การศึกษาของตัวแบบอนุกรมเวลาที่มีค่าเป็นจำนวนเต็มแบบปัวซงผสม

นางสาวปาณิสรา แย้มสุข

โครงงานนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2561

A study on Mixed Poisson integer-valued times series models

Miss Panisara Yamsook

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	การศึกษาของตัวแบบอนุกรมเวลาที่มีค่าเป็นจำนวนเต็มแบบปัวซงผสม
โดย	นางสาวปาณิสรา แย้มสุข รหัสประจำตัวนิสิต 5833532123
สาขาวิชา	คณิตศาสตร์
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.จิราพรรณ สุนทรโชติ

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับโครงงานฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาบัณฑิตในรายวิชา 2301499 โครงงานวิทยาศาสตร์ (Senior Project)

.... หัวหน้าภาควิชาคณิตศาสตร์
และวิทยาการคอมพิวเตอร์

(ศาสตราจารย์ ดร.กฤษณะ เนียมมณี)

คณะกรรมการสอบโครงงาน

.... อาจารย์ที่ปรึกษาโครงงานหลัก

(ผู้ช่วยศาสตราจารย์ ดร.จิราพรรณ สุนทรโชติ)

.... กรรมการ

(รองศาสตราจารย์ ดร.ณัฐพันธ์ กิติสิน)

.... กรรมการ

(รองศาสตราจารย์ ดร.อมร วาสนาวิจิตร์)

ปาณิสรา แย้มสุข : การศึกษาของตัวแบบอนุกรมเวลาที่มีค่าเป็นจำนวนเต็มแบบปัวซงผสม. (A STUDY ON MIXED POISSON INTEGER-VALUED TIMES SERIES MODELS)
อ.ที่ปรึกษาโครงงานหลัก : ผู้ช่วยศาสตราจารย์ ดร.จิราพรรณ สุนทรโชติ, 54 หน้า.

ในปี 2017 แบทเตโต-ชูชา ได้ทำการขยายตัวแบบปัวซงที่มีการถดถอยในตัวอันดับหนึ่งไปยังตัวแบบปัวซงแบบผสมที่มีการถดถอยในตัวอันดับหนึ่ง สำหรับข้อมูลที่มีการกระจายเกินเกณฑ์ โดยพวกเขาได้พิจารณาตัวแบบ ค่าจำนวนเต็มเกาส์ผกผัน-ปัวซงที่มีการถดถอยในตัวอันดับหนึ่ง ในโครงการนี้ จะทำการขยายตัวแบบปัวซงแบบผสมที่มีการถดถอยในตัวอันดับหนึ่ง เพื่อสร้างตัวแบบปัวซงผสมที่มีการเฉลี่ยเคลื่อนที่อันกับ q และศึกษาคุณสมบัติความน่าจะเป็นของตัวแบบนี้ เช่น ค่าเฉลี่ย ค่าความแปรปรวน และค่าความแปรปรวนร่วม ยิ่งไปกว่านั้นเราได้นำเสนอกราฟการกระจายของข้อมูลในแต่ละการตั้งค่าที่แตกต่างกัน

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์ ลายมือชื่อนิสิต . ปาณิสรา แย้มสุข . . . .
สาขาวิชา . คณิตศาสตร์ . ลายมือชื่อ อ.ที่ปรึกษาโครงงานหลัก . . . .
ปีการศึกษา . . . . 2561 . . . .

5833532123 : MAJOR MATHEMATICS.
KEYWORDS : MIXED POISSON; INTEGER-VALUED TIME SERIES; MOVING AVERAGE; INVERSE-GAUSSIAN.

PANISARA YAMSOOK : A STUDY ON MIXED POISSON INTEGER-VALUED TIMES SERIES MODELS. ADVISOR : ASST. JIRAPHAN SUNTORNCHOST, Ph.D., 54 pp.

In 2017, Batteto-Souza extended the Poisson INAR(1) to the Mixed Poisson INAR(1) model to accommodate overdispersion data. In their study, they considered the Poisson inverse-gaussian INAR(1) model. In this project, we will extend the study of the Mixed Poisson INAR(1) model to construct a Mixed Poisson INMA(q) model and derive their probabilistic properties such as mean, variance and covariance. Moreover, we present distribution plots of such data in many different settings.

Department Mathematics and Computer Science Student's Signature Panisara Yamsook
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 Assoc. Prof. Dr. Nataphan Kitisin and Assoc. Prof. Amorn Wasanawichit for thier suggestions and comments.

This project is funded by Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, and the His Royal Highness Crown Prince Maha Vajiralongkorn Fund.

Contents

List of Figures

Chapter 1

Introduction

Count time series arise naturally in many practical situations, for example, the insurance claim counts and the number of stock transactions. Therefore, increases in interest in the modelling have been observed. The most common distribution considered in count time series data is the Poisson distribution. The model is refered as the Poisson INAR(1) model which is a stationary integer-valued time series with lag-one dependence. The Poisson INAR model has been applied in many applications since it was introduced by McKenzie in 1985.

However, the property of the Poisson models having equal mean and variance is rarely found in applications. Many real-world data examples exhibit overdispersion, i.e., the variance is larger than the mean. Therefore, the integer-valued autoregressive (INAR) process with Poisson marginals is not adequate for modelling overdispersed counts. Consequently, several alternative distributions have been proposed for the integer-valued time series models, for example, geometric distribution and negative binomial distribution. Recently, in 2017, Batteto-Souza extended the Poisson INAR(1) to the mixed Poisson INAR(1) model to accommodate overdispersion data. In their study, they considered the inverse-gaussian Poisson INAR(1) model.

In this project, we will extend the study of the Mixed Poisson INAR(1) model to construct a Mixed Poisson INMA(1) model and derive their probabilistic properties such as mean, variance and covariance. Moreover, we present distribution plots of such data in many different settings.

Chapter 2

Preliminary

In this chapter, we present some basics of probability, an overview of count data and definition and properties of binomial thinning operator.

2.1 Basic of probability theory

In this section, we present definitions of distribution, expectation, independence, covariance, variance, generating function, moment generating function and conditional probability, and their properties.

Definition 1. Let (Ω, F, P) be a probability space and X be a random variable. Then the function $F_X : \mathbb{R} \to [0, 1]$ defined by

$$F_X(x) = P(\{\omega \in \Omega : X(\omega) \leq x\}) = P(X \leq x) \quad for\ x \in \mathbb{R},$$

is called "the distribution function of X".

Definition 2. Let X be a random variable and $g : \mathbb{R} \to \mathbb{R}$.
If X is a discrete random variable, the expectation of $g(X), E(g(X))$, is defined as

$$E(g(X)) = \sum_{x \in ImX} g(x)P(X = x).$$

If X is a continuous random variable, the expectation of $g(X), E(g(X))$, is defined as

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

Theorem 1. *Let $a, b \in \mathbb{R}$ and X be a random variable, the properties of expectation are given as follows.*

1.) $E(a) = a$*,*

2.) $E(aX) = aE(X)$*,*

3.) $E(aX + b) = aE(X) + b$*,*

4.) $E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$*.*

Definition 3. Let X and Y be any two random variables, A and B be any two subsets of real number. Then, we say that X and Y are independent random variables if and only if

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Definition 4. Let X and Y be random variables. The covariance of X and Y, denoted by $Cov(X, Y)$, is defined as

$$Cov(X, Y) = E\left((X - E(X))(Y - E(Y))\right).$$

Remark 1. 1.) The covariance function has an alternative expression as $Cov(X, Y) = E(XY) - E(X)E(Y)$.

2.) If X, Y are independent, then $Cov(X, Y) = 0$.

3.) $Cov(X, X) = Var(X)$.

Definition 5. Let X be a random variable. The variance of X, or $Var(X)$, is defined as

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

Theorem 2. *Let* $a, b \in \mathbb{R}$ *and* X *and* Y *be random variables, the properties of variance are*

1.) $Var(X + a) = Var(X)$*,*

2.) $Var(aX) = a^2 Var(X)$*,*

3.) $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y)$*.*

Definition 6. Let X be a random variable. The generating function of X, $G_X(t)$, is defined as

$$G_X(t) = E(t^X).$$

Definition 7. Let X be a random variable. The moment generating function of X or $M_X(t)$, is defined as

$$M_X(t) = E(e^{tX}).$$

Definition 8. If X and Y are discrete random variables, then the probability mass function of X given $Y = y$ is defined as

$$p_{X|Y}(x|y) = \frac{P(X = x, Y = y)}{P(Y = y)}.$$

Definition 9. The conditional expectation of X given $Y = y$ is defined by

$$E(X|Y = y) = \sum_x x p_{X|Y}(x|y)$$

Note that for any random varibles X and Y,

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

If Y is a discrete random variable, then the above formula is equivalent to

$$E(X) = \sum_y E(X|Y = y)P(Y = y).$$

Definition 10. For any random variables X and Y:

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

2.2 Count distribution

We now present some count distributions such as Bernoulli distribution, Binomial distribution and Poisson distribution.

Definition 11. (Bernoulli) A random variable X is said to have a Bernoulli distribution with parameter p, $X \sim Ber(p)$, if

$$P(X = 1) = p = 1 - P(X = 0) \ \ for\ p \in [0, 1].$$

The probability mass function f of this distribution over possible outcomes k can be alternatively written as

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

Theorem 3. *Properties of Bernoulli distribution with parameter $p \in [0,1]$ are given as follows.*

1.) $E(X) = p$,

2.) $Var(X) = p(1-p)$,

3.) $G_X(t) = 1 - p + pt$,

4.) $M_X(t) = 1 - p + pe^t$.

Definition 12. (Binomial) A random variable X is said to have the binomial distribution with parameters $n \in \mathbb{N}$ and $p \in [0,1]$, $X \sim B(n,p)$, if

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

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Remark 2. If $X \sim B(n,p)$, then $P(X = k)$ is the probability of getting exactly k successes in n trials.

Theorem 4. *Let X be a binomial random variable with parameters n and p. Then the following properties hold.*

1.) $E(X) = np$,

2.) $Var(X) = np(1-p)$,

3.) $G_X(t) = (1 - p + pt)^n$,

4.) $M_X(t) = (1 - p + pe^t)^n$.

Definition 13. (Poisson) A random variable X is said to have a Poisson distribution with parameter $\lambda \in [0,\infty)$, $X \sim Poi(\lambda)$,

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \ \text{ for } x = 0, 1, \ldots$$

Theorem 5. *Let X be a poisson random variable with parameters λ. Then the following properties hold.*

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

2.) $Var(X) = \lambda$,

3.) $G_X(t) = e^{\lambda(t-1)}$,

4.) $M_X(t) = e^{\lambda(e^t-1)}$.

Definition 14. (Inverse-gaussian) A random variable X is said to have an Inverse Gaussian distribution with parameter $\phi > 0$ and $\mu > 0$, $X \sim IG(\mu, \phi)$, if its density function is defined as

$$f(x; \mu, \phi) = [\frac{\phi}{2\pi x^3}]^{\frac{1}{2}} \exp\left(\frac{-\phi(x-\mu)^2}{2\mu^2 x}\right)$$

for $t < \phi/2$.

Theorem 6. *Let X be an inverse-gaussian random variable with parameters $\phi > 0$ and $\mu > 0$. Then the following properties hold.*

1.) $E(X) = \mu$,

2.) $Var(X) = \frac{\mu^3}{\phi}$,

3.) $M_X(t) = \exp\{\frac{\phi}{\mu}(1 - \sqrt{1 - 2\mu^2\phi^{-1}t})\}$.

Theorem 7. *Let X be an inverse-gaussian random variable with parameters $\phi > 0$ and $\mu = 1$. Then the following properties hold.*

1.) $E[X] = 1$,

2.) $Var[X] = \frac{1}{\phi}$,

3.) $M_X(t) = \exp\{\phi(1 - \sqrt{1 - 2\phi^{-1}t})\}$.

Definition 15. (Compound random variable) Let N be a nonnegative integer-valued random variable and let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed(i.i.d) positive random variables that are independent of N. The random variable

$$S_N = \sum_{i=1}^{N} X_i \tag{2.2.1}$$

is called a compound random variable.

Theorem 8. *Properties of compound random variable S_N defined in Definition 15 are given as follows.*

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

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

3.) $G_{S_N}(t) = G_N(G_X(t))$.

2.3 Binomial Thinning Operator

In this section, we introduce the definition of the binomial thinning operator which is the main tool in constructing an integer valued time series. Moreover, we present its properties with their proofs.

Definition 16. Let M be a non-negative integer-valued random variable and $\alpha \in [0, 1]$. The operator $\alpha\circ$ on M is referred to as the binomial thinning of M and is defined as

$$\alpha \circ M = \sum_{i=1}^{M} Y_i, \tag{2.3.1}$$

where $\{Y_i, i = 1, 2, \ldots\}$ is a sequence of i.i.d Berniulli random variables with mean α and is independent of M.

Theorem 9. *Let Z be an non-negative integer-valued random variable. The following properties hold.*

1.) $E(\alpha \circ Z) = \alpha E(Z)$,

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

3.) $M_{\alpha\circ Z}(t) = G_Z(1 - \alpha + \alpha e^t)$, *4.)* $E(Z(\alpha \circ Z)) = \alpha E(Z^2)$,

5.) $E((\alpha_1 \circ Z)(\alpha_2 \circ Z)) = \alpha_1 \alpha_2 E(Z^2)$,

6.) $Cov(\alpha \circ Z, Z) = \alpha Var(Z)$,

7.) $Cov(\alpha_1 \circ Z, \alpha_2 \circ Z) = \alpha_1 \alpha_2 Var(Z)$.

Proof. 1.)

$$\begin{aligned} E(\alpha \circ Z) &= E\left(\sum_{i=1}^{Z} Y_i\right) \\ &= E\left(E\left(\sum_{i=1}^{Z} Y_i \Big| Z\right)\right) \\ &= E\left(\sum_{i=1}^{Z} E(Y_i \big| Z)\right) \\ &= E\left(\sum_{i=1}^{Z} \alpha\right) \qquad (2.3.2) \\ &= \alpha E(Z), \end{aligned}$$

where we use the fact that $Y_i \sim Ber(\alpha)$ (for $i \geq 1$) to obtain (2.3.2).

2.) By the definition of conditional variance, we can show that

$$\begin{aligned} Var(\alpha \circ Z) &= Var\left(E(\alpha \circ Z|Z)\right) + E\left(Var(\alpha \circ Z|Z)\right) \\ &= Var\left(E\left(\sum_{i=1}^{Z} Y_i \Big| Z\right)\right) + E\left(Var\left(\sum_{i=1}^{Z} Y_i \Big| Z\right)\right) \\ &= Var\left(\sum_{i=1}^{Z} E(Y_i \big| Z)\right) + E\left(\sum_{i=1}^{Z} Var(Y_i \big| Z)\right) \\ &= Var\left(\sum_{i=1}^{Z} \alpha\right) + E\left(\sum_{i=1}^{Z} \alpha(1-\alpha)\right) \qquad (2.3.3) \\ &= \alpha^2 Var(Z) + \alpha(1-\alpha)E(Z), \end{aligned}$$

where we use the fact that $Y_i \sim Ber(\alpha)$ (for $i \geq 1$) to obtain (2.3.3).

3.)

$$\begin{aligned}
M_{\alpha\circ Z}(t) &= E\left(e^{(\alpha\circ Z)t}\right) \\
&= E\left(e^{t\sum_{i=1}^{Z} Y_i}\right) \\
&= E\left(E\left(e^{t\sum_{i=1}^{Z} Y_i}\middle| Z\right)\right) \\
&= E\left(\prod_{i=1}^{Z} E\left(e^{tY_i}\right)\right) \\
&= E\left(\prod_{i=1}^{Z} \left(1-\alpha+\alpha e^t\right)\right) \qquad (2.3.4) \\
&= E\left((1-\alpha+\alpha e^t)^Z\right) \\
&= G_Z(1-\alpha+\alpha e^t),
\end{aligned}$$

where we use the fact that $Y_i \sim Ber(\alpha)$ to obtain (2.3.4).

4.)

$$\begin{aligned}
E(Z(\alpha\circ Z)) &= E\left(Z\sum_{i=1}^{Z} Y_i\right) \\
&= E\left(E\left(Z\sum_{i=1}^{Z} Y_i \middle| Z\right)\right) \\
&= E\left(Z\sum_{i=1}^{Z} E(Y_i)\right) \\
&= E\left(Z\sum_{i=1}^{Z} \alpha\right) \qquad (2.3.5) \\
&= \alpha E\left(Z^2\right),
\end{aligned}$$

where we use the fact that $Y_i \sim Ber(\alpha)$ to obtain (2.3.5).

5.) Define $\alpha_1 \circ Z = \sum_{i=1}^{Z} Y_i$ and $\alpha_2 \circ Z = \sum_{j=1}^{Z} W_j$, $Y_i \sim Ber(\alpha_1)$ and $W_j \sim Ber(\alpha_2)$

then

$$
\begin{aligned}
E((\alpha_1 \circ Z)(\alpha_2 \circ Z)) &= E\left(\left(\sum_{i=1}^{Z} Y_i\right)\left(\sum_{j=1}^{Z} W_j\right)\right) \\
&= E\left(E\left(\left(\sum_{i=1}^{Z} Y_i\right)\left(\sum_{j=1}^{Z} W_j\right)\Big| Z\right)\right) \\
&= E\left(\sum_{i=1}^{Z}\sum_{j=1}^{Z} E(Y_i W_j | Z)\right) \\
&= E\left(\sum_{i=1}^{Z}\sum_{j=1}^{Z} E(Y_i)E(W_j)\right) \\
&= E\left(\sum_{i=1}^{Z}\sum_{j=1}^{Z} \alpha_1\alpha_2\right) \qquad (2.3.6) \\
&= \alpha_1\alpha_2 E\left(\sum_{i=1}^{Z}\sum_{j=1}^{Z} 1\right) \\
&= \alpha_1\alpha_2 E\left(Z^2\right),
\end{aligned}
$$

where we use the fact that $Y_i \sim Ber(\alpha)$ to obtain (2.3.6).

6.) By the definition of covariance function and 1.), we can show that

$$
\begin{aligned}
Cov(\alpha \circ Z, Z) &= E(Z(\alpha \circ Z)) - E(Z)E(\alpha \circ Z) \\
&= \alpha E(Z^2) - \alpha E(Z)E(Z) \\
&= \alpha(E(Z^2) - E^2(Z)) \\
&= \alpha Var(Z).
\end{aligned}
$$

7.) By the definition of covariance function, 1.) and 5.), we can show that

$$
\begin{aligned}
Cov(\alpha_1 \circ Z, \alpha_2 \circ Z) &= E((\alpha_1 \circ Z)(\alpha_2 \circ Z)) - E(\alpha_1 \circ Z)E(\alpha_2 \circ Z) \\
&= \alpha_1\alpha_2 E(Z^2) - \alpha_1\alpha_2 E(Z)E(Z) \\
&= \alpha_1\alpha_2(E(Z^2) - E^2(Z)) \\
&= \alpha_1\alpha_2 Var(Z).
\end{aligned}
$$

□

Chapter 3

Main work

In this chapter, we construct a new moving average time series model based on the Poisson-Inverse Gaussion model. The organization of this chapter is as follows. We first discuss the definition and properties of the Poisson-Inverse Gaussion distribution, in Section 3.1. The construction of the new model, Poisson-Inverse Gaussian, is given is Section 3.2. Finally, numerical simulation of distribution of such data are given in Section 3.3.

3.1 Poisson Inverse Gaussian

Definition 17. (Mixed Poisson distribution) A random variable Y follows a mixed Poisson distribution if $Y|Z = z \sim Poisson\ (\lambda z)$, for $\lambda > 0$, where Z is some non-negative random variable.

Consequently, the probability mass function of Y for $y \geq 0$, can be derived

$$\begin{aligned} P(Y = y) &= \int_0^\infty P(Y = y|Z = z) g_\phi(z) dz \\ &= \int_0^\infty \frac{e^{-\lambda z}(\lambda z)^y}{y!} dG_\phi(z), \end{aligned}$$

where $G_\phi(\cdot)$ is the distribution function of Z. The parameter ϕ denotes the parameter vector associated to the distribution of Z. We denote $Y \sim MP(\lambda, \phi)$.

Theorem 10. *Properties of mixed Poisson distribution with parameter ϕ and $\lambda > 0$ defined in Definition 17 are given as follows.*

1.) $E(Y) = \lambda E(Z)$*,*

2.) $Var(Y) = \lambda E(Z) + \lambda^2 Var(Z)$*,*

3.) $M_Y(t) = M_Z(\lambda(e^t - 1))$*.*

for t *belonging some interval containing the value zero.*

Proof. 1.)

$$\begin{aligned} E(Y) &= E(E(Y|Z)) \\ &= E(\lambda Z) \\ &= \lambda E(Z). \end{aligned}$$

2.)

$$\begin{aligned} Var(Y) &= Var(E(Y|Z)) + E(Var(Y|Z)) \\ &= Var(\lambda Z) + E(\lambda Z) \\ &= \lambda E(Z) + \lambda^2 Var(Z) \end{aligned}$$

3.)

$$\begin{aligned} M_Y(t) &= E(e^{Yt}) \\ &= E(E(e^{Yt}|Z)) \\ &= E(e^{\lambda Z(e^t - 1)}) \\ &= M_Z(\lambda(e^t - 1)). \end{aligned}$$

□

Theorem 11. *Let* $Y \sim MP(\lambda, \phi)$*. Then,* $\alpha \circ Y \sim MP(\alpha\lambda, \phi)$ *for* $\alpha \in [0, 1)$*.*

Proof. Consider $\alpha \circ Y|Z$. By the Theorem 9(3), we can show that

$$\begin{aligned} M_{\alpha\circ Y|Z}(t) &= E((1 - \alpha + \alpha e^t)^Y |Z) \qquad (3.1.1) \\ &= e^{\lambda z(1-\alpha+\alpha e^t - 1)} \\ &= e^{\lambda z(\alpha e^t - \alpha)} \\ &= e^{\lambda z\alpha(e^t - 1)}, \end{aligned}$$

Then $\alpha \circ Y|Z \sim Poisson(\lambda z\alpha)$.

$$\begin{aligned} M_{\alpha\circ Y}(t) &= E(e^{t(\alpha\circ Y)}) \\ &= E(E(e^{t(\alpha\circ Y)}|Z)) \\ &= E(e^{\lambda z\alpha(e^t-1)}) \qquad (3.1.2) \\ &= M_Z(\lambda\alpha(e^t-1)), \end{aligned}$$

where we use the fact that $\alpha \circ Y|Z \sim Poisson(\lambda z\alpha)$ in (3.1.2). □

Definition 18. (Poisson-inverse Gaussian) A random variable X is said to have a Poisson-inverse Gaussian distribution of $X|Z = z \sim Poisson(\lambda z)$ and $Z \sim IG(\mu, \phi)$. We denote $X \sim PIG(\mu, \phi)$.

Remark 3. A random variable X is said to have a Poisson-Inverse Gaussian distribution with parameter $\lambda > 0$, $\phi > 0$ and $\mu > 0$, $X \sim PIG(\mu, \phi)$, if its density function, $f(x; \lambda, \mu, \phi)$ is defined as

$$f(x; \lambda, \mu, \phi) = \int_0^\infty \frac{e^{-\lambda z}(\lambda z)^x}{x!}[\frac{\phi}{2\pi z^3}]^{\frac{1}{2}} \exp\left(\frac{-\phi(z-\mu)^2}{2\mu^2 z}\right) dz$$

for $t < \phi/2$.

Theorem 12. *Let X be a Poisson-Inverse Gaussian distribution with parameter $\lambda > 0$, $\mu = 1$ and $\phi > 0$. Then the following properties hold.*

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

2.) $Var(X) = \lambda(1 + \lambda\phi^{-1})$,

3.) $M_X(t) = \exp\{\phi(1 - \sqrt{1 - 2\phi^{-1}\lambda(e^t - 1)})\}$, *for* $t < log(1 + \phi/(2\lambda))$.

Proof. 1.) By Theorem 7(1) and Theorem 10(1), we can show that

$$\begin{aligned} E(X) &= \lambda E(Z) \\ &= \lambda. \end{aligned}$$

2.) By Theorem 7(2), Theorem 10(1) and theorem 29(2), we can show that

$$\begin{aligned} Var(X) &= \lambda E(Z) + \lambda^2 Var(Z) \\ &= \lambda + \lambda^2\phi^{-1} \\ &= \lambda(1+\lambda\phi^{-1}). \end{aligned}$$

3.) By Theorem 7(3) and Theorem 10(3), we can show that

$$\begin{aligned} M_X(t) &= M_Z(\lambda(e^t-1)) \\ &= \exp\{\phi(1-\sqrt{1-2\phi^{-1}\lambda(e^t-1)})\}. \end{aligned}$$

□

3.2 (Poisson-inverse Gaussian INMA(q))

In order to present our class of Poisson-Inverse Gaussian INMA(q) processes and discuss its properties. We begin this section by giving the definition of the Integer-Valued moving average model.

Definition 19. The Integer-Valued moving average model of order q, denoted by INMA(q), is defined as

$$X_n = \alpha_1 \circ \epsilon_{n-1} + \alpha_2 \circ \epsilon_{n-2} + \cdots + \alpha_q \circ \epsilon_{n-q} + \epsilon_n$$

for $n \geq 1$, with $\{\epsilon_n\}_{n=1}^{\infty}$ be a sequence of i.i.d random variables independent of ϵ_j, for $j \leq n$, for all n.

Definition 20. Poisson-Inverse Gaussian INMA(q) The sequence $\{X_n\}_{n=1}^{\infty}$ is said to be a Poisson-inverse Gaussian INMA(q) process, defined in Definition 19 if $\{\epsilon_i\}_{i=1}^{\infty}$ is a sequence of i.i.d $PIG(\mu, \phi)$.

Theorem 13. *The mgf of X_n is defined as*

$$M_{X_n}(t) = \exp\left((q+1)\phi - \sum_{i=1}^{q} \phi\sqrt{1-2\phi^{-1}\lambda\alpha_i(e^t-1)} - \sqrt{1-2\phi^{-1}\lambda(e^t-1)}\right).$$

and $\{X_n, n \in \mathbb{N}\}$ is a stationary process.

Proof. Since $\{\epsilon_n\}_{n=0}^{\infty}$ is a sequence independent random variable, from theorem 9(3),

$$\begin{aligned} M_{X_n}(t) =& M_{\sum_{i=1}^{q}\alpha_i\circ\epsilon_{n-i}+\epsilon_n}(t) \\ =& E(e^{(\sum_{i=1}^{q}\alpha_i\circ\epsilon_{n-i}+\epsilon_n)t}) \\ =& E(e^{t\sum_{i=1}^{q}\alpha_i\circ\epsilon_{n-i}})E(e^{\epsilon_n t}) \\ =& \prod_{i=1}^{q} E(e^{t(\alpha_i\circ\epsilon_{n-i})})E(e^{\epsilon_n t}) \\ =& \prod_{i=1}^{q} M_{\alpha_i\circ\epsilon_{n-i}}(t)M_{\epsilon_n t}(t) \\ =& \prod_{i=1}^{q} M_Z(\lambda\alpha_i(e^t-1))\times M_{\epsilon_n}(t) \end{aligned} \tag{3.2.1}$$

where we use the fact that theorem 10(3) and theorem 11 in 3.2.1. Substitute $M_Z(t)$ and $M_\epsilon(t)$ from Theorem 7(3) and Theorem 12(3) respectively, then the mgf of X_n is

$$\begin{aligned} M_{X_n}(t) &= \left(\prod_{i=1}^{q}\exp\{\phi(1-\sqrt{1-2\phi^{-1}\lambda\alpha_i(e^t-1)})\}\right)\times\exp\{\phi(1-\sqrt{1-2\phi^{-1}\lambda(e^t-1)})\} \\ &= \exp\left((q+1)\phi-\sum_{i=1}^{q}\phi\sqrt{1-2\phi^{-1}\lambda\alpha_i(e^t-1)}-\sqrt{1-2\phi^{-1}\lambda(e^t-1)}\right). \end{aligned}$$

Since the moment generating function does not depend on n, so $\{X_n, n\in\mathbb{N}\}$ is a stationary process. □

Theorem 14. *Let a sequence $\{X_n\}_{n=0}^{\infty}$ be a PIGINMA(q) defined in Definition 20, if $Z\sim IG(1,\phi)$. The mean, variance and covariance of (X_n) are as follows.*

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

2.) $Var(X_n)=\sum_{i=1}^{q}[\alpha_i^2\lambda(1+\phi^{-1}\lambda)+\alpha_i(1-\alpha_i)\lambda]+\lambda(1+\phi^{-1}\lambda)$,

3.) $Cov(X_n,X_{n-k})=\begin{cases}\lambda(1+\phi^{-1}\lambda)\left(\alpha_k+\sum_{i=1}^{q-k}\alpha_i\alpha_{k+i}\right) & \text{for } k<q,\\ \alpha_q\lambda(1+\phi^{-1}\lambda) & \text{for } k=q\\ 0 & \text{for } k>q.\end{cases}$

4.) $Corr(X_n, X_{n-k}) = \begin{cases} \frac{\lambda(1+\phi^{-1}\lambda)\left(\alpha_k+\sum_{i=1}^{q-k}\alpha_i\alpha_{k+i}\right)}{\sum_{i=1}^{q}[\alpha_i^2\lambda(1+\phi^{-1}\lambda)+\alpha_i(1-\alpha_i)\lambda]+\lambda(1+\phi^{-1}\lambda)} & \text{for } k<q \\ \frac{\alpha_q\lambda(1+\phi^{-1}\lambda)}{\sum_{i=1}^{q}[\alpha_i^2\lambda(1+\phi^{-1}\lambda)+\alpha_i(1-\alpha_i)\lambda]+\lambda(1+\phi^{-1}\lambda)} & \text{for } k=q \\ 0 & \text{for } k>q. \end{cases}$

Proof. 1.) From the Definition 19 and Theorem 9(1), we can show that

$$\begin{aligned} &E(X_n) \\ &=E(\alpha_1\circ\epsilon_{n-1}+\alpha_2\circ\epsilon_{n-2}+\cdots+\alpha_q\circ\epsilon_{n-q}+\epsilon_n) \\ &=E(\alpha_1\circ\epsilon_{n-1})+E(\alpha_2\circ\epsilon_{n-2})+\cdots+E(\alpha_q\circ\epsilon_{n-q})+E(\epsilon_n) \\ &=\sum_{i=1}^{q}\alpha_i E(\epsilon_{n-i})+E(\epsilon_n) \\ &=\sum_{i=1}^{q}\alpha_i\lambda+\lambda \\ &=\lambda\left(\sum_{i=1}^{q}\alpha_i+1\right). \end{aligned}$$

2.) From the Definition 19 and Theorem 9(2), we can show that

$$\begin{aligned} &Var(X_n) \\ &=Var(\alpha_1\circ\epsilon_{n-1}+\alpha_2\circ\epsilon_{n-2}+\cdots+\alpha_q\circ\epsilon_{n-q}+\epsilon_n) \\ &=\sum_{i=1}^{q}Var(\alpha_i\circ\epsilon_{n-i})+Var(\epsilon_n) \\ &=\sum_{i=1}^{q}[\alpha_i^2Var(\epsilon_{n-i})+\alpha_i(1-\alpha_i)E(\epsilon_{n-i})]+\mu(1+\phi^{-1}\mu) \\ &=\sum_{i=1}^{q}[\alpha_i^2\mu(1+\phi^{-1}\mu)+\alpha_i(1-\alpha_i)\mu]+\mu(1+\phi^{-1}\mu). \end{aligned}$$

3.) To find $Cor(X_n, X_{n-k})$, we consider four different cases which are $k=1$, $k<q, k=q$, and $k>q$.

For $k=1$, since $\{\epsilon_i\}$ are independent, theorem 9(5),(6) and theorem 12, we have

$Cov(X_n, X_{n-k}) = Cov(X_n, X_{n-1})$

$=Cov(\alpha_1\circ\epsilon_{n-1}+\alpha_2\circ\epsilon_{n-2}+\cdots+\alpha_q\circ\epsilon_{n-q}+\epsilon_n, \alpha_1\circ\epsilon_{n-2}+\alpha_2\circ\epsilon_{n-3}+\cdots+\alpha_q\circ\epsilon_{n-q-1}+\epsilon_{n-1})$

$=Cov(\alpha_1\circ\epsilon_{n-1},\epsilon_{n-1})+Cov(\alpha_2\circ\epsilon_{n-2},\alpha_1\circ\epsilon_{n-2})+\cdots+Cov(\alpha_q\circ\epsilon_{n-q},\alpha_{q-1}\circ\epsilon_{n-q})$

$=\alpha_1Var(\epsilon_{n-1})+\alpha_1\alpha_2Var(\epsilon_{n-2})+\cdots+\alpha_{q-1}\alpha_qVar(\epsilon_{n-q})$

$=\lambda(1+\phi^{-1}\lambda)(\alpha_1+\alpha_1\alpha_2+\cdots+\alpha_{q-1}\alpha_q).$

For $k < q$

Since $\{\epsilon_i\}$ are independent, Theorem 9(5-6) and Theorem 12, we have

$Cov(X_n, X_{n-k})$

$=Cov(\alpha_1 \circ \epsilon_{n-1} + \cdots + \alpha_q \circ \epsilon_{n-q} + \epsilon_n, \alpha_1 \circ \epsilon_{n-k-1} + \cdots + \alpha_q \circ \epsilon_{n-k-q} + \epsilon_{n-k})$

$=Cov(\alpha_k \circ \epsilon_{n-k}, \epsilon_{n-k}) + Cov(\alpha_{k+1} \circ \epsilon_{n-k-1}, \alpha_1 \circ \epsilon_{n-k-1}) + \cdots + Cov(\alpha_{q-k} \circ \epsilon_{n-q}, \alpha_q \circ \epsilon_{n-q})$

$=\alpha_k Var(\epsilon_{n-k}) + \alpha_1\alpha_{k+1} Var(\epsilon_{n-k-1}) + \cdots + \alpha_{q-k}\alpha_q Var(\epsilon_{n-q})$

$=\lambda(1 + \phi^{-1}\lambda)(\alpha_k + \alpha_1\alpha_{k+1} + \alpha_2\alpha_{k+2} + \cdots + \alpha_{q-k}\alpha_q).$

For $k = q$

Since $\{\epsilon_i\}$ are independent, Theorem 9(5) and Theorem 12, we have

$Cov(X_n, X_{n-q})$

$=Cov(\alpha_1 \circ \epsilon_{n-1} + \cdots + \alpha_q \circ \epsilon_{n-q} + \epsilon_n, \alpha_1 \circ \epsilon_{n-q-1} + \cdots + \alpha_q \circ \epsilon_{n-q-q} + \epsilon_{n-q})$

$=Cov(\alpha_q \circ \epsilon_{n-q}, \epsilon_{n-q})$

$=\alpha_q Var(\epsilon_{n-q})$

$=\lambda\alpha_q(1 + \phi^{-1}\lambda).$

Consider $k > q$

Since $\{\epsilon_i\}$ are independent, Theorem 9(5) and Theorem 12, we have

$Cov(X_n, X_{n-k})$

$=Cov(\alpha_1 \circ \epsilon_{n-1} + \cdots + \alpha_q \circ \epsilon_{n-q} + \epsilon_n, \alpha_1 \circ \epsilon_{n-k-1} + \cdots + \alpha_q \circ \epsilon_{n-k-q} + \epsilon_{n-k})$

$=0.$

Hence

$$Cov(X_n, X_{n-k}) = \begin{cases} \lambda(1+\phi^{-1}\lambda)\left(\alpha_k + \sum_{i=1}^{q-k} \alpha_i\alpha_{k+i}\right) & \text{for } k < q, \\ \alpha_q\lambda(1+\phi^{-1}\lambda) & \text{for } k = q \\ 0 & \text{for } k > q. \end{cases}$$

4.) From 2.) and 3.), the autocorrelation function of X is

$$
\begin{aligned}
Corr(X_n, X_{n-k}) &= \frac{Cov(X_n, X_{n-k})}{\sqrt{Var(X_n)Var(X_{n-k})}} \\
&= \frac{Cov(X_n, X_{n-k})}{Var(X_n)} \\
&= \begin{cases} \frac{\lambda(1+\phi^{-1}\lambda)\left(\alpha_k+\sum_{i=1}^{q-k}\alpha_i\alpha_{k+i}\right)}{\sum_{i=1}^{q}[\alpha_i^2\lambda(1+\phi^{-1}\lambda)+\alpha_i(1-\alpha_i)\lambda]+\lambda(1+\phi^{-1}\lambda)} & for\ k < q \\ \frac{\alpha_q\lambda(1+\phi^{-1}\lambda)}{\sum_{i=1}^{q}[\alpha_i^2\lambda(1+\phi^{-1}\lambda)+\alpha_i(1-\alpha_i)\lambda]+\lambda(1+\phi^{-1}\lambda)} & for\ k = q \\ 0 & for\ k > q. \end{cases}
\end{aligned}
$$

□

3.3 Data Simulation

In this section, we study numerical simulation of data distribution of the Poisson-Inverse Gaussion INMA(q) model in different settings.

3.3.1 Data Generation Algorithm

For our simulation experiment, we use the following algorithm to generate data. Following Poisson-Inverse Gaussian INMA(q) model:

$$X_n = \alpha_1 \circ \epsilon_{n-1} + \alpha_2 \circ \epsilon_{n-2} + \cdots + \alpha_q \circ \epsilon_{n-q} + \epsilon_n$$

1. Set parameters $\lambda, \mu, \phi, \alpha, n, q$ in INMA(q) model where $\epsilon_n \sim PIG(\lambda\mu, \phi)$ with parameter $\alpha_i \in [0, 1]$ where $\alpha_i = \alpha_j$ for all $i, j = 1, 2, \ldots, q$, $\mu = 1$ and λ, ϕ, n, q are a positive integer.

2. Generate a positive finite integer matrix

$$Eps = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \ldots & \epsilon_{1n} \\ \epsilon_{21} & \epsilon_{22} & \ldots & \epsilon_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{t1} & \epsilon_{t2} & \ldots & \epsilon_{tn} \end{bmatrix}$$

where $\epsilon_i,\ i = 0, 1, \cdots, n$ is generated from the Poisson-inverse Gaussian distribution with parameter λ, ϕ.

3. Generate a positive finite integer matrix

$$\text{Thin} = \begin{bmatrix} T_{11} & T_{12} & \dots & T_{1n} \\ T_{21} & T_{22} & \dots & T_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ T_{t1} & T_{t2} & \dots & T_{tn} \end{bmatrix}$$

   where $T_{i,j},\ i = 1, 2, \dots, t$ and $j = 1, 2, \dots, n$ is generated from Binomial distribution with parameter α.

4. Generate a positive finite integer matrix

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{t1} & X_{t2} & \dots & X_{tn} \end{bmatrix}$$

   where $X_{i,j} = \sum_{l=i}^{i+q-1} T_{l,j} + \epsilon_{i+q,j},\ i = 1, 2, \dots, t$ and $j = 1, 2, \dots, n$.

3.3.2 Numerical Simulation of PIGINMA(1) model

In this section, we study numerical simulation of the PIGINMA(1) model:

$$X_n = \alpha_1 \circ \epsilon_{n-1} + \epsilon_n.$$

In our study, we consider 3 different settings,

1. Fix $\phi = 0.5, \alpha = 0.5, n = 1, 2, \dots, 100$ and $\mu = 1$, consider $\lambda \in \{1, 1.4, 1.8, 2.4, 8, 16, 32\}$.

2. Fix $\phi = 0.5$, $n = 1, 2, \dots, 100$ and $\mu = 1$

   a) $\lambda = 1$, consider $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$.

   b) $\lambda = 32$, consider $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$.

3. We generate 100 series of PIGINMA(1). Fix $\phi = 0.5$, $n = 1, 2, \dots, 100$ and $\mu = 1$.

   a) Consider $\lambda = 1$

   b) Consider $\lambda = 32$

4. We generate 100 series of PIGINMA(1). Fix $\phi = 0.5$, $n = 1, 2, \ldots, 100$ and $\mu = 1$, consider X_{10} when $\lambda \in \{1, 4, 32\}$.

1. Comparison the Data of PIGINMA(1) model with different λ

For PIGINMA(1) model we set the parameters $\phi = 0.5, \alpha = 0.5$, $n = 100$ and $t = 1$.

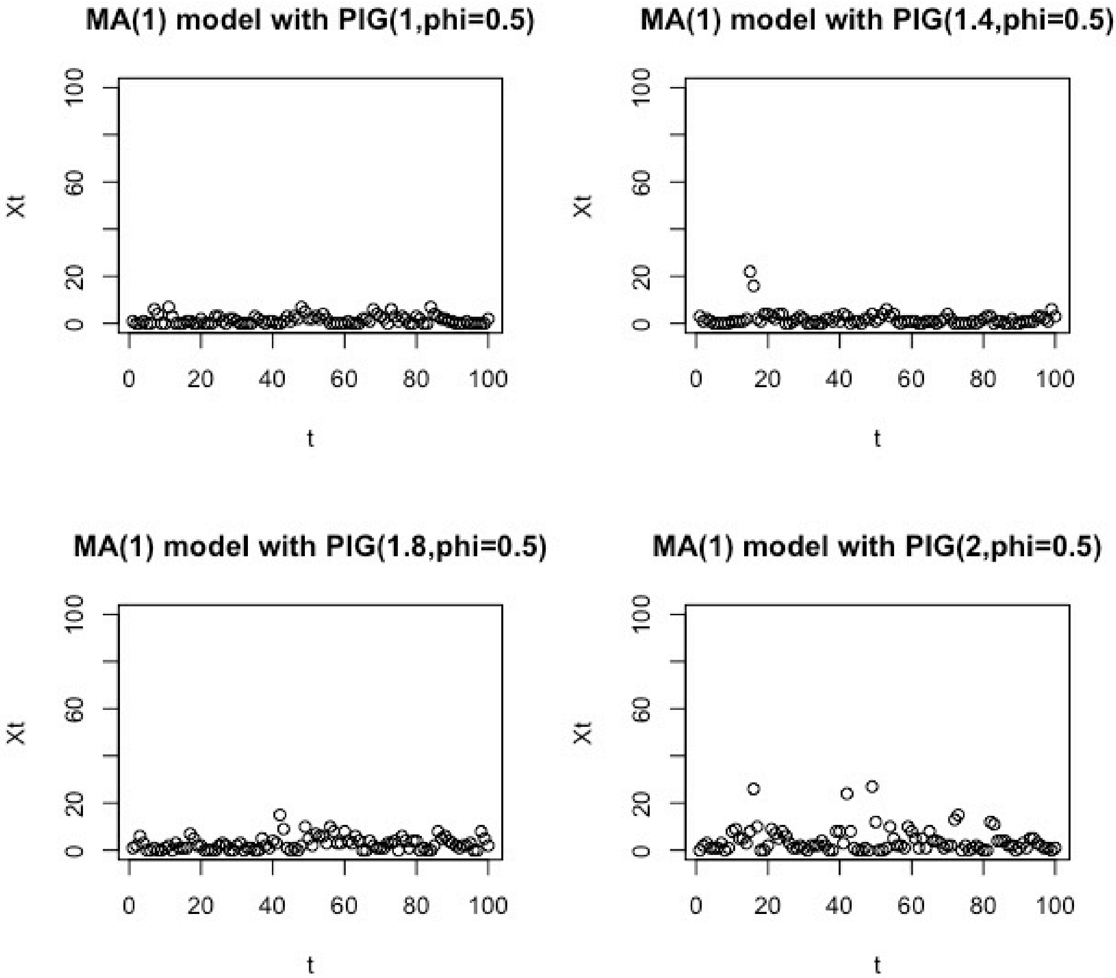

Figure 3.1: Scatter plot of data generated from PIGINMA(1) model with parameters $\phi = 0.5$, $\mu = 1$ and $\lambda \in \{1, 1.4, 1.8, 2\}$

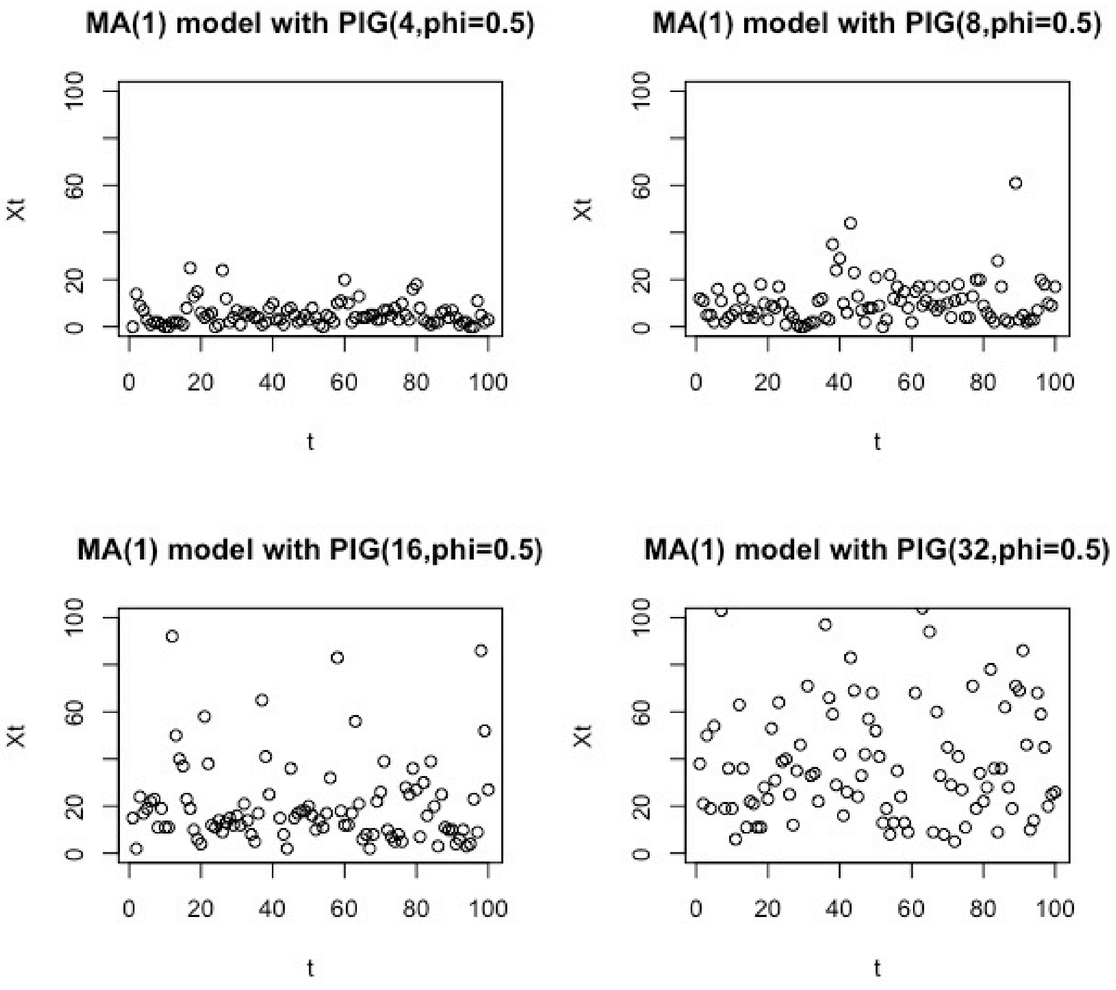

Figure 3.2: Scatter plot of data generated from PIGINMA(1) model with parameters $\phi = 0.5$, $\mu = 1$ and $\lambda \in \{4, 8, 16, 32\}$

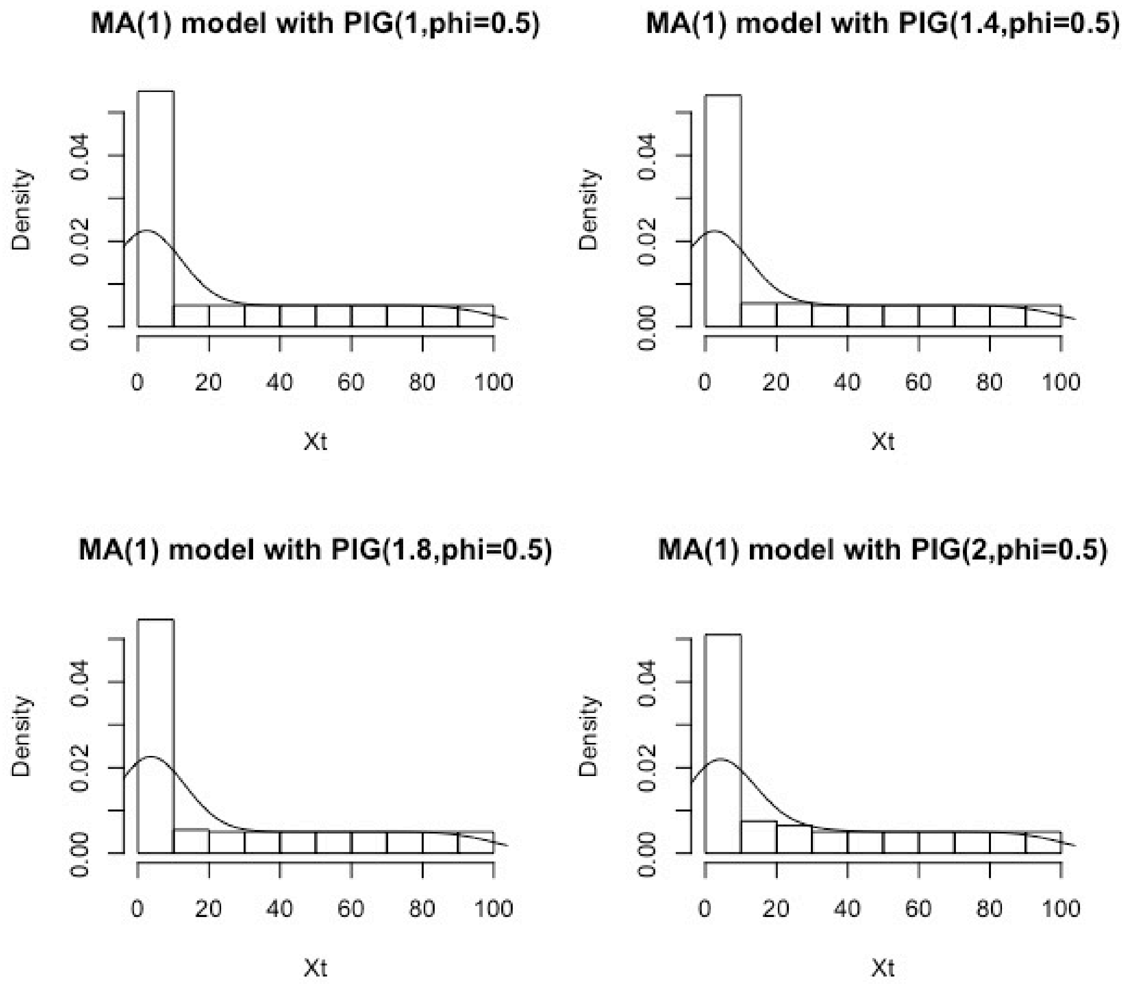

Figure 3.3: Histograms of data generated from PIGINMA(1) model with parameters $\phi = 0.5$, $\mu = 1$ and $\lambda \in \{1, 1.4, 1.8, 2\}$

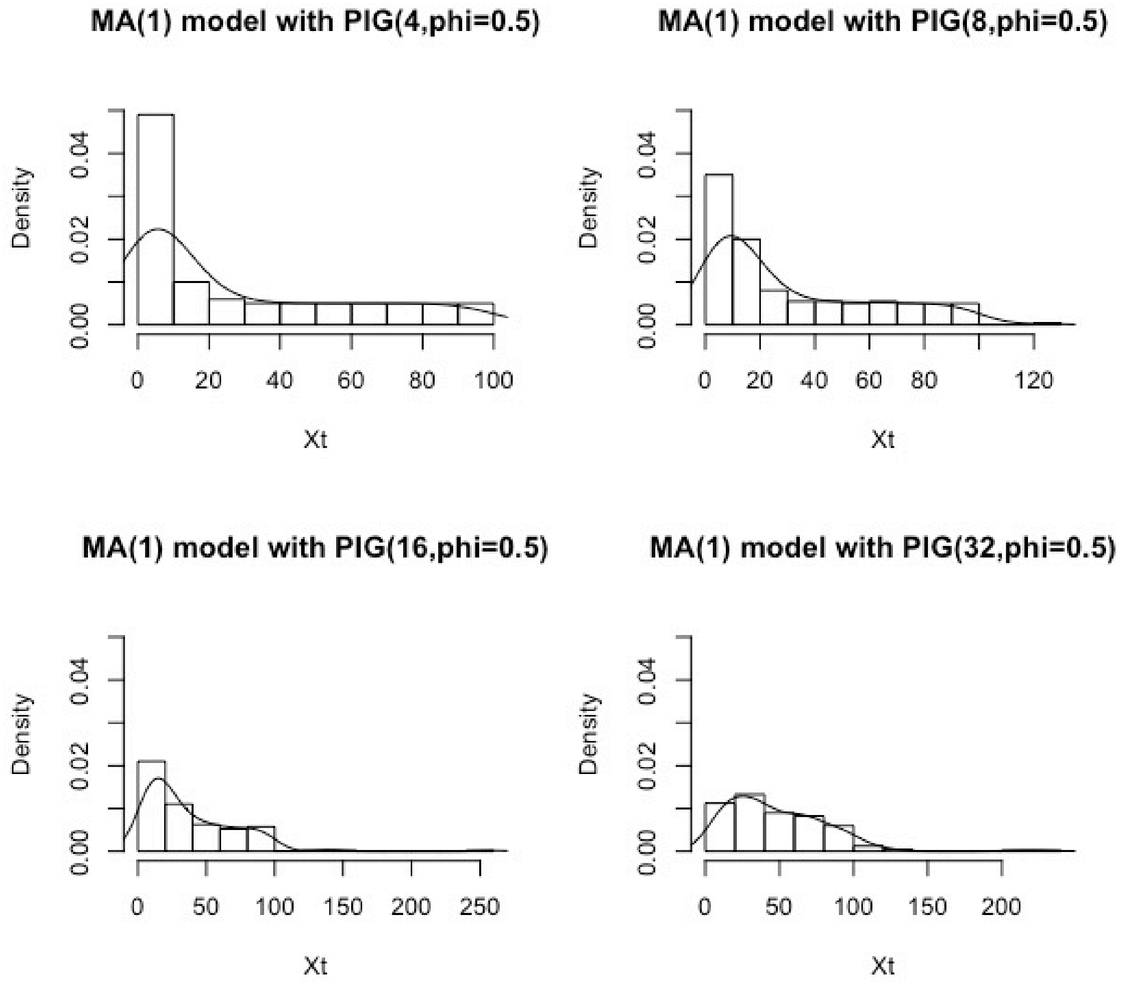

Figure 3.4: Histograms of data generated from PIGINMA(1) model with parameters $\phi = 0.5$, $\mu = 1$ and $\lambda \in \{4, 8, 16, 32\}$

By the distribution plots and histograms, we can see that data skewed to the left and the frequency of the data with small value decreases when λ increases.

2. Comparision the Data of PIGINMA(1) model with different λ and α

For PIGINMA(1) model we set the parameter $\phi = 0.5, \mu = 1$ and $n = 1, 2, \ldots, 100$.

a) $\lambda = 1$, consider $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

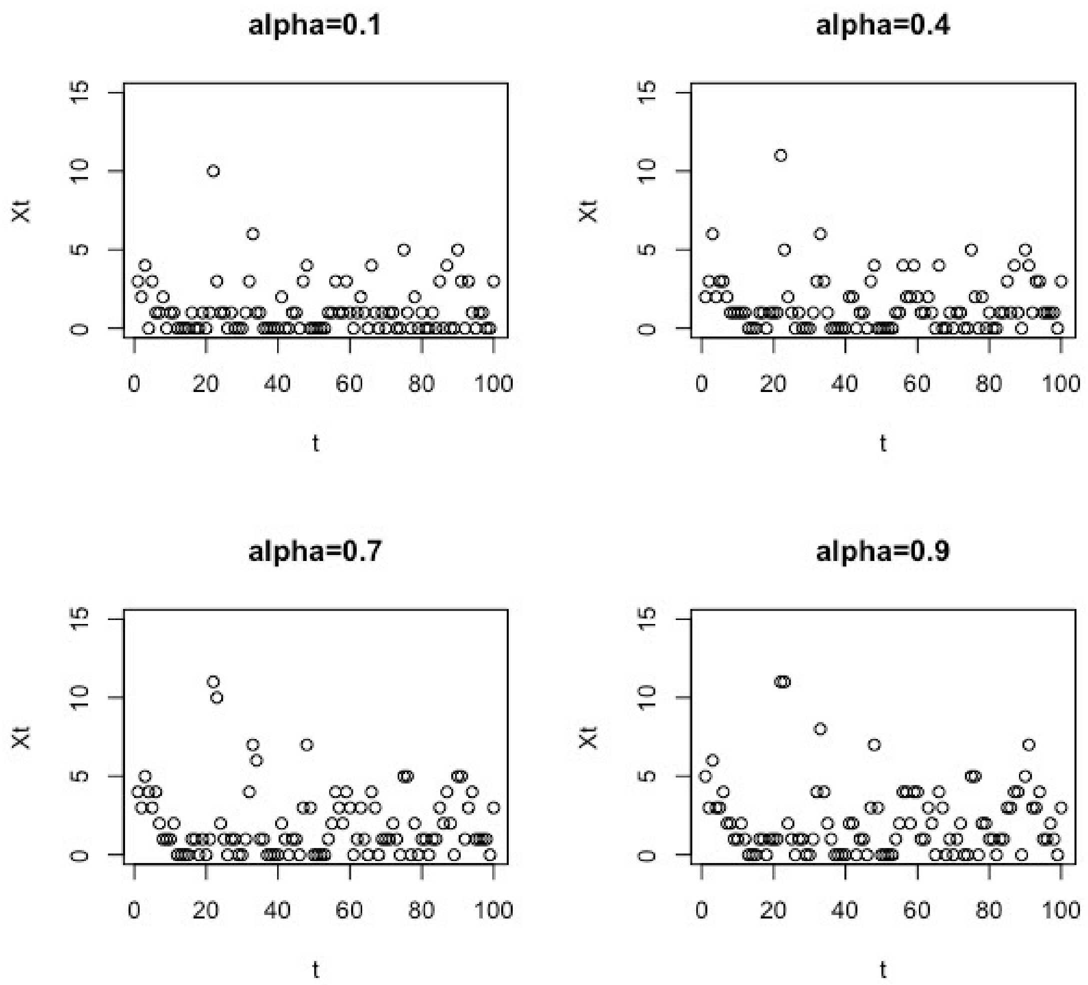

Figure 3.5: Scatter plot of data generated from PIGINMA(1) model with parameters $\lambda = 1$, $\phi = 0.5$, $\mu = 1$ and $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

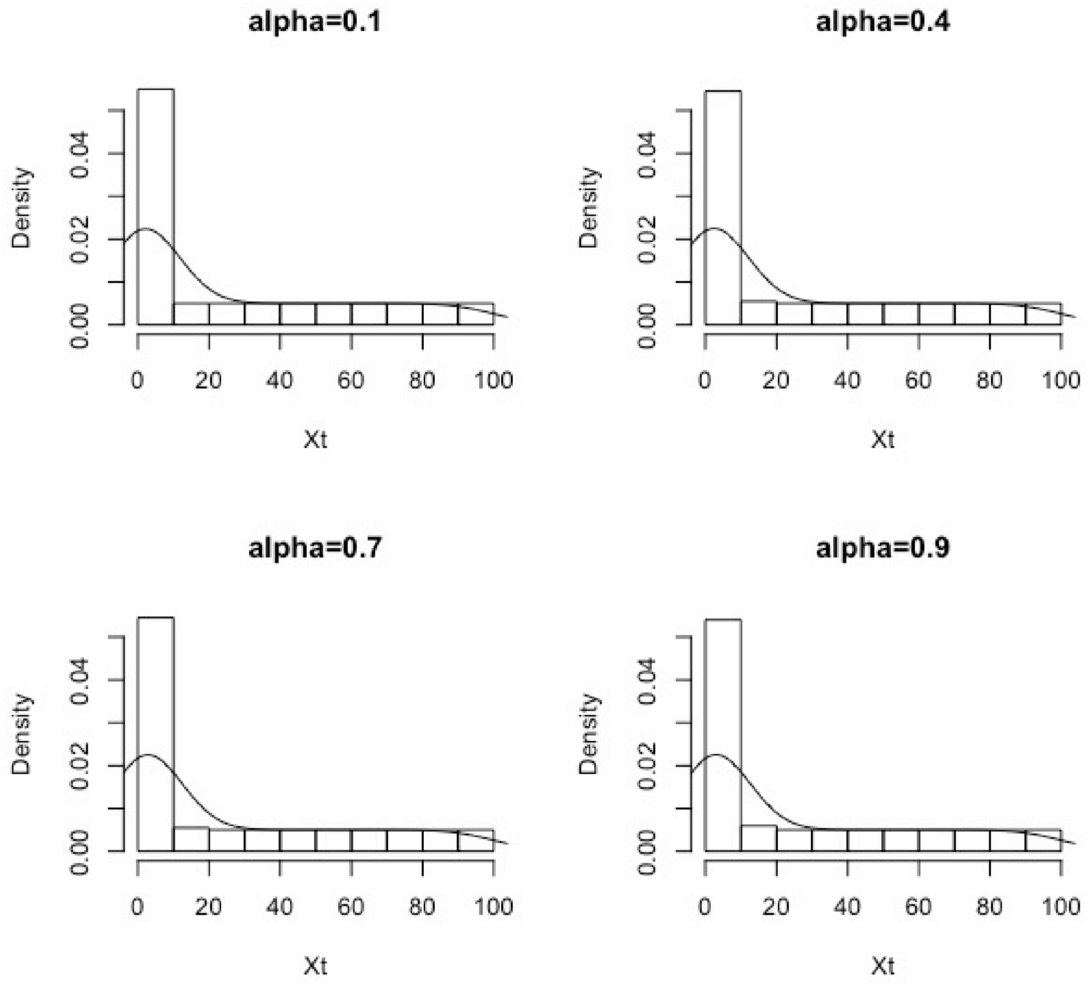

Figure 3.6: Histograms of data generated from PIGINMA(1) model with parameters $\lambda = 1$, $\phi = 0.5$, $\mu = 1$ and $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

By the distribution plots and histograms, we can see that data remains constant when λ increases.

b) $\lambda = 32$, consider $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

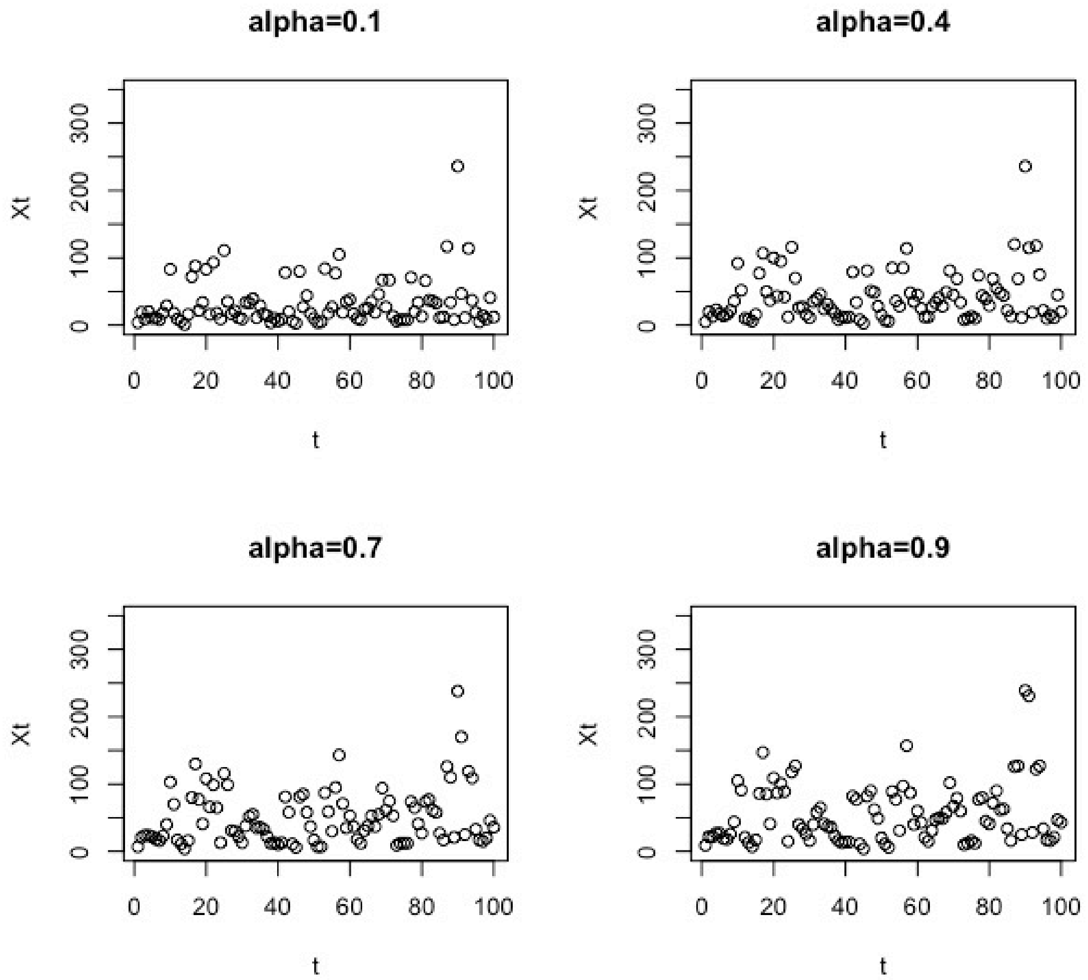

Figure 3.7: Scatter plot of data generated from PIGINMA(1) model with parameters $\lambda = 32$, $\phi = 0.5$, $\mu = 1$ and $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

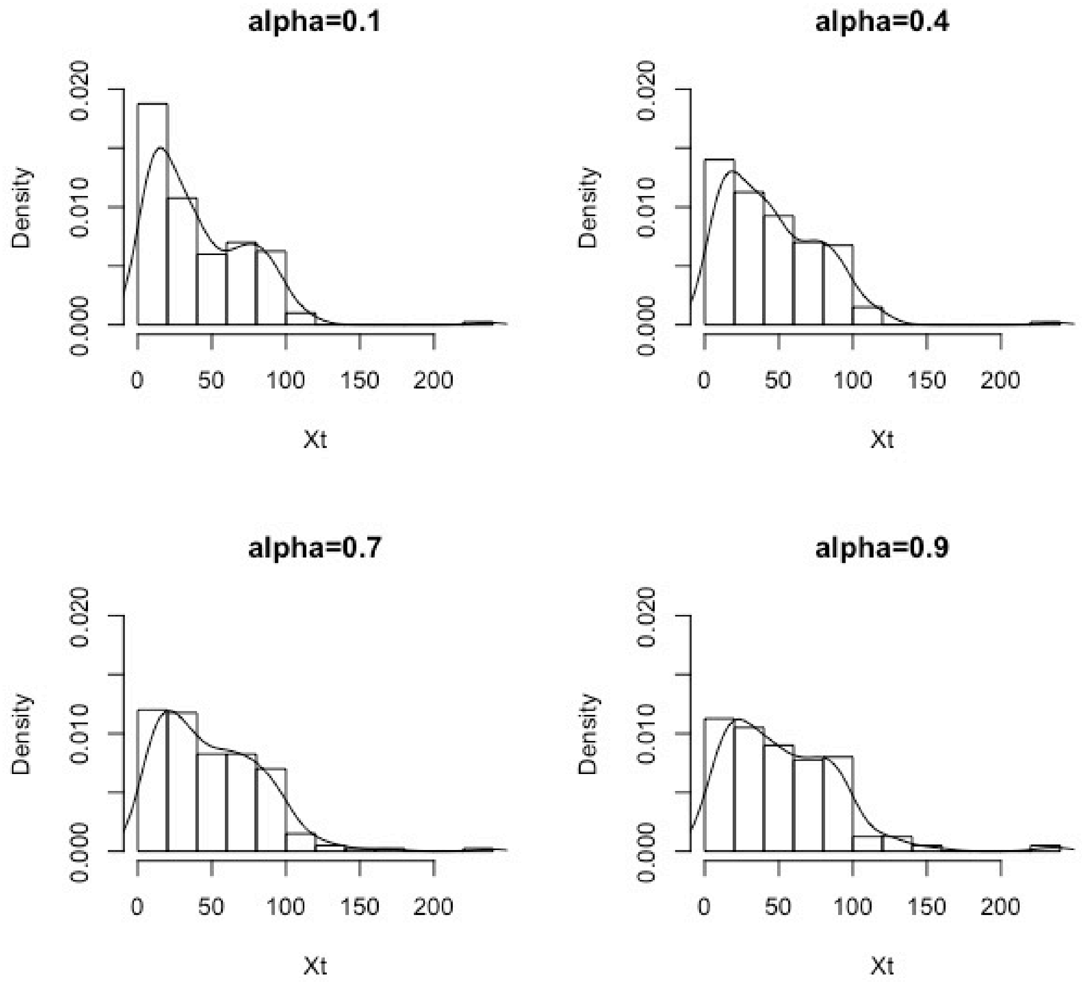

Figure 3.8: Histograms of data generated from PIGINMA(1) model with parameters $\lambda = 32$, $\phi = 0.5$, $\mu = 1$ and $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

By the distribution plots and histograms, we can see that data remains constant when λ increases.

By Figure 3.5 - 3.8, we can see that the data distribution for the case $\lambda = 1$ is more heavily weighted on small values than the case $\lambda = 32$. The data for the case $\lambda = 32$ is more spread than the case $\lambda = 1$.

3. Comparison the generate 100 series of PIGINMA(1) with different λ

We generate 100 series of PIGINMA(1) model parameter $\lambda \in \{1, 32\}, \phi = 0.5, \mu = 1$. The histograms for $X_{10}, X_{20}, \ldots, X_{100}$ are given in Figure 3.9 and Figure 3.10.

a) Consider $\lambda = 1$,

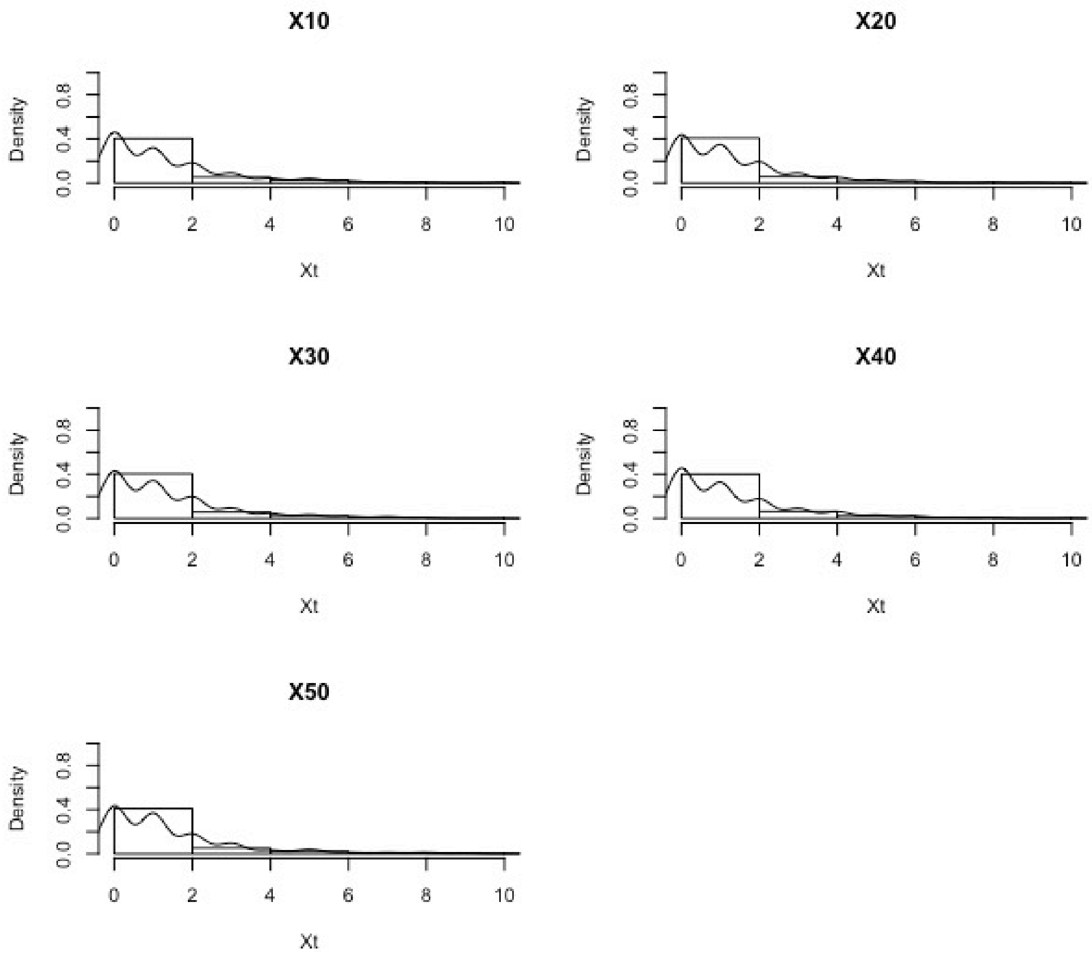

Figure 3.9: Histograms of data $X_{10}, X_{20}, X_{30}, X_{40}, X_{50}$ of PIGINMA(1) model with parameters $\lambda = 1$, $\phi = 0.5$, $\mu = 1$ and $\alpha = 0.5$

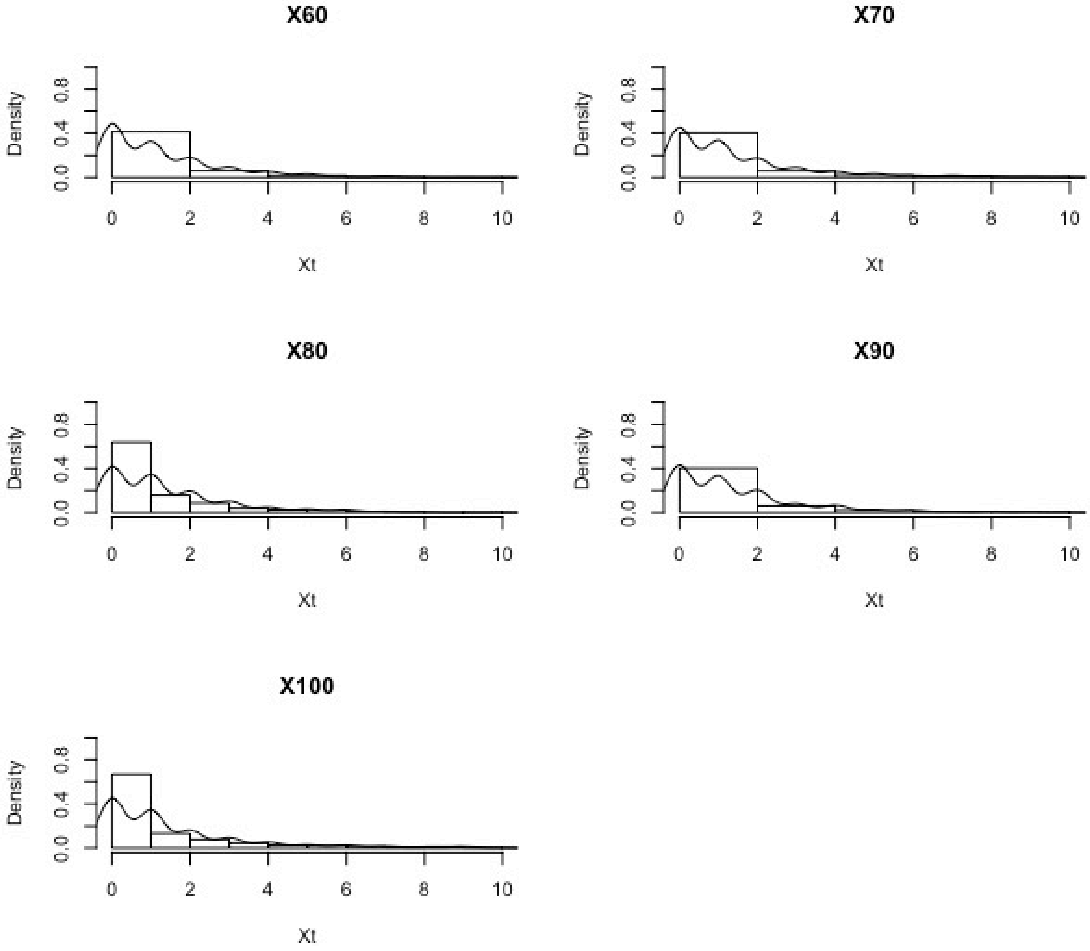

Figure 3.10: Histograms of data $X_{60}, X_{70}, X_{80}, X_{90}, X_{100}$ of PIGINMA(1) model with parameters $\lambda = 1$, $\phi = 0.5$, $\mu = 1$ and $\alpha = 0.5$

By the distribution plots and histograms, we can see that data remains constant when λ increases.

b) Consider $\lambda = 32$,

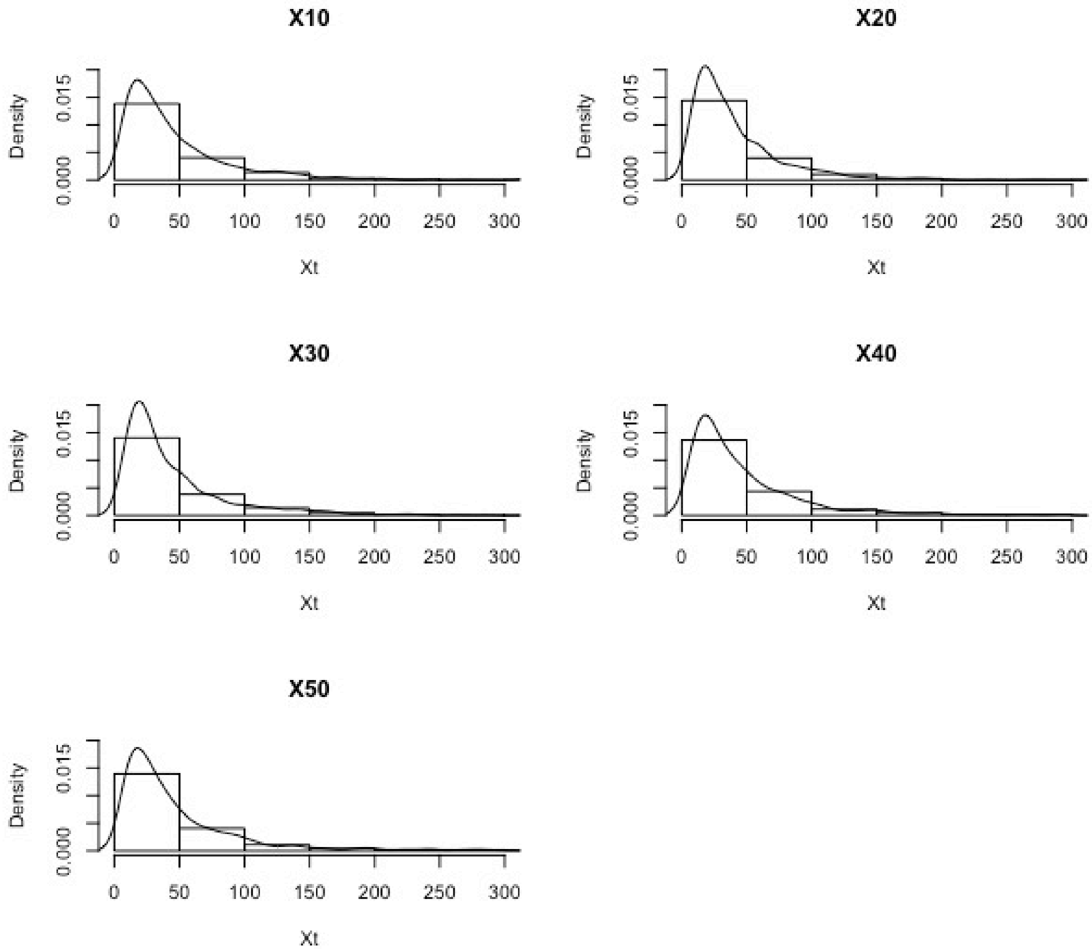

Figure 3.11: Histograms of data $X_{10}, X_{20}, X_{30}, X_{40}, X_{50}$ of PIGINMA(1) model with parameters $\lambda = 32$, $\phi = 0.5$, $\mu = 1$ and $\alpha = 0.5$

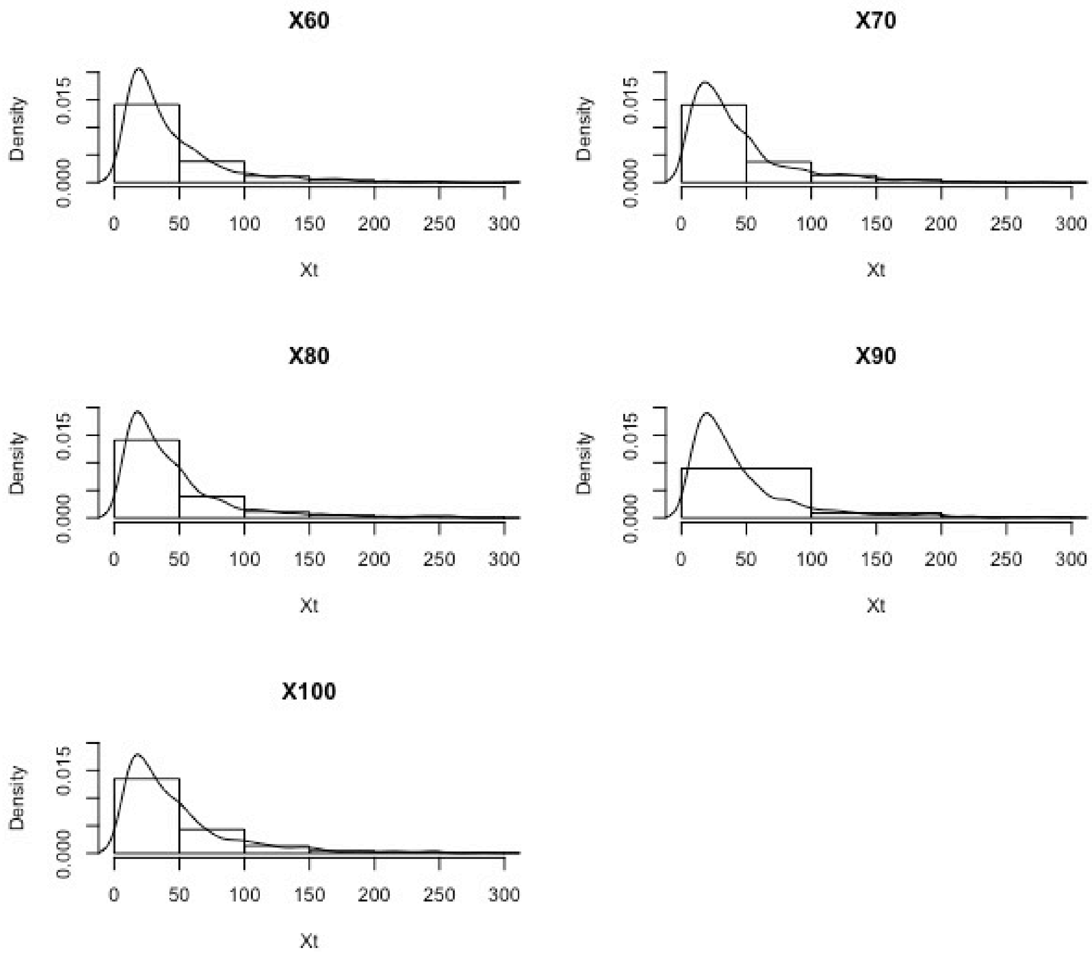

Figure 3.12: Histograms of data $X_{60}, X_{70}, X_{80}, X_{90}, X_{100}$ of PIGINMA(1) model with parameters $\lambda = 32$, $\phi = 0.5$, $\mu = 1$ and $\alpha = 0.5$

By the distribution plots and histograms, we can see that data remains constant when λ increases.

By Figure 3.9 - 3.12, we can see that the data distribution for the case $\lambda = 1$ is more heavily weighted on small values than the case $\lambda = 32$. The data for the case $\lambda = 32$ is more spread than the case $\lambda = 1$.

4. Comparison the generate 100 series of PIGINMA(1) model with different λ and consider X_{10}

We generate 100 series of PIGINMA(1). Fix $\phi = 0.5$, $n = 1, 2, \ldots, 100$ and $\mu = 1$, consider X_{10} when $\lambda \in \{1, 4, 32\}$. The histograms for X_{10} are given in Figure 3.13.

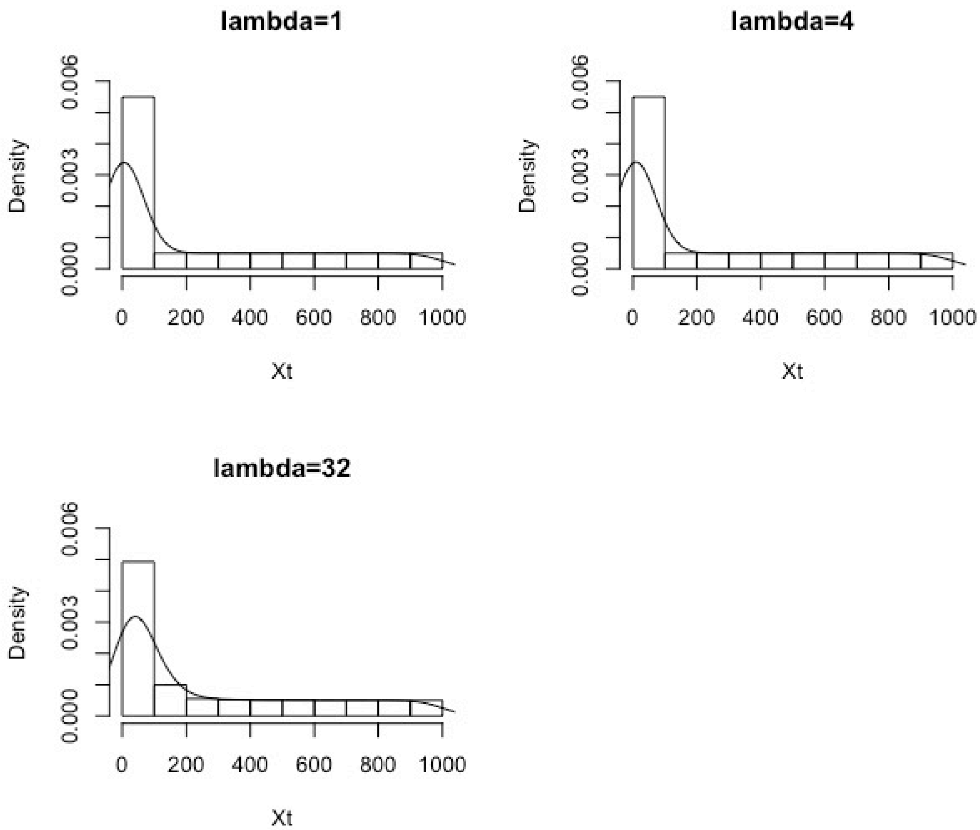

Figure 3.13: Histograms of data X_1 when generated 100 series of PIGINMA(1) model with parameters $\lambda \in \{1, 4, 32\}$, $\phi = 0.5$, $\mu = 1$ and $\alpha = 0.5$

By the distribution plots and histograms, we can see that data skewed to the left and the frequency of the data with small value decreases when λ increases.

3.3.3 Numerical Simulation of PIGINMA(5) model

In this section, we study numerical simulation of the PIGINMA(5) model:

$$X_n = \alpha_1 \circ \epsilon_{n-1} + \alpha_2 \circ \epsilon_{n-2} + \cdots + \alpha_5 \circ \epsilon_{n-5} + \epsilon_n.$$

In our study, we consider 3 different settings,

1. Fix $\phi = 0.5, \alpha = 0.5, n = 1, 2, \ldots, 100$ and $\mu = 1$, consider $\lambda \in \{1, 1.4, 1.8, 2.4, 8, 16, 32\}$.

2. Fix $\phi = 0.5$, $n = 1, 2, \ldots, 100$ and $\mu = 1$

   a) $\lambda = 1$, consider $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$.

   b) $\lambda = 32$, consider $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$.

3. We generate 100 series of PIGINMA(1). Fix $\phi = 0.5$, $n = 1, 2, \ldots, 100$ and $\mu = 1$.

   a) Consider $\lambda = 1$

   b) Consider $\lambda = 32$

4. We generate 100 series of PIGINMA(1). Fix $\phi = 0.5$, $n = 1, 2, \ldots, 100$ and $\mu = 1$, consider X_{10} when $\lambda \in \{1, 4, 32\}$.

1. Comparison the Data of PIGINMA(5) model with different λ

For PIGINMA(5) model we set the parameters $\phi = 0.5, \alpha = 0.5$, $n = 100$ and $t = 1$.

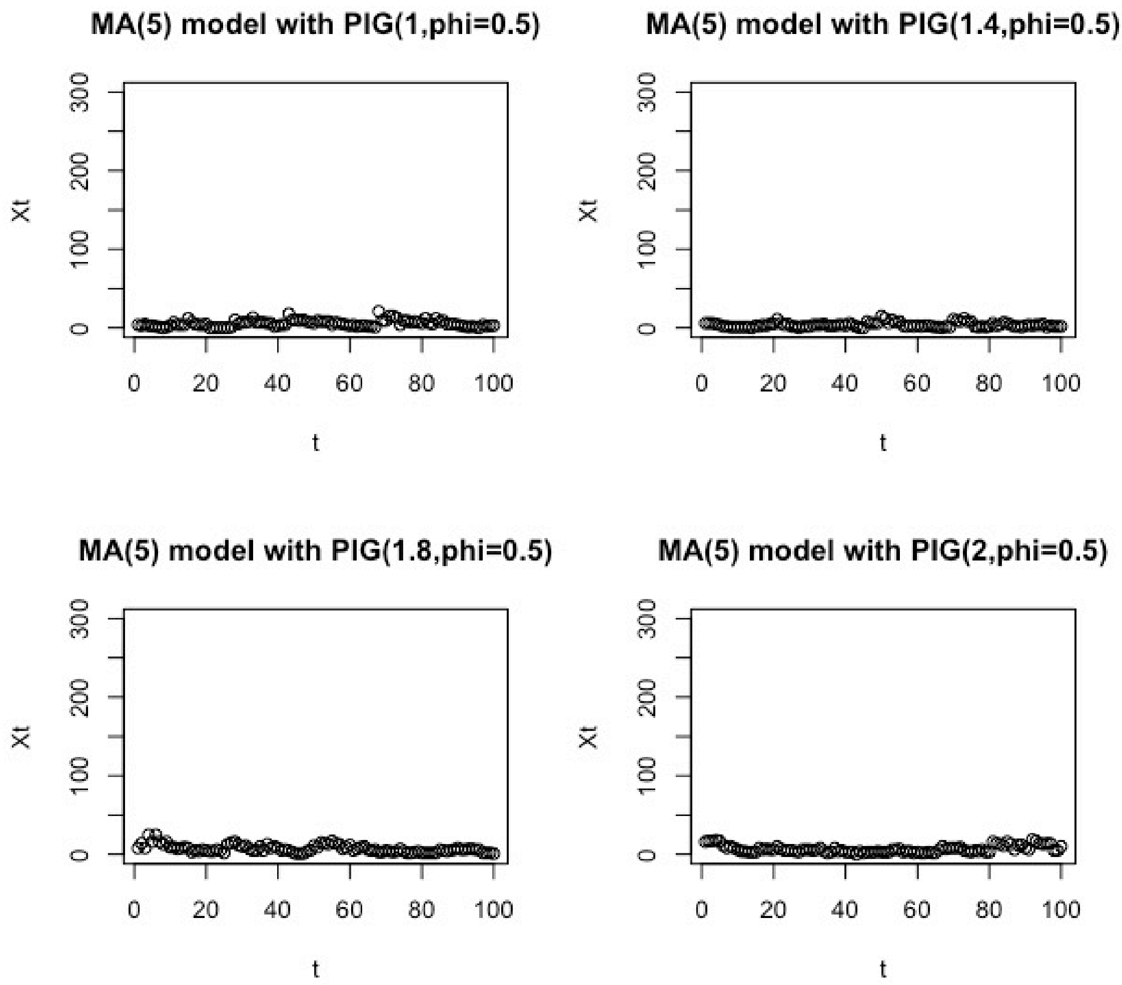

Figure 3.14: Scatter plot of data generated from PIGINMA(5) model with parameters $\phi = 0.5$, $\mu = 1$ and $\lambda \in \{1, 1.4, 1.8, 2\}$

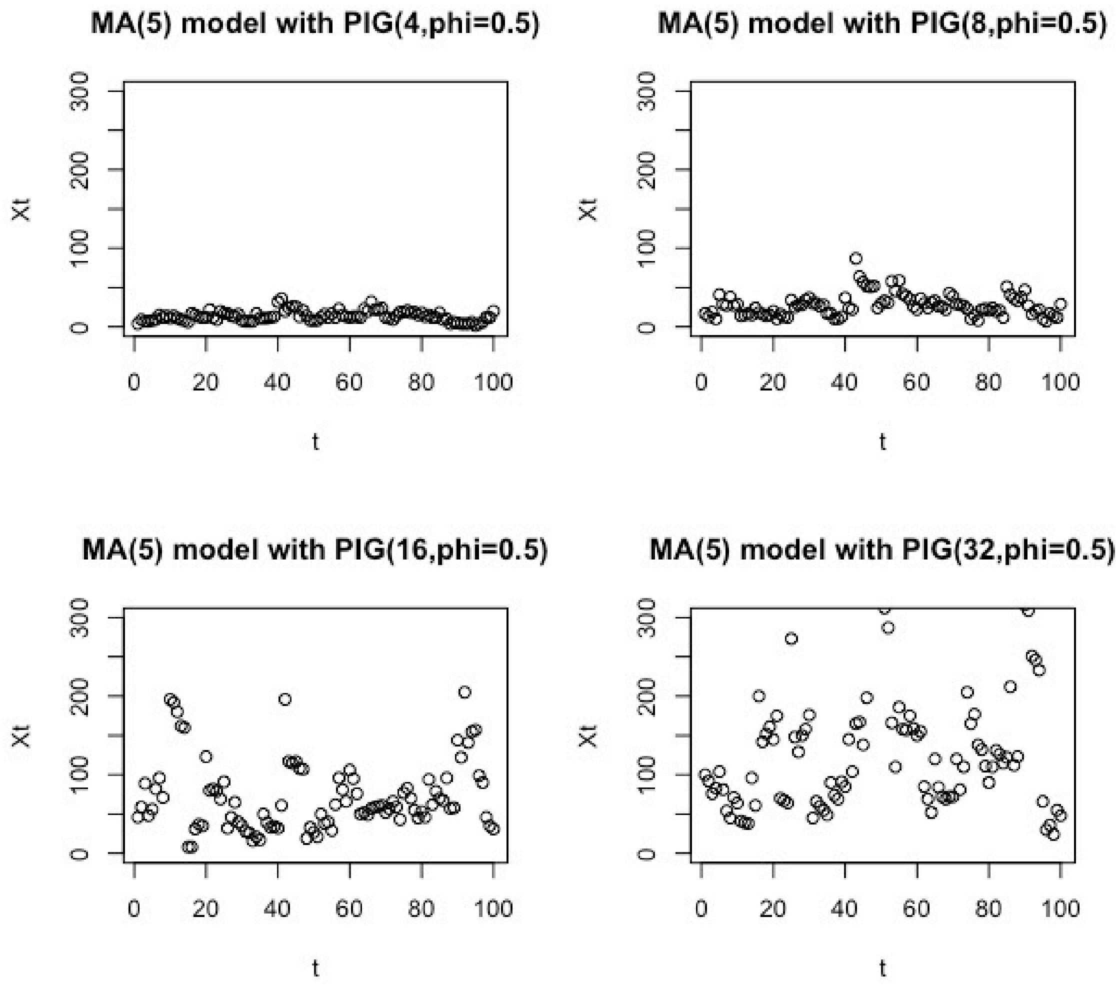

Figure 3.15: Scatter plot of data generated from PIGINMA(5) model with parameters $\phi = 0.5$, $\mu = 1$ and $\lambda \in \{4, 8, 16, 32\}$

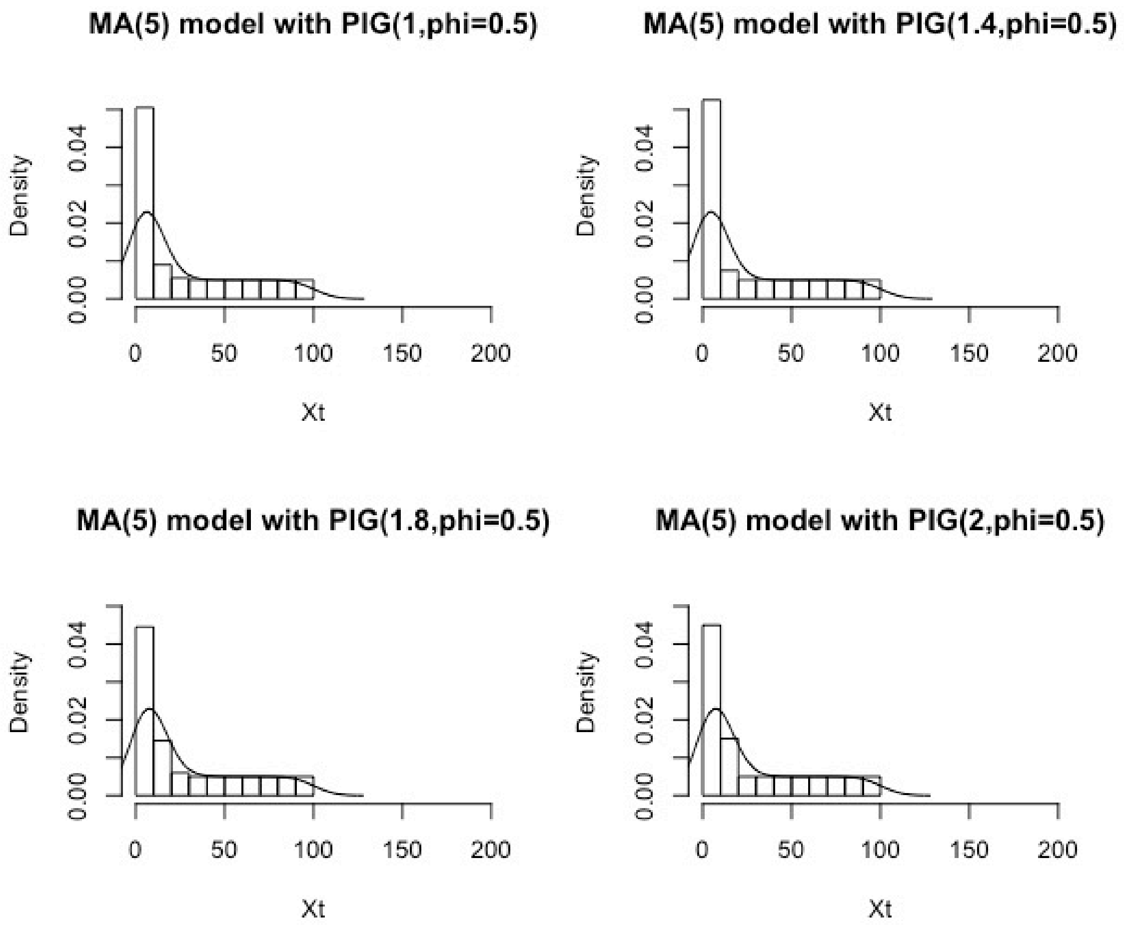

Figure 3.16: Histograms of data generated from PIGINMA(5) model with parameters $\phi = 0.5$, $\mu = 1$ and $\lambda \in \{1, 1.4, 1.8, 2\}$

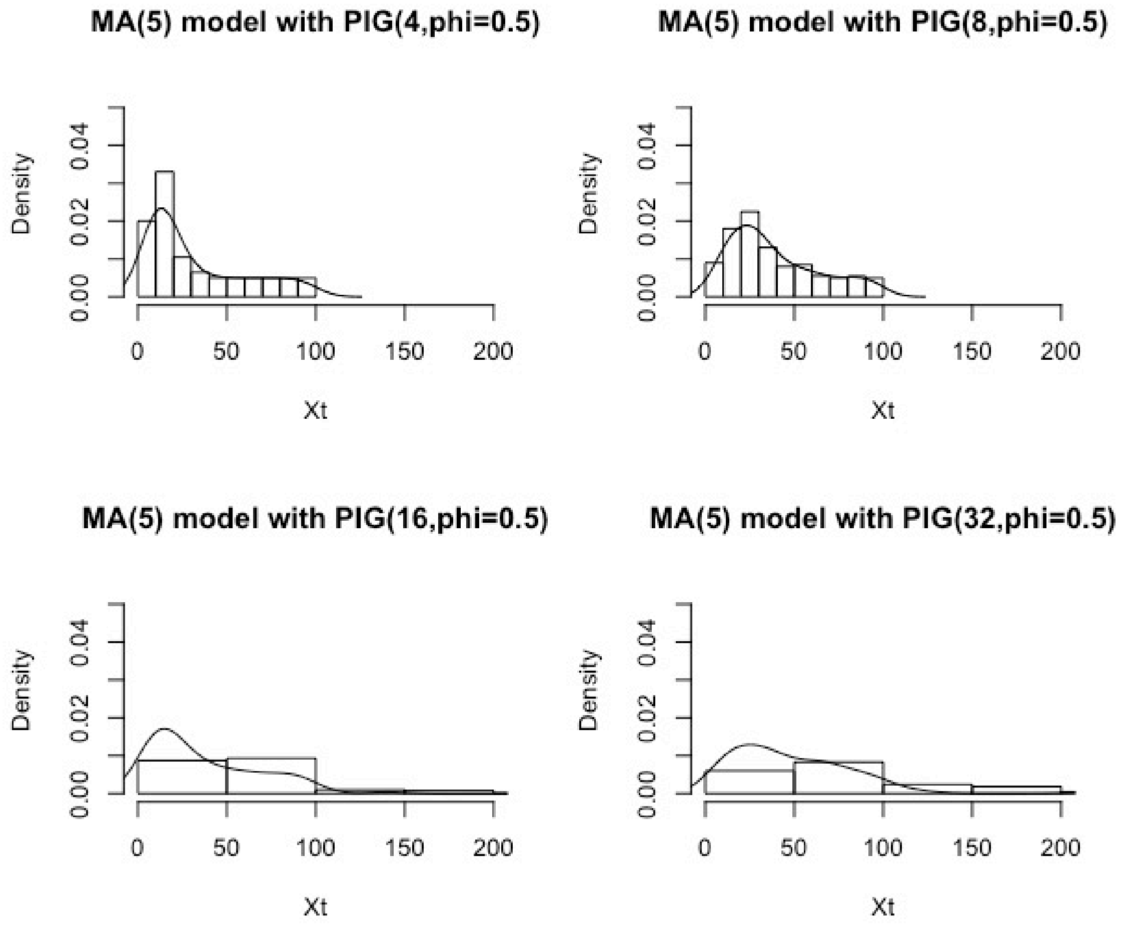

Figure 3.17: Histograms of data generated from PIGINMA(5) model with parameters $\phi = 0.5$, $\mu = 1$ and $\lambda \in \{4, 8, 16, 32\}$

By the distribution plots and histograms, we can see that data skewed to the left and the frequency of the data with small value decreases when λ increases.

2. Comparision the Data of PIGINMA(5) model with different λ and α

For PIGINMA(5) model we set the parameter $\phi = 0.5, \mu = 1$ and $n = 1, 2, \ldots, 100$.

a) $\lambda = 1$, consider $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

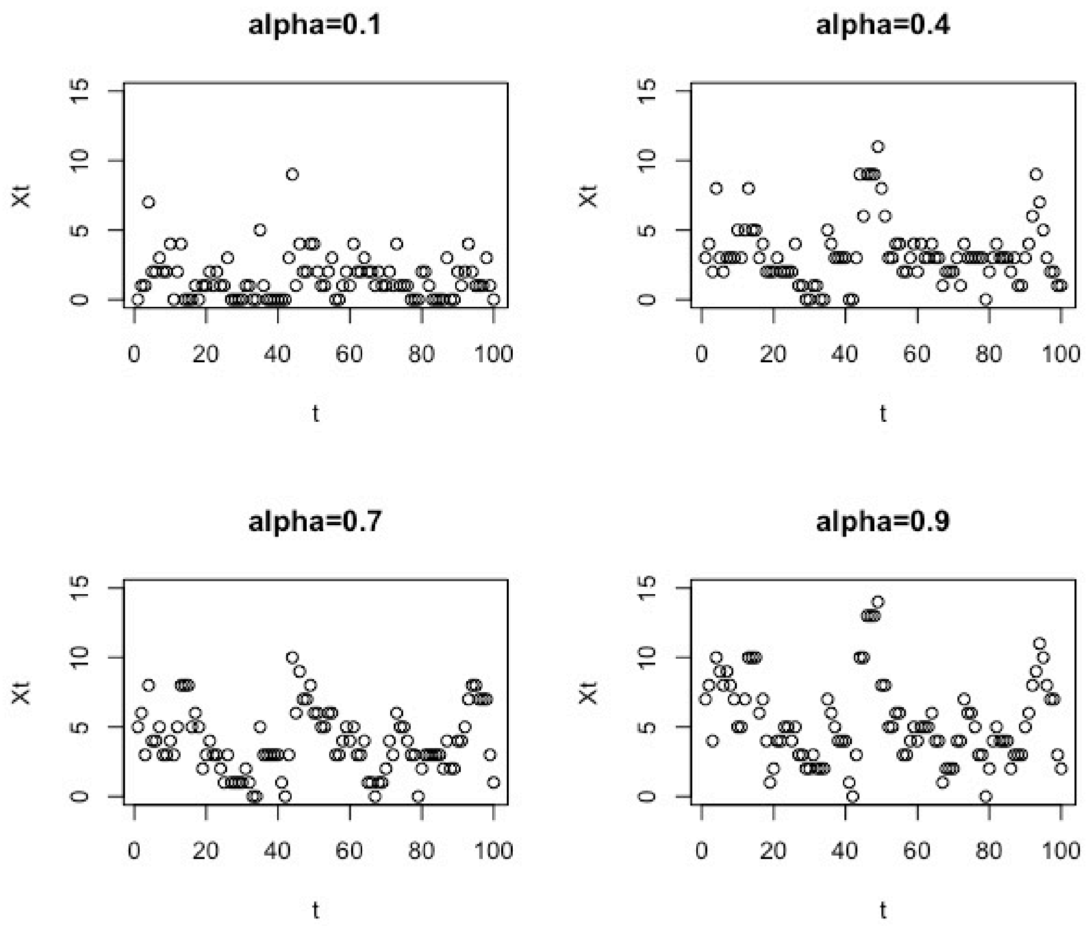

Figure 3.18: Scatter plot of data generated from PIGINMA(5) model with parameters $\lambda = 1$, $\phi = 0.5$, $\mu = 1$ and $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

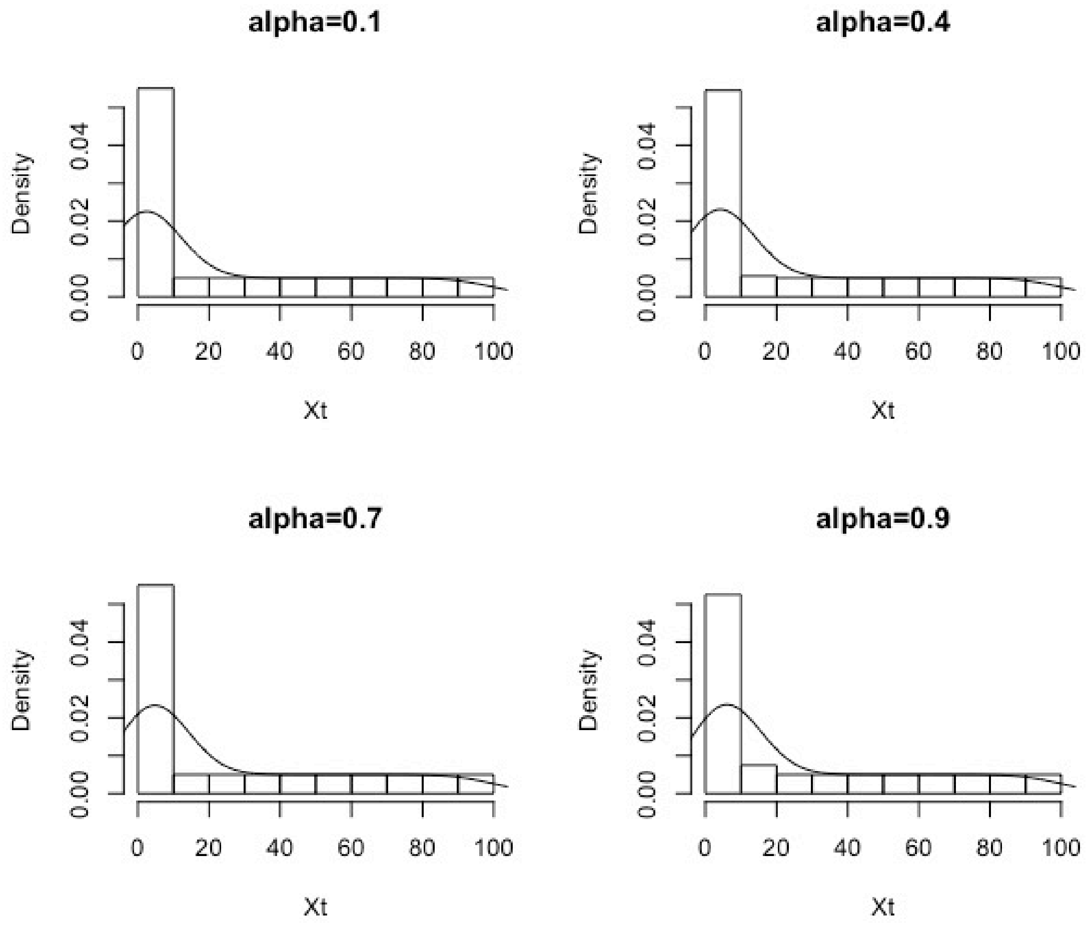

Figure 3.19: Histograms of data generated from PIGINMA(5) model with parameters $\lambda = 1$, $\phi = 0.5$, $\mu = 1$ and $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

By the distribution plots and histograms, we can see that data skewed to the left.

b) $\lambda = 32$, consider $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

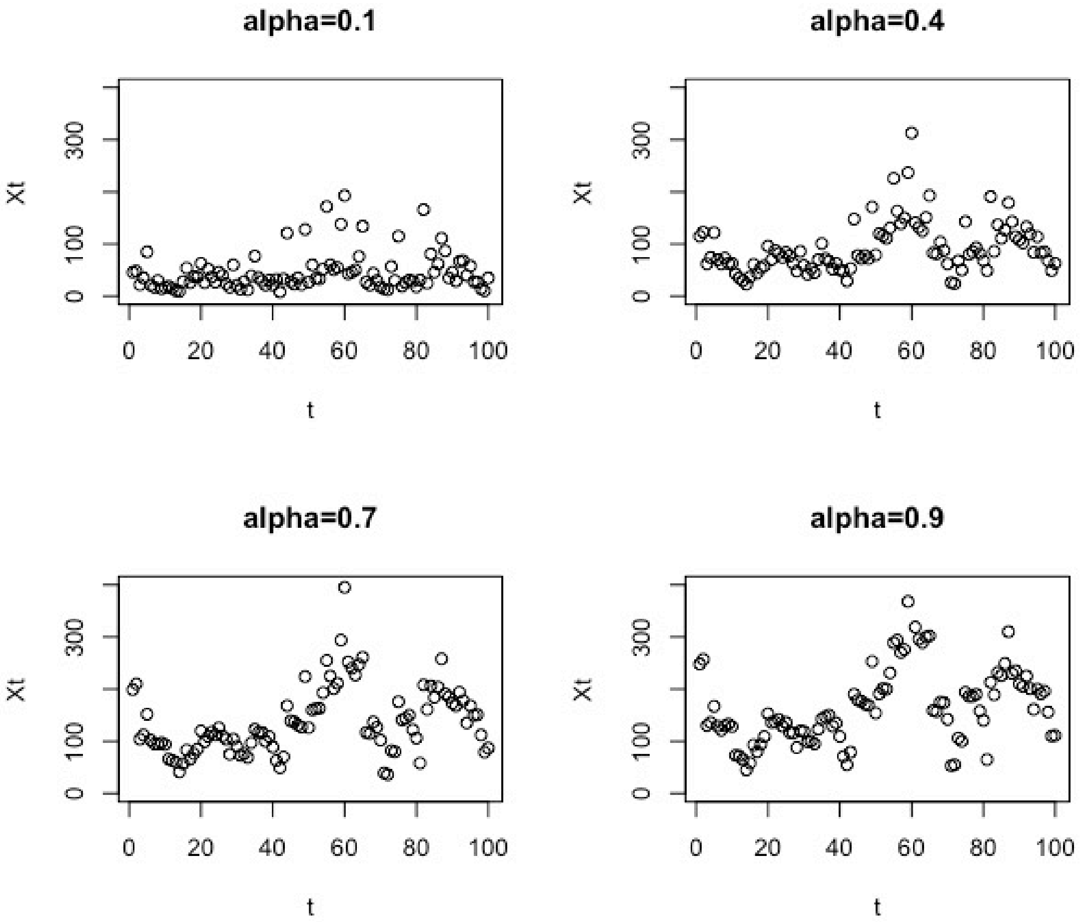

Figure 3.20: Scatter plot of data generated from PIGINMA(5) model with parameters $\lambda = 32$, $\phi = 0.5$, $\mu = 1$ and $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

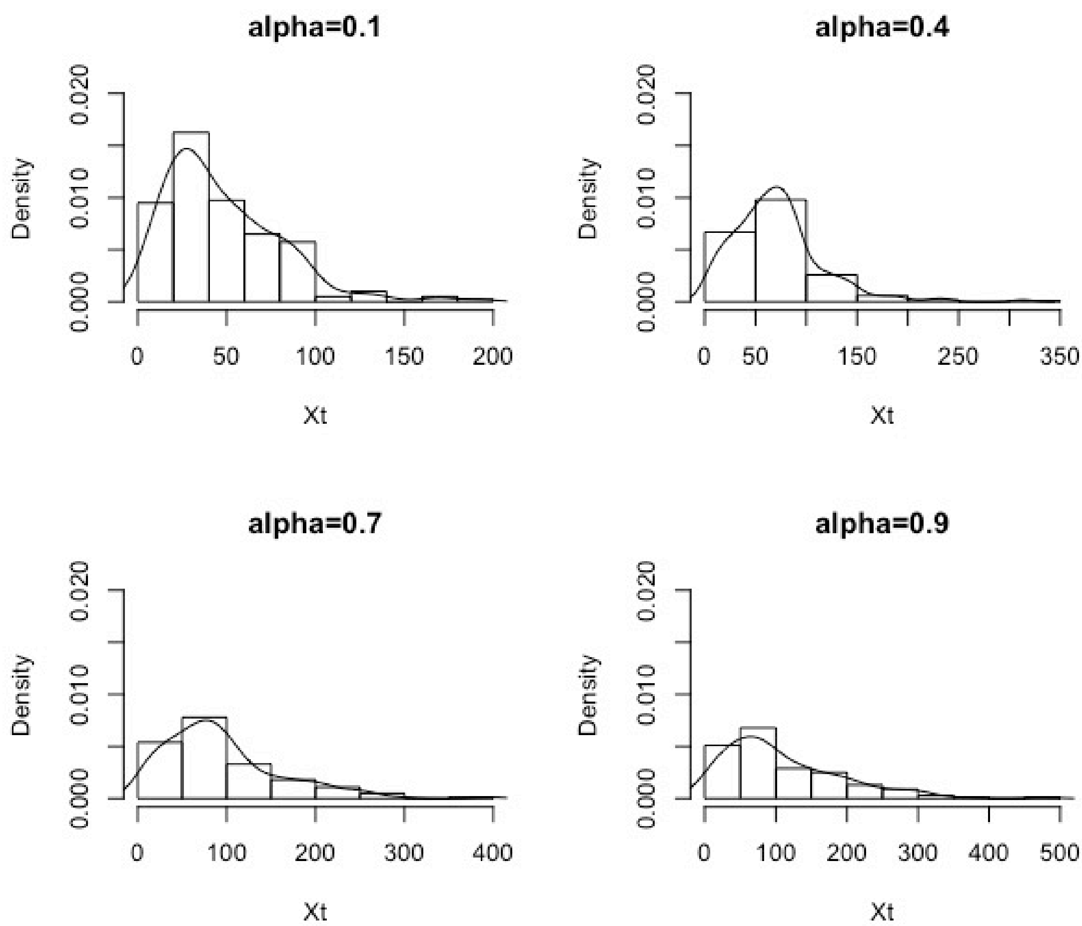

Figure 3.21: Histograms of data generated from PIGINMA(5) model with parameters $\lambda = 32$, $\phi = 0.5$, $\mu = 1$ and $\alpha \in \{0.1, 0.4, 0.7, 0.9\}$

By the distribution plots and histograms, we can see that data remains constant when λ increases.

By Figure 3.19 and Figure 3.21, we can see that the data distribution for the case $\lambda = 1$ is more heavily weighted on small values than the case $\lambda = 32$. The data for the case $\lambda = 32$ is more spread than the case $\lambda = 1$.

3. Comparison the generate 100 series of PIGINMA(5) with different λ

We generate 100 series of PIGINMA(5) model parameter $\lambda \in \{1, 32\}, \phi = 0.5, \mu = 1.$ The histograms for $X_{10}, X_{20}, \ldots, X_{100}$ are given in Figure 3.22 and Figure 3.23.

a) Consider $\lambda = 1,$

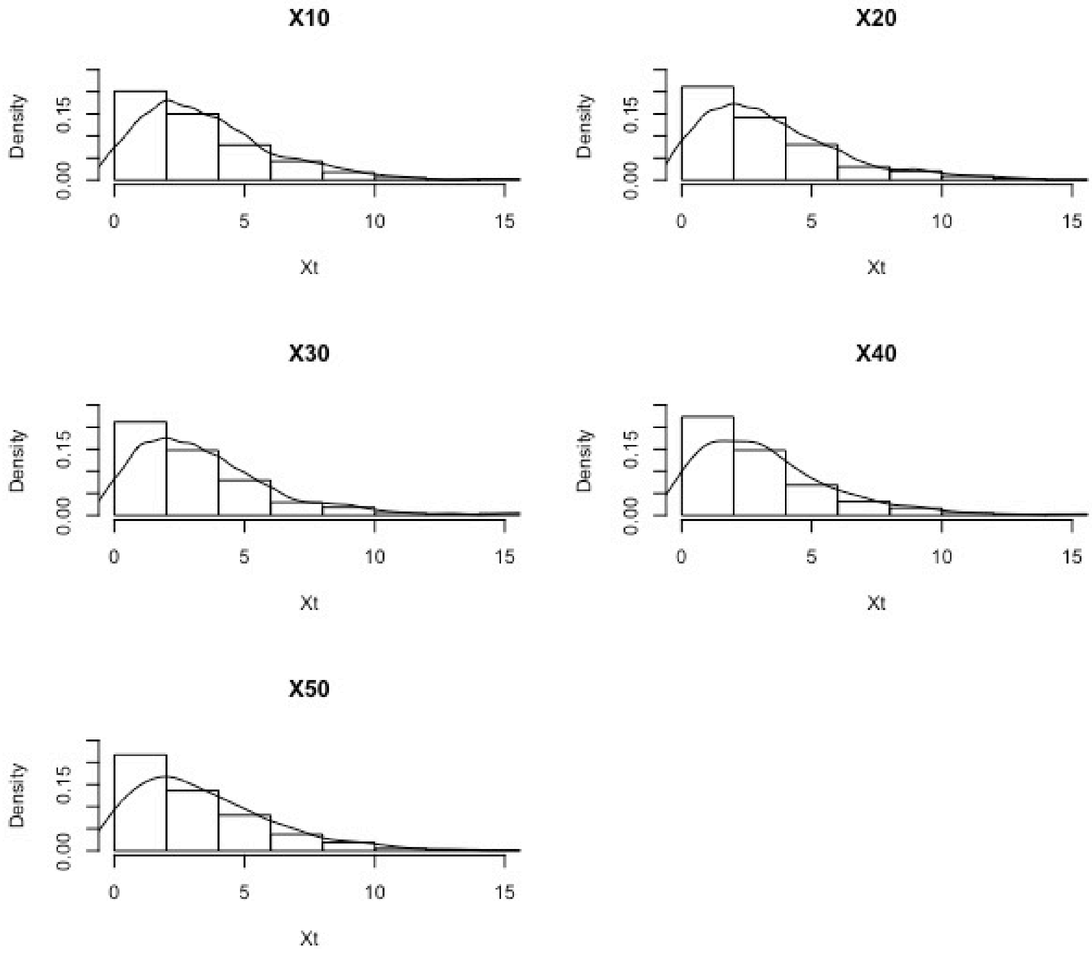

Figure 3.22: Histograms of data $X_{10}, X_{20}, X_{30}, X_{40}, X_{50}$ of PIGINMA(5) model with parameters $\lambda = 1$, $\phi = 0.5$, $\mu = 1$ and $\alpha = 0.5$

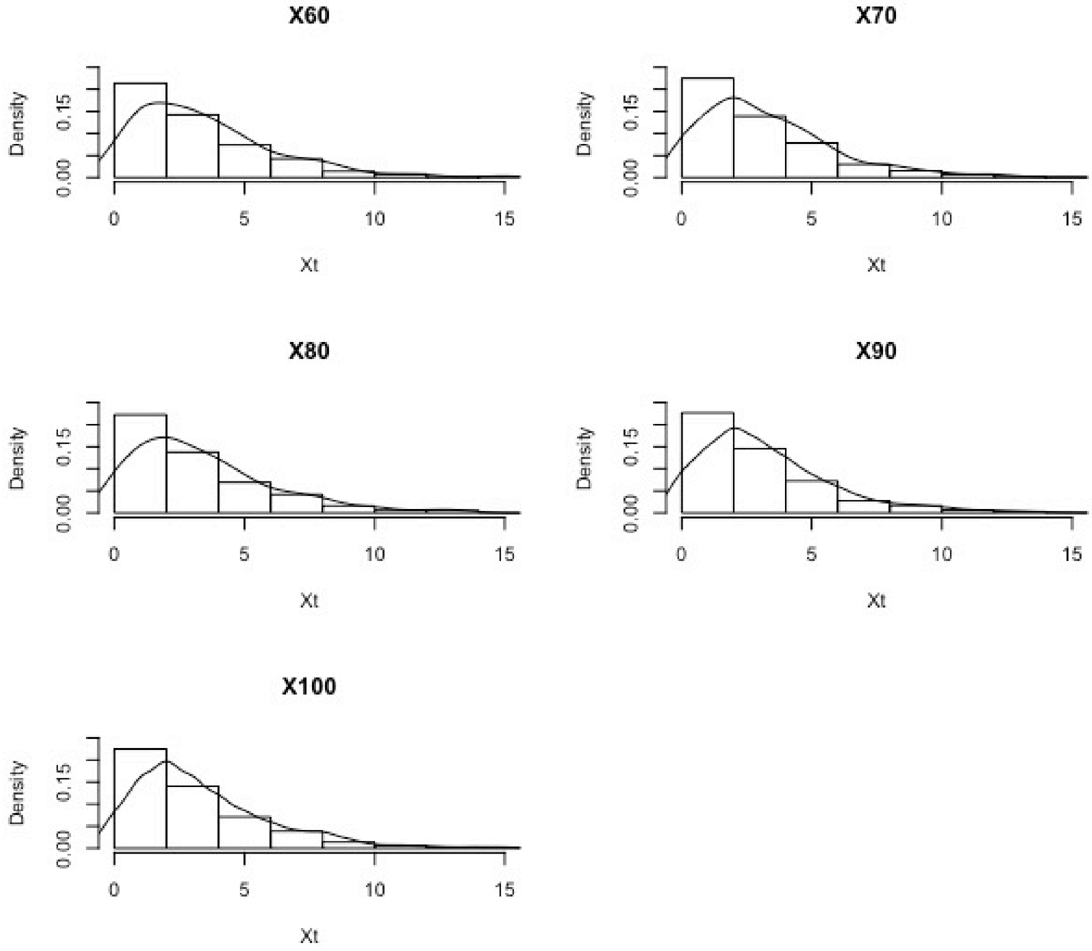

Figure 3.23: Histograms of data $X_{60}, X_{70}, X_{80}, X_{90}, X_{100}$ of PIGINMA(5) model with parameters $\lambda = 1$, $\phi = 0.5$, $\mu = 1$ and $\alpha = 0.5$

By the distribution plots and histograms, we can see that data remains constant when λ increases.

b) Consider $\lambda = 32$,

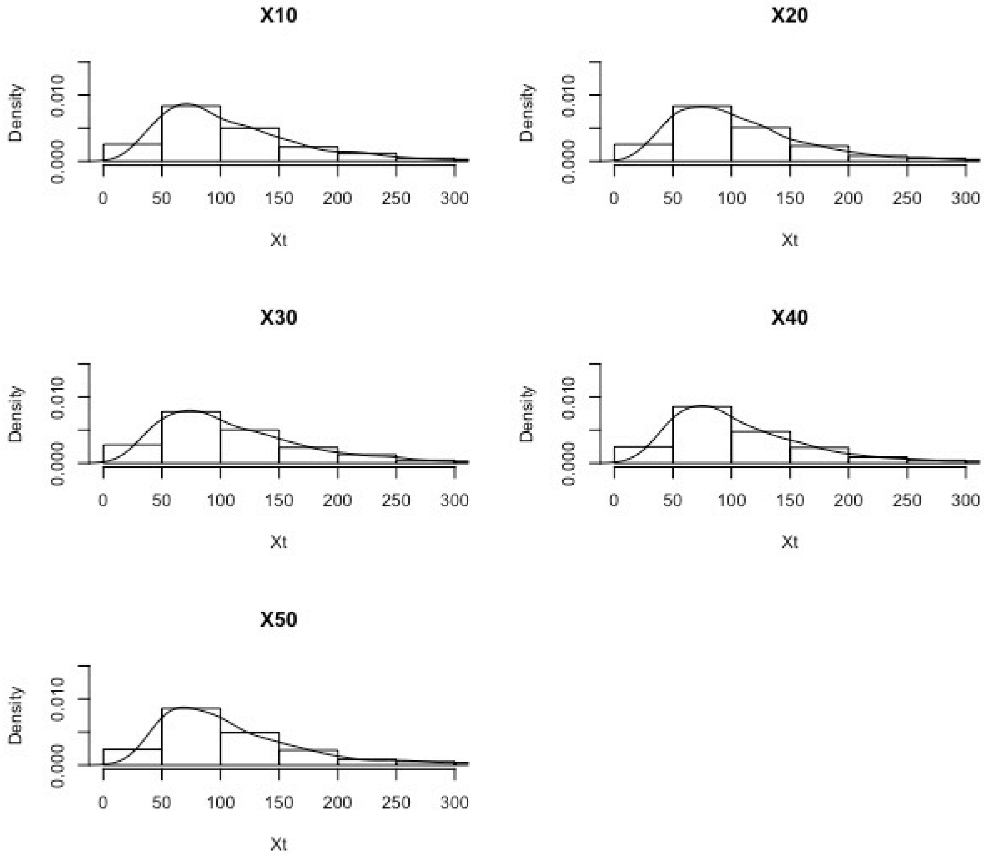

Figure 3.24: Histograms of data $X_{10}, X_{20}, X_{30}, X_{40}, X_{50}$ of PIGINMA(5) model with parameters $\lambda = 32$, $\phi = 0.5$, $\mu = 1$ and $\alpha = 0.5$

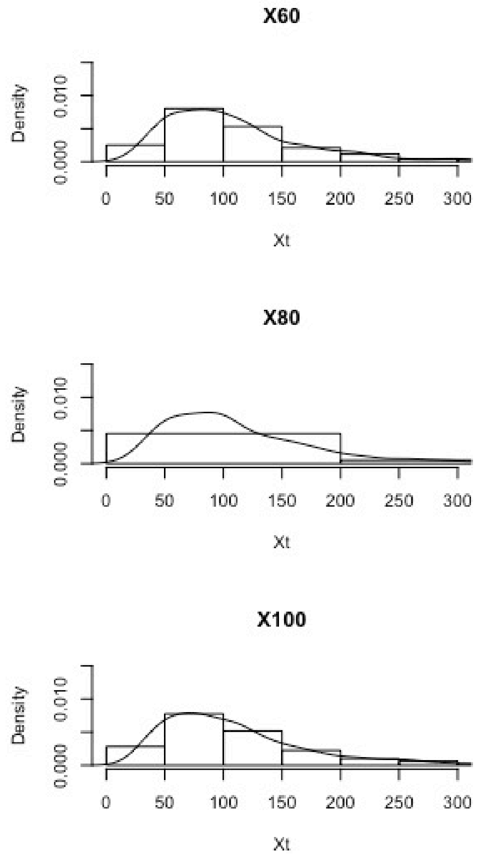

Figure 3.25: Histograms of data $X_{60}, X_{70}, X_{80}, X_{90}, X_{100}$ of PIGINMA(5) model with parameters $\lambda = 32$, $\phi = 0.5$, $\mu = 1$ and $\alpha = 0.5$

By the distribution plots and histograms, we can see that data remains constant when λ increases.

By Figure 3.22 - 3.25, we can see that the data distribution for the case $\lambda = 1$ is more heavily weighted on small values than the case $\lambda = 32$. The data for the case $\lambda = 32$ is more spread than the case $\lambda = 1$.

4. Comparison the generate 100 series of PIGINMA(5) model with different λ and consider X_{10}

We generate 100 series of PIGINMA(5). Fix $\phi = 0.5$, $n = 1, 2, \ldots, 100$ and $\mu = 1$, consider X_{10} when $\lambda \in \{1, 4, 32\}$. The histograms for X_{10} are given in Figure 3.26.

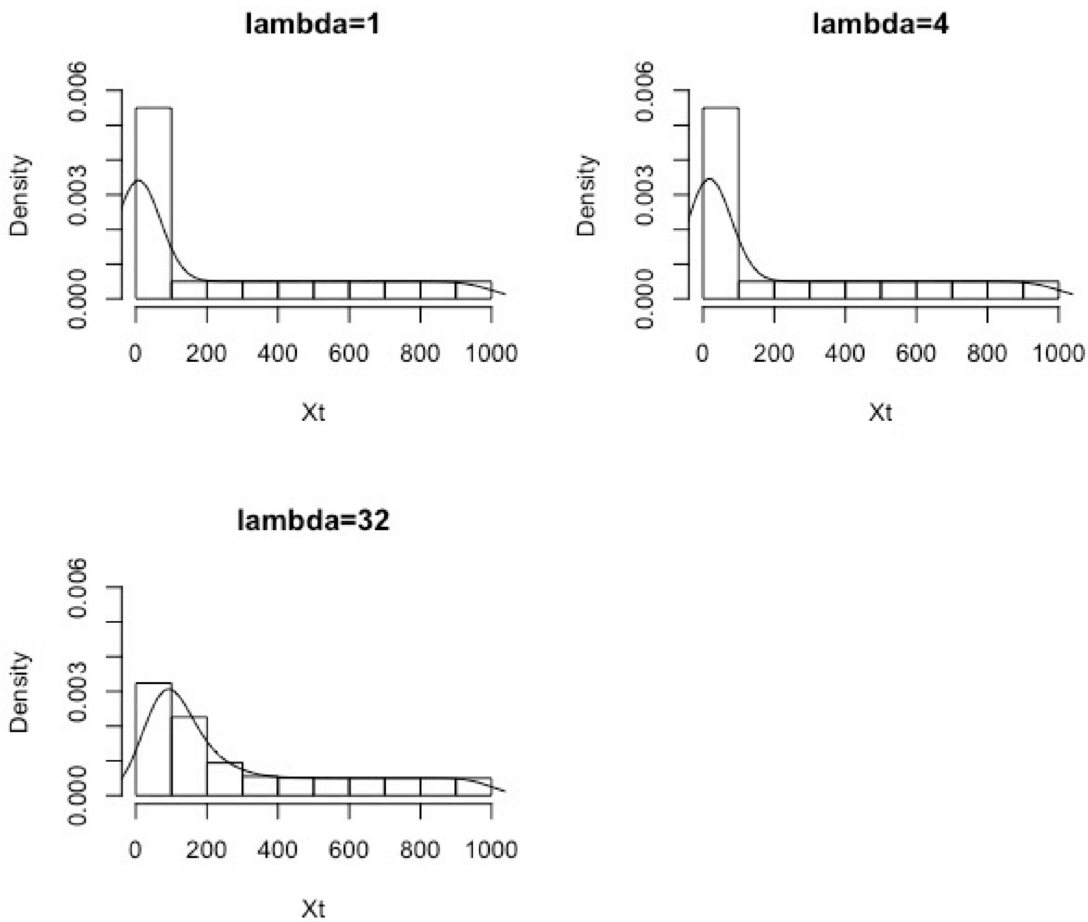

Figure 3.26: Histograms of data X_1 when generated 100 series of PIGINMA(5) model with parameters $\lambda \in \{1, 4, 32\}$, $\phi = 0.5$, $\mu = 1$ and $\alpha = 0.5$

By the distribution plots and histograms, we can see that data skewed to the left and the frequency of the data with small value decreases when λ increases.

Chapter 4
Conclusion

In this project, we extend the study of the Mixed Poisson INAR(1) model to construct a Mixed Poisson INMA(1) model and derive their probabilistic properties such as mean, variance and covariance. Moreover, we presented distribution plots of such data in many different settings. From our simulation study, we found that the distribution of data is skewed to the left as the value of λ increases. However, the distribution is hardly affected by the value of α.

References

1. Alosh, M. and Alzaid, A. Integer-valued moving average (INMA) process. **Statistical Papers Statistische Heffe.** 29(1), 281-300, (1988).

2. Alosh, M. and Alzaid, A. First-order integer-valued autoregressive (INAR(1)) process. **Journal Of Time Series Analysis.** 8(3), 261-275, (1987).

3. JinGuan, D. and Yuan, L. The integer-valued autoregressive (INAR(p)) model. **Journal Of Time Series Analysis.** 12(2), 129-142, (1989).

4. Karlis, D. and Xekalaki, E. Mixed Poisson Distributions, **Revue Internationale De Statistique.** 73(1), 35-58, (2005).

5. Khodabin, M. and Ahmadabadi, A. Some properties of generalized gamma distribution. **Mathematical Sciences.** 4(1), 9-28, (2010).

6. Schweer, S. and Weiss, C. Compound Poisson INAR(1) processes: Stochastic properties and testing for overdispersion. **Computational Statistics And Data Analysis.** 77(C), 267-284, (2014).

Appendix

Coding for PIGINMA(1)

```
n=1000#number of simulation
t=100#number of Xt
q=1#MA(1)
mu=32
phi=0.5*mu
alpha=0.5

fEps<-function(t,q,n,mu,phi){
    Eps=rep(1,((t+q)*n))
    for(i in 1:((t+q)*n)){Eps[i]=rpoisinvgauss(1,mu,dis=(1/phi))}
return(Eps)}
A=fEps(t,q,n,mu,phi)
MEps=matrix(A,nrow = n)

Thin=matrix(MEps,n,(t+q))
for (i in 1:n){
    for(j in 1:(t+q-1)){Thin[i,j]=rbinom(1,MEps[i,j],alpha)}
}

Xmat=matrix(1,n,t)
for (j in 1:n){
    for(i in 1:t){Xmat[j,i]=sum(Thin[j,(i:(i+q-1))])+MEps[j,(i+q)]}
}
```

The Project Proposal of Course 2301399 Project Proposal Academic Year 2018

Project Tittle (Thai)	การศึกษาของตัวแบบอนุกรมเวลาที่มีค่าเป็นจำนวนเต็มแบบปัวซงผสม
Project Tittle (English)	A study on Mixed Poisson integer-valued times series models
Advisor	Assistant Prof.Jiraphan Suntornchost, Ph.D.
By	Panisara Yamsook ID 5833532123
	Mathematics, Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University

Background and Rationale

Count time series arise naturally in many practical situations, for example, the insurance claim counts and the number of stock transactions. Therefore, increases in interest in the modelling have been observed. The most common distribution considered in count time series data is the Poisson distribution. The model is refered as the Poisson INAR(1) model which is a stationary integer-valued time series with lag-one dependence. The Poisson INAR model has been applied in many applications since it was introduced by McKenzie in 1985.

However, the property of the Poisson models having equal mean and variance is rarely found in applications. Many real-world data examples exhibit overdispersion, i.e., the variance is larger than the mean. Therefore, the integer-valued autoregressive (INAR) process with Poisson marginals is not adequate for modelling overdispersed counts. Consequently, several alternative distributions have been proposed for the integer-valued time series models, for example, geometric distribution and negative binomial distribution. Recently, in 2017, Batteto-Souza extended the Poisson INAR(1) to the mixed Poisson INAR(1) model to accommodate overdispersion data. In their study, they considered the inverse-gaussian Poisson INAR(1) model.

In this project, we will extend the study of the Mixed Poisson INAR(1) model to construct a Mixed Poisson INMA(q) model and derive their probabilistic properties such as mean, variance and covariance. Moreover, we present distribution plots of such data in many different settings.

Objectives

To extend the one order integer-valued time series model based on the mixed Poisson distribution, Mixed Poisson INAR(1),to construct a Mixed Poisson INMA(1) model and derive their probabilistic properties such as mean, variance and covariance. Moreover, we present distribution plots of such data in many different settings.

Scope

In this project, we consider the integer-valued time series models based on the Mixed Poisson distribution.

Project Activities

1. Study fundamental concepts of probability theory and integer-valued time series models.
2. Study properties and constructions of Mixed Poisson distributions.
3. Study Mixed Poisson INAR(1) processes.
4. Construct Mixed Poisson INMA(q) model and study its properties.
5. Summarize and write the report.

Scheduled Operations

Procedures	Months								
	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.
1. Study fundamental concepts of probability theory integer-valued time series models.									
2. Study properties and constructions of Mixed poisson distributions.									
3. Study Mixed Poisson INAR(1) processes.									
4. Construct Mixed Poisson INMA (q) model and study its properties.									
5. Summarize and write the report.									

Benefits

The benefits for student who implement this project.

1. To learn properties of integer-valued time series for count data.
2. To gain knowledge in probability theory and apply the models to suitable applications.

The benefits for users of the project.

To have more general Mixed Poisson integer-valued time series processes for wider applications.

Equipment

Software

1. Mathematica

2. RStudio
3. Adobe PDF
4. Latex

Hardware

1. Printer
2. Computer

References

1. Alosh, M. and Alzaid, A. Integer-valued moving average (INMA) process. **Statistical Papers Statistische Heffe.** 29(1), 281-300, (1988).
2. Alosh, M. and Alzaid, A. First-order integer-valued autoregressive (INAR(1)) process. **Journal Of Time Series Analysis.** 8(3), 261-275, (1987).
3. JinGuan, D. and Yuan, L. The integer-valued autoregressive (INAR(p)) model. **Journal Of Time Series Analysis.** 12(2), 129-142, (1989).
4. Karlis, D. and Xekalaki, E. Mixed Poisson Distributions, **Revue Internationale De Statistique.** 73(1), 35-58, (2005).
5. Khodabin, M. and Ahmadabadi, A. Some properties of generalized gamma distribution. **Mathematical Sciences.** 4(1), 9-28, (2010).
6. Schweer, S. and Weiss, C. Compound Poisson INAR(1) processes: Stochastic properties and testing for overdispersion. **Computational Statistics And Data Analysis.** 77(C), 267-284, (2014).

Author Profile

Miss Panisara Yamsook ID 583 35321 23

Branch of Mathematics

Department of Mathematics and Computer Science

Chulalongkorn University