



N° Réf :.....

Centre Universitaire de Mila

Institut des Sciences et de la Technologie

Département de Mathématique et Informatique

Mémoire préparé En vue de l'obtention du diplôme de licence

En : - Filière : Mathématiques Fondamentales

Thème

Estimation de l'ARIMA avec méthode de vraisemblance

Préparé par :

**Boumenakh Ameul
Mouchmouche Aicha
Belmehboul Rofia
Hamoudi Mounya**

Encadré par :

BOULAROUK YAKOUB

Année universitaire : 2013/2014

Remerciements

Nous souhaitons la réussite à tous les étudiants des sections En premier lieu, on tenant compte à témoigner nos reconnaissances à dieu

Tout puissant, de nos avoir donné la possibilité de terminer ce travail

On adresse nos chaleureux remerciements à notre encadreur de mémoire Mr BOULAROUK YAKOUB, enseignant au C, U de MILLA pour son attention de tout instant sur nos travaux, pour ses conseils avisés et son écoute qui on été prépondérants pour la bonne réussite de ce mémoire.

Son énergie et sa confiance ont été des éléments moteurs pour nous on' a pris un grand plaisir à travailler avec lui.

Il est important pour nous de remercions nos famille nos très chers Parents, frères, sœurs, collègues et amis, qui on toujours été une source

Inépuisable d'encouragements.

Et la plus importante, nous remercions aussi tout éducateurs, enseignants

Et nous orienté sur

Le chemin de la connaissance et du savoir depuis le cycle primaire Jusqu'au cycle universitaire.

Un grand merci à tous ceux qui ont participé de près ou de loin à L'abonnissement de ce travail

Nous souhaitons la réussite à tous les étudiants des sections Mathématique et Informatique.

Table des matières

Introduction	3
1 Les variables aléatoires et les séries Chronologiques	4
1.1 Les variables aléatoires :	5
1.1.1 Variables aléatoires discrètes :	5
1.1.2 Variables aléatoires continues :	5
1.1.3 Fonction densité de probabilité :	6
1.1.4 Fonction de répartition :	6
1.1.5 Loi de probabilité :	7
1.1.6 Espérance et variance :	7
1.1.7 Couples des variables aléatoires :	9
1.2 Loi de laplace (probabilités) :	10
1.2.1 Densité de probabilités :	10
1.2.2 Fonction de répartition :	11
1.3 Processus Stochastique :	12
1.4 Série Chronologique :	14
1.4.1 Description d'une série chronologique :	14
2 les modèles et les fonctions de vraisemblance	17
2.1 Le modèle linéaire :	18
2.1.1 Le modèle de régression linéaire :	18

2.1.2	Modèle linéaire gaussien :	19
2.1.3	Fonction de vraisemblance du modèle linéaire :	20
2.2	Modèle ARMA :	21
2.2.1	Processus auto-regressif AR(p) :	21
2.2.2	Processus a moyenne mobile MA(q) :	23
2.2.3	Processus auto-régressif moyenne mobile ARMA (p,q) :	23
2.2.4	Les fonctions d'auto-correlatin et d'auto-covariance :	25
2.3	Vecteurs Gaussien :	30
2.4	Fonction de vraisemblance du modèle ARMA :	31
2.4.1	La fonction de vraisemblance inconditionnelle :	31
	Bibliographie	33

Introduction

La notion de variable aléatoire (v.a) est au coeur des probabilités et statistiques .La régression l'étude des séries chronologiques , l'analyse de données, les processus stochastiques , ou bien encore les tests d'hypothèses sont autant de ramifications des méthodes statistiques ayant pour socle l'étude des variables aléatoires .Malheureusement la définition moderne d'une v.a. ne peut être exposée rigoureusement sans faire appel a la théorie de la mesure et de l'intégration au sens de Lebergue .

Depuis toujours, une méthode de prévision très populaire est basée sur l'étude rigoureuse de série chronologique et un processus, cette approche permet de prédire ,par exemple (de nombreux phénomènes naturels et financiers) une série chronologique est constitué de valeurs observées à des intervalles de temps réguliers et cette prévision dépend fortement de la qualité du modèle choisi .

Ici on étudient les modèles (ARMA , linéaire) qui peuvent donner une bonne représentation de certaines séries chronologique , processus avec l'utilisation d'un vecteur aléatoire (Gaussien) permet donner la présentation du fonction de vraisemblance .

Chapitre 1

Les variables aléatoires et les séries Chronologiques

1.1 Les variables aléatoires :

on appelle variables aléatoires tout nombres réel aléatoire. c'est-a'-dire dont la valeur dépend du résultat d'une expérience probabilité.

1.1.1 Variables aléatoires discrètes :

une variable aléatoire est dite discrète si elle ne prend que des valeurs discontinues dans un intervalle donné (borné ou non borné) ,l'ensemble des nombres entiers est discrète ,en règle générale ,toutes les variables qui résultent d'un dénombrement ou d'une numération sont de type discrètes.

les variables aléatoires discrètes :

-le nombre de petits par portée pour une espèce animale donnée (chat,marmotte,...)

-la nombre de bactéries dans 100 ml de préparation.

-le nombre de mutations dans une séquence d'ADN de 10kb.

sont des variables aléatoires discrètes.

1.1.2 Variables aléatoires continues :

une variable aléatoire est dite continue si elle peut prendre toutes les valeurs dans un intervalle donné (borné ou non borné).en règle générale, toutes les variables qui résultent d'une mesure sont de type continue.

des variables aléatoires continue :

-le masse corporelle des individus pour une espèce animale donnée.

-taux de glucose dans le sang.

sont des variables aléatoires continues.

1.1.3 Fonction densité de probabilité :

on appelle densité de probabilité toute application continue par morceaux :

$$f : \mathbb{R} \rightarrow \mathbb{R}$$
$$x \rightarrow f(x)$$

tq :

i- $\forall x \in \mathbb{R} \quad f(x) \geq 0$

ii- $\int_{-\infty}^{+\infty} f(x) dx = 1$

1.1.4 Fonction de répartition :

Si comme pour les variables aléatoires discrète

on peut définir la fonction de répartition de x

$$F_x : \mathbb{R} \rightarrow \mathbb{R}$$
$$t \rightarrow F_x(t) = p(x \prec t)$$

alors la relation entre la fonction de répartition F_x et la fonction densité de probabilité $f(x)$ est la suivante :

$$\forall t \in \mathbb{R}. F_x(t) = p(x \prec t) = \int_{-\infty}^t f(x) dx$$

1.1.5 Loi de probabilité :

une variable aléatoire est caractérisée par l'ensemble des valeurs qu'elle peut prendre et par l'expression mathématique de la probabilité de ces valeurs, cette expression s'appelle la loi de probabilité (ou distribution de probabilité) de la variable aléatoire.

La loi de probabilité d'une variable aléatoire discrète :

est entièrement déterminée par les probabilités p_i des événements $\{X = x_i\}$ la loi de probabilité est donnée par le (x_i, p_i) , $p(X = x_i)$. Il y a aussi plusieurs lois de probabilité sont : loi binomiale, bernoulli, uniforme, géométrie et poisson et la loi normal....

1.1.6 Espérance et variance :

Espérance mathématique :

l'espérance d'une variable aléatoire $E(X)$ correspond à la moyenne des valeurs possibles de X pondérées par les probabilités associées à ces valeurs, c'est un paramètre de position qui correspond au moment d'ordre 1 de la variable aléatoire X c'est l'équivalent de moyenne.

Espérance d'une variable aléatoire discrète : X est une variable aléatoire discrète de loi de probabilité (x_i, p_i) définie sur un nombre fini d'événements élémentaires alors :

$$E(X) = \sum_{i=1}^n x_i p_i$$

Espérance d'une variable aléatoire continue : X est une variable aléatoire absolument continue de densité f , on appelle espérance de X , le réel $E(X)$, défini par :

$$E(X) = \int_{-\infty}^{+\infty} X f(x) dx$$

si cette intégrale est convergente.

Propriétés de l'espérance : si X et Y sont deux variables aléatoires admettant une espérance alors :

$$i-\mathbf{E}(\mathbf{X} + \mathbf{Y}) = \mathbf{E}(\mathbf{X}) + \mathbf{E}(\mathbf{Y})$$

$$ii-\mathbf{E}(\mathbf{aX}) = \mathbf{aE}(\mathbf{X}), \forall a \in \mathbb{R}$$

$$iii\text{-si } X \geq 0 \text{ alors } E(X) \geq 0$$

Variance :

la variance d'une variable aléatoire $V(X)$ est l'espérance mathématique du carré de l'écart à l'espérance mathématique c'est un paramètre de dispersion qui correspond au moment centre d'ordre 2 de la variable aléatoire X .

X est une variable aléatoire ayant une espérance $E(X)$, on appelle variance de X le réel : $V(X) = E(X^2) - E(X)^2$

on appelle écart-type de X , le réel $\delta(X) = \sqrt{V(X)}$

La variance d'une variable aléatoire discrète : X est une variable aléatoire discrète de loi de probabilité (X_i, p_i) définie sur un nombre fini (n) d'événements élémentaires alors la variance est égale à :

$$V(X) = \sum_{i=1}^n (X_i - E(X))^2 p_i = \sum_{i=1}^n X_i^2 p_i - E(X)^2$$

La variance d'un variable aléatoire continue : X est une variable aléatoire continue donnée par sa densité de probabilité alors la variance de X et le nombre réel positive tel que :

$$V(X) = \int_{-\infty}^{+\infty} (X - E(X))^2 f(x) dx = \int_{-\infty}^{+\infty} x^2 f(x) dx - E(X)^2$$

Propriétés de la variance : X est une variable aléatoire admettant une variance alors :

$$i-\forall a \in \mathbb{R}, V(aX) = a^2 V(X)$$

$$ii-\forall (a, b) \in \mathbb{R}, V(aX + b) = a^2 V(X) \quad iii-V(X) = 0 \iff X = E(X)$$

1.1.7 Couples des variables aléatoires :

Loi jointe :

les définition portant sur la loi jointe entre deux variable aléatoire X et Y impliquent que ces dernières soient définies sur le même espace fondamental Ω . si X et Y sont définies respectivement sur les espaces fondamentaux Ω_1 et Ω_2 alors, il faut envisager un espace qui englobe Ω_1 et Ω_2 appelle (espace-produit)

L'indépendance entre variables aléatoires :

La espérance : X et Y sont deux variables aléatoires indépendantes définies sur le même univers Ω alors :

$$E(XY) = E(X)E(Y)$$

La variance : X et Y sont deux variables aléatoires indépendantes définies sur le même univers Ω alors :

$$V(X + Y) = V(X) + V(Y)$$

Covariance et corrélation : X et Y sont deux variables aléatoires définies sur le même univers Ω , on appelle covariance de ces deux variables, le réel :

$$cov(X, Y) = E(XY) - E(X)E(Y)$$

et coefficient de corrélation, le réel :

$$\rho(X, Y) = \frac{cov(X, Y)}{\delta(X)\delta(Y)}$$

• X et Y sont deux variables aléatoires définies sur le même univers Ω et indépendantes :

$$cov(X, Y) = 0$$

1.2 Loi de laplace (probabilités) :

Dans la théorie des probabilités et en statistique , la loi(distribution)de laplace ,est une densité de probabilité continues nommée d'après pierre-simon de laplace , on connaît aussi sous le nom de loi double exponentielle , car sa densité peut être vue comme l'association des densité de deux lois exponentielle accolées des à dos , la loi de laplace s'obtient aussi comme le résultat de la difference de deux variables exponentielle indépendantes.

1.2.1 Densité de probabilités :

une variable aléatoire possède une distribution laplace (μ, b) si sa densité de probabilité est :

$$\begin{aligned} f\left(\frac{x}{\mu}, b\right) &= \frac{1}{2b} \exp\left(-\frac{|x-\mu|}{b}\right) \\ &= \frac{1}{2b} \left\{ \begin{array}{ll} \exp\left(-\frac{\mu-x}{b}\right) & \text{si } x < \mu \\ \exp\left(-\frac{x-\mu}{b}\right) & \text{si } x \geq \mu \end{array} \right\} \end{aligned}$$

le réel μ est un paramètre de position et $b > 0$ un paramètre d'échelle . Si $\mu = 0$ et $b = 1$, la loi de laplace est dite standart et sa restriction à la dernie-droite réel positive est la loi exponentielle de paramètre $\frac{1}{2}$.

la densité rappelle aussi de la loi normale :toute fois tandis que la loi normal est exprimée en termes de la difference au carré $(x - \mu)^2$, la loi de laplace fai tintervenir la différence absolue $|x - \mu|$,la loi de laplace présente alors des queues plus epaisses que la loi normale.

1.2.2 Fonction de répartition :

la densité de la loi de laplace s'intégraisément grâce à la prèssence de la valeur absolue et sa fonction de répartition est :

$$F(x) = \int_{-\infty}^x f(\mu) d\mu$$

alors :

$$F(x) = \begin{cases} \frac{1}{2} \exp\left(-\frac{\mu-x}{b}\right) & \text{si } x < \mu \\ 1 - \frac{1}{2} \exp\left(-\frac{-x-\mu}{b}\right) & \text{si } x \geq \mu \end{cases}$$
$$= 0.5 \left[1 + \operatorname{sgn}(x - \mu) \left(1 - \exp\left(-\left|\frac{x-\mu}{b}\right|\right) \right) \right]$$

la réciproque de la fonction de répartition est :

$$F^{-1}(p) = \mu - b \operatorname{sgn}(p - 0.5) \ln(1 - 2|p - 0.5|).$$

1.3 Processus Stochastique :

Definition 1 un processus stochastique est une famille $\{X_t\}_{t \in T}$ de variables aléatoires indexées par le temps t , les mots processus et stochastique signifient respectivement fonction et aléatoire. alors qu'une variable aléatoire X associée à chaque $w \in \Omega$ une réalisation $X(w)$, une processus stochastique $\{X_t\}_{t \in T}$ associée à chaque w une fonction $\{X_t\}_{t \in T}$:

$$\begin{aligned} T &\rightarrow E \\ t &\rightarrow X_t(w) \end{aligned}$$

$t \in T$ l'ensemble des temps.

lorsque l'ensemble temps T est au plus dénombrable (par exemple $T = \mathbb{N}$), on parle de processus stochastique à temps discret, lorsqu'il est continu (ie $T = [0, t_0]$ ou $T = \mathbb{R}_+$) on parle processus stochastique à temps continu.

les situations réelles pouvant être modélisées par des processus stochastiques sont nombreuses, en voici quelques exemples :

-Problème de la ruine de joueur : considérons une suite de variable aléatoire $(Y_n)_{n \geq 1}$ dont la loi comme est définie par $P(Y_1 = 1) = p$ et $P(Y_1 = -1) = 1 - p$, et une quantité $Y_0 \in \mathbb{Z}$ indépendante de variable aléatoire, on définit la marche aléatoire simple par :

$$\begin{aligned} X_{n+1} &= X_n + Y_{n+1} \\ &= Y_0 + Y_1 + \dots + Y_n + Y_{n+1} \end{aligned}$$

-Problème de l'extinction d'une population : considérons une suite doublement indexée de variable aléatoire $\{Y_{n,m}, n \in \mathbb{N}, m \in \mathbb{N}^*\}$ à valeurs entières, la variable $Y_{n,m}$ représente le nombre de fils du m^e individu dans la n^e génération posons $X_0 = 1$, il y a initialement un seul individu, puis pour tout n :

$$X_{n+1} = \sum_{m=1}^{X_n} Y_{n,m}$$

-Séries temporelles ou chronologiques : Les séries temporelles sont des processus stochastiques. Elles peuvent illustrer le nombre de morts suite à des accidents de la route un pays donné durant un intervalle de temps.

-File d'attente : la salle de réservation d'une grande gare SNTV donne une bonne représentation d'une file d'attente elle comprend un certain nombre de guichets et des clients qui sont soit en train d'être servis ,soit en attente qu'un guichet se libère .

Le nombre total de ces clients présent dans la salle de réservation au temps t et noté N_t .La suite $(N_t)_{t \geq 0}$ est un processus stochastique à temps continue et à valeurs dans $E = N$, l'objectif est d'étudier l'évolution N_t au cours du temps afin d'optimiser le nombre de guichets nécessaires pour satisfaire en un temps raisonnables les clients.

1.4 Série Chronologique :

La théorie des séries chronologique (ou temporelles) abordées dans ce exposé est appliqué de nos jours dans des domaines aussi variés que l'économétrie ,la médecine ou la démographie , pour n'eniter qu'une petite partie ,ou s'intéresse à l'évolution au cours du temps d'un phénomène , dans le but décrire expliquer puis prévoir ce phénomène dans le future ,on dispose aissi d'observations à des dates différentes c'est-à-dire d'une suite de valeurs numériques indicées par le temps.

1.4.1 Description d'une série chronologique :

on considère qu'une série chronologique (X_t) est la resultante de différents composantes fondamentales :

-La tandance (Z_t) représente l'évolution à long terme de la série étudiée ,elle traduit le comportement "moyen" de la série .

-La composante saisonnière (ou saisonnalité) (S_t) correspond à un phénomène qui se répète à intervalles de temps réguliers (périodiques) . En générale c'est un phénomène saisonnier d'ou le terme de variations saisnnières .

-la composante résiduelle (ou pruit ou résidu) (ε_t) correspond à des fluctuations irréguliérs , en générale de faiblsaible intensité mais de nature aléatoire , on parle aussi d'aléas.

-Des phénomènes accidentels (grèves , conditions météorologiques exceptionnelles , crash financier) peuvent notamment intervenir .

Objectifs principaux :

L'étude d'une série chronologique permet d'analyser , de décrire et d'expliquer un phénomène au cours du temps et d'en tirer des conséquences pour des prises de décisions (marketing).cette étude permet aussi de faire un contrôle .

Description schématique de l'étude complétée d'une série chronologique :

Comme nous venons de le voir l'un des objectifs principaux de l'étude d'une série chronologique est la prévision des valeurs futures de cette série ,pour cela ,on a besoin de connaître ou tout au moins de modéliser le mécanisme de production de la série chronologique .

Notons que les variables X_t ne sont le plus souvent ni indépendantes(on peut s'attendre en effet à ce que des observations relativement proches dans le temps soient liées) ni indépendamment distribuées (dans la plupart des cas , le phénomène évolue ,se modifie au cours du temps ce qui entraîne que les variables de description ne sont pas équidistribuées),cela nécessite des méthodes statistiques de traitement et de modélisation spécifiques puisqu'en particulier dans un cadre standard (celui de la description d'un échantillon) les méthodes statistiques classiques sont basées sur des hypothèses d'indépendance .

Schématiquement , les principales étapes de traitement d'une série chronologique sont les suivantes :

Connexion des données :

Avant de se lancer dans l'étude d'une série chronologique ,il est souvent nécessaire de traiter ,modifier les données brutes par exemple :

- Évaluation de données manquantes ,remplacement de données accidentelles
- découpage en sous-série.

Observation de la série :

Une règle générale en statistique descriptive consiste à commencer par regarder les données avant d'effectuer le moindre calcul .

Modélisation :

Un modèle est un image simplifiée de la réalité qui vise à traduire les mécanisme de fonctionnement du phénomène étudié et permet de mieux les comprendre .

Chapitre 2

les modèles et les fonctions de vraisemblance

2.1 Le modèle linéaire :

un modèle linéaire exprime une variable aléatoire Y linéairement en fonction d'une variable explicative X selon la formule : $Y_i = \sum B_j X_j + U_i$ telle que Y_i les variables aléatoires réelle réponse et U_i est la variable aléatoire réel erreur , les U_i étant indépendants , les B_j sont coefficients des paramètres inconnus à estimer les X_j sont les valeurs explicatives.

2.1.1 Le modèle de régression linéaire :

On cherche à modéliser une variable quantitative Y en fonction de variables explicatives quantitatives X^1, X^2, \dots, X^p sous l'hypothèse gaussienne , le modèle de régression linéaire s'écrit :

$$y_i = b_0 + b_1 x_i^1 + \dots + b_p x_i^p + e_i$$

avec b_0, b_1, \dots, b_p inconnus et e_1, \dots, e_n observations indépendantes d'une loi $N(0, \delta^2)$ avec δ^2 inconnue.

Le modèle de régression linéaire simple :

La relation entre y_i et x_i s'écrit sous la forme d'un modèle de régression simple :

$$y_i = b_0 + b_1 x_i + e_i \quad , \quad \forall i = 1, \dots, n$$

Le modèle de régression linéaire multiple :

aussi de la façon suivante : le modèle s'écrit :

$$y_i = b_0 + b_1 x_i^1 + \dots + b_p x_i^p + e_i \quad i = 1, \dots, n$$

2.1.2 Modèle linéaire gaussien :

C'est un modèle linéaire dans lequel on fait l'hypothèse supplémentaire que la variable aléatoire réelle est gaussienne c'est-à-dire normale on pose donc $U \sim N_n(\theta, \delta_{In}^2)$ et $Y \sim N_n(XB, \delta_{In}^2)$ entraînant la normalité de Y

ona :

$$Y_i = \sum_{j=1}^p B_j X_i^j + U_i \quad (\text{égalité entre variable aléatoire réel})$$

$$y_i = \sum_{j=1}^p b_j x_i^j + u_i \quad (\text{égalité entre nombre réel})$$

2.1.3 Fonction de vraisemblance du modèle linéaire :

Estimation par maximum de vraisemblance :

On se place dans cadre paramétrique la loi jointe d'un vecteur $Y' = (y_1, \dots, y_n)'$ possède une densité de probabilité $f(y, \theta)$, θ vecteur de paramètre à estimer, appartient à un ensemble ouvert, borné Θ de \mathbb{R}^k

Le principe : On observe une réalisation du vecteur y de taille n :

$$y = (y_1, \dots, y_n)$$

On appelle fonction de vraisemblance la fonction de y et de θ :

$$V(y, \theta) = V(y_1, \dots, y_n, \theta) = f(y_1, \dots, y_n, \theta)$$

On appelle estimateur du maximum de vraisemblance, une valeur $\widehat{\theta}(y)$ solution du problème de maximisation : $\max_{\theta \in \Theta} V(y_1, \dots, y_n, \theta)$

une transformation strictement croissante ne changeant pas un maximum (et la tradition gaussienne), on considère souvent le logarithme népérien de la vraisemblance,

d'où une forme équivalente $\max \ln(V(y_1, \dots, y_n, \theta))$

On notera dans la suite cette fonction à maximum $LV(y, \theta)$ ou lorsque qu'il n'y a pas d'ambiguïté, y fait référence à une réalisation de taille n .

Dans le cas particuliers d'un échantillon de taille n : $V(y_1, \dots, y_n, \theta) = \prod_{i=1}^n f(y_i, \theta)$

$$LV(y_1, \dots, y_n, \theta) = \sum_{i=1}^n \ln(f(y_i, \theta))$$

Estimation :

θ est les vecteurs des paramètre à estimer dans le cas général o est un vecteur à k composantes : $\theta_1, \theta_2, \dots, \theta_k$ on note :

Y : la variable aléatoire à expliquer, y une réalisation de cette v, a Y .

θ : la vraie valeur théorique du vecteur des paramètre du modèle.

$\widehat{\theta}$: l'estimateur de θ .

$\widehat{\theta}(y)$: une réalisation de v, a θ (une estimation de θ à partir des données observées).

2.2 Modèle ARMA :

Les modèles ARMA (modèle autorégressifs et moyenne mobile) ou modèle de BOX-JenKins, sont les principaux modèles de séries temporelles. Étant donnée une série temporelle X_t , le modèle ARMA est un outil pour comprendre les valeurs futures de cette série, le modèle est composé de deux processus : processus à auto-régressif AR et processus moyenne mobile MA.

2.2.1 Processus auto-régressif AR(p) :

AR d'ordre 1 :

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

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

où la série α_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.

Fonction de transfert :

on peut écrire :

$$\begin{aligned}X_t &= \alpha_t + \phi X_{t-1} \\ X_t &= \alpha_t + \phi \alpha_{t-1} + \phi^2 X_{t-2} \\ X_t &= \alpha_t + \phi \alpha_{t-1} + \phi^2 X_{t-2} + \dots + \phi^k \alpha_{t-k} + \dots \\ X_t &= \sum_{i=0}^{+\infty} \phi^i \alpha_{t-i} \\ X &= \sum_{i=0}^{+\infty} \phi^i B^i \alpha_t = \psi(B) \alpha_t\end{aligned}$$

la fonction de transfert a donc une infinité de termes si on revient à notre écriture formelle sous forme de polynôme

on remarque qu'on a calculé l'inverse du polynôme $(1 - \phi B)$ et que :

$$(1 - \phi B)^{-1} = \sum_{i=0}^{+\infty} \phi^i B^i$$

—variance de X par calculs simples , on peut montrer que :

$$Var(X) = \frac{\delta_\alpha^2}{1-\phi^2}$$

ici aussi , on diminue la variance en modélisant à condition que ϕ soit en valeur absolue inférieure à 1 .

—autocorrélation de X de même on peut montrer que :

$$P(k) = \phi^k \text{ pour } k \geq 1$$

on peut obtenir deux sortes de corrélogramme suivant : si ϕ est positif ou négatif.

Stationnarité :

un processus AR(1) est stationnaire si et seulement si la valeur absolue de ϕ est inférieure à 1

Processus autoregressif AR d'ordre p :

une série X suit un processus autoregressif d'ordre p (AR(p)) si on peut écrire :

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \alpha_t$$

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \alpha_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = \alpha_t$$

$$\phi(B) X_t = \alpha_t$$

où la série α_t est un bruit blanc ici encore , on fait une régression des p série décalées sur la série elle-même .

Autocorrélation pour un processus AR(p) on ne peut rien dire de sa fonction ACF , si ce n'est que :

$$P(k) \neq 0 \text{ pour tout } k .$$

C'est pour pouvoir reconnaître un processus autoregressif qu'on a introduit la fonction d'auto-correlation partielle.

2.2.2 Processus a moyenne mobile MA(q) :

En anglais , moyenne mobile se dit " moving average" c'est pourquoi nous dirons que ces séries sont des MA .

MA d'ordre 1 :

On dit que la série X_t suit un processus de moyenne mobile d'ordre (MA(1)) si elle est générée par un bruit blanc α_t sous la forme :

$$\begin{aligned} X_t &= \delta_t - \theta\alpha_{t-1} \\ X_t &= (1 - \theta B)\alpha_t \end{aligned}$$

La fonction de transfert du filtre se réduit à un seul terme. Variance de X par calculs simples , on peut montrer :

$$var(X) = (1 + \theta^2)\delta_\alpha^2, var(X) \succ \delta_\alpha^2$$

c'est-à-dire qu'en modélisant , on diminue la variance du phénomène , par nature le propre de toute modélisation et autocorrélation de X de même , on peut montrer que :

$$P(k) = \begin{cases} \frac{-\theta}{1+\theta^2} & \text{si } K + 1 \\ 0 & \text{si } K \geq 2 \end{cases}$$

MA d'ordre q :

On suppose que la série X est générée par un bruit blanc α sous la forme :

$$\begin{aligned} X_t &= \alpha_t - \theta_1\alpha_{t-1} - \theta_2\alpha_{t-2} - \dots - \theta_q\alpha_{t-q} \\ X_t &= \theta(B)\alpha_t \end{aligned}$$

Auto-corrélation de X on peut montrer que :

$$P(k) = \begin{cases} \neq 0 & \text{si } k \leq q \\ = 0 & \text{si } k \succ q \end{cases}$$

2.2.3 Processus auto-régressif moyenne mobile ARMA (p,q) :

Pour mieux représenter le processus stochastique il est parfois utile de combiner les deux processus AR(p) (auto-régressif) et MA(q)(moyenne mobile) ce que donne le processus

autorégressif moyenne mobile ARMA(p,q).

Donc pour que le processus X_t soit un ARMA(p,q) il faut il suffit qu'il existe deux polynôme ϕ , θ et un bruit blanc ε_t , de sorte que

$$\phi(B)X_t = \theta(B)\delta_t$$

autrement dit si :

$$\theta(B) = \sum_{j=1}^q \theta_j B^j, \phi(B) = 1 - \sum_{j=1}^p \phi_j B^j \text{ et } \widehat{X}_t = \theta(B)\varepsilon_t$$

le processus X_t est un ARMA(p,q) qui se represente sous la forme :

$$\widehat{X}_t = \phi_1 \widehat{X}_{t-1} + \phi_2 \widehat{X}_{t-2} + \dots + \phi_p \widehat{X}_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

si en outre le processus X_t est centré ($\mu = 0$) alors :

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

Autrement, on modélisés le processus $\widehat{X}_t = X_t - \mu$ au lieu du processus X_t et l'on écrit :

$$\widehat{X}_t = \phi_1 \widehat{X}_{t-1} + \dots + \phi_p \widehat{X}_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

régressif et l'on a :

$$\left\{ \begin{array}{ll} \phi_{kk} = 0 & \forall k \geq p+1 \\ \phi_{kk} \neq 0 & \text{si } k = p \end{array} \right\}$$

2.2.4 Les fonctions d'auto-correlatin et d'auto-covariance :

Dans ce qui suit , que X_t est un processus stochastique stationnaire centré de fonction d'auto-covariance γ_x .

Fonction d'auto-corrélation AR(p) : L'auto-corrélation partielle au décalage k ($PACF(k)$) est définie comme étant la corrélation entre :

-les résidu de la régression de la série $X_{t+1}, X_{t+2} \dots X_{t+k-1}$ et

-les résidu de la régression de la série X_{t+k} par les série $X_{t+1}, X_{t+2}, X_{t+k-1}$, en d'autre termes :

$$X_{t+k} = \alpha_1 X_{t+1} + \alpha_2 X_{t+2} + \dots + \alpha_{k-1} X_{t+k-1} + U$$

$$X_t = \beta_1 X_{t+1} + \beta_2 X_{t+2} + \dots + \beta_{k-1} X_{t+k-1} + V$$

et :
$$PACF(k) = cov(U, V)$$

il faut comprendre que l'autocorrélation partielle est la corrélation entre X_t et X_{t+k} . une fois qu'on à explique ceux-ci par les valeurs entre aux deux $X_{t+1}, X_{t+2} \dots X_{t+k-1}$.

ACF et PACF d'un AR(1) : prenons une série X_t suivant un processus autorégressif d'ordre 1, par exemple :

$$X_t = 0.9X_{t-1} + \alpha_t.$$

– $ACF(1)$ la corrélation entre X_t et X_{t-1} peut se calculer facilement

$$\begin{aligned} ACF(1) &= cov(X_{t-1}, X_t) \\ &= cov(0.9X_t + \alpha_{t+1}, X_t) \\ &= 0.9cov(X_t, X_t) + cov(\alpha_{t+1}, X_t) \\ &= 0.9 + 0 \\ &= 0.9 \end{aligned}$$

ceci est vrai car X_t et α_{t+1} sont indépendants

– $ACF(2)$ en écrivant la relation entre un valeur à un instant et la valeur précédent à l'ordre $t - 1$ on à : $X_{t-1} = 0.9X_{t-2} + \alpha_{t-1}$

d'ou :
$$\begin{aligned} X_t &= 0.9(0.9X_{t-2} + \alpha_{t-1}) + \alpha_t \\ &= 0.81X_{t-2} + 0.9\alpha_{t-1} + \alpha_t \end{aligned}$$

comme X_{t-2} est indépendant à la fois de X_{t-1} et α_t et on déduit facilement que :

$$cov(X_t, X_{t-2}) = 0.81$$

$ACF(p)$ en continuant ce type de raisonnement on montre que :

$$cov(X_t, X_{t-k}) = (0.9)^k$$

On voit ici où se situe le problème de la fonction d'auto-corrélation des processus auto-régressif : il semblerait par la valeur de cette fonction qu'il y ait une relation entre X_t et X_{t-2} . On cette relation n'est qu'indirecte puisqu'elle n'existe que par X_{t-1} et par les relations de celui-ci avec X_t et X_{t-2}

-la fonction d'auto-corrélation partielle existe justement pour connaître jusqu'à quel niveau de décalage il existe une relation directe entre X_t et les valeurs précédentes .

$PACF(1)$ l'auto-corrélation partielle au décalage 1 est :

$$PACF(1) = cov(X_{t+1}, X_t)$$

l'autocorrélation partielle au décalage 1 est :

$$PACF(1) = cov(X_{t+1}, X_t) = 0.9$$

Rappele dans une régression : $Y_i = \alpha X_i + e_i$, le coefficient α peut être calculé comme :

$$\alpha = \frac{cov(X,Y)}{var(X)} = cov(X, Y) \frac{\delta(X)}{var(X)} = cov(X, Y) \frac{\delta(Y)}{\delta(X)}$$

c'est -à-dire que si les deux variable X et Y ont même écart-type , que l'on fasse la régression de X sur Y ou la régression de Y sur X on retrouve le même coefficient α

$$\alpha = cov(X, Y)$$

Pour résumer on peut dire que si on dispose de deux variables X et Y de moyennes nulles et de même écart-type , si on fait la régression de X sur Y :

$$Y = \alpha X + \varepsilon$$

alors la régression de Y sur X est :

$$X = \alpha Y + \delta \quad , \text{ et non}$$

comme on aurait pu le penser :

$$X = \frac{1}{\alpha Y} + \delta .$$

$PACF(2)$ il faut pour calculer $PACF(2)$ faire la régression de X_{t+2} et sur X_t on sait déjà que :

$$X_{t+2} = 0.9X_{t+1} + \alpha_{t+2}$$

Du rappèle précédent , on déduit que si on fait la régression de X_{t+1} et sur X_t on obtient : $X_t = 0.9X_{t+1} + V_t$, pour le calcul de $PACF(2)$

On calcule la corrélation entre les 2 résidus de ce régression :

$$\begin{aligned} PACF(2) &= corr(X_{t+2} - 0.9X_{t+1}, X_t - 0.9X_{t+1}) \\ &= corr(\alpha_{t+2}, X_t - 0.9X_{t+1}) \\ &= corr(\alpha_{t+2}, X_t) - 0.9corr(\alpha_{t+2}, X_{t+1}) \\ &\Rightarrow PACF(2) = 0 \end{aligned}$$

$PACF(K)$ par de calculs similaires on peut montrer que : $PACF(K) = 0$ pour tout $K > 1$.

$PACF$ d'un $AR(p)$

Nous ne nous attarderons pas sur cette fonction d'outocorrelation partielle mais on peut montrer que $PACF(K) = 0$.Pour tout $K > p$

Fonction d'auto-covariance d'un processus $AR(p)$: Si X_t est un processus autorégressif , alors :

$$\begin{aligned} \forall k \geq 1, \delta_x(k) &= cov(X_{t+k}, X_t) = E(X_{t+k}, X_t) \\ &= E(X_{t-k}, X_t) \end{aligned}$$

(à cause de la stationnaire du processus)

$$\begin{aligned} X_{t-k}X_t &= \phi_1 X_{t-k}X_{t-1} + \phi_2 X_{t-k}X_{t-2} + \dots + \phi_p X_{t-k}X_{t-p} + X_{t-k}\varepsilon_t \\ \Rightarrow E(X_{t-k}X_t) &= \phi_1 E(X_{t-k}X_{t-1}) + \phi_2 E(X_{t-k}X_{t-2}) + \dots + \phi_p E(X_{t-k}X_{t-p}) + E(X_{t-k}\varepsilon_t) \\ \Rightarrow y_k &= \phi_1 y_x(k-1) + \phi_2 y_x(k-2) + \dots + \phi_p y_x(k-p) \end{aligned}$$

(relation entre auto-covariance) en devisant cette dernière relation par γ_0 on obtient :

$$\rho_x(k) = \phi_1 \rho_x(k-1) + \phi_2 \rho_x(k-2) + \dots + \phi_p \rho_x(k-p)$$

La quantité ϕ_k notée ϕ_{kk} en tant que fonction de k est appelle fonction d'auto-corrélation partielle , cette fonction condition alors un outil pour reconnaître l'ordre (p)

d'un processus AR(p) . Puisque la fonction d'auto-corrélation est la dernière pondération d'un AR(p) .

La quantité ϕ_{kk} définit , alors l'ordre du processus autorégressif et l'on a :

$$\left\{ \begin{array}{ll} \phi_{kk} = 0 & \forall k \geq p + 1 \\ \phi_{kk} \neq 0 & \text{si } k = p \end{array} \right\}$$

Fonction d'auto-covariance d'un processus MA(q) : Si X_t est un processus moyenne mobile alors :

$$\begin{aligned} y_x(k) &= cov(X_{t+k}, X_t) = E(X_{t+k} - X_t) \\ &= E \left[\left(\sum_{j=0}^q \Theta_j \varepsilon_{t+k-j} \right) \left(\sum_{i=0}^q \Theta_i \varepsilon_{t-i} \right) \right] \\ &= \sum_{j=0}^q \sum_{i=1}^q \Theta_j \Theta_i E(\varepsilon_{t-i} \varepsilon_{t+k-j}) \\ &= \left\{ \begin{array}{ll} \sigma_\varepsilon^2 \sum_{j=0}^q \Theta_j & \text{si } k \leq q \\ 0 & \text{si non} \end{array} \right\} \end{aligned}$$

Dans le cas où X_t n'est pas centré on le remplace par :

$$\hat{X}_t = X_t - E(X_t) \text{ et } y_x \text{ sera égale à } y_{\hat{x}}$$

Fonction d'auto-covariance d'un bruit blanc : A partir de la définition d'un bruit blanc on écrit :

$$\gamma_\varepsilon(h) = \left\{ \begin{array}{ll} \gamma_\varepsilon^2 & \text{si } h = 0 \\ 0 & \text{si non} \end{array} \right\}$$

Présentation spectrale du processus ARMA :

Densité spectrale : soit X_t un processus stationnaire de moyenne nulle et de fonction d'auto-covariance $\gamma_x(\cdot)$, satisfaisant $\sum_{h=-\infty}^{+\infty} |\gamma_x(h)| < \infty$ la densité spectrale du processus X_t est alors définie par : $f_x(\mathfrak{J}) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \exp(ih\mathfrak{J}) \gamma_x(h)$ ou $-\infty < \mathfrak{J} < +\infty$

Densité spectrale pour un bruit blanc : Le bruit blanc étant un processus stationnaire de moyenne non nulle alors , sa densité spectrale s'écrit sous la forme :

$$\begin{aligned}
f_\varepsilon(\lambda) &= \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} (\exp(ih\lambda)\gamma_\varepsilon(h)) \\
&= \frac{1}{2\pi}\gamma_\varepsilon(0) = \frac{\gamma_\varepsilon^2}{2\pi}
\end{aligned}$$

Proposition 2 Soit X_x un processus stationnaire de moyenne mobile et de spectrale $f_x(\lambda)$, si $(\psi_j)_{j \geq 0}$ est une suite numérique vérifiant :

$$\sum_{j=0}^{+\infty} |\psi_j| < +\infty \quad , \text{ alors le processus :}$$

$$Y_t = \sum_{j=0}^{+\infty} \psi_j X_{t-j}$$

est stationnaire de moyenne nulle et de densité spectrale :

$$f_y(\lambda) = |\psi(\exp(-i\lambda))|^2 f_x(\lambda) = \psi(\exp(-i\lambda))\psi(\exp(i\lambda))f_x(\lambda)$$

Alors les fonctions de vraisemblance de modèle ARMA peuvent donner la bonne représentation des certaines séries chronologiques et un processus .

2.3 Vecteurs Gaussien :

Dans toute la suite (Ω, F, P) est un espace de probabilité .

Definition 3 Un vecteur aléatoire est une va à valeurs dans $\mathbb{R}^d, V(\Omega, F, \mathbb{R}) \rightarrow (\mathbb{R}^d, B(\mathbb{R}^d))$ L'espérance d'un vecteur aléatoire $V \in L^1(\mathbb{R}^d)$ est l'espérance coordonnée par coordonnée : si $V(v_1, \dots, v_d)$ dans un base de \mathbb{R}^d , alors $E(V) = (E(v_1), \dots, E(v_d))$ l'espérance est une application linéaire sur $L^1((\Omega, F, \mathbb{R}); \mathbb{R}^d)$.

Definition 4 Si V est un vecteur aléatoire de carré intégrable on appelle matrice de covariance (ou de dispersion) de V la matrice K_V définie par $E((V - EV)(V - EV)^t)$ si V est un vecteur colonne (ou $E((V - EV)^t(V - EV))$ si V est un vecteur ligne) .

Definition 5 Un vecteur aléatoire $V \in \mathbb{R}^d$ est gaussienne lorsque pour tout $U \in \mathbb{R}^d$ la var $U^t V$ est gaussienne. Autrement dit , tout combinaison linéaire à coefficient constants des coordonnées de V est gaussienne , on note $N(m, K)$ la loi gaussienne sur \mathbb{R}^d , d'espérance $m \in \mathbb{R}^d$ et de matrice de covariance K .

Proposition 6 Soit V un vecteur gaussienne à valeur \mathbb{R}^d , de loi $N(m, K)$ avec K inversible alors V admet la densité suivante par rapport à la mesure de lebergue :

$$f(x) = (2\pi)^{\frac{-d}{2}} (\det K)^{\frac{-1}{2}} \exp \left\{ \frac{-1}{2} (x - m)^t k^{-1} (x - m) \right\}$$

2.4 Fonction de vraisemblance du modèle ARMA :

2.4.1 La fonction de vraisemblance inconditionnelle :

Soit r_t le processus présenté par le modèle ARMA(p,q) de moyenne $\mu = 0$, qui s'écrit sous la forme :

$$\begin{aligned} r_t &= \phi_1 r_{t-1} + \dots + \phi_p r_{t-p} + \varepsilon_t + \Theta_1 \varepsilon_{t-1} + \Theta_2 \varepsilon_{t-2} + \dots + \Theta_q \varepsilon_{t-q} \\ \Rightarrow \varepsilon_t &= r_t - \phi_1 r_{t-1} - \dots - \phi_p r_{t-p} - \Theta_1 \varepsilon_{t-1} - \Theta_2 \varepsilon_{t-2} - \dots - \Theta_q \varepsilon_{t-q} \end{aligned}$$

et suppose que les variables du vecteur $\varepsilon = (\varepsilon_i)_{i=1, \dots, n}$, i.i.d selon une loi normale centré et d'écart type δ_ε la densité conjointe de ce vecteur sera d'un vecteur gaussien définie par :

$$f(\varepsilon_1, \dots, \varepsilon_n) = \frac{1}{(2\pi\sigma_\varepsilon^2)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=1}^n \varepsilon_t^2\right\}$$

On pose :

$$\begin{aligned} r &= (r_1 \dots, r_n), \varepsilon = (\varepsilon_1, \dots, \varepsilon_n), e_* = (r_{1-p} \dots, r_0, \varepsilon_{1-q}) \\ e_* &= (r_*, \varepsilon_{1-q}, \dots, \varepsilon_0) \end{aligned}$$

La fonction de vraisemblance exacte est de la forme :

$$L(B) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}} |\Omega|^{\frac{1}{2}} |D|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} S(\phi, \theta)\right\}$$

et le logarithme de la fonction de vraisemblance exacte (inconditionnelle) pour le processus ARMA aura pour expression :

$$L\left(\frac{r}{B}\right) = \left(\sum_{i=1}^n f\left(\frac{r_i}{r_{t-1}, \dots, r_1, B}\right) \right)$$

Posons :

$$\begin{aligned} \hat{r}_t &= E\left(\frac{r_t}{r_{t-1}, \dots, r_1, B}\right) \\ \delta_\varepsilon^2 V_t &= E(r_t - \hat{r}_t)^2 \\ \det(M_n^{(p,q)}) &= v_1 \dots v_n \end{aligned}$$

Telle que : $M_n^{(p,q)}$ étant la matrice d'auto-covariance du n-uple r .

Si on suppose que le processus r_t est la loi normale, on aura :

$$f\left(\frac{r_t}{r_{t-1}, \dots, r_1, B}\right) = \frac{1}{[2\pi\delta_\varepsilon^2 v_t]^{\frac{1}{2}}} \exp\left(-\frac{(r_t - \hat{r}_t)^2}{v_t}\right)$$

Et par la suite :

$$f\left(\frac{r}{B}\right) = \prod_{t=1}^n \frac{1}{[2\pi\delta_\varepsilon^2 v_t]^{\frac{1}{2}}} \exp\left(-\frac{1}{2\delta_\varepsilon^2} \sum_{t=1}^n \frac{(r_t - \hat{r}_t)^2}{v_t}\right)$$

de cette dernière relation on conclut que :

$$\begin{aligned}
 L(\phi, \theta, \delta_\varepsilon^2) &= \prod_{t=1}^n \frac{1}{[2\pi\delta_\varepsilon^2 v_t]^{\frac{1}{2}}} \exp\left(\frac{-1}{2\delta_\varepsilon^2} \sum_{t=1}^n \frac{(r_t - \hat{r}_t)^2}{v_t}\right) \\
 \Rightarrow L(B) = \ln[L(\phi, \theta, \delta_\varepsilon^2)] &= \frac{n}{2} \ln(2\pi) - \frac{n}{2} \sum_{t=1}^n \delta_\varepsilon^2 - \frac{1}{2} \sum_{t=1}^n \ln(v_t) - \frac{1}{2\delta_\varepsilon^2} S^2(\phi, \theta) \\
 &= \frac{n}{2} \ln(2\pi) - \frac{n}{2} \sum_{t=1}^n \delta_\varepsilon^2 - \frac{1}{2} \sum_{t=1}^n \ln \left[\det \left(M_n^{(p,q)} \right) \right] - \frac{1}{2\delta_\varepsilon^2} S^2(\phi, \theta)
 \end{aligned}$$

avec :

$$S^2(\phi, \theta) = \sum_{t=1}^n \frac{(r_t - \hat{r}_t)^2}{v_t}$$

Par la méthode précédente on a calculé la fonction de vraisemblance exacte utilisant un changement de variable , alors La fonction de densité conjointe du vecteur aléatoire $\varepsilon = (\varepsilon_t)_{i=1, \dots, n}$ vue comme fonction de B s'appelle fonction de vraisemblance.

Enfin , pour terminer la justification d'un modèle, il faut vérifier si les prévisions que nous pouvons faire , a partir des modèles (ARMA , linéaire).

Bibliographie

1. Les variables aléatoires. Mathématique : outils pour la biologie-Deng SUI-UcBl , D.Mouchiroud 10-10-2002 .
2. Loi de Laplace .Wikipédia
3. Cours processus stochastique . David Coupier Plytech'Lille.
4. Série chronologique , Agnès Lagnoux , Université de Toulouse le Mirail.
5. poly (les modèle linéaires) C.Chouquet Laboratoire de Statistique et Probabilités - Université Paul Sabatier - Toulouse
M1 IMAT, Année 2009-2010.
6. Box . Bernard Rapacchi Centre Interuniversitaire de Calcul de Grenoble 18-08-1993
7. Modèle stochastique et leur prevision par des processus Auto-régressifs -Mémoire présenté pour L'obtention du Diplome de Magister en Mathématique par Boularouk Yakoub-(Université Mentouri Constantine)15 /06 / 2010 . Page 10-14 .
8. Maximisation. 1980-2007 Philippe JOLIVALDT PARIS1. Permission d'utiliser ce document pour un usage étudiant et strictement personnel (usage professionnel et/ou commercial interdit).