แบบจำลองการถดถอยในตัวอันดับหนึ่งที่มีค่าเป็นจำนวนเต็มด้วยการแจกแจงแบบนัยทั่วไปของปัวซงลินด์เลย์สองพารามิเตอร์

นางสาวสุชิราภรณ์ บุนยะริส

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาคณิตศาสตร์ประยุกต์และวิทยาการคณนา
ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2562

FIRST ORDER INTEGER-VALUED AUTOREGRESSIVE MODEL WITH A TWO-PARAMETER GENERALIZED POISSON-LINDLEY DISTRIBUTION

Miss Suchiraporn Bunyaris

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Applied Mathematics and Computational Science

Department of Mathematics and Computer Science

Faculty of Science

Chulalongkorn University

Academic Year 2019

Thesis Title	FIRST ORDER INTEGER-VALUED AUTOREGRESSIVE MODEL WITH A TWO-PARAMETER GENERALIZED POISSON-LINDLEY DISTRIBUTION
By	Miss Suchiraporn Bunyaris
Field of Study	Applied Mathematics and Computational Science
Thesis Advisor	Assistant Professor Jiraphan Suntornchost, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

.... Dean of the Faculty of Science
(Professor Polkit Sangvanich, Ph.D.)

THESIS COMMITTEE

.... Chairman
(Assistant Professor Krung Sinapiromsaran, Ph.D.)

.... Thesis Advisor
(Assistant Professor Jiraphan Suntornchost, Ph.D.)

.... Examiner
(Associate Professor Nattakarn Chaidee, Ph.D.)

.... External Examiner
(Assistant Professor Angkana Boonyued, Ph.D.)

สุชิราภรณ์ บุนยะริส : แบบจำลองการถดถอยในตัวอันดับหนึ่งที่มีค่าเป็นจำนวนเต็มด้วยการแจกแจงแบบนัยทั่วไปของปัวซงลินด์เลย์สองพารามิเตอร์. (FIRST ORDER INTEGER-VALUED AUTOREGRESSIVE MODEL WITH A TWO-PARAMETER GENERALIZED POISSON-LINDLEY DISTRIBUTION) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ.ดร.จิราพรรณ สุนทรโชติ, 85 หน้า.

ในการศึกษานี้ เราได้แนะนำแบบจำลองการถดถอยในตัวอันดับหนึ่งที่มีค่าเป็นจำนวนเต็มสำหรับข้อมูลการนับด้วยการแจกแจงแบบนัยทั่วไปของปัวซงลินด์เลย์สองพารามิเตอร์ ตัวดำเนินการที่พิจารณาในการศึกษานี้คือ ตัวดำเนินการแบบทวินาม และแบบทวินามลบ สมบัติที่สำคัญทางความน่าจะเป็นและสถิติ เช่น โมเมนต์ ความหยุดนิ่ง และฟังก์ชันสัมพันธ์อัตโนมัติ ถูกสร้างขึ้น นอกจากนี้ยังทำการประมาณค่าพารามิเตอร์ และการจำลองเชิงตัวเลขเพื่อตรวจสอบแบบจำลองที่สร้างขึ้น สุดท้ายเราอภิปรายการประยุกต์ใช้ของแบบจำลองสำหรับชุดข้อมูลจริง

ภาควิชา	คณิตศาสตร์และวิทยาการคอมพิวเตอร์	ลายมือชื่อนิสิต ....
		ลายมือชื่อ อ.ที่ปรึกษาหลัก ....
สาขาวิชา	คณิตศาสตร์ประยุกต์	
ปีการศึกษา	2562	

6072005723 : MAJOR APPLIED MATHEMATICS AND COMPUTATIONAL SCIENCE

KEYWORDS : MIXED POISSON DISTRIBUTION/COUNT DATA/FIRST ORDER AUTOREGRESSIVE PROCESS/ TWO-PARAMETER/GENERALIZED POISSON-LINDLEY/ THINNING.

SUCHIRAPORN BUNYARIS : FIRST ORDER INTEGER-VALUED AUTOREGRESSIVE MODEL WITH A TWO-PARAMETER GENERALIZED POISSON-LINDLEY DISTRIBUTION. ADVISOR : ASST. PROF. JIRAPHAN SUNTORNCHOST, Ph.D., 85 pp.

In this study, we introduce the first order integer-valued autoregressive models for count data with the two-parameter generalized Poisson-Lindley distribution. The thinning operators considered in this study are the binomial and negative binomial thinning operators. Some important probabilistic and statistical properties such as moments, stationarity and autocorrelation functions are derived. Moreover, parameter estimations and numerical simulations to examine the constructed model are performed. Finally, applications of the models for real data sets are also discussed.

Department	: Mathematics and Computer Science	Student's Signature	....
Field of Study	: Applied Mathematics and Computational Science	Advisor's Signature	....
Academic Year	: 2019		

ACKNOWLEDGEMENTS

First of all, I would like to thank you for my thesis advisor, Assistant Professor Dr. Jiraphan Suntornchost, for perceptive suggestions and kindness that always help me to get through all of the research problems. She inspirits and recommends me to succeed in the thesis process.

Moreover, I would like to thank and deep appreciation for my thesis committees, Assistant Professor Dr. Krung Sinapiromsaran, Associate Professor Dr. Nattakarn Chaidee and Assistant Professor Dr. Angkana Suntadkarn, for their comments and suggestions. In particular, I would like to thank all teachers who have educated and taught me important knowledge.

Finally, I would like to thank my friends for giving me many supports and experiences at Chulalongkorn University. I would like to express my gratitude to my family for their love and supports throughout my undergraduate and graduate studies.

CONTENTS

CHAPTER

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Recently, the stationary time series models with discrete marginal distributions attract attentions from researchers in many fields. These models are found to be superior for count data in many applications such as the number of road accidents [9], the number of insurance claim counts [16], the number of stock transactions [8] and the number of submissions to animal health laboratories [5]. The model was first introduced by McKenzie [10] as the first order non-negative integer-valued autoregressive model (INAR(1)) by applying the binomial thinning operator introduced in Steutel and van Harn [15]. In 1988, Alzaid and Al-Osh [3] constructed the integer-valued time series model with Poisson marginal. They applied the model to count data time series. However, the Poisson marginal has equidispersion. This restriction might not be suitable to non-equidispersed data [1]. Therefore, alternative distributions have been explored in literatures such as generalized Poisson, geometric and Poisson-Lindley distribution. The generalized Poisson model introduced by Alzaid and Al-Osh [4] as an extension of the Poisson distribution of accommodate non-equidispersed data by McKenzie and et al [13]. Recently, Poisson-Lindley first order integer-valued autoregressive model was introduced by Mohammadpour and Shirozhan [5]. The model has the Poisson-Lindley marginal distribution introduced in Sankaran [12]. This distribution belongs to a compound Poisson family which is obtained from Poisson distribution when its parameter follows a Lindley distribution. Later in 2018, Rostami and Roozegar [9] used the distribution to introduce the INAR(1) model with Poisson-Lindley innovations based on the binomial and the negative binomial thinning operators. Integer-valued autoregressive models has been applied in many applications with different discrete marginal distributions such as Poisson, geometric, negative binomial and Poisson-Lindley distribution. In 2015, Bhati and Qadri [7] introduced a two-parameter generalized Poisson-Lindley distribution with parameter θ and β which are obtained from a mixed Poisson distribution when its mixing parameters follow a two-parameter Lindley distribution introduced in Shanker and Sharma [14]. This distribution is unimodal and over-dispersed. The distribution is more flexible than the Poisson-Lindley distribution. The two-parameter generalized Poisson-Lindley distribution is stated as follows. A random variable X is said to have a two-parameter generalized Poisson-Lindley distribution with parameter θ and β , denoted as $X \sim NGPL(\theta, \beta)$, if

$$P(X = x) = \frac{\theta^2}{(\theta+\beta)(1+\theta)^{x+1}}\left(1+\frac{\beta(x+1)}{1+\theta}\right) \qquad x = 0, 1, \ldots \text{ and } \theta, \beta \geq 0.$$

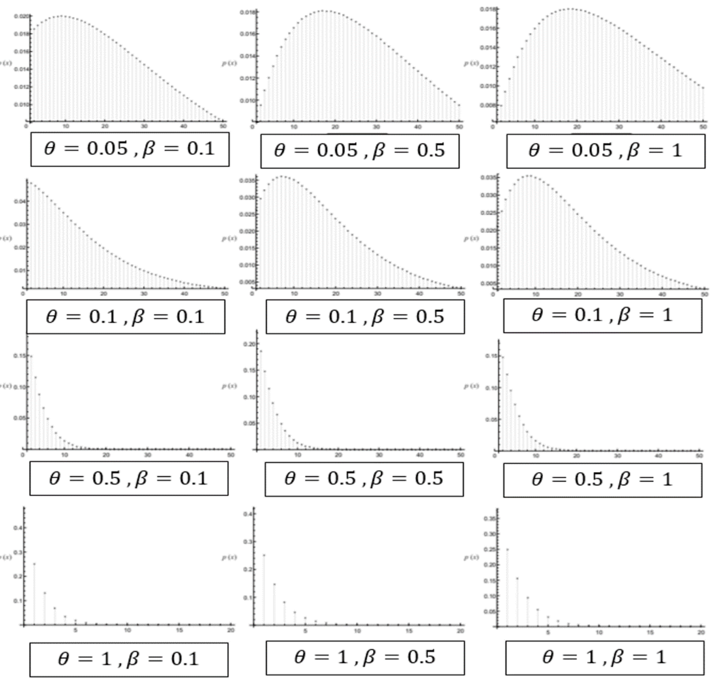


Figure 1.1: Probability density curves of the two-parameter generalized Poisson-Lindley distribution for different values of θ and β

For different values of two parameters, the probability function is evaluated and presented in Figure 1.1. From the figure, we can see that the distribution condenses and the right tail approaches to zero at a faster rate than the Poisson-Lindley distribution when θ increases for any fixed β. Therefore, the distribution is suitable for data having the right tail approaches to zero at a faster rate than the Poisson-Lindley distribution. Such data sets are commonly found in insurance business [7]. Therefore, the two-parameter generalized Poisson-Lindley distribution attracts attentions from many researchers. Development of the family of the Poisson-Lindley first order integer-valued autoregressive models are presented in Figure 1.2.

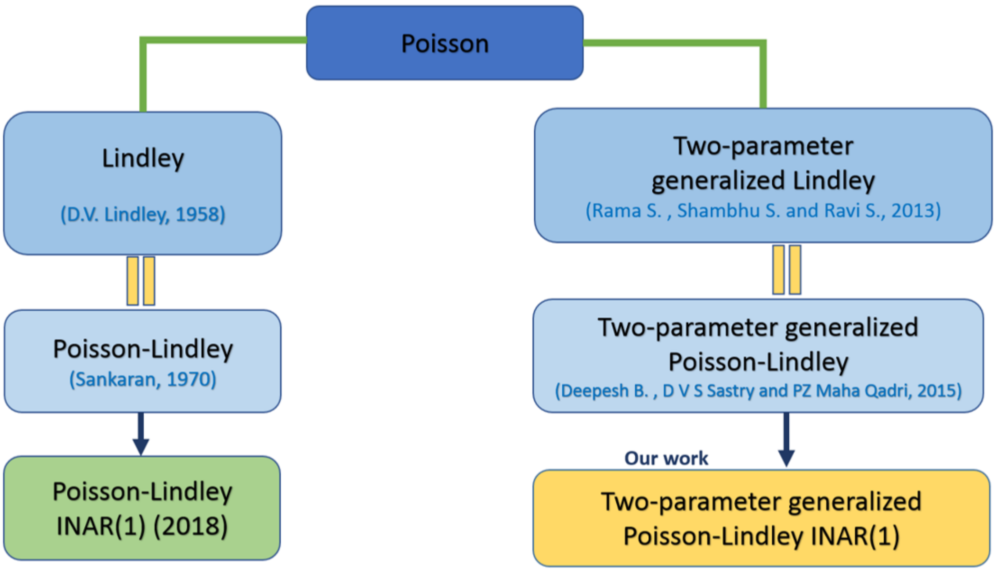


Figure 1.2: Development of the family of the Poisson-Lindley first order integer-valued autoregressive models

In this thesis, we apply the two-parameter generalized Poisson-Lindley to construct four new autoregressive models : (1) the first order integer-valued autoregressive models with a two-parameter generalized Poisson-Lindley distribution based on the binomial thinning operator, (2) the first order integer-valued autoregressive models with a two-parameter generalized Poisson-Lindley distribution based on the negative binomial thinning operator, (3) the first order integer-valued autoregressive model with a two-parameter generalized Poisson-Lindley innovation based on the binomial thinning operator and (4) the first order integer-valued autoregressive model with a two-parameter generalized Poisson-Lindley innovation based on the negative binomial thinning operator. The structure of these models are presented in Figure 1.3.

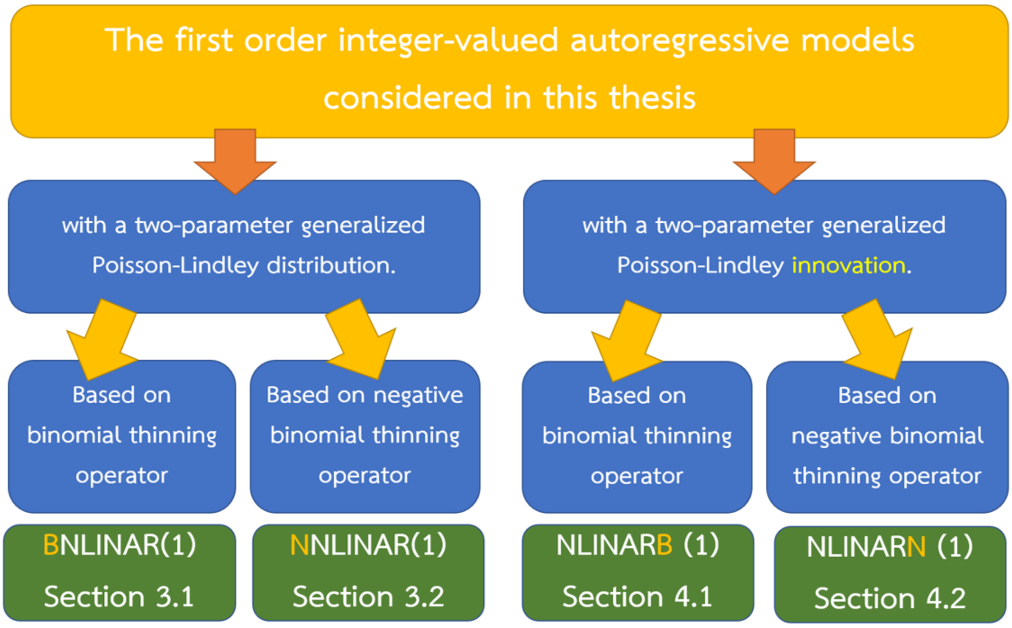


Figure 1.3: The first order integer-valued autoregressive models

In Chapter 3, we construct two first order integer-valued autoregressive models with a two-parameter generalized Poisson-Lindley distributions based on two thinning operators : (1) the binomial thinning operator, called the BNLINAR(1) model and (2) the negative binomial thinning operator, called the NNLINAR(1) model. The two models are discussed in Section 3.1 and Section 3.2, respectively. In this chapter, we derive some probabilistic properties of the constructed models such as moments and parameter estimations of the unknown parameters in the model by using the conditional least square estimator (CLS) and the Yule-Walker estimator (YW). These estimators are compared via Monte Carlo simulations in terms of their means and variances. Moreover, we discuss some possible applications of the BNLINAR(1) model and the NNLINAR(1) model for two real count time series.

In Chapter 4, we construct two first order integer-valued autoregressive models with a two-parameter generalized Poisson-Lindley innovations based on two thinning operators : (1) the binomial thinning operator, called the NLINARB(1) model and (2) the negative binomial thinning operator, called the NLINARN(1) model. The two models are discussed in Section 4.1 and Section 4.2, respectively. In this chapter, we derive some probabilistic properties of the constructed models such as moments and parameter estimations of the unknown parameters in the model by using the conditional least square estimator and the Yule-Walker estimator. These estimators are compared via Monte Carlo simulations in terms of their means and variances. Moreover, we discuss some possible applications of the NLINARB(1) model and the NLINARN(1) model for two real count time series. Finally, we discuss some possible applications of the NLINARB(1) model and the NLINARN(1) model for two real count time series. Conclusions of our study are provided in Chapter 5.

CHAPTER II

BACKGROUND KNOWLEDGE

In this chapter, we recall some important definitions and theorems that will be used repeatedly throughout this thesis. We start with the definitions of some necessary distributions and their properties.

2.1 Distribution

In this part, we discuss some background knowledge in probability theory such as distributions and moments. Since our work will cover only discrete distributions, all properties will be discussed in the setting of discrete random variables.

Definition 2.1.1. Let X be a discrete random variable with space R_X and probability mass function $f(\cdot)$. The expectation of X, denoted as $E(X)$, is defined by

$$E(X) = \sum_{x \in R_X} x f(x).$$

Definition 2.1.2. Let X be a discrete random variable with space R_X. The variance of X, denoetd as $Var(X)$, is defined by

$$Var(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2.$$

Definition 2.1.3. Let X be a discrete random variable with space R_X. The probability generating function of X, denoted as $\Phi_X(\cdot)$, is defined by

$$\Phi_X(s) = E(s^X) = \sum_{x \in R_X} s^x f(x).$$

for $s \in \mathbb{R}$.

Definition 2.1.4. Let X be a discrete random variable with space R_X. The moment generating function of X, denoted as $M_X(\cdot)$, is defined by

$$M_X(s) = E(e^{sX}) = \sum_{x \in R_X} e^{sx} f(x).$$

for $s \in \mathbb{R}$.

Definition 2.1.5. A random variable X is said to have the Bernoulli distribution with parameter p $(0 < p < 1)$, denoted as $X \sim Ber(p)$, if

$$P(X = x) = p^x q^{1-x},$$

for $x \in \{0, 1, \ldots\}$, where $q = 1 - p$.

Theorem 2.1.1. The Bernoulli random variable X with parameter p has the following properties

1. $E(X) = p$,
2. $Var(X) = pq$,
3. $\Phi_X(s) = q + ps$,
4. $M_X(s) = q + pe^s$.

Definition 2.1.6. A random variable X is said to have the binomial distribution with parameters n $(n \in \mathbb{N})$ and p $(0 < p < 1)$, denoted as $X \sim Bi(n, p)$, if

$$P(X = x) = \binom{n}{x} q^{n-x} p^x,$$

for $x \in \{0, 1, \ldots, n\}$, where $q = 1 - p$.

Theorem 2.1.2. The binomial random variable X with parameters n and p has the following properties:

1. $E(X) = np$,
2. $Var(X) = npq$,
3. $\Phi_X(s) = (q + ps)^n$,
4. $M_X(s) = (q + pe^s)^n$.

Definition 2.1.7. A random variable X is said to have the geometric distribution with parameter p $(0 < p < 1)$, denoted as $X \sim Geo(p)$, if

$$P(X = x) = q^x p,$$

for $x \in \{0, 1, \ldots\}$, where $q = 1 - p$.

Theorem 2.1.3. The geometric random variable X with parameter p has the following properties:

1. $E(X) = \frac{q}{p}$,
2. $Var(X) = \frac{q}{p^2}$,
3. $\Phi_X(s) = \frac{p}{1 - qs}$, for $|s| < \frac{1}{q}$,
4. $M_X(s) = \frac{p}{1 - qe^s}$, for $|s| < \frac{1}{q}$,

where $q = 1 - p$.

Definition 2.1.8. A random variable X is said to have the negative binomial distribution with parameters r ($r \in \mathbb{N}$) and p ($0 < p < 1$), denoted as $X \sim NB(r,p)$, if

$$P(X = x) = \binom{r+x-1}{x} q^x p^r,$$

for $x \in \{0, 1, \ldots\}$, where $q = 1 - p$.

Theorem 2.1.4. The negative binomial random variable X with parameters r and p has the following properties

1. $E(X) = \dfrac{rq}{p}$,
2. $Var(X) = \dfrac{rq}{p^2}$,
3. $\Phi_X(s) = (\dfrac{p}{1-qs})^r$, for $|s| < \dfrac{1}{q}$,
4. $M_X(s) = (\dfrac{p}{1-qe^s})^r$, for $|s| < \dfrac{1}{q}$,

where $q = 1 - p$.

Definition 2.1.9. (Bhati and Qadri [7]). A random variable X is said to have the two-parameter generalized Poisson-Lindley distribution with parameters θ and β, denoted as $X \sim NGPL(\theta, \beta)$, if

$$P(X = x) = \frac{\theta^2}{(\theta+\beta)(1+\theta)^{x+1}} \left(1 + \frac{\beta(x+1)}{1+\theta}\right)$$

for $x \in \{0, 1, \ldots\}$ and $\theta, \beta > 0$.

Theorem 2.1.5. (Bhati and Qadri [7]). Properties of the two-parameter generalized Poisson-Lindley random variable X with parameters β and θ defined in Definition 2.1.9 are as follows.

1. $E(X) = \dfrac{2\beta + \theta}{\theta(\beta + \theta)}$,

2. $Var(X) = \dfrac{2\beta^2(1+\theta)+\theta^2(1+\theta)+\beta\theta(4+3\theta)}{\theta^2(\beta+\theta)^2}$,

3. $\Phi_X(s) = \dfrac{\theta^2(\beta+\theta-s+1)}{(\beta+\theta)(\theta-s+1)^2}$, for $s \in \mathbb{R}$,

4. $M_X(s) = \dfrac{\theta^2(\beta+\theta-e^s+1)}{(\beta+\theta)(\theta-e^s+1)^2}$, for $s \in \mathbb{R}$.

Definition 2.1.10. Let X be a discrete random variable on space R_X. Then X has a degenerate distribution with parameter r if X is degenerated at r. That is

$$P(X=x) = \begin{cases} 1, & \text{for} \quad x = r \\ 0, & \text{for} \quad x \neq r. \end{cases}$$

Definition 2.1.11. We say that the distribution function $F(\cdot)$ is a generalized mixture of the distribution functions $F(\cdot\,;1), F(\cdot\,;2), \ldots$ if

$$F(x) = \sum_{i\geq 1} w_i F(x;i),$$

for all $x \in \mathbb{R}$ where $w_1, w_2, \ldots$ are real numbers such that $\sum_{i\geq 1} w_i = 1$, $\sum_{i\geq 1} |w_i| < \infty$ and for some index i, $w_i < 0$.

2.2 Conditional distribution

In this part, we discuss the definitions and properties of conditional distribution, conditional mean and conditional variance.

Definition 2.2.1. Let X and Y be two random variables with joint density $f(x,y)$ and marginals $f_1(x)$ and $f_2(y)$, respectively. The conditional probability density function g of X, given (the event) $Y = y$, is defined as

$$g(x|y) = \frac{f(x,y)}{f_2(y)},$$

where $f_2(y) > 0$. Similary, the conditional probability density function h of X, given (the event) $X = x$, is defined as

$$h(y|x) = \frac{f(x,y)}{f_1(x)},$$

where $f_1(x) > 0$.

Definition 2.2.2. Let X and Y be discrete random variables with space R_X. The conditional expectation of X, given that $Y = y$, is

$$E(X|Y = y) = \sum_{x \in R_X} x f_{X|Y}(x|y).$$

Definition 2.2.3. Let X and Y be discrete random variables with space R_X. The conditional variance of X, given that $Y = y$, is

$$Var(X|Y = y) = E(X^2|Y = y) - (E(X|Y = y))^2.$$

Theorem 2.2.1. Properties of conditional expectation and conditional variance are as follows.

1. $E(a|Y) = a$ if $a \in \mathbb{R}$,
2. $E(X) = E(E(X|Y))$,
3. $E(X|Y) = E(X)$ if X and Y are independent,
4. $Var(X) = E(Var(X|Y)) + Var(E(X|Y))$.

2.3 Time series and stationary process

In this part, we state some important the concepts of time series that are necessary such as stationarity and autocorrelation function.

Definition 2.3.1. Time series $\{Z_t; t \in \mathbb{N}\}$ is a series of data points indexed in $\{1, 2, \ldots\}$. If Z_t has an integer value, the time series is called the integer-valued time series.

Definition 2.3.2. A process $\{Z_t; t \in \mathbb{N}\}$ is said to be the first-order stationary if

$$F_{Z_{t_1}}(z_1) = F_{Z_{t_1+k}}(z_1)$$

any $t_1 \in \mathbb{N}$ and $k \in \mathbb{Z}$. A process $\{Z_t; t \in \mathbb{N}\}$ is said to be the second-order stationary if

$$F_{Z_{t_1},Z_{t_2}}(z_1, z_2) = F_{Z_{t_1+k,t_2+k}}(z_1, z_2)$$

any $t_1, t_2 \in \mathbb{N}$ and $k \in \mathbb{Z}$. A process $\{Z_t; t \in \mathbb{N}\}$ is said to be stationary if

$$F_{Z_{t_1},Z_{t_2},\ldots,Z_{t_s}}(z_1, z_2, \ldots, z_s) = F_{Z_{t_1+k},Z_{t_2+k},\ldots,Z_{t_s+k}}(z_1, z_2, \ldots, z_s)$$

for any finite set of indices $\{t_1, t_2, \ldots, t_s\} \subset \mathbb{N}$ with $s \in \mathbb{N}$, and $k \in \mathbb{Z}$.

Definition 2.3.3. For a given real-valued process $\{Z_t; t \in \mathbb{N}\}$, the mean function of the process is

$$\mu_t = E(Z_t).$$

The variance function of the process is

$$\sigma_t^2 = E(Z_t - \mu_t)^2.$$

The covariance function of the between Z_{t_1} and Z_{t_2} is

$$\gamma(t_1, t_2) = Cov(Z_{t_1}, Z_{t_2}) = E[(Z_{t_1} - \mu_{t_1})(Z_{t_2} - \mu_{t_2})].$$

The correlation function between Z_{t_1} and Z_{t_2} is

$$\rho(t_1, t_2) = \frac{Cov(Z_{t_1}, Z_{t_1})}{\sqrt{Var(Z_{t_1})}\sqrt{Var(Z_{t_2})}} = \frac{\gamma(t_1, t_2)}{\sigma_{t_1}\sigma_{t_2}}.$$

For a stationary process, the mean function and the variance function are constant. That is $\mu_t = \mu$ and $\sigma_t^2 = \sigma$, respectively. Then, the autocovariance function, γ_k, of Z_t and Z_{t+k} for any $k \in \mathbb{N}$ is defined as

$$\gamma_k = \gamma(t-k, t) = \gamma(t, t+k).$$

The corresponding autocorrelation function, ρ_k, is

$$\rho_k = \rho(t-k, t) = \rho(t, t+k).$$

2.4 Integer-valued time series model

In this part, we give the definition of the first order integer-valued autoregressive model (INAR(1)). Since the integer-valued time series discussed in our study are based on the binomial thinning operator and the negative binomial thinning operator. We first discuss the definitions and properties of the two thinning operators.

Definition 2.4.1. The binomial thinning operator, $\alpha\circ$, is defined as

$$\alpha \circ X = \sum_{i=1}^{X} W_i, \quad \alpha \in (0, 1), \tag{2.1}$$

where X is a non negative integer-valued random variable and $\{W_i\}_{i\geq 1}$ is a sequence of independent and identically distributed random variables with $Ber(\alpha)$ distribution and is independent of X.

Theorem 2.4.1. (Steutel and van Harn [15]). The properties of the binomial thinning operator defined in Definition 2.4.1 are as follows.

1. $E(\alpha \circ X) = \alpha E(X)$,
2. $Var(\alpha \circ X) = \alpha(1-\alpha)E(X) + \alpha^2 Var(X)$,
3. $\Phi_{\alpha \circ X}(s) = \Phi_X(1-\alpha+\alpha s)$ for $s \in \mathbb{R}$,
4. $M_{\alpha \circ X}(s) = M_X(1-\alpha+\alpha e^s)$ for $s \in \mathbb{R}$.

Definition 2.4.2. The negative binomial thinning operator, $\alpha *$, is defined as

$$\alpha * X = \sum_{i=1}^{X} Z_i, \quad \alpha \in (0,1), \tag{2.2}$$

where X is a non negative integer-valued random variable and $\{Z_i\}_{i \geq 1}$ is a sequence of independent and identically distributed random variables with $Geo(\frac{1}{1+\alpha})$ distribution and is independent of X.

Theorem 2.4.2. (Ristic and Nastic [6]). The properties of the negative binomial thinning operator defined in Definition 2.4.2 are as follows.

1. $E(\alpha * X) = \alpha E(X)$,
2. $Var(\alpha * X) = \alpha(1-\alpha)E(X) + \alpha^2 Var(X)$,
3. $\Phi_{\alpha * X}(s) = \Phi_X((1+\alpha-\alpha s)^{-1})$ for $s \in \mathbb{R}$,
4. $M_{\alpha * X}(s) = M_X((1+\alpha-\alpha e^s)^{-1})$ for $s \in \mathbb{R}$.

Definition 2.4.3. The first order integer-valued autoregressive model based on the binomial thinning operator is defined as

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \tag{2.3}$$

where the binomial thinning $\alpha\circ$ is defined in Definition 2.4.1 and the innovation process $\{\epsilon_t\}_{t\geq 1}$ is a sequence of independent and identically distributed (i.i.d.) random variables such that $\alpha \circ X_{t-1}$ and ϵ_t are independent.

Definition 2.4.4. The first order integer-valued autoregressive model based on the negative binomial thinning operator is defined as

$$X_t = \alpha * X_{t-1} + \varepsilon_t, \tag{2.4}$$

where the negative binomial thinning $\alpha*$ is defined in Definition 2.4.2 and the innovation process $\{\varepsilon_t\}_{t\geq 1}$ is a sequence of independent and identically distributed (i.i.d.) random variables such that $\alpha * X_{t-1}$ and ε_t are independent.

Definition 2.4.5. The integer-valued time series $\{X_t; t \in \mathbb{N}\}$ is said to be stationary if

$$F_{X_{t_1},X_{t_2},\ldots,X_{t_s}}(x_1, x_2, \ldots, x_s) = F_{X_{t_1+k},X_{t_2+k},\ldots,X_{t_s+k}}(x_1, x_2, \ldots, x_s)$$

for any finite set of indices $\{t_1, t_2, \ldots, t_s\} \subset \mathbb{N}$ with $s \in \mathbb{N}$, and $k \in \mathbb{Z}$.

2.5 Parameter estimation

In this part, we discussion two Parameter estimations for INAR(1) model which are the conditional least squares estimators and the Yule-Walker estimators.

Definition 2.5.1. The conditional least squares estimators of the parameters are obtained by minimizing the function

$$Q_n = \sum_{t=2}^{n}(X_t - E(X_t|X_{t-1}))^2,$$

where $n \geq 2$ and $E(X_t|X_{t-1})$ is the conditional mean of X_t.

Definition 2.5.2. The Yule-Walker equation for $INAR(1)$ defined as

$$\alpha\hat{\gamma}(0) = \hat{\gamma}(1)$$

where $\hat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(X_t - \bar{X})(X_{t+k} - \bar{X})$ and $\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$ is the sample mean.

CHAPTER III

INAR(1) MODEL WITH THE TWO-PARAMETER GENERALIZED POISSON-LINDLEY DISTRIBUTION

In this chapter, we construct two first order integer-valued autoregressive models with a two-parameter generalized Poisson-Lindley distribution based on (1) the binomial thinning operator and (2) the negative binomial thinning operator. Moreover, probabilistic properties of the constructed models and parameter estimation are demonstrated.

3.1 Construction of the first order integer-valued autoregressive models with the two-parameter generalized Poisson-Lindley distribution based on the binomial thinning operator (BNLINAR(1))

In this section, we construct the first order integer-valued autoregressive models with two-parameter generalized Poisson-Lindley distribution based on the binomial thinning operator model (BNLINAR(1)). Moreover, we investigate many properties of the constructed model such as moments, parameter estimations and perform some numerical studies.

Definition 3.1.1. The first order integer-valued autoregressive model with two-parameter generalized Poisson-Lindley distribution based on the binomial thinning operator (BNLINAR(1)) $\{X_t\}_{t\geq 1}$ is defined as

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \tag{3.1}$$

where the binomial thinning $\alpha\circ$ is defined in Definition 2.4.1, $\{X_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$ distribution and $\{\epsilon_t\}_{t\geq 1}$ is a sequence of i.i.d. random variables such that $\alpha \circ X_{t-1}$ and ϵ_t are independent.

Theorem 3.1.1. The innovation process $\{\epsilon_t\}_{t\geq 1}$ has the probability generating function

$$\Phi_{\epsilon_t}(s) = \frac{(\beta + \theta - s + 1)(\theta + \alpha(1-s))^2}{(\theta - s + 1)^2(\theta + \beta + \alpha(1-s))}, \tag{3.2}$$

for $s \in \mathbb{R}$.

Proof. Since $\{X_t\}_{t\geq 1}$ is a stationary process with $NGPL(\theta, \beta)$, from Theorem 2.4.1(5),

$$\Phi_{\alpha\circ X_{t-1}}(s) = E(s^{(\alpha \circ X_{t-1})}) = \Phi_{X_t}(1 - \alpha + \alpha s).$$

From (3.1) and the property that $\alpha \circ X_{t-1}$ and ϵ_t are independent for $s \in \mathbb{R}$,

$$\begin{aligned}
\Phi_{X_t}(s) &= E(s^{X_t}) \\
&= E(s^{\alpha \circ X_{t-1} + \epsilon_t}) \\
&= E(s^{\alpha \circ X_{t-1}} s^{\epsilon_t}) \\
&= E(s^{(\alpha \circ X_{t-1})}) E(s^{\epsilon_t}) \\
&= \Phi_{\alpha\circ X_{t-1}}(s)\Phi_{\epsilon_t}(s) \\
&= \Phi_{X_t}(1 - \alpha + \alpha s)\Phi_{\epsilon_t}(s).
\end{aligned}$$

From Theorem 2.1.5(3), the innovation process $\{\epsilon_t\}_{t\geq 1}$ has the probability generating function (pgf)

$$\begin{aligned}
\Phi_{\epsilon_t}(s) &= \frac{\Phi_{X_t}(s)}{\Phi_{X_t}(1-\alpha+\alpha s)} \\
&= \frac{\theta^2(\beta+\theta-s+1)}{(\beta+\theta)(\theta-s+1)^2}\frac{(\beta+\theta)(\theta-1+\alpha-\alpha s+1)^2}{\theta^2(\beta+\theta-1+\alpha-\alpha s+1)} \\
&= \frac{(\beta+\theta-s+1)(\theta+\alpha-\alpha s)^2}{(\theta-s+1)^2(\beta+\theta+\alpha-\alpha s)}.
\end{aligned}$$

□

Lemma 3.1.1. The generalized mixture

$$\begin{aligned}
g(x) = &\left[\frac{\alpha^2\theta(\theta-\beta)+\theta(\theta+\beta)+2\alpha(\beta^2-\theta^2)}{(\beta+\theta(1-\alpha))^2}\right]\frac{\theta}{1+\theta}\left(1-\frac{\theta}{1+\theta}\right)^x \\
&+\frac{\beta(1-\alpha)}{\beta+\theta(1-\alpha)}(x+1)\left(\frac{\theta}{1+\theta}\right)^2\left(1-\frac{\theta}{1+\theta}\right)^x \\
&-\frac{\alpha\beta^2}{(\beta+\theta(1-\alpha))^2}\frac{\theta+\beta}{\theta+\beta+\alpha}\left(1-\frac{\theta+\beta}{\theta+\beta+\alpha}\right)^x
\end{aligned} \tag{3.3}$$

is a probability mass function where $0<\alpha<1$, $\theta\geq 1$, and $\beta>0$.

Proof. Let

$$w_1 = \frac{\alpha^2\theta(\theta-\beta)+\theta(\theta+\beta)+2\alpha(\beta^2-\theta^2)}{(\beta+\theta(1-\alpha))^2},$$

$$w_2 = \frac{\beta(1-\alpha)}{\beta+\theta(1-\alpha)},$$

and

$$w_3 = -\frac{\alpha\beta^2}{(\beta+\theta(1-\alpha))^2},$$

and let

$$g_1(x) = \frac{\theta}{1+\theta}\left(1-\frac{\theta}{1+\theta}\right)^x,$$
$$g_2(x) = (x+1)\left(\frac{\theta}{1+\theta}\right)^2\left(1-\frac{\theta}{1+\theta}\right)^x,$$
$$g_3(x) = \frac{\theta+\beta}{\theta+\beta+\alpha}\left(1-\frac{\theta+\beta}{\theta+\beta+\alpha}\right)^x.$$

Then $w_1 + w_2 + w_3 = 1$, $g_1(\cdot)$ is the probability mass function of $Geo(\frac{\theta}{1+\theta})$, $g_2(\cdot)$ is the probability mass function of $NB(2, \frac{\theta}{1+\theta})$ and $g_3(\cdot)$ is the probability mass function of $Geo(\frac{\theta+\beta}{\beta+\theta+\alpha})$. Thus,

$$\sum_{x=0}^{\infty} g(x) = w_1 \sum_{x=0}^{\infty} g_1(x) + w_2 \sum_{x=0}^{\infty} g_2(x) + w_3 \sum_{x=0}^{\infty} g_3(x) = 1.$$

Following Mohammadpour and Shirozhan [5], we next show that $g(x) \geq 0$ for $x \in \{0, 1, \ldots\}$ and the function $g(x)$ can be written as

$$g(x) = \left(\frac{1}{1+\theta}\right)^x r(x),$$

where

$$r(x) = w_1\left(\frac{\theta}{1+\theta}\right) + w_2\left(\frac{\theta}{1+\theta}\right)^2(x+1) + w_3\left(\frac{\theta+\beta}{\theta+\beta+\alpha}\right)\left(\frac{\alpha(1+\theta)}{\theta+\beta+\alpha}\right)^x.$$

First, we show that $(r(x))' > 0$ for $x \in \{0, 1, \ldots\}$.
Since $w_3 < 0$ and $\ln\left(\frac{\alpha(1+\theta)}{\theta+\beta+\alpha}\right) < 0$,

$$(r(x))' = w_2\left(\frac{\theta}{1+\theta}\right)^2 + w_3\left(\frac{\theta+\beta}{\theta+\beta+\alpha}\right)\left(\frac{\alpha(1+\theta)}{\theta+\beta+\alpha}\right)^x \ln\left(\frac{\alpha(1+\theta)}{\theta+\beta+\alpha}\right)$$
$$> 0.$$

Then, $(r(x))' > 0$ for $x \in \{0, 1, \ldots\}$.

Moreover,

$$\begin{aligned}
\lim_{x\to\infty} r(x) &= \lim_{x\to\infty}\left[w_1\left(\frac{\theta}{1+\theta}\right)+w_2\left(\frac{\theta}{1+\theta}\right)^2(x+1)+w_3\left(\frac{\theta+\beta}{\theta+\beta+\alpha}\right)\left(\frac{\alpha(1+\theta)}{\theta+\beta+\alpha}\right)^x\right]\\
&= \lim_{x\to\infty}\left[w_1\left(\frac{\theta}{1+\theta}\right)+w_2\left(\frac{\theta}{1+\theta}\right)^2(x+1)\right]\\
&= +\infty.
\end{aligned}$$

Since $\theta \geq 1$, we can show that

$$\begin{aligned}
r(0) &= w_1\left(\frac{\theta}{1+\theta}\right)+w_2\left(\frac{\theta}{1+\theta}\right)^2+w_3\left(\frac{\theta+\beta}{\theta+\beta+\alpha}\right)\\
&\geq \frac{w_1}{2}+\frac{w_2}{4}+w_3,
\end{aligned}$$

and

$$\begin{aligned}
\frac{w_1}{2}+\frac{w_2}{4}+w_3 &= \frac{\alpha^2\theta(\theta-\beta)+\theta(\theta+\beta)+2\alpha(\beta^2-\theta^2)}{2(\beta+\theta(1-\alpha))^2}+\frac{\beta(1-\alpha)}{4(\beta+\theta(1-\alpha))}\\
&\quad -\frac{\alpha\beta^2}{(\beta+\theta(1-\alpha))^2}\\
&= \frac{(1-\alpha)(\beta^3+4\beta^2\theta+\beta\theta(1-\alpha)(2-\alpha(\theta-2)+3\theta)+2\theta^2(1-\alpha)^2)}{4(1-\alpha+\beta)(\beta+\theta-\alpha\theta)^2}\\
&\geq 0.
\end{aligned}$$

Therefore, $g(\cdot)$ is a probability mass function. □

Theorem 3.1.2. The innovation sequence $\{\epsilon_t\}_{t\geq 1}$ of the $BNLINAR(1)$ model defined by (3.1) process has the probability mass function

$$f_\epsilon(x) = \alpha h(x) + (1-\alpha)g(x),$$

where $h(\cdot)$ is the degenerate distribution function at zero defined in Definition 2.1.10 and $g(\cdot)$ is the probability mass function defined in Lemma 3.1.1.

Proof. From Theorem 3.1.1, the generating function of ϵ is

$$\Phi_\epsilon(s) = \frac{(\beta+\theta-s+1)(\theta+\alpha(1-s))^2}{(\theta-s+1)^2(\theta+\beta+\alpha(1-s))}.$$

The function can be written in terms of a weighted sum of three probability generating functions as follows.

$$\Phi_\epsilon(s) = \alpha + (1-\alpha)\left[w_1\Phi_1(s) + w_2\Phi_2(s) + w_3\Phi_3(s)\right],$$

where the weights are

$$w_1 = \frac{\alpha^2\theta(\theta-\beta)+\theta(\theta+\beta)+2\alpha(\beta^2-\theta^2)}{(\beta+\theta(1-\alpha))^2},$$

$$w_2 = \frac{\beta(1-\alpha)}{\beta+\theta(1-\alpha)},$$

and

$$w_3 = -\frac{\alpha\beta^2}{(\beta+\theta(1-\alpha))^2},$$

and the probability generating functions are

$$\Phi_1(s) = \frac{\theta}{1+\theta-s},$$

$$\Phi_2(s) = \frac{\theta^2}{(1+\theta-s)^2},$$

and

$$\Phi_3(s) = \frac{(\beta+\theta)}{\beta+\theta+\alpha(1-s)},$$

respectively. We can see that $w_1+w_2+w_3=1$. Moreover, $\Phi_1(\cdot)$ is the probability generating function of $Geo(\frac{\theta}{1+\theta})$, $\Phi_2(\cdot)$ is the probability generating function of $NB(2,\frac{\theta}{1+\theta})$ and $\Phi_3(\cdot)$ is the probability generating function of $Geo(\frac{\theta+\beta}{\beta+\theta+\alpha})$. Thus, the probability density function of ϵ can be written as

$$f_\epsilon(x) = \alpha h(x) + (1-\alpha)g(x),$$

where $h(\cdot)$ is the degenerate distribution function at zero and $g(\cdot)$ is the probability mass function defined in Lemma 3.1.1. □

Definition 3.1.2. [15] A discrete distribution with the probability generating function Φ is called self-decomposable if

$$\Phi(z) = \Phi(1-\alpha+\alpha z)\Phi_\alpha(z) \quad ; \alpha \in (0,1).$$

with Φ_α the probability generating function.

Corollary 3.1.1. The two-parameter generalized Poisson-Lindley distribution is self-decomposable.

Proof. From Theorem 3.1.2, the mass function of ϵ is

$$f_\epsilon(x) = \alpha h(x) + (1-\alpha)g(x)$$

where $g(\cdot)$ is the probability mass function defined in Lemma 3.1.1 and $h(\cdot)$ is the degenerate distribution function at zero. Thus, the two-parameter generalized Poisson-Lindley distribution is self-decomposable defined in Definition 3.1.2 □

Theorem 3.1.3. The process $\{X_t\}_{t\geq 1}$ defined in (3.1) can be rewritten as

$$X_t = \begin{cases} \alpha \circ X_{t-1}, & \text{w.p.} \quad \alpha, \\ \alpha \circ X_{t-1} + \epsilon_t, & \text{w.p.} \quad 1-\alpha, \end{cases}$$

where w.p. stands for "with probability".

Thus, we can write the process X_t as

$$X_t = \alpha \circ X_{t-1} + I_t H_t, \tag{3.4}$$

where H_t has the probability mass function $g(\cdot)$ defined in Lemma 3.1.1, I_t is the Bernoulli with parameter α and I_tH_t is independent of X_{t-k} for $k \geq 1$.

Proof. From Theorem 3.1.2 and the process (3.1), the probability mass function of ϵ is

$$f_\epsilon(x) = \alpha h(x) + (1-\alpha)g(x)$$

where $h(\cdot)$ is the degenerate distribution function at zero and $g(\cdot)$ is the probability mass function defined in Lemma 3.1.1. Then, the mass function of ϵ can be rewritten as

$$f_\epsilon(x) = \begin{cases} \alpha, & \text{for} \quad x = 0 \\ (1-\alpha)g(x), & \text{for} \quad x \neq 0. \end{cases}$$

Let I_t is the Bernoulli with parameter α and H_t has the probability mass function $g(\cdot)$ defined in Lemma 3.1.1. Thus, we can write X_t as

$$X_t = \alpha \circ X_{t-1} + I_tH_t,$$

where I_tH_t is independent of X_{t-k} for $k \geq 1$. □

Remark 3.1.1. The marginal distribution of the model (3.4) based on the binomial thinning operator can be expressed in terms of the innovation sequence $\{I_tH_t\}$ as

$$X_t \stackrel{d}{=} \sum_{j=0}^{\infty} \alpha^j \circ (I_{t-j}H_{t-j}),$$

where "$\stackrel{d}{=}$" means equal in distribution.

Proof. Note that

$$\begin{aligned}
X_t &= \alpha \circ X_{t-1} + I_t H_t \\
&= \alpha \circ (\alpha \circ X_{t-2} + I_{t-1} H_{t-1}) + I_t H_t \\
&= \alpha \circ (\alpha \circ X_{t-2}) + \alpha \circ (I_{t-1} H_{t-1}) + I_t H_t \\
&= \alpha \circ (\alpha \circ (\alpha \circ X_{t-3} + I_{t-2} H_{t-2})) + \alpha \circ (I_{t-1} H_{t-1}) + I_t H_t \\
&= \alpha^3 \circ X_{t-3} + \alpha^2 \circ I_{t-2} H_{t-2} + \alpha \circ I_{t-1} H_{t-1} + I_t H_t. \qquad (3.5) \\
&\stackrel{d}{=} \sum_{j=1}^{\infty} \alpha^j \circ (I_{t-j} H_{t-j}) + I_t H_t \\
&\stackrel{d}{=} \sum_{j=0}^{\infty} \alpha^j \circ (I_{t-j} H_{t-j}),
\end{aligned}$$

where we use (3.4) to obtain (3.5).

□

Theorem 3.1.4. The autocovariance function, γ_k $(k \geq 1)$, of the $BNLINAR(1)$ model $\{X_t\}_{t \geq 1}$ defined in Definition 3.1.1 is given by

$$\gamma_k = Cov(X_t, X_{t-k}) = \alpha^k \gamma_0, \qquad (3.6)$$

where γ_0 is the variance of X_t.

Consequently, the autocorrelation function of order k, ρ_k, of the $BNLINAR(1)$ model is

$$\rho_k = \alpha^k. \qquad (3.7)$$

Proof. From (3.4) and the property that I_tH_t and X_{t-k} are independent, for $k \geq 1$,

$$\begin{aligned}
\gamma_k &= Cov(X_t, X_{t-k}) \\
&= Cov(\alpha \circ X_{t-1} + I_tH_t, X_{t-k}) \\
&= Cov(\alpha \circ X_{t-1}, X_{t-k}) + Cov(I_tH_t, X_{t-k}) \\
&= \alpha Cov(X_{t-1}, X_{t-k}) \\
&= \alpha Cov(\alpha \circ X_{t-2} + I_{t-1}H_{t-1}, X_{t-k}) \\
&= \alpha^{k-1} Cov(\alpha \circ X_{t-k}, X_{t-k}) \qquad (3.8) \\
&= \alpha^k \gamma_0.
\end{aligned}$$

By applying (3.4) recursively to obtain (3.8). Consequently, the correlation function ρ_k can be written as

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \alpha^k.$$

□

Remark 3.1.2. From (3.7), the autocorrelation function declines exponentially as k converges to infinity.

3.1.1 Probabilistic properties of the BNLINAR(1) model

In this section, we investigate many conditional properties such as conditional expectation and conditional variance of the constructed model. Since $\{X_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$. From Theorem 2.1.5, expectation and variance of X_t for the $BNLINAR(1)$ model are given respectively

$$E(X_t) = \frac{2\beta+\theta}{\theta(\beta+\theta)}, \tag{3.9}$$

$$Var(X_t) = \frac{2\beta^2(1+\theta)+\theta^2(1+\theta)+\beta\theta(4+3\theta)}{\theta^2(\beta+\theta)^2}. \tag{3.10}$$

Theorem 3.1.5. The expectation of H_t defined in (3.4) is

$$E(H_t) = E(X_t) = \frac{2\beta+\theta}{\theta(\beta+\theta)}.$$

Proof. From (3.4), since $\{X_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$ and the fact that I_t and H_t are independent,

$$\begin{aligned} E(X_t) &= E(\alpha \circ X_{t-1} + I_t H_t) \\ &= \alpha E(X_{t-1}) + E(I_t)E(H_t) \\ &= \alpha E(X_t) + E(I_t)E(H_t). \end{aligned}$$

Then $(1-\alpha)E(X_t) = E(I_t)E(H_t)$. Since $E(I_t) = 1-\alpha$,

$$E(H_t) = E(X_t).$$

□

Theorem 3.1.6. The expectation of H_t^2 and variance of $I_t H_t$ defined in (3.4) are

1. $E(H_t^2) = \dfrac{\theta^2(2+\theta)+\beta^2(6-2\alpha+2\theta)+\beta\theta(8+3\theta)}{\theta^2(\beta+\theta)^2}$,

2. $Var(I_t H_t) = \dfrac{(1-\alpha)(2\beta^2(1+\alpha+\theta)+\theta^2(1+\alpha+\theta)+\beta\theta(4+4\alpha+3\theta))}{\theta^2(\beta+\theta)^2}$,

respectively.

Proof. From Theorem 2.4.1(4), the stationary property of $\{X_t\}_{t\geq 1}$ and the fact that $\alpha \circ X_{t-1}$ and I_tH_t are independent,

$$\begin{aligned}
Var(X_t) &= Var(\alpha \circ X_{t-1} + I_tH_t) \\
&= Var(\alpha \circ X_{t-1}) + Var(I_tH_t) \\
&= \alpha(1-\alpha)E(X_{t-1}) + \alpha^2 Var(X_{t-1}) + Var(I_tH_t) \\
&= \alpha(1-\alpha)E(X_{t-1}) + \alpha^2 Var(X_t) + E(I_t^2H_t^2) - (E(I_tH_t))^2 \\
&= \alpha(1-\alpha)E(X_{t-1}) + \alpha^2 Var(X_t) + E(I_t^2)E(H_t^2) - (E(I_t)E(H_t))^2 \\
&= \alpha(1-\alpha)E(X_t) + \alpha^2 Var(X_t) + (1-\alpha)E(H_t^2) - (1-\alpha)^2(E(H_t))^2.
\end{aligned} \tag{3.11}$$

From (3.9) – (3.11) and Theorem 3.1.5, we have

$$\begin{aligned}
E(H_t^2) &= \frac{Var(X_t) - \alpha(1-\alpha)E(X_t) - \alpha^2 Var(X_t) + (1-\alpha)^2(E(H_t))^2}{(1-\alpha)} \\
&= \frac{\theta^2(2+\theta) + \beta^2(6-2\alpha+2\theta) + \beta\theta(8+3\theta)}{\theta^2(\beta+\theta)^2}.
\end{aligned} \tag{3.12}$$

Then we consider the variance of I_tH_t and the fact that I_t and H_t are independent,

$$\begin{aligned}
Var(I_tH_t) &= E(I_t^2H_t^2) - (E(I_tH_t))^2 \\
&= E(I_t^2)E(H_t^2) - (E(I_t)E(H_t))^2 \\
&= (1-\alpha)E(H_t^2) - (1-\alpha)^2(E(H_t))^2,
\end{aligned} \tag{3.13}$$

we substitute $E(H_t^2)$ and $E(H_t)$ defined in (3.12) and Theorem 3.1.5 respectively, we have

$$Var(I_tH_t) = \frac{(1-\alpha)}{\theta^2(\beta+\theta)^2}(2\beta^2(1+\alpha+\theta) + \theta^2(1+\alpha+\theta) + \beta\theta(4+4\alpha+3\theta)).$$

□

Theorem 3.1.7. The $(k+1)$-step ahead conditional expectation of the $BNLINAR(1)$ model is

$$E(X_{t+k}|X_{t-1}=x)=\alpha^{k+1}x+(1-\alpha^{k+1})\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right),$$

for $x\in\{0,1,2,\ldots\}$.

Proof.

$$\begin{aligned}&E(X_{t+k}|X_{t-1}=x)\\&=E(\alpha\circ X_{t+k-1}+I_{t+k}H_{t+k}|X_{t-1}=x)\\&=E(\alpha\circ(\alpha\circ X_{t+k-2}+I_{t+k-1}H_{t+k-1})+I_{t+k}H_{t+k}|X_{t-1}=x),\end{aligned}$$

by using (3.4) to obtain the last equality. Applying (3.4) to $\{X_t\}_{t\geq 1}$ recursively,

$$\begin{aligned}&E(X_{t+k}|X_{t-1}=x)\\&=E(\alpha^{k+1}\circ X_{t-1}+\alpha^{k}\circ I_tH_t+\alpha^{k-1}\circ I_{t+1}H_{t+1}+\cdots+I_{t+k}H_{t+k}|X_{t-1}=x)\\&=E(\alpha^{k+1}\circ X_{t-1}|X_{t-1}=x)+\sum_{h=0}^{k}E(\alpha^{h}\circ I_{t+k-h}H_{t+k-h}|X_{t-1}=x)\\&=\alpha^{k+1}x+\sum_{h=0}^{k}\alpha^{h}E(I_{t+k-h}H_{t+k-h}) \qquad (3.14)\\&=\alpha^{k+1}x+\left(\frac{1-\alpha^{k+1}}{1-\alpha}\right)E(I_tH_t),\end{aligned}$$

where we use Theorem 2.4.1(3) to obtain (3.14).

Then, we substitute $E(H_t)$ defined in Theorem 3.1.5.

$$E(X_{t+k}|X_{t-1}=x)=\alpha^{k+1}x+(1-\alpha^{k+1})\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right).$$

□

Remark 3.1.3. The conditional expectation $E(X_{t+k}|X_{t-1}=x)$ converges to the unconditional expectation $\dfrac{2\beta+\theta}{\theta(\beta+\theta)}$ as k converges to infinity.

Proof. Since $0<\alpha<1$,

$$\begin{aligned}\lim_{k\to\infty} E(X_{t+k}|X_{t-1}=x) &= \lim_{k\to\infty}\left(\alpha^{k+1}x+(1-\alpha^{k+1})\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right)\right)\\ &= \frac{2\beta+\theta}{\theta(\beta+\theta)}.\end{aligned}$$

□

Theorem 3.1.8. The $(k+1)$-step ahead conditional variance for the $BNLINAR(1)$ model is

$$\begin{aligned}Var(X_{t+k}|X_{t-1}=x) &= \alpha^{k+1}(1-\alpha^{k+1})x+\frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(I_tH_t)\\ &\quad+\frac{(1-\alpha^k)(\alpha-\alpha^{k+2})}{1-\alpha^2}E(I_tH_t),\end{aligned} \tag{3.15}$$

for $x\in\{0,1,2,\ldots\}$.

Proof. From (3.4),

$$\begin{aligned}&Var(X_{t+k}|X_{t-1}=x)\\ &= Var(\alpha\circ X_{t+k-1}+I_{t+k}H_{t+k}|X_{t-1}=x)\\ &= Var(\alpha\circ(\alpha\circ X_{t+k-2}+I_{t+k-1}H_{t+k-1})+I_{t+k}H_{t+k}|X_{t-1}=x)\\ &= Var(\alpha^{k+1}\circ X_{t-1}+\alpha^k\circ I_tH_t+\alpha^{k-1}\circ I_{t+1}H_{t+1}+\cdots+I_{t+k}H_{t+k}|X_{t-1}=x)\end{aligned} \tag{3.16}$$

$$
\begin{aligned}
&= Var(\alpha^{k+1} \circ X_{t-1}) + \sum_{h=0}^{k} Var(\alpha^h \circ I_{t+k}H_{t+k} | X_{t-1} = x) \\
&= \alpha^{k+1}(1-\alpha^{k+1})x + \sum_{h=0}^{k} Var(\alpha^h \circ I_{t+k}H_{t+k}) \qquad (3.17) \\
&= \alpha^{k+1}(1-\alpha^{k+1})x + \sum_{h=0}^{k}(\alpha^h(1-\alpha^h)E(I_{t+k}H_{t+k} + \alpha^{2h})Var(I_{t+k}H_{t+k})) \\
&= \alpha^{k+1}(1-\alpha^{k+1})x + Var(I_{t+k}H_{t+k})\sum_{h=0}^{k}\alpha^{2h} + E(I_{t+k}H_{t+k})\sum_{h=0}^{k}\alpha^h(1-\alpha^h) \\
&= \alpha^{k+1}(1-\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(I_tH_t) + \frac{(1-\alpha^k)(\alpha-\alpha^{k+2})}{1-\alpha^2}E(I_tH_t),
\end{aligned}
$$

where we use (3.4) to obtain (3.16) and Theorem 2.4.1(4) to obtain (3.17). □

Remark 3.1.4. The conditional variance $Var(X_{t+k}|X_{t-1} = x)$ converges to the unconditional variance $\dfrac{2\beta^2(1+\theta)+\theta^2(1+\theta)+\beta\theta(4+3\theta)}{\theta^2(\beta+\theta)^2}$ as k converges to infinity.

Proof. Since $0 < \alpha < 1$,

$$
\begin{aligned}
&\lim_{k\to\infty} Var(X_{t+k}|X_{t-1} = x) \\
&= \lim_{k\to\infty}\left[\alpha^{k+1}(1-\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(I_tH_t) + \frac{(1-\alpha^k)(\alpha-\alpha^{k+2})}{1-\alpha^2}E(I_tH_t)\right] \\
&= \frac{Var(I_tH_t)}{1-\alpha^2} + \frac{\alpha E(I_tH_t)}{1-\alpha^2} \\
&= \frac{2\beta^2(1+\theta)+\theta^2(1+\theta)+\beta\theta(4+3\theta)}{\theta^2(\beta+\theta)^2}.
\end{aligned}
$$

□

Theorem 3.1.9. The partial autocorrelation function of the $BNLINAR(1)$ model at lag h is 0 where $h > 1$.

Proof. For $h > 1$,

$$\begin{aligned}
\beta(h) &= Corr(X_{h+1} - E(X_{h+1}|X_2, \ldots, X_h), X_1) \\
&= Corr(X_{h+1} - \alpha X_h + (1-\alpha)\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right), X_1) \quad \text{by using Theorem 3.1.7} \\
&= Corr(X_{h+1}, X_1) - Corr(\alpha X_h, X_1) + Corr((1-\alpha)\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right), X_1) \\
&= \alpha^h - \alpha(\alpha^{h-1}) \\
&= 0.
\end{aligned}$$

□

Theorem 3.1.10. The Markov process with transition probabilities of the $BNLINAR(1)$ model is

$$p_{lk} = \sum_{m=0}^{\min(l,k)} \binom{l}{m} \alpha^m (1-\alpha)^{l-m} P(I_t H_t = k - m)$$

where the process $I_t H_t$ is defined in (3.4).

Proof.

$$\begin{aligned}
p_{lk} &= P(X_t = k | X_{t-1} = l) \\
&= P(\alpha \circ X_{t-1} + I_t H_t = k | X_{t-1} = l) \\
&= \sum_{m=0}^{\min(l,k)} P(\alpha \circ X_{t-1} = m | X_{t-1} = l) P(I_t H_t = k - m | X_{t-1} = l) \\
&= \sum_{m=0}^{\min(l,k)} P(\alpha \circ X_{t-1} = m | X_{t-1} = l) P(I_t H_t = k - m) \\
&= \sum_{m=0}^{\min(l,k)} \binom{l}{m} \alpha^m (1-\alpha)^{l-m} P(I_t H_t = k - m),
\end{aligned}$$

by using Definition 2.4.1 and the process $I_t H_t$ is defined in (3.4). □

3.1.2 Estimation and inference of the BNLINAR(1) model

In this section, we consider parameter estimation methods of the unknown parameters by (1) the conditional least squares estimator (CLS) and (2) the Yule-Walker estimator (YW). These estimators are compared via Monte Carlo simulations in terms of their means and variances by using the statistical software R [11].

3.1.2.1 Conditional least squares estimation

The conditional least squares estimators of the parameters α and μ of the $BNLINAR(1)$ model are obtained by minimizing the function defined in Definition 2.5.1. Let $k = 0$ in the expression in Theorem 3.1.7, the conditional expectation is

$$E(X_t|X_{t-1}) = \alpha X_{t-1} + \mu(1-\alpha),$$

where $\mu = E(X_t)$. Then

$$Q_n = \sum_{t=2}^{n}(X_t - E(X_t|X_{t-1}))^2 = \sum_{t=2}^{n}(X_t - \alpha X_{t-1} - \mu(1-\alpha))^2.$$

Equating the first order partial derivatives of Q_n with respect to μ and α to zero, we have

$$\frac{\partial Q_n}{\partial \mu}\Big|_{\mu=\hat{\mu},\alpha=\hat{\alpha}} = -\sum_{t=2}^{n} 2(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))(1-\hat{\alpha}) = 0. \tag{3.18}$$

$$\frac{\partial Q_n}{\partial \hat{\alpha}}\Big|_{\mu=\hat{\mu},\alpha=\hat{\alpha}} = \sum_{t=2}^{n} 2(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))(\hat{\mu} - X_{t-1}) = 0. \tag{3.19}$$

From (3.18),

$$\sum_{t=2}^{n} X_t - \hat{\alpha}\sum_{t=2}^{n} X_{t-1} - \hat{\mu}(n-1)(1-\hat{\alpha}) = 0. \tag{3.20}$$

By solving equation (3.20), the estimation of μ can be computed as

$$\hat{\mu} = \frac{\sum_{t=2}^{n} X_t - \hat{\alpha} \sum_{t=2}^{n} X_{t-1}}{(n-1)(1-\hat{\alpha})}. \tag{3.21}$$

From (3.19),

$$\begin{aligned} 0 = \hat{\mu} \sum_{t=2}^{n} X_t - \hat{\alpha}\hat{\mu} \sum_{t=2}^{n} X_{t-1} - \hat{\mu}^2(1-\hat{\alpha})(n-1) - \sum_{t=2}^{n} X_{t-1}X_t + \hat{\alpha} \sum_{t=2}^{n} X_{t-1}^2 \\ + (1-\hat{\alpha})\hat{\mu} \sum_{t=2}^{n} X_{t-1}. \end{aligned} \tag{3.22}$$

By solving equation (3.22) and substitute $\hat{\mu}$ in equation (3.21), the estimation of α can be computed as

$$\hat{\alpha}_{CLS} = \frac{(n-1) \sum_{t=2}^{n} X_{t-1}X_t - \sum_{t=2}^{n} X_t \sum_{t=2}^{n} X_{t-1}}{(n-1) \sum_{t=2}^{n} X_{t-1}^2 - \left(\sum_{t=2}^{n} X_{t-1} \right)^2}.$$

From (3.21) and (3.9), we have

$$\frac{2\hat{\beta}_{CLS} + \hat{\theta}_{CLS}}{\hat{\theta}_{CLS}(\hat{\theta}_{CLS} + \hat{\beta}_{CLS})} = \hat{\mu}_{CLS} = \frac{\sum_{t=2}^{n} X_t - \hat{\alpha}_{CLS} \sum_{t=2}^{n} X_{t-1}}{(n-1)(1-\hat{\alpha}_{CLS})}. \tag{3.23}$$

The conditional least squares estimator of the parameters σ^2 is obtained by minimizing the function defined in Abdulhamid et al. [2]. First, substitute $k = 0$ into (3.15),

$$Var(X_t | X_{t-1}) = \alpha(1-\alpha)X_{t-1} + Var(I_t H_t). \tag{3.24}$$

Substitute $Var(I_tH_t)$ from (3.13) into (3.24), the conditional variance is

$$Var(X_t|X_{t-1}) = \alpha(1-\alpha)X_{t-1} + (1-\alpha^2)\sigma^2 - \alpha(1-\alpha)\mu. \qquad (3.25)$$

To obtain $\hat{\sigma}^2$, we follow Abdulhamid et al. [2] by minimizing the function S_n defined as

$$\begin{aligned} S_n &= \sum_{t=2}^{n}[(X_t - E(X_t|X_{t-1}))^2 - Var(X_t|X_{t-1})]^2 \\ &= \sum_{t=2}^{n}[(X_t - \alpha X_{t-1} - \mu(1-\alpha))^2 - \alpha(1-\alpha)X_{t-1} - (1-\alpha^2)\sigma^2 + \alpha(1-\alpha)\mu]^2. \end{aligned}$$

Taking the first order partial derivative of S_n with respect to σ^2 and equating it to zero, we get

$$\begin{aligned} 0 &= \frac{\partial S_n}{\partial \sigma^2}\Big|_{\sigma^2=\hat{\sigma}^2,\mu=\hat{\mu},\alpha=\hat{\alpha}} \\ &= \sum_{t=2}^{n} 2[(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1-\hat{\alpha})X_{t-1} - (1-\hat{\alpha}^2)\hat{\sigma}^2 + \hat{\alpha}(1-\hat{\alpha})\hat{\mu}](\hat{\alpha}^2 - 1). \end{aligned}$$

Then

$$\sum_{t=2}^{n}[(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1-\hat{\alpha})X_{t-1} - (1-\hat{\alpha}^2)\hat{\sigma}^2 + \hat{\alpha}(1-\hat{\alpha})\hat{\mu}] = 0. \qquad (3.26)$$

By solving the equation (3.26), the estimation of σ^2 can be obtained as

$$\hat{\sigma}^2 = \frac{\sum_{t=2}^{n}[(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1-\hat{\alpha})X_{t-1} + \hat{\alpha}(1-\hat{\alpha})\hat{\mu}]}{(1-\hat{\alpha}^2)(n-1)}. \qquad (3.27)$$

From (3.10) and (3.27),

$$\begin{aligned} &\frac{2\hat{\beta}^2(1+\hat{\theta}) + \hat{\theta}^2(1+\hat{\alpha}) + \hat{\beta}\hat{\theta}(4+3\hat{\theta})}{\hat{\theta}^2(\hat{\beta}+\hat{\theta})^2} \\ &= \hat{\sigma}^2_{CLS} = \frac{\sum_{t=2}^{n}[(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1-\hat{\alpha})X_{t-1} + \hat{\alpha}(1-\hat{\alpha})\hat{\mu}]}{(1-\hat{\alpha}^2)(n-1)}. \end{aligned}$$

3.1.2.2 The Yule-Walker estimation

In this part, the Yule-Walker estimation for α, μ and σ^2 are obtained. By using Definition 2.3.3 then the sample autocovariance function of X_t

$$\hat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(X_t - \bar{X})(X_{t+k} - \bar{X}), \tag{3.28}$$

where $0 \leq k < n$ and $\bar{X} = \dfrac{1}{n}\displaystyle\sum_{t=1}^{n} X_t$ is the sample mean.
From the Yule-Walker equation defined in Definition 2.5.2 and equation (3.28), the Yule-Walker estimator of α is

$$\hat{\alpha}_{YW} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\displaystyle\sum_{t=2}^{n}(X_t - \bar{X})(X_{t-1} - \bar{X})}{\displaystyle\sum_{t=1}^{n}(X_t - \bar{X})^2}.$$

Consider $\mu = E(X_t)$ defined in (3.9) and $\sigma^2 = Var(X_t)$ defined in (3.10) and note that $S^2 = \dfrac{\displaystyle\sum_{t=1}^{n}(X_t - \bar{X})}{n-1}$. The Yule-Walker estimators of μ and σ^2 are

$$\hat{\mu}_{YW} = \bar{X} = \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})},$$

$$\hat{\sigma^2}_{YW} = S^2 = \frac{2\hat{\beta}^2(1 + \hat{\theta}) + \hat{\theta}^2(1 + \hat{\alpha}) + \hat{\beta}\hat{\theta}(4 + 3\hat{\theta})}{\hat{\theta}^2(\hat{\beta} + \hat{\theta})^2},$$

respectively.

3.1.3 Simulation Results

In this section, we produce 10,000 samples from the $BNLINAR(1)$ model for true parameter values in different settings (1) $\alpha = 0.1, \beta = 1, \theta = 1$; (2) $\alpha = 0.3, \beta = 2, \theta = 2$; (3) $\alpha = 0.5, \beta = 3, \theta = 3$ of different sample sizes $n = 50, 100, 500, 1000, 5000$ and 10000 by using the statistical software R and obtain estimators of parameters from two methods described in Section 3.1.2. Then we compare the obtained estimators in terms of their means and variances. Table 3.1 shows mean and variance (in brackets) of the estimators for different values of the parameters α, μ and σ^2.

From Table 3.1, we observe that the estimators obtained from the two estimation methods converge to the true parameters. In addition, increasing the sample size yields smaller variance. The conditional least squares estimate (CLS) and the Yule-Walker estimate (YW) are approximately the same. Considering the variance we can see that the CLS estimators have smaller variance than the YW estimators for parameters α and μ. However, the YW has smaller variance than the CLS for the parameter σ^2. Considering the mean, we can see that the CLS estimators converge to the true parameter faster than the YW estimators for parameters α and μ.

3.1.4 Real data

In this section, we apply the two models with two real data sets : (1) the numbers of Skin-lesions and (2) the numbers of Anorexias.

3.1.4.1 The numbers of Skin-lesions

The first example considers the numbers of Skin-lesions monthly from January 2003 to December 2009 from a region in New Zealand. The data was original introduced in Aghababaei et al. [1]. Sample mean and variance are 1.43 and 3.36, respectively. The fitted $BNLINAR(1)$ model is

$$X_t = 0.2365 \circ X_{t-1} + I_t H_t,$$

The predicted values of the numbers of Skin-lesions series are given by

$$\hat{X}_1 = \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})} = 1.4142,$$
$$\hat{X}_t = \hat{\alpha}\hat{X}_{t-1} + (1 - \hat{\alpha})\frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})}.$$

$(\alpha, \theta, \beta, \mu, \sigma^2) = (0.1, 1, 1, 1.5, 3.25)$						
n	$\hat{\alpha}_{CLS}$	$\hat{\mu}_{CLS}$	$\hat{\sigma}^2_{CLS}$	$\hat{\alpha}_{YW}$	$\hat{\mu}_{YW}$	$\hat{\sigma}^2_{YW}$
50	0.0838 (0.0193)	1.4796 (0.0742)	3.1156 (1.3730)	0.0707 (0.0193)	1.4730 (0.0766)	3.1751 (1.3920)
100	0.0943 (0.0101)	1.4855 (0.0372)	3.1740 (0.7122)	0.0848 (0.0101)	1.4854 (0.0382)	3.2074 (0.7310)
500	0.0982 (0.0021)	1.4969 (0.0070)	3.2364 (0.1418)	0.0970 (0.0021)	1.4964 (0.0079)	3.2386 (0.1454)
1000	0.0990 (0.0011)	1.4998 (0.0032)	3.2453 (0.0723)	0.0982 (0.0011)	1.4983 (0.0040)	3.2455 (0.0717)
5000	0.0999 (0.0002)	1.4996 (0.0007)	3.2493 (0.0154)	0.0998 (0.0002)	1.4996 (0.0008)	3.2497 (0.0144)
10000	0.0999 (0.0001)	1.4999 (0.0003)	3.2495 (0.0077)	0.0999 (0.0001)	1.4999 (0.0004)	3.2497 (0.0073)
$(\alpha, \theta, \beta, \mu, \sigma^2) = (0.3, 2, 2, 0.75, 1.1875)$						
n	$\hat{\alpha}_{CLS}$	$\hat{\mu}_{CLS}$	$\hat{\sigma}^2_{CLS}$	$\hat{\alpha}_{YW}$	$\hat{\mu}_{YW}$	$\hat{\sigma}^2_{YW}$
50	0.2649 (0.0223)	0.7035 (0.0413)	1.0840 (0.2564)	0.2515 (0.0218)	0.7018 (0.0412)	1.1044 (0.2535)
100	0.2815 (0.0120)	0.7094 (0.0214)	1.1120 (0.1271)	0.2758 (0.0117)	0.7092 (0.0202)	1.1182 (0.1228)
500	0.2956 (0.0026)	0.7157 (0.0042)	1.1371 (0.0267)	0.2948 (0.0026)	0.7143 (0.0043)	1.1372 (0.0268)
1000	0.2988 (0.0013)	0.7270 (0.0022)	1.1416 (0.0136)	0.2971 (0.0013)	0.7166 (0.0021)	1.1430 (0.0137)
5000	0.2999 (0.0003)	0.7366 (0.0004)	1.1435 (0.0032)	0.2997 (0.0003)	0.7166 (0.0004)	1.1445 (0.0027)
10000	0.3000 (0.0000)	0.7399 (0.0002)	1.1698 (0.0015)	0.2999 (0.0001)	0.7378 (0.0002)	1.1699 (0.0013)
$(\alpha, \theta, \beta, \mu, \sigma^2) = (0.5, 3, 3, 0.5, 0.6944)$						
n	$\hat{\alpha}_{CLS}$	$\hat{\mu}_{CLS}$	$\hat{\sigma}^2_{CLS}$	$\hat{\alpha}_{YW}$	$\hat{\mu}_{YW}$	$\hat{\sigma}^2_{YW}$
50	0.4363 (0.0266)	0.4396 (0.0496)	0.5286 (0.6079)	0.4221 (0.0258)	0.4328 (0.0351)	0.5801 (0.1272)
100	0.4680 (0.0130)	0.4407 (0.0180)	0.5970 (0.0647)	0.4604 (0.0135)	0.4403 (0.0184)	0.6024 (0.0671)
500	0.4925 (0.0028)	0.4453 (0.0037)	0.6161 (0.0140)	0.4911 (0.0028)	0.4442 (0.0037)	0.6166 (0.0139)
1000	0.4967 (0.0015)	0.4459 (0.0018)	0.6207 (0.0072)	0.4959 (0.0014)	0.4460 (0.0019)	0.6217 (0.0072)
5000	0.4993 (0.0003)	0.4562 (0.0004)	0.6215 (0.0019)	0.4992 (0.0003)	0.4462 (0.0004)	0.6229 (0.0014)
10000	0.5000 (0.0001)	0.4772 (0.0001)	0.6458 (0.0016)	0.5000 (0.0001)	0.4729 (0.0001)	0.6464 (0.0007)

Table 3.1: Mean and variance (in brackets) of the estimators for different values of the parameters α, μ and σ^2 for the $BNLINAR(1)$ model

Substituting parameter estimates $\hat{\alpha} = 0.2365$ and $\frac{2\hat{\beta}+\hat{\theta}}{\hat{\theta}(\hat{\beta}+\hat{\theta})} = 1.4142$.

$$\hat{X}_t = 0.2365\hat{X}_{t-1} + 1.0797, \quad t = 2, 3, \ldots, 72.$$

The expectation and variance computed from the $BNLINAR(1)$ model are 1.414 and 3.356, respectively. We can see that the model can capture the sample mean and variance of the data set. Therefore, the model is reasonable to this data set.

3.1.4.2 The numbers of Anorexias

The second example considers the numbers of Anorexias monthly from January 2003 to December 2009 from a region in New Zealand. Sample mean and variance are 0.82 and 2.90, respectively. The fitted $BNLINAR(1)$ is

$$X_t = 0.4909 \circ X_{t-1} + I_t H_t,$$

The predicted values of the numbers of Anorexias series are given by

$$\hat{X}_1 = \frac{2\beta+\theta}{\theta(\beta+\theta)} = 0.8313,$$
$$\hat{X}_t = \hat{\alpha}\hat{X}_{t-1} + (1-\hat{\alpha})\frac{2\beta+\theta}{\theta(\beta+\theta)}.$$

Substituting parameter estimates $\hat{\alpha} = 0.4909$ and $\frac{2\hat{\beta}+\hat{\theta}}{\hat{\theta}(\hat{\beta}+\hat{\theta})} = 0.8313$.

$$\hat{X}_t = 0.4909\hat{X}_{t-1} + 0.4232, \quad t = 2, 3, \ldots, 72.$$

The expectation and variance computed from the $BNLINAR(1)$ model are 0.831 and 2.900, respectively. We can see that the model can capture the sample mean and variance of the data set. Therefore, the model is reasonable to this data set.

3.2 Construction of the first order integer-valued autoregressive models with the two-parameter generalized Poisson-Lindley distribution based on the negative binomial thinning operator model (NNLINAR(1))

In this section, we construct the first order integer-valued autoregressive models with two-parameter generalized Poisson-Lindley distribution based on the negative binomial thinning operator model (NNLINAR(1)). Moreover, we investigate many properties of the constructed model such as moments, parameter estimations and perform some numerical studies.

Definition 3.2.1. The first order integer-valued autoregressive model with two-parameter generalized Poisson-Lindley distribution based on the negative binomial thinning operator (NNLINAR(1)) $\{X_t\}_{t\geq 1}$ is defined as

$$X_t = \alpha * X_{t-1} + \varepsilon_t, \tag{3.29}$$

where the negative the binomial thinning $\alpha*$ defined in Definition 2.4.2, $\{X_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$ distribution and $\{\varepsilon_t\}_{t\geq 1}$ is a sequence of i.i.d. random variables such that $\alpha * X_{t-1}$ and ε_t are independent.

Theorem 3.2.1. The innovation process $\{\varepsilon_t\}_{t\geq 1}$ has the probability generating function

$$\Phi_{\varepsilon_t}(s) = \frac{(\beta + \theta - s + 1)(\theta + \alpha\theta - \alpha\theta s + \alpha - \alpha s)^2}{(\theta - s + 1)^2(1 + \alpha - \alpha s)(\theta + \beta + \alpha\theta + \alpha\beta - \alpha(\beta + \theta + 1)s)}, \tag{3.30}$$

for $s \in \mathbb{R}$.

Proof. Since $\{X_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$, from Theorem 2.4.2(5),

$$\Phi_{\alpha * X_{t-1}}(s) = E(s^{(\alpha * X_{t-1})}) = \Phi_{X_t}((1 + \alpha - \alpha s)^{-1}).$$

From (3.29) and the property that $\alpha * X_{t-1}$ and ε_t are independent, for $s \in \mathbb{R}$,

$$\begin{aligned}
\Phi_{X_t}(s) &= E(s^{X_t}) \\
&= E(s^{\alpha * X_{t-1}+\varepsilon_t}) \\
&= E(s^{\alpha * X_{t-1}} s^{\varepsilon_t}) \\
&= E(s^{(\alpha * X_{t-1})}) E(s^{\varepsilon_t}) \\
&= \Phi_{\alpha * X_{t-1}}(s) \Phi_{\varepsilon_t}(s) \\
&= \Phi_{X_t}((1+\alpha-\alpha s)^{-1}) \Phi_{\varepsilon_t}(s).
\end{aligned}$$

From Theorem 2.1.5(3), the innovation process $\{\varepsilon_t\}_{t\geq 1}$ has the probability generating function

$$\begin{aligned}
\Phi_{\varepsilon_t}(s) &= \frac{\Phi_{X_t}(s)}{\Phi_{X_t}((1+\alpha-\alpha s)^{-1})} \\
&= \frac{\theta^2(\beta+\theta-s+1)}{(\beta+\theta)(\theta-s+1)^2} \frac{(\beta+\theta)(\theta-(1+\alpha-\alpha s)^{-1}+1)^2}{\theta^2(\theta-(1+\alpha-\alpha s)^{-1}+1+\beta)} \\
&= \frac{\beta+\theta-s+1}{(\theta-s+1)^2} \frac{(\theta+\alpha\theta-\alpha\theta s-1+1+\alpha-\alpha s)^2}{(1+\alpha-\alpha s)(\beta+\alpha\beta-\alpha\beta s+\theta+\theta\alpha-\alpha\theta s-1+1+\alpha-\alpha s)} \\
&= \frac{(\beta+\theta-s+1)(\theta+\alpha\theta-\alpha\theta s+\alpha-\alpha s)^2}{(\theta-s+1)^2(1+\alpha-\alpha s)(\theta+\beta+\alpha+\alpha\theta+\alpha\beta-\alpha(\beta+\theta+1)s)}.
\end{aligned}$$

□

Theorem 3.2.2. The autocovariance function, γ_k $(k \geq 1)$, of the $NNLINAR(1)$ model $\{X_t\}_{t\geq 1}$ defined in Definition 3.2.1 is given by

$$\gamma_k = Cov(X_t, X_{t-k}) = \alpha^k \gamma_0, \qquad (3.31)$$

where γ_0 is the variance of X_t.

Consequently, the autocorrelation function of order k, ρ_k, of the $NNLINAR(1)$ model is

$$\rho_k = \alpha^k. \qquad (3.32)$$

Proof. From (3.29) and the property that ε_t and X_{t-k} are independent, for $k \geq 1$,

$$\begin{aligned}
\gamma_k &= Cov(X_t, X_{t-k}) \\
&= Cov(\alpha * X_{t-1} + \varepsilon_t, X_{t-k}) \\
&= Cov(\alpha * X_{t-1}, X_{t-k}) + Cov(\varepsilon_t, X_{t-k}) \\
&= \alpha Cov(X_{t-1}, X_{t-k}) \\
&= \alpha Cov(\alpha * X_{t-2} + \varepsilon_{t-1}, X_{t-k}) \\
&= \alpha^{k-1} Cov(\alpha * X_{t-k}, X_{t-k}) \qquad (3.33) \\
&= \alpha^k \gamma_0.
\end{aligned}$$

By applying (3.29) recursively to obtain (3.33). Consequently, the correlation function ρ_k can be written as

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \alpha^k.$$

□

Remark 3.2.1. From (3.32), the autocorrelation function declines exponentially as k converges to infinity.

3.2.1 Probabilistic properties of the NNLINAR(1) model

In this section, we investigate many conditional properties such as conditional mean and conditional variance of the constructed model. Since $\{X_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$. From Theorem 2.1.5, mean and variance of X_t for the $NNLINAR(1)$ model are given respectively by

$$E(X_t) = \frac{2\beta + \theta}{\theta(\beta + \theta)}, \qquad (3.34)$$

$$Var(X_t) = \frac{2\beta^2(1+\theta) + \theta^2(1+\theta) + \beta\theta(4+3\theta)}{\theta^2(\beta+\theta)^2}. \qquad (3.35)$$

Theorem 3.2.3. The expectation of ε_t defined in Definition 3.2.1 is

$$E(\varepsilon_t) = (1-\alpha)E(X_t),$$

where $E(X_t)$ is defined in (3.34).

Proof. From (3.29), since $\{X_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$,

$$\begin{aligned} E(X_t) &= E(\alpha * X_{t-1} + \varepsilon_t) \\ &= \alpha E(X_t) + E(\varepsilon_t). \end{aligned}$$

Then $\quad E(\varepsilon_t) = (1-\alpha)E(X_t)$. □

Theorem 3.2.4. The variance of ε_t defined in (3.29) is

$$Var(\varepsilon_t) = (1-\alpha^2)Var(X_t) - \alpha(1+\alpha)E(X_t),$$

where $E(X_t)$ and $Var(X_t)$ are defined in (3.34) and (3.35), respectively.

Proof. From (3.29), since $\{X_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$,

$$\begin{aligned} Var(X_t) &= Var(\alpha * X_{t-1} + \varepsilon_t) \\ &= Var(\alpha * X_{t-1}) + Var(\varepsilon_t) \\ &= \alpha(1+\alpha)E(X_{t-1}) + \alpha^2 Var(X_{t-1}) + Var(\varepsilon_t). \end{aligned}$$

Then, $\quad (1-\alpha^2)Var(X_t) = \alpha(1+\alpha)E(X_t) + Var(\varepsilon_t)$.

Consequently, $\quad Var(\varepsilon_t) = (1-\alpha^2)Var(X_t) - \alpha(1+\alpha)E(X_t)$. □

Theorem 3.2.5. The $(k+1)$-step ahead conditional expectation of the $NNLINAR(1)$ model is

$$E(X_{t+k}|X_{t-1}=x) = \alpha^{k+1}x + (1-\alpha^{k+1})\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right),$$

for $x \in \{0,1,2,\ldots\}$.

Proof.

$$\begin{aligned} E(X_{t+k}|X_{t-1}=x) &= E(\alpha * X_{t+k-1} + \varepsilon_{t+k}|X_{t-1}=x) \\ &= E(\alpha * (\alpha * X_{t+k-2} + \varepsilon_{t+k-1}) + \varepsilon_{t+k}|X_{t-1}=x), \end{aligned}$$

by using (3.29) to obtain the last equality. Applying (3.29) to $\{X_t\}_{t\geq 1}$ recursively,

$$\begin{aligned} E(X_{t+k}|X_{t-1}=x) &= E(\alpha^{k+1} * X_{t-1} + \alpha^{k} * \varepsilon_t + \alpha^{k-1} * \varepsilon_{t+1} + \cdots + \varepsilon_{t+k}|X_{t-1}=x) \\ &= E(\alpha^{k+1} * X_{t-1}|X_{t-1}=x) + \sum_{h=0}^{k} E(\alpha^{h} * \varepsilon_{t+k-h}|X_{t-1}=x) \\ &= \alpha^{k+1}x + \sum_{h=0}^{k} \alpha^{h} E(\varepsilon_{t+k-h}) \qquad (3.36) \\ &= \alpha^{k+1}x + \left(\frac{1-\alpha^{k+1}}{1-\alpha}\right) E(\varepsilon_t) \\ &= \alpha^{k+1}x + (1-\alpha^{k+1})\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right), \end{aligned}$$

where we use Theorem 2.4.2(3) to obtain (3.36). □

Remark 3.2.2. The conditional expectation $E(X_{t+k}|X_{t-1}=x)$ converges to the unconditional expectation $\dfrac{2\beta+\theta}{\theta(\beta+\theta)}$ as k converges to infinity.

Proof. Since $0 < \alpha < 1$,

$$\begin{aligned}\lim_{k\to\infty} E(X_{t+k}|X_{t-1}=x) &= \lim_{k\to\infty}\left(\alpha^{k+1}x + (1-\alpha^{k+1})\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right)\right) \\ &= \frac{2\beta+\theta}{\theta(\beta+\theta)}.\end{aligned}$$

□

Theorem 3.2.6. The $(k+1)$-step ahead conditional variance for the $NNLINAR(1)$ model is

$$\begin{aligned}Var(X_{t+k}|X_{t-1}=x) =&\alpha^{k+1}(1+\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(\varepsilon_t) \\ &+ \frac{\alpha(1-\alpha^k)(1-\alpha^{k+1})}{1-\alpha^2}E(\varepsilon_t),\end{aligned} \tag{3.37}$$

for $x \in \{0, 1, 2, \ldots\}$.

Proof. From (3.29),

$$\begin{aligned}&Var(X_{t+k}|X_{t-1}=x) \\ &= Var(\alpha * X_{t+k-1} + \varepsilon_{t+k}|X_{t-1}=x) \\ &= Var(\alpha * (\alpha * X_{t+k-2} + \varepsilon_{t+k-1}) + \varepsilon_{t+k}|X_{t-1}=x) \qquad (3.38) \\ &= Var(\alpha^{k+1} * X_{t-1} + \alpha^k * \varepsilon_t + \alpha^{k-1} * \varepsilon_{t+1} + \cdots + \varepsilon_{t+k}|X_{t-1}=x) \\ &= Var(\alpha^{k+1} * X_{t-1}) + \sum_{h=0}^{k} Var(\alpha^h * \varepsilon_{t+k}|X_{t-1}=x) \\ &= \alpha^{k+1}(1+\alpha^{k+1})x + \sum_{h=0}^{k} Var(\alpha^h * \varepsilon_{t+k-h}) \qquad (3.39) \\ &= \alpha^{k+1}(1+\alpha^{k+1})x + \sum_{h=0}^{k}(\alpha^h(1+\alpha^h)E(\varepsilon_{t+k-h}) + \alpha^{2h}Var(\varepsilon_{t+k-h})) \\ &= \alpha^{k+1}(1+\alpha^{k+1})x + Var(\varepsilon_t)\sum_{h=0}^{k}\alpha^{2h} + E(\varepsilon_t)\sum_{h=0}^{k}\alpha^h(1-\alpha^h) \\ &= \alpha^{k+1}(1+\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(\varepsilon_t) + \frac{\alpha(1-\alpha^k)(1-\alpha^{k+1})}{1-\alpha^2}E(\varepsilon_t),\end{aligned}$$

where we use (3.29) to obtain (3.38) and Theorem 2.4.2(4) to obtain (3.39). □

Remark 3.2.3. The conditional variance $Var(X_{t+k}|X_{t-1} = x)$ converges to the unconditional variance $\frac{2\beta^2(1+\theta)+\theta^2(1+\theta)+\beta\theta(4+3\theta)}{\theta^2(\beta+\theta)^2}$ as k converges to infinity.

Proof. Since $0 < \alpha < 1$,

$$\begin{aligned}
&\lim_{k\to\infty} Var(X_{t+k}|X_{t-1} = x) \\
&= \lim_{k\to\infty}\left[\alpha^{k+1}(1+\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(\varepsilon_t) + \frac{\alpha(1-\alpha^k)(1-\alpha^{k+1})}{1-\alpha^2}E(\varepsilon_t)\right] \\
&= \frac{Var(\varepsilon_t)}{1-\alpha^2} + \frac{\alpha E(\varepsilon_t)}{1-\alpha^2} \\
&= \frac{2\beta^2(1+\theta)+\theta^2(1+\theta)+\beta\theta(4+3\theta)}{\theta^2(\beta+\theta)^2}.
\end{aligned}$$

□

Theorem 3.2.7. The Markov process with transition probabilities of the $NNLINAR(1)$ model is

$$p_{lk} = \sum_{m=0}^{k}\binom{l+m-1}{m}\left(\frac{1}{1+\alpha}\right)^l\left(\frac{\alpha}{1+\alpha}\right)^m P(\varepsilon_t = k-m)\mathbf{I}(l\neq 0) + P(\varepsilon_t = k)\mathbf{I}(l=0), \tag{3.40}$$

where the process ε_t is defined in (3.29).

Proof.

$$\begin{aligned}
p_{lk} &= P(X_t = k|X_{t-1} = l) \\
&= P(\alpha * X_{t-1} + \varepsilon_t = k|X_{t-1} = l). \\
&= \sum_{m=0}^{k} P(\alpha * X_{t-1} = m|X_{t-1} = l)P(\varepsilon_t = k-m)\mathbf{I}(l\neq 0) + P(\varepsilon_t = k)\mathbf{I}(l=0) \\
&= \sum_{m=0}^{k}\binom{l+m-1}{m}\left(\frac{1}{1+\alpha}\right)^l\left(\frac{\alpha}{1+\alpha}\right)^m P(\varepsilon_t = k-m)\mathbf{I}(l\neq 0) + P(\varepsilon_t = k)\mathbf{I}(l=0).
\end{aligned}$$

□

3.2.2 Estimation and inference of the NNLINAR(1) model

In this section, we consider parameter estimation methods of the unknown parameters by (1) the conditional least squares estimator (CLS) and (2) the Yule-Walker estimator (YW). These estimators are compared via Monte Carlo simulations in terms of their means and variances by using the statistical software R [11].

3.2.2.1 Conditional least squares estimation

The conditional least squares estimators of the parameters α and μ of the $NNLINAR(1)$ model are obtained by minimizing the function defined in Definition 2.5.1. Let $k = 0$ in the expression in Theorem 3.2.5, the conditional expectation is

$$E(X_t|X_{t-1}) = \alpha X_{t-1} + \mu(1-\alpha),$$

where $\mu = E(X_t)$. Then

$$Q_n = \sum_{t=2}^{n}(X_t - E(X_t|X_{t-1}))^2 = \sum_{t=2}^{n}(X_t - \alpha X_{t-1} - \mu(1-\alpha))^2.$$

Equating the first order partial derivatives of Q_n with respect to μ and α to zero, then

$$\frac{\partial Q_n}{\partial \mu}\Big|_{\mu=\hat{\mu},\alpha=\hat{\alpha}} = -\sum_{t=2}^{n} 2(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))(1-\hat{\alpha}) = 0. \qquad (3.41)$$

$$\frac{\partial Q_n}{\partial \hat{\alpha}}\Big|_{\mu=\hat{\mu},\alpha=\hat{\alpha}} = \sum_{t=2}^{n} 2(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))(\hat{\mu} - X_{t-1}) = 0. \qquad (3.42)$$

From (3.41),

$$\sum_{t=2}^{n} X_t - \hat{\alpha}\sum_{t=2}^{n} X_{t-1} - \hat{\mu}(n-1)(1-\hat{\alpha}) = 0. \qquad (3.43)$$

By solving the equation (3.43), the estimation of μ can be computed as

$$\hat{\mu} = \frac{\sum_{t=2}^{n} X_t - \hat{\alpha}\sum_{t=2}^{n} X_{t-1}}{(n-1)(1-\hat{\alpha})}. \tag{3.44}$$

From (3.42)

$$\begin{aligned} 0 = \hat{\mu}\sum_{t=2}^{n} X_t - \hat{\alpha}\hat{\mu}\sum_{t=2}^{n} X_{t-1} - \hat{\mu}^2(1-\hat{\alpha})(n-1) - \sum_{t=2}^{n} X_{t-1}X_t + \hat{\alpha}\sum_{t=2}^{n} X_{t-1}^2 \\ + (1-\hat{\alpha})\hat{\mu}\sum_{t=2}^{n} X_{t-1}. \end{aligned} \tag{3.45}$$

By solving equation (3.45) and substitute $\hat{\mu}$ in equation (3.44), the estimation of α can be computed as

$$\hat{\alpha} = \frac{(n-1)\sum_{t=2}^{n} X_{t-1}X_t - \sum_{t=2}^{n} X_t \sum_{t=2}^{n} X_{t-1}}{(n-1)\sum_{t=2}^{n} X_{t-1}^2 - \left(\sum_{t=2}^{n} X_{t-1}\right)^2}.$$

From (3.42) and (3.34), we have

$$\frac{2\hat{\beta}_{CLS} + \hat{\theta}_{CLS}}{\hat{\theta}_{CLS}(\hat{\theta}_{CLS} + \hat{\beta}_{CLS})} = \hat{\mu}_{CLS} = \frac{\sum_{t=2}^{n} X_t - \hat{\alpha}_{CLS}\sum_{t=2}^{n} X_{t-1}}{(n-1)(1-\hat{\alpha}_{CLS})}.$$

The conditional least squares estimator of parameter σ^2 is obtained by minimizing the function defined in Abdulhamid et al. [2]. First, substitute $k = 0$ into (3.37),

$$Var(X_t|X_{t-1}) = \alpha(1-\alpha)X_{t-1} + Var(\varepsilon_t). \tag{3.46}$$

Substitute $Var(\varepsilon_t)$ from Theorem 3.2.4 in (3.46), the conditional variance is

$$Var(X_t|X_{t-1}) = \alpha(1+\alpha)X_{t-1} + (1-\alpha^2)\sigma^2 - \alpha(1+\alpha)\mu. \tag{3.47}$$

To obtain $\hat{\sigma}^2$, we follow Abdulhamid et al. [2] by minimizing the function S_n defined as

$$\begin{aligned} S_n &= \sum_{t=2}^{n}[(X_t - E(X_t|X_{t-1}))^2 - Var(X_t|X_{t-1})]^2 \\ &= \sum_{t=2}^{n}[(X_t - \alpha X_{t-1} - \mu(1-\alpha))^2 - \alpha(1+\alpha)X_{t-1} - (1-\alpha^2)\sigma^2 + \alpha(1+\alpha)\mu]^2. \end{aligned}$$

Taking the first order partial derivative of S_n with respect to σ^2 and equating it to zero, we get

$$\begin{aligned} 0 &= \frac{\partial S_n}{\partial \sigma^2}\Big|_{\sigma^2=\hat{\sigma}^2,\mu=\hat{\mu},\alpha=\hat{\alpha}} \\ &= \sum_{t=2}^{n} 2[(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1+\hat{\alpha})X_{t-1} - (1-\hat{\alpha}^2)\hat{\sigma}^2 + \hat{\alpha}(1+\hat{\alpha})\hat{\mu}](\hat{\alpha}^2 - 1). \end{aligned}$$

Then

$$\sum_{t=2}^{n}[(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1+\hat{\alpha})X_{t-1} - (1-\hat{\alpha}^2)\hat{\sigma}^2 + \hat{\alpha}(1+\hat{\alpha})\hat{\mu}] = 0. \quad (3.48)$$

By solving the equation (3.48), the estimation of σ^2 can be obtained as

$$\hat{\sigma}^2 = \frac{\sum_{t=2}^{n}[(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1+\hat{\alpha})X_{t-1} + \hat{\alpha}(1+\hat{\alpha})\hat{\mu}]}{(1-\hat{\alpha}^2)(n-1)}. \quad (3.49)$$

From (3.35) and (3.49)

$$\begin{aligned} &\frac{2\hat{\beta}^2(1+\hat{\theta}) + \hat{\theta}^2(1+\hat{\theta}) + \hat{\beta}\hat{\theta}(4+3\hat{\theta})}{\hat{\theta}^2(\hat{\beta}+\hat{\theta})^2} \\ &= \frac{\sum_{t=2}^{n}[(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1+\hat{\alpha})X_{t-1} + \hat{\alpha}(1+\hat{\alpha})\hat{\mu}]}{(1-\hat{\alpha}^2)(n-1)}. \end{aligned}$$

3.2.2.2 The Yule-Walker estimation

In this part, the Yule-Walker estimation for α, μ and σ^2 are obtained. By using Definition 2.3.3, the sample autocovariance function of X_t

$$\hat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(X_t - \bar{X})(X_{t+k} - \bar{X}), \tag{3.50}$$

where $0 \leq k < n$ and $\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$ is the sample mean.
From the Yule-Walker equation defined in Definition 2.5.2 and equation (3.50), the Yule-Walker estimator of α is

$$\hat{\alpha}_{YW} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\sum_{t=2}^{n}(X_t - \bar{X})(X_{t-1} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}.$$

Consider $\mu = E(X_t)$ defined in (3.34) and $\sigma^2 = Var(X_t)$ defined in (3.35) and note that $S^2 = \frac{\sum_{t=1}^{n}(X_t - \bar{X})^2}{n-1}$. The Yule-Walker estimators of μ and σ^2 are

$$\hat{\mu}_{YW} = \bar{X} = \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})},$$

$$\hat{\sigma^2}_{YW} = S^2 = \frac{2\hat{\beta}^2(1+\hat{\theta}) + \hat{\theta}^2(1+\hat{\theta}) + \hat{\beta}\hat{\theta}(4+3\hat{\theta})}{\hat{\theta}^2(\hat{\beta}+\hat{\theta})^2},$$

respectively.

3.2.3 Real data

In this section, we apply the two models with two real data sets : (1) the numbers of Skin-lesions and (2) the numbers of Anorexias.

3.2.3.1 The numbers of Skin-lesions

The first example considers the numbers of Skin-lesions monthly from January 2003 to December 2009 from a region in New Zealand. The data was originally introduced in Aghababaei et al. [1]. Sample mean and variance are 1.43 and 3.36, respectively. The fitted $NNLINAR(1)$ model is

$$X_t = 0.2365 * X_{t-1} + \varepsilon_t,$$

The predicted values of the numbers of Skin-lesions series are given by

$$\hat{X}_1 = \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})} = 1.4142$$
$$\hat{X}_t = \hat{\alpha}\hat{X}_{t-1} + (1 - \hat{\alpha})\frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})}.$$

Substituting parameter estimates $\hat{\alpha} = 0.2365$ and $\frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})} = 1.4142$.

$$\hat{X}_t = 0.2365\hat{X}_{t-1} + 1.0797, \quad t = 2, 3, \ldots, 72.$$

The expectation and variance computed from the $NNLINAR(1)$ model are 1.4142 and 3.356, respectively. We can see that the model can capture the sample mean and variance of the data set. Therefore, the model is reasonable to this data set.

3.2.3.2 The numbers of Anorexias

The second example considers the numbers of Anorexias monthly from January 2003 to December 2009 from a region in New Zealand. The data was original introduced in Aghababaei et al. [1]. Sample mean and variance are 0.82 and 2.90, respectively. The fitted $NNLINAR(1)$ is

$$X_t = 0.4909 * X_{t-1} + \varepsilon_t,$$

The predicted values of the numbers of Anorexias series are given by

$$\hat{X}_1 = \frac{2\beta + \theta}{\theta(\beta + \theta)} = 0.8313,$$
$$\hat{X}_t = \hat{\alpha}\hat{X}_{t-1} + (1 - \hat{\alpha})\frac{2\beta + \theta}{\theta(\beta + \theta)}.$$

Substituting parameter estimates $\hat{\alpha} = 0.4909$ and $\frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})} = 0.8313$.

$$\hat{X}_t = 0.4909\hat{X}_{t-1} + 0.4232, \quad t = 2, 3, \ldots, 72.$$

The expectation and variance computed from the $NNLINAR(1)$ model are 0.831 and 2.900, respectively. We can see that the model can capture the sample mean and variance of the data set. Therefore, the model is reasonable to this data set.

CHAPTER IV

INAR(1) MODEL WITH A TWO-PARAMETER GENERALIZED POISSON-LINDLEY INNOVATION

In this chapter, we construct two first order integer-valued autoregressive models with a two-parameter generalized Poisson-Lindley innovation based on (1) the binomial thinning operator and (2) the negative binomial thinning operator. Moreover, probabilistic properties of the constructed models and parameter estimation are demonstrated.

4.1 Construction of the first order integer-valued autoregressive model with a two-parameter generalized Poisson-Lindley innovation based on the binomial thinning operator (NLINARB(1))

In this section, we construct the first order integer-valued autoregressive model with two-parameter generalized Poisson-Lindley innovation based on the binomial thinning operator model ($NLINARB(1)$). Moreover, we investigate many properties of the constructed model such as expectations, parameter estimations and perform some numerical studies.

Definition 4.1.1. The first order integer-valued autoregressive model with two-parameter generalized Poisson-Lindley innovation based on the binomial thinning operator (NLINARB(1)) $\{X_t\}_{t\geq 1}$ is defined as

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \tag{4.1}$$

where the binomial thinning $\alpha\circ$ is defined in Definition 2.4.1 and $\{\epsilon_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta,\beta)$ defined in Definition 2.1.9 such that $\alpha\circ X_{t-1}$ and ϵ_t are independent.

Theorem 4.1.1. The autocovariance function, γ_k $(k \geq 1)$, of the $NLINARB(1)$ model $\{X_t\}_{t\geq 1}$ defined in Definition 4.1.1 is given by

$$\gamma_k = Cov(X_t, X_{t-k}) = \alpha^k \gamma_0. \tag{4.2}$$

where γ_0 is the variance of X_t.

Consequently, the autocorrelation function of order k, ρ_k, of the $NLINARB(1)$ model is

$$\rho_k = \alpha^k. \tag{4.3}$$

Proof. From (4.1) and the property that ϵ_t and X_{t-k} are independent, for $k \geq 1$,

$$\begin{aligned}
\gamma_k &= Cov(X_t, X_{t-k}) \\
&= Cov(\alpha\circ X_{t-1} + \epsilon_t, X_{t-k}) \\
&= Cov(\alpha\circ X_{t-1}, X_{t-k}) + Cov(\epsilon_t, X_{t-k}) \\
&= \alpha Cov(X_{t-1}, X_{t-k}) \\
&= \alpha Cov(\alpha\circ X_{t-2} + \epsilon_{t-1}, X_{t-k}) \\
&= \alpha^{k-1} Cov(\alpha\circ X_{t-k}, X_{t-k}) \qquad (4.4) \\
&= \alpha^k \gamma_0.
\end{aligned}$$

By applying (4.1) recursively to obtain (4.4). Consequently, the correlation function ρ_k can be written as

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \alpha^k.$$

□

Remark 4.1.1. From (4.3), the autocorrelation function declines exponentially as k converges to infinity.

4.1.1 Probabilistic properties of the NLINARB(1) model

In this section, we investigate many conditional properties such as conditional expectation and conditional variance of the constructed model. Since $\{\epsilon_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta,\beta)$. From Theorem 2.1.5, expectation and variance of ϵ_t for the $NLINARB(1)$ model are given respectively by

$$E(\epsilon_t) = \frac{2\beta+\theta}{\theta(\beta+\theta)}, \tag{4.5}$$

$$Var(\epsilon_t) = \frac{2\beta^2(1+\theta)+\theta^2(1+\theta)+\beta\theta(4+3\theta)}{\theta^2(\beta+\theta)^2}. \tag{4.6}$$

Theorem 4.1.2. The expectation of X_t defined in (4.1) is

$$E(X_t) = \frac{2\beta+\theta}{\theta(\beta+\theta)(1-\alpha)}.$$

Proof. From (4.1), since $\{\epsilon_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta,\beta)$,

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

Then, $\quad (1-\alpha)E(X_t) = E(\epsilon_t)$.

Therefore, $\quad E(X_t) = \dfrac{E(\epsilon_t)}{1-\alpha} = \dfrac{2\beta+\theta}{\theta(\beta+\theta)(1-\alpha)}.$ □

Theorem 4.1.3. The variance of X_t defined in (4.1) is

$$Var(X_t) = \frac{\theta^2(1+\alpha+\alpha\theta)+2\beta^2(1+\theta+\alpha\theta)+\beta\theta(4+3(1+\alpha)\theta)}{(1-\alpha^2)\theta^2(\beta+\theta)^2}.$$

Proof. From (4.1), since $\{\epsilon_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$,

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

Then
$$(1-\alpha^2)Var(X_t) = \alpha(1-\alpha)E(X_t) + Var(\epsilon_t).$$

Consequently,
$$\begin{aligned} Var(X_t) &= \frac{\alpha(1-\alpha)E(X_t) + Var(\epsilon_t)}{(1-\alpha^2)} \\ &= \frac{\alpha E(\epsilon_t) + Var(\epsilon_t)}{(1-\alpha^2)}. \end{aligned} \tag{4.7}$$

By using (4.5)–(4.7), we have

$$Var(X_t) = \frac{\theta^2(1+\alpha+\alpha\theta) + 2\beta^2(1+\theta+\alpha\theta) + \beta\theta(4+3(1+\alpha)\theta)}{(1-\alpha^2)\theta^2(\beta+\theta)^2}.$$

□

Theorem 4.1.4. The $(k+1)$-step ahead conditional expectation of the $NLINARB(1)$ model is

$$E(X_{t+k}|X_{t-1} = x) = \alpha^{k+1}x + \frac{(1-\alpha^{k+1})}{(1-\alpha)}\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right),$$

for $x = 0, 1, 2, \ldots$.

Proof.

$$\begin{aligned} E(X_{t+k}|X_{t-1} = x) &= E(\alpha \circ X_{t+k-1} + \epsilon_{t+k}|X_{t-1} = x) \\ &= E(\alpha \circ (\alpha \circ X_{t+k-2} + \epsilon_{t+k-1}) + \epsilon_{t+k}|X_{t-1} = x), \end{aligned}$$

by using (4.1) to obtain the last equality. Applying (4.1) to $\{X_t\}_{t\geq 1}$ recursively,

$$\begin{aligned}
E(X_{t+k}|X_{t-1}=x) &= E(\alpha^{k+1}\circ X_{t-1}+\alpha^{k}\circ\epsilon_t+\alpha^{k-1}\circ\epsilon_{t+1}+\cdots+\epsilon_{t+k}|X_{t-1}=x)\\
&= E(\alpha^{k+1}\circ X_{t-1}|X_{t-1}=x)+\sum_{h=0}^{k}E(\alpha^{h}\circ\epsilon_{t+k-h}|X_{t-1}=x)\\
&= \alpha^{k+1}x+\sum_{h=0}^{k}\alpha^{h}E(\epsilon_{t+k-h}) \qquad (4.8)\\
&= \alpha^{k+1}x+\left(\frac{1-\alpha^{k+1}}{1-\alpha}\right)E(\epsilon_t), \qquad (4.9)
\end{aligned}$$

where we use Theorem 2.4.2(3) to obtain (4.8).

Substitute $E(\epsilon_t)$ in (4.9), then

$$E(X_{t+k}|X_{t-1}=x)=\alpha^{k+1}x+\frac{(1-\alpha^{k+1})}{(1-\alpha)}\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right).$$

□

Remark 4.1.2. The conditional expectation $E(X_{t+k}|X_{t-1}=x)$ converges to the unconditional expectation $\frac{2\beta+\theta}{\theta(\beta+\theta)(1-\alpha)}$ as k converges to infinity.

Proof. Since $0<\alpha<1$,

$$\begin{aligned}
\lim_{k\to\infty}E(X_{t+k}|X_{t-1}=x) &= \lim_{k\to\infty}\left(\alpha^{k+1}x+\frac{(1-\alpha^{k+1})}{(1-\alpha)}\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right)\right)\\
&= \frac{2\beta+\theta}{\theta(\beta+\theta)(1-\alpha)}.
\end{aligned}$$

□

Theorem 4.1.5. The $(k+1)$-step ahead conditional variance for the $NLINARB(1)$ model is

$$\begin{aligned} Var(X_{t+k}|X_{t-1}=x) &= \alpha^{k+1}(1-\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(\epsilon_t) \\ &\quad + \frac{(1-\alpha^k)(\alpha-\alpha^{k+2})}{1-\alpha^2}E(\epsilon_t), \end{aligned} \tag{4.10}$$

for $x = \{0, 1, 2, \ldots\}$.

Proof.

$$\begin{aligned} &Var(X_{t+k}|X_{t-1}=x) \\ &= Var(\alpha\circ X_{t+k-1} + \epsilon_{t+k}|X_{t-1}=x) \\ &= Var(\alpha\circ(\alpha\circ X_{t+k-2} + \epsilon_{t+k-1}) + \epsilon_{t+k}|X_{t-1}=x) \\ &= Var(\alpha^{k+1}\circ X_{t-1} + \alpha^k\circ\epsilon_t + \alpha^{k-1}\circ\epsilon_{t+1} + \cdots + \epsilon_{t+k}|X_{t-1}=x) \qquad (4.11) \\ &= Var(\alpha^{k+1}\circ X_{t-1}) + \sum_{h=0}^{k} Var(\alpha^h\circ\epsilon_{t+h}|X_{t-1}=x) \\ &= \alpha^{k+1}(1-\alpha^{k+1})x + \sum_{h=0}^{k} Var(\alpha^h\circ\epsilon_{t+h}) \\ &= \alpha^{k+1}(1-\alpha^{k+1})x + \sum_{h=0}^{k}(\alpha^h(1-\alpha^h)\epsilon_{t+h} + \alpha^{2h})Var(\epsilon_{t+h}) \qquad (4.12) \\ &= \alpha^{k+1}(1-\alpha^{k+1})x + Var(\epsilon_t)\sum_{h=0}^{k}\alpha^{2h} + E(\epsilon_t)\sum_{h=0}^{k}\alpha^h(1-\alpha^h) \\ &= \alpha^{k+1}(1-\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(\epsilon_t) + \frac{(1-\alpha^k)(\alpha-\alpha^{k+2})}{1-\alpha^2}E(\epsilon_t), \end{aligned}$$

where we use (4.1) to obtain (4.11) and Theorem 2.4.1(4) to obtain (4.12). □

Remark 4.1.3. The conditional variance $Var(X_{t+k}|X_{t-1}=x)$ converges to the unconditional variance $\frac{\theta^2(1+\alpha+\alpha\theta)+2\beta^2(1+\theta+\alpha\theta)+\beta\theta(4+3(1+\alpha)\theta)}{(1-\alpha^2)\theta^2(\beta+\theta)^2}$ as k converges to infinity.

Proof. Since $0 < \alpha < 1$,

$$\begin{aligned}
&\lim_{k\to\infty} Var(X_{t+k}|X_{t-1} = x) \\
&= \lim_{k\to\infty}\left[\alpha^{k+1}(1-\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(\epsilon_t) + \frac{(1-\alpha^k)(\alpha-\alpha^{k+2})}{1-\alpha^2}E(\epsilon_t)\right] \\
&= \frac{Var(\epsilon_t)}{1-\alpha^2} + \frac{\alpha E(\epsilon_t)}{1-\alpha^2} \\
&= \frac{\theta^2(1+\alpha+\alpha\theta) + 2\beta^2(1+\theta+\alpha\theta) + \beta\theta(4+3(1+\alpha)\theta)}{(1-\alpha^2)\theta^2(\beta+\theta)^2}.
\end{aligned}$$

□

Theorem 4.1.6. The Markov process with transition probabilities of the $NLINARB(1)$ model is

$$p_{lk} = \sum_{m=0}^{\min(l,k)} \binom{l}{m} \alpha^m (1-\alpha)^{l-m} P(\epsilon_t = k-m),$$

where the process ϵ_t is defined in (4.1).

Proof.

$$\begin{aligned}
p_{lk} &= P(X_t = k|X_{t-1} = l) \\
&= P(\alpha \circ X_{t-1} + \epsilon_t = k|X_{t-1} = l) \\
&= \sum_{m=0}^{\min(l,k)} P(\alpha \circ X_{t-1} = m|X_{t-1} = l)P(\epsilon_t = k-m|X_{t-1} = l) \\
&= \sum_{m=0}^{\min(l,k)} P(\alpha \circ X_{t-1} = m|X_{t-1} = l)P(\epsilon_t = k-m) \\
&= \sum_{m=0}^{\min(l,k)} \binom{l}{m} \alpha^m (1-\alpha)^{l-m} P(\epsilon_t = k-m),
\end{aligned}$$

by using Definition 2.4.1 and the process ϵ_t is defined in (4.1). □

4.1.2 Estimation and inference of the NLINARB(1) model

In this section, we consider parameter estimation methods of the unknown parameters by (1) the conditional least squares estimator (CLS) and (2) the Yule-Walker estimator (YW). These estimators are compared via Monte Carlo simulations in terms of their means and variances by using the statistical software R [11].

4.1.2.1 Conditional least squares estimation

The conditional least squares estimators of the parameters α and μ of the $NLINARB(1)$ model are obtained by minimizing the function defined in Definition 2.5.1. Let $k = 0$ in the expression in Theorem 4.1.4, the conditional expectation is

$$E(X_t|X_{t-1}) = \alpha X_{t-1} + \mu(1-\alpha),$$

where $\mu = E(X_t)$. Then

$$Q_n = \sum_{t=2}^{n}(X_t - E(X_t|X_{t-1}))^2 = \sum_{t=2}^{n}(X_t - \alpha X_{t-1} - \mu(1-\alpha))^2.$$

Equating the first order partial derivatives of Q_n with respect to μ and α to zero, then

$$\frac{\partial Q_n}{\partial \mu}\Big|_{\mu=\hat{\mu},\alpha=\hat{\alpha}} = -\sum_{t=2}^{n} 2(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))(1-\hat{\alpha}) = 0, \qquad (4.13)$$

$$\frac{\partial Q_n}{\partial \hat{\alpha}}\Big|_{\mu=\hat{\mu},\alpha=\hat{\alpha}} = \sum_{t=2}^{n} 2(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))(\hat{\mu} - X_{t-1}) = 0. \qquad (4.14)$$

From (4.13),

$$\sum_{t=2}^{n} X_t - \hat{\alpha}\sum_{t=2}^{n} X_{t-1} - \hat{\mu}(n-1)(1-\hat{\alpha}) = 0. \qquad (4.15)$$

By solving equation (4.15), the estimation of μ can be computed as

$$\hat{\mu} = \frac{\sum_{t=2}^{n} X_t - \hat{\alpha} \sum_{t=2}^{n} X_{t-1}}{(n-1)(1-\hat{\alpha})}. \tag{4.16}$$

From (4.14),

$$\begin{aligned} 0 = \hat{\mu} \sum_{t=2}^{n} X_t - \hat{\alpha}\hat{\mu} \sum_{t=2}^{n} X_{t-1} - \hat{\mu}^2(1-\hat{\alpha})(n-1) - \sum_{t=2}^{n} X_{t-1}X_t + \hat{\alpha} \sum_{t=2}^{n} X_{t-1}^2 \\ + (1-\hat{\alpha})\hat{\mu} \sum_{t=2}^{n} X_{t-1}. \end{aligned} \tag{4.17}$$

By solving equation (4.17) and substitute $\hat{\mu}$ in equation (4.16), the estimation of α can be computed as

$$\hat{\alpha} = \frac{(n-1)\sum_{t=2}^{n} X_{t-1}X_t - \sum_{t=2}^{n} X_t \sum_{t=2}^{n} X_{t-1}}{(n-1)\sum_{t=2}^{n} X_{t-1}^2 - \left(\sum_{t=2}^{n} X_{t-1}\right)^2}.$$

From (3.21) and Theorem 4.1.2, we have

$$\frac{2\hat{\beta}_{CLS} + \hat{\theta}_{CLS}}{\hat{\theta}_{CLS}(\hat{\theta}_{CLS} + \hat{\beta}_{CLS})(1-\hat{\alpha}_{CLS})} = \hat{\mu}_{CLS} = \frac{\sum_{t=2}^{n} X_t - \hat{\alpha}_{CLS} \sum_{t=2}^{n} X_{t-1}}{(n-1)(1-\hat{\alpha}_{CLS})}.$$

The conditional least squares estimator of the parameter σ^2 is obtained by minimizing the function defined in Abdulhamid et al. [2]. First, substitute $k = 0$ into (4.10).

$$Var(X_t|X_{t-1}) = \alpha(1-\alpha)X_{t-1} + Var(\epsilon_t). \tag{4.18}$$

Substitute $Var(\epsilon_t)$ from (4.6) into (4.18), the conditional variance is

$$Var(X_t|X_{t-1}) = \alpha(1-\alpha)X_{t-1} + (1-\alpha^2)\sigma^2 - \alpha(1-\alpha)\mu. \tag{4.19}$$

To obtain $\hat{\sigma}^2$, we follow Abdulhamid et al. [2] by minimizing the function S_n defined as

$$\begin{aligned}S_n &= \sum_{t=2}^{n}[(X_t - E(X_t|X_{t-1}))^2 - Var(X_t|X_{t-1})]^2\\ &= \sum_{t=2}^{n}[(X_t - \alpha X_{t-1} - \mu(1-\alpha))^2 - \alpha(1-\alpha)X_{t-1} - (1-\alpha^2)\sigma^2 + \alpha(1-\alpha)\mu]^2.\end{aligned}$$

Taking the first order partial derivative of S_n with respect to σ^2 and equating it to zero, we get

$$\begin{aligned}0 &= \frac{\partial S_n}{\partial \sigma^2}\Big|_{\sigma^2=\hat{\sigma}^2, \mu=\hat{\mu}, \alpha=\hat{\alpha}}\\ &= \sum_{t=2}^{n} 2[(X_t - \hat{\alpha} X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1-\hat{\alpha})X_{t-1} - (1-\hat{\alpha}^2)\hat{\sigma}^2 + \hat{\alpha}(1-\hat{\alpha})\hat{\mu}](\hat{\alpha}^2 - 1).\end{aligned}$$

Then

$$\sum_{t=2}^{n}[(X_t - \hat{\alpha} X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1-\hat{\alpha})X_{t-1} - (1-\hat{\alpha}^2)\hat{\sigma}^2 + \hat{\alpha}(1-\hat{\alpha})\hat{\mu}] = 0. \quad (4.20)$$

By solving the equation (4.20), the estimation of σ^2 can be obtained as

$$\hat{\sigma}^2 = \frac{\sum_{t=2}^{n}[(X_t - \hat{\alpha} X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1-\hat{\alpha})X_{t-1} + \hat{\alpha}(1-\hat{\alpha})\hat{\mu}]}{(1-\hat{\alpha}^2)(n-1)}. \quad (4.21)$$

From Theorem 4.1.3, and (4.21),

$$\begin{aligned}&\frac{\hat{\theta}^2(1+\hat{\alpha}+\hat{\alpha}\hat{\theta}) + 2\hat{\beta}^2(1+\hat{\theta}+\hat{\alpha}\hat{\theta}) + \hat{\beta}\hat{\theta}(4+3(1+\hat{\alpha})\hat{\theta})}{(1-\hat{\alpha}^2)\hat{\theta}^2(\hat{\beta}+\hat{\theta})^2}\\ &= \hat{\sigma}^2_{CLS} = \frac{\sum_{t=2}^{n}[(X_t - \hat{\alpha} X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1-\hat{\alpha})X_{t-1} + \hat{\alpha}(1-\hat{\alpha})\hat{\mu}]}{(1-\hat{\alpha}^2)(n-1)}.\end{aligned}$$

4.1.2.2 The Yule-Walker estimation

In this part, the Yule-Walker estimation for α, μ and σ^2 are obtained. By using Definition 2.3.3 then the sample autocovariance function of X_t

$$\hat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(X_t - \bar{X})(X_{t+k} - \bar{X}), \tag{4.22}$$

where $0 \leq k < n$ and $\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$ is the sample mean.
From the Yule-Walker equation defined in Definition 2.5.2 and equation (4.22), the Yule-Walker estimator of α is

$$\hat{\alpha}_{YW} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\sum_{t=2}^{n}(X_t - \bar{X})(X_{t-1} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}.$$

Consider $\mu = E(X_t)$ defined in Theorem 4.1.2 and $\sigma^2 = Var(X_t)$ defined in (4.6) and note that $S^2 = \frac{\sum_{t=1}^{n}(X_t - \bar{X})}{n-1}$. The Yule-Walker estimators of μ and σ^2 are

$$\hat{\mu}_{YW} = \bar{X} = \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})(1 - \hat{\alpha})}, \tag{4.23}$$

$$\hat{\sigma^2}_{YW} = S^2 = \frac{\hat{\theta}^2(1 + \hat{\alpha} + \hat{\alpha}\hat{\theta}) + 2\hat{\beta}^2(1 + \hat{\theta} + \hat{\alpha}\hat{\theta}) + \hat{\beta}\hat{\theta}(4 + 3(1 + \hat{\alpha})\hat{\theta})}{(1 - \hat{\alpha}^2)\hat{\theta}^2(\hat{\beta} + \hat{\theta})^2}, \tag{4.24}$$

respectively.

4.1.3 Simulation Results

In this section, we produce 1,000 samples from the $NLINARB(1)$ model for true parameter values in different settings (1) $\alpha = 0.1, \beta = 0.1, \theta = 0.5$; (2) $\alpha = 0.4, \beta = 0.3, \theta = 0.8$; (3) $\alpha = 0.7, \beta = 0.5, \theta = 1.5$ of different sample sizes $n = 50, 500, 1000$ by using the statistical software R and obtain estimators of parameters from two methods described in Section 4.1.2. Then we compare the obtained estimators in terms of their

$(\alpha, \theta, \beta, \mu, \sigma^2) = (0.1, 0.5, 0.1, 2.593, 7.8938)$						
n	$\hat{\alpha}_{CLS}$	$\hat{\mu}_{CLS}$	$\hat{\sigma}^2_{CLS}$	$\hat{\alpha}_{YW}$	$\hat{\mu}_{YW}$	$\hat{\sigma}^2_{YW}$
50	0.1020 (0.0181)	3.6138 (0.2874)	10.8499 (14.2988)	0.0766 (0.0184)	3.6150 (0.2812)	11.0341 (14.5398)
500	0.0936 (0.0095)	2.5714 (0.0866)	7.7373 (4.0705)	0.0970 (0.0021)	2.5917 (0.0191)	7.8660 (0.9042)
1000	0.0996 (0.0010)	2.5879 (0.0092)	7.8122 (0.4615)	0.0987 (0.0010)	2.5880 (0.0092)	7.8200 (0.4609)
$(\alpha, \theta, \beta, \mu, \sigma^2) = (0.4, 0.8, 0.3, 2.6515, 5.3879)$						
n	$\hat{\alpha}_{CLS}$	$\hat{\mu}_{CLS}$	$\hat{\sigma}^2_{CLS}$	$\hat{\alpha}_{YW}$	$\hat{\mu}_{YW}$	$\hat{\sigma}^2_{YW}$
50	0.3736 (0.0177)	2.5851 (0.2519)	5.0868 (4.4710)	0.3473 (0.0186)	2.5694 (0.2433)	5.1007 (4.3824)
500	0.3837 (0.0098)	2.6319 (0.1300)	5.2851 (2.2662)	0.3945 (0.0021)	2.6467 (0.0244)	5.3759 (0.4151)
1000	0.4005 (0.0010)	2.6563 (0.0127)	5.3936 (0.2171)	0.3993 (0.0010)	2.6554 (0.0126)	5.3944 (0.2166)
$(\alpha, \theta, \beta, \mu, \sigma^2) = (0.7, 1.5, 0.5, 2.7778, 4.0305)$						
n	$\hat{\alpha}_{CLS}$	$\hat{\mu}_{CLS}$	$\hat{\sigma}^2_{CLS}$	$\hat{\alpha}_{YW}$	$\hat{\mu}_{YW}$	$\hat{\sigma}^2_{YW}$
50	0.6542 (0.0109)	2.7112 (0.4355)	3.6773 (2.9730)	0.6219 (0.01225)	2.6145 (0.3861)	3.5899 (2.7554)
500	0.6781 (0.0063)	2.7625 (0.2360)	3.9581 (1.9262)	0.6932 (0.0012)	2.7633 (0.0458)	4.0039 (0.3713)
1000	0.6979 (0.0006)	2.7783 (0.0245)	4.0228 (0.1877)	0.6966 (0.0006)	2.7779 (0.0243)	4.0291 (0.1872)

Table 4.1: Mean and variance (in brackets) of the estimators for different values of the parameters α, μ and σ^2 for the $NLINARB(1)$ model

means and variances. Table 4.1 shows mean and variance (in brackets) of the estimators for different values of the parameters α, μ and σ^2.

From Table 4.1, we observe that the estimators obtained from the two estimation methods converge to the true parameters. In addition, increasing the sample size yields smaller variance. The conditional least squares estimate (CLS) and the Yule-Walker estimate (YW) are approximately the same. Considering the variance, we can see that the CLS estimators have smaller variance than YW estimators for the parameter α. However, the YW has smaller variance than CLS for the parameters μ and σ^2. Then, considering the mean, we can see that the YW estimators converge to the true parameters faster than the CLS estimators for the parameters σ^2 and μ.

4.1.4 Real data

In this section, we apply the two models with two real data sets : (1) the incidents of acute febrile mucocutaneous lymph node syndrome and (2) the numbers of earthquakes per year magnitude 7.0 or greater (1900-1998).

4.1.4.1 The incidents of acute febrile mucocutaneous lymph node syndrome (MCLS)

The first example considers the data give weekly counts of the incidents of acute febrile mucocutaneous lymph node syndrome (MCLS) in Totori-prefecture, Japan, during 1982. Sample mean and variance are 1.711 and 3.111, respectively. The fitted $NLINARB(1)$ model is

$$X_t = 0.5241 \circ X_{t-1} + \epsilon_t,$$

The predicted values of the numbers of MCLS series are given by

$$\hat{X}_1 = \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})(1 - \hat{\alpha})} = 1.6843,$$
$$\hat{X}_t = \hat{\alpha}\hat{X}_{t-1} + \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})}.$$

Substituting parameter estimates $\hat{\alpha} = 0.5241$ and $\frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})} = 0.8016.$

$$\hat{X}_t = 0.5241\hat{X}_{t-1} + 0.8016, \quad t = 2, 3, \ldots, 52.$$

The expectation and variance computed from the $NLINARB(1)$ model are 1.684 and 3.111, respectively. We can see that the model can capture the sample mean and variance of the data set. Therefore, the model is reasonable to this data set.

4.1.4.2 The number of earthquakes per year magnitude 7.0 or greater (1900-1998).

The second example considers the numbers of earthquakes per year magnitude 7.0 or greater (1900-1998). Sample mean and variance are 20.02 and 52.75, respectively. The fitted $NLINARB(1)$ model is

$$X_t = 0.5434 \circ X_{t-1} + \epsilon_t,$$

The predicted values of the number of earthquakes per year magnitude 7.0 or greater series are given by

$$\hat{X}_1 = \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})(1 - \hat{\alpha})} = 20.1283,$$
$$\hat{X}_t = \hat{\alpha}\hat{X}_{t-1} + \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})},$$

Substituting parameter estimates $\hat{\alpha} = 0.5434$ and $\frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})} = 9.1906$.

$$\hat{X}_t = 0.5434\hat{X}_{t-1} + 9.1906, \quad t = 2, 3, \ldots, 99.$$

The expectation and variance computed from the $NLINARB(1)$ model are 20.128 and 52.755, respectively. We can see that the model can capture the sample mean and variance of the data set. Therefore, the model is reasonable to this data set.

4.2 Construction of the first order integer-valued autoregressive model with a two-parameter generalized Poisson-Lindley innovation based on the negative binomial thinning operator (NLINARN(1))

In this section, we construct the first order integer-valued autoregressive model with two-parameter generalized Poisson-Lindley innovation based on the negative binomial thinning operator model ($NLINARN(1)$). Moreover, we investigate many properties of the constructed model such as means, parameter estimations and perform some numerical studies.

Definition 4.2.1. The first order integer-valued autoregressive model with two-parameter generalized Poisson-Lindley innovation based on the negative binomial thinning operator (NLINARN(1)) $\{X_t\}_{t\geq 1}$ is defined as

$$X_t = \alpha * X_{t-1} + \varepsilon_t, \tag{4.25}$$

where the negative binomial thinning $\alpha*$ defined in Definition 2.4.2 and $\{\varepsilon_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$ defined in Definition 2.1.9 such that $\alpha * X_{t-1}$ and ε_t are independent.

Theorem 4.2.1. The autocovariance function, γ_k $(k \geq 1)$, of the $NLINARN(1)$ model $\{X_t\}_{t\geq 1}$ defined in Definition 4.2.1 is given by

$$\gamma_k = Cov(X_t, X_{t-k}) = \alpha^k \gamma_0, \tag{4.26}$$

where γ_0 is the variance of X_t.

Consequently, the autocorrelation function of order k, ρ_k, of the $NLINARN(1)$ model is

$$\rho_k = \alpha^k. \tag{4.27}$$

Proof. From (4.25) and the property that ε_t and X_{t-k} are independent, for $k \geq 1$,

$$\begin{aligned}\gamma_k &= Cov(X_t, X_{t-k}) \\ &= Cov(\alpha * X_{t-1} + \varepsilon_t, X_{t-k}) \\ &= Cov(\alpha * X_{t-1}, X_{t-k}) + Cov(\varepsilon_t, X_{t-k}) \\ &= \alpha Cov(X_{t-1}, X_{t-k}) \\ &= \alpha Cov(\alpha * X_{t-2} + \varepsilon_{t-1}, X_{t-k}) \\ &= \alpha^{k-1} Cov(\alpha * X_{t-k}, X_{t-k}) \qquad (4.28) \\ &= \alpha^k \gamma_0.\end{aligned}$$

By applying (4.25) recursively to obtain (4.28). Consequently, the correlation function ρ_k can be written as

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \alpha^k.$$

□

Remark 4.2.1. From (4.27), the autocorrelation function declines exponentially as k converges to infinity.

4.2.1 Probabilistic properties of the NLINARN(1) model

In this section, we investigate many conditional properties such as conditional expectation and conditional variance of the constructed model. Since $\{\varepsilon_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$. From Theorem 2.1.5, expectation and variance of ε_t for the $NLINARN(1)$ model are given respectively by

$$E(\varepsilon_t) = \frac{2\beta + \theta}{\theta(\beta + \theta)}, \qquad (4.29)$$

$$Var(\varepsilon_t) = \frac{2\beta^2(1+\theta) + \theta^2(1+\theta) + \beta\theta(4+3\theta)}{\theta^2(\beta+\theta)^2}. \qquad (4.30)$$

Theorem 4.2.2. The expectation of X_t defined in (4.25) is

$$E(X_t) = \frac{2\beta + \theta}{\theta(\beta + \theta)(1 - \alpha)}.$$

Proof. From (4.25), since $\{\varepsilon_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$,

$$\begin{aligned} E(X_t) &= E(\alpha * X_{t-1} + \varepsilon_t) \\ &= \alpha E(X_t) + E(\varepsilon_t). \end{aligned}$$

Then $(1 - \alpha)E(X_t) = E(\varepsilon_t)$,

Therefore, $E(X_t) = \dfrac{E(\varepsilon_t)}{1 - \alpha} = \dfrac{2\beta + \theta}{\theta(\beta + \theta)(1 - \alpha)}$. □

Theorem 4.2.3. The variance of X_t defined in (4.25) is

$$Var(X_t) = \frac{(2\beta^2 + \theta^2)(1 - \alpha + \theta + \alpha^2\theta) + \beta\theta(4 - 4\alpha + 3\theta + 3\alpha^2\theta)}{(\alpha - 1)^2(1 + \alpha)\theta^2(\beta + \theta)^2}.$$

Proof. From (4.25), since $\{\varepsilon_t\}_{t\geq 1}$ is a stationary process with the $NGPL(\theta, \beta)$,

$$\begin{aligned} Var(X_t) &= Var(\alpha * X_{t-1} + \varepsilon_t) \\ &= Var(\alpha * X_{t-1}) + Var(\varepsilon_t) \\ &= \alpha(1 + \alpha)E(X_{t-1}) + \alpha^2 Var(X_{t-1}) + Var(\varepsilon_t) \\ &= \alpha(1 + \alpha)E(X_t) + \alpha^2 Var(X_t) + Var(\varepsilon_t). \end{aligned}$$

Then, $(1 - \alpha^2)Var(X_t) = \alpha(1 + \alpha)E(X_t) + Var(\varepsilon_t)$.

Consequently, $$Var(X_t) = \frac{\alpha(1 + \alpha)E(X_t) + Var(\varepsilon_t)}{(1 - \alpha^2)}. \quad (4.31)$$

By using Theorem 4.2.2, (4.30) and (4.31), we have

$$Var(X_t) = \frac{(2\beta^2 + \theta^2)(1 - \alpha + \theta + \alpha^2\theta) + \beta\theta(4 - 4\alpha + 3\theta + 3\alpha^2\theta)}{(\alpha - 1)^2(1 + \alpha)\theta^2(\beta + \theta)^2}.$$

□

Theorem 4.2.4. The $(k+1)$-step ahead conditional expectation of the $NLINARN(1)$ model is

$$E(X_{t+k}|X_{t-1}=x) = \alpha^{k+1}x + \frac{(1-\alpha^{k+1})}{(1-\alpha)}\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right),$$

for $x = 0, 1, 2, \ldots$.

Proof.

$$\begin{aligned} E(X_{t+k}|X_{t-1}=x) &= E(\alpha * X_{t+k-1} + \varepsilon_{t+k}|X_{t-1}=x) \\ &= E(\alpha * (\alpha * X_{t+k-2} + \varepsilon_{t+k-1}) + \varepsilon_{t+k}|X_{t-1}=x), \end{aligned}$$

by using (4.25) to obtain the last equality. Applying (4.25) to $\{X_t\}_{t\geq 1}$ recursively,

$$\begin{aligned} E(X_{t+k}|X_{t-1}=x) &= E(\alpha^{k+1} * X_{t-1} + \alpha^k * \varepsilon_t + \alpha^{k-1} * \varepsilon_{t+1} + \cdots + \varepsilon_{t+k}|X_{t-1}=x) \\ &= E(\alpha^{k+1} * X_{t-1}|X_{t-1}=x) + \sum_{h=0}^{k} E(\alpha^h * \varepsilon_{t+k-h}|X_{t-1}=x) \\ &= \alpha^{k+1}x + \sum_{h=0}^{k} \alpha^h E(\varepsilon_{t+k-h}) \qquad (4.32) \\ &= \alpha^{k+1}x + \left(\frac{1-\alpha^{k+1}}{1-\alpha}\right) E(\varepsilon_t), \qquad (4.33) \end{aligned}$$

where we use Theorem 2.4.2(3) to obtain (4.32).
Substitute $E(\varepsilon_t)$ in (4.33), then

$$E(X_{t+k}|X_{t-1}=x) = \alpha^{k+1}x + \frac{(1-\alpha^{k+1})}{(1-\alpha)}\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right).$$

□

Remark 4.2.2. The conditional expectation $E(X_{t+k}|X_{t-1}=x)$ converges to the unconditional expectation $\dfrac{2\beta+\theta}{\theta(\beta+\theta)(1-\alpha)}$ as k converges to infinity.

Proof. Since $0 < \alpha < 1$,

$$\begin{aligned}\lim_{k\to\infty} E(X_{t+k}|X_{t-1}=x) &= \lim_{k\to\infty}\left(\alpha^{k+1}x + \frac{(1-\alpha^{k+1})}{(1-\alpha)}\left(\frac{2\beta+\theta}{\theta(\beta+\theta)}\right)\right)\\ &= \frac{2\beta+\theta}{\theta(\beta+\theta)(1-\alpha)}.\end{aligned}$$

□

Theorem 4.2.5. The $(k+1)$-step ahead conditional variance for the $NLINARN(1)$ model is

$$\begin{aligned}Var(X_{t+k}|X_{t-1}=x) =& \alpha^{k+1}(1+\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(\varepsilon_t)\\ &+ \frac{\alpha(1-\alpha^k)(1-\alpha^{k+1})}{1-\alpha^2}E(\varepsilon_t),\end{aligned} \tag{4.34}$$

for $x = 0, 1, 2, \ldots$.

Proof.

$$\begin{aligned}&Var(X_{t+k}|X_{t-1}=x)\\ &= Var(\alpha * X_{t+k-1} + \varepsilon_{t+k}|X_{t-1}=x)\\ &= Var(\alpha * (\alpha * X_{t+k-2} + \varepsilon_{t+k-1}) + \varepsilon_{t+k}|X_{t-1}=x) \qquad (4.35)\\ &= Var(\alpha^{k+1} * X_{t-1} + \alpha^k * \varepsilon_t + \alpha^{k-1} * \varepsilon_{t+1} + \cdots + \varepsilon_{t+k}|X_{t-1}=x)\\ &= Var(\alpha^{k+1} * X_{t-1}) + \sum_{h=0}^{k} Var(\alpha^h * \varepsilon_{t+k}|X_{t-1}=x)\\ &= \alpha^{k+1}(1+\alpha^{k+1})x + \sum_{h=0}^{k} Var(\alpha^h * \varepsilon_{t+k-h}) \qquad (4.36)\\ &= \alpha^{k+1}(1+\alpha^{k+1})x + \sum_{h=0}^{k}(\alpha^h(1+\alpha^h)E(\varepsilon_{t+k-h}) + \alpha^{2h}Var(\varepsilon_{t+k-h}))\\ &= \alpha^{k+1}(1+\alpha^{k+1})x + Var(\varepsilon_t)\sum_{h=0}^{k}\alpha^{2h} + E(\varepsilon_t)\sum_{h=0}^{k}\alpha^h(1-\alpha^h)\\ &= \alpha^{k+1}(1+\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(\varepsilon_t) + \frac{\alpha(1-\alpha^k)(1-\alpha^{k+1})}{1-\alpha^2}E(\varepsilon_t),\end{aligned}$$

where we use (4.25) to obtain (4.35) and Theorem 2.4.2(4) to obtain (4.36). □

Remark 4.2.3. The conditional variance $Var(X_{t+k}|X_{t-1} = x)$ converges to the unconditional variance $\frac{\theta^2(1+\alpha+\alpha\theta)+2\beta^2(1+\theta+\alpha\theta)+\beta\theta(4+3(1+\alpha)\theta)}{(1-\alpha^2)\theta^2(\beta+\theta)^2}$ as k converges to infinity.

Proof. Since $0 < \alpha < 1$,

$$\begin{aligned}
&\lim_{k\to\infty} Var(X_{t+k}|X_{t-1} = x) \\
&= \lim_{k\to\infty}\left[\alpha^{k+1}(1+\alpha^{k+1})x + \frac{1-\alpha^{2(k+1)}}{1-\alpha^2}Var(\varepsilon_t) + \frac{\alpha(1-\alpha^k)(1-\alpha^{k+1})}{1-\alpha^2}E(\varepsilon_t)\right] \\
&= \frac{Var(\varepsilon_t)}{1-\alpha^2} + \frac{\alpha E(\varepsilon_t)}{1-\alpha^2} \\
&= \frac{\theta^2(1+\alpha+\alpha\theta)+2\beta^2(1+\theta+\alpha\theta)+\beta\theta(4+3(1+\alpha)\theta)}{(1-\alpha^2)\theta^2(\beta+\theta)^2}.
\end{aligned}$$

□

Theorem 4.2.6. The Markov process with transition probabilities of the $NLINARN(1)$ model is

$$p_{lk} = \sum_{m=0}^{k}\binom{l+m-1}{m}\left(\frac{1}{1+\alpha}\right)^l\left(\frac{\alpha}{1+\alpha}\right)^m p(\varepsilon_t = k-m)\mathbf{I}(l\neq 0) + p(\varepsilon_t = k)\mathbf{I}(l=0),$$

where the process ε_t is defined in (4.25).

Proof.

$$\begin{aligned}
p_{lk} &= P(X_t = k|X_{t-1} = l) \\
&= P(\alpha * X_{t-1} + \varepsilon_t = k|X_{t-1} = l) \\
&= \sum_{m=o}^{k} P(\alpha * X_{t-1} = m|X_{t-1} = l)P(\varepsilon_t = k-m)\mathbf{I}(l\neq 0) + P(\varepsilon_t = k)\mathbf{I}(l=0) \\
&= \sum_{m=0}^{k}\binom{l+m-1}{m}\left(\frac{1}{1+\alpha}\right)^l\left(\frac{\alpha}{1+\alpha}\right)^m P(\varepsilon_t = k-m)\mathbf{I}(l\neq 0) + P(\varepsilon_t = k)\mathbf{I}(l=0).
\end{aligned}$$

□

4.2.2 Estimation and inference of the NLINARN(1) model

In this section, we consider parameter estimation methods of the unknown parameters by (1) the conditional least squares estimator (CLS) and (2) the Yule-Walker estimator (YW). These estimators are compared via Monte Carlo simulations in terms of their means and variances by using the statistical software R [11].

4.2.2.1 Conditional least squares estimation

The conditional least squares estimators of the parameters α and μ of the $NLINARN(1)$ model are obtained by minimizing the function defined in Definition 2.5.1. Let $k = 0$ in the expression in Theorem 4.2.4, the conditional expectation is

$$E(X_t|X_{t-1}) = \alpha X_{t-1} + \mu(1-\alpha),$$

where $\mu = E(X_t)$. Then

$$Q_n = \sum_{t=2}^{n}(X_t - E(X_t|X_{t-1}))^2 = \sum_{t=2}^{n}(X_t - \alpha X_{t-1} - \mu(1-\alpha))^2.$$

Equating the first order partial derivatives of Q_n with respect to μ and α to zero, we then

$$\frac{\partial Q_n}{\partial \mu}\Big|_{\mu=\hat{\mu},\alpha=\hat{\alpha}} = -\sum_{t=2}^{n} 2(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))(1-\hat{\alpha}) = 0, \quad (4.37)$$

$$\frac{\partial Q_n}{\partial \hat{\alpha}}\Big|_{\mu=\hat{\mu},\alpha=\hat{\alpha}} = \sum_{t=2}^{n} 2(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))(\hat{\mu} - X_{t-1}) = 0. \quad (4.38)$$

From (4.37),

$$\sum_{t=2}^{n} X_t - \hat{\alpha}\sum_{t=2}^{n} X_{t-1} - \hat{\mu}(n-1)(1-\hat{\alpha}) = 0. \quad (4.39)$$

By solving the equation (4.39), the estimation of μ can be computed as

$$\hat{\mu} = \frac{\sum_{t=2}^{n} X_t - \hat{\alpha} \sum_{t=2}^{n} X_{t-1}}{(n-1)(1-\hat{\alpha})}. \tag{4.40}$$

From (4.38),

$$\begin{aligned} 0 = {} & \hat{\mu} \sum_{t=2}^{n} X_t - \hat{\alpha}\hat{\mu} \sum_{t=2}^{n} X_{t-1} - \hat{\mu}^2(1-\hat{\alpha})(n-1) - \sum_{t=2}^{n} X_{t-1}X_t + \hat{\alpha} \sum_{t=2}^{n} X_{t-1}^2 \\ & + (1-\hat{\alpha})\hat{\mu} \sum_{t=2}^{n} X_{t-1}. \end{aligned} \tag{4.41}$$

By solving equation (4.41) and substitute $\hat{\mu}$ in equation (4.40), the estimation of α can be computed as

$$\hat{\alpha} = \frac{(n-1) \sum_{t=2}^{n} X_{t-1}X_t - \sum_{t=2}^{n} X_t \sum_{t=2}^{n} X_{t-1}}{(n-1) \sum_{t=2}^{n} X_{t-1}^2 - \left(\sum_{t=2}^{n} X_{t-1} \right)^2}.$$

From (4.40) and Theorem 4.2.2, we have

$$\frac{2\hat{\beta}_{CLS} + \hat{\theta}_{CLS}}{\hat{\theta}_{CLS}(\hat{\theta}_{CLS} + \hat{\beta}_{CLS})(1-\hat{\alpha}_{CLS})} = \hat{\mu}_{CLS} = \frac{\sum_{t=2}^{n} X_t - \hat{\alpha}_{CLS} \sum_{t=2}^{n} X_{t-1}}{(n-1)(1-\hat{\alpha}_{CLS})}.$$

The conditional least squares estimator of the parameters σ^2 is obtained by minimizing the function defined in Abdulhamid et al. [2]. Frist, substitute $k = 0$ into (4.34),

$$Var(X_t|X_{t-1}) = \alpha(1-\alpha)X_{t-1} + Var(\varepsilon_t). \tag{4.42}$$

Substitute $Var(\varepsilon_t)$ from (4.30) in (4.42), the conditional variance is

$$Var(X_t|X_{t-1}) = \alpha(1+\alpha)X_{t-1} + (1-\alpha^2)\sigma^2 - \alpha(1+\alpha)\mu. \tag{4.43}$$

To obtain $\hat{\sigma}^2$, we follow Abdulhamid et al. [2] by minimizing the function S_n defined as

$$
\begin{aligned}
S_n &= \sum_{t=2}^{n} [(X_t - E(X_t|X_{t-1}))^2 - Var(X_t|X_{t-1})]^2 \\
&= \sum_{t=2}^{n} [(X_t - \alpha X_{t-1} - \mu(1-\alpha))^2 - \alpha(1+\alpha)X_{t-1} - (1-\alpha^2)\sigma^2 + \alpha(1+\alpha)\mu]^2.
\end{aligned}
$$

Taking the first order partial derivative of S_n with respect to σ^2 and equating it to zero, we get

$$
\begin{aligned}
0 &= \frac{\partial S_n}{\partial \sigma^2}\Big|_{\sigma^2=\hat{\sigma}^2, \mu=\hat{\mu}, \alpha=\hat{\alpha}} \\
&= \sum_{t=2}^{n} 2[(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1+\hat{\alpha})X_{t-1} - (1-\hat{\alpha}^2)\hat{\sigma}^2 + \hat{\alpha}(1+\hat{\alpha})\hat{\mu}](\hat{\alpha}^2 - 1).
\end{aligned}
$$

Then

$$
\sum_{t=2}^{n} [(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1+\hat{\alpha})X_{t-1} - (1-\hat{\alpha}^2)\hat{\sigma}^2 + \hat{\alpha}(1+\hat{\alpha})\hat{\mu}] = 0. \quad (4.44)
$$

By solving the equation (4.44), the estimation of σ^2 can be obtained as

$$
\hat{\sigma}^2 = \frac{\sum_{t=2}^{n} [(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1+\hat{\alpha})X_{t-1} + \hat{\alpha}(1+\hat{\alpha})\hat{\mu}]}{(1-\hat{\alpha}^2)(n-1)}. \quad (4.45)
$$

From Theorem 4.2.3 , and (4.45)

$$
\begin{aligned}
&\frac{(2\hat{\beta}^2 + \hat{\theta}^2)(1 - \hat{\alpha} + \hat{\theta} + \hat{\alpha}^2\hat{\theta}) + \hat{\beta}\hat{\theta}(4 - 4\hat{\alpha} + 3\hat{\theta} + 3\hat{\alpha}^2\hat{\theta})}{(\hat{\alpha}-1)^2(1+\hat{\alpha})\hat{\theta}^2(\hat{\beta}+\hat{\theta})^2} \\
&= \hat{\sigma}^2_{CLS} = \frac{\sum_{t=2}^{n} [(X_t - \hat{\alpha}X_{t-1} - \hat{\mu}(1-\hat{\alpha}))^2 - \hat{\alpha}(1+\hat{\alpha})X_{t-1} + \hat{\alpha}(1+\hat{\alpha})\hat{\mu}]}{(1-\hat{\alpha}^2)(n-1)}.
\end{aligned}
$$

4.2.2.2 The Yule-Walker estimation

In this part, the Yule-Walker estimation for α, μ and σ^2 are obtained. By using Definition 2.3.3 then the sample autocovariance function of X_t

$$\hat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(X_t - \bar{X})(X_{t+k} - \bar{X}), \tag{4.46}$$

where $0 \leq k < n$ and $\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$ is the sample mean.
From the Yule-Walker equation defined in Definition 2.5.2 and equation (3.28), the Yule-Walker estimator of α is

$$\hat{\alpha}_{YW} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\sum_{t=2}^{n}(X_t - \bar{X})(X_{t-1} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}.$$

Consider $\mu = E(X_t)$ defined in Theorem 4.2.2 and $\sigma^2 = Var(X_t)$ defined in (4.30) and note that $S^2 = \frac{\sum_{t=1}^{n}(X_t - \bar{X})^2}{n-1}$. The Yule-Walker estimators of μ and σ^2 are

$$\hat{\mu}_{YW} = \bar{X} = \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})(1 - \hat{\alpha})},$$

$$\hat{\sigma^2}_{YW} = S^2 = \frac{(2\hat{\beta}^2 + \hat{\theta}^2)(1 - \hat{\alpha} + \hat{\theta} + \hat{\alpha}^2\hat{\theta}) + \hat{\beta}\hat{\theta}(4 - 4\hat{\alpha} + 3\hat{\theta} + 3\hat{\alpha}^2\hat{\theta})}{(\hat{\alpha} - 1)^2(1 + \hat{\alpha})\hat{\theta}^2(\hat{\beta} + \hat{\theta})^2},$$

respectively.

4.2.3 Simulation Results

In this section, we produce 1,000 samples from the $NLINARN(1)$ model for true parameter values in different settings (1) $\alpha = 0.1, \beta = 0.1, \theta = 0.5$; (2) $\alpha = 0.4, \beta = 0.3, \theta = 0.8$; (3) $\alpha = 0.7, \beta = 0.5, \theta = 1.5$ of different sample sizes $n = 50, 500, 1000$ by using the statistical software R and obtain estimators of parameters from two methods described in Section 4.2.2. Then we compare the obtained estimators in terms of their

$(\alpha,\theta,\beta,\mu,\sigma^2) = (0.1, 0.5, 0.1, 2.5926, 7.9199)$						
n	$\hat{\alpha}_{CLS}$	$\hat{\mu}_{CLS}$	$\hat{\sigma}^2_{CLS}$	$\hat{\alpha}_{YW}$	$\hat{\mu}_{YW}$	$\hat{\sigma}^2_{YW}$
50	0.0957 (0.0191)	3.6178 (0.2886)	10.7572 (12.4924)	0.0706 (0.0193)	3.6174 (0.2801)	10.9366 (12.7226)
500	0.0951 (0.0091)	2.5492 (0.0934)	7.6828 (4.4768)	0.0966 (0.0021)	2.5825 (0.0172)	7.8631 (0.9130)
1000	0.0994 (0.0011)	2.5891 (0.0096)	7.8796 (0.4493)	0.0986 (0.0011)	2.5900 (0.0095)	7.8860 (0.4493)
$(\alpha,\theta,\beta,\mu,\sigma^2) = (0.4, 0.8, 0.3, 2.6515, 6.3980)$						
n	$\hat{\alpha}_{CLS}$	$\hat{\mu}_{CLS}$	$\hat{\sigma}^2_{CLS}$	$\hat{\alpha}_{YW}$	$\hat{\mu}_{YW}$	$\hat{\sigma}^2_{YW}$
50	0.3677 (0.0181)	2.6110 (0.2911)	6.0872 (6.0744)	0.3442 (0.0186)	2.6207 (0.2770)	6.1152 (5.7594)
500	0.3805 (0.0120)	2.6275 (0.1519)	6.2051 (3.0989)	0.3956 (0.0025)	2.6407 (0.0301)	6.3318 (0.6778)
1000	0.3980 (0.0011)	2.6495 (0.0157)	6.3937 (0.3335)	0.3970 (0.0011)	2.6499 (0.0157)	6.3952 (0.3324)
$(\alpha,\theta,\beta,\mu,\sigma^2) = (0.7, 1.5, 0.5, 2.7778, 9.3682)$						
n	$\hat{\alpha}_{CLS}$	$\hat{\mu}_{CLS}$	$\hat{\sigma}^2_{CLS}$	$\hat{\alpha}_{YW}$	$\hat{\mu}_{YW}$	$\hat{\sigma}^2_{YW}$
50	0.6050 (0.0201)	2.6025 (1.1181)	8.8744 (181.3861)	0.5815 (0.0204)	2.6038 (0.8418)	8.8838 (28.7368)
500	0.6485 (0.0102)	2.7759 (0.5445)	8.9004 (23.3691)	0.6854 (0.00219)	2.7777 (0.1003)	9.1168 (4.3118)
1000	0.6922 (0.0011)	2.7698 (0.0444)	9.2245 (2.1964)	0.6912 (0.0011)	2.7699 (0.0439)	9.2890 (2.1828)

Table 4.2: Mean and variance (in brackets) of the estimators for different values of the parameters α, μ and σ^2 for the $NLINARN(1)$ model

means and variances. Table 4.2 shows mean and variance (in brackets) of the estimators for different values of the parameters α, μ and σ^2.

From Table 4.2, we observe that the estimators obtained from the two estimation methods converge to the true parameters. In addition, increasing the sample size yields smaller variance. The conditional least squares estimate (CLS) and the Yule-Walker estimate (YW) are approximately the same. Considering the variance, we can see that the CLS estimators have smaller variance than the YW estimators for the parameter α. However, the YW has smaller variance than the CLS for the parameters σ^2 and μ. Then, considering the mean, we can see that the YW estimators converge to the true parameters faster than the CLS estimators for the parameters σ^2 and μ.

4.2.4 Real data

In this section, we apply the two models with two real data sets : (1) the incidents of acute febrile mucocutaneous lymph node syndrome and (2) the numbers of earthquakes per year magnitude 7.0 or greater (1900-1998).

4.2.4.1 The incidents of acute febrile mucocutaneous lymph node syndrome (MCLS)

The first example considers the data give weekly counts of the incidents of acute febrile mucocutaneous lymph node syndrome (MCLS) in Totori-prefecture, Japan, during 1982. Sample mean and variance are 1.711 and 3.111, respectively. The fitted $NLINARN(1)$ model is

$$X_t = 0.5241 * X_{t-1} + \varepsilon_t,$$

The predicted values of the numbers of MCLS series are given by

$$\hat{X}_1 = \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})(1 - \hat{\alpha})} = 1.6843,$$
$$\hat{X}_t = \hat{\alpha}\hat{X}_{t-1} + \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})}.$$

Substituting parameter estimates $\hat{\alpha} = 0.5241$ and $\dfrac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})} = 0.8016.$

$$\hat{X}_t = 0.5241\hat{X}_{t-1} + 0.8016, \quad t = 2, 3, \ldots, 52.$$

The expectation and variance computed from the $NLINARN(1)$ model are 1.684 and 3.111, respectively. We can see that the model can capture the sample mean and variance of the data set. Therefore, the model is reasonable to this data set.

4.2.4.2 The number of earthquakes per year magnitude 7.0 or greater (1900-1998).

The second example considers the numbers of earthquakes per year magnitude 7.0 or greater (1900-1998). Sample mean and variance are 20.02 and 52.75, respectively. The fitted $NLINARN(1)$ model is

$$X_t = 0.5434 * X_{t-1} + \varepsilon_t,$$

The predicted values of the numbers of earthquakes per year magnitude 7.0 or greater series are given by

$$\hat{X}_1 = \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})(1 - \hat{\alpha})} = 20.1283,$$
$$\hat{X}_t = \hat{\alpha}\hat{X}_{t-1} + \frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})},$$

Substituting parameter estimates $\hat{\alpha} = 0.5434$ and $\frac{2\hat{\beta} + \hat{\theta}}{\hat{\theta}(\hat{\beta} + \hat{\theta})} = 9.1906.$

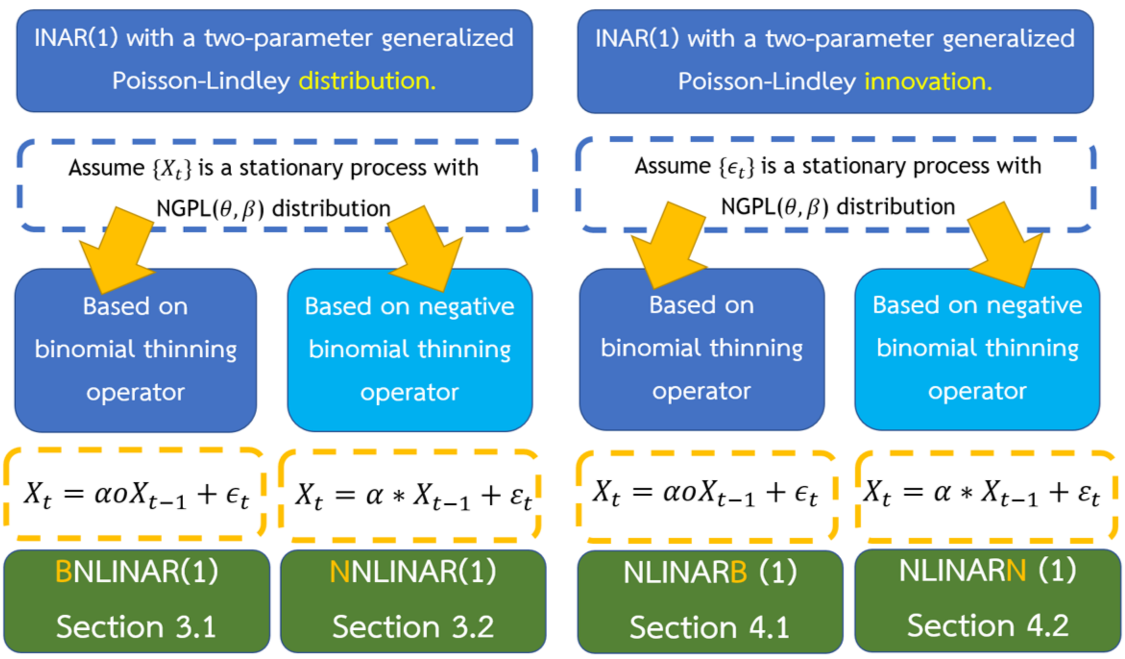
$$\hat{X}_t = 0.5434\hat{X}_{t-1} + 9.1906, \quad t = 2, 3, \ldots, 99.$$

The expectation and variance computed from the $NLINARN(1)$ model are 20.128 and 52.755, respectively. We can see that the model can capture the sample mean and variance of the data set. Therefore, the model is reasonable to this data set.

CHAPTER V

CONCLUSIONS

In this work, we apply the new generalized Poisson-Lindley distribution to construct four new autoregressive model. The first model is the first order integer-valued autoregressive model with a two-parameter generalized Poisson-Lindley distribution based on the binomial thinning operator. The second model is the first order integer-valued autoregressive model with a two-parameter generalized Poisson-Lindley distribution based on the negative binomial thinning operator. The third model is the first order integer-valued autoregressive model with a two-parameter generalized Poisson-Lindley innovation based on the binomial thinning operator. The fourth model is the first order integer-valued autoregressive model with a two-parameter generalized Poisson-Lindley innovation based on the negative binomial thinning operator. A summary of models discussed in this thesis is given below.



For each of these models from diagram, we have derived the probability mass function of the innovation and also some many properties of these models such as autocorrelation functions, multi-step ahead conditional expectation, variance and partial autocorrelation function. Moreover, we discussed estimations of the unknown parameters of the models by using the conditional least squares estimator (CLS) and the Yule-Walker estimator (YW). The estimators are compared via Monte Carlo simulations in terms of their means and variances. Applications of the models for real count time series were also discussed.

REFERENCES

[1] Aghababaei, J., Jones, M. and Lai, C.D. First-order integer valued AR processes with zero inflated Poisson innovations. *Journal of Time Series Analysis*, 33(6): 954–963, 2012.

[2] Alzaid, A.A., Abdulhamid, A. and Omair, A. Poisson difference integer valued autoregressive model of order one. *Bulletin of the Malaysian Mathematical Sciences Society*, 37(2):465–485, 2014.

[3] Alzaid, A.A. and Al-Osh, M.A. First-order integer-valued autoregressive (INAR(1)) process: distributional and regression properties. *Statistica Neerlandica*, 42:53–61, 1988.

[4] Alzaid, A.A. and Al-Osh, M.A. Some autoregressive moving average processes with generalized Poisson marginal distributions. *Annals of the Institute of Statistical Mathematics*, 45:223–232, 1993.

[5] Bakouch, H.S., Mohammadpour, M. and Shirozhan, M. Poisson-Lindley INAR(1) model with applications. *Brazilian Journal of Probability and Statistics*, 32(2): 262–280, 2018.

[6] Bakouch, H.S. Ristic, M.M. and Nastic, A.S. A new geometric first order integer-valued autoregressive (NGINAR(1)) process. *Statistical Planning and Inference*, 139:2218–2226, 2009.

[7] Bhati, D., Sastry, D. and Qadri, P. A new generalized Poisson-Lindley distribution: Applications and properties. *Austrian Journal of Statistics*, 44(4):35–51, 2015.

[8] Kurt, B. The number of traded shares: A time series modelling approach. *Umea Economic Studies*, 860, 2013.

[9] Mahmoudi, E., Rostami, A. and Roozegar, R. A new integer-valued AR(1) process with Poisson-Lindley innovation. *Statistics Applications*, Accepted:https://arxiv.org/abs/1802.00994v1, 2018.

[10] McKenzie, E. Some simple models for discrete variate time series. *JWater Resources Bulletin*, 21:645–650, 1985.

[11] R Development Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*, :ISBN 3–900051–07–0, URL http://www.R–project.org.

[12] Sankaran, M. The discrete Poisson-Lindley distribution. *Biometrics*, 26:145–149, 1970.

[13] Shanbhag, D., McKenzie, E. and N.(ed) et al. Discrete variate time series stochastic processes: Modeling and simulation. *North Holland, Handbook of Statistics*, 21: 573–606, 2003.

[14] Shanker, R. and Sharma, S. A two-parameter Lindley distribution for modeling waiting and survival times data. *Applied Mathematics*, 4:363–368, 2013.

[15] Steutel, F.W. and Van Harn, K. Discrete analogues of self-decomposability and stability. *Annals of Probability*, 7:893–899, 1979.

[16] Xiang, H., Nannan, Y. and Mi, C. Risk aggregation based on the Poisson INAR(1) process with periodic structure. *Lithuanian Mathematical Journal*, 58:505–515, 2017.

APPENDIX

VITA

Name	Miss Suchiraporn Bunyaris
Date of Birth	6 April 1995
Place of Birth	Chachoengsao, Thailand
Education	B.Sc. (Mathematics, Second Class Honors), Chulalongkorn University, 2016