

N° d'ordre:

RÉPUBLIQUE ALGERIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ MOULOU MAMMARI DE TIZI OUZOU
FACULTÉ DES SCIENCES
DÉPARTEMENT DE MATHÉMATIQUES
LABORATOIRE LMPA



MÉMOIRE DE MASTER

Filière : Mathématiques
Spécialité : Probabilités & Statistique

Par

NOURA CHITTI

SUR LES SÉRIES CHRONOLOGIQUES À VALEURS ENTIÈRES

Soutenu le 27 Juillet 2023 devant le jury :

FARIDA ACHEMINE	M.C.A.	UMMTO	Présidente
ABDELGHANI HAMAZ	M.C.A.	UMMTO	Rapporteur
HOURLA BOUZIANE	M.A.A.	UMMTO	Examinatrice

Année Universitaire : 2022/2023

*À la mémoire de ma grand mère, cette grande dame qui m'a chérie et aimée
depuis enfant, et qui serait fière de moi aujourd'hui ...*

REMERCIEMENTS

IL m'est agréable d'adresser mes premiers remerciements à mon promoteur, Monsieur Hamaz Abdelghani qui, malgré toutes les obligations liées à son poste de doyen, a toujours été de bons conseils, m'a soutenu et encouragé tout au long de ce travail.

Je tiens aussi à remercier Madame Achemine Farida d'avoir présidé le jury et Madame Bouziane Houria d'avoir examiné ce travail.

Je dédie ce travail à mes parents, sans qui rien de cela n'aurait été possible, qui ont cru en moi quand moi-même je n'y croyais plus, qui n'ont cessé de se battre pour me voir réussir.

Par la suite je tiens à remercier mes deux grands frères, Momoh et Smail ainsi que mes belles sœurs qui m'ont soutenus inconditionnellement.

Viendront ma tante Samou et mes cousins et cousines Camy, Kenza et Yacine, merci de faire partie de ma vie.

Tizi-Ouzou, le 27 Juillet 2023.

TABLE DES MATIÈRES

TABLE DES MATIÈRES	iv
LISTE DES FIGURES	vi
LISTE DES TABLEAUX	vi
INTRODUCTION	1
1 GÉNÉRALITÉS SUR LES SÉRIES CHRONOLOGIQUES	3
1.1 INTRODUCTION	3
1.2 PROCESSUS ALÉATOIRES	4
1.3 SÉRIES CHRONOLOGIQUES	4
1.3.1 Composantes principales d'une série chronologiques	4
1.3.2 Modélisation des séries chronologiques	5
1.3.3 Analyse de la composante tendance-cycle	7
1.3.4 Analyse de composante saisonnière	7
1.4 NOTIONS DE STATIONNARITÉ, AUTO-COVARIANCE, AUTO-CORRÉLATION	7
1.4.1 Bruit Blanc	8
1.4.2 Stationnarité au sens stricte	8
1.4.3 Stationnarité faible	8
1.4.4 Fonction d'auto-covariance	8
1.4.5 Fonction d'auto-corrélation	9
1.5 MODÈLE CLASSIQUE DE BOX ET JENKINS	9
1.5.1 Élimination de la tendance et de la composante saisonnière par différentiation	9
1.5.2 Étapes de la méthode de Box et Jenkins	10
1.6 MODÈLE ARMA	12
1.6.1 Modèle mixte ARMA	12
1.6.2 Modèle AR	12
1.6.3 Modèle MA	12
1.6.4 Causalité	13
1.6.5 Inversibilité	13
1.7 MÉTHODES DE VÉRIFICATION DE LA STATIONNARITÉ	13
1.7.1 Test de Dickey-Fuller	14
1.7.2 Test de Dickey-Fuller augmenté	14
CONCLUSION	17
2 MODÈLE AUTORÉGRESSIF D'ORDRE 1 À VALEURS ENTIÈRES INAR(1)	18
2.1 INTRODUCTION	18
2.2 PROCESSUS DE COMPTAGE	18

2.2.1	Notations et caractéristiques préliminaires des processus de comptages	18
2.2.2	Domaines d'application	19
2.2.3	Quelques processus de comptages	19
2.2.4	Autoregression et dépendance à 1 pas	21
2.3	LE MODELE INAR(1) (<i>the integer-valued autoregressive of order 1</i>)	21
2.3.1	L'opérateur d'amincissement binomiale "o"	21
2.3.2	Modele INAR(1)	23
2.3.3	Propriétés de distribution du modèle INAR(1)	24
2.4	LE MODÈLE POISSON INAR (1)	25
2.5	STATIONNARITÉ STRICTE	25
2.5.1	Propriétés	26
2.5.2	Simulation du modèle	26
2.6	EXEMPLE DE SIMULATION	27
2.7	MODÈLES INAR(1) AVEC DES INNOVATIONS GÉNÉRALISÉES	28
2.7.1	Le Modèle Géométrique INAR(1)	29
2.7.2	Le modèle Poisson composé INAR(1),noté CP-INAR(1)	29
2.7.3	Le modèle NB-INAR(1)	29
2.8	ESTIMATION	30
2.9	PRÉDICTION DANS LE MODÈLE INAR(1)	33
2.9.1	Prédiction avec la méthode classique	34
2.9.2	Prédiction avec la méthode bayésienne	34
2.9.3	Intervalle de prédiction	35
2.10	EXEMPLE DE DONNÉES RÉELLES	37
3	INAR(p) ET ESTIMATION	39
3.1	INTRODUCTION	39
3.2	MODÈLE INAR(p)	39
3.2.1	Propriétés du modèle	39
3.2.2	Propriétés de corrélation	40
3.3	MODÈLE INAR(p) DE DU ET LI	42
3.4	STATIONNARITÉ D'UN MODÈLE INAR(p)	43
3.5	ESTIMATION D'UN INAR(p)	43
3.5.1	Estimation de Yule Walker	44
3.5.2	Estimation par la méthode des moindres carrés	44
3.6	GÉNÉRALISATION	46
3.6.1	L'opérateur d'amincissement généralisé noté α^*	47
3.6.2	Le modèle GINAR(p)	47
3.6.3	Condition de stationnarité	47
3.6.4	Moments du premier et du second ordre	47
3.7	PRÉVISION DANS LES MODÈLES INAR(p)	48
3.7.1	Prédiction : Méthode de prédiction médiane	48
3.7.2	Approche des chaînes de Markov	49
3.8	DIAGNOSTIC DU MODÈLE	50
3.8.1	Analyse des résidus	51
3.9	EXEMPLE DE SIMULATION	51
3.10	CONCLUSION DU CHAPITRE	54
	CONCLUSION GÉNÉRALE	55
	BIBLIOGRAPHIE	56

LISTE DES FIGURES

1.1	Représentation graphiques des composantes d'une série chronologique	5
1.2	Modèle additif	6
1.3	Modèle multiplicatif	6
1.4	Série temporelle pour test de dickey et Fuller	16
2.1	Distributions de données de comptage ($E(X) = 3.5$)	21
2.2	Simulation d'un processus INAR(1)	27
2.3	Exemples de trajets simulés de processus INAR (1) Poissonien avec $\mu = 3$ et $\alpha = 0,5$	28
2.4	Exemples de trajets simulés de processus INAR (1) Poissonien avec $\mu = 3$ et $\alpha = 0 : 95$	28
2.5	Prévisions à h pas en avant des demandes mensuelles de juillet à décembre 1994.	38
3.1	Diagramme de série chronologique de 115 observations à partir de données simulées.	52
3.2	a-ACF b-PACF	52

LISTE DES TABLEAUX

2.1	Définition et notation des moments d'une variable aléatoire de comptage X	18
2.2	Définition et notation des fonctions génératrices d'une variable aléatoire de comptage X	19
3.1	Probabilité conditionnelle de X_{115+h} sachant X_{112+h} , X_{113+h} et X_{114+h} pour une série chronologique de données simulées.	53
3.2	Diagnostic du modèle AR(3) et INAR(3).	53

INTRODUCTION GÉNÉRALE

DE tout temps, l'homme a cherché à connaître l'avenir, avec des intentions diverses, que ce soit pour se prémunir des catastrophes, des guerres, ou bien tout simplement pour satisfaire sa curiosité. Les premières tentatives de prévision reposaient souvent sur des croyances ésotériques telle que l'astrologie.

Au fil du temps, la science a bien évolué sur le sujet, ouvrant la voie à des méthodes de prévision plus développées et scientifiquement fondées, l'une d'entre elles étant l'analyse des séries temporelles.

D'après Kendall (1973) les premiers à avoir fait usage des séries chronologique sont les astronomes de l'antiquité, ils enregistraient des données nombreuses et précises sur l'orbite des planètes et s'en servaient pour prédire des phénomènes astronomiques. Depuis, de nombreuses avancées théoriques et méthodologiques ont été réalisées pour mieux comprendre et analyser les séries temporelles : les développements statistiques, tels que les tests de stationnarité, les modèles de régression et les techniques de lissage, ont permis de capturer les motifs et les tendances des données.

Au cœur de l'étude formelle des séries temporelles se trouve la recherche d'un modèle mathématique qui explique le mieux les données observées. A partir de ce modèle, des prédictions peuvent être faites, mais leurs précisions dépendra fortement du modèle choisi, on privilégiera toujours celui qui minimise l'erreur. Pour le cas des séries chronologiques univariées, elles ont bénéficié d'une théorie très développée on peut citer la méthode de Box et Jenkins (1970) qui repose sur la fusion des modèles autorégressifs et des modèles à moyenne mobile. Cette approche est largement documentée et reconnue dans la communauté scientifique.

Il existe un autre type de séries chronologique qu'on peut observer dans divers domaines scientifiques, les séries chronologiques à valeurs entières, c'est à dire des séries qui prennent leurs valeurs dans un ensemble fini ou infini dénombrable.

Les deux grands exemples de séries temporelles discrètes sont les processus de naissance, mort et les séries de comptage où on s'intéresse au nombre d'occurrences d'un événement particulier dans un intervalle de temps spécifique.

On les retrouve notamment dans les disciplines telles que l'économie, la médecine, les télécommunications, les assurances et pleins d'autres encore. On peut citer pour exemple : le nombre de patients par jour dans un hôpital, le nombre de réclamations que reçoit une société d'assurance, le nombre de

lignes occupés dans un réseau téléphonique etc.. D'où l'intérêt d'avoir un modèle spécifique permettant de reproduire et d'expliquer l'évolution de ce type de séries chronologiques. Cependant avant les années 80 ce domaine de recherche n'était que peu étudié, McKenzie (2003) a donné comme explication la complexité du sujet, en effet la modélisation des séries temporelles discrètes est rendue difficile car la plupart des représentations traditionnelles de la dépendance deviennent soit impossibles, soit impraticables.

Dans certains cas, les valeurs de la séries chronologique sont assez grandes et peuvent être modélisées par des modèles à valeurs continues (modèles autorégressifs ARMA) mais pour des séries aux plus petits nombres l'efficacité de ce modèle est limitée, en effet le simple fait de multiplier une variable entière par une constante réelle conduit à une variable aléatoire non entière, c'est pour palier à ce problème que Stautel et Van Harn (1979) ont introduit une alternative à la multiplication pour le cas des valeurs entières appelé opérateur d'amincissement.

Au cours des quatre dernières décennies, l'analyse des séries chronologique a connue un intérêt croissant, plusieurs modèles basés sur cet opérateur aléatoire ont des lors été présentés, l'un des premiers est le modèle INAR(1) (Integer-valued Autoregressive of order 1) ou Modèle autorégressif d'ordre 1 à valeurs entières, introduit par McKenzie (1985), Al Osh et Al Zaid (1987), s'en est suivie une extension directe le modèle INAR(p) étudié par Al-Osh et Al-Zaid et Du & Li (1991).

L'objectif de ce mémoire est d'explorer les principaux concepts et méthodes utilisés dans l'analyse des séries chronologique à valeurs entières, en mettant l'accent sur les modèles INAR. Nous étudierons les techniques d'estimation des paramètres, les méthodes de prévision et les outils de diagnostic pour évaluer la performance des modèles, il est scindé en trois chapitres :

Dans le premier chapitre, nous allons dans un premier temps revoir des notions générales sur les séries chronologiques, en étudier les différentes composantes et présenter le modèle de régression linéaire classique, le modèle ARMA.

Dans le second chapitre, Nous nous intéressons au processus autorégressif à valeurs entières INAR(1), on passera en revue ses principales caractéristiques et nous explorerons ses méthodes d'estimation et de prévision. Nous examinerons aussi les avantages ainsi que les limites du modèle.

Dans le troisième chapitre, on abordera les INAR(p) (Integer-Valued Pth-Order Autoregressive Model), on verra la définitions des deux différents modèles, celle donné par Al-Osh et Al-Zaid et celle donné par Du et Li.

GÉNÉRALITÉS SUR LES SÉRIES CHRONOLOGIQUES



1.1 INTRODUCTION

Les séries chronologiques constituent un domaine essentiel de l'analyse de données et de la modélisation statistique. Elles sont omniprésentes dans de nombreux domaines tels que l'économie, la finance, la météorologie, la démographie et bien d'autres. L'étude des séries chronologiques permet de comprendre et de modéliser l'évolution temporelle d'une variable d'intérêt, fournissant ainsi des informations précieuses pour la prise de décisions, la prévision et la compréhension des phénomènes dynamiques.

Une série chronologique est une collection de données ordonnées dans le temps, où les observations sont enregistrées à des intervalles réguliers. Ces données peuvent représenter diverses mesures, telles que des valeurs économiques, des mesures de performances, des températures, des ventes, des taux de croissance, etc. L'analyse des séries chronologiques vise à extraire des modèles, des tendances, des motifs saisonniers, des cycles et des comportements imprévisibles de ces données.

L'objectif de ce chapitre est de :

- Présenter les concepts fondamentaux et les principaux outils d'analyse des séries chronologiques. Nous commencerons par explorer les caractéristiques spécifiques des séries chronologiques qui les distinguent des autres types de données.
- Étudier les différentes composantes qui peuvent influencer une série chronologique, notamment les tendances, les saisons et les cycles. Nous discuterons des méthodes de décomposition qui permettent de séparer ces composantes afin d'obtenir une meilleure compréhension du comportement de la série. De plus, nous aborderons les notions de stationnarité, qui sont cruciales pour choisir les méthodes d'analyse et de modélisation appropriées.
- Introduire les concepts de modélisation et de prévision des séries chronologiques. Nous présenterons les modèles classiques tels que les modèles ARIMA (AutoRegressive Integrated Moving Average) et les modèles de régression linéaire.

En conclusion, ce chapitre fournira une base solide pour la compréhension des séries chronologiques en exposant les concepts clés, les méthodes d'analyse et les modèles de prévision. Une fois les principes de base établis, les chercheurs et les praticiens seront en mesure d'appliquer ces connaissances pour explorer, analyser et anticiper les tendances futures dans leurs domaines respectifs.

1.2 PROCESSUS ALÉATOIRES

Définition 1.1 *Un processus aléatoire est une famille de variables ou de vecteurs aléatoires $(X_t)_{t \in T}$, indexée par l'ensemble T des temps dénombrable ou continu, définie sur un espace de probabilité (Ω, T, P) et à valeurs dans un espace d'états E . La variable ou le vecteur aléatoire X_t décrit l'état du processus du temps t . A toute occurrence ω on fait correspondre la trajectoire de la réalisation $(X_t(\omega))_{t \in T}$ du processus définie par l'application :*

$$t \in T \longrightarrow X_t(\omega) \in E \quad (1.1)$$

Remarque 1.1

- Si T est dénombrable ($T \subseteq \mathbb{N} \subseteq \mathbb{Z}$), $X = (X_t)_{t \in T}$ est une suite de variables aléatoires ou un processus à temps discret.
- Si $T \subseteq \mathbb{R}$ a la puissance du continu (intervalle ou combinaison d'intervalles) $(X_t)_{t \in T}$ est dit processus à temps continue ou processus stochastique.
- Si E est discret (dénombrable fini ou infini) X est dit processus discret.
- Si $E \subseteq \mathbb{R}$ a la puissance du continu, X est dit processus continu.
- par abus de langage les réalisations dans une durée finie d'un processus sont notés série chronologique.

1.3 SÉRIES CHRONOLOGIQUES

Définition 1.2 *Une série temporelle est une collection de données ordonnées dans le temps, qui sont mesurées à des intervalles réguliers ou irréguliers.*

Une série temporelle est donc une suite de variables aléatoires indexées par le temps, c'est-à-dire un ensemble de valeurs numériques X_1, X_2, \dots, X_t , où t est un entier représentant l'instant de temps correspondant à chaque valeur de la série. Les valeurs de la série temporelle peuvent être discrètes ou continues. Une série temporelle peut également avoir des propriétés telles que la stationnarité, l'autocorrélation, la périodicité, etc., qui peuvent être utilisées pour modéliser et analyser la série temporelle.

Les exemples courants de séries temporelles incluent les données économiques telles que les ventes mensuelles d'une entreprise, les données météorologiques telles que les températures quotidiennes, les données financières telles que les prix quotidien des actions, etc.

1.3.1 Composantes principales d'une série chronologiques

1. **La tendance (Trend) (Z_t)** : Représente l'évolution à long terme de la série étudiée. Elle traduit le comportement moyen de la série.
2. **Composante saisonnière (S_t)** : Correspond à un phénomène qui se répète à intervalles de temps réguliers (périodiques). En général, c'est un phénomène saisonnier dans la variation saisonnière.
3. **Composante résiduelle (bruit ou résidus) (ϵ_t)** : Correspond à des fluctuations régulières, en général de faible intensité, mais de nature aléatoire. On parle aussi d'aléas.
Des phénomènes accidentels peuvent notamment intervenir (grèves, conditions météorologiques exceptionnelles, crash financier).

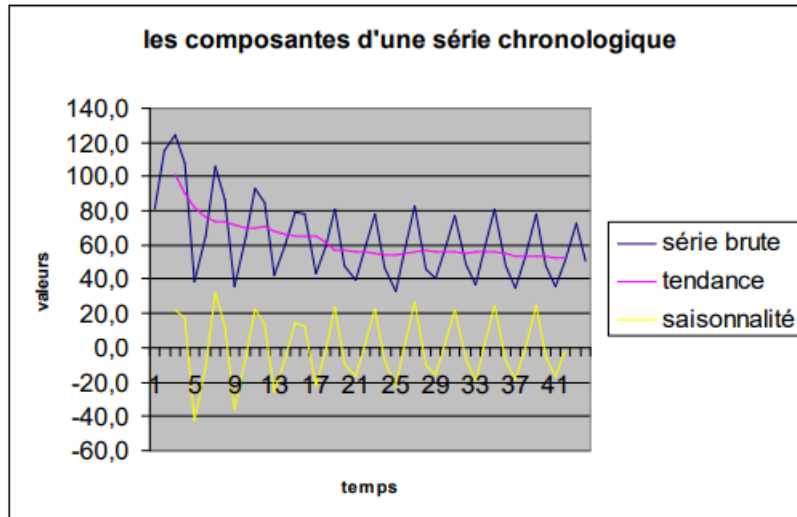


FIGURE 1.1 – Représentation graphique des composantes d'une série chronologique

1.3.2 Modélisation des séries chronologiques

Un modèle est une représentation simplifiée de la réalité, conçue pour expliquer les mécanismes sous-jacents d'un phénomène étudié et faciliter sa compréhension. Certains modèles sont plus efficaces que d'autres pour décrire la réalité, ce qui soulève des questions importantes : comment évaluer la qualité d'un modèle ? Comment diagnostiquer ses performances ?

Nous avons deux principaux modèles :

Les modèles déterministes

Les modèles déterministes supposent que l'observation d'une série temporelle à la date t , représentée par X_t , est une fonction du temps t et d'une variable d'erreur centrée, notée ϵ_t . On suppose également que les erreurs ϵ_t sont décorrélées. Ainsi, le modèle déterministe peut être exprimé par l'équation :

$$X_t = f(t, \epsilon_t) \quad (1.2)$$

Les diverses composantes peuvent être regroupées en fonction de l'un des trois modèles (descriptifs) ci-dessous :

- **Modèle additif** : Dans un modèle additif, on suppose que les 3 composantes : tendance, variations saisonnières et variations accidentelles sont indépendantes les unes des autres.

Il convient de noter que le modèle additif est approprié lorsque la magnitude des variations saisonnières est relativement constante au fil du temps. Si la magnitude des variations saisonnières change avec la tendance de la série, le modèle multiplicatif peut être plus approprié. il se présente sous la forme suivante :

$$X_t = Z_t + S_t + \epsilon_t \quad (1.3)$$

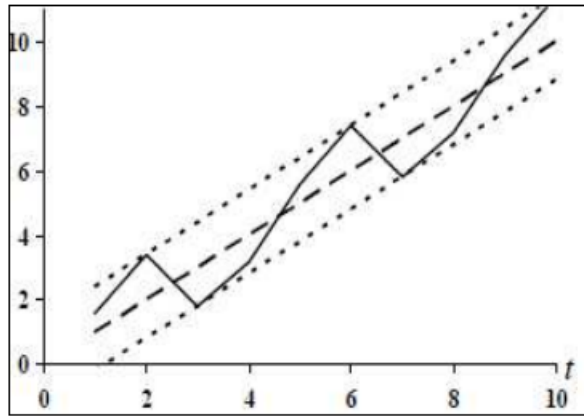


FIGURE 1.2 – *Modèle additif*

- **Modèle multiplicatif** : On suppose que les variations saisonnières dépendent de la tendance. On considère qu’il s’écrit de la manière suivante :

$$X_t = Z_t * S_t * \epsilon_t \quad (1.4)$$

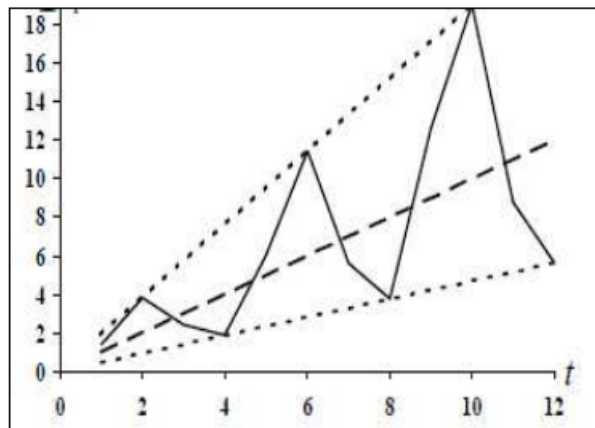


FIGURE 1.3 – *Modèle multiplicatif*

- **Modèle mixte** Le modèle mixte combine les caractéristiques des modèles additif et multiplicatif pour la décomposition des séries chronologiques. Il est utilisé lorsque la série présente à la fois des variations saisonnières dont l’amplitude change avec la tendance et des variations aléatoires. La décomposition se fait en deux étapes : une décomposition additive suivie d’un ajustement multiplicatif des variations saisonnières en fonction de la tendance. Cela permet de capturer les relations complexes entre la tendance et la saisonnalité, offrant une décomposition plus précise et réaliste de la série. Il s’écrit sous la forme suivante :

$$X_t = Z_t * S_t + \epsilon_t \quad (1.5)$$

Les modèles stochastiques

C’est du même ordre que les modèles déterministes mais à la différence de ceux ci, les variables ϵ_t ne sont pas i.i.d. mais disposent d’une structure de corrélation non nulle : ϵ_t est une fonction qui dépend des valeurs précédentes

et d'un terme d'erreur n_t .

$$\epsilon_t = g(\epsilon_{t-1}, \epsilon_{t-2}, \dots, n_t) \quad (1.6)$$

La classe des modèles de ce type la plus fréquemment utilisée est la classes des modèles SARIMA (et de ses sous-modeles ARIMA, ARMA,...).

1.3.3 Analyse de la composante tendance-cycle

Définition 1.3 *La tendance, également appelée mouvement de fond ou structurel, est un phénomène qui se manifeste comme un mouvement de longue durée, représentant l'évolution générale du dit phénomène. Elle est considérée comme la composante la plus significative au sein d'une série chronologique et est souvent exprimée en fonction du temps. La tendance peut adopter différentes formes, telles qu'une progression linéaire, exponentielle, ou présenter des modèles plus complexes, tels que la progression logistique ou d'autres.*

On parle de tendance linéaire lorsque la série peut se décomposer en :

$$x_n = a_n + b + e_n \quad \text{pour } n = 1, 2, \dots \quad (1.7)$$

Plus généralement, on parle de tendance polynomiale lorsque la série peut se décomposer en :

$$x_n = a_1 n^p + a_2 n^{p-1} + \dots + a_{p+1} + e_n \quad \text{pour } n = 1, 2, \dots \quad (1.8)$$

dans laquelle e_n est un résidu où ne figure plus la tendance et qui, de ce fait, a une allure relativement homogène dans le temps. Ceci sera précisé par la suite. De même, on peut définir des tendances logarithmiques, exponentielles, etc.

1.3.4 Analyse de composante saisonnière

Pour modéliser un effet saisonnier, admettant du bruit mais pas de tendance, nous utilisons le modèle simple suivant :

$$X_t = S_t + \epsilon_t \quad (1.9)$$

où S_t est une fonction périodique de t , de période d i.e, pour tout t , $S_{t-d} = S_t$. Un choix convenable pour S_t est une somme des fonctions harmoniques définies par :

$$S_t = \alpha_0 + \sum_{j=1}^k [\alpha_j \cos(\lambda_j t) + b_j \sin(\lambda_j t)] \quad (1.10)$$

ou $\alpha_0, \alpha_1, \dots, \alpha_k$ et b_1, b_2, \dots, b_k sont des paramètres inconnus et $\lambda_1, \dots, \lambda_k$ sont des fréquences fixes, chacune étant un multiple entier de $2\pi/d$: $\lambda_j = 2\pi j/d, \quad \forall j = 1, 2, \dots, k$.

1.4 NOTIONS DE STATIONNARITÉ, AUTO-COVARIANCE, AUTO-CORRÉLATION

La stationnarité est un concept clé dans l'analyse des séries chronologiques. Une série stationnaire suppose que les caractéristiques statistiques de la série restent constantes au fil du temps, sans tendance significative ou variations systématiques. Dans cette section, nous explorerons les critères utilisés pour

évaluer la stationnarité d'une série. Comprendre ces notions est essentiel pour analyser de manière précise et fiable les données temporelles. Avant de définir les notions de stationnarité faible et forte, nous allons d'abord expliquer ce qu'est le "bruit blanc".

1.4.1 Bruit Blanc

Définition 1.4 *Un processus $(X_t; t \in \mathbb{Z})$ est un bruit blanc faible ou bruit blanc de moyenne nulle et de variance σ^2 noté $\{X_t, t \in \mathbb{Z}\} \sim BB(0, \sigma^2)$, (BB est une abréviation pour Bruit Blanc)*

Si X_t est de moyenne nulle et de fonction d'autocovariance définie par :

$$\gamma(h) = \begin{cases} \sigma^2 & \text{si } h \in E \\ 0 & \text{sinon.} \end{cases} \quad (1.11)$$

1.4.2 Stationnarité au sens stricte

Définition 1.5 *Une série temporelle $\{X_t, t \in T\}$ est strictement ou fortement stationnaire si \forall le n -uplet du temps $t_1 < t_2 \dots < t_n$ tel que $t_i \in T$ et pour tout temps $h \in T$ avec $t_{i+h} \in T \forall i, i = 1 \dots n$. La suite $(X_{t_1+h}, \dots, X_{t_n+h})$ a la même loi de probabilité que $(X_{t_1}, \dots, X_{t_n})$. La loi de probabilité qui correspond à la suite $(X_{t_1}, \dots, X_{t_n})$ est caractérisée par sa fonction de répartition d'ou la définition équivalente de la stationnarité forte.*

$$\forall (X_{t_1}, \dots, X_{t_n}), \forall (t_1, t_2, \dots, t_n), \forall h$$

$$P[X_{t_1} < x_1, \dots, X_{t_n} < x_n] = P[X_{t_1+h} < x_1, \dots, X_{t_n+h} < x_n]. \quad (1.12)$$

Ainsi une série chronologique est strictement stationnaire si toutes ses caractéristiques, c'est-à-dire tout ses moments sont invariants pour tout changement à l'origine du temps.

1.4.3 Stationnarité faible

Définition 1.6 *Une série chronologique est faiblement stationnaire si et seulement si :*

1. $E(X_t^2) < \infty$.
2. $E(X_t) = E(X_s) = \mu$.
3. $Cov(X_s, X_t) = Cov(X_{s+k}, X_{t+k})$.

Autrement dit, X_t est faiblement stationnaire si seul les moments d'ordre 1 et 2 sont stationnaires, par exemple, si $E(X_t^3)$ dépend du temps t alors le processus est faiblement stationnaire.

1.4.4 Fonction d'auto-covariance

Définition 1.7 *La fonction d'autocovariance du processus stationnaire $X_t, t \in \mathbb{Z}$, notée $\gamma(h)$, est définie par :*

$$\gamma(h) = cov(X_t, X_{t-h}) = E[(X_t - E(X_t))(X_{t-h} - E(X_{t-h}))], \forall t, h \in \mathbb{Z}. \quad (1.13)$$

Remarque 1.2 *Nous remarquons que pour $h = 0$, $Var(X_t) = \gamma(0)$ et que :*

1. $\gamma(-h) = \gamma(h), \quad \forall h \in \mathbb{Z}$ (la fonction d'auto-covariance est symétrique).
2. $|\gamma(h)| \leq \gamma(0) = Var(X_t), \quad \forall t, h \in \mathbb{Z}$.

1.4.5 Fonction d'auto-corrélation

Définition 1.8 La fonction d'auto-corrélation au pas h ($h \in \mathbb{Z}$) d'un processus stationnaire du second ordre de moyenne $E(X_t) = \mu$, notée $\rho(h)$, est définie par :

$$\rho(h) = \frac{\text{Cov}(X_t, X_{t-h})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t-h})}} = \frac{\text{Var}(X_t)}{\text{Var}(X_{t-h})} \quad \forall h \in \mathbb{Z} \quad (1.14)$$

- Propriété 1.1**
- $\rho(0) = 1 \quad \forall h \in \mathbb{Z}$.
 - $|\rho(h)| \leq 1 \quad \forall h \in \mathbb{Z}$.
 - $\rho(h) = \rho(-h) \quad \forall h \in \mathbb{Z}$.

Remarque 1.3

- La représentation graphique de $\rho(h)$ est appelée *corrélogramme*.
- Si la fonction d'autocorrélations $\rho(h)$ décroît rapidement quand le nombre de retards augmente, cela signifie que la série est stationnaire.

1.5 MODÈLE CLASSIQUE DE BOX ET JENKINS

La méthode de Box-Jenkins est une approche largement utilisée pour l'analyse et la prévision des séries chronologiques. Cette méthode a été développée par George Box et Gwilym Jenkins dans les années 1970 et a depuis été largement appliquée dans de nombreux domaines. Elle est basée sur la modélisation des parties autorégressive (AR) et moyenne mobile (MA) des séries chronologiques et sur l'utilisation de tests statistiques pour vérifier l'adéquation du modèle aux données.

En utilisant la méthode de Box-Jenkins, il est possible de prévoir les valeurs futures de la série chronologique et d'évaluer la qualité des prévisions obtenues. Cette méthode est particulièrement adaptée aux séries chronologiques présentant une saisonnalité et/ou une tendance, car elle permet de modéliser ces composantes de manière précise.

1.5.1 Élimination de la tendance et de la composante saisonnière par différenciation

Opérateur retard et opérateur différence

L'opérateur retard B décale le processus d'une unité de temps vers le passé :

$$BX_t = X_{t-1}$$

Si on applique j fois cet opérateur, on décale le processus de j unités de temps :

$$B^j X_t = \underbrace{B(B(\dots(BX_t)\dots))}_{j \text{ fois}} = X_{t-j}$$

L'opérateur différence Δ fait la différence entre le processus et sa version décalée d'une unité de temps :

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

Élimination de la tendance

L'opérateur différence Δ élimine les tendances linéaires. Par exemple, pour un processus de la forme :

$$X_t = a + bt + e_t$$

On a

$$\Delta X_t = -b + e_t - e_{t-1}.$$

De façon générale, l'opérateur Δ^d élimine les tendances polynomiales de degré d . Par exemple, pour une tendance de degré 2

$$\Delta^2 X_t = \Delta^2(a + bt + ct^2 + e_t) = 2c + (e_t - 2e_{t-1} + e_{t-2})$$

Élimination de la composante saisonnière

L'opérateur $\Delta^d = (1 - B^d)$ élimine une saisonnalité de période d . Par exemple, pour un modèle général,

$$X_t = m_t + s_t + e_t,$$

où s_t est de période d , on obtient

$$\Delta^d X_t = m_t - m_{t-d} + e_t - e_{t-d}$$

où $m_t - m_{t-d}$ représente la tendance et $e_t - e_{t-d}$ le bruit.

1.5.2 Étapes de la méthode de Box et Jenkins

La méthode de Box et Jenkins consiste en quatre étapes principales : *l'identification, l'estimation, le diagnostic et la prévision.*

Identification du modèle

La première étape dans la méthodologie Box-Jenkins est d'examiner le graphe chronologique de la série pour déterminer son comportement (stationnarité, tendance, saisonnalité, etc.). Si la série présente une tendance et/ou une saisonnalité, des transformations sont nécessaires pour stationnariser la série. Pour identifier les ordres p et q , la comparaison entre la structure des corrélations estimées de la série à travers le corrélogramme et la structure de corrélation théorique exhibée par des modèles bien connus est effectuée. La fonction d'autocorrélation simple et la fonction d'autocorrélation partielle peuvent indiquer la présence d'un modèle de moyenne mobile ou autorégressif, respectivement. Si la fonction d'autocorrélation simple décroît rapidement vers 0 et que la fonction d'autocorrélation partielle présente un cut-off après p retards, on peut conclure que la série provient d'un processus AR d'ordre p (AR(p)). Si la fonction d'autocorrélation simple présente un cut-off après q retards et que la fonction d'autocorrélation partielle décroît rapidement vers 0, alors on peut conclure que la série est générée à partir d'un modèle de moyenne mobile d'ordre q (MA(q)). Si les fonctions d'autocorrélation simple et partielle présentent une forme exponentielle ou sinusoidale, on est en présence d'un processus autorégressif-moyenne mobile ARMA(p, q). Cette étape est complexe et nécessite de l'expertise, mais il existe des méthodes d'identification automatiques basées sur le critère d'information.

Critère d'information

Des critères d'information sont disponibles pour guider le choix du modèle, évitant ainsi la sélection arbitraire des paramètres p et q . Les critères d'information mesurent l'écart entre la loi inconnue réelle et celle du modèle proposé. Différentes estimations de la qualité de l'information ont été proposées, telles que :

1. Critère d'Akaike (1973) appelé aussi AIC tel que :

$$AIC = \log \sigma^2 + 2(p + q)/N$$

2. Critère bayésien (1978) appelé aussi BIC tel que :

$$BIC = \log \sigma^2 + (p + q) \log(N)$$

Estimation des paramètres

Après l'étape d'identification, il est nécessaire d'estimer les paramètres du modèle, tels que les coefficients des polynômes AR et MA ainsi que la variance des résidus σ^2 . La méthode d'estimation la plus courante est celle du maximum de vraisemblance ou des moindres carrés. Cette méthode consiste à maximiser le logarithme de la fonction de vraisemblance par rapport aux paramètres θ_i et θ_j (avec $i = 1 \dots p, j = 1 \dots q$), permettant ainsi de trouver les valeurs numériques les plus plausibles pour ces paramètres. Une fois l'étape d'estimation terminée, il est important de valider le modèle estimé.

Validation

Une fois l'étape d'estimation terminée, nous disposons de plusieurs modèles ARMA dont les paramètres ont été estimés. La prochaine étape consiste à valider ces modèles afin de les comparer et de déterminer lequel est le plus approprié. Pour ce faire, nous effectuons des tests sur les paramètres et les résidus de chaque modèle. Si plusieurs modèles sont validés, nous devons poursuivre la validation en comparant leur qualité respective.

Prévision

Soit $(X_t)_{t \in \mathbb{Z}}$ un processus au second ordre réel centré dont la fonction d'autocovariance est notée $\gamma(i, j) = E(X_i X_j)$. Soit (X_1, X_2, \dots, X_n) un échantillon de $(X_t)_{t \in \mathbb{Z}}$. On note $H_n = [X_1, X_2, \dots, X_n]$ le sous-espace fermé de $L^2(\Omega, \mathcal{A}, P)$ engendré par $(X_j)_{1 \leq j \leq n}$. On pose que $X = 0$ et $X_j = H_{n-1}(X_j)$.

On suppose que la matrice $K(i, j)$ est définie positive, alors on montre que $H(n) = [(X_1 - X), \dots, (X_n - X)]$ et on déduit que :

$$X_{n+1} = \sum_{j=1}^n \theta_{n,j} (X_{n+1-j} - \hat{X}_{n+1-j}) \quad \text{pour } n \geq 1 \quad (1.15)$$

(c'est-à-dire la prévision d'horizon d'ordre 1)

Pour des prévisions d'ordre $h \geq 1$, on a :

$$X_{n+h} = \sum_{j=0}^{h-1} \theta_{n+h-1,j} (X_{n+h-j} - \hat{X}_{n+h-j}) \quad \text{pour } n \geq 1 \quad (1.16)$$

Si on écrit (X_t) sous forme d'une moyenne mobile infinie :

$$X_t = \varepsilon_t + \sum_{j=1}^{\infty} b_j \varepsilon_{t-j} \quad (1.17)$$

On a l'intervalle de confiance au niveau $\alpha = 5\%$:

$$X_t \in \left[X_{t+h} - 1.96\sigma_\varepsilon \sqrt{\sum_{j=0}^{p-1} b_j^2}, X_{t+h} + 1.96\sigma_\varepsilon \sqrt{\sum_{j=0}^{p-1} b_j^2} \right]$$

1.6 MODÈLE ARMA

Dans cette section, nous examinons les modèles linéaires proposés par Box et Jenkins, qui sont couramment utilisés pour modéliser une série chronologique X_t . Ces modèles incluent l'autorégression, la moyenne mobile et l'ARMA qui combine ces deux modèles.

1.6.1 Modèle mixte ARMA

Définition 1.9 Soit $(X_t)_{t \in \mathbb{Z}}$ un processus stationnaire au second ordre et centré. $(X_t)_{t \in \mathbb{Z}}$ est un processus ARMA(p, q) s'il existe un bruit blanc ε_t défini sur (Ω, \mathcal{A}, P) et des nombres réels ϕ_j (avec $\phi_0 = 1$), ϕ_1, \dots, ϕ_p et θ_j (avec $\theta_0 = 1$), $\theta_1, \dots, \theta_q$ tels que :

$$\sum_{i=0}^p \phi_i X_{t-i} = \sum_{j=0}^q \theta_j \varepsilon_{t-j} \quad (1.18)$$

avec ε_t un bruit blanc de moyenne nulle et de variance σ_ε^2 .

1.6.2 Modèle AR

Définition 1.10 On dit que X_t est un modèle autorégressif ("Auto-Regressive") d'ordre p ($p \geq 1$), abrégé AR(p) si :

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \dots + \Phi_p X_{t-p} + \varepsilon_t = \sum_{j=1}^p \Phi_j X_{t-j} + \varepsilon_t \quad (1.19)$$

ou $\Phi_1, \Phi_2, \dots, \Phi_p$ sont des constantes ($\Phi_p \neq 0$) et $\varepsilon_t \sim BB(0, \sigma_\varepsilon^2)$. On utilise généralement la notation suivante :

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

avec $\Phi(L) = 1 - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p$

On dit dans ce cas, qu'un modèle AR(p) est stationnaire si toutes les racines du polynôme caractéristique $\Phi(z)$ sont de module supérieur strictement à 1.

1.6.3 Modèle MA

Définition 1.11 On dit qu'un processus $(X_t)_{t \in \mathbb{Z}}$ est un processus à moyenne mobile d'ordre q , noté MA(q), s'il vérifie l'équation suivante avec $p = 0$, c'est-à-dire :

$$\forall t \in \mathbb{Z}, X_t = \sum_{j=0}^q \theta_j \varepsilon_{t-j} \quad (1.20)$$

où (ε_t) est un bruit blanc de moyenne nulle et de variance σ_ε^2 .

Ce processus est toujours stationnaire. Nous allons caractériser les processus AR(p) et donner, sans démonstration, quelques résultats sur les processus ARMA(p,q).

Les processus AR(p) jouent un rôle important dans les applications car ce sont ceux qui, à tout instant t , peuvent être extrapolés linéairement à partir des p valeurs précédentes X_{t-1}, \dots, X_{t-p} à un bruit blanc près. Si A désigne le polynôme unitaire (c'est-à-dire $A(0) = 1$) de degré p

$$A(z) = 1 + \sum_{i=1}^p \phi_i z^i \quad (1.21)$$

1.6.4 Causalité

Définition 1.12 *Un modèle ARMA(p,q) est causal si la série chronologique $\{x_t, \pm 1, \pm 2, \dots\}$ peut s'écrire comme un processus linéaire unilatéral.*

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} = \psi(B)w_t \quad (1.22)$$

ou $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ et $\sum_{j=0}^{\infty} |\psi_j| \leq \infty$; On pose $\psi_0 = 1$

Propriété 1.2 *Un modèle ARMA(p, q) est causal si et seulement si $\Phi(z) \neq 0$ pour $|z| \leq 1$ Les coefficients du processus linéaire donnés dans (1.22) peuvent être déterminés en résolvant :*

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\Phi(z)}, \quad |z| \leq 1 \quad (1.23)$$

1.6.5 Inversibilité

Définition 1.13 *Un modèle ARMA(p,q) est inversible si la série chronologique $\{x_t, \pm 1, \pm 2, \dots\}$ peut s'écrire :*

$$\pi(B)x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = w_t \quad (1.24)$$

ou $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$, et $\sum_{j=0}^{\infty} |\pi_j| \leq \infty$, on pose $\pi_0 = 1$

Propriété 1.3 *Un modèle ARMA(p, q) est inversible si et seulement si $\theta(z) \neq 0$ et $|z| \leq 1$ le coefficient π_j de $\pi(B)$ donné dans (1.24) peut être déterminé en résolvant :*

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\Phi(z)}{\theta(z)}, \quad |z| \leq 1 \quad (1.25)$$

Autrement dit, le processus ARMA est inversible uniquement lorsque les racines de $\theta(z)$ se trouvent à l'extérieur du cercle unité; c'est-à-dire, $\theta(z) = 0$ uniquement lorsque $|z| \geq 1$.

1.7 MÉTHODES DE VÉRIFICATION DE LA STATIONNARITÉ

Il existe plusieurs méthodes pour vérifier la stationnarité d'une série chronologique, les plus couramment utilisés sont :

- Analyse visuelle de la série : Analyser graphiquement la série temporelle et observer si il y'a des tendances ou des variations systématiques au fil du temps.

- Décomposition de la série : Comme nous l'avons vu plus haut : on peut séparer la série en ses composantes de tendances, saisonnalité et résiduelle et ainsi examiner si la composante résiduelle présente des caractéristique de stationnarité.
- Il existe aussi des tests statistiques qui sont utilisé pour vérifier la stationnarité comme le test de Philippe-Perron et le test de Dickey-Fuller augmenté (ADF).

Dans cette section nous allons présenter comme méthode de vérification de la stationnarité le test de Dickey-Fuller augmenté (ADF).

1.7.1 Test de Dickey-Fuller

Le test de Dickey-Fuller est un test statistique couramment utilisé pour tester la présence d'une racine unitaire dans un ensemble de données de séries temporelles. L'hypothèse nulle du test est qu'il existe une racine unitaire dans la série temporelle, ce qui implique que la série est non stationnaire et présente une tendance. Le test de Dickey-Fuller est basé sur l'équation de modèle suivante :

$$Y_t = c + \beta t + \alpha Y_{t-1} + \Phi \Delta Y_{t-1} + \varepsilon_t \quad (1.26)$$

L'hypothèse nulle du test de Dickey-Fuller est que $\alpha = 1$, ce qui implique la présence d'une racine unitaire dans la série temporelle. L'hypothèse alternative est que $\alpha < 1$, ce qui indique que la série temporelle est stationnaire et ne possède pas de racine unitaire.

Le test de Dickey-Fuller augmenté (Augmented Dickey-Fuller test en anglais) est une évolution du test de Dickey-Fuller basé sur l'équation mentionnée précédemment. Il s'agit de l'une des formes les plus courantes de test de racine unitaire.

1.7.2 Test de Dickey-Fuller augmenté

Le test de Dickey-Fuller augmenté (ADF) est une extension du test de Dickey-Fuller (DF) qui tient compte des processus autorégressifs d'ordre supérieur et d'autres variables pouvant affecter la série temporelle. Le test de DF est basé sur une régression de la première différence de la série temporelle sur ses valeurs passées, et la statistique de test est comparée à des valeurs critiques d'une table pour déterminer la significativité statistique.

En revanche, le test ADF ajoute des termes retardés supplémentaires de la première différence de la série temporelle à l'équation de régression du test DF pour tenir compte des processus autorégressifs d'ordre supérieur. Le test ADF inclut également d'autres variables, telles qu'une tendance linéaire, pour prendre en compte les facteurs qui peuvent affecter la série temporelle.

L'hypothèse nulle du test ADF est similaire à celle du test DF, selon laquelle la série temporelle possède une racine unitaire, tandis que l'hypothèse alternative est que la série temporelle est stationnaire. L'équation de régression du test ADF prend la forme suivante :

$$Y_t = c + \beta t + \alpha Y_{t-1} + \Phi_1 \Delta Y_{t-1} + \Phi_2 \Delta Y_{t-2} + \dots + \Phi_p \Delta Y_{t-p} + \varepsilon_t \quad (1.27)$$

où ΔY_t représente la première différence de la série temporelle à l'instant t , Y_{t-1} est le niveau retardé de la série temporelle à l'instant $t-1$, α est le coefficient des

termes retardés, βt est un terme de tendance linéaire, Φ_1, Φ_2 et Φ_p sont les coefficients des différences des termes retardés respectifs et ε_t est l'erreur.

Les étapes du test

1. Tout d'abord, nous soustrayons la valeur retardée de Y de la valeur actuelle de Y pour obtenir la première différence, notée ΔY_t . Ensuite, nous remplaçons Y_{t-1} par ΔY_{t-1} dans l'équation de régression.
2. Nous estimons l'équation de régression à l'aide des moindres carrés ordinaires (MCO) pour obtenir les coefficients des variables dans l'équation, y compris le coefficient de tendance β_t , le coefficient de la valeur retardée de Y , et les coefficients α des différences retardées de Y , $\Phi_1, \Phi_2, \dots, \Phi_p$.
3. Nous calculons la statistique de test en utilisant la statistique t du coefficient α . La statistique de test est ajustée en fonction du nombre de différences retardées incluses dans l'équation de régression. La statistique de test est donnée ci-dessous.
4. Nous comparons la statistique de test aux valeurs critiques d'une table qui dépendent de la taille de l'échantillon et du niveau de signification choisi. Si la statistique de test est inférieure à la valeur critique, nous rejetons l'hypothèse nulle selon laquelle la série temporelle a une racine unitaire et concluons que la série temporelle est stationnaire. Si la statistique de test est supérieure à la valeur critique, nous ne rejetons pas l'hypothèse nulle et concluons que la série temporelle a une racine unitaire et est non stationnaire.

La statistique de test ADF est calculée comme suit :

$$ADF = \frac{\alpha - 1}{SE(\alpha)}$$

où $SE(\alpha)$ représente l'écart-type de β . La distribution de la statistique ADF dépend du nombre d'observations, de la présence de termes d'erreur autocorrélés et du choix du nombre de retards inclus dans le modèle.

Étapes de réalisation du test de Dickey Fuller dans R

1. Création des données de la série temporelle :

```
install.packages("tseries")
library("tseries")
# Creation des données
vect <- c(3, 8, 2, 1, 3, 3, 9, 8, 7, 3, 10, 3, 4)
```

2. Visualisation de la série temporelle : Avant de pouvoir effectuer le test augmenté de Dickey-Fuller, nous créons d'abord un graphique pour visualiser les données créées.

```
plot(vect, type='l')
```

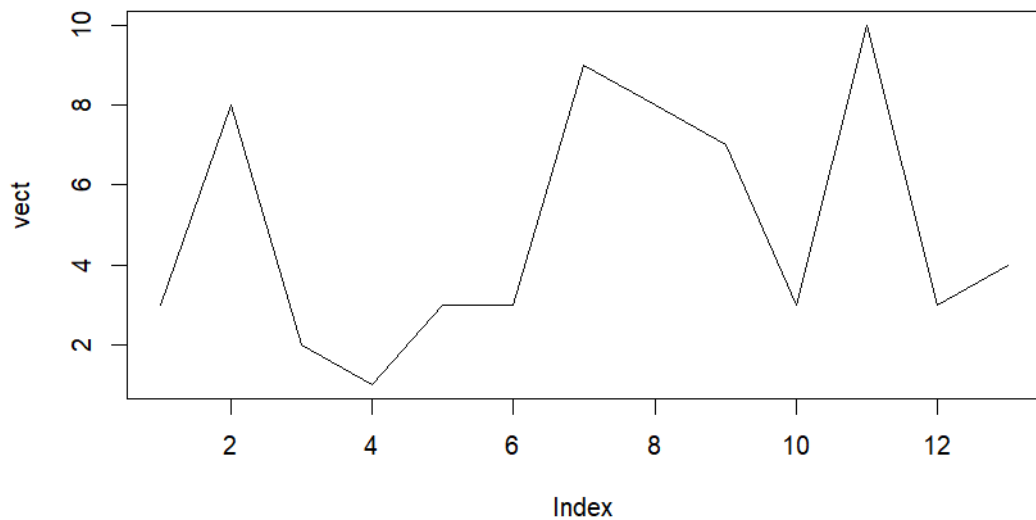


FIGURE 1.4 – Série temporelle pour test de dickey et Fuller

- Réalisation du test de Dickey-Fuller augmenté : pour effectuer le test on utilise la fonction **adf.test()** fournit dans la bibliothèque "tseries".

```
adf.test(vect)
```

- Résultats du test : On obtiens les résultats suivants :

```
Augmented Dickey-Fuller Test
```

```
data: vect
Dickey-Fuller = -1.6846, Lag order = 2, p-value = 0.6925
alternative hypothesis: stationary
```

Interprétation des résultats

La statistique de test et la valeur p s'avèrent être respectivement égales à -1,6846 et 0,6925. Étant donné que la valeur p est supérieure ou égale à 0,05, nous ne pouvons pas rejeter l'hypothèse nulle. Cela implique que la série temporelle n'est pas stationnaire. En d'autres termes, on peut dire qu'elle possède une structure dépendante du temps et ne présente pas de variance constante au fil du temps.

CONCLUSION DU CHAPITRE

En conclusion de ce premier chapitre, nous avons jeté les bases essentielles pour comprendre les séries chronologiques. Nous avons exploré en détail les notions de stationnarité, en nous concentrant notamment sur la distinction entre la stationnarité faible et forte, ainsi que l'importance cruciale du bruit blanc dans l'analyse de ces séries. Ce chapitre introductif nous a ainsi fourni une solide compréhension des principes fondamentaux nécessaires pour la suite de notre étude. Les chapitres à venir nous permettront d'approfondir notre analyse en explorant des méthodes et des techniques avancées afin de modéliser, prévoir et interpréter avec finesse les séries chronologiques à valeurs entières.

MODÈLE AUTORÉGRESSIF D'ORDRE 1 À VALEURS ENTIÈRES INAR(1)

2

2.1 INTRODUCTION

Au cours de ce deuxième chapitre, nous allons explorer une version à valeurs entières du modèle autoregressif d'ordre 1, le modèle INAR(1) introduit par McKenzie (1985, 1988) et Al-Osh et Alzaid (1987) afin d'analyser et modéliser les séries chronologiques de comptage.

2.2 PROCESSUS DE COMPTAGE

Définition 2.1 *Un processus de comptage $\{N_t; t \in \mathbb{Z}\}$ est un type de processus stochastique qui modélise le nombre d'événements se produisant dans un intervalle de temps donné $[0, t]$. Un processus stochastique est dit processus de comptage si :*

1. $N_t \in \mathbb{N}, \forall t \in \mathbb{R}^+$
2. $\forall t > s, N(t) > N(s)$

Dans un processus de comptage, les valeurs qu'on observe sont des entiers positifs, et la distribution des événements peut être dépendante du temps. Dans un processus de comptage de Poisson par exemple, le nombre d'événements se produisant dans un intervalle de temps donné suit une distribution de Poisson avec un taux moyen constant. Dans d'autres cas, la distribution des événements peut varier dans le temps, dans ce cas cela peut être modélisé à l'aide de processus de comptage plus complexes tels que les processus autorégressifs d'ordre un à valeurs entières (INAR(1)).

2.2.1 Notations et caractéristiques préliminaires des processus de comptages

Quelques notations et caractéristiques spécifiques aux processus de comptage seront retrouvées dans les deux tableaux suivants :

TABLE 2.1 – Définition et notation des moments d'une variable aléatoire de comptage X .

Notation	Moment	Description
μ_n	$E[X^n]$	Le n-ème moment de X
μ_n	$E[(X - \mu)^n]$	Le n-ème moment centré de X
$\mu(n)$	$E[X \cdots (X - n + 1)]$	Le n-ème moment factoriel de X

TABLE 2.2 – Définition et notation des fonctions génératrices d'une variable aléatoire de comptage X .

Fonction	Définition
Probabilité (pgf)	$pgf(z) = E[z^X] = \sum_{k=0}^{\infty} p_k \cdot z^k$
Moment (mgf)	$mgf(z) = pgf(e^z) = 1 + \sum_{j=1}^{\infty} \frac{\mu_j}{j!} \cdot z^j$
Cumulant (cgf)	$cgf(z) = \ln(pgf(e^z)) = \sum_{j=1}^{\infty} \frac{\kappa_j}{j!} \cdot z^j$
Factoriel-cumulant (fcgf)	$fcgf(z) = \ln(pgf(1+z)) = \sum_{j=1}^{\infty} \frac{\kappa_{(j)}}{j!} \cdot z^j$

2.2.2 Domaines d'application

Les processus de comptage ont un large domaine d'application, on peut citer par exemple :

- Les sciences économiques, la finance : Les séries de comptages peuvent être utilisées pour modéliser des données financières telles que le nombre de transactions boursières, le nombre de faillites, les sinistres d'assurance, les défaillances d'entreprises, etc.
- La médecine : Suivre le nombre de cas de maladies, le nombre de patients par jour dans un hôpital.
- Télécommunication : pour compter le nombre de communications dans un intervalle de temps donné.

Voici quelques exemples d'utilisation des séries de comptages dans la littérature :

- Kennan (1985) donne le nombre mensuel de grèves contractuelles aux États-Unis. Dans son analyse, Kennan se concentre sur la durée des grèves plutôt que sur leur nombre en soi.
- McCullagh et Nelder (1989) examinent l'incidence de certains dommages de navires causés par les vagues en utilisant les données fournies par une compagnie d'assurance. Ils modélisent le nombre d'incidents quel que soit le niveau de dommage.
- Dionne, Gagné, Gagnon et Vanasse (1997) ont étudié la fréquence des accidents (et incidents) de ligne par transporteur au Canada sur une base trimestrielle entre 1974 et 1988.

2.2.3 Quelques processus de comptages

Ci-dessous les processus de comptages les plus couramment utilisés :

Distribution de Poisson

Définition 2.2 Soit X une variable aléatoire ayant une distribution discrète définie sur $\mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$. X suit une distribution de Poisson de paramètre λ , notée $X \sim \text{Poisson}(\lambda)$, si et seulement si sa densité de probabilité est la suivante :

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad \text{pour } \lambda \in \mathbb{R}^+, k = 0, 1, 2, \dots \quad (2.1)$$

Les moments de premier et de second ordre de la distribution de Poisson sont :

$$E(X) = \lambda \quad \text{Var}(X) = \lambda$$

L'égalité entre la moyenne et la variance est une caractéristique de la distribution de Poisson. Cette égalité est appelée équidispersion. Les écarts par rapport à l'équidispersion peuvent être soit une surdispersion (la variance est supérieure à la moyenne), soit une sous-dispersion (la variance est inférieure à la moyenne). Toute violation de cette égalité indique une violation de l'hypothèse de Poisson.

Distribution binomiale négative

Définition 2.3 Une variable aléatoire X suit une distribution binomiale négative avec les paramètres $\alpha \geq 0$ et $\theta \geq 0$, notée $X \sim \text{Negbin}(\alpha, \theta)$, si la fonction de probabilité est donnée par

$$P(X = k) = \frac{\Gamma(\alpha + k)}{\Gamma(\alpha)\Gamma(k + 1)} \left(\frac{1}{1 + \theta} \right)^\alpha \left(\frac{\theta}{1 + \theta} \right)^k, \quad k = 0, 1, 2, \dots \quad (2.2)$$

où $\Gamma(\cdot)$ représente la fonction gamma telle que $\Gamma(s) = \int_0^\infty z^{s-1} e^{-z} dz$ pour $s > 0$.

Les moments de premier et de second ordre sont :

$$E(X) = \alpha\theta, \quad \text{Var}(X) = \alpha\theta(1 + \theta) = E(X)(1 + \theta)$$

Distribution Binomiale

Définition 2.4 Une variable aléatoire X suit une distribution binomiale avec les paramètres $n \in \mathbb{N}$ et $p \in (0, 1)$, notée $X \sim B(n, p)$, si la fonction de probabilité est donnée par :

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n \quad (2.3)$$

Les moments de premier et de second ordre sont :

$$E(X) = np \quad \text{Var}(X) = np(1 - p)$$

Distribution logarithmique

Définition 2.5 La variable aléatoire X suit une distribution logarithmique si :

$$P(X = k) = \frac{\alpha\theta^k}{k}, \quad k = 1, 2, \dots, \quad 0 < \theta < 1 \quad (2.4)$$

Les moments de premier et second ordre sont :

$$E(X) = \alpha\theta(1 - \theta)^{-1} \quad \text{Var}(X) = \alpha\theta(1 - \alpha\theta)(1 - \theta)^{-2}$$

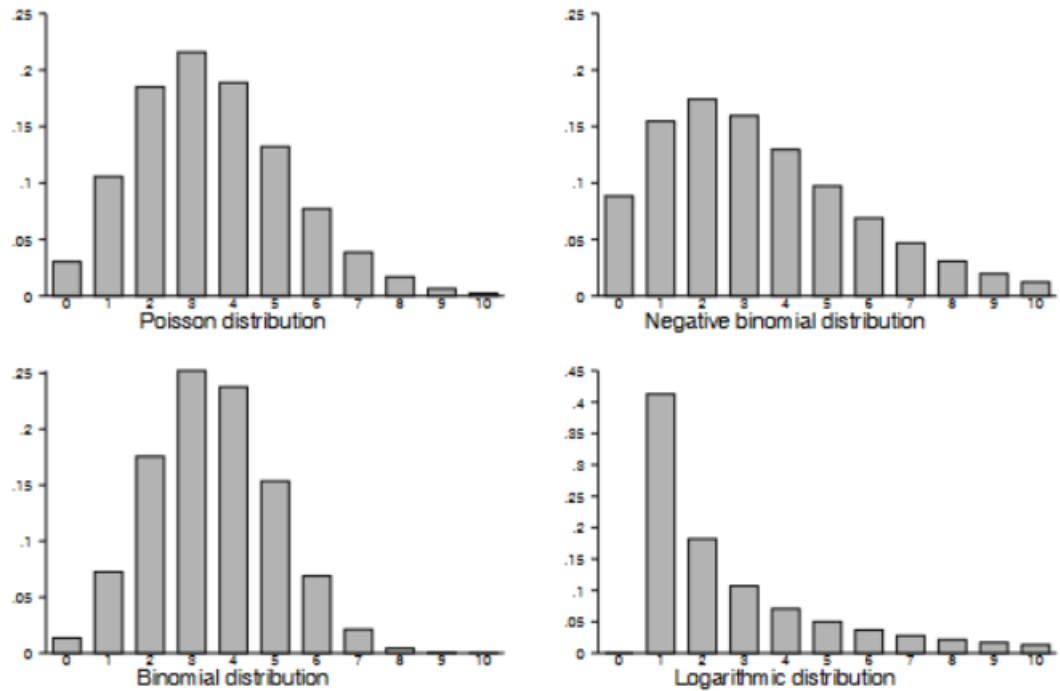


FIGURE 2.1 – Distributions de données de comptage ($E(X) = 3.5$)

2.2.4 Autoregression et dépendance à 1 pas

Une dépendance a un pas (lag 1), également appelée dépendance temporelle de premier ordre, fait référence à la relation entre une variable et sa valeur précédente immédiate. Cela signifie que la variable dépend linéairement de sa valeur au pas de temps précédent, c'est-à-dire du pas 1. Dans le contexte des séries temporelles, une dépendance à un pas peut être observée dans un modèle autorégressif (AR) d'ordre 1. Cette dépendance peut être importante dans l'analyse des séries temporelles car elle indique qu'il existe une influence significative de la valeur précédente sur la valeur actuelle. Cela peut avoir des implications sur la prévision et l'interprétation des données temporelles.

2.3 LE MODELE INAR(1) (*the integer-valued autoregressive of order 1*)

Avant de définir le modèle autorégressif d'ordre 1 à valeurs entières (INAR 1), nous devons d'abord introduire une notion fondamentale, celle de l'opérateur d'amincissement binomiale.

2.3.1 L'opérateur d'amincissement binomiale "o"

Une multitude d'opérateurs d'amincissement ont déjà été défini, on peut citer l'opérateur d'amincissement binomiale étendue, l'opérateur d'amincissement de Cox, cependant le plus connu est celui qui a été défini par Stutel et van Harn (1979). On l'appelle opérateur d'amincissement binomiale (thinning operator)

Définition 2.6 Soit X une variable aléatoire à valeurs entières non négatives ; alors pour tout $\alpha \in [0, 1]$, l'opérateur d'amincissement ' \circ ' est défini par :

$$\alpha \circ X = \sum_{i=1}^X Y_i \quad (2.5)$$

où Y_i est une séquence de variables aléatoires de Bernouilli i.i.d. indépendantes de X de paramètre $\alpha \in [0, 1]$, telles que $P(Y_i = 1) = 1 - P(Y_i = 0) = \alpha$. Cette suite est appelé série de comptage de $\alpha \circ X$. Ainsi la variable aléatoire $\alpha \circ X$ suit une loi binomiale de paramètre X et α .

Propriété 2.1 À partir de la définition de l'opérateur ' \circ ', il est clair que :

1. $0 \circ X = 0$.
2. $1 \circ X = X$.
3. $E(\alpha \circ X) = \alpha E(X)$.
4. $V(\alpha \circ X) = \alpha^2 V(X) + \alpha(1 - \alpha)E(X)$.
5. Pour tout $\beta \in [0, 1]$, $\beta \circ (\alpha \circ X) = (\beta\alpha) \circ X$.
6. $Cov(\alpha \circ X, X) = \alpha \cdot V(X)$.

Démonstration. — Les deux premières propriétés découlent de la définition même de l'opérateur d'amincissement.

en effet :

Quand $\alpha = 0$, cela veut dire que Y_i est une variable aléatoire de Bernouilli avec comme probabilité de succès 0 . alors, quelque soit la valeur de X , la somme $\sum_{i=1}^X Y_i$ sera toujours égale à zéro, car chaque terme dans la somme sera égal à zéro. Ainsi, on peut effectivement dire que $0 \circ X = 0$ lorsque $\alpha = 0$.

De la même manière, on peut démontrer que $1 \circ X = X$, Y_i étant une variable aléatoire de Bernouilli avec $\alpha = 1$ comme probabilité de succès.

— Pour prouver que $E(\alpha \circ X) = \alpha E(X)$ on va se référer à Weiss qui a pour cela utilisé la loi de l'espérance totale :

$$E[\alpha \circ X] = E[E[\alpha \circ X|X]]$$

□

Remarque 2.1 L'opération d'amincissement peut également être définie dans un contexte multivarié. Ainsi soit $X = [X_1 \dots X_p]^T$ un vecteur aléatoire et $A = a_{ij}$ une matrice $(p \times p)$ vérifiant $0 \leq a_{ij} \leq 1$ pour $i, j = 1 \dots p$ alors l'opération d'amincissement vectorielle $A \circ X$, est définie comme un vecteur aléatoire dont la i ème composante est donnée par (Franke et Subba Rao, 1995).

$$[A \circ X]_i = \sum_{j=1}^p a_{ij} \circ X_j, \quad i = 1 \dots p \quad (2.6)$$

où les séries de comptage de tout les $a_{ij} \circ X_j$, $i, j = 1 \dots p$, sont supposées indépendantes.

Notez que l'opération d'amincissement vectoriel satisfait des propriétés analogues à l'opération univariée. Pour un compte rendu de ces propriétés, voir Franke et Subba Rao (1995), Latour (1997) et Oliveira (2000)

2.3.2 Modèle INAR(1)

Maintenant, nous définissons le processus INAR(1) $\{X_t : t = 0 \pm 1 \pm 2 \dots\}$ par :

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \quad (2.7)$$

tel que :

$$\alpha \circ X_{t-1} = \sum_{j=1}^{X_{t-1}} Y_j \quad (2.8)$$

où $\alpha \in [0, 1]$ et ϵ_t est une suite de variables aléatoires à valeurs entières non négatives non corrélées, ayant une moyenne μ et une variance finie σ^2 .

Et (Y_j) est une suite de variables aléatoires i.i.d, suivant une loi de Bernoulli de paramètre α , indépendante de X_{t-1} , et (ϵ_t) est indépendante de (Y_j) . Par conséquent, $(\alpha \circ X_t | X_t)$ suit une loi binomiale de paramètres X_t et α .

Le processus INAR(1) est une chaîne de Markov avec la transition en une étape donnée par (McKenzie, 1985; Al-Osh & Alzaid, 1987)

$$\begin{aligned} p_{k|l} &= P(X_t = k | X_{t-1} = l) \\ &= \sum_{j=0}^{\min\{k,l\}} \binom{l}{j} \alpha^j \cdot (1 - \alpha)^{l-j} \cdot P(\epsilon_t = k - j) \end{aligned} \quad (2.9)$$

Pour la moyenne conditionnelle et la variance conditionnelle, nous avons (Alzaid & Al-Osh, 1988) :

$$\begin{aligned} E[X_t | X_{t-1}] &= \alpha \cdot X_{t-1} + \mu \epsilon \\ V[X_t | X_{t-1}] &= \alpha(1 - \alpha) \cdot X_{t-1} + \sigma^2 \epsilon \end{aligned} \quad (2.10)$$

Ce sont toutes les deux des fonctions linéaires de X_{t-1} . Pour la dérivation des équations (2.10), il convient de noter que $\alpha \circ X_{t-1}$ et ϵ_t sont indépendants. Étant donné que la moyenne conditionnelle est linéaire en X_{t-1} , le modèle INAR(1) appartient à la classe des modèles AR(1) linéaires conditionnels, comme discuté par Grunwald et al. (2000). Notez que la variance conditionnelle diffère du cas AR(1) car elle varie avec le temps.

Démonstration. — On a :

$$\begin{aligned} E(X_t | X_{t-1}) &= E(\alpha \circ X_{t-1} + \epsilon_t | X_{t-1}) \\ &= \alpha E(X_{t-1} | X_{t-1}) + E(\epsilon_t | X_{t-1}) \\ &= \alpha X_{t-1} + \mu \epsilon \end{aligned}$$

Ici, nous utilisons l'indépendance entre ϵ_t et X_{t-1} .

— On a :

$$\begin{aligned} \text{Var}(X_t | X_{t-1}) &= \text{Var}(\alpha \circ X_{t-1} + \epsilon_t | X_{t-1}) \\ &= \text{Var}(\alpha \circ X_{t-1} | X_{t-1}) + \text{Var}(\epsilon_t | X_{t-1}) \\ &= \text{Var}(\alpha \circ X_{t-1} | X_{t-1}) + \sigma^2 \\ &= \alpha(1 - \alpha) X_{t-1} + \sigma^2 \end{aligned}$$

Ici, nous utilisons la propriété 3 et 4 de l'opérateur d'amincissement binomial et l'indépendance entre ϵ_t et X_{t-1} .

□

Remarque 2.2 Si nous avons donné la distribution des innovations en termes de la fonction génératrice de probabilités (pgf), alors la distribution marginale stationnaire des observations est déterminée par l'équation (Alzaid & Al-Osh, 1988) :

$$pgf(z) = pgf(1 - \alpha + \alpha z) \cdot pgf_{\varepsilon}(z). \quad (2.11)$$

2.3.3 Propriétés de distribution du modèle INAR(1)

La moyenne et la variance du processus $\{X_n\}$ sont simplement :

$$E(X_t) = \alpha E(X_{t-1}) + \mu = \alpha^t E(X_0) + \mu \sum_{j=0}^{t-1} \alpha^j \quad (2.12)$$

et

$$\begin{aligned} \text{var}(X_t) &= \alpha^2 \text{var}(X_{t-1}) + \alpha(1 - \alpha)E(X_{t-1}) + \sigma^2 \\ &= \alpha^{2t} \text{var}(X_0) + (1 - \alpha) \sum_{j=1}^t \alpha^{2j-1} E(X_{t-j}) + \sigma^2 \sum_{j=1}^t \alpha^{2(j-1)} \end{aligned} \quad (2.13)$$

On peut voir d'après ces expressions que la stationnarité d'ordre deux requiert que la valeur initiale du processus, X_0 , ait :

$$E(X_0) = \frac{\mu}{1 - \alpha} \quad \text{et} \quad \text{var}(X_0) = \frac{\alpha\mu + \sigma^2}{1 - \alpha^2}. \quad (2.14)$$

Pour tout entier k non négatif, la covariance avec un décalage de k , $\gamma(k)$, est :

$$\begin{aligned} \gamma(k) &= \text{cov}(X_{t-k}, \alpha^k \circ X_{t-k} + \text{cov}(X_{t-k}, \sum_{j=0}^{k-1} \alpha^j \circ \varepsilon_{t-j})) \\ &= \alpha^k \text{Var}(X_{t-k}) + \sum_{j=0}^{k-1} \alpha^j \text{cov}(X_{t-k}, \varepsilon_{t-j}) = \alpha^k \gamma(0) \end{aligned} \quad (2.15)$$

L'équation (2.15) montre que la fonction d'autocorrélation, $\gamma(k)$, décroît exponentiellement avec le décalage k et a la même forme que l'équation de Yule-Walker dans le modèle AR(1). Cependant, contrairement à la fonction d'autocorrélation pour le modèle AR(1) standard, $\phi(k)$ est toujours positif pour le modèle INAR(1). La dérivation de $\gamma(k)$ se fait directement, après avoir noté que :

1. $\text{cov}(\alpha^k \circ \varepsilon_{t-k}, \alpha^s \circ \varepsilon_{t-i}) = 0$ pour $l \neq k$ et tout entier positif s .
2. $\alpha^{j+k} \circ \varepsilon_{t-j} = \alpha^k \circ \alpha^j \circ \varepsilon_{t-j}$.

En ce qui concerne les propriétés de distribution du processus INAR(1), on remarque que ce processus a des propriétés de distribution similaires à celles du modèle AR(1). Par exemple, lorsque $X_t = X_{t-1}$ et X_{t-1} est indépendant de ε_t dans (2.7), la validité de (2.7) pour tous les $\alpha \in (0, 1)$ implique que X_t a une distribution discrète auto-décomposable (voir par exemple Stuetal et Van Harn, 1979). Après avoir mentionné cette propriété, d'autres propriétés pour X_t en découlent, telles que l'unimodalité et le fait que la distribution de X_t est uniquement déterminée par la distribution de ε_t .

Il en résulte que X_t suit une distribution de Poisson si et seulement si ε_t

suit une distribution de Poisson. Par conséquent, le rôle joué par la distribution de ϵ_t dans la détermination de la distribution de X_t dans le processus INAR(1) est similaire au rôle joué par la distribution normale des erreurs dans le processus AR(1).

2.4 LE MODÈLE POISSON INAR (1)

Un des modèles les plus populaire de la famille des INAR(1) est le modèle Poisson INAR(1) qui a été introduit par McKenzie (1985) et Al-Osh & Alzaid (1987).

Définition 2.7 *Un processus du second ordre $\{X_t, t \in \mathbb{Z}\}$ appartient à la famille INAR (1) d'innovations poissonniennes POINAR (1), s'il est de la forme :*

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \quad t \in \mathbb{Z} \quad (2.16)$$

tel que :

$$\alpha \circ X_{t-1} = \sum_{j=1}^{X_{t-1}} Y_j \quad (2.17)$$

Et que les observation ϵ_t sont i.i.d qui suivent une loi de poisson de paramètre λ tel que :

$$\epsilon_t \sim \mathbb{P}(\lambda)$$

et $\mu_\epsilon = \sigma^2 = \lambda$

Les moments de ce modèle sont définis tel que :

$$E(X_t) = \frac{\lambda}{1 - \alpha} \quad V(X_t) = \frac{\lambda}{1 - \alpha}$$

Et sa fonction d'auto-corrélation :

$$\rho(k) = \text{Corr}(X_t, X_{t-k}) = \alpha^k, \quad \forall k \in \mathbb{N} \quad (2.18)$$

Remarque 2.3 *Étant donné que toutes les probabilités de transition $P(\epsilon_t = j)$ sont réellement positives, cela s'applique également à toutes les probabilités de transition $P_{k|l}$ à partir de l'équation (2.9). Par conséquent, un processus Poisson INAR(1) est une chaîne de Markov irréductible et apériodique.*

2.5 STATIONNARITÉ STRICTE

Un processus Poisson INAR(1) est une chaîne de Markov irréductible et apériodique, ce qui implique qu'il est strictement stationnaire avec une distribution marginale stationnaire. Il est bien connu que cette distribution marginale stationnaire est également une loi de Poisson, $P(\mu)$, avec :

$$\mu = \frac{\lambda}{1 - \alpha}$$

2.5.1 Propriétés

Les définitions précédentes découlent de deux propriétés d'invariance importantes de la distribution de Poisson :

- L'invariance par rapport à l'opérateur d'amincissement binomial : si X suit une distribution de Poisson avec un paramètre μ , alors $\alpha \circ X$ suit également une distribution de Poisson avec un paramètre $\alpha\mu$. En d'autres termes, si nous multiplions chaque occurrence d'un événement par un facteur α , la distribution de Poisson reste la même, seule la valeur du paramètre est ajustée.
- L'additivité : si $Z \sim P(\alpha\mu)$ et $\epsilon \sim P((1-\alpha)\mu)$ et si Z et ϵ sont indépendantes, alors $Z + \epsilon \sim P(\alpha\mu + (1-\alpha)\mu) = P(\mu)$

2.5.2 Simulation du modèle

Dans cette exemple, on va générer 100 observations à simuler, avec un taux moyen $\lambda = 2$ et un paramètre d'autoregression $\phi = 0,7$.

```
#Installer le package tseries
install.packages("tseries")
# Charger la bibliothèque requise
library(tseries)

# Définir les paramètres du modèle
lambda <- 2      # Paramètre de taux moyen des comptages
phi <- 0.7       # Paramètre d'autoregression

# Définir le nombre d'observations à simuler
n <- 100

# Initialiser un vecteur pour stocker les observations simulées
observations <- numeric(n)

# Générer les observations du modèle INAR(1)
observations[1] <- rpois(1, lambda) # Générer la première observation

for (i in 2:n) {
  observations[i] <- rpois(1, lambda + phi * lambda)
}

# Afficher les observations simulées
print(observations)

# Tracer le graphique de la série
plot(observations, type = "b", xlab = "Temps", ylab = "Observations",
     main = "Modèle INAR(1) ")
```

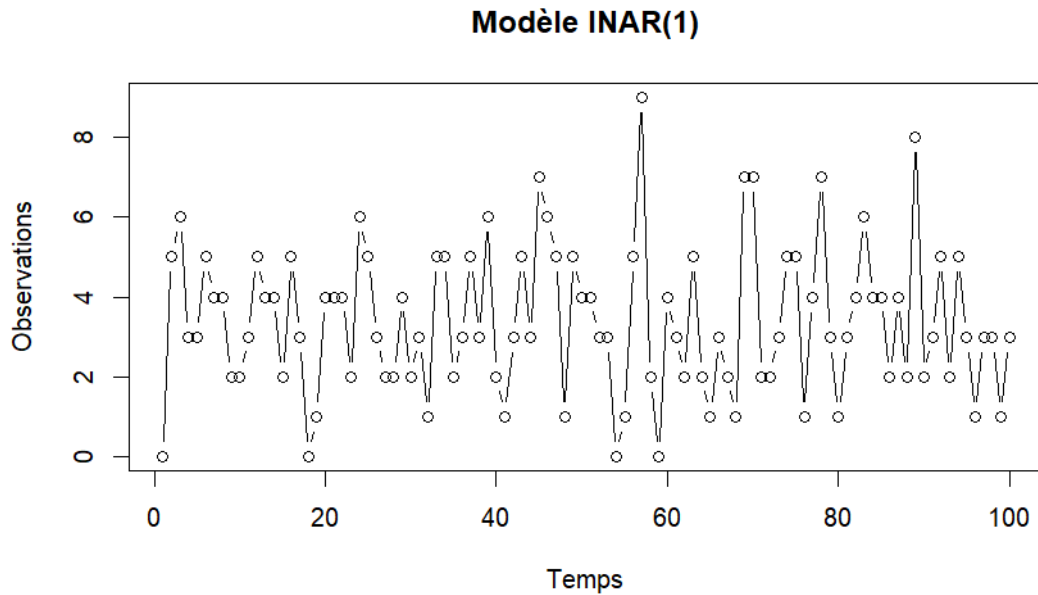


FIGURE 2.2 – Simulation d'un processus INAR(1)

2.6 EXEMPLE DE SIMULATION

(Trajectoires d'échantillon) Les Figures suivantes montre deux trajectoires d'échantillon pour des processus simulés Poisson INAR(1). Les deux modèles ont été calibrés pour donner la même moyenne observée, $\mu = 3$, mais le paramètre d'auto-corrélation α diffère, et donc la moyenne des innovations $\lambda = \mu(1 - \alpha)$. Dans la Figure 2.3, nous avons $\alpha = 0,5$ et $\lambda = 1,5$, et ce niveau modéré d'auto-corrélation devient visible à travers les mouvements à court terme en montée et en descente.

La situation dans la Figure 2.4 est beaucoup plus extrême : $\lambda = 0,15$ implique que rarement une innovation véritablement positive (et donc un mouvement ascendant) est générée. $\alpha = 0,95$ entraîne que $\alpha \circ X$ soit égal à X la plupart du temps, d'où les segments constants, et sinon à un comportement de descente lente (extinction progressive). Les segments constants sont également accompagnés d'une variance conditionnelle très petite et presque constante selon (2.6); notez que le coefficient linéaire $\alpha(1 - \alpha)$ tend vers 0 pour $\alpha \rightarrow 0$ ou $\alpha \rightarrow 1$. Ce comportement en descente lente et en morceaux constants est une caractéristique commune de nombreux modèles basés sur l'éclaircissement binomial (avec de grandes probabilités d'éclaircissement).

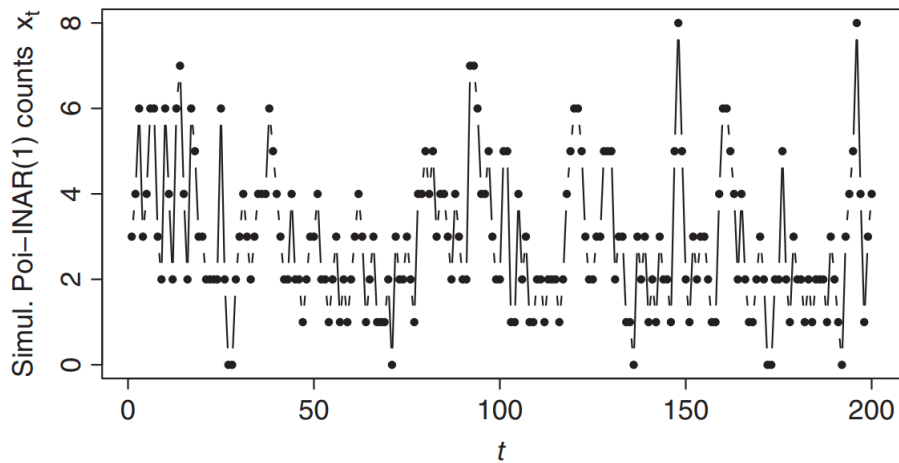


FIGURE 2.3 – Exemples de trajets simulés de processus INAR (1) Poissonien avec $\mu = 3$ et $\alpha = 0,5$

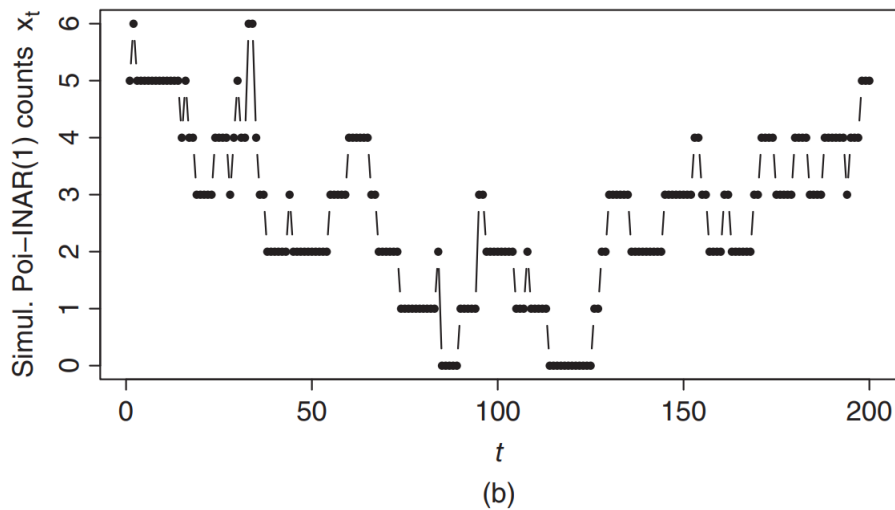


FIGURE 2.4 – Exemples de trajets simulés de processus INAR (1) Poissonien avec $\mu = 3$ et $\alpha = 0 : 95$

2.7 MODÈLES INAR(1) AVEC DES INNOVATIONS GÉNÉRALISÉES

Le modèle INAR(1) devient particulièrement simple si les innovations sont choisies selon une distribution de Poisson, Cependant, il est possible d'obtenir une plus grande flexibilité en termes de distributions marginales. Une option consiste à choisir un modèle approprié pour les observations X_t , puis à calculer la distribution correspondante des innovations à partir de l'équation (2.11); voir McKenzie (1985).

Une autre option consiste à choisir la distribution des innovations ϵ_t pour obtenir certaines propriétés pour la distribution des observations.

En général, comme indiqué dans l'équation (2.14) le comportement de dispersion des observations est facilement contrôlé par celui des innovations. De plus, la probabilité d'observer un zéro est influencée par les innovations : puisque la probabilité d'obtenir un zéro est simplement égale à $pgf(0)$. (2.11) implique que :

$$P(X = 0) = P(\epsilon = 0) \prod_{k=1}^{\infty} pgf_{\epsilon}(1 - \alpha_k). \quad (2.19)$$

Ensuite, examinons des cas particuliers. L'extension la plus naturelle au-delà des distributions de Poisson est de considérer la famille des distributions autos-décomposables discrètes (DSD) pour X_t , une distribution est dite DSD si sa fonction génératrice de probabilités (pgf) satisfait l'équation :

$$\frac{pgf(z)}{pgf(1 - \alpha + \alpha z)} \text{ est elle même une pgf pour tout } \alpha \in (0, 1). \quad (2.20)$$

D'après (2.11), (2.20) implique que les distributions DSD sont les distributions marginales d'un processus INAR(1), qui peuvent être préservées pour n'importe quel choix de $\alpha \in [0, 1]$; la pgf des innovations correspondantes est alors donnée par (2.20). Notez que toute distribution DSD est également infiniment divisible et, contrairement à la distribution de Poisson, elles présentent une surdispersion et une inflation de zéros.

2.7.1 Le Modèle Géométrique INAR(1)

La distribution géométrique $Geom(\pi)$ est DSD, ce qui en fait une distribution marginale possible pour un processus INAR(1). Si $X \sim Geom(\pi)$ alors $\alpha \circ X$ suit aussi une loi géométrique avec comme pgf :

$$pgf_{\alpha \circ X}(z) = pgf(1 - \alpha + \alpha z) = \frac{\pi}{1 - (1 - \pi)(1 - \alpha + \alpha z)} = \frac{\frac{\pi}{\pi + \alpha(1 - \pi)}}{1 - \frac{\alpha(1 - \pi)}{\pi + \alpha(1 - \pi)}z},$$

Alors $\alpha \circ X \sim Geom\left(\frac{\pi}{\pi + \alpha(1 - \pi)}\right)$.

En conséquence, les innovations (ε_t) doivent avoir une pgf donnée par :

$$pgf_{\varepsilon}(z) = \frac{pgf(z)}{pgf(1 - \alpha + \alpha z)} = \alpha + (1 - \alpha) \frac{\pi}{1 - (1 - \pi)z},$$

Cette pgf correspond à une distribution géométrique avec inflation de zéros, où la distribution parente est $Geom(\pi)$ et le paramètre d'inflation est α .

2.7.2 Le modèle Poisson composé INAR(1), noté CP-INAR(1)

Un processus INAR(1) $\{X_t, t \in \mathbf{N}\}$, est appelé processus CP-INAR(1) si les innovations $\{\varepsilon_t\}_{t \in \mathbf{N}}$ sont *i.i.d.* et suivent une loi de Poisson composé $CP_v(\lambda, H)$. De plus, les innovations $\varepsilon_s \sim CP_v(\lambda, H)$ ont une moyenne finie à condition que $H'(1) < \infty$. Étant donné qu'un processus CP-INAR(1) est également irréductible et apériodique (Schweer & Weib, 2014), on en déduit qu'un processus CP-INAR(1) avec $H'(1) < \infty$ possède une distribution marginale stationnaire unique.

Ainsi, la distribution des observations est effectivement surdispersée, mais présente également une inflation de zéros.

Une instance spéciale largement utilisée du modèle CP-INAR(1) est le modèle NB-INAR(1), dans lequel les innovations suivent une distribution binomiale négative.

2.7.3 Le modèle NB-INAR(1)

Dans le modèle NB-INAR(1), les innovations suivent une distribution binomiale négative, $\varepsilon_t \sim NB(n, \pi)$ avec un indice de dispersion $I_\varepsilon = \frac{1}{\pi}$.

La probabilité d'obtenir zéro est approximée par :

$$p_0 \approx \frac{\pi^{(M+1)^n}}{\prod_{k=1}^M (1 - (1 - \pi)(1 - \alpha_k))^n}.$$

Les probabilités de transition sont calculées à l'aide de :

$$P(\varepsilon_t = k) = \binom{n+k-1}{k} \cdot (1 - \pi)^k \cdot \pi^n, \quad \text{pour } k \in \mathbb{N}$$

2.8 ESTIMATION

Le problème d'estimation lié au processus INAR(1) est plus compliqué que celui du processus AR(1). La complication provient du fait que la distribution conditionnelle de X_t sachant X_{t-1} dans le processus INAR(1), est la convolution de la distribution de ε_t et celle d'une binomiale avec un paramètre d'échelle α et un paramètre d'indice X_{t-1} .

Tout au long de cette section, nous supposons que la séquence ε_t suit une distribution de Poisson avec un paramètre λ . Dans ce qui suit, nous présenterons plusieurs méthodes pour estimer les paramètres α et λ

Estimateur de Yule-Walker

La manière la plus simple d'obtenir un estimateur pour α est de remplacer $\gamma(k)$ par la fonction d'auto-covariance dans l'équation de Yule-Walker et la résoudre pour α afin d'obtenir

$$\hat{\alpha} = \frac{\sum_{t=0}^{n-1} (x_t - \bar{x})(x_{t+1} - \bar{x})}{\sum_{t=0}^n (x_t - \bar{x})^2} \quad (2.21)$$

où \bar{x} est la moyenne de l'échantillon.

Une estimation de λ peut être obtenue en calculant $\hat{\varepsilon}_t = x_t - \hat{\alpha}x_{t-1}$ pour $t = 1, 2, \dots, n$.

Étant donné que X_t suit une loi de Poisson de paramètre λ alors l'estimateur de λ est :

$$\hat{\lambda} = \sum_{t=1}^n \frac{\hat{\varepsilon}_t}{n} \quad (2.22)$$

Estimateur du moindre carré conditionnel

Contrairement au modèle AR(1) standard, on constate pour le processus INAR(1) que, étant donné X_{t-1} et ε_t , X_t est toujours une variable aléatoire.

La moyenne conditionnelle de X_t sachant X_{t-1} est donnée par

$$E(X_t | X_{t-1}) = \alpha X_{t-1} + \lambda = g(\theta, X_{t-1}) \quad (2.23)$$

où $\theta = (\alpha, \lambda')$ représente l'ensemble des paramètres à estimer. La procédure d'estimation que nous allons appliquer est basée sur la minimisation de la somme des écarts quadratiques par rapport à l'espérance conditionnelle. Ainsi, les estimateurs des moindres carrés conditionnels pour α et λ sont les valeurs qui minimisent :

$$Q_n(\theta) = \sum_{t=1}^n [(X_t - g(\theta, X_{t-1}))^2] \quad (2.24)$$

par rapport à θ . En évaluant les dérivées de (2.24) par rapport à α et λ , on obtient :

$$\hat{\alpha} = \frac{\sum x_t x_{t-1} - (\sum x_t \sum x_{t-1})/n}{\sum x_{t-1}^2 - (\sum x_{t-1})^2/n} \quad (2.25)$$

et

$$\hat{\lambda} = \frac{1}{n} (\sum x_t - \hat{\alpha} \sum x_{t-1}) \quad (2.26)$$

où toutes les sommes sont effectuées sur t , prenant des valeurs de 1 à n . Étant donné que l'on s'attend, pour les processus stationnaires, à ce que les estimations de la moyenne et de la variance calculées à partir des premières n observations soient très proches de celles calculées à partir de l'ensemble des $(n + 1)$ observations, on constate que les estimateurs CLE dans (2.25) et (2.26) sont très proches des estimateurs dans (2.21) et (2.22). Il est facile de vérifier que les fonctions g , $dg/d\alpha$, $dg/d\lambda$ et $d^2g/d\alpha d\lambda$ satisfont toutes les conditions de régularité du théorème de Klimko et Nelson (1978). Par conséquent, les estimateurs des moindres carrés conditionnels dans (2.25) et (2.26) sont fortement convergents.

De plus, si $E(\theta, \Sigma) < \infty$, ce qui est le cas pour une distribution de Poisson, alors selon le théorème de Klimko et Nelson, (θ, Σ) sont asymptotiquement distribués normalement lorsque :

$$n^{1/2}(\hat{\theta} - \theta^0) \sim MVN(0, V^{-1}WV^{-1})$$

où $\hat{\theta} = (\hat{\alpha}, \hat{\lambda})'$ et $\theta^0 = (\alpha^0, \lambda^0)'$ représente la valeur "vraie" des paramètres. V est une matrice 2×2 avec des éléments

$$V_{ij} = E\left(\frac{\delta g(\theta^0, X_{t-1})}{\delta \theta_i}, \frac{\delta g(\theta^0, X_{t-1})}{\delta \theta_j}\right) \quad i, j = 1, 2$$

et les éléments de W avec la forme :

$$W_{ij} = E(u_t^2(\theta^0) \frac{\delta g(\theta^0, X_{t-1})}{\delta \theta_i}, \frac{\delta g(\theta^0, X_{t-1})}{\delta \theta_j})$$

avec

$$u_t(\theta^0) = X_t - g(\theta^0, X_{t-1})$$

L'estimateur du maximum de vraisemblance

La fonction de vraisemblance d'un échantillon de $(n + 1)$ observations du processus INAR(1) peut être écrite comme suit :

$$L(x, \alpha, \lambda) = \left(\prod_{t=1}^n P_t(x_t) \right) \frac{[\lambda/(1-\alpha)]^{x_0}}{x_0!} \exp[-\lambda/(1-\alpha)] \quad (2.27)$$

où

$$P_t(x) = \exp(-\lambda) \sum_{i=0}^{\min(x_{t-1}, x_t)} \frac{\lambda^{x-i}}{(x-i)!} \binom{x_{t-1}}{i} \alpha_i (1-\alpha)^{x_{t-1}-i} \quad \text{pour } t = 1$$

et $x = (x_0, x_1, \dots, x_n)$

Nous allons d'abord considérer les estimateurs du maximum de vraisemblance conditionnelle de α et λ lorsque x_0 est donné. Dans ce cas, la fonction de vraisemblance conditionnelle est de la forme :

$$L(x, \alpha, \lambda x_0) = \prod_{t=1}^n P_t(x_t)$$

Sprott (1983) a proposé une méthode pour maximiser la fonction de vraisemblance conditionnelle $L(x, \alpha, \lambda/x_0)$ par rapport à α et λ , en supposant que $P(x)$ est indépendant de t , c'est-à-dire du paramètre d'index des observations. Cette méthode peut également être appliquée dans notre cas. En utilisant l'approche de Sprott, les dérivées de la fonction de log-vraisemblance conditionnelle par rapport à λ et α sont les suivantes :

$$S_\lambda = \sum_{t=1}^n \frac{\partial \log P_t(x_t)}{\partial \lambda} = \sum_{t=1}^n P_t(x_t - 1) / P_t(x_t) - n = \sum_{t=1}^n H(t) - n \quad (2.28)$$

et

$$S_\alpha = \sum_{t=1}^n \frac{\partial \log P_t(x_t)}{\partial \alpha} = \sum_{t=1}^n [(x_t - \alpha x_{t-1} - \lambda H(t)) / \alpha(1 - \alpha)] \quad (2.29)$$

Pour trouver l'estimateur de maximum de vraisemblance conditionnel (CMLE) on fixe $S_\lambda = 0$ et $S_\alpha = 0$ ce qui, d'après (2.29)

$$\sum x_t = \alpha \sum x_{t-1} + n\lambda \quad (2.30)$$

L'équation (2.30) peut être utilisée pour éliminer l'un des paramètres, disons α , puis S_λ dans (2.28) peut s'écrire en fonction de λ uniquement. Ainsi, les CMLE peuvent être trouvés en itérant sur S_λ . Il convient de noter que si l'on utilise (2.30) pour éliminer λ , on obtiendrait un estimateur de λ en fonction de α qui aurait la même forme que celui dans (2.26). Cependant, on ne peut rien dire analytiquement en comparant l'estimateur α dans (2.21) avec celui qui résulterait de la minimisation de (2.29). Maintenant, en considérant la fonction de vraisemblance inconditionnelle telle que donnée par (2.27), on peut voir que les dérivées de la fonction de log-vraisemblance par rapport à λ et α sont :

$$S_\lambda^0 = \frac{\partial \log L}{\partial \lambda} = S_\lambda + \frac{x_0}{\lambda} - \frac{1}{1 - \alpha} \quad (2.31)$$

Et

$$S_\alpha^0 = \frac{\partial \log L}{\partial \alpha} = S_\alpha + \frac{1}{1 - \alpha} \left(x_0 - \frac{\lambda}{1 - \alpha} \right) \quad (2.32)$$

Le MLE (Maximum Likelihood Estimator) peut être trouvé en fixant $S_\lambda^0 = 0$ et $S_\alpha^0 = 0$, à partir desquelles nous obtenons :

$$\lambda = \left(\sum_{t=0}^n x_t - \alpha \sum_{t=0}^n x_t \right) / \left(n + \frac{1 + \alpha}{1 - \alpha} \right) \quad (2.33)$$

L'équation (2.33) peut être utilisée pour éliminer λ dans (2.32), ce qui permet de trouver le MLE de α en itérant cette équation. Cependant, il est évident à partir des équations (2.31), (2.32) et (2.33) que les MLE ne sont pas faciles à manipuler comme dans le cas des CMLE.

La matrice de covariance asymptotique estimée pour les estimations des paramètres peut être calculée en prenant la deuxième dérivée de la fonction de vraisemblance. Cependant, comme l'a souligné Sprott, les estimations des paramètres dans une convolution sont généralement fortement corrélées, ce qui signifie que les inférences sur les paramètres sont fortement interdépendantes.

Dans la dérivation du MLE, nous avons supposé que (ϵ_t) suit une distribution de Poisson. Il convient de mentionner que Sprott (1983) a noté que cette procédure peut également être appliquée pour trouver le MLE des paramètres d'une convolution de deux membres de la famille de distribution de la série de puissances. Par conséquent, cette méthode peut être modifiée de manière similaire pour obtenir les estimations des paramètres du processus INAR(1) dans le cas où ϵ_t n'a pas nécessairement une distribution de Poisson.

2.9 PRÉDICTION DANS LE MODÈLE INAR(1)

Il existe deux approches principales pour faire de la prédiction dans les modèles INAR(1) :

- **La méthodologie classique** : Les méthodes classiques sont largement utilisées et relativement simples à mettre en œuvre. Elles reposent sur des calculs analytiques ou numériques pour estimer les paramètres du modèle INAR(1) à partir des données historiques. elles comprennent la méthode des moments, la méthode du maximum de vraisemblance et la méthode de moindres carrés. Une fois les paramètres estimés, les prédictions sont générées en utilisant les équations de récurrence du modèle INAR(1). Les méthodes classiques sont efficaces lorsque les hypothèses du modèle sont respectées et fournissent des résultats précis et fiables.
- **La méthodologie bayésienne** : C'est une approche assez différente de la méthode classique, ici on considère que les paramètres du modèle sont des variables aléatoires et on utilise les principes de l'inférence bayésienne. Elles nécessitent la spécification de distributions a priori pour les paramètres du modèle. En utilisant le théorème de Bayes, la distribution a posteriori des paramètres est calculée en combinant l'information a priori et les données observées. Les prédictions sont ensuite obtenues en échantillonnant à partir de la distribution a posteriori des paramètres. Les méthodes bayésiennes sont particulièrement utiles lorsque les données sont limitées ou lorsque des connaissances a priori sur les paramètres sont disponibles. Elles permettent également de quantifier l'incertitude associée aux prédictions et d'intégrer formellement des connaissances préalables.

Théorème 2.1 *La fonction génératrice des moments de X_{n+h} sachant X_n est :*

$$\rho_{X_{n+h}|X_n}(s) = [\alpha^h e^s + (1 - \alpha^h)]^{X_n} \exp \left[\lambda \frac{1 - \alpha^h}{1 - \alpha} (e^s - 1) \right] \quad (2.34)$$

L'expression (2.34) montre que la distribution de $X_{n+h}|X_n$ est une convolution d'une distribution binomiale de paramètres α^h et X_n , et d'une distribution de Poisson avec le paramètre $\lambda(1 - \alpha^h)/(1 - \alpha)$. Autrement dit, la fonction de

probabilité de $X_{n+h}|X_n$ est donnée par :

$$\begin{aligned} f(x_{n+h}) &= P(X_{n+h} = x_{n+h} | X_n = x_n) \\ &= \exp\left(-\lambda \frac{1-\alpha^h}{1-\alpha}\right) \sum_{i=0}^{M_h} \frac{1}{(x_{n+h}-i)!} \\ &\quad \times \left(\lambda \frac{1-\alpha^h}{1-\alpha}\right)^{x_{n+h}-i} \binom{x_n}{i} (\alpha^h)^i (1-\alpha^h)^{x_n-i}, \quad x_{n+h} = 0, 1, \dots, \end{aligned} \quad (2.35)$$

où $M_h = \min(X_{n+h}, X_n)$, par conséquent, nous avons le corollaire suivant :

Corollaire 1

2.9.1 Prédiction avec la méthode classique

Le prédicteur à h pas en avant basé sur l'espérance conditionnelle de l'INAR(1) :

$$\hat{X}_{n+h|X_n} = \mathbb{E}[X_{n+h}|X_n] = \alpha^h \left[X_n - \frac{\lambda}{1-\alpha} \right] + \frac{\lambda}{1-\alpha} \quad h = 1, 2, 3, \dots, \quad (2.36)$$

a été obtenu par Brännäs (1994) et Freeland et McCabe (2003), mais il produira difficilement des prévisions à valeurs entières. Afin d'obtenir des prédictions cohérentes pour X_{n+h} , Freeland et McCabe (2003) suggèrent d'utiliser la valeur qui minimise l'erreur absolue attendue donnée l'échantillon, c'est-à-dire la valeur qui minimise $\mathbb{E}[|X_{n+h} - \hat{X}_{n+h}| | X_n]$. Ainsi, ils ont conclu que $\hat{X}_{n+h} = \hat{m}_{n+h}$ est la médiane de la distribution conditionnelle à h pas en avant $f(x_{n+h}|x_n)$.

2.9.2 Prédiction avec la méthode bayésienne

La fonction de probabilité prédictive bayésienne est basée sur l'hypothèse selon laquelle l'observation future X_{n+h} et le vecteur des paramètres inconnus $\theta = (\alpha, \lambda)$ sont aléatoires. Comme nous le savons que l'information sur θ est donnée par l'échantillon observé x_n et quantifié dans le prédicteur a posteriori $\pi(\theta, x_n)$.

Définition 2.8 Soit $\theta \in \Theta$ le vecteur des paramètres inconnus. La distribution prédictive a posteriori bayésienne à h pas en avant est donnée par :

$$f(x_{n+h}|x_n) = \int_{\Theta} f(x_{n+h}; \theta | x_n) d\theta = \int_{\Theta} f(x_{n+h}|x_n; \theta) \pi(\theta | x_n) d\theta, \quad (2.37)$$

où $\pi(\theta | x_n)$ est la fonction a posteriori θ et $f(x_{n+h}|x_n; \theta)$ est la distribution prédictive (classique) donnée par (2.35).

La distribution prédictive à h pas en avant de $X_{n+h}|X_n$, donnée par l'expression (2.37).

Le prédicteur bayésien à h pas en avant peut être obtenu en utilisant la valeur attendue, la médiane ou le mode de X_{n+h} sachant X_n .

Les distributions Beta et Gamma sont utilisées comme lois a priori pour les paramètres du modèle INAR(1), $\alpha \sim \text{Beta}(a, b)$, $a, b > 0$ et $\lambda \sim \text{Gamma}(c, d)$. En considérant l'indépendance entre α et λ , la loi a priori de (α, λ) est donnée par :

$$P(\alpha, \lambda) \propto \lambda^{c-1} e^{-d\lambda} \alpha^{a-1} (1-\alpha)^{b-1}, \quad \lambda > 0, 0 < \alpha < 1, \quad (2.38)$$

Les paramètres a, b, c et d sont des paramètres connus. Remarquez que lorsque $a \rightarrow 0, b \rightarrow 0, c \rightarrow 0$ et $d \rightarrow 0$, nous avons une loi a priori vague.

La loi a posteriori de (α, λ) peut être écrite comme :

$$\begin{aligned} P(\alpha, \lambda | x_n) &\propto L(x_n, \alpha, \lambda | x_1) P(\lambda, \alpha) \\ &= \exp[-d + (n-1)\lambda] \lambda^{c-1} \alpha^{a-1} (1-\alpha)^{b-1} \\ &\quad \times \prod_{t=2}^n \sum_{i=0}^{M_t} \frac{\lambda^{x_t-i}}{(x_t-i)!} \alpha^i (1-\alpha)^{x_{t-1}-i} \end{aligned} \quad (2.39)$$

En conséquence, pour le modèle PoINAR(1), la fonction prédictive bayésienne de X_{n+h} étant donné X_n est donnée par :

$$\begin{aligned} f(x_{n+h} | x_n) &\propto \int_{\alpha} \int_{\lambda} \left(\sum_{i=0}^{M_h} \binom{x_n}{i} (\alpha^h)^i (1-\alpha)^{x_n-i} \right) \frac{1}{(x_{n+h}-i)!} \\ &\quad \times \exp\left(-\lambda \frac{1-\alpha^h}{1-\alpha}\right) \left(\lambda \frac{1-\alpha^h}{1-\alpha}\right)^{x_{n+h}-i} \exp[-(d+n)\lambda] \lambda^{c-1} \\ &\quad \times \alpha^{a-1} (1-\alpha)^{b-1} \prod_{t=2}^n \sum_{i=0}^{M_t} \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} d\alpha d\lambda \end{aligned} \quad (2.40)$$

La complexité de $f(x_{n+h} | x_n)$ ne nous permet pas de travailler directement avec cette fonction. Afin d'estimer x_{n+h} , nous pouvons utiliser l'algorithme suivant :

1. À partir de l'échantillon (X_1, X_2, \dots, X_n) , calculer (via la méthode classique) une estimation initiale pour α , notée α_0 .
2. Utilisant l'échantillonnage adaptatif de rejet dans la méthodologie Gibbs (voir Gilks et al., 1995), calculez à partir des distributions conditionnelles complètes des paramètres α et λ , un échantillon $(\alpha_1, \lambda_1), (\alpha_2, \lambda_2), \dots, (\alpha_m, \lambda_m)$.
3. Pour chaque i ($i = 1, \dots, m$), échantillonnez $X_{n+h,i}$ à partir de $f(x_{n+h} | x_n, \alpha_i, \lambda_i)$ en utilisant la méthode de transformation inverse adaptée aux variables entières, c'est-à-dire :
 - Échantillonnez u à partir d'une uniforme $U(0, 1)$.
 - Calculez le plus petit entier s tel que $\sum_{i=0}^s f(x_{n+h} | x_n, \alpha_i, \lambda_i) \geq u$.
 - Considérez $X_{n+h,i} = s$.

Après avoir échantillonné $X_{n+h,1}, X_{n+h,2}, \dots, X_{n+h,m}$ le prédicteur X_{n+h} peut être calculé à partir de la moyenne de l'échantillon \hat{X}_{n+h} , de la médiane \hat{m}_{n+h} ou du mode $\hat{m}o_{n+h}$. Pour plus de méthodes se référer à Silva et Pereira (2009). Ainsi le prédicteur peut être écrite ainsi :

$$\tilde{X}_{n+h} = X_n \left(\frac{1}{m} \sum_{i=1}^m \alpha_i^h \right)! + \left(\frac{1}{m} \sum_{i=1}^m \frac{1-\alpha_i^h}{1-\alpha^i} \lambda^i \right)! \quad (2.41)$$

2.9.3 Intervalle de prédiction

Dans cette partie, nous allons calculer les intervalles de prédiction :

Méthode Classique

Un intervalle de confiance pour le prédicteur \hat{X}_{n+h} peut être calculé en utilisant la fonction de probabilité de l'erreur de prédiction à h pas en avant, donnée par :

$$X_{n+h}|X_n = X_{n+h}|X_n - \hat{X}_{n+h}|X_n \quad (2.42)$$

En remplaçant $\hat{X}_{n+h}|X_n$ donné par (2.36), nous obtenons :

$$\hat{e}_{n+h}|X_n = X_{n+h} - \alpha^h x_n - \lambda \frac{1 - \alpha^h}{1 - \alpha} \quad (2.43)$$

Puisque $\hat{e}_{n+h}|X_n$ est une fonction de la variable aléatoire discrète X_{n+h} , nous avons :

$$\hat{e}_{n+h}|X_n = k - \alpha^h x_n - \lambda \frac{1 - \alpha^h}{1 - \alpha} \quad \text{ou } k = 0, 1, 2, \dots \quad (2.44)$$

Ainsi,

$$\begin{aligned} P(e_{n+h} = k - \alpha^h x_n - \lambda \frac{1 - \alpha^h}{1 - \alpha} | X_n) &= P(X_{n+h} = k | X_n = x_n) \\ &= \exp -\lambda \frac{1 - \alpha^h}{1 - \alpha} \sum_{i=0}^{M_h} \frac{1}{(k-i)!} \\ &\times \left(\lambda \frac{1 - \alpha^h}{1 - \alpha} \right)^{k-i} \binom{x_n}{i} (\alpha^h)^i (1 - \alpha^h)^{x_n - i}. \end{aligned} \quad (2.45)$$

À partir de l'expression (2.45), nous obtenons un intervalle de confiance de niveau γ pour X_{n+h} :

$$(\hat{X}_{n+h} + e_{t_1}, \hat{X}_{n+h} + e_{t_2}) \quad (2.46)$$

où \hat{X}_{n+h} est donné par (2.36), e_{t_1} est la plus grande valeur de e_{n+h} telle que $P(e_{n+h} \leq e_{t_1}) \leq (1 - \gamma)/2$, et e_{t_2} est la plus petite valeur de e_{n+h} telle que $P(e_{n+h} \leq e_{t_2}) \geq (1 + \gamma)/2$.

Méthode Bayésienne

Nous proposons une généralisation adaptative de la méthode utilisée pour obtenir le plus haut intervalles de densité postérieure (HPD) des paramètres du modèle, dans lesquels nous considérons la distribution prédictive au lieu de la loi a posteriori.

Définition 2.9

$$P(X_L \leq X_{n+h} \leq X_R) = \sum_{X_{n+h}=X_L}^{X_R} f(x_{n+h}|x_n) \quad (2.47)$$

Cependant, étant donné que $f(x_{n+h}|x_n)$ n'est pas toujours symétrique, les intervalles avec une probabilité prédictive postérieure maximale sont plus souhaitables que les intervalles prédictifs.

Définition 2.10 $R(\gamma) = (X_L, X_R)$ est un intervalle HPD à 100 γ % pour X_{n+h} si

$$P(X_L \leq X_{n+h} \leq X_R) = \sum_{X_{n+h}=X_L}^{X_R} f(x_{n+h}|x_n) \geq K_\gamma, \quad (2.48)$$

où K_γ est la plus grande constante telle que $P(X_{n+h} \in R(\gamma)) \geq \gamma$.

En raison de la complexité de la fonction de probabilité prédictive (2.40), il n'est pas possible de calculer l'intervalle HPD exact pour X_{n+h} . Nous pouvons fournir une approximation pour $R(\gamma)$ en utilisant l'algorithme de Chen et Shao (1999) :

1. Générer un échantillon de taille M à partir de la loi a posteriori de α et λ à l'aide d'une méthode MCMC, telle que l'algorithme de Metropolis-Hastings.
2. Pour chaque échantillon i ($i = 1, \dots, M$), générer un échantillon $X_{n+h,i}$ à partir de la distribution prédictive $f(x_{n+h}|x_n, \alpha_i, \lambda_i)$ en utilisant une méthode appropriée, par exemple la méthode de rejet adaptatif.
3. Calculer les quantiles $\alpha/2$ et $1 - \alpha/2$ de l'échantillon $X_{n+h,1}, \dots, X_{n+h,M}$ pour obtenir les limites inférieure et supérieure de l'intervalle approximatif $R(\gamma)$.

En utilisant l'algorithme de Chen et Shao, nous pouvons obtenir une approximation de l'intervalle HPD pour X_{n+h} sans avoir besoin de la forme exacte de la fonction de probabilité prédictive $f(x_{n+h}|x_n)$. Cela permet d'estimer l'intervalle avec une précision raisonnable malgré la complexité de la fonction.

2.10 EXEMPLE DE DONNÉES RÉELLES

Nous appliquons la méthodologie proposée à un ensemble de données analysées par Freeland (1998), comprenant 120 comptages mensuels de travailleurs bénéficiant d'indemnités de perte de salaire pour des blessures par brûlure. Tous les détails descriptifs de l'ensemble de données se trouvent dans l'étude de Freeland (1998), qui conclut que le modèle PoINAR(1) est un choix plausible pour modéliser les données. Afin d'évaluer et de comparer les différentes méthodes de prévision, des prévisions à h pas en avant ($h = 1, 2, 3, 4, 5, 6$) sont produites pour la période de juillet à décembre 1994, pour laquelle nous connaissons les valeurs observées. Les prévisions ponctuelles basées sur la moyenne, la médiane, le mode et les valeurs observées sont présentées dans le Tableau suivant. En général, on peut noter que les valeurs MAPE des prédictions ponctuelles classiques sont plus petites que celles des prédictions bayésiennes. Ce résultat est attendu étant donné les résultats de simulation présentés dans la dernière section, car la valeur estimée pour α est de 0,4. Des prédictions d'intervalle pour la période de juillet à décembre 1994 sont obtenues en utilisant les deux approches proposées données par (3.1). Les intervalles obtenus, présentés dans ci-dessous, sont analogues, bien que les prédictions bayésiennes aient une largeur plus petite.

h	year/ month	claims of	point prediction						interval prediction	
			classical			Bayesian			classical	Bayesian
			\hat{x}	\hat{m}	\hat{m}_0	\hat{x}	\hat{m}	\hat{m}_0		
1	94/07	11	7.89	8	7	7.67	7	7	(2.1,13.0)	(3.0,13.1)
2	94/08	12	8.24	8	8	8.01	8	7	(2.0,14.0)	(3.0,14.0)
3	94/09	11	8.38	8	8	8.14	8	7	(2.0,14.0)	(3.0,14.0)
4	94/10	12	8.44	8	8	8.21	8	7	(2.0,15.0)	(3.0,14.0)
5	94/11	7	8.46	8	8	8.22	8	7	(2.0,15.0)	(3.0,14.0)
6	94/12	11	8.47	8	8	8.22	8	7	(2.0,15.0)	(3.0,14.0)
MAPE			0.26	0.27	0.27	0.27	0.27	0.32		

FIGURE 2.5 – Prévisions à h pas en avant des demandes mensuelles de juillet à décembre 1994.

INAR(p) ET ESTIMATION

3

3.1 INTRODUCTION

l'intervalle de confiance de niveau γ pour X_{n+h} :

$$(\hat{X}_{n+h} + e_{t_1}, \hat{X}_{n+h} + e_{t_2}) \quad (3.1)$$

Dans une perspective de généralisation, nous introduirons l'opérateur d'amin-cissement généralisé \star et on présentera le modèle qui y est directement associé le modèle GINAR(p).

3.2 MODÈLE INAR(p)

Le modèle INAR(p) également connu sous le nom de Integer-valued pth-order Autoregressive process est une extension du modèle INAR(1) étudié par Al-Osh et Al-Zaid (1987).

Définition 3.1 Soit $\{X_t\}_{t \in \mathbb{Z}}$ une suite de variables aléatoires à valeurs entières positives; $\{\epsilon_t\}_{t \in \mathbb{Z}}$, une suite de variables aléatoires i.i.d. à valeurs entières positives, avec $E(\epsilon_t) = \mu_\epsilon > 0$ et $\text{var}(\epsilon_t) = \sigma^2 > 0$, $p \in \mathbb{N}^*$ et $\{\alpha_j\}_{j \in 1 \dots p}$ une suite de constantes telles que $0 \leq \alpha_j < 1$, $\forall j$ et $(\alpha_p > 0)$. Alors $\{X_t, t \in \mathbb{Z}\}$ est un processus INAR(p) s'il est de la forme :

$$X_t = \sum_{j=1}^p \alpha_j \circ X_{t-j} + \epsilon_t, \quad \forall t \in \mathbb{Z} \quad (3.2)$$

Avec :

$$\alpha_j \circ X_{t-j} = \sum_{k=0}^{X_{t-j}} Y_{jk}$$

La série de comptage (Y_{jk}) est une suite de variables aléatoires i.i.d. qui suit une loi de Bernoulli de paramètre α_j , et indépendante de X_{t-j} , de plus (ϵ_t) est indépendante de (Y_{jk})

3.2.1 Propriétés du modèle

Les caractéristiques ou propriétés du modèle INAR(p) comprennent l'espérance conditionnelle et la variance de X_t , où ces propriétés sont utilisées pour prouver l'espérance et la variance inconditionnelles. Ensuite, les fonctions d'autocovariance et d'autocorrélation seront également présentées.

Espérance conditionnelle de X_t : $E[X_t | X_{t-1}, \dots, X_{t-p}] = \mu_\epsilon + \sum_{k=1}^p \alpha_k X_{t-k}$.
 Variance conditionnelle de X_t : $\text{Var}[X_t | X_{t-1}, \dots, X_{t-p}] = \sigma_\epsilon^2 + \sum_{k=1}^p \alpha_k (1 -$

$\alpha_k)X_{t-k}$.

Autocovariance de X_t : $\gamma_k = \alpha_1\gamma_{k-1} + \alpha_2\gamma_{k-2} + \dots + \alpha_p\gamma_{k-p} = \sum_{i=1}^p \alpha_i\gamma_{k-i}$.

Autocorrélation de X_t : $\rho_k = \alpha_1\rho_{k-1} + \alpha_2\rho_{k-2} + \dots + \alpha_p\rho_{k-p} = \sum_{i=1}^p \alpha_i\rho_{k-i}$.

Pour obtenir les espérances et la variance inconditionnelles du modèle INAR(p), il est nécessaire de faire une hypothèse de stationnarité. L'hypothèse de stationnarité dans le modèle INAR(p) suit l'hypothèse de stationnarité dans le modèle AR(p), où la série temporelle est stationnaire si $\alpha_1 + \alpha_2 + \dots + \alpha_p < 1$ et $|\alpha_p| < 1$. Ainsi, les espérances et la variance inconditionnelles du modèle INAR(p) peuvent être exprimées comme suit :

$$\text{Espérance inconditionnelle : } E[X_t] = \frac{\mu_\epsilon}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_p},$$

$$\text{Variance inconditionnelle : } \text{Var}[X_t] = \frac{\sigma_\epsilon^2}{1 - \alpha_1 - \alpha_2 - \dots - \alpha_p}.$$

3.2.2 Propriétés de corrélation

En étudiant les propriétés de corrélation du processus INAR(p), On utilise l'hypothèse selon laquelle, étant donné $X_n = x_n$, la variable aléatoire $\alpha_i x_n$ ne depend pas des valeurs passées du processus (X_{n-k}) ainsi que de $\alpha_j \circ X_{n-k}$ pour $i, j = 1, 2, \dots, p$ et $k > 1$. À présent, il est possible de constater à partir de (3.2) que

$$E(X_n) = \sum_{i=1}^p \alpha_i E(X_{n-i}) + \mu_\epsilon \quad (3.3)$$

où $\mu_\epsilon = E(\epsilon_n)$. Sous l'hypothèse de stationnarité faible, nous avons :

$$\mu_x = E(X_n) = \frac{\mu_\epsilon}{1 - \sum_{i=1}^p \alpha_i}$$

Pour calculer l'auto-covariance en, $\gamma(k)$, on a :

$$\gamma(k) = \text{cov}(X_{n-k}, X_n) = \sum_{i=1}^p \text{cov}(X_{n-k}, \alpha_i \circ X_{n-i}) + \delta_k(0)\sigma_\epsilon^2 \quad (3.4)$$

$$= \sum_{i=1}^p \gamma(k-i, \alpha_i) + \delta_k(0)\sigma_\epsilon^2$$

où $\gamma(l, \alpha_i) = \text{Cov}(X_{n-l}, \alpha_i \circ X_n)$ et $\gamma_k(0) = 1$ si $k = 0$ et 0 sinon.
à partir de la définition de $\gamma(l, \alpha_i)$

$$\gamma(l, \alpha) = \alpha\gamma(l), \quad l \geq 0 \quad (3.5)$$

En utilisant un raisonnement conditionnel et en supposant que, étant donné $X = x$, le vecteur $(\alpha_1 \circ x, \alpha_2 \circ x, \dots, \alpha_p \circ x)$ suit une distribution multinomiale avec des paramètres $(\alpha_1, \alpha_2, \dots, \alpha_p, x)$, nous obtenons :

$$\text{Cov}(\alpha_j \circ X_{n-k}, \alpha_i \circ X_{n-l}) = \begin{cases} \alpha_j\gamma(k-l, \alpha_i) & \text{pour } k < l \\ \alpha_j\alpha_i\gamma(0) + \alpha_j(\delta_j(i) - \alpha_i)\mu_x & \text{pour } k = l \\ \alpha_i\gamma(l-k, \alpha_j) \text{ pour} & \text{pour } k > l \end{cases} \quad (3.6)$$

$\gamma(-l, \alpha_i)$ peut être déterminé de manière récursive en utilisant (3.6).
On commence avec $l = 1$, nous avons :

$$\begin{aligned}\gamma(-l, \alpha_i) &= \alpha_1 \alpha_i \gamma(0) + \alpha_1 (\delta_i(1) - \alpha_i) \mu_x + \alpha_i \sum_{j=2}^p \gamma(\alpha_j, j-1) \\ &= \alpha_i \gamma(-1) + \alpha_1 (\delta_i(1) - \alpha_i) \mu_x\end{aligned}\quad (3.7)$$

Soit $\mu(l, \alpha_i) \equiv \gamma(l, \alpha_i) - \alpha_i \gamma(l)$, l'équation (3.5) donne $\mu(l, \alpha_i) = 0$ pour $l \geq 0$ et d'après (3.7), on a $\mu(-1, \alpha_i) = \alpha_1 (\delta_i(1) - \alpha_i) \mu_x$

De même :

$$\begin{aligned}\mu(-2, \alpha_i) &= \sum_{j=1}^p \text{cov}(\alpha_j \circ X_{n-j}, \alpha_i \circ X_{n-2}) - \alpha_i \gamma(-2) \\ &= \alpha_1 \gamma(-1, \alpha_i) + \alpha_2 \alpha_i \gamma(0) + \alpha_2 (\delta_i(2) - \alpha_i) \mu_x \\ &\quad + \alpha_i \sum_{j=3}^p \text{cov}(\alpha_j \circ X_{n-j}, X_{n-2}) - \alpha_i \gamma(-2) \\ &= \alpha_1 \mu(-1, \alpha_i) + \alpha_2 (\delta_i(2) - \alpha_i) \mu_x\end{aligned}$$

En utilisant cette méthode itérative, nous obtenons :

$$\mu(-l, \alpha_i) = \sum_{j=1}^{l-1} \alpha_j \mu(j-l, \alpha_i) + \mu_{i,l} \quad (3.8)$$

ou $\mu_{i,l} = \alpha_i (\delta_i(l) - \alpha_l) \mu_x$. ça signifie que $\mu(-l, \alpha_i)$ représente la fonction linéaire de la moyenne du processus μ_x .

En utilisant la définition de $\mu(l, \alpha_i)$, on peut voir que $\gamma(k-i, \alpha_i)$ dans (3.4) satisfait :

$$\gamma(k-i, \alpha_i) = \alpha_i \gamma(k-i) + \mu(k-1, \alpha_i) \quad (3.9)$$

ou $\mu(k-1, \alpha_i)$ est déterminé par (3.8) pour $k < i$ et $\mu(k-1, \alpha_i) = 0$ pour $k \geq i$.
En utilisant (3.9) on a :

$$\gamma(k) = \sum_{i=1}^p \alpha_i \gamma(k-1) + \sum_{i=k+1}^p \mu(k-i, \alpha_i) + \delta_k(0) \delta_\epsilon^2 \quad (3.10)$$

Il est clair à partir de l'équation (3.10) que l'autocovariance du processus INAR(p) a la même forme que celle du processus ARMA(p, p-1) gaussien. Ce comportement de $\gamma(k)$ est dû à la structure de dépendance mutuelle entre les composantes de X_n , c'est-à-dire $a_i \circ X_{n-i}$ pour $i = 1, 2, \dots, p$ apparaissant à des moments différents comme discuté dans la section précédente.

Remarque 3.1 *La sélection des séries de comptage joue un rôle crucial dans la détermination de la structure de corrélation d'un processus. Il est primordial de faire un choix judicieux des séries de comptage afin de capturer au mieux les dépendances et les relations entre les observations.*

Ainsi le modèle INAR(p) comporte deux structures spécifique distinctes :

- *L'une proposé par Al-Osh et Al-Zaid (1990) : ce dernier étant une extension directe du modèle INAR(1), ici les auteurs estiment que la distribution conditionnelle du vecteur aléatoire $(\alpha_1 \circ X_t, \alpha_2 X_t, \dots, \alpha_p X_t)$ sachant $X_t = x_t$ est multinomiale de paramètre $(\alpha_1, \dots, \alpha_p, X_t)$ et indépendante de ses valeurs passées c'est à dire X_{t-k} et de tout les amincissement de celui-ci $\forall k > 0$. Dans cette approche, le modèle INAR(p) dispose de la même structure de covariance de le modèle ARMA(p,p-1).*

- et une alternative proposé par Du & Li (1991) où les auteurs supposent que les séries de comptage sont mutuellement indépendantes.
 Cette structure du second ordre est similaire à celle du modèle AR(p).

Dans la suite de ce travail, c'est l'approche de Du & Li qui sera adoptée.

3.3 MODÈLE INAR(P) DE DU ET LI

Du et Li (1991) définissent le modèle INAR(p) comme suit :

$$X_t = \alpha_1 \circ X_{t-1} + \alpha_2 \circ X_{t-2} + \dots + \alpha_p \circ X_{t-p} + \epsilon_t \quad (3.11)$$

où

1. $\{\epsilon_t\}$ est une séquence de variable aléatoires entières non négative i.i.d. avec $E[\epsilon_t] = \mu_\epsilon$, $Var[\epsilon_t] = \sigma^2$, $E[\epsilon_t^3] = \gamma_\epsilon$, et $E[\epsilon_t^4] = \kappa_\epsilon$.
2. Toutes les séries de comptage $\{Y_{i,j}\}$ des opérations d'amincissement :

$$\alpha_i \circ X_{t-i} = \sum_{j=0}^{X_{t-i}} Y_{j,i} \quad i = 1, \dots, p \quad (3.12)$$

Sont mutuellement indépendants, et indépendants de ϵ_t , tels que $E[Y_{j,i}] = \alpha_i$, $Var[Y_{j,i}] = \sigma^2$, $E[Y_{j,i}^3] = \gamma_i$ et $E[Y_{j,i}^4] = \kappa_i$

Une représentation alternative du processus INAR(p) en tant que processus INAR(1) p-dimensionnel a été obtenue par Franke et Subba Rao (1995). De ce fait, en utilisant l'opération d'amincissement vectoriel donnée dans (2.1), le processus INAR(p) peut s'écrire :

$$X_t = A \circ X_{t-1} + W_t \quad (3.13)$$

$$X_t = HX_t \quad (3.14)$$

où :

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{p-1} & \alpha_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \quad (3.15)$$

Pour $\alpha_i \geq 0 \quad i = 1, \dots, p-1$ et $\alpha_p \geq 0$ $H = [1 \ 0 \ \dots \ 0]$, $X_t = [X_t \ X_{t-1} \ \dots \ X_{t-p+1}]^T$ et $W_t = [\epsilon_t \ 0 \ \dots \ 0]^T$, pour $\{\epsilon_t\}$ une séquence de variables aléatoires i.i.d. avec $E[\epsilon_t] = \mu_\epsilon$ $Var[\epsilon_t] = \sigma^2$, $E[\epsilon_t^3] = \gamma_\epsilon$, et $E[\epsilon_t^4] = \kappa_\epsilon < \infty$.

De plus, le (3.13) peut être exprimé comme suit :

$$\begin{aligned} X_t &= A \circ X_{t-1} + W_t \\ &= A \circ (A \circ X_{t-2} + W_{t-1}) + W_t \\ &\stackrel{d}{=} A^2 \circ X_{t-2} + A \circ W_{t-1} + W_t \\ &\vdots \\ &\stackrel{d}{=} A^k \circ X_{t-k} + \sum_{j=0}^{k-1} A^j \circ W_{t-j}, \end{aligned}$$

où $\stackrel{d}{=}$ représente une distribution égale.
Alors les équations (3.13) et (3.14) peuvent s'écrire :

$$X_t \stackrel{d}{=} \sum_{j=0}^{k-1} A^j \circ W_{t-j} \quad (3.16)$$

$$X_t \stackrel{d}{=} HX_t \quad (3.17)$$

3.4 STATIONNARITÉ D'UN MODÈLE INAR(p)

Un modèle INAR(p) est stationnaire si il respecte les conditions suivantes :
— Les coefficients d'amincissement α_k doivent respecter la condition :

$$0 < \alpha_k < 1$$

— La somme des coefficients doit être inférieure à 1 :

$$\sum_{k=1}^p \alpha_k < 1$$

De plus, on peut dire que la condition de stationnarité pour le processus INAR(p) est que le rayon spectral de la matrice A définie en (3.15) est inférieur à 1, c'est à dire $\rho(A) < 1$.

Propriété

Soit $R(\cdot)$ la fonction d'auto-covariance du processus INAR(p) $\{X_i\}$. Alors $R(\cdot)$ satisfait un ensemble d'équations de différence de type Yule-Walker, qui peuvent être écrites, respectivement, sous forme scalaire et vectorielle, comme suit :

$$\begin{cases} R(0) = V_p + \sum_{i=1}^p \alpha_i R(i) \\ R(k) = \sum_{i=1}^p \alpha_i R(i-k) \end{cases} \Leftrightarrow \mathbf{R}_p \boldsymbol{\alpha} = \begin{bmatrix} -V_p \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3.18)$$

où

$$\begin{aligned} R_p &= [R(j, k)] = [R(p+1-k, p+1-j)] = [R(j-k)] \\ &= \begin{pmatrix} R(0) & R(1) & \cdots & R(p) \\ R(1) & R(0) & \cdots & R(p-1) \\ \vdots & \vdots & \ddots & \vdots \\ R(p) & R(p-1) & \cdots & R(0) \end{pmatrix} \end{aligned} \quad (3.19)$$

$$\boldsymbol{\alpha} = [-1 \quad \alpha_1 \quad \cdots \quad \alpha_p]^T \quad (3.20)$$

3.5 ESTIMATION D'UN INAR(p)

Pour estimer les paramètres d'un modèle INAR(p) plusieurs approches peuvent être utilisées notamment :

3.5.1 Estimation de Yule Walker

L'estimation de Yule Walker a été la première méthode proposée par (Al-Osh et Alzaid, 1987, Du and Li, 1991) pour estimer les paramètres INAR, , une méthode des moments, qui consiste à remplacer la fonction d'auto-covariance théorique par la fonction d'auto-covariance empirique dans les équations de type Yule Walker. Pour $k = 1, 2, \dots, p$ de l'équation d'autocorrélation $\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2} + \dots + \alpha_p \rho_{k-p}$, l'équation de Yule-Walker est définie comme suit :

$$\Gamma \alpha = \rho,$$

où $\Gamma = [\rho_{|i-j|}]_{p \times p}$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_p)^\top$, et $\rho = (\rho_1, \rho_2, \dots, \rho_p)^\top$.

En remplaçant l'autocorrélation théorique ρ_k par l'autocorrélation échantillonnale r_k , on obtient l'estimation des paramètres de Yule-Walker $\hat{\alpha}$ pour α :

$$\hat{\Gamma} \hat{\alpha} = \hat{\rho}. \quad (3)$$

Pour obtenir l'estimation des paramètres, en plus de résoudre l'équation matricielle ci-dessus, on peut également résoudre l'équation linéaire de Yule-Walker :

$$\rho_1 = \varphi_1 + \varphi_2 \rho_1 + \dots + \varphi_p \rho_{p-1},$$

$$\rho_2 = \varphi_1 \rho_1 + \varphi_2 + \dots + \varphi_p \rho_{p-2},$$

$$\vdots$$

$$\rho_p = \varphi_1 \rho_{p-1} + \varphi_2 \rho_{p-2} + \dots + \varphi_p.$$

Après avoir estimé les paramètres α , l'espérance inconditionnelle X_t peut être estimée afin d'obtenir une estimation de l'erreur moyenne du modèle, à savoir :

$$\mu_\varepsilon = \mu_X \left(1 - \sum_{k=1}^p \hat{\alpha}_k \right) = \mu_X (1 - \hat{\alpha}_1 - \dots - \hat{\alpha}_p).$$

Veuillez noter que φ est utilisé ici pour représenter les coefficients autorégressifs estimés, et μ_X représente la moyenne de la série temporelle.

3.5.2 Estimation par la méthode des moindres carrés

Soit

$$F_t = \mathcal{F}(X_1, X_2, \dots, X_t)$$

$$B = (\mu, \alpha_1, \dots, \alpha_p)^\top \hat{=} (\beta_1, \beta_2, \dots, \beta_{p+1})^\top$$

$$g(\beta, F_t) = E(X_t | F_{t-1})$$

$$= \mu + \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} \quad (t > p) \quad (3.21)$$

$$Q_N(\beta) = \sum_{t=p+1}^N \{X_t - g(\beta, F_t)\}^2 \quad (3.22)$$

On sélectionne la valeur de p qui minimise $Q_N(\beta)$ comme l'estimateur de B , c'est-à-dire :

$$Q_N(\hat{\beta}) = \min Q_N(\beta) \quad (3.23)$$

$\hat{\beta}$ est appelé l'estimateur des moindres carrés conditionnels (CLS) de β (Klimko et Nelson, 1978). $\hat{\beta}$ peut être résolu à partir de :

$$\frac{\partial Q_N}{\partial \mu} = 0 \quad (3.24)$$

$$\frac{\partial Q_N}{\partial \alpha_l} = 0 \quad (l = 1, 2, \dots, p) \quad (3.25)$$

ce qui donne :

$$\hat{\mu}^* = \bar{X}_N^{(0)} - \sum_{j=1}^p \hat{\alpha}_j^* \bar{X}_N^{(j)} \quad (3.26)$$

$$\hat{\Gamma}^* \hat{\alpha}^* = \hat{\rho}^* \quad (3.27)$$

avec : $\hat{\beta} = (\hat{\mu}^*, \hat{\alpha}^{t*})^t$ ou :

$$\bar{X}_N^{(j)} = \frac{1}{N-p} \sum_{t=p+1}^N X_{t-j}$$

$$\hat{\gamma}_{k-j}^* = \frac{1}{N-p} \sum_{t=p+1}^N (X_{t-j} - \bar{X}_N^{(j)})(X_{t-1} - \bar{X}_N^{(k)})$$

$$\hat{\rho}_{k-j}^* = \frac{\hat{\gamma}_{k-j}^*}{\hat{\gamma}_0^*}$$

$$\hat{\Gamma}^* = [\hat{\rho}_{|i-j|}^*]_{p \times p}$$

$$\hat{\rho} = [\hat{\rho}_1^* \hat{\rho}_2^* \dots \hat{\rho}_p^*]^t$$

On peut voir à partir de (3.26) et (3.27) que lorsque N est suffisamment grand $\hat{\Gamma}^* - \hat{\Gamma}$, $\hat{\rho}^* - \hat{\rho}$ et $\bar{X}_N^{(j)} - \bar{X}$ tendent vers 0, de sorte que $\hat{\alpha}^*$ et $\hat{\alpha}$ sont très proches, de même pour $\hat{\mu}^*$ et $\hat{\mu}$.

On définit :

$$g(t) = g(\beta, F_t) \quad (t \geq (p+1))$$

$$U_p(\beta) = X_{p+1} - \alpha_1 X_p - \dots - \alpha_p X_1 - \mu \quad (3.28)$$

Nous savons de Klimko et Nelson (1978) que l'estimation CLS est asymptotiquement normale avec la variance asymptotique $V^{-1} W V^{-1}$, où

$$W = \left[E U_p^2(\beta) \frac{\partial g_{p+1}}{\partial \beta_i} \frac{\partial g_{p+1}}{\partial \beta_j} \right]_{(p+1) \times (p+1)} \quad (3.29)$$

$$V = \left[E \frac{\partial g_{p+1}}{\partial \beta_i} \frac{\partial g_{p+1}}{\partial \beta_j} \right]_{(p+1) \times (p+1)} \quad (3.30)$$

Pour $i \geq 2, j \geq 2$

$$V_{ij} = E \frac{\partial g_{p+1}}{\partial \beta_i} \frac{\partial g_{p+1}}{\partial \beta_j} = E X_{t-i} E X_{t-j} = \gamma_{t-j} + \mu_x^2$$

où $\mu_x = E X_t, V_{1,j} = E X_{t-j} = \mu_x$ de sorte que :

$$V = \begin{pmatrix} 1 & \mu_1 \tau \\ \mu_x 1 & G + \mu_x 1 \tau \end{pmatrix} \quad (3.31)$$

Où : $\mathbf{1} = (1, 1, \dots, 1)$ et $G = [\gamma_{i-j}\tau]_{p \times p}$ de (3.31), on a :

$$V^{-1} = \begin{pmatrix} 1 + \mu_x^2 \mathbf{1} \tau G^{-1} \mathbf{1} & -\mu_x \mathbf{1} \tau G^{-1} \\ -\mu_x G^{-1} \mathbf{1} & G^{-1} \end{pmatrix} \quad (3.32)$$

Pour $2 \leq p \leq p+1, 2 \leq j \leq p+1$

$$W_{ij} \hat{=} EU_p^2(\beta) \frac{\partial g_{p+1}}{\partial \beta_i} \frac{\partial g_{p+1}}{\partial \beta_j}$$

Ce qui en le conditionnant ce réduit à :

$$\begin{aligned} W_{ij} &= \sum_{k=1}^p \alpha_k (1 - \alpha_k) EX_{p-k+1} X_{p+2-i} X_{p+2-j} + \sigma^2 EX_{p+2-i} X_{p+2-j} \\ &= \sum_{k=1}^p \alpha_k (1 - \alpha_k) EX_{p-k+1} X_{p+2-i} X_{p+2-j} + \sigma^2 V_{ij} \end{aligned} \quad (3.33)$$

De manière similaire, nous pouvons trouver :

$$\begin{aligned} W_{1j} &\hat{=} EU_p^2(\beta) \frac{\partial g_{p+1}}{\partial \beta_1} \frac{\partial g_{p+1}}{\partial \beta_j} \\ &= \sum_{k=1}^p \alpha_k (1 - \alpha_k) EX_{p-k+1} X_{p+1-j} + \sigma^2 V_{1j} \end{aligned} \quad (3.34)$$

De (3.33) et (3.34) on a :

$$W_{ij} = \sum_{k=1}^p \alpha_k (1 - \alpha_k) EX_{p-k+1} X_{p+2-i} X_{p+2-j} + \sigma^2 V_{ij} \quad (3.35)$$

Pour $1 \leq i \leq p+1, 1 \leq j \leq p+1$ avec la convention temporaire que $X_{p+1} = 1$. Alors on obtiens le théorème suivant.

Théorème 3.1

$$\hat{\beta} = (\mu^*, \hat{\alpha}^{\tau*}) \tau$$

est fortement cohérent. Si $E(\epsilon_t^3) < \infty$, alors :

$$(N - P)^{1/2} (\hat{\beta} - \beta) \longrightarrow MVN(0, V^{-1} W V^{-1}) \quad (3.36)$$

3.6 GÉNÉRALISATION

Afin d'enrichir la famille des modèles autoregressif qui permettent l'étude des séries chronologiques à valeurs entières Gautier et Latour (1994), Dion et al (1995) et Latour (1998) ont proposés une extension du modèle INAR(p), noté GINAR(p) pour Generalized Integer-valued Autoregressive of order p ou en Français (Modèle Autorégressif Généralisé à Valeurs Entières d'ordre p. Ce modèle se sert l'opérateur généralisé de Stautel et Van Harn pour offrir une flexibilité supplémentaire dans l'analyse et la modélisation des dépendances entre les observations d'une série chronologique à valeurs entières.

3.6.1 L'opérateur d'amincissement généralisé noté $\alpha \star$

Définition 3.2 Soit X une variable aléatoire à valeurs positives; Y une variable aléatoire à valeurs entières et positives de moyenne finie α et de variance finie σ^2 et (Y_k) une suite de variables aléatoires indépendantes, indépendantes de X et distribué selon la même loi que Y alors l'opérateur $\alpha \star$ est défini par :

$$\alpha \star X = \sum_{k=1}^X Y_k \quad (3.37)$$

La suite (Y_k) est aussi une série de comptage. de plus en supplant que $\beta \star$ est un autre opérateur d'amincissement généralisé de X , basé sur une série de comptage Z_k alors les opérateur $\alpha \star$ et $\beta \star$ sont indépendants si est seulement si les séries de comptage (Y_k) et (Z_k) sont mutuellement indépendantes.

3.6.2 Le modèle GINAR(p)

Soit $\{X_t, t \in \mathbb{Z}\}$ une suite de variables aléatoires à valeurs entières positives, $\{\varepsilon_t, t \in \mathbb{Z}\}$ une suite de variables aléatoires i.i.d. à valeurs entières positives, ayant une moyenne finie μ et une variance finie σ^2 :

$$E(\varepsilon_t) = \mu, \quad V(\varepsilon_t) = \sigma^2 \quad \text{pour tout } t. \quad (3.38)$$

Alors, $\{X_t, t \in \mathbb{Z}\}$ est un processus GINAR(p), s'il est de la forme :

$$X_t = \sum_{j=1}^p \alpha_j \star X_{t-j} + \varepsilon_t, \quad \forall t \in \mathbb{Z}, \quad (3.39)$$

Ici toutes les séries de comptage $(Y_{jk})_{k \in \mathbb{N}}$ associé à $\alpha_j \star$ pour $j = 1, 2, \dots, p$ sont indépendantes entre elles et indépendantes de ε_t , de plus selon Gautier et Latour elles ne doivent pas obligatoirement suivre une loi de bernouilli, ce qui différencie leurs résultats de ceux de Du & Li (1991).

3.6.3 Condition de stationnarité

Le processus Ginar(p) est stationnaire si :

$$\sum_{j=1}^p \alpha_j < 1 \quad (3.40)$$

3.6.4 Moments du premier et du second ordre

Les moments du premier et du second ordre du processus GINAR(p) sont tels que :

$$E(X_t) = \frac{\mu_{\varepsilon_t}}{1 - \sum_{j=1}^p \alpha_j} = \mu_x, \quad \text{Var}(X_t) = \mu_x \sum_{j=1}^p \lambda_j + \sum_{j=1}^p \alpha_j \gamma(j) + \sigma_{\varepsilon}^2 \quad (3.41)$$

3.7 PRÉVISION DANS LES MODÈLES INAR(p)

Dans cette section, nous abordons la problématique de la prévision dans les modèles INAR(p) en utilisant deux différentes approches :

3.7.1 Prédiction : Méthode de prédiction médiane

La méthode la plus courante dans la prévision des séries chronologiques consiste à utiliser des anticipations conditionnelles en tenant compte des observations en raison de la faible erreur quadratique qui en découle.

Cette méthode est principalement utilisée pour les modèles AR(p), cependant elle est moins pertinente pour les séries chronologiques à valeurs entières car elle produit des résultats continus.

D'autres méthodes peuvent être appliquées pour résoudre ce problème comme la méthode de prévision médiane ou la méthode bayésienne. Ici nous utiliserons la méthode de prévision médiane pour produire une prévision de la période d'étape h à venir avec une probabilité conditionnelle.

La méthode de prévision utilisant la méthode de prévision médiane utilise la probabilité conditionnelle en calculant la probabilité des valeurs qui pourraient être prédites dans la prochaine période.

On sait que $\alpha_1 \circ X_{t-1}, \alpha_2 \circ X_{t-2}, \dots, \alpha_p \circ X_{t-p}$ ont une distribution binomiale de paramètre X et α , On suppose aussi que l'innovation suit une distribution de Poisson alors, les probabilités conditionnelles pour le modèle Poisson INAR(p) sont :

$$\begin{aligned}
 P(X_t | X_{t-1}, \dots, X_{t-p}) = & \\
 & \sum_{i_1=0}^{\min(X_{t-1}, X_t)} \binom{X_{t-1}}{i_1} \alpha_1^{i_1} (1 - \alpha_1)^{X_{t-1} - i_1} \sum_{i_2=0}^{\min(X_{t-2}, X_{t-1} - i_1)} \binom{X_{t-2}}{i_2} \alpha_2^{i_2} (1 - \alpha_2)^{X_{t-2} - i_2} \dots \\
 & \sum_{i_p=0}^{[X_{t-p}, X_{t-(i_1+i_2+\dots+i_p)}]} \binom{X_{t-p}}{i_p} \alpha_p^{i_p} (1 - \alpha_p)^{X_{t-p} - i_p} \frac{e^{-\lambda X_{t-(i_1+i_2+\dots+i_p)}}}{[X_t - (i_1 + i_2 + \dots + i_p)]!}
 \end{aligned} \tag{3.42}$$

Et les probabilités conditionnelles pour les h étapes suivantes :

$$\begin{aligned}
 P(X_{t+h} | X_{t+h-1}, \dots, X_{t+h-p}) = & \\
 & \sum_{i_1=0}^{\min(X_{t+h-1}, X_{t+h})} \binom{X_{t+h-1}}{i_1} \alpha_1^{i_1} (1 - \alpha_1)^{X_{t+h-1} - i_1} \\
 & \sum_{i_2=0}^{\min(X_{t+h-2}, X_{t+h-1} - i_1)} \binom{X_{t+h-2}}{i_2} \alpha_2^{i_2} (1 - \alpha_2)^{X_{t+h-2} - i_2} \dots \\
 & \sum_{i_p=0}^{\min[X_{t+h-p}, X_{t+h} - (i_1+i_2+\dots+i_p)]} \binom{X_{t+h-p}}{i_p} \alpha_p^{i_p} (1 - \alpha_p)^{X_{t+h-p} - i_p} \\
 & \frac{e^{-\lambda X_{t+h} - (i_1+i_2+\dots+i_p)}}{[X_{t+h} - (i_1 + i_2 + \dots + i_p)]!}
 \end{aligned} \tag{3.43}$$

Cette probabilité conditionnelle sera utilisée pour calculer les résultats de prévision pour le modèle INAR(p) à h pas en avance avec la méthode de prévision médiane.

Soit m_h le plus petit entier résultant de $\sum_{x=0}^{m_h} P(X_{t+h} | X_{t+h-1}, \dots, X_{t+h-p}) \geq 0,5$, alors m_h est appelée la valeur médiane conditionnelle de X_{t+h} sachant $X_{t+h-1}, \dots, X_{t+h-p}$ et peut être utilisée comme valeur prévisionnelle pour le modèle INAR(p).

3.7.2 Approche des chaînes de Markov

Comme nous l'avons vu précédemment, le processus INAR(p) peut être considéré comme une chaîne de Markov prenant différentes valeurs dans un ensemble fini.

En principe le nombre d'états d'une chaîne de Markov est infini, cependant pour un ensemble de données donné, on suppose qu'il existe un entier positif M suffisamment grand tel que la probabilité d'observer un comptage supérieur à M est négligeable. Ainsi, nous pouvons limiter les valeurs prises par le processus à l'ensemble fini $\{0, 1, \dots, M\}$.

Exemple 3.1 Dans le cas où $p=2$, on peut considérer que les états du système sont donnés par des paires de valeurs consécutives du processus, ainsi, à l'instant $t-1$, la chaîne pourrait être dans l'un des états suivants :

$$S = \{(0,0), (0,1), \dots, (0,M), (1,0), (1,1), \dots, (1,M), (2,0), \dots\}$$

où (X_{t-2}, X_{t-1}) prend les valeurs $(i_2, i_1) \in S$ À l'instant t , le processus passe à une nouvelle paire de valeurs dans le même espace d'états, et les probabilités de transition d'un état à un autre sont données par :

$$P(X_t = j_1, X_{t-1} = j_2 | X_{t-1} = i_1, X_{t-2} = i_2) = P(X_t = j_1 | X_{t-1} = i_1, X_{t-2} = i_2) \quad (3.44)$$

Si $j_2 = i_1$ et 0 sinon.

La probabilité à droite est donnée par (3.43) avec $p = 2$ pour le modèle PoINAR(p). La probabilité zéro se produit lorsque $j_2 \neq i_1$ car les deux valeurs se réfèrent au processus X à la même période $t - 1$. Ce schéma peut être étendu pour prendre en compte des valeurs plus grandes de p . Ainsi, à n'importe quel instant t donné, il y a $(M + 1)p$ états différents dans l'ensemble S , déterminés par $\{X_{t-p+1}, X_{t-p+2}, \dots, X_t\}$. Les éléments de S sont des vecteurs de taille $p \times 1$, et pour chacun de ces vecteurs, le premier composant réfère à X_{t-p+1} , le deuxième se réfère à X_{t-p+2} , et ainsi de suite.

Nous définissons le vecteur $S(X_t)$ de dimension $(M + 1)p \times 1$ pour représenter les éléments des vecteurs de l'ensemble d'états S correspondant à X_t . Pour une chaîne de Markov à états finis, la distribution de prévision de chaque état à un instant t peut être obtenue à l'aide de la méthode de la matrice de transition.

Soit Q la matrice de transition de probabilité de dimension $(M + 1)^p \times (M + 1)^p$ d'un modèle INAR(p) avec un décompte maximal possible M . Cette matrice nous permet d'obtenir les prévisions de probabilité pour chaque état. Nous utilisons également un vecteur de probabilité π_t de dimension $(M + 1)^p \times 1$ pour représenter les probabilités de chaque état à un instant donné t . De plus, pour chaque valeur i de 0 à M , nous utilisons un vecteur de sélection S_i de dimension $(M + 1)^p \times 1$, qui identifie les positions correspondant à l'état $X_t = i$.

Ce vecteur a des entrées égales à 1 aux positions correspondant aux éléments de $S(X_t)$ pour lesquels X_t est égal à i , et des entrées nulles pour les autres

positions. Ainsi, nous pouvons exprimer la probabilité de $X_t = i$ comme $\pi'_t S_i$. Par conséquent, pour un processus INAR(p) général, les prévisions de probabilité conditionnelle pour X_{T+h} peuvent être obtenues à partir des prévisions du vecteur de probabilité π_{T+h} . Autrement dit :

$$P(X_{T+h} = i | X_T, \dots, X_{T-p+1}) = \pi'_{T+h} S_i \quad (3.45)$$

Soit Q et Q_h les matrices de transition à un pas et à h pas, respectivement, pour un système de Markov homogène d'ordre p . Alors :

$$Q^{(h)} = Q^{(h-1)}Q = Q^h, \quad (3.46)$$

et

$$\pi'_{T+h} = \pi'_{T+h-1}Q = \pi'_t Q^h. \quad (3.47)$$

L'équation (3.46) indique que la matrice de transition à h pas est égale à la puissance h de la matrice de transition à un pas, et l'équation (3.47) indique que la prévision à h pas en avant du vecteur de probabilité π_{T+h} est égale au produit du vecteur de probabilité courant π_T avec la matrice de transition à h pas. Ainsi, le vecteur de probabilité courant et la matrice de transition à un pas sont tout ce qui est nécessaire pour produire des prévisions pour n'importe quel nombre de périodes à l'avance. Nous pouvons résumer les développements précédents dans la proposition suivante :

Proposition 1 *Pour un processus INAR(p) général avec une valeur maximale possible M pour les comptages, la prévision à h pas en avant de la probabilité conditionnelle de $X_{T+h} | T = i$, pour $i = 0, 1, \dots, M$, est donnée par :*

$$P(X_{T+h} = i | X_T, \dots, X_{T-p+1}) = \pi'_T Q^h S_i; \quad (3.48)$$

Ici, π_T représente le vecteur de probabilité actuel, s_i représente le vecteur de sélection pour la valeur i et Q^h est la puissance h de la matrice de transition Q du processus.

La matrice de transition pour le processus PoINAR(p) peut être calculée à partir de l'équation (3.43), qui dépend des paramètres sous-jacents $\alpha_1, \dots, \alpha_p$ et λ .

3.8 DIAGNOSTIC DU MODÈLE

Après l'estimation des paramètres et la prévision du modèle INAR(p), il est nécessaire de déterminer si le modèle est bon et s'il décrit correctement le processus observé.

On procède en premier lieu à la sélection du meilleur modèle à l'aide d'informations des critères tels que AIC ou AICc pour un petit échantillon de données de comptage. Ensuite, il est également nécessaire de diagnostiquer les résidus du modèle.

AIC est l'estimateur du maximum de vraisemblance de la distance moyenne attendue de l'information de Kullback-Leibler du modèle estimé au modèle réel avec le paramètre K , où :

$$AIC = -2 \log(L(\hat{\theta} | x_1, x_2, \dots, x_n)) + 2K \quad (3.49)$$

L'AIC est une estimation biaisée si le modèle INAR a une petite taille d'échantillon. Ce biais peut être géré en ajoutant une composante non stochastique à

l'AIC qui se traduit par les critères de modification des informations d'Akaike ou l'Akaike corrigé Critère d'information (AICc) qui peut s'écrire comme suit :

$$AICc = AIC + \frac{2(k+1)(k+2)}{n-k-2} \quad (3.50)$$

De plus, l'estimation du paramètre θ à l'aide du maximum de vraisemblance pour le modèle INAR est analytiquement difficile à faire, de sorte que la valeur AIC est difficile à obtenir et aide à obtenir des informations sur les critères pour le modèle INAR. Ainsi, AICc peut aussi être défini comme :

$$AICc(p) = N \log(\hat{V}_p) + N \frac{(1+p/N)}{(1-p+2/N)} \quad (3.51)$$

Où :

$$\hat{V}_p = E [(X_{N+1} - \hat{X}_{N+1})^2] = \sigma_\varepsilon^2 + \frac{\mu_X}{p} \sum_{i=1}^p \sigma_i^2. \quad (3.52)$$

\hat{V}_p est la variance de l'erreur de la prévision pour la période suivante, N est le nombre de données, et σ_k^2 est la variance de $\alpha_k \circ X_{t-k}$.

3.8.1 Analyse des résidus

L'hypothèse de distribution de Poisson de son résidu ou de son innovation sur la formule de probabilité conditionnelle utilisée pour prévision, doit être testée avec le test de qualité d'ajustement.

Le résidu est obtenu en calculant la différence entre la valeur observée ou donnée et la valeur prédite précédemment obtenue en utilisant la médiane méthode de prévision de la valeur. Le résidu pour le modèle INAR(p) pourrait également être écrit comme $r_{kt} = E[\alpha_k \circ X_{t-k}] - \alpha_k X_{t-k}$.

Le test statistique du test d'ajustement est $\chi^2 = \sum \frac{(O-E)^2}{E}$ où O est la fréquence observée à partir de la fréquence de chaque catégorie de valeur possible et E est la fréquence attendue qui pourrait être calculée à l'aide de la fonction de masse de probabilité de distribution de Poisson. L'hypothèse nulle est rejetée si la valeur du test statistique calculé est supérieure à la valeur de la table du $\chi^2_{(\alpha, k-p-1)}$.

3.9 EXEMPLE DE SIMULATION

Dans cette section, les méthodes exposées précédemment sont appliquées à un ensemble de 115 données discrètes simulées, générées de manière aléatoire et dont la plage va de 0 à 10.

Le graphique de la série temporelle des données simulées est présenté dans la figure 3.1.

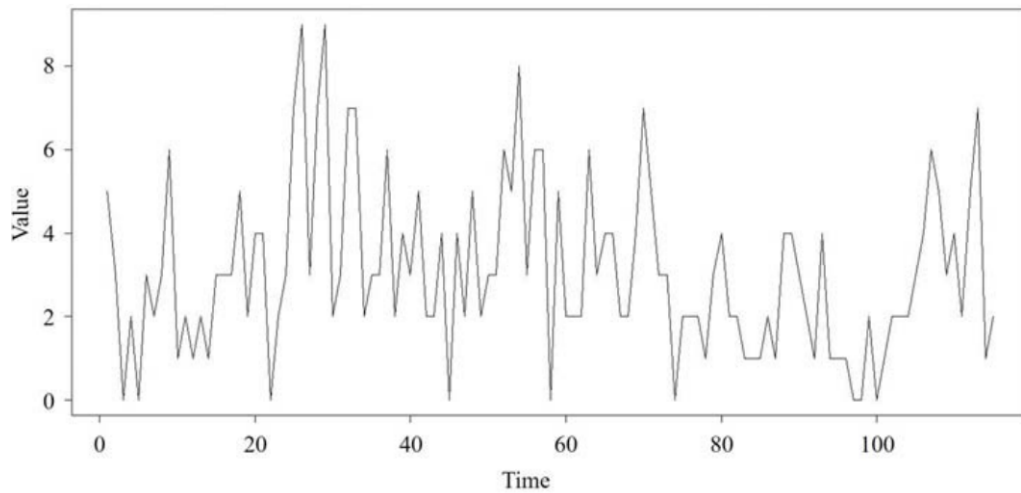


FIGURE 3.1 – Diagramme de série chronologique de 115 observations à partir de données simulées.

Selon le test de Dicky Fuller augmenté (ADF), la série temporelle est stationnaire.

Ensuite, à l’aide des fonctions d’autocorrélation (ACF) et d’autocorrélation partielle (PACF) représentées dans la Figure ci dessous on observe que la série temporelle des données est significative jusqu’à l’ordre 3 et peut être modélisée par un processus INAR(3) ou AR(3).

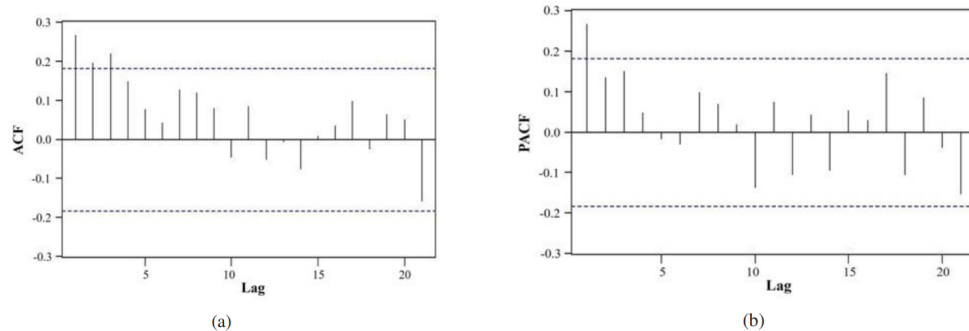


FIGURE 3.2 – a-ACF b-PACF

Le modèle de série temporelle INAR(3) estimé avec la méthode de Yule-Walker fournit les estimations suivantes pour les paramètres : $\alpha_1 = 0.2115, \alpha_2 = 0.0982, \alpha_3 = 0.1509$. Le modèle peut être écrit comme suit :

$$X_t = 0.2115 \circ X_{t-1} + 0.0982 \circ X_{t-2} + 0.1509 \circ X_{t-3} + \epsilon_t$$

Avec $\epsilon_t \sim \text{Poisson}$ Ce modèle indique que le nombre d’événements au temps $t-1$ contribue à hauteur de 0.2115 au nombre d’événements au temps t , au temps $t-2$ avec une contribution de 0.0982, et au temps $t-3$ avec une contribution de 0.1509. Au temps t , un nouvel événement se produit selon une distribution de Poisson avec une moyenne λ de 1.66.

Nous souhaitons maintenant observer les valeurs de prévision pour une période future de h étapes en utilisant le modèle AR(3).

$$\hat{X}_{115}(1) = 3.25; \quad \hat{X}_{115}(2) = 2.69; \quad \hat{X}_{115}(3) = 2.84; \quad \hat{X}_{115}(4) = 3.01; \quad \hat{X}_{115}(5) = 2.98.$$

Comme indiqué précédemment, ces résultats continus ne sont pas des résultats satisfaisants et ne peuvent pas être représentés en conditions réelles parce que nous nous attendions à des résultats discrets. Nous allons donc modéliser les données simulées en utilisant le modèle INAR(3) et calculer la probabilité conditionnelle des valeurs prédites pour la période future de h étapes, comme résumé dans le tableau 3.1 Le tableau 3.1 présente la probabilité conditionnelle de 15 valeurs discrètes possibles pour une période de cinq étapes à venir. Selon le tableau 3.1, les valeurs prévues seront la médiane ou le plus petit nombre ayant une probabilité conditionnelle cumulative supérieure ou égale à 0,5, à savoir :

De plus, les résidus des données ont été testés par un test d'adéquation aux données et ont été prouvés suivre une distribution de Poisson. Ensuite, nous pouvons diagnostiquer le modèle en comparant les modèles AR et INAR pour voir lequel est le meilleur pour les données simulées, en se basant sur les critères d'information présentés dans le Table. D'après le deuxième tableau, le modèle INAR(3) présente une valeur inférieure de l'AIC, de l'AICc et du BIC par rapport au modèle AR(3), ce qui prouve que ce modèle INAR est meilleur pour prévoir les données discrètes simulées pour les 115 périodes.

$$\hat{X}_{115}(1) = 3; \quad \hat{X}_{115}(2) = 2; \quad \hat{X}_{115}(3) = 3; \quad \hat{X}_{115}(4) = 3; \quad \hat{X}_{115}(5) = 3.$$

De plus, modèle INAR(3) présente une valeur inférieure de l'AIC, de l'AICc et du BIC par rapport au modèle AR(3), ce qui prouve que ce modèle INAR est meilleur pour prévoir les données discrètes simulées pour les 115 périodes.

TABLE 3.1 – Probabilité conditionnelle de X_{115+h} sachant X_{112+h} , X_{113+h} et X_{114+h} pour une série chronologique de données simulées.

h	$ph(X_{115+h} I_h)$				
	1	2	3	4	5
$ph(0 I_h)$	0.03407	0.06464	0.06277	0.04660	0.04950
$ph(1 I_h)$	0.12078	0.18461	0.18043	0.14966	0.15555
$ph(2 I_h)$	0.20766	0.25485	0.25169	0.23252	0.23662
$ph(3 I_h)$	0.23112	0.22735	0.22768	0.23343	0.23276
$ph(4 I_h)$	0.18752	0.14783	0.15056	0.17064	0.16687
$ph(5 I_h)$	0.11843	0.07491	0.07778	0.09707	0.09317
$ph(6 I_h)$	0.06072	0.03089	0.03276	0.04483	0.04228
$ph(7 I_h)$	0.02602	0.01068	0.01159	0.01732	0.01607
$ph(8 I_h)$	0.00953	0.00317	0.00352	0.00573	0.00523
$ph(9 I_h)$	0.00303	0.00082	0.00093	0.00017	0.00148
$ph(10 I_h)$	0.00085	0.00019	0.00022	0.00042	0.00037
$ph(11 I_h)$	0.00021	0.00004	0.00005	0.00009	0.00008
$ph(12 I_h)$	0.00005	0.00001	0.00001	0.00002	0.00002
$ph(13 I_h)$	0.00001	0.00000	0.00000	0.00000	0.00000
$ph(14 I_h)$	0.00000	0.00000	0.00000	0.00000	0.00000
$ph(15 I_h)$	0.00000	0.00000	0.00000	0.00000	0.00000

TABLE 3.2 – Diagnostic du modèle AR(3) et INAR(3).

Modèle	AIC	AICc	BIC
AR(3)	485.95	486.5	499.67
INAR(3)	242.89	243.11	119.87

3.10 CONCLUSION DU CHAPITRE

Dans ce chapitre, on a pris connaissance des modèles INAR(p) et de leurs estimations, on a examiné leurs principales propriétés, les méthodes d'estimation et de prévision. On a jeté les bases nécessaires pour une compréhension approfondie des modèles INAR(p) pour la modélisation des séries chronologique à valeurs entières.

CONCLUSION GÉNÉRALE

Au cours de ce travail, nous avons passé en revue quelques notions générales sur les séries chronologiques, pour ensuite nous attarder sur des modèles spécifiques aux séries à valeurs entières.

Nous avons exploré un type particulier de modèle, les INAR, d'ordre 1 dans un premier temps et d'ordre p dans un second temps.

Nous avons présenté différentes techniques d'estimation et de prédiction.

On a terminé l'analyse du modèle par le diagnostic de celui-ci, en mettant en évidence l'analyse des résidus pour évaluer la qualité de l'ajustement du modèle aux données.

En conclusion, cette étude a permis de comprendre l'importance de ces modèles de séries temporelles dans les domaines où les données de comptages sont couramment utilisés.

Il est important de noter que ce travail n'a fait qu'effleurer la surface d'un vaste domaine. Il existe ainsi de nombreux autres modèles de séries chronologique à valeurs entières à explorer, offrant ainsi des perspectives de recherches futures.

BIBLIOGRAPHIE

1. Akaikei, H. (1973). Information theory and an extension of maximum likelihood principle. In Proc. 2nd int. symp. on information theory (pp. 267-281).
2. Al-Osh, M. A., & Alzaid, A. A. (1987). First-order integer-valued autoregressive (INAR (1)) process. *Journal of Time Series Analysis*, 8(3), 261-275.
3. Alzaid, A., & Al-Osh, M. (1988). First-order integer-valued autoregressive (INAR (1)) process : distributional and regression properties. *Statistica Neerlandica*, 42(1), 53-61.
4. Bendjeddou, S. (2018). Inférence du quasi-maximum de vraisemblance de modèles de séries chronologiques à valeurs entières (Doctoral dissertation, Faculté de Mathématiques), USTHB.
5. Bentarzi, O. (2018). Sur des modèles de séries chronologiques périodiques à valeurs réelles ou entières (Doctoral dissertation, Faculté de Mathématiques), USTHB.
6. Bourguignon, M., & Weiß, C. H. (2017). An INAR (1) process for modeling count time series with equidispersion, underdispersion and overdispersion. *Test*, 26(4), 847-868.
7. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis : forecasting and control*. John Wiley & Sons.
8. Brännäs, K. and P. Johansson (1994). Time series count data regression. *Communications in Statistics : Theory and Methods*, 23, 2907-2925.
9. Bu, R., & McCabe, B. (2008). Model selection, estimation and forecasting in INAR (p) models : A likelihood-based Markov chain approach. *International journal of forecasting*, 24(1), 151-162.
10. Charpentier, A. (2006). *Cours de séries temporelles : théorie et applications*. Université Paris Dauphine.
11. Charemza, W. W., & Syczewska, E. M. (1998). Joint application of the Dickey-Fuller and KPSS tests. *Economics Letters*, 61(1), 17-21.
12. DE OLIVEIRA, Kenya Andresia, VANNUCCI, Alvaro, et DA SILVA, Elton Cesar. Using artificial neural networks to forecast chaotic time series. *Physica A : Statistical Mechanics and its Applications*, 2000, vol. 284, no 1-4, p. 393-404.
13. Dionne, G., & Gagné, R. (1997). The Non-Optimality of Deductible Contracts Against Fraudulent Claims : An Empirical Evidence in Automobile Insurance (cahier de recherche no 97-04, Chaire de gestion des risques, HEC, Montréal).
14. Dulkevičiūtė, J. (2022). Modelling integer-valued autoregressive processes with application to covid-19 data (Doctoral dissertation, Vilniaus universitetas).

15. Dufour, J. M. (2003). Histoire de l'analyse des séries chronologiques.
16. DU MAROC, R. O. Y. A. U. M. E. (2010). FILTRAGE DU TREND CYCLE : QUELQUES SOLUTIONS AUX PROBLEMES RENCONTRES. Séminaire sur l'extraction de la tendance-cycle, 18, 20.
17. Eduarda Da Silva, M., & Oliveira, V. L. (2004). Difference equations for the higher-order moments and cumulants of the INAR (1) model. *Journal of Time Series Analysis*, 25(3), 317-333.
18. Gardner Jr, E. S., & McKenzie, E. D. (1985). Forecasting trends in time series. *Management science*, 31(10), 1237-1246.
19. Gilks, W. R., Best, N. G., & Tan, K. K. (1995). Adaptive rejection Metropolis sampling within Gibbs sampling. *Journal of the Royal Statistical Society Series C : Applied Statistics*, 44(4), 455-472.
20. Fokianos, K. (2009). Integer-valued time series. *Wiley Interdisciplinary Reviews : Computational Statistics*, 1(3), 361-364.
21. Franke, J., & Rao Subba, T. (1993). Multivariate first-order integer-valued autoregressions.
22. Franke, J. and T. Seligmann (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. In : *Developments in Time Series Analysis : in honour of Maurice B. Priestley (T. Subba Rao, Ed.)*. Chapman & Hall. London. pp. 310-330.
23. Gouiri, S. (2021). Sur les modèles de séries chronologiques à valeurs entières. (Mémoire de master, Université de Bejaia).
24. Grunwald, G. K., Hyndman, R. J., Tedesco, L., & Tweedie, R. L. (2000). Theory & Methods : Non-Gaussian Conditional Linear AR (1) Models. *Australian & New Zealand Journal of Statistics*, 42(4), 479-495.
25. Jin-Guan, D., & Yuan, L. (1991). The integer-valued autoregressive (INAR (p)) model. *Journal of time series analysis*, 12(2), 129-142.
26. Kachour, M. (2009). Une nouvelle classe de modèles auto-régressifs à valeurs entières (Doctoral dissertation, Université Rennes 1).
27. Kendall, M. G. (1973). *Time-Series* (Griffin, London).
28. Latour, A. (1997). The multivariate GINAR(p) process. *Advances in Applied Probability*, 29, 228-248.
29. Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process. *Journal of Time Series Analysis*, 19(4), 439-455.
30. Lefebvre, M. (2005). *Processus stochastiques appliqués*. Presses inter Polytechnique.
31. McCullagh, P. & Nelder, J. A. (1989). *Generalized Linear Models*, 2nd ed. London : Chapman and Ha
32. McKenzie, E. (2003). Ch. 16. discrete variate time series. *Handbook of statistics*, 21, 573-606.
33. Novita, M., & Belinda, B. (2021, July). Integer-valued Pth-order autoregressive model. In *AIP Conference Proceedings (Vol. 2374, No. 1)*. AIP Publishing.

34. Oliveira, V.L. (2000). Autoregressive Models for Time Series of Counts : Characterization and Modeling. PhD dissertation. University of Porto.
35. Scotto, M. G., Weiss, C. H., & Gouveia, S. (2015). Thinning-based models in the analysis of integer-valued time series : a review. *Statistical Modelling*, 15(6), 590-618.
36. Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461-464.
37. Schweer, S., & Weiß, C. H. (2014). Compound Poisson INAR (1) processes : stochastic properties and testing for overdispersion. *Computational Statistics & Data Analysis*, 77, 267-284.
38. Shumway, R. H., Stoffer, D. S., & Stoffer, D. S. (2000). Time series analysis and its applications (Vol. 3). New York : springer.
39. Silva, I. M. M. D. (2005). Contributions to the analysis of discrete-valued time series.
40. Silva, N., Pereira, I., & Silva, M. E. (2009). Forecasting in INAR (1) model. *REVSTAT-Statistical Journal*, 7(1), 119-134.
41. Silva, I., & Silva, M. E. (2006). Asymptotic distribution of the Yule-Walker estimator for INAR(p) processes. *Statistics & Probability Letters*, 76(15), 1655-1663.
42. Simarmata, D. M., Novkaniza, F., & Widyaningsih, Y. (2017). A time series model : First-order integer-valued autoregressive (INAR (1)). In *AIP Conference Proceedings* (Vol. 1862, No. 1). AIP Publishing.
43. Steutel, F. W., & van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability*, 893-899
44. Taveira da Cunha, E. (2018). On INAR(1) models for integer time series (thèse de doctorat). Universidade Federal de Pernambuco.
45. Viano, M. C., & Philippe, A. (1999). *Maîtrise d'Économétrie : Cours de Séries Temporelles*. Université des Sciences et Technologies de Lille, UFR de Mathématiques Pures et Appliquées, 22-30.
46. Weiß, C. H. (2018). *An introduction to discrete-valued time series*. John Wiley & Sons.
47. Winkelmann, R. (2008). *Econometric analysis of count data*. Springer Science & Business Media.

Résumé

Nous nous proposons d'étudier des modèles de séries chronologiques à valeurs entières, principalement un modèle basé sur l'opérateur d'amincissement binomiale, le modèle autoregressif à valeurs entières d'ordre 1 (INAR(1)) et d'ordre p (INAR(p)).

On s'intéressera à leurs modèle probabiliste, méthodes d'estimations et de prédiction.

Abstract

We aim to study models of time series with integer values, primarily a model based on the binomial thinning operator, the autoregressive integer-valued model of order 1 (INAR(1)), and order p (INAR(p)).

We will focus on their probabilistic modeling, estimation methods, and prediction techniques.