1, l'opérateur inverse $\Phi(B)^{-1}$ admet un développement ne faisant intervenir que les puissances positives de B . On a alors :

$$X_t = \sum_{i=0}^{+\infty} h_i \varepsilon_{t-i}.$$

Dans ce cas, on montre que l'on a

$$\sum_{i=0}^{+\infty} |h_i| < +\infty, \quad h_0 = 1$$

Dans ce cas, X_{t-1}, X_{t-2}, \dots sont fonctions linéaires de $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$

Exemple 2.1.1

$$X_t = \phi X_{t-1} + \varepsilon_t \text{ avec } \varepsilon_t \sim BB(0, \sigma_\varepsilon^2) \quad \text{''modèle AR(1)'}$$

La racine de $\Phi(Z) = 1 - \phi Z$ est $Z = 1/\phi$.

- Si $|Z| = 1$ et donc $|\phi| = 1$, par exemple $\phi = 1$ alors :

$$X_t = X_{t-1} + \varepsilon_t = \dots = X_{t-n} + \varepsilon_{t-n+1} + \dots + \varepsilon_{t-1} + \varepsilon_t,$$

et la variance dépend de t donc il n'existe pas de solution stationnaire.

- Si $|Z| > 1$ et donc $|\phi| < 1$, alors :

$$X_t = (1 - \phi B)^{-1} \varepsilon_t = \sum_{i=0}^{+\infty} \phi^i B^i \varepsilon_t = \varepsilon_t + \sum_{i=0}^{+\infty} \phi^i \varepsilon_{t-i}.$$

La représentation est alors causale, en plus d'être stationnaire et inversible, elle est donc canonique, ε_t est le processus d'innovation de X_t , puisque le passé de X_t dépend du passé de ε_t .

- Si $|Z| < 1$ et donc $|\phi| > 1$, alors :

$$X_t = (1 - \phi B)^{-1} \varepsilon_t = - \sum_{i=1}^{+\infty} \phi^{-i} B^{-i} \varepsilon_t = - \left(\frac{\varepsilon_{t+1}}{\phi} + \frac{\varepsilon_{t+2}}{\phi^2} + \dots \right).$$

Cette forme n'est pas tournée vers le passé et la représentation n'est donc pas canonique.

Conclusion 2.1.2 La représentation $\Phi(B)X_t = \varepsilon_t$ est canonique si les racines de $\Phi(B)$ sont supérieures à 1 en module.

2.1.4 Les fonctions auto-covariance, auto-corrélations et équivalence de Yule-Walker :

On considère le cas $\Phi(B)X_t = \varepsilon_t$ avec $E(X_t) = 0$.

L'auto-covariance

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

On a

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

Donc

$$X_t^2 = \phi_1 X_t X_{t-1} + \dots + \phi_p X_t X_{t-p} + X_t \varepsilon_t.$$

Alors

$$\gamma(0) = \phi_1 \gamma(1) + \dots + \phi_p \gamma(p) + E(X_t \varepsilon_t).$$

Or

$$E(X_t \varepsilon_t) = E[(\phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t) \varepsilon_t] + E(\varepsilon_t^2).$$

D'où

$$\gamma(0) = \phi_1 \gamma(1) + \dots + \phi_p \gamma(p) + \sigma_\varepsilon^2.$$

Si $h > 0$, on procède de la même façon :

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

Donc

$$\gamma(h) = \phi_1 \gamma(h-1) + \dots + \phi_p \gamma(h-p) + \underbrace{E(\varepsilon_t X_{t-h})}_{=0 \text{ car } \varepsilon_t \perp X_{t-h}}.$$

Les auto-corrélations

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

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

$$\rho(h) = \phi_1 \rho(h-1) + \dots + \phi_p \rho(h-p), \quad \forall h \geq 0.$$

Ces dernières équations sont appelées **équations de Yule – Walker**.

Pour $h > 0$, les $\gamma(h)$ et les $\rho(h)$ vérifient une relation de récurrence d'ordre p et

$$1 = \phi_1\rho(1) + \dots + \phi_p\rho(p) + \frac{\sigma_\varepsilon^2}{\gamma(0)},$$

Donc

$$\gamma(0) = \sigma_\varepsilon^2 \frac{1}{1 - (\phi_1\rho(1) + \dots + \phi_p\rho(p))}.$$

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

$$\begin{pmatrix} 1 & \rho(1) & \rho(p-1) \\ \rho(1) & 1 & \rho(p-2) \\ \vdots & \dots & \vdots \\ \vdots & \dots & \vdots \\ \rho(p-1) & \dots & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \vdots \\ \vdots \\ \phi_p \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \vdots \\ \vdots \\ \rho(p) \end{pmatrix}$$

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 ϕ_1, \dots, ϕ_p .

$$\Leftrightarrow \begin{cases} \rho(1) = \phi_1 + \phi_2\rho(1) + \dots + \phi_p\rho(p-1) \\ \vdots \\ \vdots \\ \rho(p) = \phi_1\rho(p-1) + \dots + \phi_{p-1}\rho(1) + \phi_p \\ \vdots \\ \vdots \\ \phi_1 = (1 - \phi_2)\rho(1) - \dots - \phi_p\rho(p-1) \\ \vdots \\ \vdots \\ \phi_p = \rho(p) - \phi_1\rho(p-1) - \dots - \phi_{p-1}\rho(1) \end{cases}$$

On peut donc aussi obtenir $\rho(1), \dots, \rho(p)$ en fonction de ϕ_1, \dots, ϕ_p .

Auto-corrélations partielle

Dans un processus auto-régressif d'ordre p ; il est possible de montrer que

$$\Gamma(h) = \begin{cases} 1 & si & h = 0, \\ \rho(1) & si & h = 1, \\ \phi_p & si & h = p, \\ 0 & si & h > p. \end{cases}$$

Estimation des paramètres

On dispose d'une observation $\{x_1, \dots, x_T\}$ de longueur T d'un processus stationnaire X_t supposé suivre un modèle $AR(p)$, c'est à dire vérifiant :

$$X_t = \phi_0 + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t,$$

avec $t \in \mathbb{Z}$ et ϕ_1, \dots, ϕ_p , des paramètres inconnus. On cherche alors à estimer ces paramètres à l'aide des observations disponibles.

Estimateurs de Yule-Walker La méthode consiste à reprendre les équations de Yule-Walker en inversant les relations : on exprime les coefficients en fonction des auto-corrélations, on trouve les paramètres estimés d'après les auto-corrélations estimées.

On a vu précédemment que

$$\rho(h) = \phi_1 \rho(h-1) + \dots + \phi_p \rho(h-p).$$

En prenant l'équation sous sa forme matricielle

$$\begin{cases} \phi_1 \rho(0) + \phi_2(1) + \dots + \phi_p \rho(p-1) = \rho(1), h = 1 \\ \phi_1 \rho(1) + \phi_2(0) + \dots + \phi_p \rho(p-2) = \rho(2), h = 2 \\ \vdots \\ \vdots \\ \phi_1 \rho(p-1) + \phi_2(p-2) + \dots + \phi_p \rho(0) = \rho(p), h = p \end{cases}$$

que l'on écrira

$$R_p \phi = \rho.$$

Alors

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

avec

$$\hat{R}_p = \begin{pmatrix} 1 & \hat{\rho}(1) & \hat{\rho}(2) & \dots & \hat{\rho}(p-1) \\ \hat{\rho}(1) & 1 & \hat{\rho}(1) & \dots & \hat{\rho}(p-2) \\ \hat{\rho}(2) & \hat{\rho}(1) & 1 & \dots & \hat{\rho}(p-3) \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \hat{\rho}(p-1) & \hat{\rho}(p-2) & \hat{\rho}(p-3) & \dots & 1 \end{pmatrix}$$

$\hat{\rho} = ((\hat{\rho}(i))_{i=1,\dots,p})$ et $\hat{\phi}$ le vecteur des paramètres

$$\hat{\phi} = \begin{pmatrix} \hat{\phi}_1 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\phi}_p \end{pmatrix}$$

2.1.5 Prédiction dans le modèle AR(p)

Dans ce paragraphe, on suppose que $\{X_t\}$ est un processus stationnaire qui suit un modèle $AR(p)$, c'est à dire vérifie :

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

avec ε_t un bruit blanc.

Objectif : on cherche à prédire la valeur prise par le processus aux instants $T+1, T+2, \dots$ à partir de la connaissance des valeurs prises par ce processus jusqu'à l'instant T , c'est à dire de x_1, \dots, x_T .

Le modèle se réécrit sous la forme

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

La prévision à la date $T+1$ est

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

Alors

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

Pour prédire X_{T+2} à partir de X_1, \dots, X_T on a

$$\hat{X}_{T+2} = \phi_1 \hat{X}_{T+1} + \dots + \phi_p X_{T-p+2}.$$

Et de façon générale

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

Remarque 2.1.1 Dans le cas des modèles d'ordre 1, on a $\hat{X}_{T+1} = \phi_1 X_T$ et $\hat{X}_{T+2} = \phi_1 \hat{X}_{T+1} = \phi_1^2 X_T, \dots$ On vérifie aisément par récurrence que

$$\hat{X}_{T+h} = \phi_1^h X_T,$$

donc en particulier que $\hat{X}_{T+h} \rightarrow 0$ quand h tend vers l'infini.

2.2 Les modèles moyennes mobiles, (MA)

Définition 2.2.1 On appelle processus moyenne mobile d'ordre q , noté $MA(q)$ pour Moving Average, un processus $\{X_t\}_{t \in \mathbb{Z}}$ défini par :

$$X_t = \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad \forall t \in \mathbb{Z}.$$

où les θ_i sont des réels, $\theta_i \neq 0$ et $\{\varepsilon\}_{t \in \mathbb{Z}}$ est un processus bruit blanc de variance σ_ε^2 .

cette relation équivalente à

$$X_t = (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t.$$

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

$\Theta(B)$: est une equation caractéristique de B (L'opérateur retard).
avec $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$.

2.2.1 stationnarité

La définition d'un $MA(q)$ est explicite et ne pose donc pas de problème : le processus X_t est parfaitement défini et est automatiquement stationnaire.

2.2.2 causalité

La représentation est causale par définition.

2.2.3 inversibilité

$\Theta(B)$ est un polynôme en B de degré q , que l'on peut factoriser en ayant calculé ses racines $Z_i = 1/\lambda_i, i = 1, \dots, q$:

$$\begin{aligned}\Theta(B) &= 1 - \theta_1 B - \dots - \theta_q B^q. \\ &= (1 - \lambda_1 B)(1 - \lambda_2 B) \dots (1 - \lambda_q B).\end{aligned}$$

Si $\Theta(Z)$ n'a pas de racine de module égal à 1, on peut calculer l'inverse de $\Theta(B)$, qui est alors donnée par :

$$\begin{aligned}\Theta(B)^{-1} &= (1 - \lambda_1 B)^{-1} (1 - \lambda_2 B)^{-1} \dots (1 - \lambda_q B)^{-1}. \\ &= \frac{k_1}{1 - \lambda_1 B} + \dots + \frac{k_q}{1 - \lambda_q B}.\end{aligned}$$

Si toutes les racines de (c'est-à-dire $1/\lambda_i, i = 1, \dots, q$) sont distinctes et où k_1, \dots, k_q sont des paramètres qui dépendent de $\lambda_1, \dots, \lambda_q$.

On obtient alors l'expression suivante :

$$\Theta(B)^{-1} X_t = \sum_{i=-\infty}^{+\infty} \pi_i X_{t-i} = \varepsilon_t.$$

où les π_i sont fonctions des paramètres θ_j , et on peut montrer que

$$\sum_{i=-\infty}^{+\infty} |\pi_i| < +\infty.$$

Si les racines de $\Theta(Z) = 0$ sont toutes de module supérieur à 1, on peut montrer que $\pi_i = 0 \forall i < 0$. On dit que le processus est inversible.

2.2.4 La fonction d'autocorrélation

Comme $MA(q)$ est stationnaire, on peut étudier sa structure d'autocorrélation. En effet, les autocorrélations existent toujours puisque le modèle est stationnaire,

$$\gamma(h) = E[X_t X_{t+h}],$$

ce qui est plus adéquat ici, car il y a non-corrélation des ε_t avec le futur (on suppose par simplicité que $\mu = 0$) :

$$\begin{aligned} X_t &= \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}. \\ X_{t+h} &= \varepsilon_{t+h} - \theta_1 \varepsilon_{t+h-1} - \theta_2 \varepsilon_{t+h-2} - \dots - \theta_q \varepsilon_{t+h-q}. \end{aligned}$$

Comme les ε_t sont un bruit blanc, les ε_t sont des variables aléatoires non-corrélées, autrement dit

$$\begin{cases} E[\varepsilon_{t+i} \varepsilon_{t+i}] = \sigma_\varepsilon^2, \\ E[\varepsilon_{t+i} \varepsilon_{t+j}] = 0, i \neq j, \end{cases}$$

On obtient

$$\gamma(h) = \begin{cases} \sigma_\varepsilon^2(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) & h = 0, \\ \sigma_\varepsilon^2(-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_q \theta_{h+q}) & 1 \leq h \leq q, \\ 0 & h > q. \end{cases}$$

La fonction d'autocorrélation vaut donc

$$\rho(h) = \begin{cases} 1 & h = 0 \\ \frac{-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_q \theta_{h+q}}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & 1 \leq h \leq q, \\ 0 & h > q. \end{cases}$$

2.2.5 Prédiction dans le modèle MA(q)

On supposera là aussi que l'on s'est ramené à un processus centré (X_t), satisfaisant :

$$\begin{aligned} X_t &= \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}. \\ &= \Theta(B) \varepsilon_t. \end{aligned}$$

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

$$X_t = 0 - \theta_1 \varepsilon_T - \dots - \theta_q \varepsilon_{T+1-q}.$$

De façon analogue, X_{T+h} est estimé par

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

et donc

$$\hat{X}_{T+h} = \begin{cases} \theta_h \hat{X}_{T+1} + \dots + \theta_q \hat{X}_{T+h-q} & \text{pour } h \leq q, \\ 0 & \text{pour } h > q. \end{cases}$$

Toutefois, cette méthode présente le désavantage d'estimer X_{T+h} à partir des résidus passés, a priori non observables, et non pas du passé de la variable.

2.3 Les modèles ARMA(p, q) (Autorégressive-Moving Average)

Les modèles ARMA sont représentatifs de processus générés par une combinaison des valeurs passées et des erreurs passées.

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \dots - \theta_q \varepsilon_{t-q}, t \in \mathbb{Z}.$$

On peut aussi écrire le modèle $ARMA(p, q)$ sous la forme :

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

où B est l'opérateur de décalage et (ε_t) est le processus de bruit blanc.

2.3.1 Inversibilité et stationnarité en terme de valeurs passées

Pour que le processus puisse s'écrire :

$$\Pi(B)X_t = \varepsilon_t,$$

où

$$\Pi(B)X_t = \Theta_q^{-1}(B)\Phi_p(B),$$

il faut que toutes les racines de $\Theta_q(z)$ soient à l'extérieur du cercle unité. Le processus est alors inversible.

Pour que le processus puisse s'écrire

$$X_t = \Psi(B)\varepsilon_t,$$

où

$$\Psi(B) = \Theta_q(B)\Phi_p^{-1}(B),$$

il faut que toutes les racines de $\Phi_p(z)$ soient à l'extérieur du cercle unité. Le processus est alors stationnaire.

On suppose donc que $\Phi_p(z)$ et $\Theta_q(z)$ ont leurs racines à l'extérieur du cercle unité de manière à avoir la stationnarité et l'inversibilité.

2.3.2 Auto-corrélation

La covariance vaut

$$\begin{aligned} \gamma(h) &= E[X_t X_{t-h}] \\ &= E \left[\left(\sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \right) X_{t-h} \right]. \end{aligned}$$

Si on note

$$\gamma_{x_\varepsilon}(h) = E[\varepsilon_t X_{t-h}].$$

la covariance croisée entre X_{t-h} et ε_t On a

$$\begin{aligned} \gamma(h) &= \sum_{j=1}^p \phi_j E[X_{t-j} X_{t-h}] - \sum_{i=1}^q \theta_i E[\varepsilon_{t-i} X_{t-h}] + \gamma_{x_\varepsilon}(h) \\ &= \sum_{j=1}^p \phi_j \gamma(h-j) - \sum_{i=1}^q \theta_i \gamma_{x_\varepsilon}(h-i) + \gamma_{x_\varepsilon}(h), h \geq 0 \end{aligned}$$

Or

$$E(X_t \varepsilon_t) = E \left[\left(\sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \right) \varepsilon_t \right] = E[\varepsilon_t \varepsilon_t] = \sigma_\varepsilon^2,$$

et ε_t n'est pas corrélé avec le passé.

$$\begin{cases} \gamma_{x_\varepsilon}(h) = 0 & h > 0, \\ \gamma_{x_\varepsilon}(0) = \sigma_\varepsilon^2, \\ \gamma_{x_\varepsilon}(h) \neq 0 & h < 0. \end{cases}$$

La fonction $\gamma_{x_\varepsilon}(h)$ n'est pas paire.

Si $h > q$, tous les $\gamma_{x_\varepsilon}(h) = 0$, et on a

$$\begin{cases} \gamma(h) = \sum_{j=1}^p \phi_j \gamma(h-j), \\ \rho(h) = \sum_{j=1}^p \phi_j \rho(h-j). \end{cases}$$

Ce qui est un système d'équation de type Yule-Walker. En effet, la partie MA ne joue que jusqu'au délai q , après elle ne joue plus. Il est donc normal de retrouver des équations ressemblant à celles trouvées dans les $AR(p)$ à partir de $h > q$.

Pour un $ARMA(p, q)$ la structure d'auto-corrélation ne suit pas un schéma connu jusqu'au délai q mais ensuite le comportement est le même que celui d'un $AR(p)$.

Il y a donc une infinité de $\gamma(h)$ non-nuls, mais à partir de $h > q$, la fonction d'auto-corrélation a le même comportement que pour $AR(p)$.

2.3.3 Auto-corrélations partielles

Ici la partie, AR ne se fait sentir que jusqu'à l'ordre p . Pour $h \leq p$, il n'y a pas de schéma bien précis. Pour $h > p$, on retrouve le même comportement que pour un $MA(q)$.

2.3.4 Prédiction dans modèle ARMA(p,q)

On supposera là aussi que l'on s'est ramené à un processus centré (X_t) , satisfaisant :

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

Sous cette forme $ARMA$, alors

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

et donc

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

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

On peut noter que pour $h > q$

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

2.4 Les modèles ARIMA(p, d, q)

Une série (X_t) suit un processus *ARIMA* (Auto Regressive Integrated Moving Average) d'ordre (p, d, q) si elle écrit sous la forme :

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

On modélise alors la série sous la forme :

$$\begin{aligned} \Phi(B)(1-B)^d X_t &= \Theta(B)\varepsilon_t, \\ \Phi(B)\Delta^d X_t &= \Theta(B)\varepsilon_t. \end{aligned}$$

où le polynôme $\Phi(B)$ est de degré p et le polynôme $\Theta(B)$ est de degré q . On écrit que la série X_t suit un processus *ARIMA* (p, d, q) .

• Prévisions dans le cas d'un processus ARIMA(p,d,q)

On considérons ici (X_t) satisfaisant une équation de la forme :

$$\Phi(B)(1-B)^d X_t = \Theta(B)\varepsilon_t,$$

Posons alors $\Psi(B) = \Phi(B)(1-B)^d$: La forme *ARIMA* (p, d, q) peut s'écrire

$$X_t = \sum_{j=1}^{p+d} \Psi_j X_{t-j} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i};$$

et donc

$$X_{t+h} = \sum_{j=1}^{p+d} \Psi_j X_{t+h-j} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t+h-i}.$$

Notons \hat{X}_{T+h} la prévision faite à la date $T+h$

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

Alors

$$\hat{X}_{T+h} = \sum_{j=1}^{p+d} \Psi_j \hat{X}_{T+h-j} + 0 + \sum_{i=1}^q \theta_i \hat{\varepsilon}_{T+h-i},$$

Où

$$\hat{X}_{T+h-j} = X_{T+h-j} \text{ pour } j \geq h,$$

et

$$\hat{\varepsilon}_{T+h-i} = \begin{cases} 0 & \text{pour } i < h, \\ \varepsilon_{T+h-i} & \text{pour } i \geq h. \end{cases}$$

En particulier, pour $h \geq q$, on obtient une relation de récurrence de la forme

$$\hat{X}_{T+h} = \sum_{j=1}^{p+d} \Psi_j \hat{X}_{T+h-j}.$$

2.5 Les modèles ARIMA saisonniers (SARIMA)

Si l'on veut en même temps traiter les saisonnalités de période s (sans supposer une répétition exacte, déterministe des données), on est amené à définir les processus *SARIMA*.

2.6 Les processus SARIMA

Les modèles *SARIMA* (S pour Seasonal) permettent de rendre compte des variations saisonnières dans la série considérée, ces variations pouvant elles-mêmes présenter un caractère aléatoire. On trouve successivement :

- un processus ARMA présentant une composante saisonnière de période 20.
- le même processus intégré une fois, à la période de l'échantillon.
- le processus initial intégré une fois, à la période de la saison .
- le processus a été intégré une fois à la période de l'échantillon, et une fois à la période de la saison.

Définition 2.6.1 X_t est appelé processus *autorégressif moyenne mobile et saisonnalité intégrés* avec période s ($X_t \rightsquigarrow \text{SARIMA}(p, d, q) \times (P, D, Q)_s$) (*Seasonal AutoRegressive Integrated Moving Average*), si

$$Y_t = \Delta^d \Delta_s^D X_t = (1 - B)^d (1 - B^s)^D X_t,$$

est un processus *ARMA* stationnaire de la forme

$$A(B) \times F(B^s)Y_t = \Theta(B) \times G(B^s)\varepsilon_t,$$

où $A(z)$ est le polynôme générateur d'un $AR(p)$

$$A(z) = 1 - \sum_{k=1}^p a_k z^k,$$

$\Theta(z)$ est le polynôme générateur d'un $MA(q)$

$$\Theta(z) = 1 + \sum_{k=1}^q \theta_k z^k,$$

et où, pour la saisonnalité $Y_t - Y_{t-s}$, $F(z)$ est le polynôme générateur d'un $AR(P)$

$$F(z) = 1 - \sum_{k=1}^P \phi_k z^k,$$

et $G(z)$ est le polynôme générateur d'un $MA(Q)$

$$G(z) = 1 + \sum_{k=1}^Q \gamma_k z^k,$$

Remarque 2.6.1 Y_t est un processus particulier du type $ARMA(p + P_s, q + Q_s)$ stationnaire .

1) Soit le processus (X_t) défini par

$$X_t = \mu + S_t + \varepsilon_t,$$

où $S_t = S_{t+s} \forall t$, est un processus $SARIMA(0, 0, 0) \times (0, 1, 1)_s$. Donc ,

$$Y_t = (1 - B^s)X_t = \varepsilon_t - \varepsilon_{t-s},$$

est un $ARMA(0, s)$ (non inversible).

2) $s = 12$ (données mensuelles),

$p = 1 = P, q = 0 = Q, d = D = 1$:

soit l'opérateur

$$A(z)F(z^s) = (1 - az)(1 - \phi z^{12}) = (1 - az - \phi z^{12} - az^{13}),$$

$$\implies Y_t = aY_{t-1} + \phi Y_{t-12} - a\phi Y_{t-13} + \varepsilon_t,$$

donc un $AR(p + Ps) = AR(13)$ ou un $AR(1)$ pour $Y_{t-1} - Y_{t-12}$.

3) Soit le processus (X_t) défini par :

$$X_t = a \sin \left(\frac{\pi t}{6} + 4 \right) + \varepsilon_t,$$

ou (ε_t) est un bruit blanc.

On a

$$E(X_t) = a \sin \left(\frac{\pi t}{6} + 4 \right).$$

(X_t) n'est pas stationnaire en moyenne, la partie déterministe est périodique de période 12.

Proposition 2.6.1 *Inclusion des trends ($d > 0$) et des composantes déterministes périodiques ($D > 0$)*

$$\implies SARIMA(p, d, q) \times (P, D, Q)_s.$$

Proposition 2.6.2 *Des autocorrélations oscillantes \implies existence des saisonnalités (processus SARIMA).*

Note :

La notion de processus SARIMA peut s'étendre à plusieurs périodes différentes pour un même processus.

Par exemple, l'observation des températures en un lieu donné fait apparaître :

- des variations journalières,
- des variations annuelles.

Il est nécessaire de disposer de séries de grande taille pour analyser ces processus. Par exemple, observer une température pendant trente ans avec une période de une heure conduit à une taille de l'ordre de 260000.

2.7 Box-Jenkins pour des modèles saisonniers

L'ouvrage de Box et Jenkins "Time series analysis, forecasting and control", publié en 1970 a proposé une démarche de prévision pour les séries univariées, fondée sur l'utilisation de processus SARIMA .

Les étapes pour l'estimation des coefficients d'un processus SARIMA sont les suivantes :

2.7.1 Identification et ajustement des modèles SARIMA aux données

Plan : X_1, \dots, X_T données

0. Transformation

Applications des transformation ("Box-Cox") pour stabiliser la variance transformation de type $(\log, \exp, \sqrt{\cdot})$.

Rappel : Transformations de Box-Cox

Il arrive fréquemment que des séries chronologiques aient une variance de bruit et une amplitude de variations saisonnières qui sont des fonctions croissantes de m . Pour y remédier, une solution est d'utiliser une transformation de Box-Cox pour les séries strictement positives. Cela consiste à transformer une série x_t en $x_t^{(a)}$ où :

$$x_t^{(a)} = \begin{cases} \frac{x_t^a - 1}{a} & \text{si } a \neq 0, \\ \log(x_t) & \text{si } a = 0. \end{cases}$$

où a est une constante à choisir. En pratique, un bon choix de a permet souvent de stabiliser les fluctuations des résidus et de la saisonnalité.

Remarque 2.7.1 *On peut utiliser les tests dits tests de racines unitaire pour vérifier si la série est stationnaire au non.*

1. Choisir d , D , s

(souvent $0 \leq d \leq 2, 0 \leq D \leq 1$) de sorte que $(1 - B)^d(1 - B^s)^D X_t = Y_t$ stationnaire.

Remarque 2.7.2 *Role de D (souvent $0 \leq D \leq 1$) : Si $D > 1$, la tendance n'est pas seulement dans X_t , mais encore dans $(\nabla^d X)_t - (\nabla^d X)_{t-s}$. C.a.d., les différences de l'ordre s , $(\nabla^d \nabla_s X)_t$ ne sont pas stationnaires, il reste une tendance (faible, alors que $D = 2$ suffira).*

Détermination de l'ordre de différenciation d

1. Une série stationnaire fluctue autour d'une valeur moyenne et sa fonction d'autocorrélation décline rapidement vers zéro. Si une série présente des autocorrélations positives pour un grand nombre de décalages (par exemple 10 ou plus), alors elle nécessite d'être différenciée. La différenciation tend à introduire des autocorrélations négatives.

Remarque 2.7.3 Si (X_t) admet une racine unité, la proposition $\rho(h)$ décroît exponentiellement vers 0 avec h n'est plus vraie : c'est la persistance des chocs .

choix de d : combien de fois faut-il différencier pour obtenir une série stationnaire (autocorrélogrammes, tests statistiques...). Donc un estimateur de d est le nombre total de fois ou on rejette la stationnarité.

2. L'ordre optimal de différenciation est souvent celui pour lequel l'écart-type est minimal. Un accroissement de l'écart-type doit donc être considéré comme un symptôme de surdifférenciation.
3. Si l'auto-corrélation de décalage 1 est égale à 0 ou négative, la série n'a pas besoin d'être différenciée. Si l'auto-corrélation de décalage 1 est inférieure à -0.5 , la série est sur-différenciée.
4. Un troisième symptôme de sur-différenciation est un changement systématique de signe d'une observation à l'autre.

2. Stationnarisation par filtrage

- L'opérateur $(1 - B^s)^D$, élimine la périodicité (désaisonnalisation).
- L'opérateur de différenciation ∇ , (ou ∇^d) élimine les tendances.

3. Estimation des paramètres P, Q

Calculer $ACF \hat{\rho}_t$ et $PACF \hat{r}_t$ empirique et examiner $\hat{\rho}_{ks}$, $k \geq 0$, et \hat{r}_{ks} , $k \geq 0$, pour trouver P, Q de sorte que $\hat{\rho}_{ks}$ et \hat{r}_{ks} correspond à $ARMA(P, Q)$ (et décroissent exponentiellement).

Exemple :

$$ACF (s = 12) : |\hat{\rho}_{12}| > IC \implies P \geq 1.$$

$$PACF |\hat{r}_{12}| > IC \implies P \geq 1.$$

4. Estimation des paramètres p, q

Choix du p, q tel que $\hat{\rho}_1, \dots, \hat{\rho}_{s-1}$ et $\hat{r}_1, \dots, \hat{r}_{s-1}$ sont compatibles avec modèle ARMA(p, q).

Méthode pratique d'estimation des ordres p, q

pour estimer les ordres p ou q , on utilise les propriétés vues précédemment sur les formes autocorrélogrammes $\rho(h)$ ou des autocorrélogrammes partiels $r(h)$.

En particulier

Cas d'un processus MA(q)

Soit $X = (X_t, t \in \mathbb{Z})$ un processus MA(q). pour tout $r > q$, alors l'autocorrélation empirique $\rho_X^{(N)}(r)$ est approximativement une gaussienne centrée de variance

$$\frac{1}{N} \left(1 + \sum_{i=1}^q \rho_X(i)^2 \right) = \frac{1}{N} \left(1 + \sum_{i=1}^r \rho_X(i)^2 \right)$$

En pratique, on fait donc l'approximation

$$\forall r > q, \sqrt{N} \frac{\rho_X^{(N)}(r)}{\sqrt{1 + \sum_{i=1}^r \rho_X^{(N)}(i)^2}} \rightsquigarrow \mathcal{N}(0, 1)$$

On en déduit un test (pratique mais peu regoureux) de $H_0 : X$ est un MA(q), si pour (presque) tout $r > q$ et $r < N/4$, on a

$$\left| \rho_X^{(N)}(r) \right| < 1,96 \sqrt{\frac{1 + \sum_{i=1}^r \rho_X^{(N)}(i)^2}{N}}$$

alors on accepte H_0 . Si l'on se limite à $r < N/4$, c'est que l'estimation de $\rho_X^{(N)}(r)$ devient de moins en moins précise quand r croît vers N .

Cas d'un processus AR(p)

Soit $X = (X_t, t \in \mathbb{Z})$ un processus AR(p). pour tout $r > p$, alors l'autocorrélation partielle empirique $r_X^{(N)}(r)$ est approximation une gaussienne centrée de variance $1/N$. On en déduit comme précédemment un test de $H_0 : X$ est un AR(p), pour (presque) tout $r > p$ et $r < N/4$, on a

$$r_X^{(N)}(r) < 1,96 \sqrt{\frac{1}{N}}.$$

alors on accepte H_0 .

Cas d'un processus ARMA(p, q)

On pourrait s'inspirer du critère du coin pour construire un test sur les paramètres p et q . En pratique, on procède différemment et plus simplement en vertu par :

Si X un processus $ARMA(p, q)$ vérifiant

$$\Phi(B)(X) = \Theta(B)(\varepsilon),$$

avec $\Theta(z)$ et $\Phi(z)$ sans racine commune, et ε bruit blanc d'innovation de variance σ^2 , alors il peut se mettre sous forme $MA(\infty)$

$$X_t = \varepsilon_t + \sum_{i \geq 1} \psi_i \varepsilon_{t-i}$$

ou $AR(\infty)$

$$X_t + \sum_{j \geq 1} \eta_j X_{t-j} = \varepsilon_t$$

Comme on a les inégalités

$$E[(X_t - \varepsilon_t - \sum_{i=1}^{q'} \psi_i \varepsilon_{t-i})^2] \leq \sigma^2 \left(\sum_{i \geq q'+1} |\psi_i| \right)^2$$

$$E[(X_t - \varepsilon_t - \sum_{j=1}^{p'} \eta_j \varepsilon_{t-j})^2] \leq \gamma_X(0) \left(\sum_{j \geq p'+1} |\eta_j| \right)^2$$

et que $\sum_{i > 0} |\psi_i|$ comme $\sum_{j > 0} |\eta_j|$ sont finis, on en déduit qu'il est possible d'approcher d'aussi près qu'on le souhaite X par un processus de type $MA(q')$ ou $AR(p')$ avec $p' \geq p$ et $q' \geq q$: En pratique, on choisit pour valeurs de p' et q' celles que l'on détermine par les méthodes vues précédemment, en supposant le processus AR puis MA . Cela fournit une première modélisation de X comme processus $ARMA(p', q')$, avec en générale plusieurs couples (p', q') possibles. On verra par la suite comment améliorer cette première estimation majoration des paramètres du processus $ARMA$.

5. Estimation des paramètres du filtre ARMA

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

Estimation préliminaire

(i) Estimation de la partie autorégressive : Si X est un processus $ARMA(p, q)$ causal et inversible, vérifiant $\Phi(B)X = \Theta(B)\varepsilon$, alors on a vu que sa fonction d'autocovariance vérifie la relation de récurrence suivante :

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

et que l'on peut en déduire le système d'équations dit aussi de Yule-Walker

$$\begin{pmatrix} \rho_X(q) & \dots & \rho_X(q - p + 1) \\ \vdots & \ddots & \vdots \\ \rho_X(q + p - 1) & \dots & \rho_X(q) \end{pmatrix} \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_p \end{pmatrix} = \begin{pmatrix} \rho_X(q + 1) \\ \vdots \\ \rho_X(q + p) \end{pmatrix}$$

Les estimateurs de Yule Walker de ϕ_1, \dots, ϕ_p sont donc naturellement définis par

$$\begin{pmatrix} \hat{\rho}_X^{(N)}(q) & \dots & \hat{\rho}_X^{(N)}(q - p + 1) \\ \vdots & \ddots & \vdots \\ \hat{\rho}_X^{(N)}(q + p - 1) & \dots & \hat{\rho}_X^{(N)}(q) \end{pmatrix} \begin{pmatrix} \hat{\phi}_1^{(N)} \\ \vdots \\ \hat{\phi}_p^{(N)} \end{pmatrix} = \begin{pmatrix} \hat{\rho}_X^{(N)}(q + 1) \\ \vdots \\ \hat{\rho}_X^{(N)}(q + p) \end{pmatrix}$$

(ii) Estimation de la partie moyenne mobile et de la variance du bruit blanc : On s'appuie sur la remarque simple suivant : $Y = \hat{\phi}^{(N)}(B)(X)$ est un processus $MA(q)$, de paramètres $\theta_1, \dots, \theta_q$: Une fois calculé $\hat{\phi}^{(N)}(z) = 1 + \hat{\phi}_1^{(N)}z + \dots + \hat{\phi}_q^{(N)}z^q$, on pose

$$Y = \hat{\phi}^{(N)}(B)(X)$$

et l'on détermine les estimateurs $\hat{\theta}_1^{(N)}, \dots, \hat{\theta}_q^{(N)}$ on appliquant à Y l'algorithme des innovations.

On en déduit également l'estimateur $\sigma_{(N)}^2$ de la variance du bruit blanc σ^2 .

6. Validation

Il s'agit de vérifier notamment que les résidus du modèle $ARMA$ estimé, résidus notés $\hat{\varepsilon}_t$, vérifient les propriétés requises pour que l'estimation soit valide, à savoir qu'ils suivent un processus BB , non autocorrélé et de même variance, et qu'ils

suivent une loi normale. Si ces hypothèses ne sont pas rejetées, on peut alors mener des tests sur les paramètres.

Tests de bruit blanc et de stationnarité

1. Analyse des fonctions d'autocorrélation

i. Regarder le graphique des résidus estimés pour voir s'il apparaît des points aberrants, une tendance, une rupture, de l'autocorrélation, etc. Ceci n'est évidemment qu'indicatif.

ii. Regarder les autocorrélations simples et partielles. Elles doivent être significativement nulles si les résidus sont un bruit blanc.

2. Test du portemanteau

Afin de tester que les résidus estimés suivent un BB , on teste l'hypothèse d'absence d'autocorrélation jusqu'à l'ordre m . En pratique, on utilise deux tests :

a) Le test de Box-Pierce :

Ce test a pour objet de tester le caractère non autocorrélé des résidus.

b) Le test de Ljung et Box :

Ce test est à appliquer, de préférence au test de Box-Pierce, lorsque l'échantillon est de petite taille, la statistique de Ljung-Box, donnée par :

$$LB = T(T+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{T-k}$$

où les coefficients d'autocorrélation $\hat{\rho}_k$ sont calculés sur les résidus estimés $\hat{\varepsilon}_t$. Cette statistique, sous l'hypothèse que les résidus suivent un BB , suit une loi du χ_m^2 .

3. Test de normalité

Il s'agit de tester que les résidus estimés $\hat{\varepsilon}_t$ suivent une loi normale, c'est-à-dire ne présentent pas d'asymétrie (Skewness) ni d'aplatissement (kurtosis).

Le coefficient de Skewness est donné par :

$$\beta_1^{1/2} = \frac{\mu_3}{\mu_2^{3/2}}$$

et le coefficient de kurtosis est donné par :

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

où

$$\mu_k = \frac{1}{T} \sum_{t=1}^T (\hat{\varepsilon}_t - \bar{\hat{\varepsilon}})^k$$

est le moment centré d'ordre k de la variable $\hat{\varepsilon}_t$.

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

$$\beta_1^{1/2} \rightsquigarrow \mathcal{N}(0, \sqrt{6/T}) \beta_2 \rightsquigarrow \mathcal{N}(3, \sqrt{24/T})$$

Tests sur les paramètres

On vérifie tout d'abord que les racines des polynômes AR et MA ne sont pas égales à 1.

Si les hypothèses testées sur les résidus ne sont pas rejetées, on teste la significativité des retards du modèle ARMA par des tests de Student.

Choix d'un modèle parmi plusieurs

Si, à la suite de ces étapes, il reste plusieurs modèles valides, on peut choisir parmi ces modèles, soit celui qui donne les meilleurs critères d'ajustement, soit celui qui donne les meilleures performances en prévision. Concernant les critères d'ajustement, on retient le modèle qui minimise les critères d'information (AIC et BIC). Ces critères sont tous basés sur les deux idées suivantes :

1. Étant donné que σ_2 est la variance d'erreur de prévision à horizon 1, on aimerait choisir, parmi les modèles estimés, celui qui fournit la plus petite valeur.

2. On ne peut pas, pour des raisons statistiques, choisir un modèle présentant un grand nombre de paramètres (p et q grands).

Concernant les performances en prévision des modèles, on utilise couramment les critères suivants (que l'on cherche bien entendu à minimiser) :

$$RMSE = \sqrt{\frac{1}{T-k+1} \sum_{t=k}^T (Y_t - \hat{Y}_t(h))^2}$$

et

$$MAE(h) = \frac{1}{T-k+1} \sum_{t=k}^T |Y_t - \hat{Y}_t(h)|$$

où K est le nombre d'observations minimales pour mener une estimation du modèle. On peut calculer ces critères, soit sur la base de prévisions in-sample (toutes les observations ont été utilisées pour estimer le modèle et on calcule les prévisions sur cet même ensemble d'observations), soit sur la base de prévisions

out-of-sample (on estime le modèle sur un ensemble d'observations et on mène la prévision sur le reste).

2.7.2 Prédiction des processus SARIMA

Si les données X_1, \dots, X_T suivent un modèle $SARIMA(p, d, q) \times (P, D, Q)_s$, on définit

$$Y_t = (I - B)^d (I - B^s)^D X_t,$$

Si par exemple, $(d, D, s) = (1, 1, 12)$,

$$Y_t = (X_t - X_{t-12}) - (X_{t-1} - X_{t-13}),$$

On traite le problème de prédiction du processus Y_t du type $ARMA(p + P_s, q + Q_s)$,

(a) Traitement de partie $AR(p + P_s)$ de Y_t , par exemple :

$$\begin{aligned} Y_t &= aY_{t-1} + \phi Y_{t-12} - a\phi Y_{t-13} + \varepsilon_t, \\ \implies \hat{Y}_{T+1} &= aY_T + \phi Y_{T-11} - a\phi Y_{T-12}, \\ \implies \hat{Y}_{T+2} &= a\hat{Y}_{T+1} + \phi Y_{T-10} - a\phi Y_{T-11}, \\ &etc. \end{aligned}$$

$$\implies \hat{X}_T = \hat{Y}_{T+1} + X_{T-11} + X_T - Y_{T-12},$$

(b) Traitement de partie $MA(q + Q_s)$ de Y_t comme vu avec les résidus, par exemple

pour un $SARIMA(1, 1, 1) \times (1, 1, 1)_{12}$:

$$e_t = Y_t - aY_{t-1} + \phi Y_{t-12} - a\phi Y_{t-13} - \theta e_{t-1} - \gamma e_{t-12} - \theta\gamma e_{t-13},$$

Prédiction selon Box-Jenkins si T est suffisamment grand et l'horizon h n'est pas trop grand. Sinon : méthodes non paramétriques, par exemple lissage exponentiel. et on rentre après au niveau du processus X_t du type $SARIMA$ en exprimant $\hat{X}_T(1)$ en fonction de $\hat{Y}_T(1)$ et les valeurs observées X_t pour $t \leq T$.

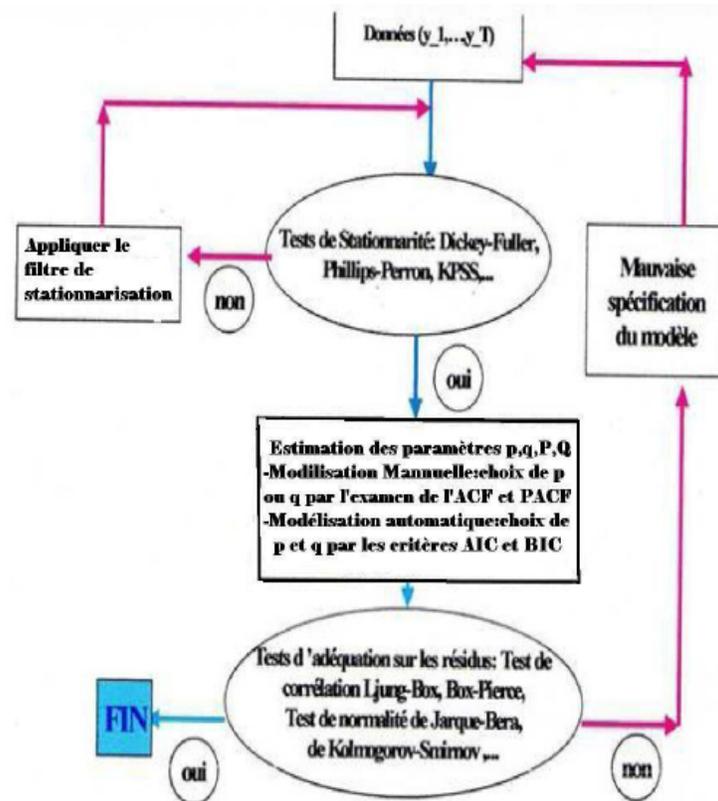


FIG. 2.1 – Schéma général de la modélisation d'une série temporelle par un modèle SARIMA.

Chapitre 3

Application

Le tableau suivant présente l'évolution du nombre des naissances à l'hôpital de Sidi Abd-elkader de Ouargla depuis *Janvier* 2015 jusqu'à *Août* 2020 :

DATE	2015	2016	2017	2018	2019	2020
Jan	619	620	902	975	907	901
fév	743	730	748	769	715	765
Mars	717	741	826	771	777	741
Avr	644	771	839	889	874	881
Mai	678	722	753	775	798	731
Juin	628	728	806	794	743	837
Juil	740	734	829	835	803	801
Août	739	678	858	760	771	772
Sept	744	787	793	862	746	
Oct	661	667	877	945	901	
Nov	742	783	894	800	809	
Déc	753	615	881	855	856	

TAB. 3.1 – L'évolution du nombre des naissances à l'hôpital de Sidi Abd-elkader de Ouargla (Janvier 2015 - Aout 2020).

3.1 Identification de la série

Pour identifier le type de chronique, nous allons procéder à des tests formels et informels afin de connaître la tendance et la stationnarité.

3.1.1 Les tests informels

L'analyse graphique de la série

L'analyse graphique nous permet de visualiser l'évolution temporelle de la série du nombre des naissances à l'hôpital de Sidi Abd-elkader de Ouargla (*Janvier 2015 - Août 2020*), Ainsi, à partir des données de la série, nous obtenons le graphique suivant :

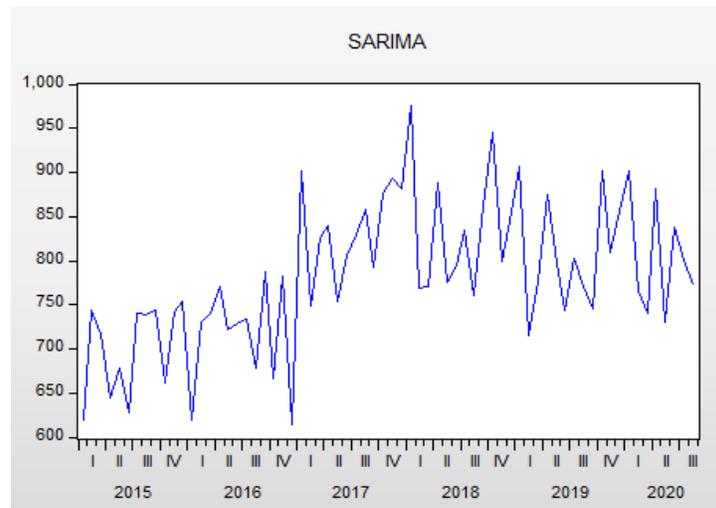


FIG. 3.1 – Graphique de l'évolution du nombre des naissances à l'hôpital de Sidi Abd-elkader de Ouargla (Janvier 2015 - Août 2020)

De ce graphique, nous observons que les données contiennent un élément saisonnier toutes les 12 *périodes*.

et avec :

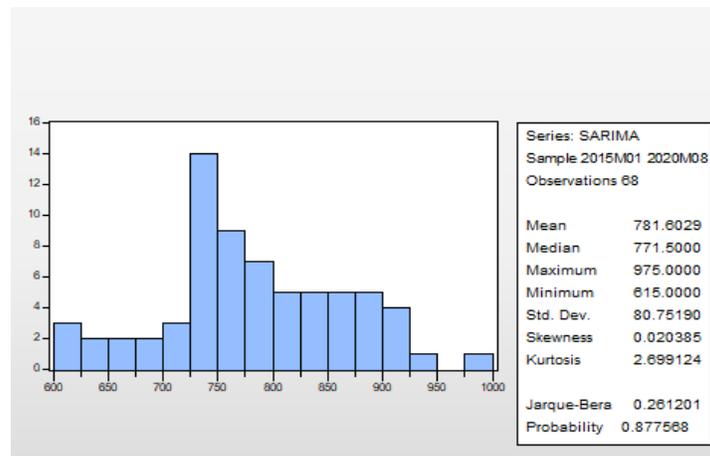


FIG. 3.2 – Histogramme et statistiques.

Ce qui nous permet de tirer quelques conclusions :

- La présomption d'une tendance de la série.
- La série n'est pas stationnaire.

Corrélogramme de la série initiale

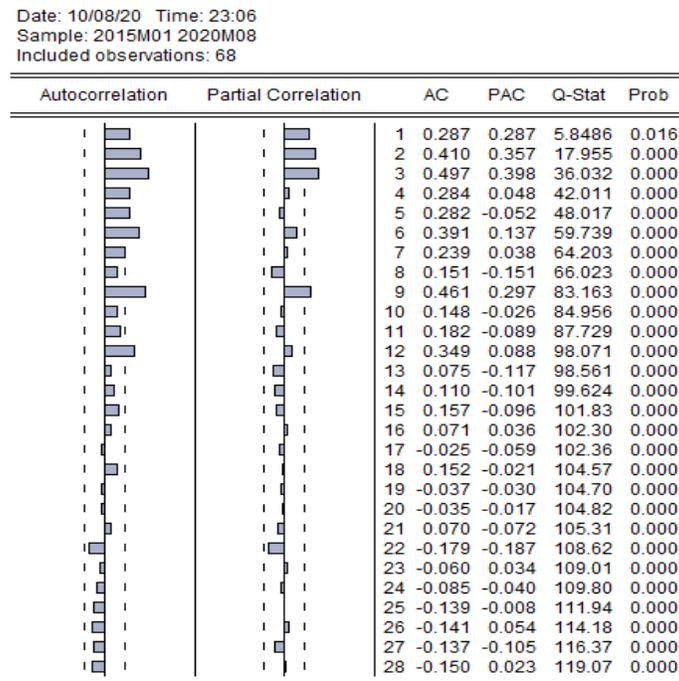


FIG. 3.3 – Corrélogramme de la série SARIMA

✓ Notez qu'il y a des fluctuations et une saisonnalité dans Les coefficients d'autocorrélation simples et coefficient d'autocorrélation partielle, ce qui appuie notre affirmation sur la non stationnarité de la série.

✓ Seulement ces indicateurs sont nécessaires mais non-suffisants.

✓ Par conséquent, nous utilisons le test de Dickey-Fuller Augmenté (ADF) qui est plus performant dans l'étude de la stationnarité des séries.

3.1.2 Tests formels : Test de racine unitaire (ADF)

Ce test a un double objectif :

-Il permet de vérifier la stationnarité d'une série.

-Il donne une idée sur la structure de la série.

Nous avançons les hypothèses suivantes pour notre test :

$$\left\{ \begin{array}{l} H_0 : \rho = 1, \text{ présence de la racine unitaire c\`ad la s\`erie est non stationnaire,} \\ H_1 : \rho \neq 1, \text{ absence de la racine unitaire c\`ad la s\`erie est stationnaire.} \end{array} \right.$$

Ainsi, nous rencontrons les cas de figures ci-après :

-Si $|t_ADF| < |VCM|$, on accepte l'hypothèse nulle, la série est non stationnaire.

-Si $|t_ADF| > |VCM|$, on rejette l'hypothèse nulle, la série est stationnaire. (Avec VCM =Valeur Critique de MacKinnon au seuil de 5%) :

- **Significativité du Trend**

$$\left\{ \begin{array}{l} H_0 : b = 0, \text{ le trend est non significatif,} \\ H_1 : b \neq 0, \text{ le trend est significatif.} \end{array} \right.$$

Ainsi, nous rencontrons les cas de figures ci-après :

-Si $|t - stat| < 2$, on accepte l'hypothèse nulle ,

-Si $|t - stat| > 2$,on rejette l'hypothèse nulle.

- **Significativité de l'intercept**

$$\left\{ \begin{array}{l} H_0 : c = 0, \text{ la constante est non significatif, modèle sans dérive ,} \\ H_1 : c \neq 0, \text{ la constante est significatif, modèle avec dérive.} \end{array} \right.$$

Ainsi, nous rencontrons les cas de figures ci-après :

-Si $|t - stat| < 2$, on accepte l'hypothèse nulle ,

-Si $|t - stat| > 2$,on rejette l'hypothèse nulle.

Null Hypothesis: SARIMA has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 2 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.469081	0.3420
Test critical values:		
1% level	-4.105534	
5% level	-3.480463	
10% level	-3.168039	

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

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(SARIMA)
 Method: Least Squares
 Date: 10/09/20 Time: 20:22
 Sample (adjusted): 2015M04 2020M08
 Included observations: 65 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
SARIMA(-1)	-0.471808	0.191087	-2.469081	0.0164
D(SARIMA(-1))	-0.561718	0.169030	-3.323191	0.0015
D(SARIMA(-2))	-0.351477	0.120859	-2.908167	0.0051
C	343.6900	135.7794	2.531238	0.0140
@TREND("2015M01")	0.832377	0.608649	1.367582	0.1765

R-squared	0.557503	Mean dependent var	0.846154
Adjusted R-squared	0.528003	S.D. dependent var	95.23462
S.E. of regression	65.42811	Akaike info criterion	11.27358
Sum squared resid	256850.2	Schwarz criterion	11.44084
Log likelihood	-361.3915	Hannan-Quinn criter.	11.33958
F-statistic	18.89854	Durbin-Watson stat	1.994787
Prob(F-statistic)	0.000000		

FIG. 3.4 – Test ADF sur premier série.

Il ressort de cet output du test d'ADF, que la série est non stationnaire ($|t_{ADF}| = 2.469081 < |VCM| = 3.480463$, le trend est non significatif ($|t - stat| = 1.367582 < 2$), et l'intercept est significatif ($|t - stat| = 2.531238 > 2$).

3.2 Stationnarisation de la série

•Transformation

On peut conclure que les données ne sont pas normalement distribuées. Par conséquent, il est nécessaire de transformer les données sous forme logarithmique. Puisque les données ont été transformées en forme \ln , vérifiez la stationnarité (en variance), le résultat est le suivant :

Null Hypothesis: LOGSARIMA has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 2 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.543291	0.3071
Test critical values: 1% level	-4.105534	
5% level	-3.480463	
10% level	-3.168039	

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

FIG. 3.5 – Test ADF sur la série LOGSARIMA.

• En utilisant un niveau de signification de 5%, on peut voir que les résultats des données converties ($ADF > t - \text{tabel}$) sous forme logarithmique sont produits pour nous. Cela signifie que les données ne sont toujours pas constantes en termes de variance, ou en d'autres termes, les données ne sont pas normales. Ensuite, l'étape suivante consiste à stabiliser les données sur la moyenne par la première différence étant la variation saisonnière, En utilisant la différence saisonnière, nous trouvons :

Null Hypothesis: DLOGSARIMA 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	-3.908544	0.0183
Test critical values: 1% level	-4.137279	
5% level	-3.495295	
10% level	-3.176618	

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

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(DLOGSARIMA)
 Method: Least Squares
 Date: 10/10/20 Time: 22:46
 Sample (adjusted): 2016M03 2020M08
 Included observations: 54 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLOGSARIMA(-1)	-0.746348	0.190953	-3.908544	0.0003
D(DLOGSARIMA(-1))	-0.287020	0.132991	-2.158187	0.0357
C	0.099787	0.044930	2.220941	0.0309
@TREND("2015M01")	-0.001855	0.000982	-1.889882	0.0646

R-squared	0.565845	Mean dependent var	0.000351
Adjusted R-squared	0.539795	S.D. dependent var	0.150659
S.E. of regression	0.102205	Akaike info criterion	-1.652491
Sum squared resid	0.522290	Schwarz criterion	-1.505159
Log likelihood	48.61726	Hannan-Quinn criter.	-1.595671
F-statistic	21.72205	Durbin-Watson stat	2.019084
Prob(F-statistic)	0.000000		

FIG. 3.6 – Test ADF sur la série DLOGSARIMA.

- On observe que la série est stationnaire ($ADF < t - \text{tabel}$).

La représentation graphique de la série stationnarisée est donnée sur la figure ci-dessous.

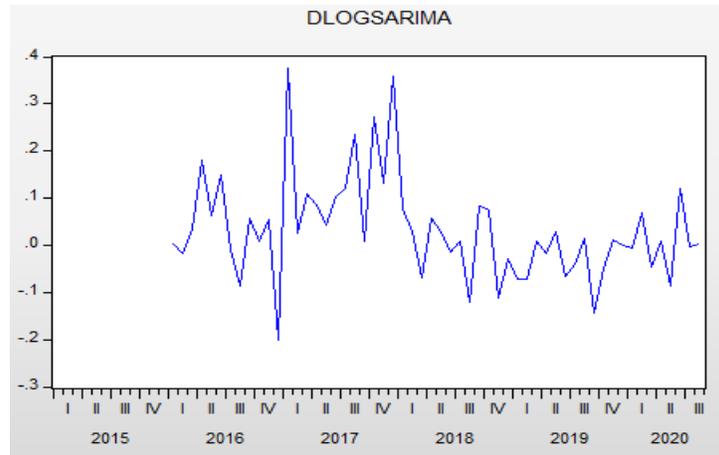


FIG. 3.7 – Graphique de la série stationnaire.

Identification de la série stationnaire

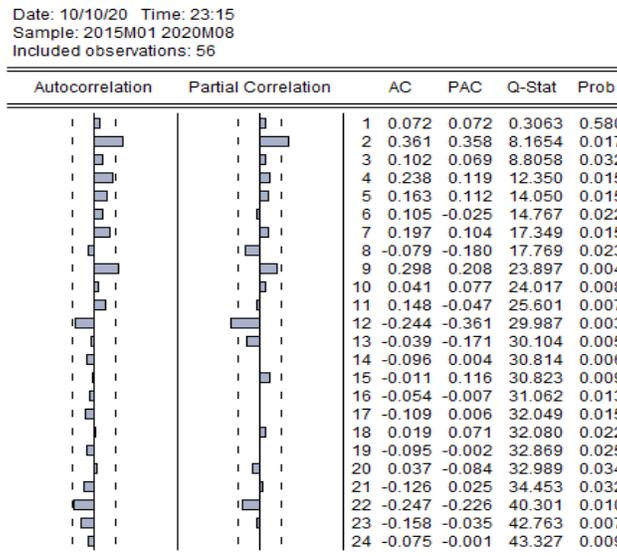


FIG. 3.8 – Correlogramme de la série stationnaire.

3.3 Identification du modèle optimal provisoire

Cette phase consiste à déterminer les modèles adéquats dans la famille SARIMA. Elle est fondée sur l'étude des corrélogrammes simples et partiels de la série étudiée (éventuellement stationnarisée).

Sur la base du diagramme de corrélation ci-dessus, la valeur \mathbf{p} est vue depuis le **PACF**, et \mathbf{q} est vu depuis l'**ACF**. On sait qu'en **ACF** il y a un premier et un deuxième retard ($\mathbf{q} = \mathbf{2}$), en plus de **PACF** il y a un premier retard ($\mathbf{p} = \mathbf{1}$). Ensuite, la valeur \mathbf{P} est vue à partir de la valeur **PACF** saisonnière et de ses multiples, puis dans ce cas, elle est ($\mathbf{P} = \mathbf{1}$). Pendant ce temps, la valeur \mathbf{Q} est obtenue à partir de la valeur saisonnière **ACF** et de ses multiples, de sorte que dans ce cas, elle est en retard de $\mathbf{12}$ ($\mathbf{Q} = \mathbf{0}$). Sur la base de ces informations, le modèle **SARIMA** $(\mathbf{2}, \mathbf{1}, \mathbf{1})(\mathbf{1}, \mathbf{1}, \mathbf{0})^{12}$ est obtenu.

3.4 Estimation

Après avoir identifié le modèle, un modèle important a été identifié, de sorte que le modèle spécifique est **SARIMA** $(\mathbf{1}, \mathbf{1}, \mathbf{2})(\mathbf{1}, \mathbf{1}, \mathbf{0})^{12}$, il convient de passer à la phase d'estimation des paramètres de ce modèle, nous avons :

```

Automatic ARIMA Forecasting
Selected dependent variable: DLOGSARIMA
Date: 10/16/20 Time: 00:28
Sample: 2015M01 2020M08
Included observations: 56
Forecast length: 0
-----
Number of estimated ARMA models: 12
Number of non-converged estimations: 0
Selected ARMA model: (1,2)(1,0)
AIC value: -1.69134138448
-----

```

FIG. 3.9 – le meilleur modèle.

Estimation de l'équation

Dependent Variable: DLOGSARIMA
 Method: ARMA Maximum Likelihood (BFGS)
 Date: 10/08/20 Time: 02:09
 Sample: 2016M01 2020M08
 Included observations: 56
 Convergence achieved after 19 iterations
 Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.035312	0.031709	1.113626	0.2708
AR(1)	0.900446	0.166408	5.411080	0.0000
SAR(12)	-0.411597	0.127810	-3.220380	0.0023
MA(1)	-0.903128	0.230843	-3.912301	0.0003
MA(2)	0.223689	0.147536	1.516164	0.1358
SIGMASQ	0.008319	0.001649	5.044616	0.0000
R-squared	0.282604	Mean dependent var		0.031489
Adjusted R-squared	0.210864	S.D. dependent var		0.108658
S.E. of regression	0.096524	Akaike info criterion		-1.691341
Sum squared resid	0.465845	Schwarz criterion		-1.474339
Log likelihood	53.35756	Hannan-Quinn criter.		-1.607210
F-statistic	3.939295	Durbin-Watson stat		2.048544
Prob(F-statistic)	0.004341			

FIG. 3.10 – Estimer l'équation.

3.5 Validation du modèle

On peut voir dans la sortie ci-dessus, que toutes les valeurs sont toutes significatives. Ensuite, ce modèle est utilisé pour la prévision. Avant de faire des prévisions, il est nécessaire de tester d'abord les hypothèses.

- **Test d'hypothèse classique**
 - a. **Test de normalité résiduelle**

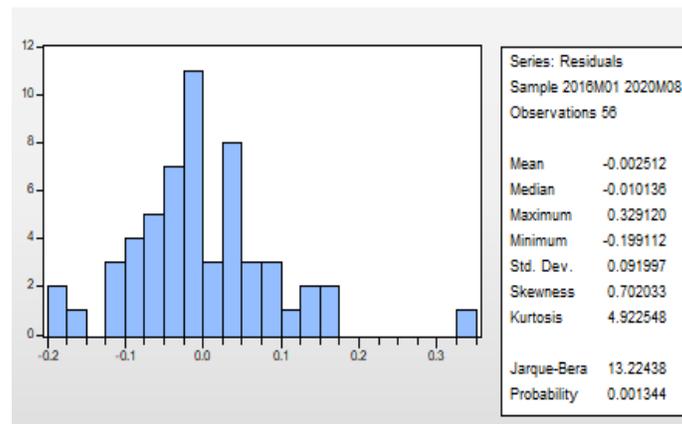


FIG. 3.11 – Test de normalité.

sur la base de l'histogramme ci-dessus, on peut voir que les résidus ont tendance à être normaux.

- b. **Test d'autocorrélation**

sur la base du corrélogramme ci-dessus, on peut voir qu'il n'y a pas de décalage qui a une valeur de probabilité $< 0,05$. On peut donc conclure que le résidu ne contient pas d'autocorrélation.

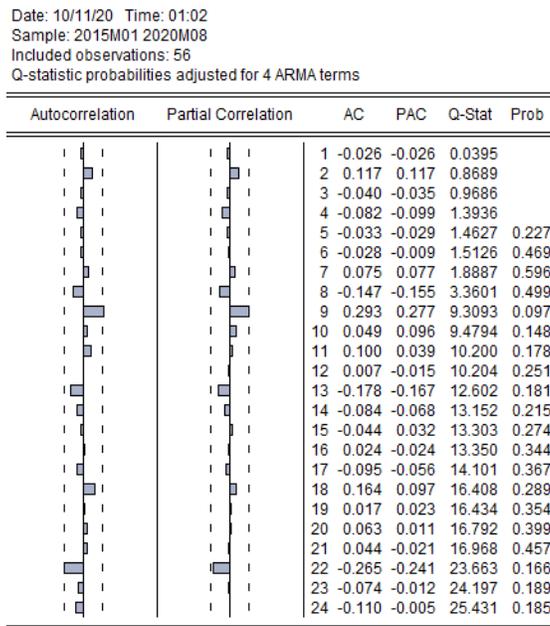


FIG. 3.12 – Correlogramme Q-Statistics.

c. Test d'hétéroscédasticité

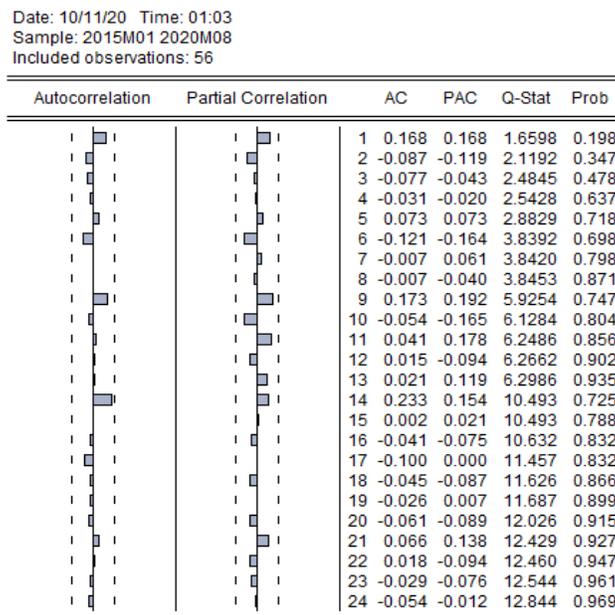


FIG. 3.13 – Correlogram Squared Residuals.

sur la base du corrélogramme ci-dessus, on peut voir que la valeur de probabilité n'est pas $< 0,05$. On peut donc conclure que les données sont l'homoscédasticité.

Une fois que toutes les hypothèses ont été remplies, on peut s'assurer que le modèle $SARIMA(1, 1, 2)(1, 1, 0)^{12}$ est un modèle approprié pour la prévision.

3.6 La prévision

D'après le programme Eviews, les prévisions de (*septembre 2020 – Juin 2021*) sont dans le tableau suivant :

DATE	Sept	Oct	Nov	Dec		
2020	783.9	834.4	804.4	819.7		
DATE	Jan	Fév	Mars	Avril	Mai	Juin
2021	834.6	790.1	782.2	827.9	779.0	813.5

TAB. 3.2 – Prévision du nombre de naissances.

Conclusion

Au terme de cette étude sur la modélisation des séries chronologiques, par l'utilisation des méthodes de Box et Jenkins, nous avons épousé volontairement une approche pragmatique de Box Jenkins généralisé par les processus SARIMA. L'hypothèse de stationnarité est fondamentale dans la modélisation d'une série temporelle. Elle s'avère être un outil puissant et performant pour faire de la prévision.

Le but de notre étude est de trouver le meilleur modèle pour prédire le nombre des naissances à l'hôpital de Sidi Abd-elkader de Ouargla(*Janvier 2015 - Août 2020*) ,les résultats de l'analyse de notre série montre que le modèle approprié pour cette étude est SARIMA(1,1,2)(1,1,0)¹² . Finalement, nous obtenons les prévisions du nombre des naissances (*Septembre 2020 - Juin 2021*) .

Bibliographie

- [1] Agnès Lagnoux .Séries Chronologiques . Universite Le DE TOULOUSE MI-RAIL.
- [2] Anderson O.D. (1976) : "Time Series Analysis and Forecasting : The BOX JENKINS Approach", Butterworth, London.
- [3] ANDERSON, T. W. (1994) The Statistical Analysis of Time Series New York, J. Wiley.
- [4] Alvian Imron Rosadi, "PERAMALAN JUMLAH PENUMPANG KERETA API DENGAN METODE SARIMA (SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE)", Université Islam Indonesia.
- [5] Chaido Dritsaki, Forecast of Sarima Models : An Application to Unemployment Rates of Greece American Journal of Applied Mathematics and Statistics, 2016, Vol. 4, No. 5, 136-148.
- [6] Chawsheen, T.A. & Broom, M. (2017). Seasonal time-series modeling and forecasting of monthly mean temperature for decision making in the Kurdistan Region of Iraq. Journal of Statistical Theory and Practice. University of London Institutional Repository.
- [7] Dégerine, S. (2005). Cours de série chronologique. Université Joseph Fourier.
- [8] Eric. Moulines et Fran cois Roue : Analyse des Séries Temporelles et Applications, 15 septembre 2010.
- [9] Florenece, N. (2006). Généralités sur les séries chronologiques.
- [10] Girard, Y. (2011). Série chronologiques à une et plusieurs variables : synthèse des méthodes classiques et modèles à base de copules. Université du Québec.
- [11] Guy.MELARD. (1990). Méthodes de Prévision I (Stat-D-205). Université Libre de Bruxelles.
- [12] HANANE IDRISSE. Prévisions par l'utilisation de série chronologique et la régression linéaire simple .UNIVERSITE SIDI MOHAMED BEN ABDELLAH .
- [13] Jaiw, C. (2015). Séries chronologiques saisonnières et prévision.

-
- [14] J. C. Ramesh Reddy, T. Ganesh, M. Venkateswaran³, PRS Reddy Forecasting of Monthly Mean Rainfall in Coastal Andhra International Journal of Statistics and Applications 2017, 7(4) : 197-204.
 - [15] Lévy, C.L. (2013). Introduction à l'étude des séries temporelles.
 - [16] Marine, E. (2015). Centre de Formation aux Métier de la Montagne.
 - [17] Ralf Vandenhouten : Non-stationary Time Series Analysis of Complex Systems and Applications in Physiology, Aachen, Shaker Verlag GmbH 1998, (ISBN 3-8265-3814-5).
 - [18] Serigne, M. (2007). Analyse et prévision des série temporelles par la méthode de Box & Jenkins.

Résumé

Les modèles de séries temporelles mixtes saisonnières SARIMA sont les modèles les plus importants de la série temporelle linéaire aléatoire. Il prend en compte l'instabilité et les variations saisonnières. L'objectif de ce travail est de tester l'exactitude de ces modèles dans l'estimation de la série mensuelle du nombre des naissances à l'hôpital de Sidi Abdelkader de la mère et de l'enfant de janvier 2015 à août 2020. puis de tester leur exactitude dans la prédiction du nombre futur de naissances futures pendant dix mois, de septembre 2020 à juin 2021. et connaître l'exactitude de Modèles SARIMA .

Mots clés: Modèles SARIMA, Séries chronologiques, Méthode de Box-Jenkins.

Abstract

The seasonal mixed time series models SARIMA are the most important models of the random linear time series .it is taking in to account the instability and the Seasonal variations .The objective of this work is to test the accuracy of those models in the estimation of the monthly number of births series at Sidi Abdelkader Hospital of Mother and Child from January 2015 to August 2020. and then testing their accuracy in predicting of the future numbers of future births for ten months, from September 2020 to June 2021. and know the accuracy of SARIMA models .

Keywords: SARIMA models, Time series, Box-Jenkins method.

ملخص

تعتبر نماذج السلاسل الزمنية الموسمية المخلطة SARIMA من أهم نماذج السلاسل الزمنية العشوائية الخطية فهي تأخذ بعين الاعتبار عدم الاستقرار والتغيرات الفصلية في السلسلة الزمنية، حيث أردنا من خلال هذه المذكرة اختبار دقة هذه النماذج في تقدير سلسلة عدد المواليد الشهرية لمستشفى سيدي عبد القادر للام والطفل من جانفي 2015 إلى أوت 2020، و من ثم اختبار دقتها في التنبؤ بعدد المواليد المستقبلية لعشر أشهر القادمة من سبتمبر 2020 إلى جوان 2021، من أجل استنتاج مدى دقة نماذج SARIMA .

الكلمات المفتاحية: نموذج SARIMA, سلاسل زمنية, طريقة بوكس-جنكينز.