โครงการ

การเรียนการสอนเพื่อเสริมประสบการณ์

ชื่อโครงการ การศึกษาตัวแบบค่าจำนวนเต็มการเฉลี่ยเคลื่อนที่บนฐานของตัวดำเนินการทวินามแบบวางนัยทั่วไป

A study on an integer-valued moving average model based on generalized Binomial thinning operator

ชื่อนิสิต นายนำชัย ผิวอ่อน **เลขประจำตัว** 5833529323

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์

ปีการศึกษา 2561

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การศึกษาตัวแบบค่าจำนวนเต็มการเฉลี่ยเคลื่อนที่บนฐานของตัวดำเนินการทวินามแบบวางนัยทั่วไป

นายนำชัย ผิวอ่อน

โครงงานนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2561

A study on an integer-valued moving average models based on generalized Binomial thinning operator

Mr. Namchai Pew-On

A Project Submitted in Partial Fulfillment of the Requirements
for the Degree of Bachelor of Science Program in Mathematics
Department of Mathematics and Computer Science
Faculty of Science
Chulalongkorn University
Academic Year 2018

โครงงาน	การศึกษาตัวแบบค่าจำนวนเต็มการเฉลี่ยเคลื่อนที่บนฐานของตัวดำเนินการทวินามแบบวางนัยทั่วไป
โดย	นายนำชัย ผิวอ่อน รหัสประจำตัวนิสิต 5833529323
สาขาวิชา	คณิตศาสตร์
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.จิราพรรณ สุนทรโชติ

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับโครงงานฉบับนี้เป็นส่วนหนึ่ง ของการศึกษาตามหลักสูตรปริญญาบัณฑิต ในรายวิชา 2301499 โครงงานวิทยาศาสตร์ (Senior Project)

.... หัวหน้าภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

(ศาสตราจารย์ ดร.กฤษณะ เนียมมณี)

คณะกรรมการสอบโครงงาน

.... อาจารย์ที่ปรึกษาโครงงานหลัก

(ผู้ช่วยศาสตราจารย์ ดร.จิราพรรณ สุนทรโชติ)

.... กรรมการ

(ผู้ช่วยศาสตราจารย์ สุรชัย สมบัติบริบูรณ์)

.... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.รตินันท์ บุญเคลือบ)

นำชัย ผิวอ่อน : การศึกษาตัวแบบค่าจำนวนเต็มการเฉลี่ยเคลื่อนที่บนฐานของตัวดำเนินการทวินามแบบวางนัยทั่วไป. (A STUDY ON AN INTEGER-VALUED MOVING AVERAGE MODELS BASED ON GENERALIZED BINOMIAL THINNING OPERATOR) อ.ที่ปรึกษาโครงงานหลัก : ผู้ช่วยศาสตราจารย์ ดร.จิราพรรณ สุนทรโชติ, 54 หน้า.

ในปี พ.ศ.2556 ริสติกและคณะได้นำเสนอตัวดำเนินการทวินามแบบวางนัยทั่วไปและตัวแบบค่าจำนวนเต็มถดถอยในตัวแบบใหม่ โดยลดเงื่อนไขของความเป็นอิสระต่อกันของตัวดำเนินการทวินาม ในโครงงานนี้เราจะขยายแนวคิดดังกล่าวเพื่อสร้างตัวแบบค่าจำนวนเต็มการเฉลี่ยเคลื่อนที่แบบใหม่โดยใช้ตัวดำเนินการทวินามแบบวางนัยทั่วไป

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์ ลายมือชื่อนิสิต ...นำชัย ผิวอ่อน...

สาขาวิชา ..คณิตศาสตร์.. ลายมือชื่อ อ.ที่ปรึกษาโครงงานหลัก ....

ปีการศึกษา .....2561.....

5833529323 : MAJOR MATHEMATICS.
KEYWORDS : BINOMIAL THINING OPERATOR / INTEGER-VALUED AUTOREGRESSIVE MODEL

NAMCHAI PEW-ON : A STUDY ON AN INTEGER-VALUED MOVING AVERAGE MODELS BASED ON GENERALIZED BINOMIAL THINNING OPERATOR. ADVISOR : ASST. PROF. JIRAPHAN SUNTORNCHOST, Ph.D., 54 pp.

Ristic et al. (2013) introduced a generalized binomial thinning operater and a new integer-valued autoregressive model, by relaxing the independence assumption of the binomial thinning operator. In this project, we extend the generalized binomial thinning operator to construct a new integer-valued moving average model.

Department ..Mathematics and Computer Science.. Student's Signature ..Namchai Pew-on..
Field of Study ...Mathematics... Advisor's Signature ..J. Suntornchost..
Academic Year ....2018.......

Acknowledgments

I would like to express my sincere thanks to my project advisor, Asst. Prof. Dr. Jiraphan Suntornchost for her invaluable help and constant encouragement throughout the course of this project. I am most grateful for her teaching and advice. I would not have achieved this far and this project would not have been completed without all the support that I have always received from her. Moreover, I would like to thank Asst. Prof. Dr. Ratinan Boomklurb and Asst. Prof. Surachai Sombatboriboon for thier suggestions and comments.

This project is funded by Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University.

Contents

List of Figures

Chapter 1

Introduction

The integer-valued time series models for time series count data play important roles in many applications. For examples, claim counts in insurance business and the number of stock transactions in stock market. The model was first introduced by McKenzie (1985) for the lag-one dependence model which is referred nowadays as the integer-valued autoregressive of order one (INAR(1)) process.

The original integer-valued time series model is based on a binomial thinning operator which is a compound sum of independent and identically distributed Bernoulli random variables. Later, Ristic et al. (2013) generalized the binomial thinning operator to construct a new INAR model, by relaxing the assumption of independence in the compound sum to a sum of dependence Bernoulli random variables, called the generalized binomial thinning operator. The new thinning operator can be applied to a wider class of applications. For example, survival or collapse of some companies in economy discussed in Ristic et al. (2013) since all companies operate in the same macroeconomic and may affect to each others. Therefore, such problem is more reasonable to use the dependent thinning operator than the independent thinning operator.

In this study, we extend the dependent binomial thinning operator introduced in Ristic et al. (2013) to construct a generalized integer-valued moving average model.

Chapter 2

Preliminary

In this chapter, we provide some basic concepts, definitions and theorems in probability theory used in our studies.

2.1 Basic of Probability Theory [Prasanna S. (2013)]

Definition 2.1. The set of all possible outcomes of an experiment, denoted by Ω, is called the sample space of the experiment.

Definition 2.2. Consider a random experiment whose sample space is Ω. A random variable X is a function from the sample space Ω into the set of real numbers $\mathbb{R}$ such that for interval I in $\mathbb{R}$, the set $\{s \in \Omega \mid X(s) \in I\}$ is an event in Ω. If the sample space is either finite or countably infinite, the random variable is said to be **discrete**.

Definition 2.3. Let Ω be the sample space of a random experiment. A probability measure $P : \mathcal{P}(\Omega) \to [0, 1]$ is a set function which assigns real numbers to the various events of Ω satisfying

1) $0 \leq P(A) \leq 1$ for any event $A \in \mathcal{P}(\Omega)$.

2) $P(\Omega) = 1$.

3) $P(\emptyset) = 0$.

Theorem 2.4. Let $\{A_1, A_2, A_3, \ldots, A_n\}$ be a finite collection of n events such that $A_i \cap A_j = \emptyset$ for $i \neq j$, then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

Theorem 2.5. Let A be any event of sample space Ω. Then,

$$P(A^c) = 1 - P(A).$$

Theorem 2.6. If A and B are any two events. Then,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Definition 2.7. Let Ω be a sample space. The conditional probability of an event A, given that event B has occured, is defined by

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

provided that $P(B) > 0$.

Definition 2.8. Two events A and B of sample space Ω are said to be independent if and only if

$$P(A \cap B) = P(A)P(B).$$

Theorem 2.9. Let $A, B \subseteq \Omega$. If A and B are independent and $P(B) > 0$, then

$$P(A \mid B) = P(A).$$

Definition 2.10. The set $\{x \in \mathbb{R} \mid x = X(s), s \in \Omega\}$ is called the space of random variable X.

Definition 2.11. Let Ω be the sample space of the random variable X. The function $f : \Omega \to \mathbb{R}$ defined by

$$f(x) = P(X = x)$$

is called the probability mass function (pdf) of X.

Theorem 2.12. If X is a discrete random variable with the sample space Ω and probability mass function $f(x)$, then

1) $f(x) \geq 0$ for all $x \in \Omega$, and

2) $\sum_{x\in\Omega} f(x) = 1$.

Definition 2.13. The cumulative distribution function $F(x)$ of a random variable X is defined by

$$F(x) = P(X \leq x)$$

for any real number x.

Definition 2.14. Let X and Y be two discrete random variables. Then, X and Y are said to be independent if

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

Expected Value of Discrete Random Variables

Definition 2.15. Let X be a numerically-valued discrete random variable with sample space Ω and probability mass function $f(x)$. The expected value of X, $E(X)$, is defined by

$$E(X) = \sum_{x\in\Omega} xf(x) \ ,$$

provided that this sum converges absolutely. We often refer to the expected value as mean, and denote $E(X)$ by μ for short. If the above sum does not converge absolutely, then we say that X does not have an expected value.

Theorem 2.16. If X is a discrete random variable with sample space Ω and distribution function $f(x)$, and if g: $\Omega \longrightarrow \mathbb{R}$ is a function. Then,

$$E(g(X)) = \sum_{x\in\Omega} g(x)f(x) \ ,$$

provided the series converges absolutely.

Theorem 2.17. Let X and Y be random variables with finite expected values. Then,

$$E(X+Y) = E(X) + E(Y) \ ,$$

and if c is a constant. Then,

$$E(cX) = cE(X).$$

Theorem 2.18. Let X and Y be independent random variables with finite expected values. Then,

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

Theorem 2.19. If F is any event and X is a random variable with sample space $\Omega = \{x_1, x_2, x_3 \ldots\}$, then the conditional expectation of X given F is defined by

$$E(X \mid F) = \sum_j x_j P(X = x_j \mid F).$$

Theorem 2.20. Let X and Y be two random variables. Then,

$$E(X) = E(E(X \mid Y)).$$

Variance of Discrete Random Variables

Definition 2.21. Let X be a numerically-valued random variable with expected value $\mu = E(X)$. Then, the variance of X, denote by $Var(X)$, is

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

Theorem 2.22. Let X be any random variable with $E(X) = \mu$. Then,

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

Theorem 2.23. Let X be any random variable and c is any constant. Then,

$$Var(cX) = c^2 Var(X)$$

and

$$Var(X+c) = Var(X).$$

Theorem 2.24. Let X and Y be two independent random variables. Then,

$$Var(X+Y) = Var(X) + Var(Y).$$

Theorem 2.25. Let X and Y be two random variables. Then,

$$Var(X) = E(Var(X \mid Y)) + Var(E(X \mid Y)).$$

Covariance and Correlation

Definition 2.26. Let X and Y be two random variables with expected values μ_X and μ_Y, respectively. The covariance $Cov(X,Y)$ is defined by

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)].$$

Corollary 2.27. Let X and Y be two random variables with finite expected values. Then,

$$Cov(X,Y) = E(XY) - E(X)E(Y).$$

Corollary 2.28. Let X and Y be two independent random variables. Then,

$$Cov(X,Y) = 0.$$

Theorem 2.29. Let X and Y be two random variables. Then,

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y).$$

Definition 2.30. Let X and Y be two random variables with positive variances. The correlation of X and Y is defined as

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}.$$

Generating Function

Definition 2.31. Let X be a random variable. The generating function of X, denote by G_X, is

$$G_X(t) = E(t^X), \quad \text{for } t \in \mathbb{R}.$$

Theorem 2.32. Let X and Y be two independent random variables. Then,

$$G_{X+Y}(t) = G_X(t)G_Y(t), \quad \text{for } t \in \mathbb{R}.$$

Theorem 2.33. Let $G_X(t)$ and $G_Y(t)$ be generating functions of X and Y, respectively. Then

$$G_X(t) = G_Y(t), \quad \forall t \in \mathbb{R}$$

if and only if X and Y have the same distribution.

Moment Generating Function

Definition 2.34. Let X be a random variable. The moment generating function of X, denote by M_X, is

$$M_X(t) = E(e^{tX}), \quad \text{for } t \in \mathbb{R}.$$

Theorem 2.35. Let X be a random variable, and a and b are constants. Then,

$$M_{aX+b}(t) = e^{bt}M_X(at), \quad \text{for } t \in \mathbb{R}.$$

Theorem 2.36. Let X and Y be two indepentdent random variables. Then,

$$M_{X+Y}(t) = M_X(t)M_Y(t), \quad \text{for } t \in \mathbb{R}.$$

Theorem 2.37. Let $M_X(t)$ and $M_Y(t)$ are moment generating functions of X and Y, respectively. Then,

$$M_X(t) = M_Y(t), \quad \forall t \in \mathbb{R}$$

if and only if X and Y have the same distribution.

Mixture Distribution

Definition 2.38. Let a finite set of probability mass functions $f_1(x), f_2(x), f_3(x)$ $, \ldots, f_n(x)$, or cumulative distribution functions $F_1(x), F_2(x), F_3(x), \ldots, F_n(x)$ and weights $w_1, w_2, w_3, \ldots, w_n$ such that $w_i \geq 0$ and $\sum_{i=1}^{n} w_i = 1$, the mixture distribution is defined by

$$F(x) = \sum_{i=1}^{n} w_i F_i(x),$$

or, equivalently,

$$f(x) = \sum_{i=1}^{n} w_i f_i(x).$$

Theorem 2.39. Let $X_1, X_2, X_3, \ldots, X_n$ denote random variables, and let X denote a random variable from the mixture distribution. For the gernerating function $G_{X_i}(t),\ i \in \{1, 2, 3, \ldots, n\}$. Then,

$$G_X(t) = \sum_{i=1}^{n} w_i G_{X_i}(t),\ t \in \mathbb{R}.$$

Some Discrete Distributions

Definition 2.40 (Bernoulli distribution). A random variable X has the Bernoulli diistribution with parameter $p \in [0, 1]$ if

$$P(X = k) = p^k (1 - p)^{1-k} \qquad \text{for} \quad k \in \{0, 1\}.$$

We write X as $X \sim \text{Ber}(p)$.

Theorem 2.41. Let X be a random variable having the Bernoulli distribution with parameter $p \in [0,1]$. The properties of X are as follows.

1) $E(X) = p$.

2) $E(X^2) = p$.

3) $Var(X) = p(1-p)$.

4) $G_X(t) = (1-p) + pt, \ t \in \mathbb{R}$.

5) $M_X(t) = (1-p) + pe^t, \ t \in \mathbb{R}$.

Definition 2.42 (Binomial distribution). A random variable X has the Binomial distribution with parameters n and $p \in [0,1]$ if

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{for} \quad k \in \{0, 1, 2, \ldots, n\}.$$

We write X as $X \sim Bi(n, p)$.

Theorem 2.43. Let X be a random variable having the Binomial distribution with parameters n and $p \in [0,1]$. The properties of X are as follows.

1) $E(X) = np$.

2) $E(X^2) = np(1-p+np)$.

3) $Var(X) = np(1-p)$.

4) $G_X(t) = [(1-p) + pt]^n, \ t \in \mathbb{R}$.

5) $M_X(t) = [(1-p) + pe^t]^n, \ t \in \mathbb{R}$.

Definition 2.44 (Poisson distribution). A random variable X has the Poisson distribution with parameter $\lambda > 0$ if

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} \qquad \text{for} \quad k \in \{0, 1, 2, \ldots\}.$$

We write X as $X \sim Poi(\lambda)$.

Theorem 2.45. Let X be a random variable having the Poisson distribution with parameter $\lambda > 0$. The properties of X are as follows.

1) $E(X) = \lambda$.

2) $E(X^2) = \lambda^2 + \lambda$.

3) $Var(X) = \lambda$.

4) $G_X(t) = e^{\lambda(t-1)},\ t \in \mathbb{R}$.

5) $M_X(t) = e^{\lambda(e^t-1)},\ t \in \mathbb{R}$.

Definition 2.46. A sequence of random variables is said to be **independent and identically distributed** (i.i.d.) if all of them have the same distribution and are mutually independent.

Definition 2.47. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. random varibles and N be a nonnegative integer valued random variable and independent of $X_1, X_2, X_3, \ldots$. The randon variable $S_N = \sum_{n=1}^{N} X_n$ is said to be a **compound random variable.**

Theorem 2.48. Let S_N be the compound random variable defined in Definition 2.47 and $E(\mid X_n \mid) < \infty$, $E(N) < \infty$, respectively. Then, the following properties hold.

1) $E(S_N) = E(N)E(X)$.

2) $Var(S_N) = E(N)Var(X) + Var(N)E(X^2)$.

3) $G_{S_N}(t) = G_N(G_X(t)),\ t \in \mathbb{R}$.

4) $M_{S_N}(t) = G_N(M_X(t)),\ t \in \mathbb{R}$.

Definition 2.49. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. random varibles and N be a Poisson random variable which is independent of $X_1, X_2, X_3, \ldots$. The randon variable $S_N = \sum_{n=1}^{N} X_n$ is said to be **compound Poisson random variable.**

Theorem 2.50. Let $\{X_i\}_{i\in\mathbb{N}}$ and $\{Y_i\}_{i\in\mathbb{N}}$ be sequences of i.i.d. Bernoulli random variables with means α_1 and α_2, respectively. If $\{X_i\}_{i\in\mathbb{N}}$ and $\{Y_i\}_{i\in\mathbb{N}}$ are independent, then

$$Cov\left(\sum_{i=1}^{N_1} X_i, \sum_{i=1}^{N_2} Y_i\right) = \alpha_1\alpha_2 Cov(N_1, N_2).$$

Integer-valued Time Series Model

Definition 2.51 (Binomial thinning operator)**.** Let $\{B_i\}_{i\in\mathbb{N}}$, be a counting sequence of i.i.d. Bernoulli random variables with mean $\alpha \in [0,1]$ and $X > 0$ be an integer-valued random variable which is independent of the counting sequence. The binomial thinning operator, $\alpha\circ$, is defined by

$$\alpha \circ X = \sum_{i=1}^{X} B_i.$$

Theorem 2.52 (Properties of the Binomial thinning operator)**.** Let X be a random variable having values in $\mathbb{N}$. The binomial thinning $\alpha\circ X$ has the following properties, for $\alpha \in [0,1]$,

1) $E(\alpha \circ X) = \alpha E(X)$,

2) $Var(\alpha \circ X) = \alpha^2 Var(X) + \alpha(1-\alpha)E(X)$.

Definition 2.53. A process of X_t is stationary if it has time invariant first and second moments, i.e., if for any choice of $t \in \mathbb{N}$, the following conditions hold:

1) $E(X_t) = \mu_X$ with $|\mu_X| < \infty$.

2) $E[(X_t - \mu_X)^2] = \sigma_X^2 < \infty$.

3) $E[(X_t - \mu_X)(X_{t-h} - \mu_X)] = \gamma_h, \ \forall h \in \mathbb{N}$ with $|\gamma_h| < \infty$.

Definition 2.54. [The Integer-Valued Autoregressive Model : INAR(1)] A discrete time non-negative, integer-valued process $\{X_t\}_{t\in\mathbb{N}}$ is called a stationary INAR(1) process, if it follows the recursion

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

where $\epsilon_t, t \in \mathbb{N}$, is a non-negative integer-valued random variable which is independent from X_s for $s \leq t$, with $E(\epsilon_t) = \mu_\epsilon > 0$ and $Var(\epsilon_t) = \sigma_\epsilon^2 \geq 0$.

Theorem 2.55. If $\{X_t\}_{t\in\mathbb{N}}$ is a stationary process defined in Definition 2.54. Then,

1) $E(X_t) = \dfrac{\mu_\epsilon}{1-\alpha}$,

2) $Var(X_t) = \dfrac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1-\alpha^2}$,

3) $E(X_t \mid X_{t-1}) = \alpha X_{t-1} + \mu_\epsilon$,

4) $Var(X_t \mid X_{t-1}) = \alpha(1-\alpha)X_{t-1} + \sigma_\epsilon^2$.

Definition 2.56. [The Integer-Valued Moving Average Model : INMA(1)] A discrete time non-negative, integer-valued process $\{X_t\}_{t\in\mathbb{N}}$ is called INMA(1) process, if it follows the recursion

$$X_t = \alpha \circ \epsilon_{t-1} + \epsilon_t,$$

where $\epsilon_t, t \in \mathbb{N}$, is a non-negative integer-valued random variable which is independent from X_s for $s \leq t$, with $E(\epsilon_t) = \mu_\epsilon > 0$ and $Var(\epsilon_t) = \sigma_\epsilon^2 \geq 0$.

Theorem 2.57. Let $\{X_t\}_{t\in\mathbb{N}}$ be INMA(1) process defined in Definition 2.56. Then,

1) $E(X_t) = (1+\alpha)\mu_\epsilon$,

2) $Var(X_t) = \alpha(1-\alpha)\mu_\epsilon + (1+\alpha^2)\sigma_\epsilon^2$.

Chapter 3

Main Result

The Binomial thinning operator is defined as a compound sum of identically distributed and independent Bernoulli random variables. The thinning operator has been applied to many applications. However, the independence assumption in the thinning operator may not be realistic in some applications. For example, Vagrants usually live together and support themselves to do vagrancy offences. Then, the Binomial thinning operator is not suitable for the model used to count the number of vagrants who do vagrancy offences. Therefore, a more general thinning operator relaxing independence assumption, called generalized Binomial thinning operator, has been proposed in literature. (Ristic et al, 2013)

In this chapter, we apply the generalized binomial thinning operator to construct a generalized integer-valued moving average model. The organization of this chapter is as follows. First we state the generalized Binomial thinning operator and discussion its properties in Section 3.1. We then apply it to construct the generalized integer-valued moving average model with order 1 and general order q in Section 3.2 and Section 3.3, respectively. Finally, numerical results and simulations are given in Section 3.4.

Now, we first introduce the new binomial thinning operator by relaxing the independence assumption Bernoulli random variables, called **generalized Binomial thinning operator**.

3.1 Generalized Binomial Thinning Operater ($\circ_\theta$)

In this section, we consider a sequence of random variables, $\{U_i\}_{i\in\mathbb{N}}$, introduced by Ristic et al. (2013). The random variable U_i defined as

$$U_i = (1 - V_i)W_i + V_i Z, \tag{1}$$

where $\{V_i\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. Bernoulli (θ), $\{W_i\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. Bernoulli (α), Z is Bernoulli (α) and the random variables V_i, W_j, Z are independent for all $i \in \mathbb{N}$ and $j \in \mathbb{N}$.

Theorem 3.1. Let $\{U_i\}_{i\in\mathbb{N}}$ be the sequence of random variables defined in (1). Then, the following properties hold.

1) U_i has the Bernoulli distribution with parameter α.

2) $E(U_iU_j) = \alpha(\alpha + (1-\alpha)\theta^2)$ for $i \neq j$.

3) $Cov(U_i, U_j) = \alpha(1-\alpha)\theta^2$ for $i \neq j$.

Proof. 1) Consider

$$\begin{aligned}
P(U_i = x) &= P(U_i = x \mid V_i = 0)P(V_i = 0) + P(U_i = x \mid V_i = 1)P(V_i = 1) \\
&= P(W_i = x)P(V_i = 0) + P(Z = x)P(V_i = 1) \\
&= \alpha^x(1-\alpha)^{1-x}(1-\theta) + \alpha^x(1-\alpha)^{1-x}\theta \\
&= \alpha^x(1-\alpha)^{1-x}.
\end{aligned}$$

Then, U_i has the Bernoulli distribution with parameter α.

2)

$$\begin{aligned}
E(U_iU_j) &= E\big(((1-V_i)W_i + V_iZ)((1-V_j)W_j + V_jZ)\big) \\
&= E\big((1-V_i)(1-V_j)W_iW_j + (1-V_i)V_jW_iZ + V_i(1-V_j)W_jZ + V_iV_jZ^2\big) \\
&= E\big((1-V_i)(1-V_j)W_iW_j\big) + E\big((1-V_i)V_jW_iZ\big) + E\big(V_i(1-V_j)W_jZ\big) \\
&\quad + E\big(V_iV_jZ^2\big)
\end{aligned}$$

$$
\begin{aligned}
&= E(1-V_i)E(1-V_j)E(W_i)E(W_j) + E(1-V_i)E(V_j)E(W_i)E(Z) \\
&\quad + E(V_i)E(1-V_j)E(W_j)E(Z) + E(V_i)E(V_j)E(Z^2) \\
&= (1-\theta)^2\alpha^2 + (1-\theta)\theta\alpha^2 + \theta(1-\theta)\alpha^2 + \theta^2\alpha \\
&= \alpha(\alpha + (1-\alpha)\theta^2).
\end{aligned}
$$

3) By 2), we have

$$
\begin{aligned}
Cov(U_i, U_j) &= E(U_iU_j) - E(U_i)E(U_j) \\
&= \alpha(\alpha + (1-\alpha)\theta^2) - \alpha^2 \\
&= \alpha^2 + \alpha(1-\alpha)\theta^2 - \alpha^2 \\
&= \alpha(1-\alpha)\theta^2.
\end{aligned}
$$

□

Theorem 3.2. [Ristic et al. (2013)] Let $\{U_i\}_{i\in\mathbb{N}}$ be a sequence of random variables defined in (1). Then the random variable $U_1 + U_2 + U_3 + \cdots + U_n$ is distributed as

$$
U_1 + U_2 + U_3 + \cdots + U_n = \begin{cases} Bi(n, \alpha(1-\theta)), & \text{w.p. } 1-\alpha, \\ Bi(n, \alpha+\theta-\alpha\theta), & \text{w.p. } \alpha, \end{cases}
$$

where w.p. refers to "with probability".

Proof. For $s \in \mathbb{R}$,

$$
\begin{aligned}
E(s^{U_1+U_2+U_3+\cdots+U_n}) &= \sum_{i=1}^{n} \binom{n}{i} (1-\theta)^{n-i}\theta^i E(s^{W_1+W_2+W_3+\cdots+W_{n-i}+iZ}) \\
&= \sum_{i=1}^{n} \binom{n}{i} (1-\theta)^{n-i}\theta^i (1-\alpha+\alpha s)^{n-i}(1-\alpha+\alpha s^i) \\
&= (1-\alpha)(1-\alpha(1-\theta)(1-s))^n + \alpha(1-(\alpha+\theta-\alpha\theta)(1-s))^n.
\end{aligned} \tag{2}
$$

Then, the random variable $U_1 + U_2 + U_3 + \cdots + U_n$ is distributed as follows

$$
U_1 + U_2 + U_3 + \cdots + U_n = \begin{cases} Bi(n, \alpha(1-\theta)), & \text{w.p. } 1-\alpha, \\ Bi(n, \alpha+\theta-\alpha\theta), & \text{w.p. } \alpha. \end{cases}
$$

□

Definition 3.3. [Ristic et al. (2013)] **Generalized Binomial Thinning Operator.** Let $\{U_i\}_{i\in\mathbb{N}}$ be a sequence of Bernoulli random variables and X be a non-negative integer-valued random variable which is independent of the counting sequence. The generalized binomial thinning operator, $\alpha\circ_\theta$, is defined by

$$\alpha \circ_\theta X = \sum_{i=1}^{X} U_i \ ,$$

where $U_i = (1 - V_i)Wi + V_i Z$ is defined in (1).

Next, we state Properties of the Generalized Binomial Thinning operator defined in Definition 3.3 given in Ristic et al. 2013. However, the proof is not provided in the paper. Therefore, we give the proof in the following theorem.

Theorem 3.4. [Ristic et al. (2013)] Properties of Generalized Binomial Thinning Operator.

1) $E(\alpha \circ_\theta X \mid X) = \alpha X$.

2) $E((\alpha \circ_\theta X)^2 \mid X) = \alpha(\alpha + (1-\alpha)\theta^2)X^2 + \alpha(1-\alpha)(1-\theta^2)X$.

3) $Var(\alpha \circ_\theta X \mid X) = \alpha(1-\alpha)(\theta^2 X^2 + (1-\theta^2)X)$.

4) $E(\alpha \circ_\theta X) = \alpha E(X)$.

5) $E(\alpha \circ_\theta X)^2 = \alpha(1-\alpha)(1-\theta^2)E(X) + \alpha(\alpha + (1-\alpha)\theta^2)E(X^2)$.

6) $E(X(\alpha \circ_\theta Y)) = \alpha E(XY)$.

7) $Cov(X, \alpha \circ_\theta Y) = \alpha Cov(X, Y)$.

Proof. 1) Since the random variable $\{U_i\}_{i\in\mathbb{N}}$ have the Bernoulli distribution parameter α,

$$\begin{aligned} E(\alpha \circ_\theta X \mid X) &= E\left(\sum_{i=1}^{X} U_i \mid X\right) \\ &= XE(U_1) \\ &= \alpha X. \end{aligned}$$

2) By 2) of Theorem 3.1, we have

$$
\begin{aligned}
E((\alpha \circ_\theta X)^2 \mid X) &= E\Big((\sum_{i=1}^{X} U_i)^2 \mid X \Big) \\
&= E\Big(\sum_{i=1}^{X} U_i^2 + \sum_{i\neq 1}^{X} U_1 U_i + \sum_{i\neq 2}^{X} U_2 U_i + \cdots + \sum_{i=1}^{X-1} U_X U_i \mid X \Big) \\
&= E\Big(\sum_{i=1}^{X} U_i^2 \mid X \Big) + E\Big(\sum_{i\neq 1}^{X} U_1 U_i \mid X \Big) + E\Big(\sum_{i\neq 2}^{X} U_2 U_i \mid X \Big) + \cdots \\
&\quad + E\Big(\sum_{i=1}^{X-1} U_X U_i \mid X \Big) \\
&= \alpha X + \alpha(\alpha + (1-\alpha)\theta^2)(X-1) + \alpha(\alpha + (1-\alpha)\theta^2)(X-1) + \cdots \\
&\quad + \alpha(\alpha + (1-\alpha)\theta^2)(X-1) \\
&= \alpha X + \alpha(\alpha + (1-\alpha)\theta^2)(X-1)X \\
&= \alpha(\alpha + (1-\alpha)\theta^2)X^2 + \alpha(1-\alpha)(1-\theta^2)X.
\end{aligned}
$$

3) By 1) and 2), we have

$$
\begin{aligned}
Var(\alpha \circ_\theta X \mid X) &= E((\alpha \circ_\theta X)^2 \mid X) - (E(\alpha \circ_\theta X \mid X))^2 \\
&= \alpha(\alpha + (1-\alpha)\theta^2)X^2 + \alpha(1-\alpha)(1-\theta^2)X - (\alpha X)^2 \\
&= \alpha(1-\alpha)(\theta^2 X^2 + (1-\theta^2)X).
\end{aligned}
$$

4) By 1), we have

$$
\begin{aligned}
E(\alpha \circ_\theta X) &= E\big(E(\alpha \circ_\theta X \mid X)\big) \\
&= E(\alpha X) \\
&= \alpha E(X).
\end{aligned}
$$

5) By 2), we have

$$
\begin{aligned}
E((\alpha \circ_\theta X)^2) &= E\big(E((\alpha \circ_\theta X)^2 \mid X)\big) \\
&= E\big(\alpha(\alpha + (1-\alpha)\theta^2)X^2 + \alpha(1-\alpha)(1-\theta^2)X\big) \\
&= E\big(\alpha(\alpha + (1-\alpha)\theta^2)X^2\big) + E\big(\alpha(1-\alpha)(1-\theta^2)X\big) \\
&= \alpha(\alpha + (1-\alpha)\theta^2)E(X^2) + \alpha(1-\alpha)(1-\theta^2)E(X).
\end{aligned}
$$

6) By using Theorem 2.20 and 1), we have

$$\begin{aligned}E(X(\alpha \circ_\theta Y)) &= E\big(E(X(\alpha \circ_\theta Y) \mid X)\big)\\&= E\big(XE(\alpha \circ_\theta Y \mid X)\big)\\&= E\big(XE(E(\alpha \circ_\theta Y \mid Y) \mid X)\big)\\&= E\big(X\alpha E(Y \mid X)\big)\\&= \alpha E\big(E(XY \mid X)\big)\\&= \alpha E(XY).\end{aligned}$$

7) By using Corollary 2.27, 4) and 6), we have

$$\begin{aligned}Cov(X, \alpha \circ_\theta Y) &= E(X(\alpha \circ_\theta Y)) - E(X)E(\alpha \circ_\theta Y)\\&= \alpha E(XY) - \alpha E(X)E(Y)\\&= \alpha\big(E(XY) - E(X)E(Y)\big)\\&= \alpha Cov(X, Y).\end{aligned}$$

□

3.2 Model: DBINMA(1)

Now, we construct the integer-valued moving average model based on the generalized binomial thinning operator order 1, called DBINMA(1) model.

Definition 3.5. DBINMA(1)
A discrete time non-negative, integer-valued process $\{X_t\}_{t\in\mathbb{N}}$ is called a stationary DBINMA(1) process, if it follows the recursion

$$X_t = \alpha \circ_\theta \epsilon_{t-1} + \epsilon_t$$

where $\{\epsilon_t\}_{t\in\mathbb{N}}$ is a sequence of i.i.d. Poisson random variables with mean λ.

Theorem 3.6. Let $\{X_t\}_{t\in\mathbb{N}}$ be the DBINMA(1) defined in Definition 3.5 , then

$$X_t = \begin{cases} Poi(\lambda(\alpha(1-\theta)+1)), & \text{w.p. } 1-\alpha \ , \\ Poi(\lambda(\theta+\alpha(1-\theta)+1)), & \text{w.p. } \alpha. \end{cases}$$

Proof. By (2) of Theorem 3.2 and Definition 3.3, for $s \in \mathbb{R}$, we have

$$\begin{aligned} G_{X_t}(s) &= E(s^{\alpha\circ_\theta\epsilon_{t-1}+\epsilon_t}) \\ &= E(E(s^{\sum_{i=1}^{\epsilon_{t-1}} U_i+\epsilon_t} \mid \epsilon_{t-1})) \\ &= E(E(s^{\sum_{i=1}^{\epsilon_{t-1}} U_i} \mid \epsilon_{t-1})) \cdot E(s^{\epsilon_t}) \\ &= E\big((1-\alpha)(1-\alpha(1-\theta)(1-s))^{\epsilon_{t-1}} + \alpha(1-(\alpha+\theta-\alpha\theta)(1-s))^{\epsilon_{t-1}}\big) \\ &\quad \cdot G_{\epsilon_t}(s) \\ &= \big[(1-\alpha)G_{\epsilon_{t-1}}(1-\alpha(1-\theta)(1-s)) + \alpha G_{\epsilon_{t-1}}(1-(\alpha+\theta-\alpha\theta)(1-s)\big] \\ &\quad \cdot G_{\epsilon_t}(s) \\ &= \big[(1-\alpha)e^{\lambda[\alpha(1-\theta)](s-1)]} + \alpha e^{\lambda[\theta+\alpha(1-\theta)](s-1)]}\big] \cdot e^{\lambda(s-1)} \\ &= (1-\alpha)e^{\lambda[\alpha(1-\theta)+1](s-1)} + \alpha e^{\lambda[\theta+\alpha(1-\theta)+1](s-1)}. \end{aligned}$$

From Theorem 2.39, we have

$$X_t = \begin{cases} Poi(\lambda(\alpha(1-\theta)+1)), & \text{w.p. } 1-\alpha \ , \\ Poi(\lambda(\theta+\alpha(1-\theta)+1)), & \text{w.p. } \alpha. \end{cases}$$

□

Theorem 3.7. The Properties of the DBINMA(1) defined in Definition 3.5 are as follows

1) $E(X_t) = \lambda(\alpha + 1)$.

2) $Var(X_t) = \lambda(1 + \alpha + \alpha(1 - \alpha)\theta^2)$.

3) $Cov(X_t, X_{t-k}) = \begin{cases} \alpha\lambda, & k = 1, \\ 0, & k > 1. \end{cases}$

4) $Corr(X_t, X_{t-k}) = \begin{cases} \dfrac{\alpha}{1 + \alpha + \alpha(1 - \alpha)\theta^2}, & k = 1, \\ 0, & k > 1. \end{cases}$

Proof. 1)

$$\begin{aligned} E(X_t) &= E(\alpha \circ_\theta \epsilon_{t-1} + \epsilon_t) \\ &= E(\alpha \circ_\theta \epsilon_{t-1}) + E(\epsilon_t) \\ &= \alpha E(\epsilon_{t-1}) + E(\epsilon_t) \\ &= \alpha\lambda + \lambda \\ &= \lambda(\alpha + 1). \end{aligned}$$

2) By 4) and 5) of Theorem 3.4, we have

$$\begin{aligned} Var(X_t) &= Var(\alpha \circ_\theta \epsilon_{t-1} + \epsilon_t) \\ &= Var(\alpha \circ_\theta \epsilon_{t-1}) + Var(\epsilon_t) \\ &= E((\alpha \circ_\theta \epsilon_{t-1})^2) - (E(\alpha \circ_\theta \epsilon_{t-1}))^2 + Var(\epsilon_t) \\ &= \alpha(1 - \alpha)(1 - \theta^2)E(\epsilon_{t-1}) + \alpha(\alpha + (1 - \alpha)\theta^2)E(\epsilon_{t-1}^2) - (E(\alpha \circ_\theta \epsilon_{t-1}))^2 \\ &\quad + Var(\epsilon_t) \\ &= \alpha(1 - \alpha)(1 - \theta^2)\lambda + \alpha(\alpha + (1 - \alpha)\theta^2)(\lambda^2 + \lambda) - (\lambda\alpha)^2 + \lambda \\ &= \lambda(1 + \alpha + \alpha(1 - \alpha)\theta^2). \end{aligned}$$

3) Since $\{\epsilon_t\}_{t\in\mathbb{N}}$ is a sequence of i.i.d. Poisson random variables with mean λ and 1), we have

$$\begin{aligned}
Cov(X_t, X_{t-k}) &= E(X_t X_{t-k}) - E(X_t)E(X_{t-k}) \\
&= E\big((\alpha \circ_\theta \epsilon_{t-1} + \epsilon_t)(\alpha \circ_\theta \epsilon_{t-k-1} + \epsilon_{t-k})\big) - E(X_t)E(X_{t-k}) \\
&= E\big((\alpha \circ_\theta \epsilon_{t-1})(\alpha \circ_\theta \epsilon_{t-k-1}) + (\alpha \circ_\theta \epsilon_{t-1})(\epsilon_{t-k}) + (\alpha \circ_\theta \epsilon_{t-k-1})(\epsilon_t) \\
&\quad + \epsilon_t \epsilon_{t-k}\big) - E(X_t)E(X_{t-k}) \\
&= E\big((\alpha \circ_\theta \epsilon_{t-1})(\alpha \circ_\theta \epsilon_{t-k-1})\big) + E\big((\alpha \circ_\theta \epsilon_{t-1})(\epsilon_{t-k})\big) \\
&\quad + E\big((\alpha \circ_\theta \epsilon_{t-k-1})(\epsilon_t)\big) + E(\epsilon_t \epsilon_{t-k}) - E(X_t)E(X_{t-k}) \\
&= E(\alpha \circ_\theta \epsilon_{t-1})E(\alpha \circ_\theta \epsilon_{t-k-1}) + E\big((\alpha \circ_\theta \epsilon_{t-1})(\epsilon_{t-k})\big) \\
&\quad + E(\alpha \circ_\theta \epsilon_{t-k-1})E(\epsilon_t) + E(\epsilon_t)E(\epsilon_{t-k}) - E(X_t)E(X_{t-k}) \\
&= (\alpha\lambda)(\alpha\lambda) + (\alpha\lambda)\lambda + \lambda^2 - (\lambda(\alpha+1))^2 + E\big((\alpha \circ_\theta \epsilon_{t-1})(\epsilon_{t-k})\big) \\
&= E\big((\alpha \circ_\theta \epsilon_{t-1})(\epsilon_{t-k})\big) - \alpha\lambda^2.
\end{aligned}$$

If $k = 1$, then by 6), we have

$$\begin{aligned}
Cov(X_t, X_{t-k}) &= E\big((\alpha \circ_\theta \epsilon_{t-1})(\epsilon_{t-1})\big) - \alpha\lambda^2 \\
&= \alpha E(\epsilon_{t-1}^2) - \alpha\lambda^2 \\
&= \alpha(\lambda^2 + \lambda) - \alpha\lambda^2 \\
&= \alpha\lambda.
\end{aligned}$$

If $k > 1$, since ϵ_{t-1} and ϵ_{t-k} are independent, then

$$\begin{aligned}
Cov(X_t, X_{t-k}) &= E\big((\alpha \circ_\theta \epsilon_{t-1})(\epsilon_{t-k})\big) - \alpha\lambda^2 \\
&= E(\alpha \circ_\theta \epsilon_{t-1})E(\epsilon_{t-k}) - \alpha\lambda^2 \\
&= (\alpha\lambda)\lambda - \alpha\lambda^2 \\
&= 0.
\end{aligned}$$

4) If $k = 1$, by Definition 2.30, we have

$$\begin{aligned} Corr(X_t.X_{t-1}) &= \frac{Cov(X_t, X_{t-1})}{\sqrt{Var(X_t)Var(X_{t-1})}} \\ &= \frac{\alpha\lambda}{\lambda(1+\alpha+\alpha(1-\alpha)\theta^2)} \\ &= \frac{\alpha}{1+\alpha+\alpha(1-\alpha)\theta^2}. \end{aligned}$$

If $k > 1$, by Definition 2.30, we have

$$Corr(X_t.X_{t-k}) = \frac{Cov(X_t, X_{t-k})}{\sqrt{Var(X_t)Var(X_{t-k})}} = 0.$$

□

3.3 Model: DBINMA(q)

In this section, we extend the DBINMA(1) model studied in the previous section to a more general model by extending the lag-dependence to general order q. The definition of the model is given as follows.

Definition 3.8. DBINMA(q).
A discrete time non-negative, integer-valued process $\{X_t\}_{t\in\mathbb{N}}$ is called a stationary DBINMA(q) process, if it follows the recursion

$$X_t = \alpha_1 \circ_\theta \epsilon_{t-1} + \alpha_2 \circ_\theta \epsilon_{t-2} + \alpha_3 \circ_\theta \epsilon_{t-3} + \cdots + \alpha_q \circ_\theta \epsilon_{t-q} + \epsilon_t$$

where $\{\epsilon_t\}_{t\in\mathbb{N}}$ is an i.i.d. sequence of Poisson random variables with mean λ.

Theorem 3.9. Let $\{X_t\}_{t\in\mathbb{N}}$ be the DBINMA(q) defined in Definition 3.8. Then

$$X_t = \begin{cases} Poi\left(\lambda\left((1-\theta)\sum_{i=1}^{q}\alpha_i + 1\right)\right), & \text{w.p. } \prod_{i=1}^{q}(1-\alpha_i)\ , \\ Poi\left(\lambda\left((1-\theta)\sum_{i=1}^{q}\alpha_i + \theta + 1\right)\right), & \text{w.p. } \sum_{i=1}^{q}\alpha_i\prod_{j\neq i}^{q}(1-\alpha_j), \\ \qquad\vdots \\ Poi\left(\lambda\left((1-\theta)\sum_{i=1}^{q}\alpha_i + q\theta + 1\right)\right), & \text{w.p. } \prod_{i=1}^{q}\alpha_i. \end{cases}$$

Proof. Since $\{\epsilon_t\}_{t\in\mathbb{N}}$ is a sequence of i.i.d. Poisson random variables with mean λ, using Theorem 2.32, for $s \in \mathbb{R}$, we have

$$\begin{aligned}
G_{X_t}(s) &= \prod_{i=1}^{q} G_{\alpha_i \circ_\theta \epsilon_{t-i}}(s) \cdot G_{\epsilon_t}(s) \\
&= e^{\lambda(s-1)} \cdot \prod_{i=1}^{q} \left[(1-\alpha_i) G_{\alpha_i(1-\theta)\circ\epsilon_{t-i}}(s) + \alpha_i G_{(\theta+\alpha_i(1-\theta))\circ\epsilon_{t-i}}(s)\right] \\
&= e^{\lambda(s-1)} \cdot \Bigg[\prod_{i=1}^{q}(1-\alpha_i) e^{\lambda((1-\theta)\sum_{i=1}^{q}\alpha_i)(s-1)} \\
&\quad + \sum_{i=1}^{q}\alpha_i \prod_{i\neq j}^{q}(1-\alpha_j) e^{\lambda((1-\theta)\sum_{i=1}^{q}\alpha_i+\theta)(s-1)} \\
&\quad + \cdots + \prod_{i=1}^{q}\alpha_i e^{\lambda((1-\theta)\sum_{i=1}^{q}\alpha_i + q\theta)(s-1)}\Bigg] \\
&= \Bigg[\prod_{i=1}^{q}(1-\alpha_i) e^{\lambda((1-\theta)\sum_{i=1}^{q}\alpha_i+1)(s-1)} \\
&\quad + \sum_{i=1}^{q}\alpha_i \prod_{i\neq j}^{q}(1-\alpha_j) e^{\lambda((1-\theta)\sum_{i=1}^{q}\alpha_i+\theta+1)(s-1)} \\
&\quad + \cdots + \prod_{i=1}^{q}\alpha_i e^{\lambda((1-\theta)\sum_{i=1}^{q}\alpha_i+q\theta+1)(s-1)}\Bigg].
\end{aligned}$$

From theorem 2.39, we have

$$X_t = \begin{cases}
Poi\left(\lambda\left((1-\theta)\sum_{i=1}^{q}\alpha_i + 1\right)\right), & \text{w.p. } \prod_{i=1}^{q}(1-\alpha_i)\ , \\
Poi\left(\lambda\left((1-\theta)\sum_{i=1}^{q}\alpha_i + \theta + 1\right)\right), & \text{w.p. } \sum_{i=1}^{q}\alpha_i \prod_{j\neq i}^{q}(1-\alpha_j), \\
\qquad\vdots & \\
Poi\left(\lambda\left((1-\theta)\sum_{i=1}^{q}\alpha_i + q\theta + 1\right)\right), & \text{w.p. } \prod_{i=1}^{q}\alpha_i.
\end{cases}$$

□

Theorem 3.10. The Properties of the DBINMA(q) defined in Definition 3.8 are as follows

1) $E(X_t) = \lambda \left(\sum_{i=1}^{q} \alpha_i + 1 \right).$

2) $Var(X_t) = \lambda \left[\lambda\theta^2 \left(\sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{q} \alpha_i^2 \right) + \sum_{i=1}^{q} \alpha_i + 1 \right].$

3) $$Cov(X_t, X_{t-k}) = \begin{cases} (\lambda^2+\lambda)\left(\alpha_1 + \sum_{i=1}^{q-1} \alpha_i\alpha_{i+1}\right), & k=1, \\ (\lambda^2+\lambda)\left(\alpha_2 + \sum_{i=1}^{q-2} \alpha_i\alpha_{i+2}\right), & k=2, \\ \vdots & \\ \alpha_q(\lambda^2+\lambda), & k=q, \\ 0, & k>q. \end{cases}$$

4) $$Corr(X_t, X_{t-k}) = \begin{cases} \dfrac{(\lambda+1)\left(\alpha_1 + \sum_{i=1}^{q-1} \alpha_i\alpha_{i+1}\right)}{\lambda\theta^2 \left(\sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{q} \alpha_i^2 \right) + \sum_{i=1}^{q} \alpha_i + 1}, & k=1, \\ \dfrac{(\lambda+1)\left(\alpha_2 + \sum_{i=1}^{q-2} \alpha_i\alpha_{i+2}\right)}{\lambda\theta^2 \left(\sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{q} \alpha_i^2 \right) + \sum_{i=1}^{q} \alpha_i + 1}, & k=2, \\ \vdots & \\ \dfrac{\alpha_q(\lambda+1)}{\lambda\theta^2 \left(\sum_{i=1}^{q} \alpha_i - \sum_{i=1}^{q} \alpha_i^2 \right) + \sum_{i=1}^{q} \alpha_i + 1}, & k=q, \\ 0, & k>q. \end{cases}$$

Proof. 1)

$$
\begin{aligned}
E(X_t) &= E(\alpha_1 \circ_\theta \epsilon_{t-1} + \alpha_2 \circ_\theta \epsilon_{t-2} + \alpha_3 \circ_\theta \epsilon_{t-3} + \cdots + \alpha_q \circ_\theta \epsilon_{t-q} + \epsilon_t) \\
&= E(\alpha_1 \circ_\theta \epsilon_{t-1}) + E(\alpha_2 \circ_\theta \epsilon_{t-2}) + E(\alpha_3 \circ_\theta \epsilon_{t-3}) + \cdots + E(\alpha_q \circ_\theta \epsilon_{t-q}) \\
&\quad + E(\epsilon_t) \\
&= \alpha_1 E(\epsilon_{t-1}) + \alpha_2 E(\epsilon_{t-2}) + \alpha_3 E(\epsilon_{t-3}) + \cdots + \alpha_q E(\epsilon_{t-q}) + E(\epsilon_t) \\
&= \alpha_1 \lambda + \alpha_2 \lambda + \alpha_3 \lambda + \cdots + \alpha_q \lambda + \lambda \\
&= \lambda\Big(\sum_{i=1}^{q} \alpha_i + 1\Big).
\end{aligned}
$$

2) Claim that $Var(\alpha \circ_\theta \epsilon_t) = \alpha\lambda + \alpha\lambda^2\theta^2 - \alpha^2\lambda^2\theta$.

$$
\begin{aligned}
Var(\alpha \circ_\theta \epsilon_t) &= Var(\alpha \circ_\theta \epsilon_t) \\
&= E((\alpha \circ_\theta \epsilon_t)^2) - (E(\alpha \circ_\theta \epsilon_t))^2 \\
&= \alpha(1-\alpha)(1-\theta^2)E(\epsilon_t) + \alpha(\alpha + (1-\alpha)\theta^2)E(\epsilon_t^2) - (E(\alpha \circ_\theta \epsilon_t))^2 \\
&= \alpha(1-\alpha)(1-\theta^2)\lambda + \alpha(\alpha + (1-\alpha)\theta^2)(\lambda^2 + \lambda) - (\lambda\alpha)^2 \\
&= \alpha\lambda + \alpha\lambda^2\theta^2 - \alpha^2\lambda^2\theta^2.
\end{aligned}
$$

$$
\begin{aligned}
Var(X_t) &= Var(\alpha_1 \circ_\theta \epsilon_{t-1} + \alpha_2 \circ_\theta \epsilon_{t-2} + \alpha_3 \circ_\theta \epsilon_{t-3} + \cdots + \alpha_q \circ_\theta \epsilon_{t-q} + \epsilon_t) \\
&= Var(\alpha_1 \circ_\theta \epsilon_{t-1}) + Var(\alpha_2 \circ_\theta \epsilon_{t-2}) + Var(\alpha_3 \circ_\theta \epsilon_{t-3}) + \cdots \\
&\quad + Var(\alpha_q \circ_\theta \epsilon_{t-q}) + Var(\epsilon_t) \\
&= (\alpha_1\lambda + \alpha_1\lambda^2\theta^2 - \alpha_1^2\lambda^2\theta^2) + (\alpha_2\lambda + \alpha_2\lambda^2\theta^2 - \alpha_2^2\lambda^2\theta^2) \\
&\quad + (\alpha_3\lambda + \alpha_3\lambda^2\theta^2 - \alpha_3^2\lambda^2\theta^2) + \cdots + (\alpha_q\lambda + \alpha_q\lambda^2\theta^2 - \alpha_q^2\lambda^2\theta^2) + \lambda \\
&= (\alpha_1\lambda + \alpha_2\lambda + \alpha_3\lambda + \cdots + \alpha_q\lambda) + (\alpha_1\lambda^2\theta^2 + \alpha_2\lambda^2\theta^2 + \alpha_3\lambda^2\theta^2 + \cdots \\
&\quad + \alpha_q\lambda^2\theta^2) - (\alpha_1^2\lambda^2\theta^2 + \alpha_2^2\lambda^2\theta^2 + \alpha_3^2\lambda^2\theta^2 + \cdots + \alpha_q^2\lambda^2\theta^2) + \lambda \\
&= \lambda\left[\lambda\theta^2\left(\sum_{i=1}^{q}\alpha_i - \sum_{i=1}^{q}\alpha_i^2\right) + \sum_{i=1}^{q}\alpha_i + 1\right].
\end{aligned}
$$

3) Since ϵ_i and ϵ_j are independent, by Corollary 25, then

$$Cov(\epsilon_i, \epsilon_j) = 0.$$

If $k = 1$,

$$
\begin{aligned}
Cov(X_t, X_{t-1}) &= Cov(\alpha_1 \circ_\theta \epsilon_{t-1} + \alpha_2 \circ_\theta \epsilon_{t-2} + \alpha_3 \circ_\theta \epsilon_{t-3} + \cdots + \alpha_q \circ_\theta \epsilon_{t-q} + \epsilon_t, \\
&\quad \alpha_1 \circ_\theta \epsilon_{t-2} + \alpha_2 \circ_\theta \epsilon_{t-3} + \alpha_3 \circ_\theta \epsilon_{t-4} + \cdots + \alpha_q \circ_\theta \epsilon_{t-q-1} + \epsilon_{t-1}) \\
&= Cov(\alpha_1 \circ_\theta \epsilon_{t-1}, \alpha_1 \circ_\theta \epsilon_{t-2}) + Cov(\alpha_1 \circ_\theta \epsilon_{t-1}, \alpha_2 \circ_\theta \epsilon_{t-3}) + \cdots \\
&\quad + Cov(\alpha_1 \circ_\theta \epsilon_{t-1}, \epsilon_{t-1}) + Cov(\alpha_2 \circ_\theta \epsilon_{t-2}, \alpha_1 \circ_\theta \epsilon_{t-2}) \\
&\quad + Cov(\alpha_2 \circ_\theta \epsilon_{t-2}, \alpha_2 \circ_\theta \epsilon_{t-3}) + \cdots + Cov(\alpha_2 \circ_\theta \epsilon_{t-2}, \epsilon_{t-1}) + \cdots \\
&\quad + Cov(\epsilon_t, \alpha_1 \circ_\theta \epsilon_{t-2}) + Cov(\epsilon_t, \alpha_2 \circ_\theta \epsilon_{t-3}) + \cdots + Cov(\epsilon_t, \epsilon_{t-1}) \\
&= Cov(\alpha_1 \circ_\theta \epsilon_{t-1}, \epsilon_{t-1}) + Cov(\alpha_2 \circ_\theta \epsilon_{t-2}, \alpha_1 \circ_\theta \epsilon_{t-2}) \\
&\quad + Cov(\alpha_3 \circ_\theta \epsilon_{t-3}, \alpha_2 \circ_\theta \epsilon_{t-3}) + \cdots \\
&\quad + Cov(\alpha_{q-1} \circ_\theta \epsilon_{t-q-1}, \alpha_q \circ_\theta \epsilon_{t-q-1}) \\
&= \alpha_1 Cov(\epsilon_{t-1}, \epsilon_{t-1}) + \alpha_1\alpha_2 Cov(\epsilon_{t-2}, \epsilon_{t-2}) + \alpha_2\alpha_3 Cov(\epsilon_{t-3}, \epsilon_{t-3}) \\
&\quad + \cdots + \alpha_{q-1}\alpha_q Cov(\epsilon_{t-q-1}, \epsilon_{t-q-1}) \\
&= \alpha_1(\lambda^2 + \lambda) + \alpha_1\alpha_2(\lambda^2 + \lambda) + \alpha_2\alpha_3(\lambda^2 + \lambda) + \cdots + \alpha_{q-1}\alpha_q(\lambda^2 + \lambda) \\
&= (\lambda^2 + \lambda)\left(\alpha_1 + \sum_{i=1}^{q-1} \alpha_i \alpha_{i+1}\right).
\end{aligned}
$$

In case of $k \in \{2, 3, 4, \ldots, q\}$, follow the same method as in the case $k = 1$, then we can show that

$$
Cov(X_t, X_{t-k}) = (\lambda^2 + \lambda)\left(\alpha_k + \sum_{i=1}^{q-k} \alpha_i \alpha_{i+k}\right).
$$

for $k \in \{1, 2, 3, \ldots, q\}$.

If $k > q$, then it is clear that $Cov(X_t, X_{t-k}) = 0$.

4) From Definition 2.30,

$$Corr(X_t, X_{t-k}) = \begin{cases} \dfrac{(\lambda+1)\left(\alpha_1 + \sum_{i=1}^{q-1}\alpha_i\alpha_{i+1}\right)}{\lambda\theta^2\left(\sum_{i=1}^{q}\alpha_i - \sum_{i=1}^{q}\alpha_i^2\right) + \sum_{i=1}^{q}\alpha_i + 1}, & k = 1, \\ \dfrac{(\lambda+1)\left(\alpha_2 + \sum_{i=1}^{q-2}\alpha_i\alpha_{i+2}\right)}{\lambda\theta^2\left(\sum_{i=1}^{q}\alpha_i - \sum_{i=1}^{q}\alpha_i^2\right) + \sum_{i=1}^{q}\alpha_i + 1}, & k = 2, \\ \vdots & \\ \dfrac{\alpha_q(\lambda+1)}{\lambda\theta^2\left(\sum_{i=1}^{q}\alpha_i - \sum_{i=1}^{q}\alpha_i^2\right) + \sum_{i=1}^{q}\alpha_i + 1}, & k = q, \\ 0, & k > q. \end{cases}$$

$\square$

3.4 Data Simulation

3.4.1 Data Simulation for DBINMA(1)

1. Data Generation Algorithm

For our simulation experiment, we use the following algorithm to generate data. Following DBINMA(1) model defined in Definition 3.5:

$$X_t = \alpha \circ_\theta \epsilon_{t-1} + \epsilon_t,$$

for $t \in \{1, 2, 3, \ldots, n\}$, and the thinning operator $\circ_\theta$ defined in Definition 3.3.

1. Set parameters α, θ, λ, n, where α, $\theta \in [0, 1]$, $\lambda > 0$ and n is a positive integer.

2. Generate a positive finite integer vector $\mathbf{Eps} = [\epsilon_0, \epsilon_1, \epsilon_2, \ldots, \epsilon_n]$, where ϵ_t, $t \in \{0, 1, 2, 3, \ldots, n\}$ are generated from the Poisson distribution with parameter λ.

3. Generate a positive finite vector $\mathbf{W} = [w_0, w_1, w_2, \ldots, w_n]$, where $w_0 = 0$ and $w_t,\ t \in \{1, 2, 3, \ldots, n\}$ are generated from the Uniform distribution on $[0, 1]$.

4. Generate a positive finite integer vector $\mathbf{Thin} = [T_0, T_1, T_2, \ldots, T_n]$, where $T_0 = 0$ and $T_t = \alpha \circ_\theta \epsilon_{t-1},\ t \in \{1, 2, 3, \ldots, n\}$ are generated from the Binomial distribution with parameters ϵ_{t-1} and $\alpha(1 - \theta)$ if $w_t \leq \alpha$, and if $w_t > \alpha$, $T_t = \alpha \circ_\theta \epsilon_{t-1}$ are generated from the the Binomial distribution with parameters ϵ_{t-1} and $\theta + \alpha(1 - \theta)$.

5. Generate a positive finite integer vector $\mathbf{X} = [X_0, X_1, X_2, \ldots, X_n]$, where $X_0 = 0$ and $X_t = T_t + \epsilon_t,\ t \in \{1, 2, 3, \ldots, n\}$.

2. Comparison the Data of INMA(1) and DBINMA(1) model

Figure 3.1 below shows comparisons of data generated from INMA(1) model defined in Definition 2.56 and DBINMA(1) model defined in Definition 2.56 for different values of θ ($\theta = 0$, 0.25, 0.5, 0.75 and 1). The corresponding histograms are presented in Figure 3.2.

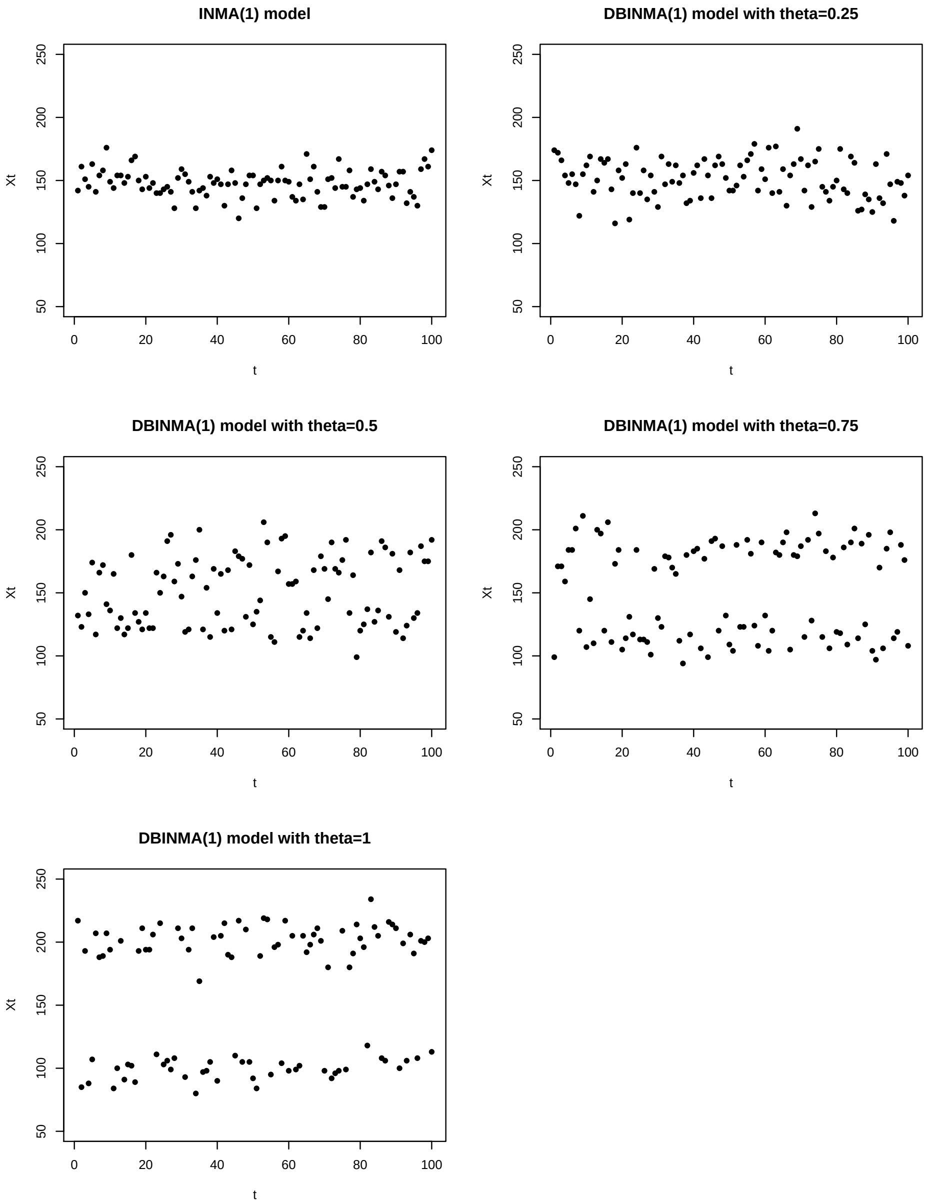

Figure 3.1: Scatter plots of data generated from DBINMA(1) model with parameters $\lambda = 100$, $\alpha = 0.5$ and $\theta = 0$ (INMA(1) model), 0.25, 0.5, 0.75 and 1.

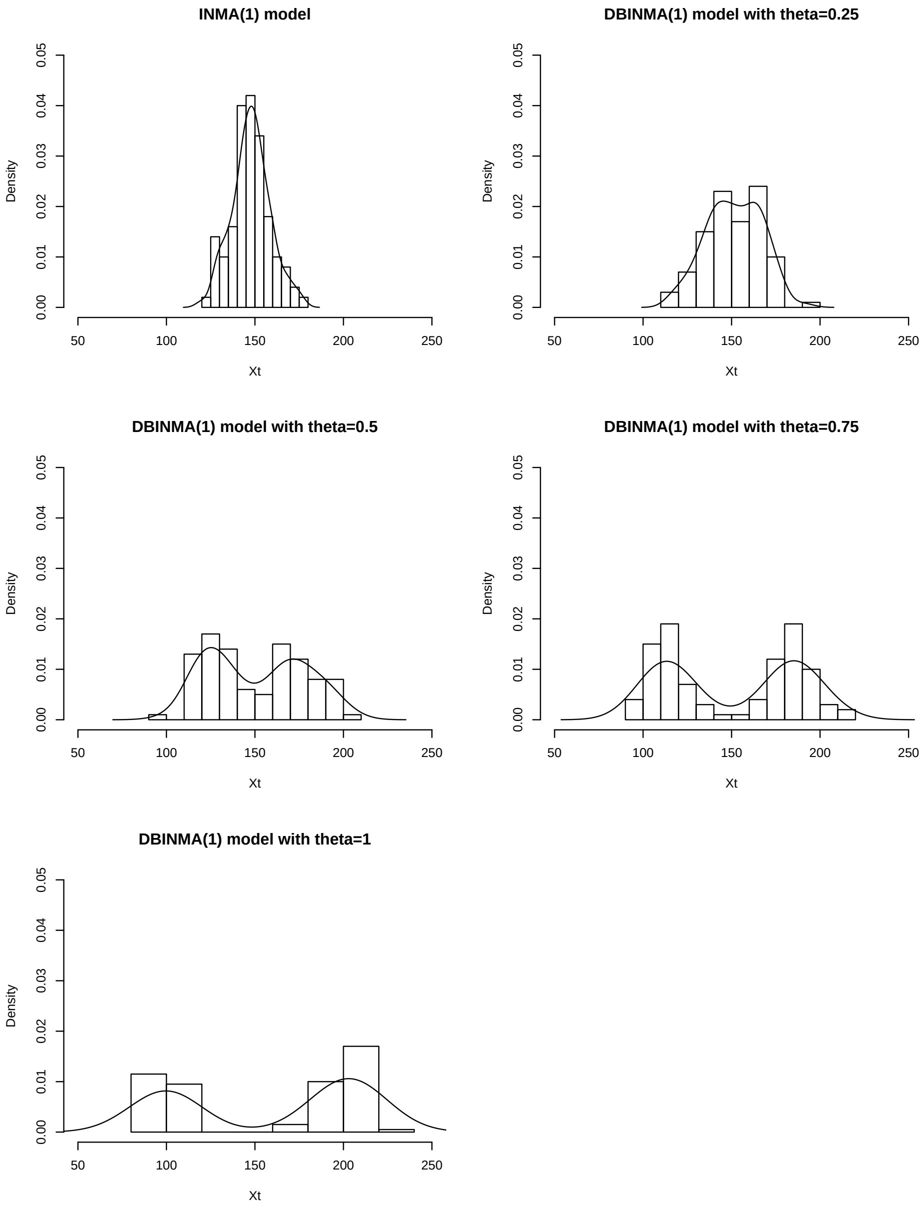

Figure 3.2: Histograms of data generated from DBINMA(1) model with parameters $\lambda = 100$, $\alpha = 0.5$ and $\theta = 0$ (INMA(1) model), 0.25, 0.5, 0.75 and 1.

From Figures 3.1 and 3.2, the parameter θ affects the aggregation of X_t. We can see that the data generated from INMA(1) model is unimodal, while the data generated from DBINMA(1) model is bimodal. Moreover, the distance between the two models of DBINMA(1) model is greater when θ increases.

3. Comparison the Data of DBINMA(1) with different values of α.

Figures 3.3 and 3.4 show respective the data plots and histograms of data generated from DBINMA(1) model for different values of α when θ and λ are fixed to be 0.5 and 100, respectively. The values of α considered in these figures are $\alpha = 0$, 0.25, 0.5, 0.75 and 1.

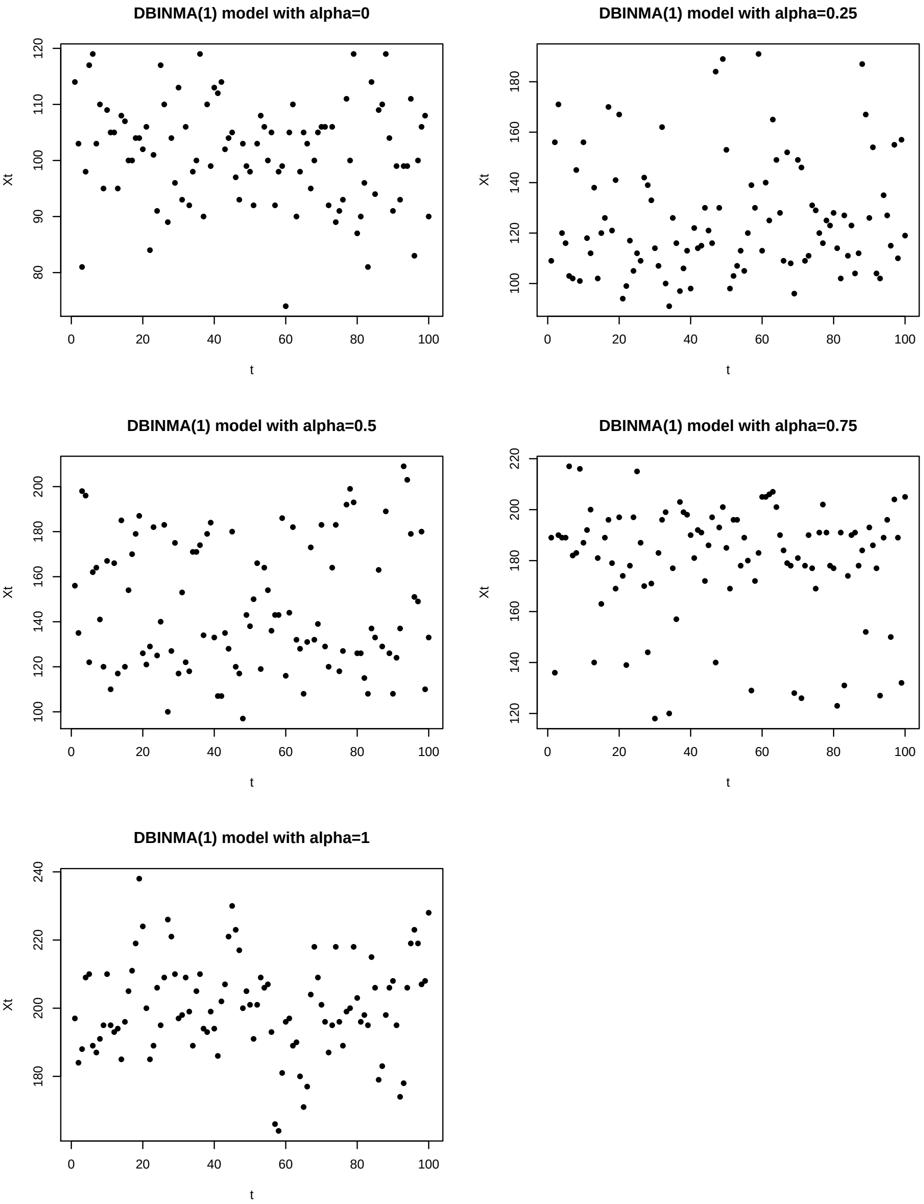

Figure 3.3: Scatter plots of data generated from DBINMA(1) model with parameters $\lambda = 100$, $\theta = 0.5$ and $\alpha = 0$, 0.25, 0.5, 0.75 and 1.

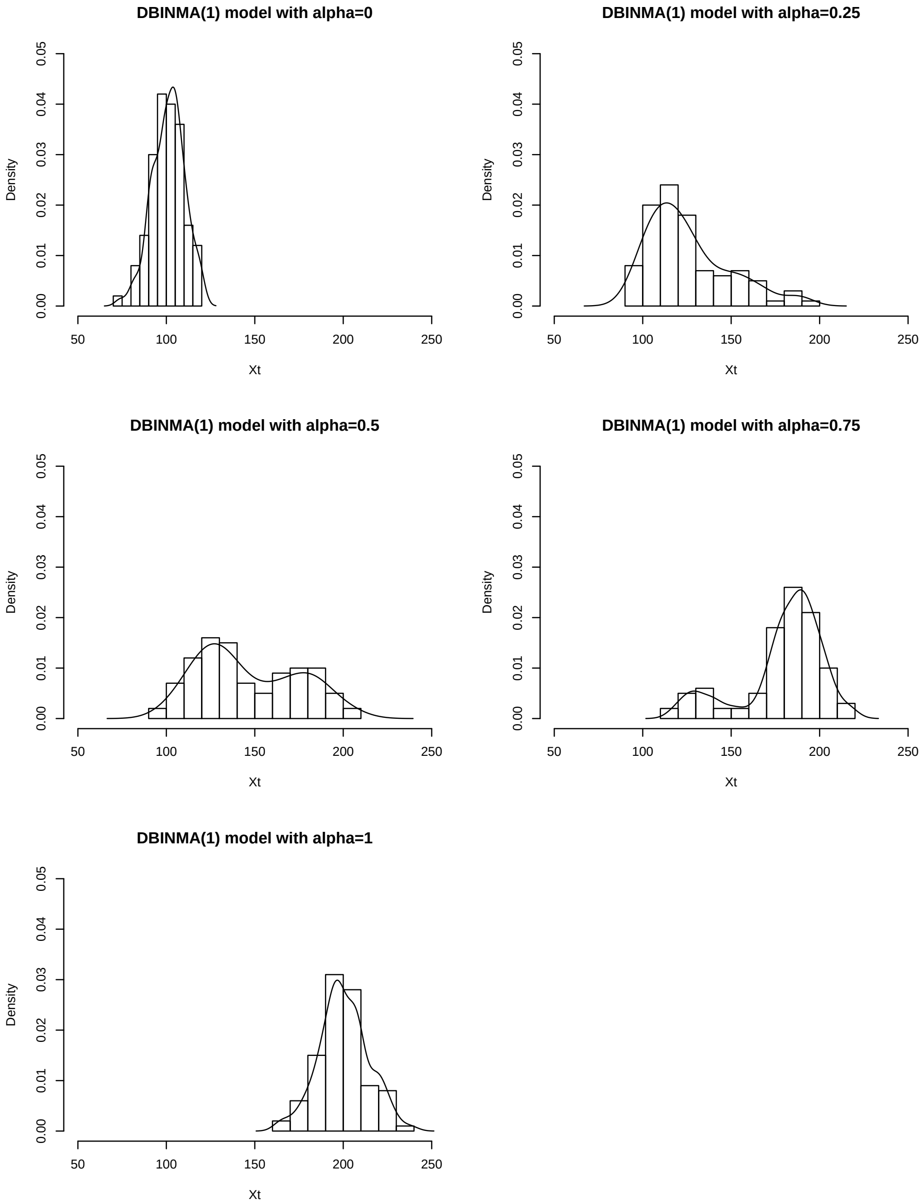

Figure 3.4: Histograms of data generated from DBINMA(1) model with parameters $\lambda = 100$, $\theta = 0.5$ and $\alpha = 0$, 0.25, 0.5, 0.75 and 1.

From Figures 3.3 and 3.4, the parameter α affects values of X_t becuase the expectation of X_t is $\lambda(\alpha + 1)$. Therefore, we can see that the modes of data are higher when α increases. The data show unimodal pattern for cases when $\alpha = 0$ and $\alpha = 1$.

4. Comparison the generate 100 series of DBINMA(1) model

Figures 3.5 and 3.6 show histograms of X_{10}, X_{20}, X_{30}, …, X_{100} of data generated from DBINMA(1) model when θ, α and λ are fixed to be 0.5, 0.5 and 100, respectively.

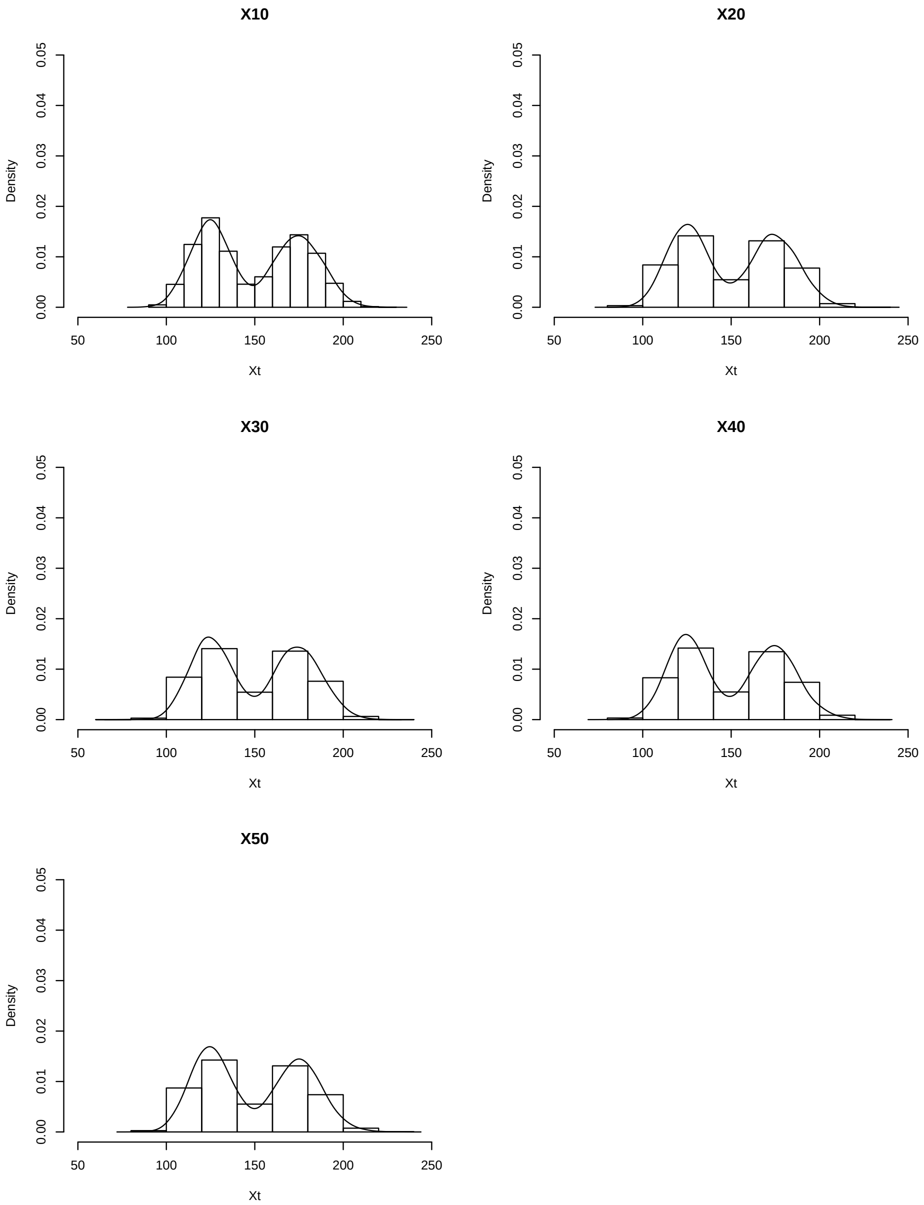

Figure 3.5: Histograms of data X_{10}, X_{20}, X_{30}, X_{40} and X_{50} of DBINMA(1) model with parameters $\lambda = 100$, $\alpha = 0.5$ and $\theta = 0.5$.

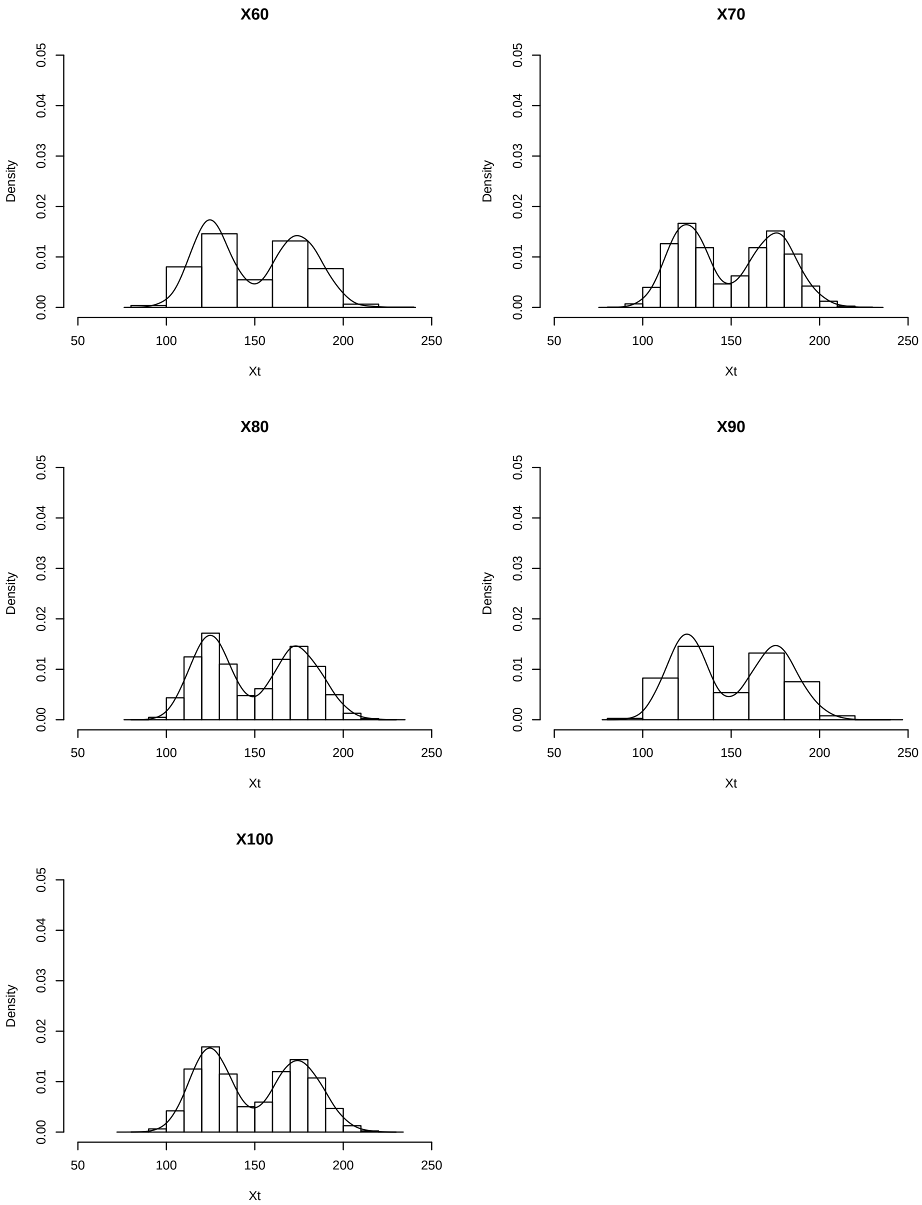

Figure 3.6: Histograms of data X_{60}, X_{70}, X_{80}, X_{90} and X_{100} of DBINMA(1) model with parameters $\lambda = 100$, $\alpha = 0.5$ and $\theta = 0.5$.

From Figures 3.5 and 3.6, X_{10}, X_{20}, X_{30}, $\ldots$, X_{100} show the same pattern of mixed distributions with two modes which agrees to the Theorem 3.6.

3.4.2 Data Simulation for DBINMA(2)

1. Data Generation Algorithm

For our simulation experiment, we use the following algorithm to generate data. Following DBINMA(2) model defined in Definition 3.8:

$$X_t = \alpha_1 \circ_\theta \epsilon_{t-1} + \alpha_2 \circ_\theta \epsilon_{t-2} + \epsilon_t,$$

for $t \in \{2, 3, 4, \dots, n\}$ and the thinning operator $\circ_\theta$ defined in Definition 3.3.

1. Set parameters $\alpha_1,\ \alpha_2,\ \theta,\ \lambda,\ n,$ where $\alpha_1,\ \alpha_2,\ \theta \in [0, 1],\ \lambda > 0$ and n is a positive integer.

2. Generate a postive finite integer vector $\mathbf{Eps} = [\epsilon_0, \epsilon_1, \epsilon_2, \dots, \epsilon_n]$, where ϵ_t, $t \in \{0, 1, 2, \dots, n\}$ are generated from the Poisson distribution with parameter λ.

3. Generate a postive finite vector $\mathbf{W1} = [w_{10}, w_{11}, w_{12}, \dots, w_{1n}]$, where $w_{10} = 0$ and $w_{1t},\ t \in \{1, 2, 3, \dots, n\}$ are generated from the Uniform distribution on $[0, 1]$

4. Generate a postive finite vector $\mathbf{W2} = [w_{20}, w_{21}, w_{22}, \dots, w_{2n}]$, where $w_{20} = 0$ and $w_{2t},\ \ t \in \{1, 2, 3, \dots, n\}$ are generated from the Uniform distribution on $[0, 1]$.

5. Generate a postive finite integer vector $\mathbf{Thin1} = [T_{10}, T_{11}, T_{12}, \dots, T_{1n}]$, where $T_{10} = 0$ and $T_{1t} = \alpha_1 \circ_\theta \epsilon_{t-1},\ \ t \in \{1, 2, 3, \dots, n\}$ are generated from the Binomial distribution with parameters ϵ_{t-1} and $\alpha_1(1-\theta)$ if $w_{1t} \leq \alpha_1$, and if $w_{1t} > \alpha_1$, $T_{1t} = \alpha_1 \circ_\theta \epsilon_{t-1}$ are generated from the Binomial distribution with parameters ϵ_{t-1} and $\theta + \alpha_1(1-\theta)$.

6. Generate a postive finite integer vector $\mathbf{Thin2} = [T_{20}, T_{21}, T_{22}, \cdots, T_{2n}]$, where $T_{20} = 0$, $T_{21} = 0$ and $T_{2t} = \alpha_2 \circ_\theta \epsilon_{t-2},\ \ t \in \{2, 3, 4, \dots, n\}$ are generated from the Binomial distribution with parameters ϵ_{t-2} and $\alpha_2(1-\theta)$ if $w_{2t} \leq \alpha_2$, and if $w_{2t} > \alpha_2$, $T_{2t} = \alpha_2 \circ_\theta \epsilon_{t-2}$ are generated from the Binomial distribution with parameters ϵ_{t-2} and $\theta + \alpha_2(1-\theta)$.

7. Generate a postive finite integer vector $\mathbf{X} = [X_0, X_1, X_2, \ldots, X_n]$, where $X_0 = 0$, $X_1 = 0$ and $X_t = T_{1t} + T_{2t} + \epsilon_t,\ t \in \{2, 3, 4, \ldots, n\}$.

2. Comparison the Data of INMA(2) and DBINMA(2) model

Figure 3.7 below shows comparisons of data generated from DBINMA(2) model for different values of θ ($\theta = 0$, 0.25, 0.5, 0.75 and 1). The corresponding histograms are presented in Figure 3.8.

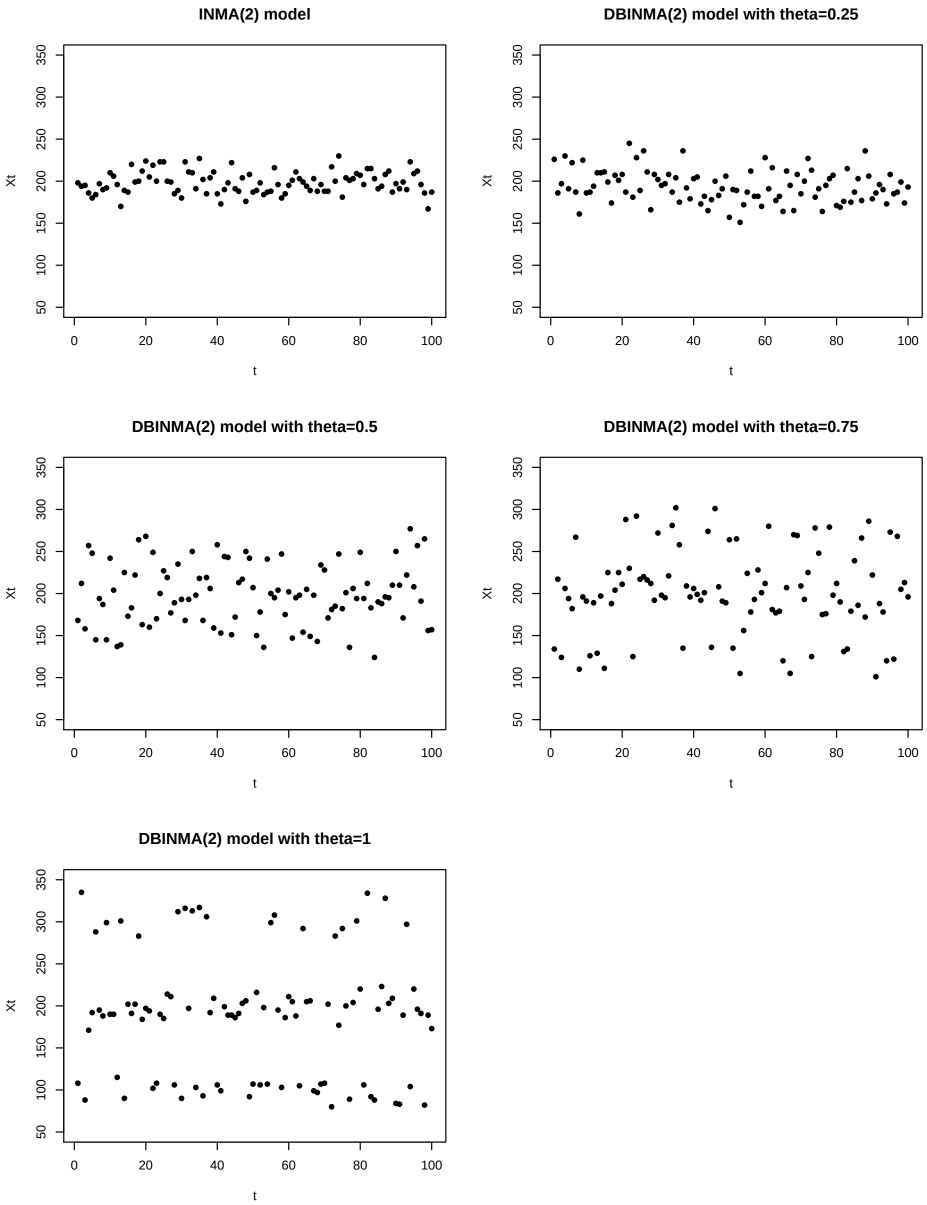

Figure 3.7: Scatter plots of data generated from DBINMA(2) model with parameters $\lambda = 100$, $\alpha_1 = \alpha_2 = 0.5$ and $\theta = 0$ (INMA(2) model), 0.25, 0.5, 0.75 and 1.

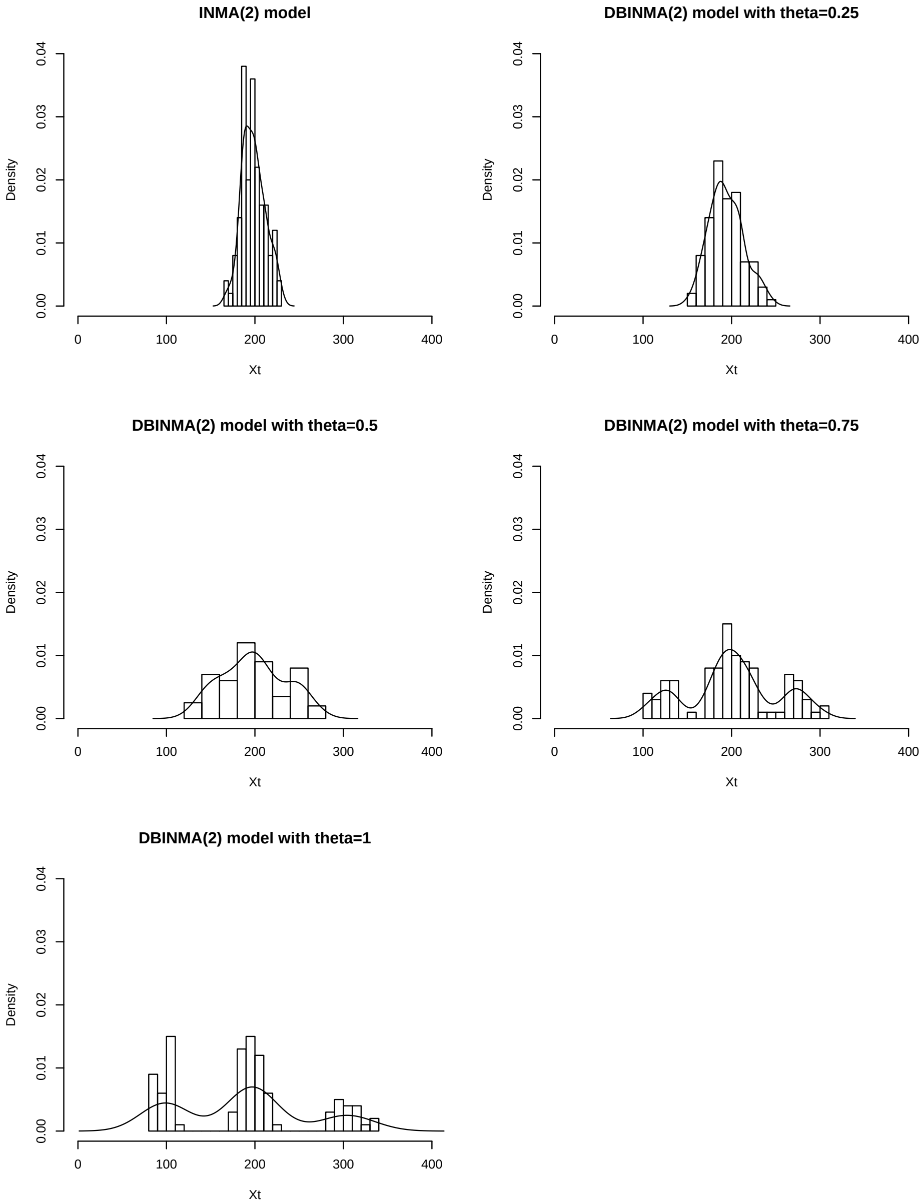

Figure 3.8: Histograms of data generated from DBINMA(2) model with parameters $\lambda = 100$, $\alpha_1 = \alpha_2 = 0.5$ and $\theta = 0$ (INMA(2) model), 0.25, 0.5, 0.75 and 1.

From Figures 3.7 and 3.8, we can see that the data generated from INMA(2) model is unimodal while the data genereted from DBINMA(2) model is trimodal. Moreover, the distances between the three models are greater when θ increases.

3. Comparison the Data of DBINMA(2) with different values of α.

Figures 3.9 and 3.10 show respective the data plots and histograms of data generated from DBINMA(2) model from different values of α_1 and α_2 when θ and λ are fixed to be 0.5 and 100, respectively. The values of α_1 and α_2 considered in these figures are $\alpha_1 = \alpha_2 = 0$, 0.25, 0.5, 0.75 and 1.

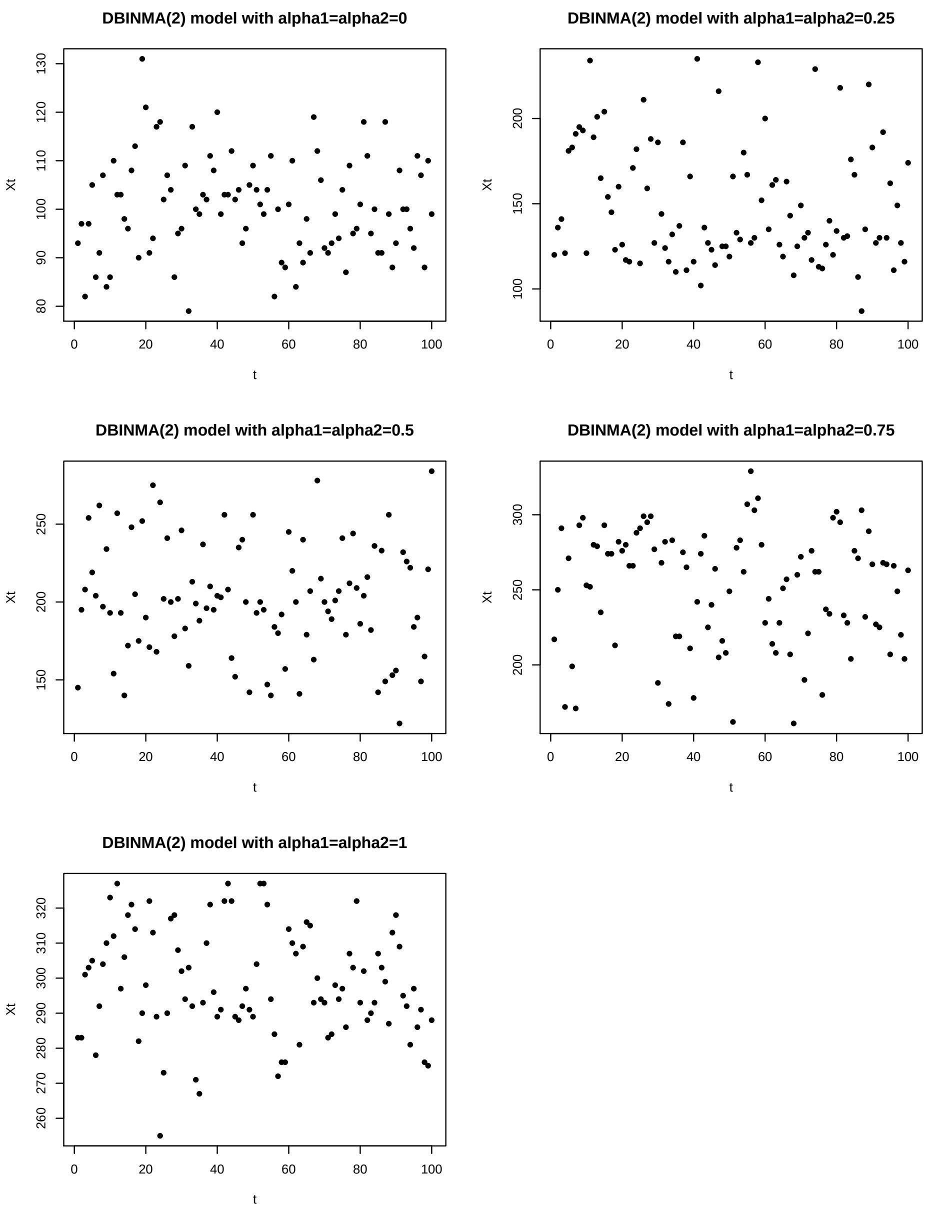

Figure 3.9: Scatter plots of data generated from DBINMA(2) model with parameters $\lambda = 100$, $\theta = 0.5$ and $\alpha_1 = \alpha_2 = 0$, 0.25, 0.5, 0.75 and 1.

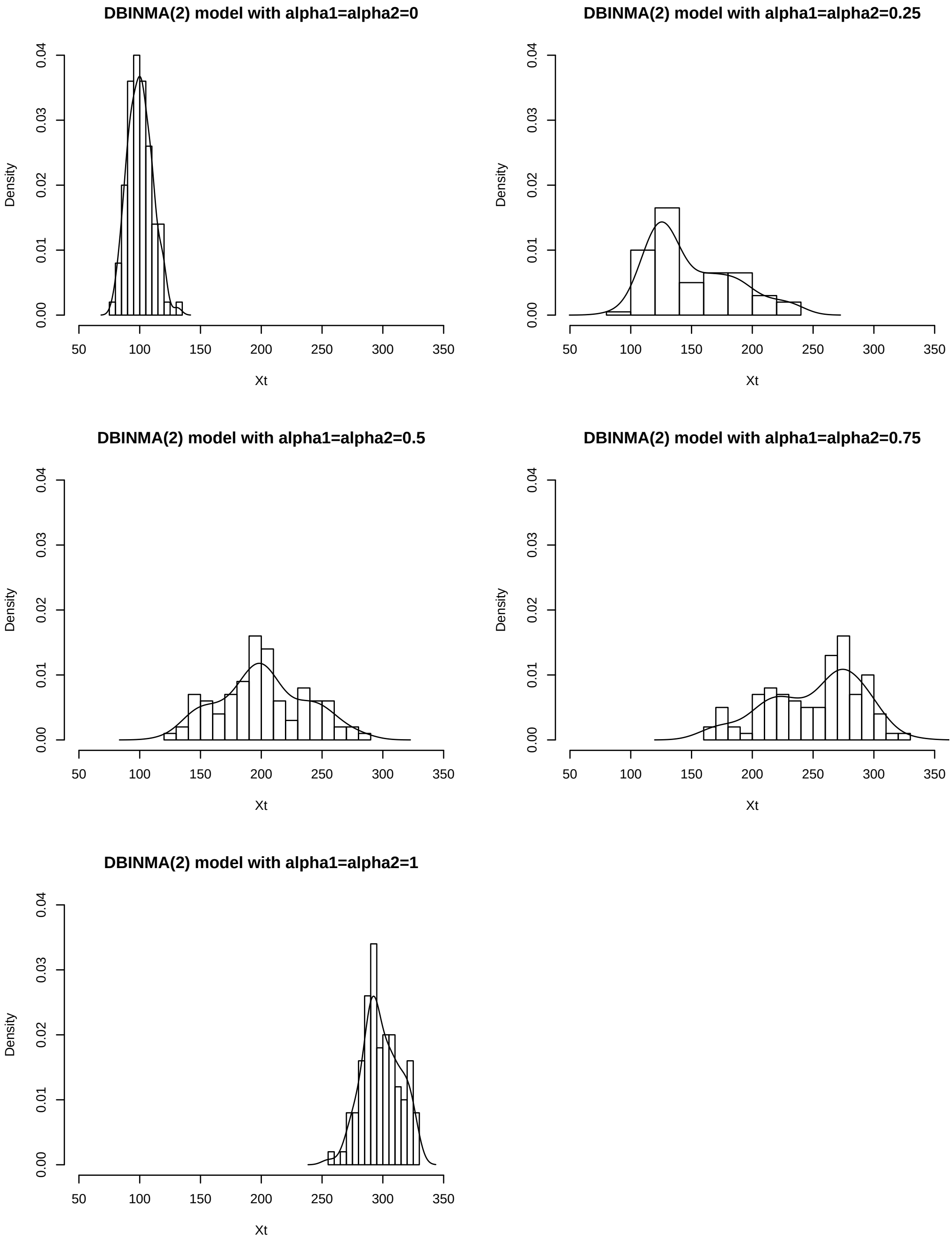

Figure 3.10: Histograms of data generated from DBINMA(2) model with parameters $\lambda = 100$, $\theta = 0.5$ and $\alpha_1 = \alpha_2 = 0$, 0.25, 0.5, 0.75 and 1.

From Figures 3.9 and 3.10, the parameters α_1 and α_2 affect values of X_t becuase the expectation of X_t is $\lambda(\alpha_1 + \alpha_2 + 1)$. Therefore, we can see that the modes of data are higher when α_1 and α_2 increases. The data show unimodal pattern for cases when $\alpha_1 = \alpha_2 = 0$ and $\alpha_1 = \alpha_2 = 1$.

4. Comparision the generate 100 series of DBINMA(2) model

Figures 3.11 and 3.12 show histograms of X_{10}, X_{20}, X_{30}, ..., X_{100} of data generated from DBINMA(2) model when θ, α_1, α_2 and λ are fixed to be 0.5, 0.5, 0.5 and 100, respectively.

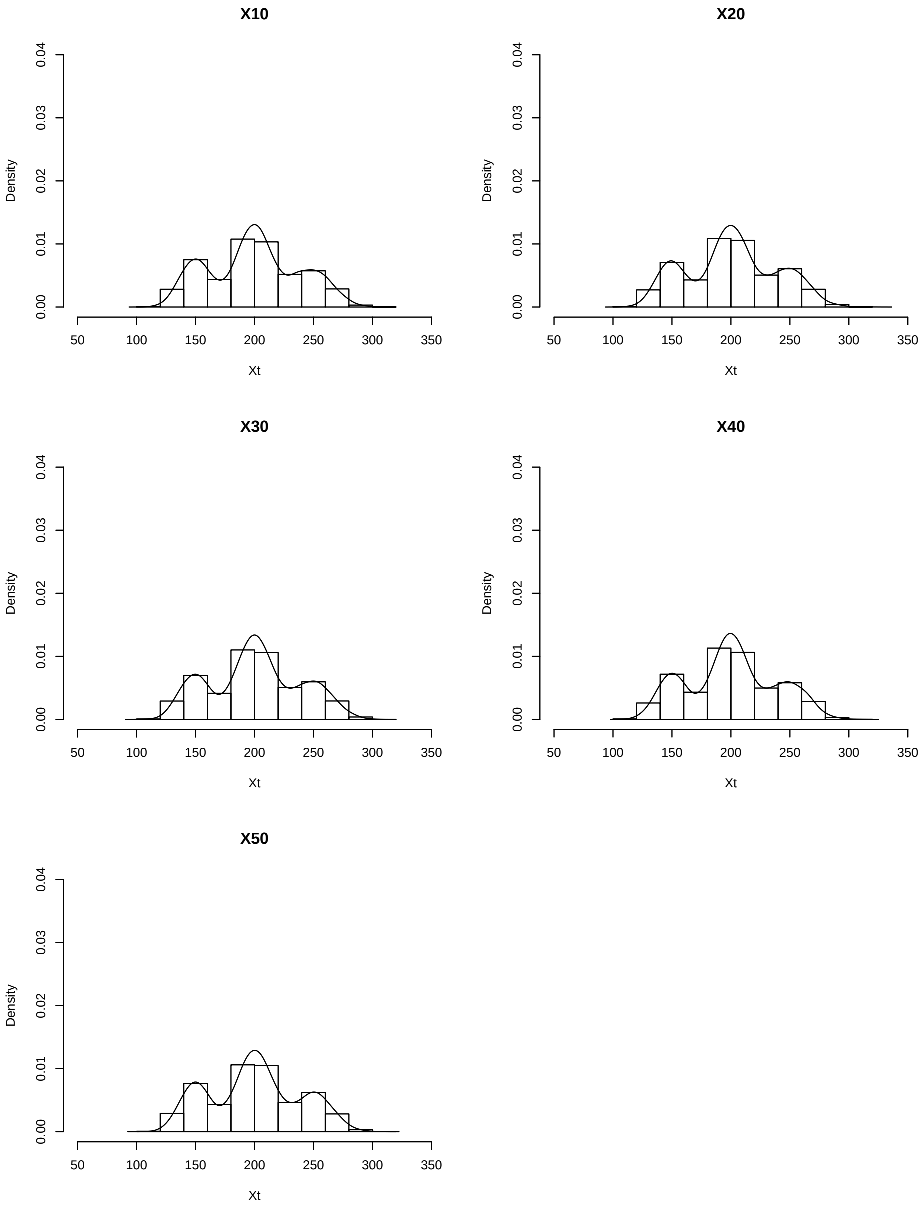

Figure 3.11: Histograms of data X_{10}, X_{20}, X_{30}, X_{40} and X_{50} of DBINMA(2) model with parameters $\lambda = 100$, $\alpha_1 = \alpha_2 = 0.5 = 0.5$ and $\theta = 0.5$.

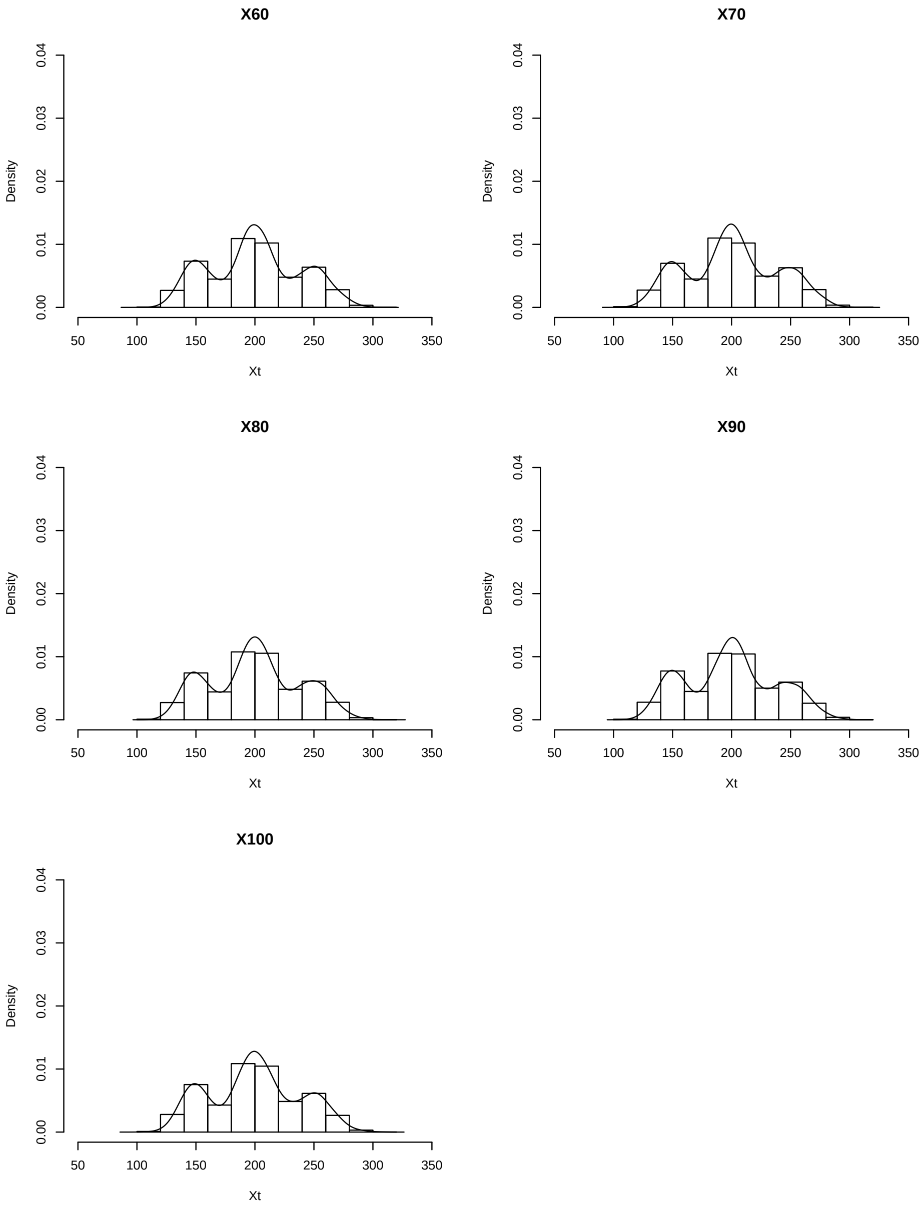

Figure 3.12: Histograms of data X_{60}, X_{70}, X_{80}, X_{90} and X_{100} of DBINMA(2) model with parameters $\lambda = 100$, $\alpha_1 = \alpha_2 = 0.5 = 0.5$ and $\theta = 0.5$.

From Figures 3.11 and 3.12, X_{10}, X_{20}, X_{30}, $\ldots$, X_{100} show the same pattern of mixed distributions with three modes which agrees to the Theorem 3.9.

Chapter 4

Conclusion

In this project, we extended the study of the regular INMA model to construct a new INMA model with generalized binomial thinning operator, called DBINMA model and derived thier probabilistic properties such as mean, variance and covariance. Moreover, we presented the data and distribution plots in different setting. For our simulation study, we found that the parameter θ affects the distribution. The histograms show multimodel of the mixture distribution the parameter θ increases. However, the distribution is hardly affected by the parameter α.

References

[1] Al-Osh, M. A., and Alzaid, A. A. (1987). First-oder integer-valued autoregressive (INAR(1)) process. **Journal of Time Series Analysis** 8(3), 261-275.

[2] Al-Osh, M., and Alzaid, A. A. (1988). Integer-valued moving average (INMA) process. **Statistical Papers** 29, 281-300.

[3] Aly, E. E., and Bouzar, N. (1994). On some integer-valued autoregressive moving average models. **Journal of Multivariate Analysis** 50, 132-151.

[4] McKenzie, E. (1985). Some simple models for discrete variate series. **Water Resources Bulltin** 21, 640-650.

[5] Prasanna, S. (2013). Probability and mathematical statistics. **Louisvile: University of Louisvile**.

[6] Ristic, M. M., Nastic, A. S., and Miletic Ilic, A. V. (2013). A geometric time series model with dependent Bernoulli counting series. **Journal of Time Series Analysis** 34, 466-476.

Appendix

Coding for INMA(1)

```
alpha = 0.5
theta = 0.5
lambda = 100
n = 101 #number of Xt+1
N = 10000 #number of simulation
Eps=c(numeric(n))
Thin = c(numeric(n))
w = c(numeric(n))
Xt = c(numeric(n))

sim = array(numeric(n*4*N),dim = c(n,4,N))
for (k in 1:N){
  for (i in 1:n){
    Eps[i] = rpois(1,lambda)
   }
for (i in 1:n-1){
  w[1] = 0
  Thin[1] = 0
  Xt[1] = 0
  w[i+1] = runif(1,0,1)
   if (w[i+1] <= alpha) { Thin[i+1] = rbinom(1,Eps[i],alpha+theta-
alpha*theta)}
  else{Thin[i+1]= rbinom(1,Eps[i], alpha*(1-theta))}
  Xt[i+1] = Thin[i+1] + Eps[i+1]
M = cbind(w,Thin,Eps,Xt)
}
for (i in 1:n){
```

```
  for (j in 1:4){
   sim[i,j,k] = M[i,j]
  }
 }
}

rownames(sim)<-rownames(sim, do.NULL = FALSE, prefix = "index")
colnames=c("w","Thin","Eps","Xt")
dimnames(sim)<-list(rownames(sim), colnames)
```

Coding for INMA(2)

```
alpha1 = 0.5
alpha2 = 0.5
theta = 0.5
lambda = 10
n = 101 #number of Xt+1
N = 10000 #number of simulation
Eps = c(numeric(n))
w1 = c(numeric(n))
w2 = c(numeric(n))
Thin1 = c(numeric(n))
Thin2 = c(numeric(n))
Xt = c(numeric(n))

sim = array(numeric(n*6*N),dim = c(n,6,N))
for (k in 1:N){
  for (i in 1:n){
   Eps[i] = rpois(1,lambda)
  }
  for (i in 1:n-1){
    w1[1] = 0
    Thin1[1] = 0
    w1[i+1] = runif(1,0,1)
   if (w1[i+1] <= alpha1) { Thin1[i+1] = rbinom(1,Eps[i],alpha1+theta-
alpha1*theta)}
    else{Thin1[i+1]= rbinom(1,Eps[i], alpha1*(1-theta))}
  }
  for (i in 1:n-2){
    w2[1] = 0
    w2[2] = 0
    Thin2[1] = 0
```

```
    Thin2[2] = 0
    Xt[1] = 0
    Xt[2] = 0
    w2[i+2] = runif(1,0,1)
   if (w2[i+2] <= alpha2) { Thin2[i+2] = rbinom(1,Eps[i],alpha2+theta-
alpha2*theta)}
    else{Thin2[i+2]= rbinom(1,Eps[i], alpha2*(1-theta))}
    Xt[i+2] = Thin1[i+2] + Thin2[i+2] + Eps[i+2]
    M = cbind(w1,w2,Thin1,Thin2,Eps,Xt)
  }
 for (i in 1:n){
  for (j in 1:6){
    sim[i,j,k] = M[i,j]
  }
 }
}
rownames(sim)<-rownames(sim, do.NULL = FALSE, prefix = "index")
colnames=c("w1","w2","Thin1","Thin2","Eps","Xt")
dimnames(sim)<-list(rownames(sim), colnames)
```

The Project Proposal of Course 2301399 Project Proposal First Semester, Academic Year 2018

Project Tittle (Thai)	การศึกษาตัวแบบค่าจำนวนเต็มการเฉลี่ยเคลื่อนที่บนฐานของตัวดำเนินการทวินามแบบวางนัยทั่วไป
Project Tittle (English)	A study on an integer-valued moving average models based on generalized Binomial thinning operator
Advisor	Asst. Prof. Jiraphan Suntornchost, Ph.D.
By	Mr. Namchai Pew-On ID 5833529323 Mathematics, Department of Mathematics and Computer Science Faculty of Science, Chulalongkorn University

---

Background and Rationale

The integer-valued time series models for time series count data play important roles in many applications. For examples, claim counts in insurance business, and the number of stock transactics in stock market. The model was first introduced by McKenzie (1985) for the lag-one dependence model which is referred nowadays as the integer-valued autoregressive of order one (INAR(1)) process.

The original integer-valued time series model is based on a Binomial thinning operator which is a compound sum of independent and identically distributed Bernoulli random variables. Later, Ristic et al. (2013) generalized the Binomial thinning operator to construct a new INAR model, by relaxing the assumption of independence in the compound sum to a sum of dependence Bernoulli random variables, called the generalized Binomial thinning operator. The new thinning operator can be applied to a wider class of applications. For example, survival or collapse of some companies in economy discussed in Ristic et al. (2013) since all companies operate in the same macroeconomic and may affect to each others. The problem is more reasonable to use the dependent thinning operator than the independent thinning operator.

In this study, we extend the generalized Binomial thinning operator introduced in of Ristic et al. (2013) to construct more general integer-valued moving average models, called DBINMA.

Objectives

To extend the INAR(1) model based on the generalized Binomial thinning operator to DBINMA models and study its properties and applications.

Scope

In this project we will study properties and applications of DBINMA models based on the generalized binomial thinning operator.

Project Activities

1. Study probabilistic properties of integer-valued time series models and binomial thinning operators.

2. Study probabilistic properties of the Poisson INAR model.

3. Study probabilistic properties of Poisson INMA model.

4. Study probabilistic properties of INAR(1) model with dependent Bernoulli thinning operator.

5. Construct DBINMA models and derive their probabilistic properties.

6. Summarize the project and write a report.

Scheduled Operations

Procedures	Months								
	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.
1. Study Integer-valued time series models and binomial thinning operators.	■	■							
2. Study the Poisson INAR models.		■	■	■					
3. Study Poisson INMA models.		■	■	■					
4. Study INAR models with dependent Bernoulli thinning operator.		■	■	■					
5. Construct DBINMA models and derive their properties.			■	■	■	■	■		
6. Summarize the project and write a report.							■	■	■

Benefits

The benefits for student who implement this project.

1. To learn properties and applications of the proposed model.

2. To gain knowledge in probability theory and to apply the model to suitable application.

The benefits for users of the project

1. To have an alternative Integer-valued time series model for wider applications.

Equipment

Software

1. Microsoft Word
2. Mathematica
3. Program R
4. Latex

Hardware

1. Computer
2. Printer

References

1. McKenzie, E. (1985). Some simple models for discrete variate series. **Water Resources Bulltin** 21, 640-650.
2. Ristic, M. M., Nastic, A. S., and Miletic Ilic, A. V. (2013). A geometric time series model with dependent Bernoulli counting series. **Journal of Time Series Analysis** 34, 466-476.

Author Profile

Mr. Namchai Pew-On ID 583 35293 23

Branch of Mathematics

Department of Mathematics and Computer Science

Chulalongkorn University