Order Identification and Estimation of Moving Average and Auto-Regressive Dynamic Models for Count Time Series

by

© **Kirushanthini Balakrishnan**

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science.

Department of Mathematics and Statistics

Memorial University

May 2019

St. John's, Newfoundland and Labrador, Canada

Abstract

Time series of count data occur frequently in practice such as in medical studies and life sciences. Identification of a proper model for time series data is extremely important as it reflects the underlying structure of the time series and the fitted model will be used for forecasting. We first discuss the basic properties of the stationary Poisson moving average (MA) and stationary Poisson auto-regressive (AR) processes up to order 3 with the intention of finding a method for model identification. Some authors have derived and discussed the basic properties of stationary Poisson MA(q) process and stationary Poisson AR(1) and AR(2) processes for analysis of count time series data. We have extended to stationary Poisson AR(3) process and derived the basic properties of mean, variance, covariance and correlation of it. We discussed auto-correlation function (ACF) of stationary Poisson MA and AR processes up to order 3 and derived partial auto-correlation function (PACF) of stationary Poisson MA up to order 2 and stationary Poisson AR processes up to order 3 in order to find the theoretical patterns of the processes for model identification purposes. The patterns and behaviour of ACF and PACF of these stationary MA and AR Poisson models have been discussed in detail. In each of the cases, the accuracy of the patterns of ACF and PACF are examined through simulation studies. We found that patterns in the ACF and PACF of these models are similar to those of AR and MA models for Gaussian processes. We have also proposed a model for non stationary Poisson AR(3) process and the basic properties have been derived. Model parameters are estimated using generalized quasi-likelihood (GQL) and generalized method of moment (GMM) methods. The performance of the estimation methods have been examined through simulation studies.

This thesis is dedicated to the memory of my beloved father, S. Balakrishnan.

Lay summary

Time series analysis of count data is a dynamic research area which has attracted the attention of researchers over the last few decades. There are two main goals of time series analysis such as, identifying the nature of the phenomenon represented by the sequence of observations, and forecasting (prediction of short-term trends from previous patterns). Both of these goals require the correct identification of patterns in observed time series data correctly. To the best of our knowledge there are currently no existing studies in the literature on the identification of dynamic Poisson models for count time series. In this study, we have derived some useful theoretical patterns and proposed a method for model identification of stationary count data. We also found basic properties of a higher order (order 3) of Auto-Regressive model for non stationary count time series and found estimates of the model parameters using statistical methods, which can be used for forecasting of time series data.

Acknowledgements

Completion of this dissertation was possible with the support of several people. I would like to express my sincere gratitude to all of them. First and foremost, I am extremely grateful to my co-supervisor Prof. Alwell Julius Oyet, whose expertise, encouragement, and patience, added considerably to my graduate experience. I appreciate his vast knowledge and skill in many areas, and his assistance in programming language and writing reports. He has always made himself available to clarify my doubts regardless of his busy schedules and I consider it as a great opportunity to do my research under his guidance and to learn from his research expertise. Thank you, for all your help and support.

My sincere gratitude goes to my supervisor Prof. Asokan Variyath, who has enormously supported me throughout my thesis with his patience and knowledge whilst giving advice, guidance and financial support during my Masters program. Your advice and encouragement on both research as well as on my career have been invaluable.

During my Masters program, I had the opportunity to learn and work with many great professors in the Department of Statistics and I indebted to them. Many thanks to all the faculty members in the department for providing an excellent atmosphere to grow both professionally and personally.

I would also like to express my deepest gratitude to my family and friends. This accomplishment would not have come true without their warm love, continued patience, and endless support.

In conclusion, I recognize that this research would not have been possible without the financial assistance of Department of Mathematics and Statistics and Graduate Studies of Memorial University of Newfoundland, and express my gratitude to them.

Statement of contribution

Prof. Alwell Julius Oyet proposed the research question that was investigated throughout this thesis. The overall study was jointly designed by Prof. Alwell Julius Oyet, Prof. Asokan Variyath and Kirushanthini Balakrishnan. All the theoretical derivations and simulation study was conducted and the manuscript was drafted by Kirushanthini Balakrishnan. Prof. Alwell Julius Oyet and Prof. Asokan Variyath supervised the study and contributed to the final manuscript.

Table of contents

List of tables

List of figures

List of abbreviations

INAR	Integer valued Auto-Regressive
INMA	Integer valued Moving Average
INARMA	Integer valued Auto-Regressive and Moving Average
AR	Auto-Regressive
MA	Moving Average
AIC	Akaike Information Criterion
ACF	Auto-Correlation Function
PACF	Partial Auto-Correlation Function
SACF	Sample Auto-Correlation Function
SPACF	Sample Partial Auto-Correlation Function
GEE	Generalized Estimating Equation
GQL	Generalized Quasi-Likelihood
GMM	Generalized Method of Moment
SM	Simulated Mean
SSE	Simulated Standard Error
TSE	True Standard Error
ESE	Estimated Standard Error
JSE	Jackknife Standard Error

Chapter 1

Introduction

The modeling in time series involves information about the mathematical model of the process. However, in practice, patterns of the data are unclear. Hence, we first need to discover the hidden patterns in the data. Determining the order of an auto-regressive moving average (ARMA) process is an important and often problematic part of time series analysis. The autoregressive integrated moving average (ARIMA) methodology developed by Box and Jenkins (1976) has gained enormous popularity in many areas and research practices confirmed its power and flexibility (Hoff, 1983; Pankratz, 2009; Vandaele, 1983). The Box-Jenkins methodology involves the steps of model identification, parameter estimation and model validation (Box et al., 2015). Identification of suitable ARIMA model can be done by choosing the model with lowest AIC (Akaike Information Criterion) (Bozdogan, 1987). Ozaki (1977) discussed the difficulty in deciding the order of an ARIMA model and the possibility of removing this difficulty by using the MAICE (minimum AIC estimation) procedure which has been studied with the numerical examples treated in the book by Box and Jenkins (1976).

Several order idntification methods used in time series analysis are reviewed by de Gooijer et al. (1985) and their theoretical and practical relevance are discussed. According to de Gooijer et al. (1985), an order determination method which minimizes the one-step-ahead quadratic forecasting error is more preferable than methods which merely lead to fitted models that provide a good representation of the data in the sample. Koreisha and Pukkila (1995) also compared the small-sample performance of several order-determination criteria for ARIMA models using simulated and economic data. They have demonstrated how the residual white-noise autoregressive order-determination criterion can be used to identify unit roots in nonstationary data.

Pukkila et al. (1990) proposed a simple and powerful procedure for determining the order of ARMA (p, q) models having small sets of observations. It is based on an autoregressive order determination criterion and linear estimation method to identify the order of ARMA processes from a finite set of observations. The performance of their identification procedure has been demonstrated and found powerful based on simulations of several model structures with varying number of observations. Garel and Hallin (1999) investigated finite-sample performances of an optimal rank-based procedures in the context of AR order identification and compared to those of classical (partial correlograms and Lagrange multipliers) method.

However, all the order identification procedures discussed above are for continuous time series processes. In addition, there are more literature available for the order identification of continuous AR and MA processes. Recently, there has been an increasing interest in modeling time series of count data as there are many situations in practice where the response of interest is discrete. In the past several years, various models have been proposed for discrete time stationary processes as well as for non stationary processes. But to the best of our knowledge, there is no procedures proposed for order identification of AR and MA processes for count time series.

Generalization of integer valued auto-regressive of order 1 (INAR(1)) to INAR(p) and integer valued auto-regressive moving average of order (p,q) (INARMA(p,q)) models are given in various papers by Al-Osh and Alzaid (1987,1988 and 1990) and by McKenzie(1985 and 1988). According to McKenzie (1988), an immediate limitation of all these models is that all correlations are positive which is an unavoidable property when joint distribution are multivariate Poisson. In the following Sections 1.1 to 1.4 we will discuss the literature about the models for stationary Poisson AR and MA processes.

1.1 Stationary Poisson MA(q) Process

An integer-valued moving average (INMA) process and several properties for the special case of Poisson INMA (1) process, such as the joint distribution, regression, time reversibility, along with the conditional and partial correlations, are discussed in details by Al-Osh and Alzaid (1988). INMA(q) process X_t was given by,

$$X_t = \sum_{i=1}^{q} b_i \circ W_{t-i} + W_t \tag{1.1}$$

where, W_t are i.i.d Poisson random variables, the constants $b_1, b_2, ..., b_q$ each lie in [0,1] and $b_i \circ W_{t-1} = \sum_{j=1}^{W_{t-i}} Y_j(b_i)$ is the Binomial thinning operation where $Y_j(b_i)$ is a sequence of independent identically distributed (i.i.d) binary random variables with $Pr[Y_j(b_i) = 1] = b_i$ and $Pr[Y_j(b_i) = 0] = 1 - b_i$. Binomial thinning operation is used for model building in Poisson processes instead of the scalar multiplication in the standard auto-regressive moving average (ARMA) processes. The definition of thinning operator ascends from the work of Steutel and Van Harn (1979).

McKenzie (1988) also developed and investigated a family of ARMA models for discrete-time stationary processes with Poisson marginal distributions. Binomial thinning ($*$ instead of $\circ$) has been used for model building purpose. McKenzie (1988) discussed the form of stationary Poisson MA(1) process and generalized that to stationary Poisson MA(q) process. The structure of the MA(q) process X_t was given by

$$X_t = Y_t + \rho_1 * Y_{t-1} + \rho_2 * Y_{t-2} + ... + \rho_q * Y_{t-q} \tag{1.2}$$

where Y_t are i.i.d Poisson random variates, the constants $\rho_1, \rho_2, ..., \rho_q$ each lie in [0,1] and the Binomial thinning operations are performed independently. That is to say, the random variable $\rho_i * Y_{t-i}$ is independent of Y_{t-j} and $\rho_j * Y_{t-j}$ for $i, j = 1, 2, ..., q$ where $i \neq j$. In addition, he derived the auto-correlation function (ACF) for the MA(q) process and found joint distribution for stationary Poisson MA(1) process.

1.2 Stationary Poisson AR(1) Process

Al-Osh and Alzaid (1987) constructed INAR(1) model. They dealt with discrete time stationary processes with Poisson marginal distribution using first order branching process with the assumption of binary offspring. INAR(1) process X_t was given by,

$$X_t = \alpha \circ X_{t-1} + \varepsilon_t \tag{1.3}$$

where, $\alpha \circ X_{t-1} = \sum_{j=1}^{X_{t-1}} Y_j(\alpha)$ is the binomial thinning operation which is defined as $Pr[Y_j(\alpha) = 1] = \alpha$ and $Pr[Y_j(\alpha) = 0] = 1 - \alpha$. It has been assumed that an element of the Poisson process at time $t-1$, X_{t-1} is independent of ε_t and X_{t-1} is distributed as of X_t.

Al-Osh and Alzaid (1987) also found that the correlation structure and the distributional properties of the INAR(1) model are similar to those of the stationary Gaussian autoregressive (AR) of order 1 process. Various methods for estimating the parameters of the model have been discussed by them. McKenzie (1988) also investigated stationary Poisson AR(1) model and discussed the ACF. To our knowledge, some other authors who have studied the Poisson AR(1) model are Alzaid and Al-Osh (1993); Park and Oh (1997); Böckenholt (1998); Ispány et al. (2003); Freeland and McCabe (2005); Jung and Tremayne (2006); Weiß (2007) and Bakouch and Ristić (2010).

1.3 Stationary Poisson AR(2) Process

Zhang and Oyet (2014) proposed Auto-Regressive (AR) model of order 2 for longitudinal count time series which was given by,

$$\begin{aligned} y_{i1} &\sim Poi(\mu_{i1}) \quad with \quad \mu_{i1} = exp(\boldsymbol{x}'_{i1}\boldsymbol{\beta}), \\ y_{i2} &= \sum_{j=1}^{y_{i1}} b_{1j}(\rho_1) + d_{i2}, \\ y_{it} &= \sum_{l=1}^{2} \sum_{j=1}^{y_{i,t-l}} b_{lj}(\rho_l) + d_{it}, \quad for \quad t = 3, 4, ..., T, \end{aligned} \tag{1.4}$$

with the following assumptions:

i $d_{i2} \sim Poi(\mu_{i2} - \rho_1\mu_{i1})$.

ii $d_{it} \sim Poi(\mu_{it} - \rho_1\mu_{i,t-1} - \rho_2\mu_{i,t-2})$, for $t = 3, 4, ..., T$.

iii d_{it} and $y_{i,t-1}$ are independent for $t = 2, 3, ..., T$.

iv d_{it} and $y_{i,t-2}$ are independent for $t = 3, 4, ..., T$.

where, $\boldsymbol{x}_{it}$ is a vector of covariates, $\boldsymbol{\beta}$ is a measure of the covariate effect and ρ_1 and ρ_2 are correlation index parameters. Basic properties of mean, variance, covariance and correlation of the model were derived by them. They also estimated the model parameters using generalized quasi-likelihood (GQL) (Wedderburn, 1974) and generalized method of moment (GMM) (Hansen, 1982) estimation and examined the performance of the methods through simulation studies. The stationary Poisson AR(2) model with binary offsprings can be obtained from their model by setting $n_t = 1$.

1.4 Stationary Poisson AR(p) Process

The INAR(1) model is further generalized to a p^{th} order integer valued auto-regressive process (INAR(p)) by Alzaid and Al-Osh (1990) as,

$$X_n = \sum_{i=1}^{p} \alpha_i \circ X_{n-1} + \varepsilon_n \quad \text{for} \quad n = 0, \pm 1, ..., \tag{1.5}$$

where, ε_n is a non-negative integer valued random variable with finite mean and variance. They have discussed some assumptions and peculiarities as well as the limiting distribution of the process and given a detailed derivation of the auto-correlation function.

1.5 Estimation Methods

For the forecasting purposes, it is important to accurately estimate the regression parameters of the identified model. Several methods have been proposed for estimating the parameters of Poisson AR and MA models.

Generalized estimating equation (GEE) approach was introduced by Liang and

Zeger (1986) based on a working correlation matrix to obtain consistent and efficient estimators of regression parameters in the class of generalized linear models for repeated measures data. Crowder (1995) has demonstrated that the Liang-Zeger approach may in some cases encounter to a complete breakdown of the estimation of the regression parameters because of the uncertainty of definition of the working correlation matrix. Sutradhar and Das (1999) showed that, while the Liang-Zeger approach in many situations yields consistent estimators to the regression parameters, these estimators were usually inefficient as compared to the regression estimators obtained by using the independence estimating equations approach.

Barron (1992) described maximum likelihood-Poisson and negative binomial regression, quasi-likelihood (QL), generalized quasi-likelihood (GQL) methods that can be used to analyze count data and found that quasi-likelihood methods have some advantages over Poisson and negative binomial regression methods, especially in the presence of auto-correlation. They have also investigated the small-sample properties of these estimators in the presence of over-dispersion and auto-correlation by means of Monte Carlo simulations. Maximum likelihood (ML) estimation of both Poisson and negative binomial regressions typically require independent observations and this assumption will often not be true in time-series data, and Poisson and negative binomial regression are then problematic (Barron, 1992). Sutradhar and Bari (2007) used GQL estimation approach in estimating the regression effects of the time dependent covariates on the responses, the over dispersion parameter, as well as the serial correlation parameter of the longitudinal mixed model for count data.

Oyet and Sutradhar (2013) and Zhang and Oyet (2014) have demonstrated that the GQL method performs well in estimating the parameters of the longitudinal model of infectious disease and they have used generalized method of moment (GMM) estimation for estimating the correlation parameters. Sutradhar et al. (2008) also used the same estimation methods for familial-longitudinal non stationary count data.

1.6 Motivation

Fitting an adequate model to the underlying time series should be done carefully due to the necessary importance of time series forecasting in numerous practical fields such as business, finance, economics, science and engineering, etc. First of all we noticed the patterns in the continuous time series which are used for model identification purposes. While studying Poisson AR and MA process, we noticed that the same patterns seem to exist for the count MA series. We found that this problem has not been studied in past to the best of our knowledge, so we decided to study further as it's an interesting problem. When it comes to count AR series, Poisson AR(1) and AR(2) processes have been proposed and the basic properties have been derived which are discussed in Sections 1.2 and 1.3. We noticed the same patterns in those as well. Therefore we proposed Poisson AR(3) model and derived the basic properties as there were no existing studies regarding that in the literature to our knowledge. In addition, since same patterns were observed in the models, these patterns can be used for the model identification of Poisson AR and MA process.

One of the main purpose of our research was to discuss the patterns in the ACF and PACF of stationary AR and MA Poisson models and propose a method for order identification of the models. In Chapter 2 we discuss the structure of stationary AR and MA Poisson models upto order three and their basic properties mean, variance, co-variance and correlation. In order to find the patterns in these models, it is also important to know the theoretical structure of partial auto-correlation function (PACF) of the considered models. However, it is complicated to find PACF at all lag points. Hence, we derived first few theoretical PACF of these models. In Section 2.3, we have discussed and examined the patterns and behavior of ACF and PACF of the models through simulation studies. In the simulation studies we have used Jackknife estimation of the standard error to find the significance of ACF and PACF and detailed explanation is provided in Section 2.3.1.

In Chapter 3, we derive basic properties (mean, variance and correlation) of non stationary Poisson AR(3) process. We obtain the estimator for the parameter of the covariate ($\boldsymbol{\beta}$) using GQL and the estimator for the correlation index parameters of the model (ρ_1, ρ_2 and ρ_3) using GMM estimation which are detailed in Sections 3.2.1 and 3.2.2. The small sample performance of the estimation methods has been examined through a simulation study in Section 3.3. Chapter 4 concludes the thesis with summary of our findings and some directions for future work.

Chapter 2

Patterns and Order Identification of Stationary AR and MA Poisson Models for Count Time Series

The identification of a proper model for time series data is extremely important as it reflects the underlying structure of the series and the fitted model will be used for forecasting. In order to determine the proper model for a given time series data, it is sometimes necessary to carry out the ACF and PACF analysis. It is often very useful to plot the ACF and PACF against consecutive time lags. In this study, these plots will be used to determine the order of AR and MA process. In our study we have considered stationary Poisson AR and MA models upto order three. Our model building approach is also parallel to that of Al-Osh and Alzaid (1987, 1988), Alzaid and Al-Osh (1988, 1990) and McKenzie (1985, 1988) which dealt with discrete time stationary processes with Poisson marginal distribution.

In Chapter 1, we mentioned that the Poisson MA(q) stationary models and stationary

Poisson AR(1) models have been discussed by McKenzie (1988) and non-stationary Poisson AR(2) models for longitudinal data have been discussed by Zhang and Oyet (2014). In both cases, they have derived only the theoretical ACF of the models. In this chapter, we will derive first few theoretical PACF of AR and MA stationary models upto order 3 for count data with Poisson marginal distribution. We will also discuss and examine the patterns and behavior of the ACF and PACF through simulation study. We will propose a method for determining the significance of ACF and PACF for time series of count data. Finally, we will examine the performance of our proposed method through an extensive simulation study.

2.1 Stationary MA models

2.1.1 Stationary Poisson MA(1) Process

For a correlation index parameter $0 < \rho < 1$,

$$d_t \overset{iid}{\sim} Poi\Big(\frac{\mu}{1+\rho}\Big), \quad for \quad t = 1, 2, ..., T,$$

where $\mu = exp(\boldsymbol{x}'\boldsymbol{\beta})$. Now, suppose that the response y_t is related to d_t (McKenzie, 1988) through the model,

$$y_t = \rho * d_{t-1} + d_t \quad , \tag{2.1}$$

where, $\rho * d_{t-1} = \sum_{j=1}^{d_{t-1}} b_j(\rho)$ is the binomial thinning operation with $Pr[b_j(\rho) = 1] = \rho$ and $Pr[b_j(\rho) = 0] = 1 - \rho$.

Basic Properties:

Mean and Variance

McKenzie (1988) has shown that the basic properties of the stationary Poisson MA(1) process (2.1) are given by,

$$E\,(y_t) = \mu \qquad \text{and} \qquad Var\,(y_t) = \mu. \tag{2.2}$$

Covariance and correlation

McKenzie (1988) also found that the lag k auto-covariance and auto-correlation functions are given by,

$$Cov(y_t, y_{t-k}) = c(k) = \begin{cases} \dfrac{\rho\mu}{1+\rho}, & k = 1 \\ 0, & k > 1, \end{cases} \tag{2.3}$$

and

$$Corr(y_t, y_{t-k}) = \gamma(k) = \frac{c(k)}{c(0)} = \begin{cases} \dfrac{\rho}{1+\rho}, & k = 1 \\ 0, & k > 1, \end{cases} \tag{2.4}$$

respectively. Following Wei (2006), we obtain the PACF of y_t at lag k, denoted by ϕ_{kk}, as follows.

Partial Autocorrelation

At $k = 1$, we find that the PACF at lag 1, denoted by ϕ_{11} is given by,

$$\phi_{11} = \gamma(1) = \frac{\rho}{1+\rho}.$$

Clearly, $\gamma(1) < 1$.

At $k = 2$, we obtain

$$\phi_{22} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(2) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{vmatrix}} = \frac{\gamma(2) - \gamma^2(1)}{1 - \gamma^2(1)} = \frac{-\gamma^2(1)}{1 - \gamma^2(1)} = \frac{-\rho^2}{1 + 2\rho} \qquad ; \gamma(2) = 0.$$

At $k = 3$, we find

$$\phi_{33} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(1) \\ \gamma(1) & \gamma(0) & \gamma(2) \\ \gamma(2) & \gamma(1) & \gamma(3) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(2) & \gamma(1) & \gamma(0) \end{vmatrix}} = \frac{\gamma^3(1)}{1 - 2\gamma^2(1)} \qquad ; \gamma(2) = \gamma(3) = 0$$

$$= \frac{\rho^3}{1 + 3\rho + \rho^2 - \rho^3}.$$

Similarly we find that the PACF at $k = 4$ and $k = 5$,

$$\phi_{44} = \frac{-\gamma^4(1)}{1 - 2\gamma^2(1) + \gamma^4(1)} \qquad ; \gamma(2) = \gamma(3) = \gamma(4) = 0.$$

$$\phi_{55} = \frac{-\gamma^5(1)}{1 - 2\gamma^2(1) + \gamma^4(1)} \qquad ; \gamma(2) = \gamma(3) = \gamma(4) = \gamma(5) = 0.$$

Now, using mathematical induction we write a general expression for PACF at lag k,

$$\phi_{kk} = \begin{cases} (-1)^{k+1}\gamma^1(1), & k=1 \\ \dfrac{(-1)^{k+1}\gamma^2(1)}{1-\gamma^2(1)}, & k=2 \\ \dfrac{(-1)^{k+1}\gamma^3(1)}{1-2\gamma^2(1)}, & k=3 \\ \dfrac{(-1)^{k+1}\gamma^k(1)}{(1-\gamma^2(1))^2}, & k>3 \end{cases}$$

where $\gamma(1) = \dfrac{\rho}{1+\rho} < 1$. Since $\gamma(1) < 1$, it is clear that $\gamma^k(1) \to 0$ as $k \to \infty$. Therefore, ϕ_{kk} will decay exponentially as $k \to \infty$, where as $\gamma(k)$ will cut-off after lag k. These patterns in the ACF and PACF of MA(1) count time series are clearly identical to the patterns in MA(1) Gaussian time series.

2.1.2 Stationary Poisson MA(2) Process

Following McKenzie (1988), we write the structure of the Stationary Poisson MA(2) process with correlation index parameters ρ_1 and ρ_2, as

$$y_t = \rho_1 * d_{t-1} + \rho_2 * d_{t-2} + d_t \quad , \tag{2.5}$$

where, $\rho_1 * d_{t-1} = \sum_{j=1}^{d_{t-1}} b_{1j}(\rho_1)$, $\rho_2 * d_{t-2} = \sum_{j=1}^{d_{t-2}} b_{2j}(\rho_2)$ with $d_t \overset{iid}{\sim} Poi\Big(\dfrac{\mu}{1+\rho_1+\rho_2}\Big)$, for $\quad t = 1, 2, ..., T; \quad \mu = exp(\boldsymbol{x}'\boldsymbol{\beta})$,

with

$Pr[b_{1j}(\rho_1) = 1] = \rho_1$ and $Pr[b_{1j}(\rho_1) = 0] = 1 - \rho_1$

$Pr[b_{2j}(\rho_2) = 1] = \rho_2$ and $Pr[b_{2j}(\rho_2) = 0] = 1 - \rho_2$

Basic Properties:

Mean and Variance

McKenzie (1988) has shown that the basic properties of the stationary Poisson MA(2) process (2.5) are

$$E\left(y_t\right) = \mu \quad \text{and} \qquad Var\left(y_t\right) = \mu. \tag{2.6}$$

Covariance and correlation

Auto-covariance and auto-correlation structures are given by,

$$Cov(y_t, y_{t-k}) = c(k) = \begin{cases} \dfrac{(\rho_1 + \rho_1\rho_2)\mu}{1+\rho_1+\rho_2}, & k = 1 \\ \dfrac{\rho_2\mu}{1+\rho_1+\rho_2}, & k = 2 \\ 0, & k > 2, \end{cases} \tag{2.7}$$

and

$$Corr(y_t, y_{t-k}) = \gamma(k) = \frac{c(k)}{c(0)} = \begin{cases} \dfrac{\rho_1 + \rho_1\rho_2}{1+\rho_1+\rho_2}, & k = 1 \\ \dfrac{\rho_2}{1+\rho_1+\rho_2}, & k = 2 \\ 0, & k > 2, \end{cases} \tag{2.8}$$

respectively. Now, using the results we derive the PACF of y_t for the first few lags using the ACF function $\gamma(k)$ given in equation (2.8).

Partial Autocorrelation

Now, at $k = 1$, we find

$$\phi_{11} = \gamma(1) = \frac{\rho_1 + \rho_1\rho_2}{1+\rho_1+\rho_2}.$$

At $k = 2$, we obtain

$$\phi_{22} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(2) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{vmatrix}} = \frac{\gamma(2) - \gamma^2(1)}{1 - \gamma^2(1)}$$

$$= \frac{\rho_2 + \rho_1\rho_2 + \rho_2^2 - \rho_1^2 - \rho_1^2\rho_2^2 - 2\rho_1^2\rho_2}{1 + \rho_2^2 + 2\rho_1 + 2\rho_2 + 2\rho_1\rho_2 - \rho_1^2\rho_2^2 - 2\rho_1^2\rho_2}.$$

At $k = 3$, we find

$$\phi_{33} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(1) \\ \gamma(1) & \gamma(0) & \gamma(2) \\ \gamma(2) & \gamma(1) & \gamma(3) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(2) & \gamma(1) & \gamma(0) \end{vmatrix}}$$

$$= \frac{-2\gamma(1)\gamma(2) + \gamma(1)\gamma^2(2) + \gamma^3(1)}{1 - 2\gamma^2(1) - \gamma^2(2) + 2\gamma^2(1)\gamma(2)}, \qquad \gamma(3) = 0.$$

At $k = 4$, we find

$$\phi_{44} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \gamma(1) \\ \gamma(1) & \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(2) & \gamma(1) & \gamma(0) & \gamma(3) \\ \gamma(3) & \gamma(2) & \gamma(1) & \gamma(4) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \gamma(3) \\ \gamma(1) & \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(2) & \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(3) & \gamma(2) & \gamma(1) & \gamma(0) \end{vmatrix}}$$

$$= \frac{3\gamma^2(1)\gamma(2) - 2\gamma^2(1)\gamma^2(2) - \gamma^2(2) - \gamma^4(1) + \gamma^4(2)}{1 - 3\gamma^2(1) - \gamma^2(2) + 3\gamma^2(1)\gamma(2) - \gamma^2(1)\gamma^2(2) - 2\gamma(1)\gamma^2(2) + \gamma^3(1)\gamma(2) + \gamma(1)\gamma^3(2) + \gamma^4(1)}$$

where $\gamma(3) = \gamma(4) = 0$. The basic properties of the MA(2) process given in this section show that the ACF cuts off after lag 2.

2.1.3 Stationary Poisson MA(3) Process

Following McKenzie (1988), we write the structure of the Stationary Poisson MA(3) process with correlation index parameters ρ_1 ,ρ_2 and ρ_3 , as

$$y_t = \rho_1 * d_{t-1} + \rho_2 * d_{t-2} + \rho_3 * d_{t-3} + d_t \quad , \tag{2.9}$$

where, $\rho_1 * d_{t-1} = \sum_{j=1}^{d_{t-1}} b_{1j}(\rho_1)$, $\rho_2 * d_{t-2} = \sum_{j=1}^{d_{t-2}} b_{2j}(\rho_2)$, $\rho_3 * d_{t-3} = \sum_{j=1}^{d_{t-3}} b_{3j}(\rho_3)$,
$d_t \overset{iid}{\sim} Poi(\frac{\mu}{1+\rho_1+\rho_2+\rho_3}), \quad for \quad t = 1, 2, ..., T, \quad \mu = exp(\boldsymbol{x}'\boldsymbol{\beta})$

with

$Pr[b_{1j}(\rho_1) = 1] = \rho_1$ and $Pr[b_{1j}(\rho_1) = 0] = 1 - \rho_1$,

$Pr[b_{2j}(\rho_2) = 1] = \rho_2$ and $Pr[b_{2j}(\rho_2) = 0] = 1 - \rho_2$,

$Pr[b_{3j}(\rho_3) = 1] = \rho_3$ and $Pr[b_{3j}(\rho_3) = 0] = 1 - \rho_3$.

Basic Properties:

Mean and Variance

McKenzie (1988) has shown that the basic properties of the stationary Poisson MA(3) process (2.9) are

$$E(y_t) = \mu \qquad \text{and} \qquad Var(y_t) = \mu. \tag{2.10}$$

Covariance and correlation

Auto-covariance and auto-correlation structures are given by,

$$Cov(y_t, y_{t-k}) = c(k) = \begin{cases} \dfrac{(\rho_1 + \rho_1\rho_2 + \rho_2\rho_3)\mu}{1 + \rho_1 + \rho_2 + \rho_3}, & k = 1 \\ \dfrac{(\rho_2 + \rho_1\rho_3)\mu}{1 + \rho_1 + \rho_2 + \rho_3}, & k = 2 \\ \dfrac{\rho_3\mu}{1 + \rho_1 + \rho_2 + \rho_3}, & k = 3 \\ 0, & k > 4, \end{cases} \tag{2.11}$$

and

$$Corr(y_t, y_{t-k}) = \gamma(k) = \frac{c(k)}{c(0)} = \begin{cases} \dfrac{\rho_1 + \rho_1\rho_2 + \rho_2\rho_3}{1 + \rho_1 + \rho_2 + \rho_3}, & k = 1 \\ \dfrac{\rho_2 + \rho_1\rho_3}{1 + \rho_1 + \rho_2 + \rho_3}, & k = 2 \\ \dfrac{\rho_3}{1 + \rho_1 + \rho_2 + \rho_3}, & k = 3 \\ 0, & k > 4, \end{cases} \tag{2.12}$$

respectively. The basic properties of the MA(3) process given in this section show that the ACF cuts off after lag 3.

2.1.4 Stationary Poisson MA(q) Process

For some correlation index parameters $\rho_1, \rho_2, ..., \rho_q$, structure of the Stationary Poisson MA(q) process (McKenzie, 1988) can be written by,

$$y_t = \rho_1 * d_{t-1} + \rho_2 * d_{t-2} + ... + \rho_q * d_{t-q} + d_t \quad , \tag{2.13}$$

where $\rho_k * d_{t-k} = \sum_{j=1}^{d_{t-k}} b_{kj}(\rho_k)$ is the binomial thinning operation for $k = 1, 2, ..., q$. That is, $Pr[b_{kj}(\rho_k) = 1] = \rho_k$ and $Pr[b_{kj}(\rho_k) = 0] = 1 - \rho_k$

and

$$d_t \overset{iid}{\sim} Poi(\frac{\mu}{1 + \rho_1 + \rho_2 + ... + \rho_q}), \quad for \quad t = 1, 2, ..., T,$$

with $\mu = exp(\boldsymbol{x}'\boldsymbol{\beta})$.

Basic Properties:

Mean and Variance

According to McKenzie (1988), mean and variance of MA(q) stationary Poisson process are,

$$E\left(y_t\right)=\mu \qquad \text{and} \qquad Var\left(y_t\right)=\mu. \tag{2.14}$$

Co-variance and Correlation

Auto-covariance and auto-correlation of MA(q) stationary Poisson process (McKenzie, 1988) are given by,

$$Cov(y_t, y_{t-k}) = c(k) = \begin{cases} \dfrac{\sum_{i=0}^{q-k}\rho_i\rho_{i+k}\mu}{\sum_{i=0}^{q}\rho_i}, & k=1,2,...,q \quad with \quad \rho_0=1 \\ 0, & k>q, \end{cases} \tag{2.15}$$

and

$$Corr(y_t, y_{t-k}) = \gamma(k) = \frac{c(k)}{c(0)} = \begin{cases} \dfrac{\sum_{i=0}^{q-k}\rho_i\rho_{i+k}}{\sum_{i=0}^{q}\rho_i}, & k=1,2,...,q \quad with \quad \rho_0=1 \\ 0, & k>q, \end{cases} \tag{2.16}$$

respectively. We can see that the theoretical ACF ($\gamma(k)$) are 0 after lag q. We will examine the accuracy of this pattern using simulation study for MA(1), MA(2) and MA(3) stationary Poisson processes.

2.2 Stationary AR models

2.2.1 Stationary Poisson AR(1) Process

Let $y_1 \sim Poi(\mu)$, where $\mu = exp(\boldsymbol{x}'\boldsymbol{\beta})$. For a correlation index parameter ρ_1, suppose the response y_t is related to y_{t-1} as,

$$y_t = \rho * y_{t-1} + d_t, \quad for \quad t = 2, 3, ..., T, \tag{2.17}$$

where, $d_t \sim Poi(\mu(1-\rho))$ and $\rho * y_{t-1} = \sum_{j=1}^{y_{t-1}} b_j(\rho)$ is the binomial thinning operation(McKenzie, 1988). That is, $Pr[b_j(\rho) = 1] = \rho$ and $Pr[b_j(\rho) = 0] = 1 - \rho$. Here we assume that d_t and y_{t-1} are independent for all $t = 2, 3, ..., T$.

Basic Properties:

Mean and Variance

McKenzie (1988) has shown that the basic properties of the stationary Poisson AR(1) process (2.17),

$$E(y_t) = \mu \qquad \text{and} \qquad Var(y_t) = \mu. \tag{2.18}$$

Covariance and correlation

Covariance and correlation structure will be in the following form.

$$Cov(y_t, y_{t-k}) = \rho^k \mu \qquad \text{and} \qquad Corr(y_t, y_{t-k}) = \gamma(k) = \rho^k. \tag{2.19}$$

Now, we find the form of the PACF at each lag k, denoted by ϕ_{kk}, using ACF function $\gamma(k)$ given in equation (2.19).

Partial Autocorrelation

Now, at $k = 1$, we find

$$\phi_{11} = \gamma(1) = \rho.$$

At $k = 2$, we find

$$\phi_{22} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(2) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{vmatrix}} = \frac{\begin{vmatrix} 1 & \rho \\ \rho & \rho^2 \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{vmatrix}} = 0,$$

since columns in the numerator are linearly dependent.

At $k = 3$, we find

$$\phi_{33} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(1) \\ \gamma(1) & \gamma(0) & \gamma(2) \\ \gamma(2) & \gamma(1) & \gamma(3) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(2) & \gamma(1) & \gamma(0) \end{vmatrix}} = \frac{\begin{vmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho^2 \\ \rho^2 & \rho & \rho^3 \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(2) & \gamma(1) & \gamma(0) \end{vmatrix}} = 0,$$

since 1st column and 3rd column in the numerator are linearly dependent.

Similarly we can find PACF at any lag k. Now, we write a general expression for PACF at lag k,

$$\phi_{kk} = \begin{cases} \rho, & k = 1 \\ 0, & k > 1. \end{cases} \tag{2.20}$$

The basic properties of the AR(1) process given in this section show that the ACF decays exponentially as k increases while the PACF cuts off after lag 1.

2.2.2 Stationary Poisson AR(2) Process

Zhang and Oyet (2014) introduced a non-stationary second order dynamic model for longitudinal count data with Binomial offsprings. The stationary Poisson AR(2) model with Binary offsprings can be obtained from the above model by setting $n_t = 1$,

$$\begin{aligned} &y_1 \sim Poi(\mu) \quad with \quad \mu = exp(\boldsymbol{x}'\boldsymbol{\beta}) \\ &y_2 = \rho_1 * y_1 + d_2 \\ &y_t = \rho_1 * y_{t-1} + \rho_2 * y_{t-2} + d_t, \quad for \quad t = 3, 4, ..., T, \end{aligned} \tag{2.21}$$

where ρ_1 and ρ_2 are correlation index parameters with the following assumptions:

i $d_2 \sim Poi((1 - \rho_1)\mu)$.

ii $d_t \sim Poi((1 - \rho_1 - \rho_2)\mu)$, for $t = 3, 4, ..., T$.

iii d_t and y_{t-1} are independent for $t = 2, 3, ..., T$.

iv d_t and y_{t-2} are independent for $t = 3, 4, ..., T$.

where, ρ_1 and ρ_2 are correlation index parameters.

In assumptions (i) and (ii), since the mean of a Poisson random variable is nonnegative, the parameter ρ_1 in model (2.21) must satisfy the conditions $\rho_1 < 1$ and $\rho_1 + \rho_2 < 1$. These two conditions imply that for a fixed value of $\rho_2 \geq 0$ and $t \geq 3$,

$$0 \leq \rho_1 \leq (1 - \rho_2) \tag{2.22}$$

in model (2.21). Now we obtain the basic properties of the stationary Poisson AR(2) process (2.21).

Basic Properties:

Mean

From the AR(2) model (2.21), we see that $E(y_1) = \mu$.

Now, at $t = 2$, we find that

$$\begin{aligned} E(y_2) &= E\Big[\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big] \\ &= E_{y_1} E\Big[\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2 \Big| y_1\Big] \\ &= E_{y_1}\Big[y_1\rho_1 + \mu - \rho_1\mu\Big] \\ &= \rho_1\mu + \mu - \rho_1\mu \\ &= \mu. \end{aligned}$$

In general, we can show that the mean of y_t for $t = 3, 4, ..., T$ is

$$
\begin{aligned}
E(y_t) &= E\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1) + \sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2) + d_t\Big] \\
&= E\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1)\Big] + E\Big[\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big] + E\Big[d_t\Big] \\
&= E_{y_{t-1}} E\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1)\Big|y_{t-1}\Big] + E_{y_{t-2}} E\Big[\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big|y_{t-2}\Big] + E\Big[d_t\Big] \\
&= \rho_1 E[y_{t-1}] + \rho_2 E[y_{t-2}] + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2}.
\end{aligned}
$$

Now, by taking successive expectation we find that $E(y_t) = \mu = exp(\boldsymbol{x}'\boldsymbol{\beta})$ for all $t = 1, 2, ..., T$.

Variance

In order to obtain the $Var(y_t)$, we first note that $b_{ij}(\rho_i)$ is a sequence of independent identically distributed (i.i.d) Binary random variables with $Pr[b_{ij}(\rho_i) = 1] = \rho_i$ and $Pr[b_{ij}(\rho_i) = 0] = 1 - \rho_i$ for all $i = 1, 2$. Therefore the covariance term will vanish when we take variance to the response y_t, that is

$$
Cov\Big(\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1), \sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big) = 0.
$$

We also note that from the assumptions 4 and 5 of model (2.21),

$$
Cov\Big(\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1), d_t\Big) = Cov\Big(\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2), d_t\Big) = 0.
$$

From model (2.21), we see that $\sigma_{1,1} = Var(y_1) = \mu_1$. Now at $t = 2$, we find that

$$\begin{aligned}
\sigma_{2,2} &= Var(y_2) \\
&= Var\Big[\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big] \\
&= E_{y_1}\Big[Var\Big(\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big|y_1\Big)\Big] + Var_{y_1}\Big[E\Big(\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big|y_1\Big)\Big] \\
&= E_{y_1}\Big[y_1\rho_1(1-\rho_1) + \mu - \rho_1\mu\Big] + Var_{y_1}\Big[y_1\rho_1 + \mu - \rho_1\mu\Big] \\
&= \rho_1(1-\rho_1)E_{y_1}\Big[y_1\Big] + \mu - \rho_1\mu + \rho_1^2 Var_{y_1}\Big[y_1\Big] \\
&= \rho_1(1-\rho_1)\mu + \mu - \rho_1\mu + \rho_1^2\mu = \mu.
\end{aligned}$$

In general, to obtain the $Var(y_t)$ for $t = 4, 5, ..., T$, we first find

$$\begin{aligned}
Var\Big(y_t\Big|y_{t-1}, y_{t-2}\Big) &= Var\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1) + \sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2) + d_t\Big|y_{t-1}, y_{t-2}\Big] \\
&= Var\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1)\Big|y_{t-1}\Big] + Var\Big[\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big|y_{t-2}\Big] + Var\Big[d_t\Big] \\
&= y_{t-1}\rho_1(1-\rho_1) + y_{t-2}\rho_2(1-\rho_2) + (1-\rho_1-\rho_2)\mu.
\end{aligned}$$

Then, using the conditional variance property we find

$$\begin{aligned}
&Var\Big(y_t\Big|y_{t-2}\Big) \\
&= E_{y_{t-1}}\Big[Var\Big(y_t\Big|y_{t-1}, y_{t-2}\Big)\Big] + Var_{y_{t-1}}\Big[E\Big(y_t\Big|y_{t-1}, y_{t-2}\Big)\Big] \\
&= E_{y_{t-1}}\big[\rho_1(1-\rho_1)y_{t-1} + \rho_2(1-\rho_2)y_{t-2} + (1-\rho_1-\rho_2)\mu\big] \\
&\quad + Var_{y_{t-1}}\big[\rho_1 y_{t-1} + \rho_2 y_{t-2} + (1-\rho_1-\rho_2)\mu\big] \\
&= \rho_1(1-\rho_1)\mu + \rho_2(1-\rho_2)y_{t-2} + \rho_1^2\sigma_{t-1,t-1} + (1-\rho_1-\rho_2)\mu.
\end{aligned}$$

Next, by using the property that

$$Var(y_t) = E_{y_{t-2}}\Big[Var\Big(y_t\Big|y_{t-2}\Big)\Big] + Var_{y_{t-2}}\Big[E\Big(y_t\Big|y_{t-2}\Big)\Big]$$

we can obtain a recursive relation variance formula for t = 3,4,...,T.

$$\begin{aligned}
\sigma_{t,t} &= Var(y_t) \\
&= E_{y_{t-2}}\Big[Var\Big(y_t\Big|y_{t-2}\Big)\Big] + Var_{y_{t-2}}\Big[E\Big(y_t\Big|y_{t-2}\Big)\Big] \\
&= E_{y_{t-2}}\Big[\rho_1(1-\rho_1)\mu + \rho_2(1-\rho_2)y_{t-2} + \rho_1^2\sigma_{t-1,t-1} + (1-\rho_1-\rho_2)\mu\Big] \\
&\quad + Var_{y_{t-2}}\Big[\rho_1\mu + \rho_2 y_{t-2} + (1-\rho_1-\rho_2)\mu\Big] \\
&= \rho_1(1-\rho_1)\mu + \rho_2(1-\rho_2)\mu + (1-\rho_1-\rho_2)\mu + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2} \\
&= (1-\rho_1^2-\rho_2^2)\mu + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2}.
\end{aligned}$$

Hence, the recursive relation formula for the variance of y_t is

$$Var(y_t) = (1-\rho_1^2-\rho_2^2)\mu + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2} \quad for \quad t = 3, 4, ..., T$$

Now, by taking variance successively we show that $\sigma_{t,t} = Var(y_t) = \mu = exp(\boldsymbol{x'\beta})$ for all $t = 1, 2, ..., T$.

Covariance and Correlation structure:

For t=2, co-variance and correlation can be obtained from AR(1) process as,

$$\sigma_{t,t-k} = Cov(y_t, y_{t-k}) = \rho_1^k \mu,$$
$$\gamma(k) = Corr(y_t, y_{t-k}) = \rho_1^k. \tag{2.23}$$

Next, for $t = 3, 4, ..., T$

$$\begin{aligned}
\sigma_{t,t-k} &= Cov(y_t, y_{t-k}) \\
&= E\Big(y_t, y_{t-k}\Big) - E\Big(y_t\Big)E\Big(y_{t-k}\Big) \\
&= E_{y_{t-k}}\Big[E\Big(y_t y_{t-k}\Big|y_{t-k}\Big)\Big] - \mu^2 \\
&= E_{y_{t-k}}\Big[y_{t-k}\Big(y_{t-1}\rho_1 + y_{t-2}\rho_2 + (1-\rho_1-\rho_2)\mu\Big)\Big] - \mu^2 \qquad (2.24) \\
&= \rho_1\Big[E\Big(y_{t-1}y_{t-k}\Big) - \mu^2\Big] + \rho_2\Big[E\Big(y_{t-2}y_{t-k}\Big) - \mu^2\Big] \\
&= \rho_1 Cov(y_{t-1}, y_{t-k}) + \rho_2 Cov(y_{t-2}, y_{t-k}) \\
&= \rho_1\sigma_{t-1,t-k} + \rho_2\sigma_{t-2,t-k}.
\end{aligned}$$

Hence, for $t = 3, 4, ..., T$, we obtain the covariances from the recursive relation

$$\sigma_{t,t-k} = Cov(y_t, y_{t-k}) = \rho_1\sigma_{t-1,t-k} + \rho_2\sigma_{t-2,t-k}.$$

It follows that the lag k auto-correlation are given by,

$$\begin{aligned}
Corr(y_t, y_{t-k}) = \gamma(k) &= \frac{Cov(y_t, y_{t-k})}{\sqrt{Var(y_t)}\sqrt{Var(y_{t-k})}} \\
&= \frac{Cov(y_t, y_{t-k})}{\sqrt{\mu}\sqrt{\mu}} \\
&= \frac{\rho_1 \sigma_{t-1,t-k} + \rho_2 \sigma_{t-2,t-k}}{\sqrt{\mu\mu}} \qquad (2.25) \\
&= \rho_1 Corr(y_{t-1}, y_{t-k}) + \rho_2 Corr(y_{t-2}, y_{t-k}) \\
&= \rho_1 \gamma(k-1) + \rho_2 \gamma(k-2).
\end{aligned}$$

Now, we find the form of PACF at each lag k, denoted by ϕ_{kk} using the ACF function $\gamma(k)$ given in equation (2.24).

Partial Autocorrelation

The PACF at lag $k = 1$ is given by

$$\phi_{11} = \gamma(1) = \frac{\rho_1}{1 - \rho_2}.$$

At $k = 2$, we find

$$\phi_{22} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(2) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{vmatrix}} = \frac{\gamma(2) - \gamma^2(1)}{1 - \gamma^2(1)}$$

$$= \frac{\rho_2 + \dfrac{\rho_1^2}{1-\rho_2} - \dfrac{\rho_1^2}{(1-\rho_2)^2}}{1 - \dfrac{\rho_1^2}{(1-\rho_2)^2}}$$

$$= \frac{\rho_2(1-\rho_2)^2 + \rho_1^2(1-\rho_2) - \rho_1^2}{(1-\rho_2)^2 - \rho_1^2}.$$

At $k = 3$, we find

$$\phi_{33} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(1) \\ \gamma(1) & \gamma(0) & \gamma(2) \\ \gamma(2) & \gamma(1) & \gamma(3) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(2) & \gamma(1) & \gamma(0) \end{vmatrix}} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) & \rho_1\gamma(0) + \rho_2\gamma(1) \\ \gamma(1) & \gamma(0) & \rho_1\gamma(1) + \rho_2\gamma(0) \\ \gamma(2) & \gamma(1) & \rho_1\gamma(2) + \rho_2\gamma(1) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(2) & \gamma(1) & \gamma(0) \end{vmatrix}} = 0,$$

since the columns in the numerator are linearly dependent. Similarly we can find PACF at any lag k. Now, we write a general expression for PACF at lag k,

$$\phi_{kk} = \begin{cases} \dfrac{\rho_1}{1-\rho_2}, & k=1 \\ \dfrac{\rho_2(1-\rho_2)^2+\rho_1^2(1-\rho_2)-\rho_1^2}{(1-\rho_2)^2-\rho_1^2}, & k=2 \\ 0, & k>2. \end{cases} \tag{2.26}$$

The basic properties of the AR(2) process obtained in this section show that the ACF decays exponentially while the PACF cuts off after lag 2.

2.2.3 Stationary Poisson AR(3) Process

One of the main contributions of our research was to derive the properties of stationary Poisson AR(3) process and we will use that for discussing the patterns and model identifying purpose. For some correlation index parameters ρ_1, ρ_2 and ρ_3, the model for stationary Poisson AR(3) process can be defined by,

$$\begin{aligned} & y_1 \sim Poi(\mu) \quad with \quad \mu = exp(\boldsymbol{x}'\boldsymbol{\beta}), \\ & y_2 = \rho_1 * y_1 + d_2, \\ & y_3 = \rho_1 * y_2 + \rho_2 * y_1 + d_3, \\ & y_t = \rho_1 * y_{t-1} + \rho_2 * y_{t-2} + \rho_3 * y_{t-3} + d_t, \quad for \quad t = 4, 5, ..., T. \end{aligned} \tag{2.27}$$

We make the following assumptions:

i $d_2 \sim Poi((1-\rho_1)\mu)$.

ii $d_3 \sim Poi((1-\rho_1-\rho_2)\mu)$.

iii $d_t \sim Poi((1-\rho_1-\rho_2-\rho_3)\mu)$, for $t = 4, 5, ..., T$.

iv d_t and y_{t-1} are independent for $t = 2, 3, ..., T$.

v d_t and y_{t-2} are independent for $t = 3, 4, ..., T$.

vi d_t and y_{t-3} are independent for $t = 4, 5, ..., T$.

We obtain the basic properties of the stationary Poisson AR(3) process (2.24).

Basic properties:

Mean

From the above AR(3) model (2.24), we can see that $E(y_1) = \mu$.

Now, when $t = 2$, we find that

$$\begin{aligned} E(y_2) &= E\Big[\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big] \\ &= E_{y_1} E\Big[\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2 \Big| y_1\Big] \\ &= E_{y_1}\Big[y_1\rho_1 + \mu - \rho_1\mu\Big] \\ &= \rho_1\mu + \mu - \rho_1\mu \\ &= \mu. \end{aligned}$$

Next at t = 3, we find that

$$
\begin{aligned}
E(y_3) =& E\Big[\sum_{j=1}^{y_2} b_{1j}(\rho_1) + \sum_{j=1}^{y_1} b_{2j}(\rho_2) + d_3\Big] \\
&= E\Big[\sum_{j=1}^{y_2} b_{1j}(\rho_1)\Big] + E\Big[\sum_{j=1}^{y_1} b_{2j}(\rho_2)\Big] + E\Big[d_3\Big] \\
&= E_{y_2} E\Big[\sum_{j=1}^{y_2} b_{1j}(\rho_1)\Big| y_2\Big] + E_{y_1} E\Big[\sum_{j=1}^{y_1} b_{2j}(\rho_2)\Big| y_1\Big] + E\Big[d_3\Big] \\
&= E_{y_2}\Big[y_2\rho_1\Big] + E_{y_1}\Big[y_1\rho_2\Big] + E\Big[d_3\Big] \\
&= \rho_1\mu + \rho_2\mu + \mu - \rho_1\mu - \rho_2\mu \\
&= \mu.
\end{aligned}
$$

In general, we can show that the mean of y_t for t = 4,5,...,T is

$$
\begin{aligned}
E(y_t) &= E\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1) + \sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2) + \sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3) + d_t\Big] \\
&= E\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1)\Big] + E\Big[\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big] + E\Big[\sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3)\Big] + E\Big[d_t\Big] \\
&= E_{y_{t-1}} E\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1)\Big| y_{t-1}\Big] + E_{y_{t-2}} E\Big[\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big| y_{t-2}\Big] \\
&\quad + E_{y_{t-3}} E\Big[\sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3)\Big| y_{t-3}\Big] + E\Big[d_t\Big] \\
&= \rho_1 E[y_{t-1}] + \rho_2 E[y_{t-2}] + \rho_3 E[y_{t-3}] + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu.
\end{aligned}
$$

Now, by taking successive expectation we find that $E(y_t) = \mu$ for all $t = 1, 2, ..., T$

Variance

As we discussed in Section 2.2.2, since $b_{ij}(\rho_i)$ is a sequence of independent identically distributed (i.i.d) binary random variables with $Pr[b_{ij}(\rho_i) = 1] = \rho_i$ and $Pr[b_{ij}(\rho_i) = 0] = 1 - \rho_i \quad$ for all $\quad i = 1, 2, 3$, the following covariance terms will vanish when we take variance with respect to the response y_t.

$Cov\Big(\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1), \sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big) = Cov\Big(\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1), \sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3)\Big)$

$= Cov\Big(\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2), \sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3)\Big) = 0.$

We also note that from assumptions 4, 5 and 6 of model (2.24),

$$Cov\Big(\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1), d_t\Big) = Cov\Big(\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2), d_t\Big) = Cov\Big(\sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3), d_t\Big) = 0.$$

Now, from the AR(3) model (2.24), $\sigma_{1,1} = Var(y_1) = \mu$. Then, at t = 2, we find that,

$$\begin{aligned}
\sigma_{2,2} &= Var(y_2) = Var\Big[\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big] \\
&= E_{y_1}\Big[Var\Big(\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big|y_1\Big)\Big] + Var_{y_1}\Big[E\Big(\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big|y_1\Big)\Big] \\
&= E_{y_1}\Big[y_1\rho_1(1-\rho_1) + \mu - \rho_1\mu\Big] + Var_{y_1}\Big[y_1\rho_1 + \mu - \rho_1\mu\Big] \\
&= \mu.
\end{aligned}$$

Similarly we find the $Var(y_t)$ at $t = 3$,

$$\begin{aligned}
\sigma_{3,3} &= Var(y_3) \\
&= Var\Big[\sum_{j=1}^{y_2} b_{1j}(\rho_1) + \sum_{j=1}^{y_1} b_{2j}(\rho_2) + d_3\Big] \\
&= Var\Big[\sum_{j=1}^{y_2} b_{1j}(\rho_1)\Big] + Var\Big[\sum_{j=1}^{y_1} b_{2j}(\rho_2)\Big] + Var\Big[d_3\Big] \\
&= E_{y_2}\Big[Var\Big(\sum_{j=1}^{y_2} b_{1j}(\rho_1)\Big|y_2\Big)\Big] + Var_{y_2}\Big[E\Big(\sum_{j=1}^{y_2} b_{1j}(\rho_1)\Big|y_2\Big)\Big] \\
&\quad + E_{y_1}\Big[Var\Big(\sum_{j=1}^{y_1} b_{2j}(\rho_2)\Big|y_1\Big)\Big] + Var_{y_1}\Big[E\Big(\sum_{j=1}^{y_1} b_{2j}(\rho_2)\Big|y_1\Big)\Big] + Var\Big[d_3\Big] \\
&= E_{y_2}\Big[y_2\rho_1(1-\rho_1)\Big] + Var_{y_2}\Big[y_2\rho_1\Big] \\
&\quad + E_{y_1}\Big[y_1\rho_2(1-\rho_2)\Big] + Var_{y_1}\Big[y_1\rho_2\Big] + Var\Big[d_3\Big] \\
&= \rho_1(1-\rho_1)E_{y_1}\Big[y_1\Big] + \mu - \rho_1\mu + \rho_1^2 Var_{y_1}\Big[y_1\Big] \\
&= \rho_1(1-\rho_1)\mu + \rho_1^2\mu + \rho_2(1-\rho_2)\mu + \rho_2^2\mu + \mu - \rho_1\mu - \rho_2\mu \\
&= \mu.
\end{aligned}$$

In general, to obtain the $Var(y_t)$ for $t = 4, 5, ..., T$, we first find

$$\begin{aligned}
Var\Big(y_t\Big|y_{t-1}, y_{t-2}, y_{t-3}\Big) &= Var\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1) + \sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2) + \sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3) + d_t\Big|y_{t-1}, y_{t-2}, y_{t-3}\Big] \\
&= Var\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1)\Big|y_{t-1}\Big] + Var\Big[\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big|y_{t-2}\Big] \\
&\quad + Var\Big[\sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3)\Big|y_{t-3}\Big] + Var\Big[d_t\Big] \\
&= y_{t-1}\rho_1(1-\rho_1) + y_{t-2}\rho_2(1-\rho_2) + y_{t-3}\rho_3(1-\rho_3) \\
&\quad + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu.
\end{aligned}$$

Then, using the conditional variance property we find

$$Var\Big(y_t\Big|y_{t-2}, y_{t-3}\Big)$$

$$= E_{y_{t-1}}\left[Var\Big(y_t\Big|y_{t-1}, y_{t-2}, y_{t-3}\Big)\Big|y_{t-2}, y_{t-3}\right] + Var_{y_{t-1}}\left[E\Big(y_t\Big|y_{t-1}, y_{t-2}, y_{t-3}\Big)\Big|y_{t-2}, y_{t-3}\right]$$

$$= E_{y_{t-1}}\left[y_{t-1}\rho_1(1-\rho_1) + y_{t-2}\rho_2(1-\rho_2) + y_{t-3}\rho_3(1-\rho_3) + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu\Big|y_{t-2}, y_{t-3}\right]$$

$$+ Var_{y_{t-1}}\left[y_{t-1}\rho_1 + y_{t-2}\rho_2 + y_{t-3}\rho_3 + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu\Big|y_{t-2}, y_{t-3}\right]$$

$$= \mu\rho_1(1-\rho_1) + y_{t-2}\rho_2(1-\rho_2) + y_{t-3}\rho_3(1-\rho_3) + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu + Var_{y_{t-1}}\left[y_{t-1}\rho_1\right] + 0$$

$$= \rho_1(1-\rho_1)\mu + y_{t-2}\rho_2(1-\rho_2) + y_{t-3}\rho_3(1-\rho_3) + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu + \rho_1^2\sigma_{t-1,t-1}.$$

Again, using the conditional variance property we find

$$Var\Big(y_t\Big|y_{t-3}\Big) = E_{y_{t-2}}\left[Var\Big(y_t\Big|y_{t-2}, y_{t-3}\Big)\Big|y_{t-3}\right] + Var_{y_{t-2}}\left[E\Big(y_t\Big|y_{t-2}, y_{t-3}\Big)\Big|y_{t-3}\right]$$

$$= E_{y_{t-2}}\Big[\rho_1(1-\rho_1)\mu + y_{t-2}\rho_2(1-\rho_2) + y_{t-3}\rho_3(1-\rho_3) + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu$$

$$+ \rho_1^2\sigma_{t-1,t-1}\Big|y_{t-3}\Big]$$

$$+ Var_{y_{t-2}}\left[\rho_1\mu + y_{t-2}\rho_2 + y_{t-3}\rho_3 + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu\Big|y_{t-3}\right]$$

$$= \rho_1(1-\rho_1)\mu + \rho_2(1-\rho_2)\mu + y_{t-3}\rho_3(1-\rho_3) + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu + \rho_1^2\sigma_{t-1,t-1}$$

$$+ 0 + Var_{y_{t-2}}\left[y_{t-2}\rho_2\right] + 0$$

$$= \rho_1(1-\rho_1)\mu + \rho_2(1-\rho_2)\mu + y_{t-3}\rho_3(1-\rho_3) + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu$$

$$+ \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2}.$$

Next, by using the property that

$$Var(y_t) = E_{y_{t-3}}\Big[Var\Big(y_t\Big|y_{t-3}\Big)\Big] + Var_{y_{t-3}}\Big[E\Big(y_t\Big|y_{t-3}\Big)\Big]$$

we can obtain a recursive relation variance formula for t = 4,5,...,T.

$$\begin{aligned}
\sigma_{t,t} &= Var(y_t) \\
&= E_{y_{t-3}}\Big[Var\Big(y_t\Big|y_{t-3}\Big)\Big] + Var_{y_{t-3}}\Big[E\Big(y_t\Big|y_{t-3}\Big)\Big] \\
&= E_{y_{t-3}}\Big[\rho_1(1-\rho_1)\mu + \rho_2(1-\rho_2)\mu + y_{t-3}\rho_3(1-\rho_3) + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu \\
&\quad + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2}\Big] \\
&\quad + Var_{y_{t-3}}\Big[\rho_1\mu + \rho_2\mu + y_{t-3}\rho_3 + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu\Big] \\
&= \rho_1(1-\rho_1)\mu + \rho_2(1-\rho_2)\mu + \rho_3(1-\rho_3)\mu + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu \\
&\quad + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2} + 0 + 0 + Var_{y_{t-3}}\Big[y_{t-3}\rho_3\Big] + 0 \\
&= \rho_1(1-\rho_1)\mu + \rho_2(1-\rho_2)\mu + \rho_3(1-\rho_3)\mu + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu \\
&\quad + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2} + \rho_3^2\sigma_{t-3,t-3} \\
&= (\mu - \rho_1^2\mu - \rho_2^2\mu - \rho_3^2\mu) + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2} + \rho_3^2\sigma_{t-3,t-3}.
\end{aligned}$$

Hence, the recursive relation formula for $\sigma_{tt} = Var(y_t)$ is

$Var(y_t) = (\mu - \rho_1^2\mu - \rho_2^2\mu - \rho_3^2\mu) + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2} + \rho_3^2\sigma_{t-3,t-3} \quad for \quad t = 4, 5, ..., T.$

Now, by taking variance successively we can see that $\sigma_{t,t} = Var(y_t) = \mu = exp(\boldsymbol{x}_t'\boldsymbol{\beta})$ for all $t = 1, 2, ..., T$.

Co-variance and Correlation

For t=2, the auto-covariance and auto-correlation can be obtained from AR(1) process as,

$$\begin{aligned} \sigma_{2,1} &= Cov(y_2, y_1) = \rho_1\mu, \\ \gamma(1) &= Corr(y_2, y_1) = \rho_1. \end{aligned} \tag{2.28}$$

At $t = 3$, from the properties of AR(2) Poisson process we derived in Section 2.2.2,

$$\begin{aligned} \sigma_{3,2} &= Cov(y_3, y_2) = \rho_1\mu + \rho_1\rho_2\mu, \\ \gamma(1) &= Corr(y_3, y_2) = \rho_1 + \rho_1\rho_2. \\ \sigma_{3,1} &= Cov(y_3, y_1) = \rho_1^2\mu + \rho_2\mu, \\ \gamma(2) &= Corr(y_3, y_1) = \rho_1^2 + \rho_2. \end{aligned} \tag{2.29}$$

Next, for $t = 4, 5, ..., T$

$$\begin{aligned} \sigma_{t,t-k} &= Cov(y_t, y_{t-k}) \\ &= E\Big(y_t, y_{t-k}\Big) - E\Big(y_t\Big)E\Big(y_{t-k}\Big) = E_{y_{t-k}}\Big[E\Big(y_t y_{t-k}\Big|y_{t-k}\Big)\Big] - \mu\mu \\ &= E_{y_{t-k}}\Big[y_{t-k}\Big(y_{t-1}\rho_1 + y_{t-2}\rho_2 + y_{t-3}\rho_3 + \mu - \rho_1\mu - \rho_2\mu - \rho_3\mu\Big)\Big] - \mu\mu \\ &= \rho_1\Big[E\Big(y_{t-1}y_{t-k}\Big) - \mu\mu\Big] + \rho_2\Big[E\Big(y_{t-2}y_{t-k}\Big) - \mu\mu\Big] + \rho_3\Big[E\Big(y_{t-3}y_{t-k}\Big) - \mu\mu\Big] \\ &= \rho_1 Cov(y_{t-1}, y_{t-k}) + \rho_2 Cov(y_{t-2}, y_{t-k}) + \rho_3 Cov(y_{t-3}, y_{t-k}) \\ &= \rho_1\sigma_{t-1,t-k} + \rho_2\sigma_{t-2,t-k} + \rho_3\sigma_{t-3,t-k}. \end{aligned} \tag{2.30}$$

Hence, for $t = 4, 5, ..., T$, we found that the covariances can be obtained from the recursive relation $\sigma_{t,t-k} = Cov(Y_t, Y_{t-k}) = \rho_1\sigma_{t-1,t-k} + \rho_2\sigma_{t-2,t-k} + \rho_3\sigma_{t-3,t-k}$. It follows that the lag k auto-correlation is given by,

$$\begin{aligned} Corr(y_t, y_{t-k}) = \gamma(k) &= \frac{Cov(y_t, y_{t-k})}{\sqrt{Var(y_t)}\sqrt{Var(y_{t-k})}} \\ &= \frac{Cov(y_t, y_{t-k})}{\sqrt{\mu}\sqrt{\mu}} \\ &= \frac{\rho_1\sigma_{t-1,t-k} + \rho_2\sigma_{t-2,t-k} + \rho_3\sigma_{t-3,t-k}}{\sqrt{\mu\mu}} \\ &= \rho_1 Corr(y_{t-1}, y_{t-k}) + \rho_2 Corr(y_{t-2}, y_{t-k}) + \rho_3 Corr(y_{t-3}, y_{t-k}) \\ &= \rho_1\gamma(k-1) + \rho_2\gamma(k-2) + \rho_3\gamma(k-3), \quad t = 4, 5, ..., T. \end{aligned} \tag{2.31}$$

Now, we find the form of PACF at each lag k, denoted by ϕ_{kk} using the ACF function $\gamma(k)$ given in equation (2.28).

Partial Autocorrelation

Now, at $k = 1$, we find that the PACF at lag 1, denoted by ϕ_{11} is given by,

$$\phi_{11} = \gamma(1) = \frac{\rho_1\rho_2}{1 - \rho_2 - \rho_1(\rho_1 + \rho_3)}.$$

At $k = 2$, we find

$$\phi_{22} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(2) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{vmatrix}} = \frac{\gamma(2) - \gamma^2(1)}{1 - \gamma^2(1)} = \frac{\dfrac{\rho_2(1-\rho_2)}{1-\rho_2-\rho_1(\rho_1+\rho_3)} - \dfrac{\rho_1^2\rho_2^2}{(1-\rho_2-\rho_1(\rho_1+\rho_3))^2}}{1 - \dfrac{\rho_1^2\rho_2^2}{(1-\rho_2-\rho_1(\rho_1+\rho_3))^2}}$$

$$= \frac{\rho_2(1-\rho_2)(1-\rho_2-\rho_1(\rho_1+\rho_3))^2 - \rho_1^2\rho_2^2}{(1-\rho_2-\rho_1(\rho_1+\rho_3))^2 - \rho_1^2\rho_2^2}.$$

Similarly at $k = 3$,

$$\phi_{33} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(1) \\ \gamma(1) & \gamma(0) & \gamma(2) \\ \gamma(2) & \gamma(1) & \gamma(3) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(2) & \gamma(1) & \gamma(0) \end{vmatrix}},$$

where $\gamma(0) = 1, \gamma(1) = \dfrac{\rho_1\rho_2}{1-\rho_2-\rho_1(\rho_1+\rho_3)}, \gamma(2) = \dfrac{\rho_2(1-\rho_2)}{1-\rho_2-\rho_1(\rho_1+\rho_3)}$ and $\gamma(3) = \dfrac{\rho_1\rho_2}{1-\rho_2-\rho_1(\rho_1+\rho_3)} + \rho_3$.

Similarly we find PACF at $k = 4$,

$$\phi_{44} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \gamma(1) \\ \gamma(1) & \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(2) & \gamma(1) & \gamma(0) & \gamma(3) \\ \gamma(3) & \gamma(2) & \gamma(1) & \gamma(4) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \gamma(3) \\ \gamma(1) & \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(2) & \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(3) & \gamma(2) & \gamma(1) & \gamma(0) \end{vmatrix}} = \frac{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \rho_1\gamma(0) + \rho_2\gamma(1) + \rho_3\gamma(2) \\ \gamma(1) & \gamma(0) & \gamma(1) & \rho_1\gamma(1) + \rho_2\gamma(0) + \rho_3\gamma(1) \\ \gamma(2) & \gamma(1) & \gamma(0) & \rho_1\gamma(2) + \rho_2\gamma(1) + \rho_3\gamma(0) \\ \gamma(3) & \gamma(2) & \gamma(1) & \rho_1\gamma(3) + \rho_2\gamma(2) + \rho_3\gamma(1) \end{vmatrix}}{\begin{vmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \gamma(3) \\ \gamma(1) & \gamma(0) & \gamma(1) & \gamma(2) \\ \gamma(2) & \gamma(1) & \gamma(0) & \gamma(1) \\ \gamma(3) & \gamma(2) & \gamma(1) & \gamma(0) \end{vmatrix}} = 0,$$

since 4^{th} column in the numerator, is a linear combination of first three columns. Similarly, we can find PACF at any lag k. Now, we write a general expression for PACF at lag k,

$$\phi_{kk} = \begin{cases} \phi_{11}, & k = 1 \\ \phi_{22}, & k = 2 \\ \phi_{33}, & k = 3 \\ 0, & k > 3. \end{cases} \tag{2.32}$$

The basic properties of the AR(3) process obtained in this section show that the ACF decays exponentially while the PACF cuts off after lag 3.

2.3 Simulation Study

In this section we examine the validity of the theoretical patterns we derived in Sections 2.1 and 2.2 for order identification purposes.

2.3.1 General Simulation Set up

We set the number of responses $T = 250$ for all the MA and AR models. We considered the number of time independent co-variates $p = 2$ which we simulated from Uniform distribution $U(-1, 1)$ with fixed coefficients $\boldsymbol{\beta}' = (0.5, 1)$. The mean μ_t of all the processes was stationary since the co-variates are time independent where the mean is $\mu_t = exp(\boldsymbol{x}_t'\boldsymbol{\beta})$ which is a constant value at all time points. Using the defined co-variate vector and fixed combination of parameter $\boldsymbol{\beta}$ values, we compute the mean μ and choose the scale parameters (ρ's) to satisfy the conditions $0 < \rho_1 < 1$ and $\rho_1 \geq \rho_2 \geq \rho_3$. It is reasonable to assume that $\rho_1 \geq \rho_2 \geq \rho_3$, since it is more likely for the correlation between responses 1 distance apart to be higher than the correlation between the responses 2 distances apart which is more than the correlation between the responses 3 distances apart.

We generate d_t from Poisson distribution with corresponding mean of each model. Using that, we generate all the y_t for $t = 1, 2, ..., 250$ from the model equations and computed SACF (Sample Auto-Correlation Function) and SPACF (Sample Partial Auto-Correlation Function) at each lag. We repeated the same procedure 1,000 times for the same values of $\boldsymbol{\beta}$ and ρ's and computed the average values of SACF and SPACF at each lag. Using the simulated mean values of SACF and SPACF, we construct SACF plot and SPACF plot. The standard error of SACF and SPACF will be shown by dashed lines in SACF and SPACF plots. This procedure was repeated for different combination of ρ values.

For continuous time series processes to test whether the hypothesis

$H_0 : \gamma(1) = \gamma(2) = ... = \gamma(k) = 0,$

it has been shown that the 95% asymptotic confidence band for ACF at lag k ($\gamma(k)$) is approximately $r_k \pm \frac{2}{\sqrt{T}}$, where T is the total number of time points in the given time series data. Under H_0, standard error of r_k equals $S_{r_k} \approx \sqrt{\frac{1}{T}}$. To find the AR and MA order for poisson process we need a procedure that help us to determine when r_k is not significantly different from zero. Since standard error of $r_k = S_{r_k}$ is unknown for count time series we use the Jackknife method to estimate S_{r_k}. Our results will show that the standard error S_{r_k} for count time series is about the same as that of continuous Gaussian time series.

2.3.2 Jackknife Estimation of the Standard Error of Sample ACF for Stationary Poisson Process

We first define r_k (Boslaugh, 2012) as

$$r_k = \frac{\sum_{t=1}^{T-k}(y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^{T-k}(y_t - \bar{y})^2},$$

where, y_t = response at time t and $\bar{y}$ = mean of the time series.

Let $Z_T = Z(y_1, y_2, ..., y_t, ..., y_T)$ be an estimator of SACF which is equal to r_k and $Z_{(-i)}$ denote the statistic with $(y_i - \bar{y})(y_{i+k} - \bar{y})$ removed. For example,

$$Z_{(-2)} = \frac{(y_1 - \bar{y})(y_{1+k} - \bar{y}) + \sum_{t=3}^{T-k}(y_t - \bar{y})(y_{t+k} - \bar{y})}{(y_1 - \bar{y})^2 + \sum_{t=3}^{T-k}(y_t - \bar{y})^2}.$$

Then according to Wasserman (2006), the variance of the jackknife estimator i.e. $Var(Z_T) = Var(r_k)$ can be written as

$$\nu_{jack} = \frac{\hat{s}^2}{T-k}$$

where,

$$\hat{s}^2 = \frac{\sum_{i=1}^{T-k}(\tilde{Z}_i - \frac{1}{T-k}\sum_{i=1}^{T-k}\tilde{Z}_i)^2}{T-k-1} \qquad \text{and} \qquad \tilde{Z}_i = (T-k)Z_T - (T-k-1)Z_{(-i)}.$$

Thus, to test

$H_0 : \gamma(1) = \gamma(2) = ... = \gamma(k) = 0,$

we use the test statistic $Z = \dfrac{r_k}{\sqrt{\nu_{jack}}}$ and we reject H_0 at the significance level α if $|Z| > Z_{\alpha/2}$. That is if $\left|\dfrac{r_k}{\sqrt{\nu_{jack}}}\right| > Z_{\alpha/2}$ or $|r_k| > Z_{\alpha/2}\sqrt{\nu_{jack}}$.

2.3.3 Simulation Study of Stationary Poisson MA(1) Process

We generated d_t from Poisson distribution with mean $(\dfrac{\mu}{1+\rho})$; where $0 < \rho < 1$ and computed mean μ as discussed in Section 2.3.1.

Then, we generated y_t, $t = 1, 2, ..., 250$ from the model (2.1). We computed the SACF and SPACF for each of 1000 simulations. Using the simulated mean values of SACF and SPACF, we constructed plots of SACF and SPACF. Thus the procedure was repeated for different values of ρ which are 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9.

In order to test the significance of the SACF values, we first compare the jackknife standard error of SACF to the simulated standard error of SACF values obtained from 10000 simulations for some specific ρ values. Figures 2.1 and 2.2 show that the

Jackknife estimator of $s.e(r_k)$ for MA(1) Poisson process is approximately equal to the simulated $s.e(r_k)$ of the MA(1) Poisson process. We note that Jackknife estimator is also approximately equal to $1/\sqrt{T}$ which is the approximate standard error of r_k of Gaussian time series.

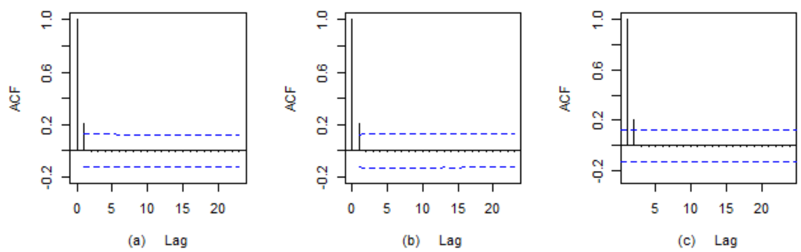

Figure 2.1: A plot of average SACF values obtained from 10,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.3$ and (a) Jackknife s.e of SACF (b) simulated s.e of SACF (c) $s.e(r_k)$ of MA(1) Gaussian process $\approx 1/\sqrt{T}$.

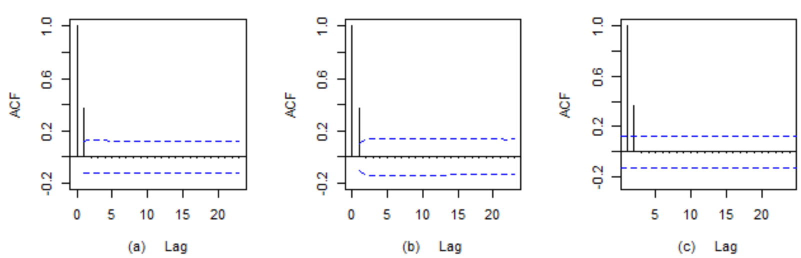

Figure 2.2: A plot of average SACF values obtained from 10,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.7$ and (a) Jackknife s.e of SACF (b) simulated s.e of SACF (c) $s.e(r_k)$ of MA(1) Gaussian process $\approx 1/\sqrt{T}$.

The standard error values of r_k at each lag are shown in Tables 2.1 and 2.2 for both Jackknife and simulation methods. For example, if we compare the Jackknife s.e of (0.06320) and simulated s.e of (0.05958) at lag 1 from Table 2.1 with the $s.e(r_k)$ of MA(1) Gaussian process $= \dfrac{1}{\sqrt{T}} = 0.06325$, it is obvious that these three values are approximately the same. Since these three methods give approximately equal results at each lag, leading to same conclusions, it is recommended to use Gaussian process approach to estimate the standard error which is $\approx \pm\dfrac{1}{\sqrt{T}}$.

Table 2.1: Standard error of SACF obtained from 10,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.3$ obtained from two different methods.

Lag	**Jackknife s.e**	**Simulated s.e**	**Lag**	**Jackknife s.e**	**Simulated s.e**
1	0.06320	0.05958	13	0.06135	0.06342
2	0.06269	0.06573	14	0.06124	0.06424
3	0.06261	0.06533	15	0.06103	0.06412
4	0.06252	0.06496	16	0.06087	0.06320
5	0.06242	0.06536	17	0.06085	0.06292
6	0.06222	0.06528	18	0.06065	0.06299
7	0.06204	0.06476	19	0.06054	0.06305
8	0.06187	0.06482	20	0.06039	0.06273
9	0.06183	0.06514	21	0.06018	0.06350
10	0.06171	0.06404	22	0.06013	0.06254
11	0.06156	0.06437	23	0.05988	0.06301
12	0.06145	0.06427			

Table 2.2: Standard error of SACF obtained from 10,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.7$ obtained from two different methods.

Lag	**Jackknife s.e**	**Simulated s.e**	**Lag**	**Jackknife s.e**	**Simulated s.e**
1	0.06150	0.05178	13	0.06120	0.06944
2	0.06254	0.07019	14	0.06099	0.06874
3	0.06242	0.07114	15	0.06089	0.06865
4	0.06237	0.07039	16	0.06080	0.06818
5	0.06216	0.06979	17	0.06071	0.06809
6	0.06204	0.07100	18	0.06053	0.06840
7	0.06191	0.07026	19	0.06038	0.06827
8	0.06185	0.07026	20	0.06030	0.06812
9	0.06174	0.07065	21	0.06008	0.06834
10	0.06153	0.06909	22	0.06001	0.06721
11	0.06144	0.06937	23	0.05983	0.06742
12	0.06137	0.06918			

The auto-correlation function is commonly used tool for model identification, specifically, in identifying the order of a moving average sequence of continuous data. From equation (2.4), we can see that theoretical ACF of a Stationary Poisson MA(1) process is non zero at lag 1 and zero at lag 2 and high order.

Now, we will discuss the patterns in ACF and PACF plots for Stationary Poisson MA(1) processes. Using those patterns and behaviors we propose a method for identifying the order of Poisson MA models for Time series of count data.

We have obtained the following SACF and SPACF plots of MA(1) stationary Poisson process through the simulation procedure we discussed earlier, for different ρ values satisfying the condition $0 < \rho < 1$.

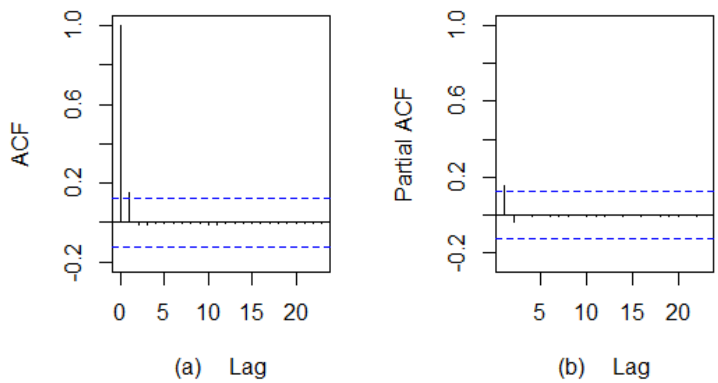

Figure 2.3: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.2$.

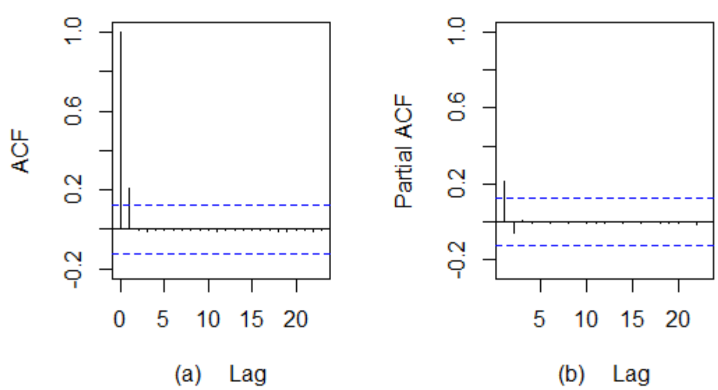

Figure 2.4: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.3$.

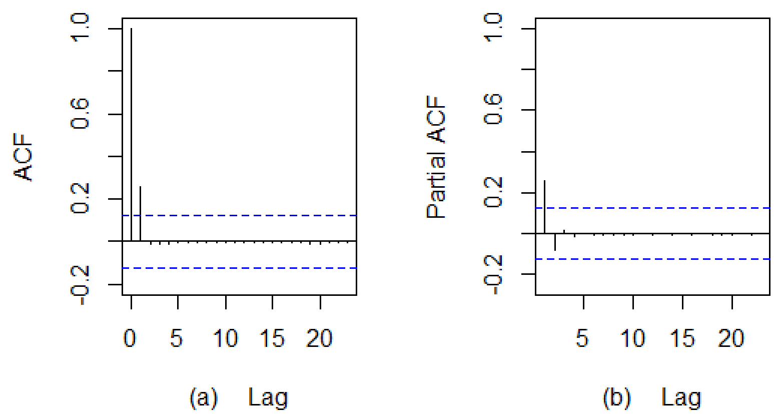

Figure 2.5: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.4$.

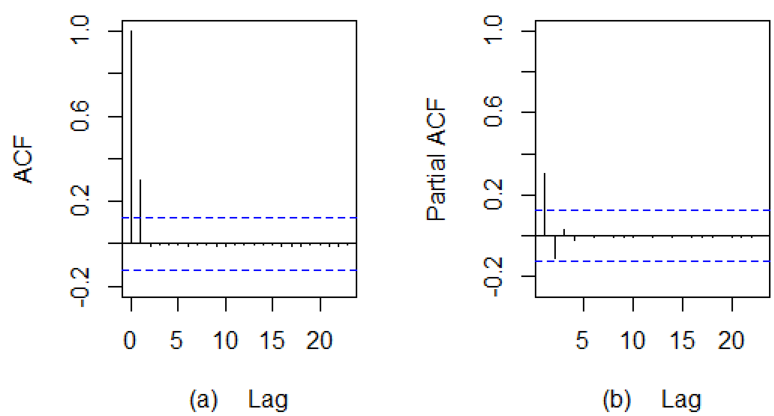

Figure 2.6: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.5$.

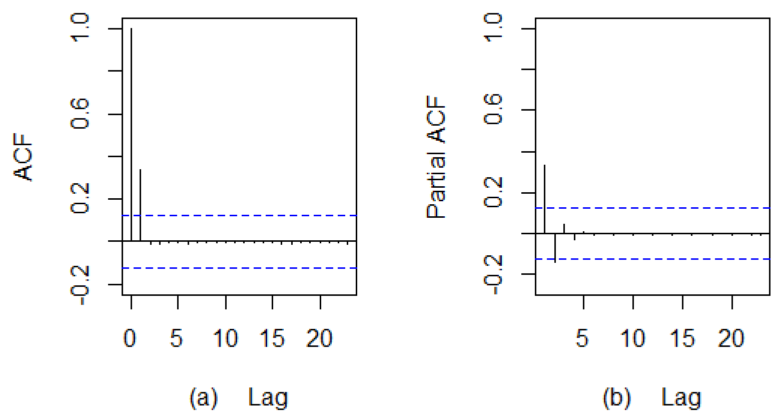

Figure 2.7: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.6$.

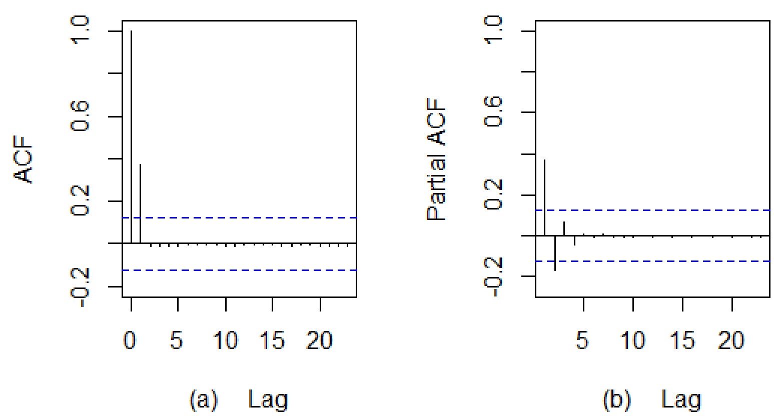

Figure 2.8: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.7$.

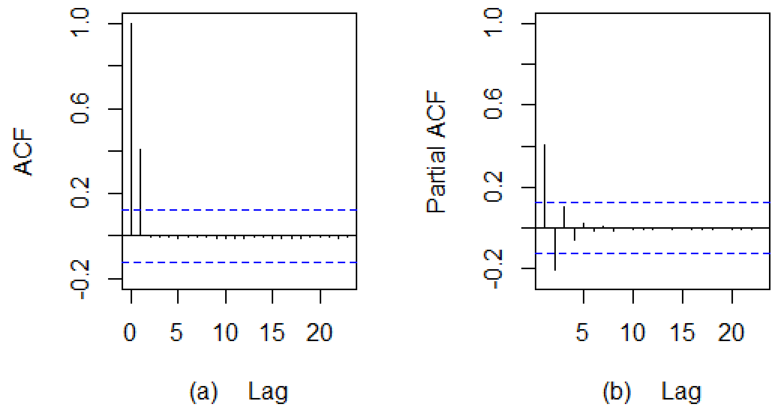

Figure 2.9: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.8$.

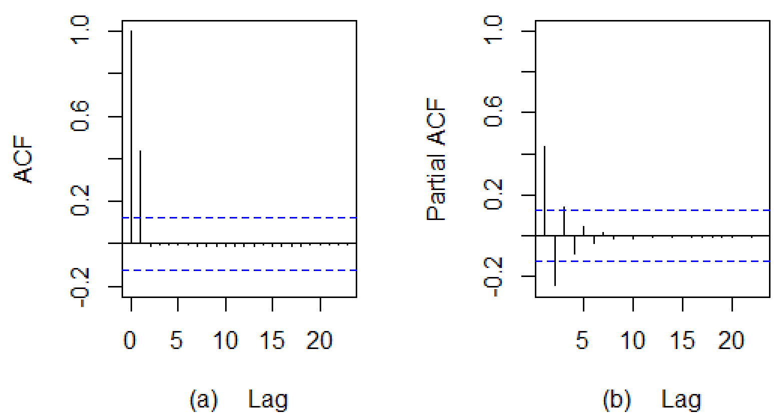

Figure 2.10: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(1) process with $\rho = 0.9$.

The SACF plots in Figures 2.3 to 2.10 show that the SACF at lag 1 is significant and it increases as ρ increases. From all of the SPACF plots in Figures 2.5 to 2.10, we can see that SPACF are decaying exponentially, alternating between positive and negative values. Therefore, we can come to a conclusion that patterns in the ACF and PACF of Stationary Poisson MA(1) processes for count data are similar to that of Stationary MA(1) models for Gaussian processes.

2.3.4 Simulation Study of Stationary Poisson MA(2) Process

For generating data from a stationary Poisson MA(2) process we choose ρ_1 and ρ_2 values satisfying the condition $0 < \rho_2 \leq \rho_1 < 1$. It is reasonable to assume that $\rho_1 \geq \rho_2$, since it is more likely that the correlation between responses 1 distance apart is will be higher than the correlation between the responses 2 distances apart.

Using the computed mean μ as discussed in Section 2.3.1, we generated d_t from Poisson distribution with mean $(\frac{\mu}{1+\rho_1+\rho_2})$ and then we generated y_t for $t = 1, 2, ..., 250$ from model (2.5). We then computed the SACF and SPACF for each of 1000 simulations. Using the simulated mean values of SACF and SPACF, we constructed plots of SACF and SPACF. The procedure was repeated for different values of (ρ_1, ρ_2) including (0.2,0.2), (0.3,0.2), (0.4,0.2), (0.5,0.2), (0.6,0.2), (0.7,0.2), (0.3,0.3), (0.4,0.3), (0.5,0.3), (0.6,0.3), (0.4,0.4) and (0.5,0.4).

In order to test the significance of the SACF values, we first compare the jackknife standard error of SACF with the simulated standard error of SACF values obtained from 10000 simulations for some specific ρ_1 and ρ_2 values. Figures 2.11 and 2.12 show that the Jackknife estimator of $s.e(r_k)$ for MA(2) Poisson process is approximately equal to the simulated $s.e(r_k)$ of the MA(2) Poisson process. We note that Jackknife estimator is also approximately equal to $1/\sqrt{T}$ which is the approximate standard error of r_k of Gaussian time series.

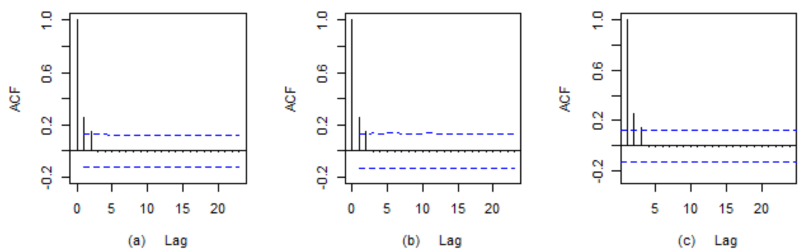

Figure 2.11: A plot of average SACF values obtained from 10,000 simulations with 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.4$ and $\rho_2 = 0.3$ using (a) Jackknife s.e of SACF (b) simulated s.e of SACF (c) $s.e(r_k)$ of MA(2) Gaussian process $\approx 1/\sqrt{T}$.

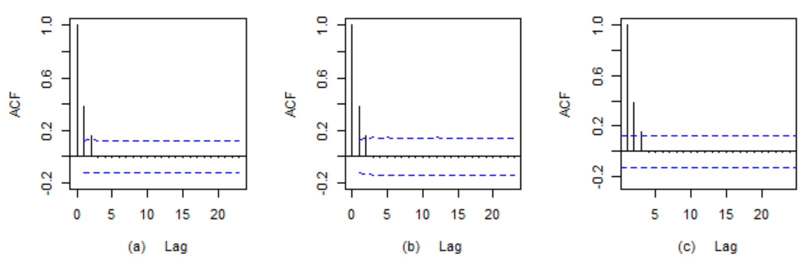

Figure 2.12: A plot of average SACF values obtained from 10,000 simulations with 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.7$ and $\rho_2 = 0.4$ using (a) Jackknife s.e of SACF (b) simulated s.e of SACF (c) $s.e(r_k)$ of MA(2) Gaussian process $\approx 1/\sqrt{T}$.

The standard error values of r_k at each lag are shown in Tables 2.3 and 2.4 for both Jackknife and simulation methods. For example if we compare the Jackknife s.e of (0.06344) and simulated s.e of (0.06432) at lag 2 from Table 2.3 with the $s.e(r_k)$ of

MA(2) Gaussian process $= \dfrac{1}{\sqrt{T}} = 0.06325$, it is obvious that these three values are approximately the same. Since these three methods give approximately equal results at each lag, leading to same conclusions, it is recommended to use Gaussian process approach to estimate the standard error which is $\approx \pm\dfrac{1}{\sqrt{T}}$.

Table 2.3: Standard error of SACF obtained from 10,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.4$ and $\rho_2 = 0.3$ obtained from two different methods.

Lag	**Jackknife s.e**	**Simulated s.e**	**Lag**	**Jackknife s.e**	**Simulated s.e**
1	0.06267	0.06698	13	0.06128	0.06626
2	0.06344	0.06432	14	0.06121	0.06655
3	0.06260	0.06801	15	0.06102	0.06644
4	0.06248	0.06720	16	0.06085	0.06537
5	0.06230	0.06755	17	0.06066	0.06533
6	0.06207	0.06777	18	0.06054	0.06578
7	0.06199	0.06690	19	0.06040	0.06506
8	0.06185	0.06676	20	0.06031	0.06581
9	0.06173	0.06666	21	0.06013	0.06527
10	0.06162	0.06712	22	0.05995	0.06494
11	0.06149	0.06770	23	0.05990	0.06509
12	0.06145	0.06624			

Table 2.4: Standard error of SACF obtained from 10,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.7$ and $\rho_2 = 0.4$ obtained from two different methods.

Lag	**Jackknife s.e**	**Simulated s.e**	**Lag**	**Jackknife s.e**	**Simulated s.e**
1	0.06103	0.05971	13	0.06106	0.07089
2	0.06356	0.06849	14	0.06093	0.07087
3	0.06235	0.07261	15	0.06073	0.07033
4	0.06232	0.07182	16	0.06065	0.06963
5	0.06215	0.07229	17	0.06052	0.06944
6	0.06206	0.07140	18	0.06041	0.06950
7	0.06190	0.07090	19	0.06025	0.06952
8	0.06174	0.07150	20	0.06016	0.06952
9	0.06160	0.07135	21	0.06002	0.06965
10	0.06153	0.07096	22	0.05982	0.06920
11	0.06142	0.07179	23	0.05972	0.06907
12	0.06123	0.07225			

Now we discuss the patterns in SACF and SPACF plots obtained for MA(2) stationary Poisson process from our simulation study and propose a method for identifying the order of MA term for Time series of count data. From equation (2.8), we can see that theoretical ACF of a MA(2) Poisson process is non zero at lags 1 , 2 and zero at lag 3 and high order.

From the simulation procedure discussed earlier, we have obtained the following SACF and SPACF plots of MA(2) stationary Poisson process for different ρ_1 and ρ_2 values satisfying the condition $0 < \rho_2 \leq \rho_1 < 1$.

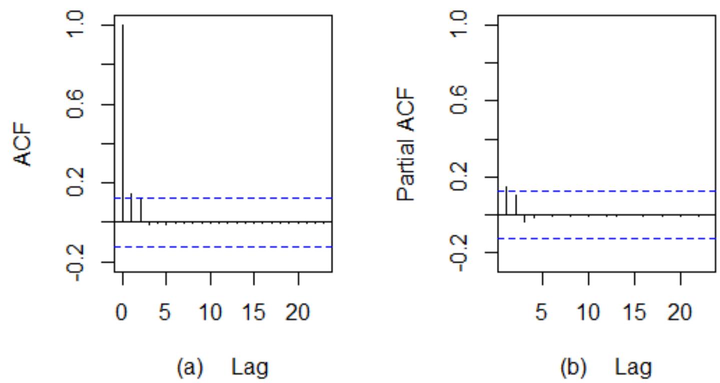

Figure 2.13: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.2$ and $\rho_2 = 0.2$.

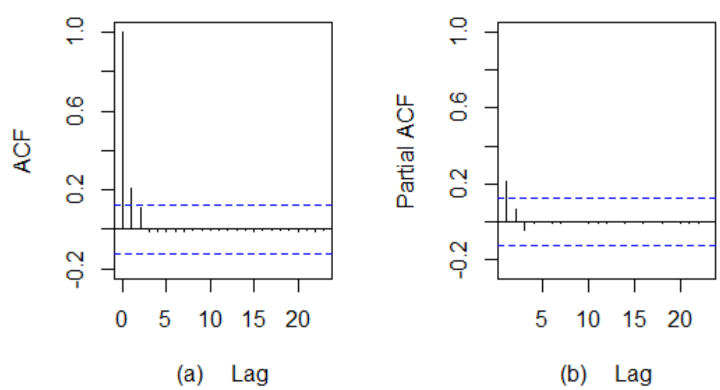

Figure 2.14: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.3$ and $\rho_2 = 0.2$.

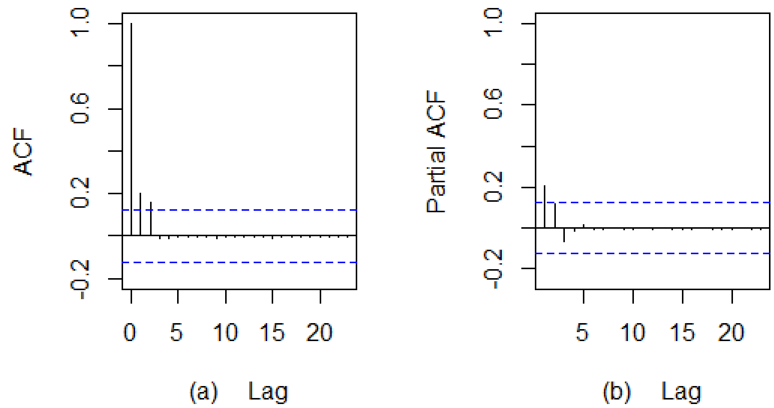

Figure 2.15: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.3$ and $\rho_2 = 0.3$.

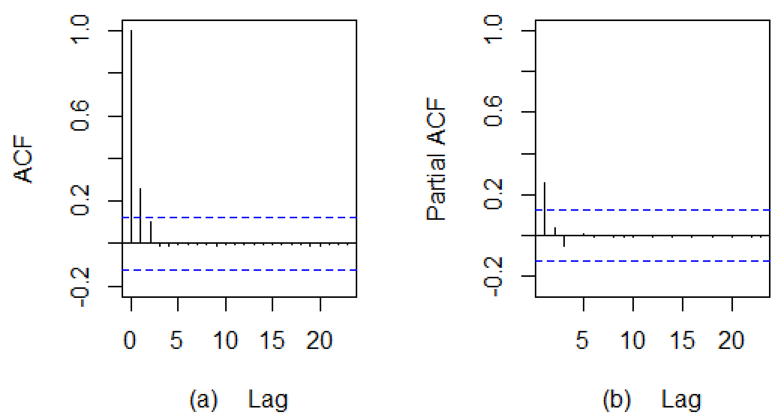

Figure 2.16: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.4$ and $\rho_2 = 0.2$.

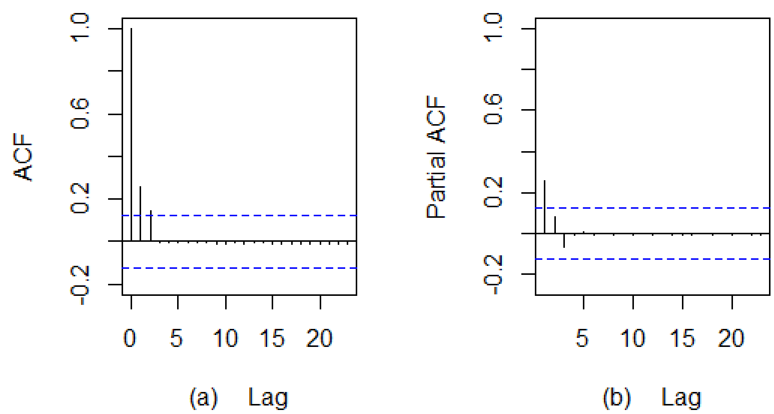

Figure 2.17: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.4$ and $\rho_2 = 0.3$.

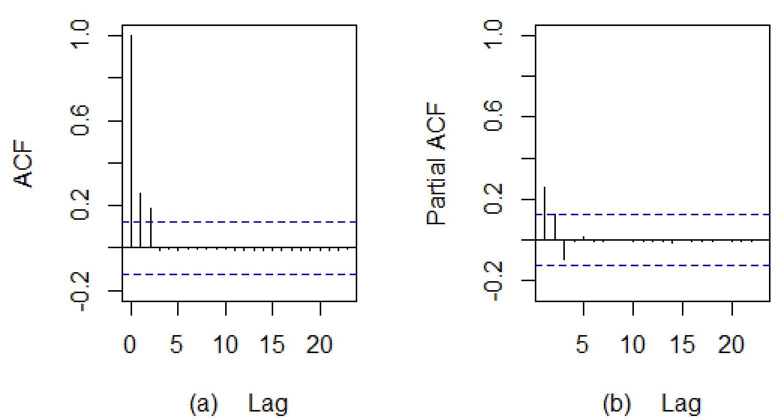

Figure 2.18: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.4$ and $\rho_2 = 0.4$.

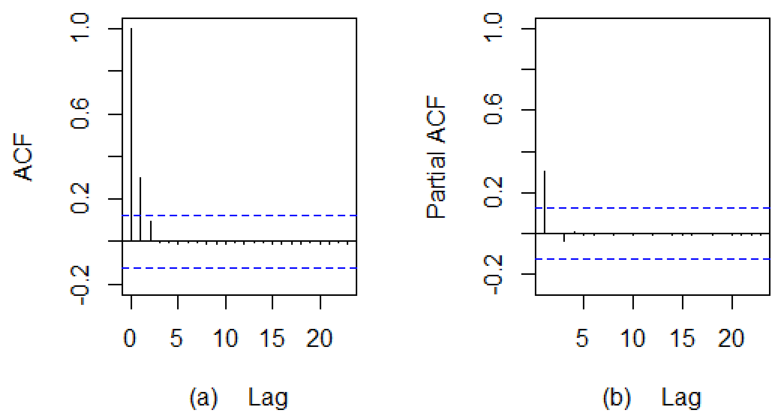

Figure 2.19: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.5$ and $\rho_2 = 0.2$.

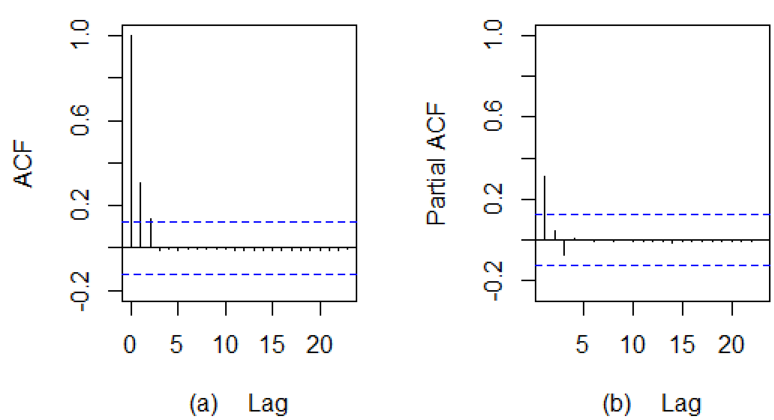

Figure 2.20: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.5$ and $\rho_2 = 0.3$.

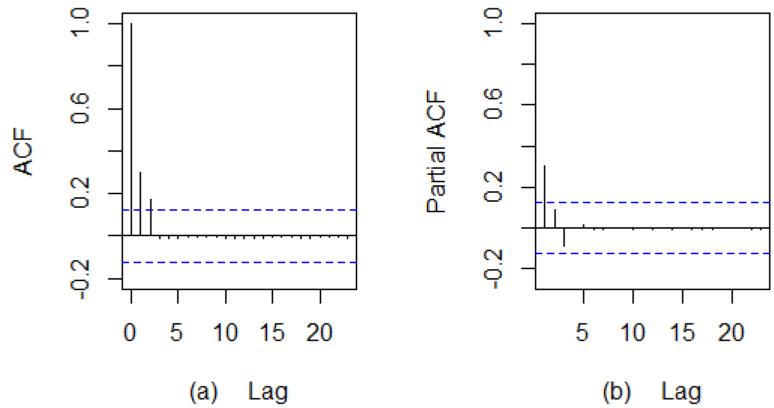

Figure 2.21: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.5$ and $\rho_2 = 0.4$.

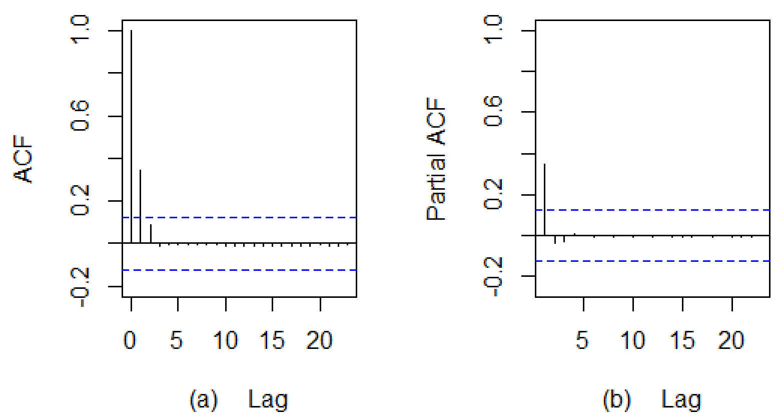

Figure 2.22: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.6$ and $\rho_2 = 0.2$.

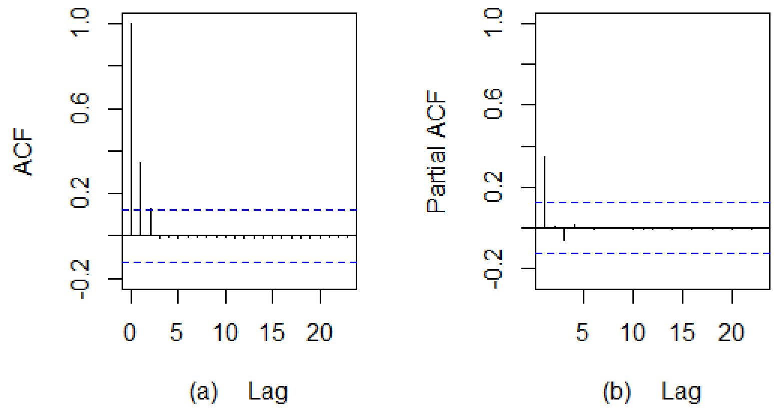

Figure 2.23: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.6$ and $\rho_2 = 0.3$.

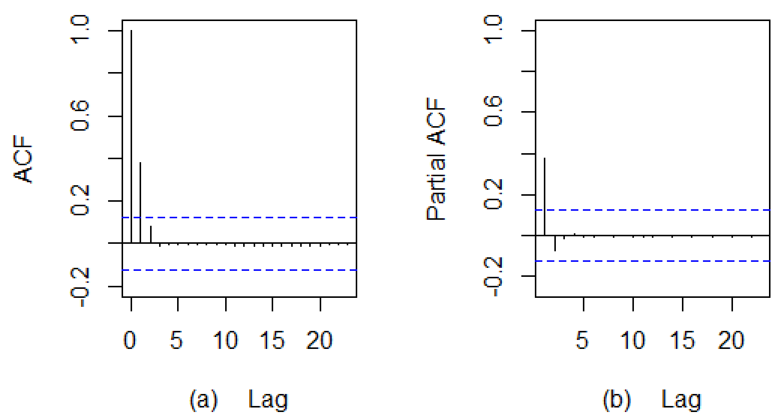

Figure 2.24: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.7$ and $\rho_2 = 0.2$.

Most of the SACF plots in Figures 2.13 to 2.24 show that the SACF at lag 1 and lag 2 are significant and SACF values increases when ρ_1 and ρ_2 increases. In some of the SPACF plots, we can clearly see that SPACF are decaying exponentially, alternating between positive and negative values.

We also considered few combinations of ρ_1 and ρ_2 values where $\rho_1 < \rho_2$. SACF plots for those combinations in Figures 2.25 to 2.27 also show that the SACF at lag 1 and lag 2 are significant and it increases when ρ_1 and ρ_2 increases respectively. All of the SPACF plots in Figures 2.25 to 2.27 clearly show that SPACF are decaying and alternating between positive and negative sides.

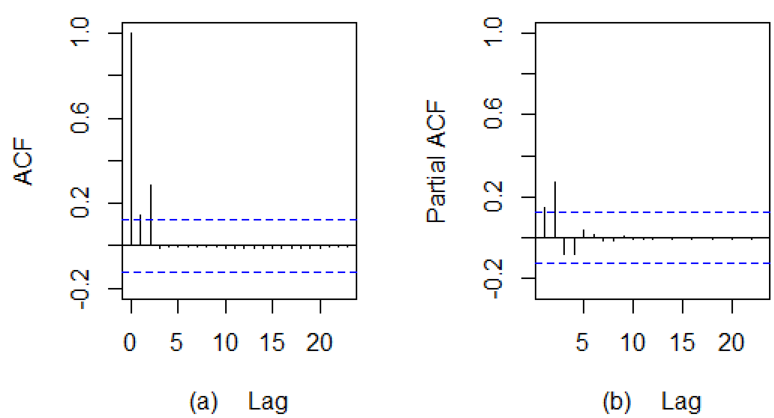

Figure 2.25: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.2$ and $\rho_2 = 0.6$.

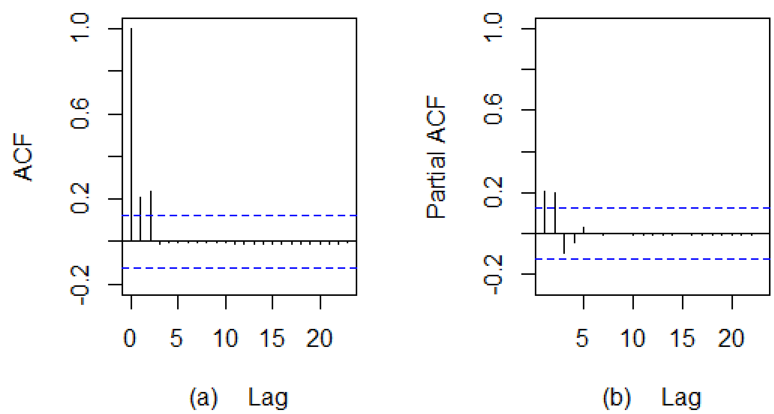

Figure 2.26: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.3$ and $\rho_2 = 0.5$.

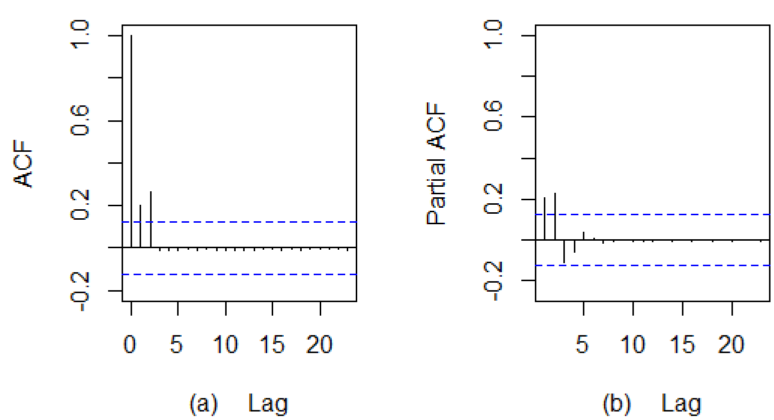

Figure 2.27: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(2) process with $\rho_1 = 0.3$ and $\rho_2 = 0.6$.

Therefore, we reach to the conclusion that patterns in the ACF and PACF of MA(2) stationary Poisson processes for count data are similar to that of MA(2) stationary models for Gaussian processes.

2.3.5 Simulation Study of Stationary Poisson MA(3) Process

First we choose ρ_1, ρ_2 and ρ_3 values satisfying the condition $0 < \rho_3 \leq \rho_2 \leq \rho_1 < 1$. It is reasonable to assume that $\rho_1 \geq \rho_2 \geq \rho_3$, since it is more likely that the correlation between responses 1 distance apart will be higher than the correlation between the responses 2 distances apart which is more than the correlation between the responses 3 distances apart.

Using the computed mean μ in Section 2.3.1, we generate d_t from Poisson distribution with mean $(\frac{\mu}{1+\rho_1+\rho_2+\rho_3})$ and we generated y_t for $t = 1, 2, ..., 250$ from model (2.9).

Then we obtain SACF plot and SPACF plots using the simulated mean values of SACF and SPACF that we found as per the procedure discussed in Section 2.3.1. The above procedure was repeated to check for different values of (ρ_1, ρ_2, ρ_3) including (0.2,0.2,0.2), (0.3,0.2,0.2), (0.3,0.3,0.2), (0.3,0.3,0.3), (0.4,0.2,0.2), (0.4,0.3,0.2), (0.5,0.2,0.2) and (0.35,0.32,0.29).

In order to test the significance of the SACF values, we first compare the jackknife standard error of SACF to the simulated standard error of SACF values obtained from 10000 simulations for some specific ρ_1, ρ_2 and ρ_3 values. Figures 2.28 and 2.29 show that the Jackknife estimator of $s.e(r_k)$ for MA(3) Poisson process is approximately equal to the simulated $s.e(r_k)$ of the MA(3) Poisson process. We note that Jackknife estimator is also approximately equal to $1/\sqrt{T}$ which is the approximate standard error of r_k of Gaussian time series.

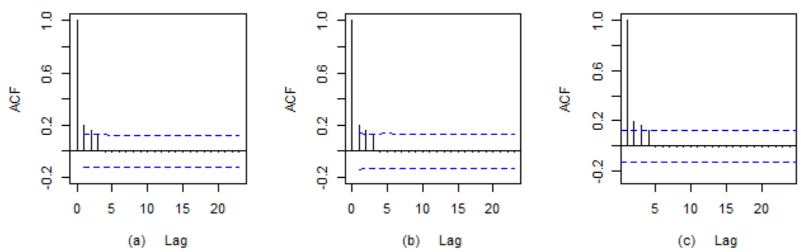

Figure 2.28: A plot of average SACF values obtained from 10,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.3$ and $\rho_3 = 0.3$ using (a) Jackknife s.e of SACF (b) simulated s.e of SACF (c) $s.e(r_k)$ of MA(3) Gaussian process $\approx 1/\sqrt{T}$.

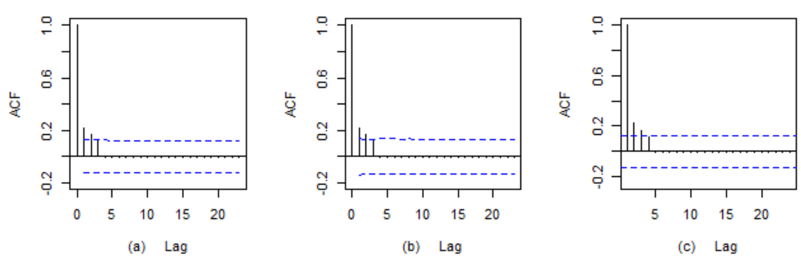

Figure 2.29: A plot of average SACF values obtained from 10,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.35$, $\rho_2 = 0.32$ and $\rho_3 = 0.29$ using (a) Jackknife s.e of SACF (b) simulated s.e of SACF (c) $s.e(r_k)$ of MA(3) Gaussian process $\approx 1/\sqrt{T}$.

The standard error values of r_k at each lag are shown in Tables 2.5 and 2.6 for both Jackknife and simulation methods. For example if we compare the Jackknife s.e of (0.06344) and simulated s.e of (0.06488) at lag 3 from Table 2.5 with the $s.e(r_k)$ of MA(3) Gaussian process $= \dfrac{1}{\sqrt{T}} = 0.06325$, it is obvious that these three values are approximately the same. Since these three methods give approximately equal results at each lag, leading to same conclusions, it is recommended to use Gaussian process approach to estimate the standard error which is $\approx \pm \frac{1}{\sqrt{T}}$.

Table 2.5: Standard error of SACF obtained from 10,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.3$ and $\rho_3 = 0.3$ obtained from two different methods.

Lag	**Jackknife s.e**	**Simulated s.e**	**Lag**	**Jackknife s.e**	**Simulated s.e**
1	0.06318	0.07338	13	0.06122	0.06663
2	0.06338	0.06430	14	0.06107	0.06631
3	0.06344	0.06488	15	0.06094	0.06585
4	0.06241	0.06682	16	0.06086	0.06478
5	0.06232	0.06770	17	0.06075	0.06569
6	0.06216	0.06634	18	0.06050	0.06494
7	0.06200	0.06663	19	0.06043	0.06507
8	0.06184	0.06601	20	0.06024	0.06581
9	0.06172	0.06668	21	0.06013	0.06462
10	0.06163	0.06563	22	0.06004	0.06440
11	0.06156	0.06610	23	0.05981	0.06480
12	0.06138	0.06544			

Table 2.6: Standard error of SACF obtained from 10,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.35$, $\rho_2 = 0.32$ and $\rho_3 = 0.29$ obtained from two different methods.

Lag	**Jackknife s.e**	**Simulated s.e**	**Lag**	**Jackknife s.e**	**Simulated s.e**
1	0.06296	0.07136	13	0.06110	0.06612
2	0.06339	0.06519	14	0.06112	0.06634
3	0.06329	0.06617	15	0.06100	0.06675
4	0.06238	0.06818	16	0.06082	0.06598
5	0.06224	0.06824	17	0.06072	0.06537
6	0.06217	0.06791	18	0.06053	0.06645
7	0.06210	0.06703	19	0.06046	0.06626
8	0.06189	0.06736	20	0.06025	0.06607
9	0.06182	0.06626	21	0.06011	0.06560
10	0.06157	0.06552	22	0.06001	0.06543
11	0.06153	0.06666	23	0.05986	0.06578
12	0.06134	0.06588			

Now we will discuss the patterns in SACF and SPACF plots we obtained for MA(3) stationary Poisson process from our simulation study and propose a method for identifying the order of MA term for Time series of count data. From equation (2.12) we can see that theoretical ACF of a MA(3) Poisson process is non zero at lags 1, 2, 3 and zero at lag 4 and high order.

From the simulation procedure we discussed earlier, we have obtained the following SACF and SPACF plots of MA(3) stationary Poisson process for different ρ_1, ρ_2 and

ρ_3 values satisfying the condition $0 < \rho_3 \leq \rho_2 \leq \rho_1 < 1$.

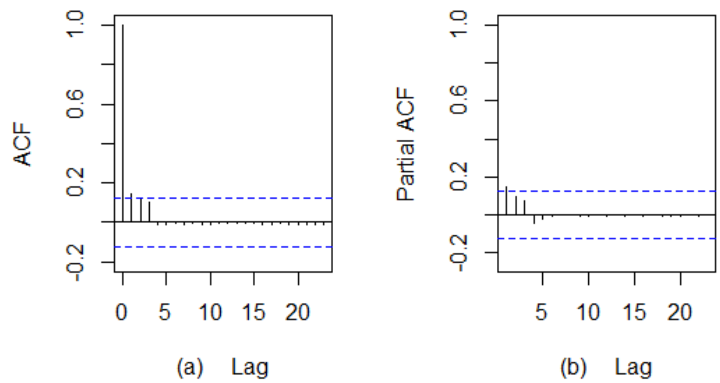

Figure 2.30: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.2$, $\rho_2 = 0.2$ and $\rho_3 = 0.2$.

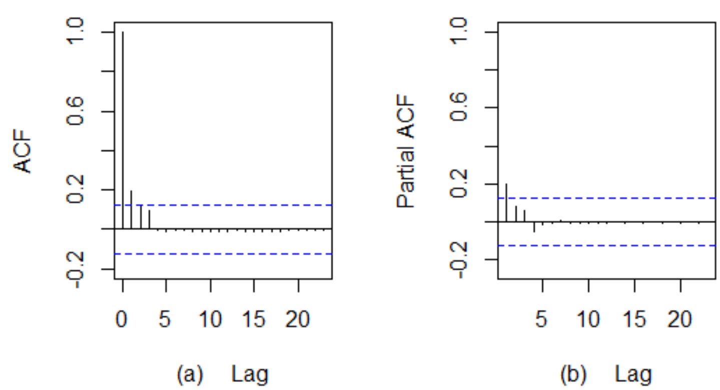

Figure 2.31: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.2$ and $\rho_3 = 0.2$.

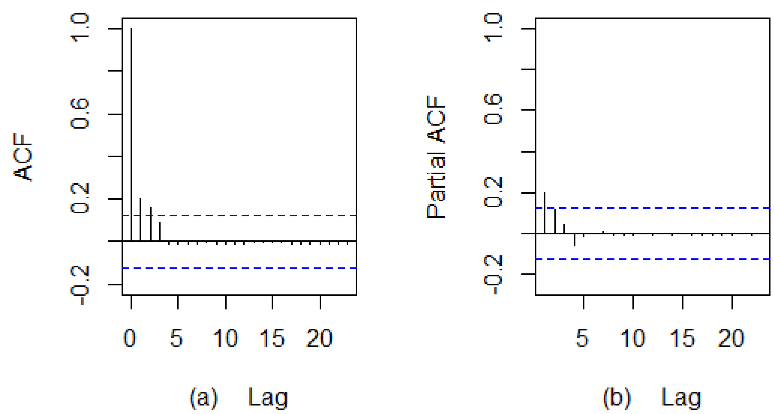

Figure 2.32: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.3$ and $\rho_3 = 0.2$.

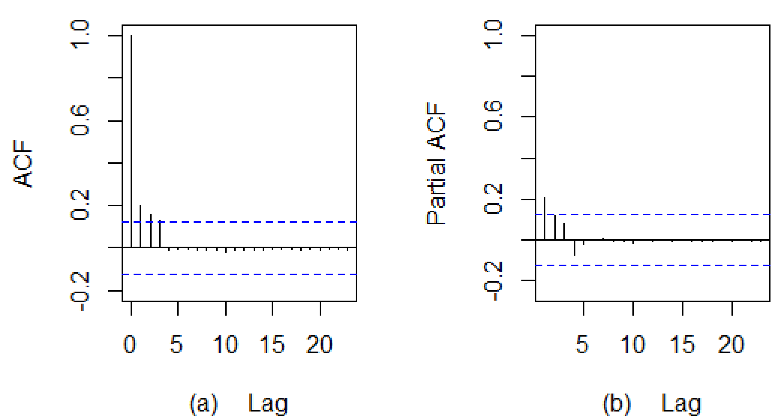

Figure 2.33: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.3$ and $\rho_3 = 0.3$.

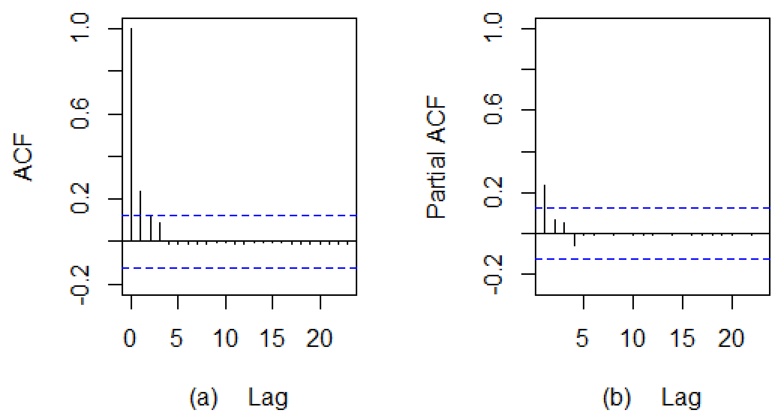

Figure 2.34: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.4$, $\rho_2 = 0.2$ and $\rho_3 = 0.2$.

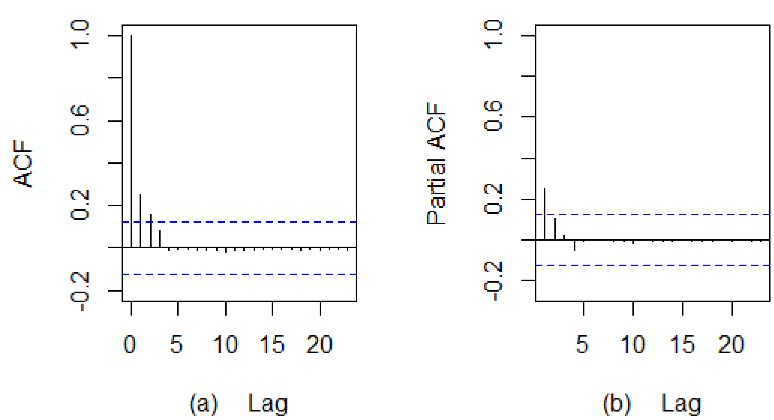

Figure 2.35: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.4$, $\rho_2 = 0.3$ and $\rho_3 = 0.2$.

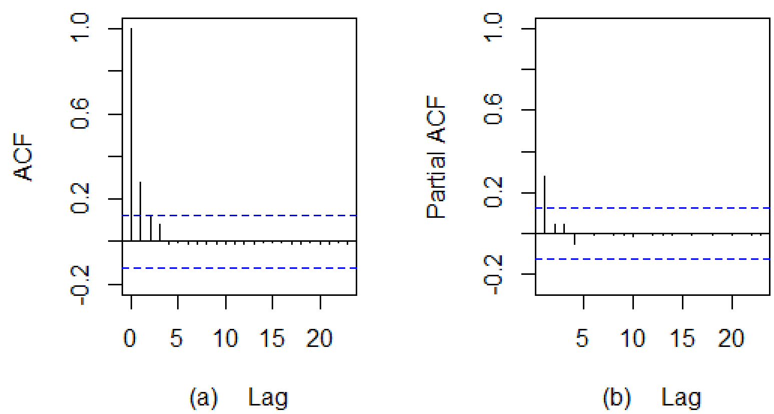

Figure 2.36: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.5$, $\rho_2 = 0.2$ and $\rho_3 = 0.2$.

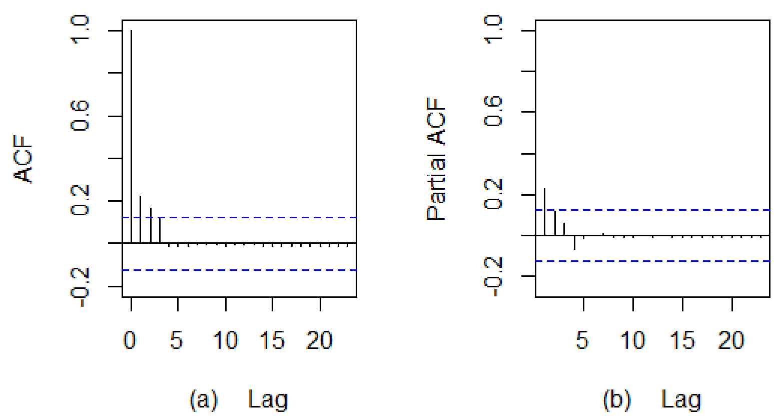

Figure 2.37: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.35$, $\rho_2 = 0.32$ and $\rho_3 = 0.29$.

Most of the SACF plots in Figures 2.30 to 2.37 show that the SACF at lag 1, lag 2 and lag 3 are significant and it increases as ρ_1 and and ρ_2 increases respectively. In the SPACF plots, we can clearly see that SPACF are decaying exponentially, alternating between positive and negative values.

We also considered few combinations of ρ_1, ρ_2 and ρ_3 values which do not satisfy the condition $\rho_1 \geq \rho_2 \geq \rho_3$. SACF plots for those combinations in Figures 2.38 to 2.40 also show that the SACF at lag 1, lag 2 and lag 3 are significant and it increases as ρ_1, ρ_2 and ρ_3 increases respectively. All of the SPACF plots in Figures 2.38 to 2.40 clearly show that SPACF are decaying and alternating between positive and negative sides.

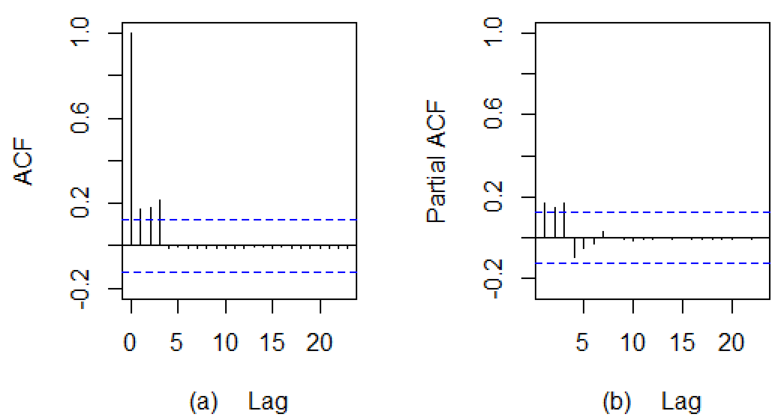

Figure 2.38: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.2$, $\rho_2 = 0.4$ and $\rho_3 = 0.6$.

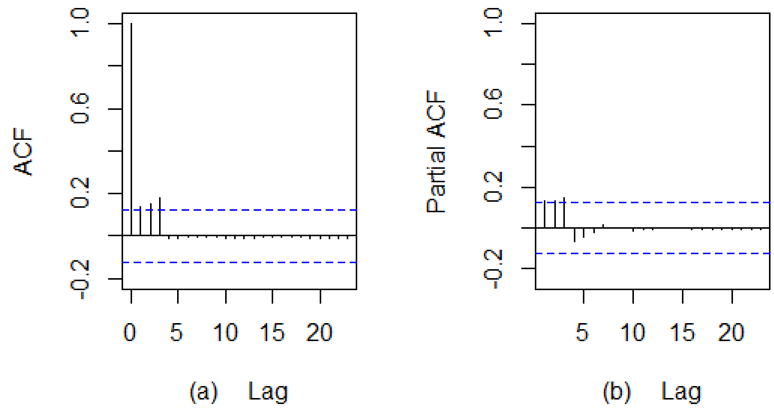

Figure 2.39: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.15$, $\rho_2 = 0.3$ and $\rho_3 = 0.4$.

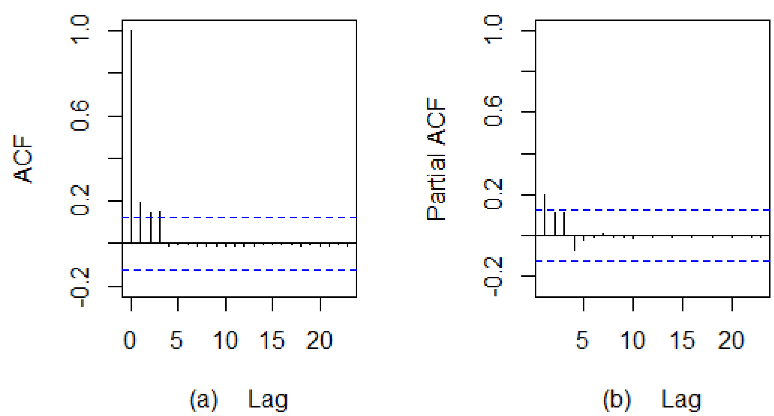

Figure 2.40: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson MA(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.25$ and $\rho_3 = 0.35$.

Therefore, we can come to a conclusion that patterns in the ACF and PACF of MA(3) stationary Poisson processes for count data are similar to that of Stationary models for Gaussian MA(3) processes.

Summary of our simulation study of stationary Poisson MA processes:

In each of the three MA stationary Poisson processes, we have explained the patterns of ACF and PACF and we also have verified that the Jackknife estimator of $s.e(r_k)$ for the stationary Poisson process is approximately equal to the $s.e_(r_k)$ for continuous time series process through our simulation study. Thus, we can conclude that the sample ACF $(r(k))$ of stationary MA(q) Poisson process will cut-off after lag q and the sample PACF will decay exponentially depending on sign and magnitude of model parameters. Hence, patterns in the ACF and PACF of stationary MA Poisson processes are similar to that of MA models for Gaussian processes. However, based on the SACF and SPACF plots especially when ρ values are small, it can be claimed that the data come from AR(q) process as well. For example, if we consider Figure 2.3, both SACF and SPACF cut-off after lag 1. Hence, we may suspect that the data com from either MA(1) or AR(1). In this case, we recommend to fit both models and identify a best model using any model selection criteria. For example, here if we use AIC, we will get lower AIC value for stationary Poisson MA(1) model.

2.3.6 Simulation Study of Stationary Poisson AR(1) Process

We consider a range of ρ value satisfying the condition $0 < \rho < 1$. Then using the computed mean μ as we discussed in the Section 2.3.1, we generated d_t from Poisson distribution with mean $(\mu(1-\rho))$. Next we generated y_t for $t = 1, 2, ..., 250$ from model (2.17). Now we get SACF plot and SPACF plot using the simulated mean values of SACF and SPACF that we found as per the procedure discussed in Section 2.3.1. The above procedure was repeated to check for different values of ρ which are 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9.

In order to test the significance of the SACF values, we first compare the jackknife standard error of SACF to the simulated standard error of SACF values obtained from 10000 simulations for some specific ρ values. Figures 2.41 and 2.42 show that Jackknife estimator of $s.e(\phi_{kk})$ and simulated $s.e(\phi_{kk})$ for AR(1) Poisson process are approximately equal to $s.e(\phi_{kk})$ of AR(1) Gaussian process $\approx \pm 1/\sqrt{T}$, which is the method generally use for continuous time series process.

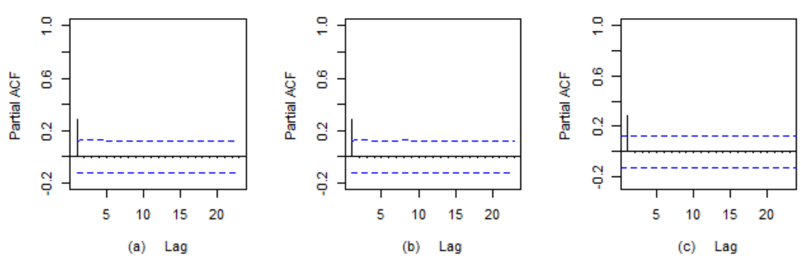

Figure 2.41: A plot of average SPACF values obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.3$ using (a) Jackknife s.e of SPACF (b) simulated s.e of SPACF (c) $s.e(\phi_{kk})$ of AR(1) Gaussian process $\approx 1/\sqrt{T}$.

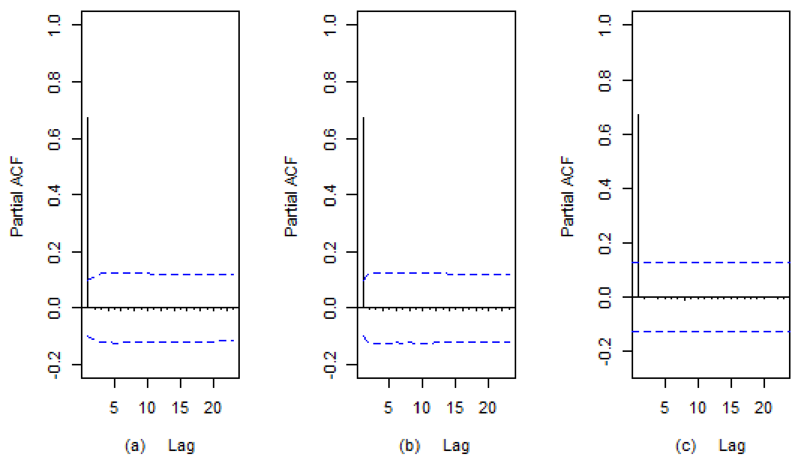

Figure 2.42: A plot of average SPACF values obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.7$ using (a) Jackknife s.e of SPACF (b) simulated s.e of SPACF (c) $s.e(\phi_{kk})$ of AR(1) Gaussian process $\approx 1/\sqrt{T}$.

The standard error values of ϕ_{kk} at each lag are shown in Tables 2.7 and 2.8 for both Jackknife and simulation methods. For example if we compare the Jackknife s.e of (0.06222) and simulated s.e of (0.06233) at lag 1 from Table 2.7 with the $s.e(r_k)$ of AR(1) Gaussian process $= \dfrac{1}{\sqrt{T}} = 0.06325$, it is obvious that these three values are approximately the same. Since these three methods give approximately equal results at each lag, leading to same conclusions, it is recommended to use Gaussian process approach to estimate the standard error which is $\approx \pm\dfrac{1}{\sqrt{T}}$.

Table 2.7: Standard error of SPACF obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.3$ obtained from two different methods.

Lag	**Jackknife s.e**	**Simulated s.e**	**Lag**	**Jackknife s.e**	**Simulated s.e**
1	0.06222	0.06233	13	0.06125	0.06031
2	0.06311	0.06362	14	0.06116	0.06149
3	0.06275	0.06268	15	0.06097	0.06150
4	0.06253	0.06202	16	0.06078	0.06050
5	0.06237	0.06191	17	0.06072	0.06108
6	0.06225	0.06221	18	0.06054	0.06018
7	0.06203	0.06237	19	0.06044	0.06040
8	0.06190	0.06243	20	0.06029	0.06086
9	0.06178	0.06177	21	0.06016	0.05999
10	0.06163	0.06180	22	0.06002	0.05928
11	0.06151	0.06116	23	0.05996	0.06054
12	0.06138	0.06160			

Table 2.8: Standard error of SPACF obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.7$ obtained from two different methods.

Lag	**Jackknife s.e**	**Simulated s.e**	**Lag**	**Jackknife s.e**	**Simulated s.e**
1	0.04908	0.04976	13	0.06026	0.06082
2	0.05771	0.06308	14	0.06013	0.06000
3	0.06069	0.06276	15	0.05997	0.06038
4	0.06170	0.06253	16	0.05986	0.06032
5	0.06189	0.06238	17	0.05965	0.06016
6	0.06173	0.06165	18	0.05945	0.06022
7	0.06152	0.06202	19	0.05933	0.05997
8	0.06124	0.06110	20	0.05922	0.05927
9	0.06099	0.06190	21	0.05909	0.05970
10	0.06073	0.06244	22	0.05899	0.06033
11	0.06051	0.06182	23	0.05888	0.05968
12	0.06040	0.06101			

Now we will discuss the patterns in SACF and SPACF plots for Stationary Poisson AR(1) process. Using those patterns and behaviors we propose a method for identifying the order of AR term for Time series of count data. From equation (2.20) we can see that theoretical PACF of a Stationary Poisson AR(1) process is non zero at lag 1 and zero at lag 2 and high order.

We have obtained the following ACF and PACF plots of Stationary Poisson AR(1) process through the simulation procedure we discussed earlier, for different ρ values satisfying the condition $0 < \rho < 1$.

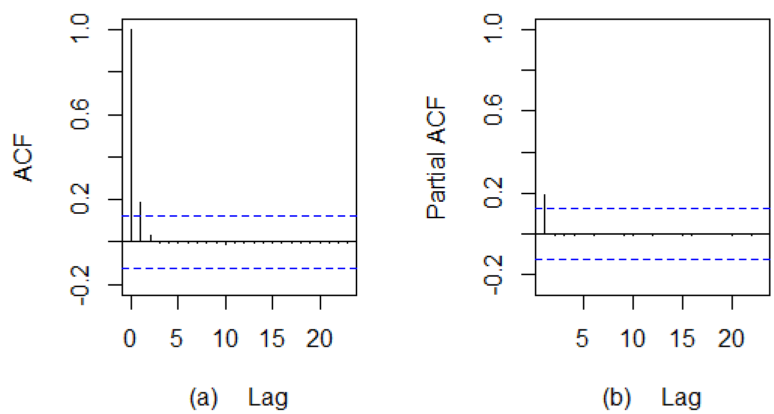

Figure 2.43: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.2$.

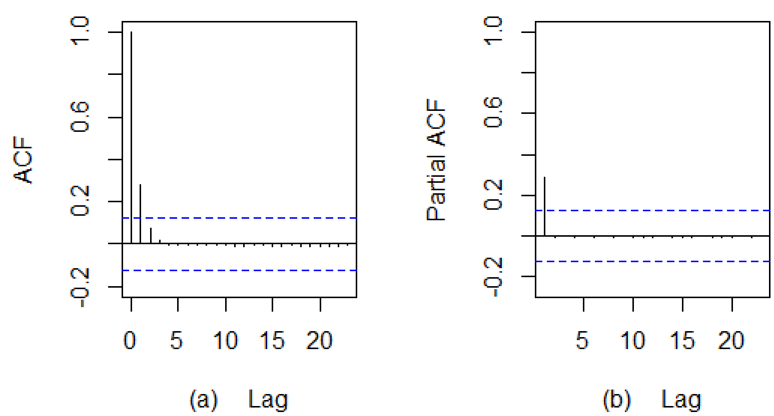

Figure 2.44: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.3$.

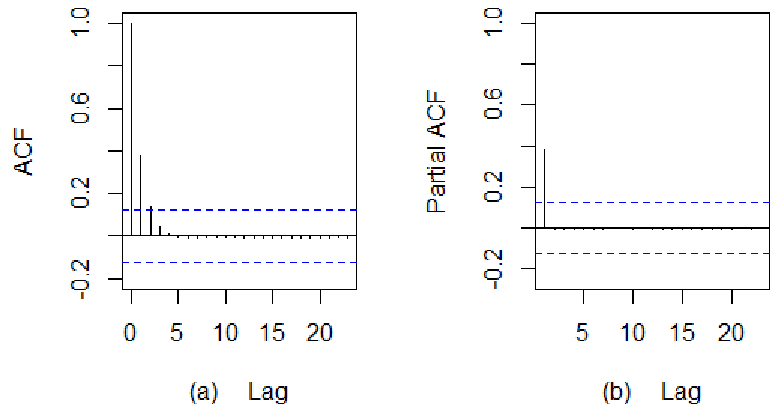

Figure 2.45: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.4$.

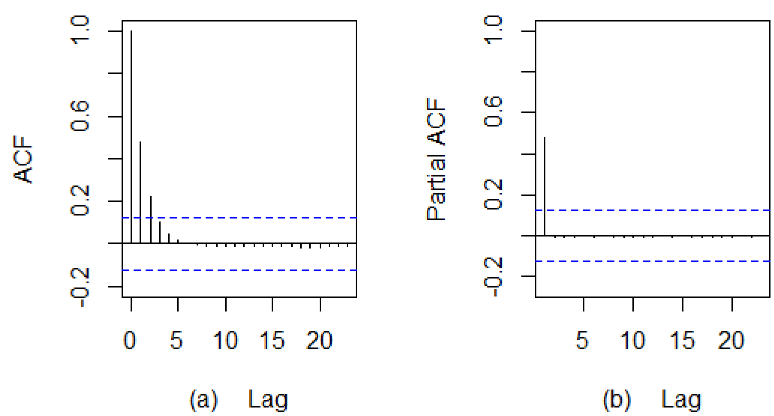

Figure 2.46: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.5$.

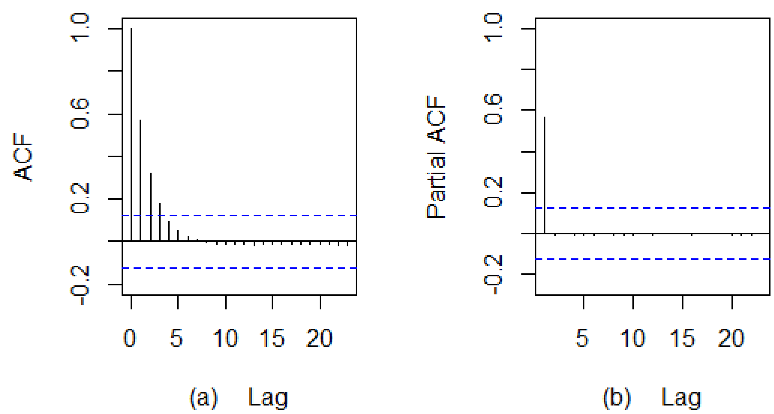

Figure 2.47: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.6$.

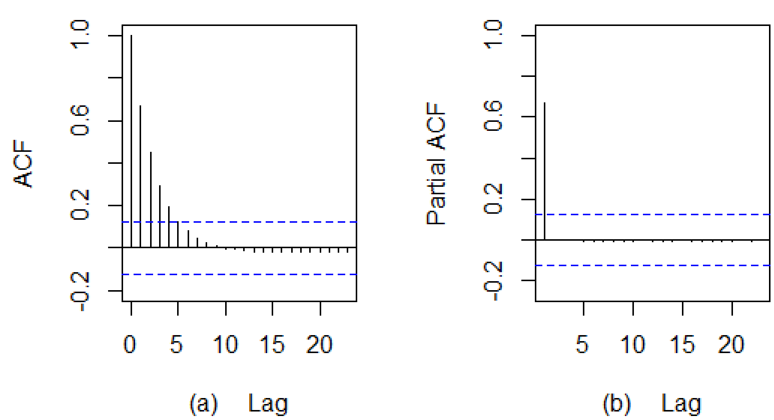

Figure 2.48: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.7$.

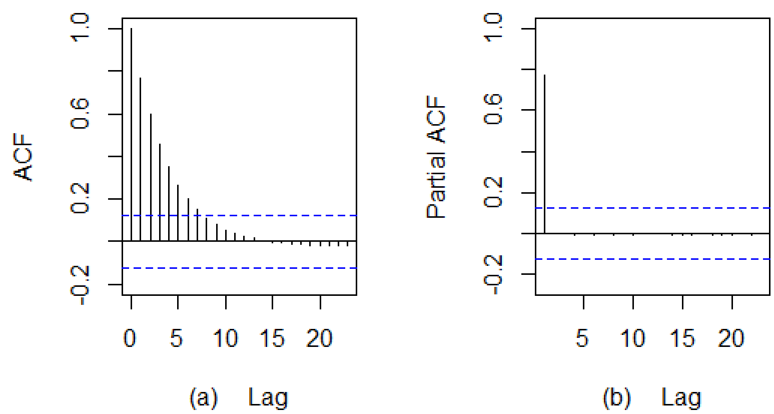

Figure 2.49: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.8$.

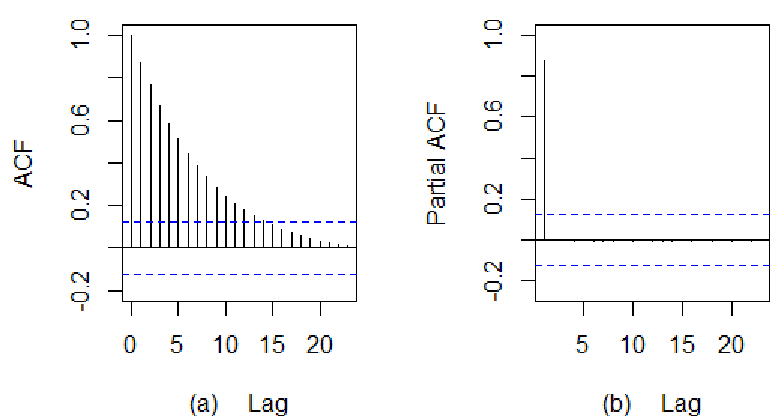

Figure 2.50: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(1) process with $\rho = 0.9$.

The SPACF plots in Figures 2.43 to 2.50 show that the SPACF at lag 1 is significant and it increases as ρ increases. From all of the SACF plots in Figures 2.43 to 2.50, we can see that SACF are decaying exponentially. Therefore, we can come to a conclusion that patterns in the ACF and PACF of Stationary Poisson AR(1) processes are similar to that of Stationary AR(1) models for Gaussian processes.

2.3.7 Simulation Study of Stationary Poisson AR(2) Process

For simulating from stationary Poisson AR(2) process, we choose ρ_1 and ρ_2 values satisfying the conditions $0 \leq \rho_1 \leq (1-\rho_2)$ as we discussed in Section 2.2.2 and $\rho_1 \geq \rho_2$. It is reasonable to assume that $\rho_1 \geq \rho_2$, since it is more likely that the correlation between responses 1 distance apart will be higher than the correlation between the responses 2 distances apart. Now using the computed mean μ as discussed in Section 2.3.1, we generate d_t from Poisson distribution with mean $\big(\mu(1-\rho_1-\rho_2)\big)$. Then we generated y_t for $t = 1, 2, ..., 250$ from model (2.21).

Next we get SACF plot and SPACF plot using the simulated mean values of SACF and SPACF that we found as per the procedure discussed in Section 2.3.1. The above procedure was repeated to check for different values of (ρ_1, ρ_2) including (0.2,0.2), (0.3,0.2), (0.4,0.2), (0.5,0.2), (0.6,0.2), (0.7,0.2), (0.3,0.3), (0.4,0.3), (0.5,0.3), (0.6,0.3), (0.4,0.4) and (0.5,0.4).

In order to test the significance of the SACF values, we first compare the jackknife standard error of SACF to the simulated standard error of SACF values obtained from 10000 simulations for some specific ρ_1 and ρ_2 values. Figures 2.51 and 2.52 show that Jackknife estimator of $s.e(\phi_{kk})$ and simulated $s.e(\phi_{kk})$ for Stationary Poisson AR(2) process are approximately equal to $s.e(\phi_{kk})$ of AR(2) Gaussian process $\approx \pm 1/\sqrt{T}$, which is the method generally use for continuous time series process.

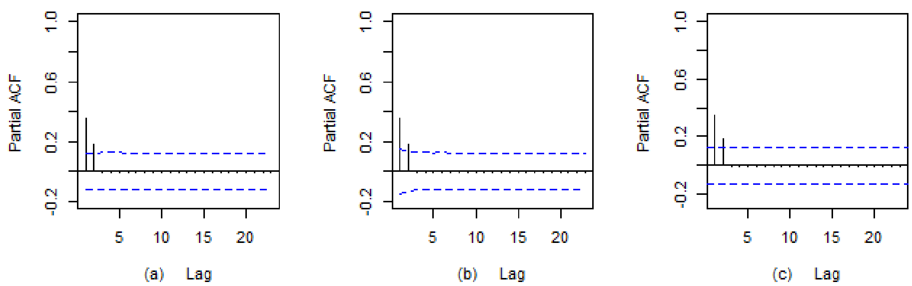

Figure 2.51: A plot of average SPACF values obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.3$ and $\rho_2 = 0.2$ using (a) Jackknife s.e of SPACF (b) simulated s.e of SPACF (c) $s.e(\phi_{kk})$ of AR(1) Gaussian process $\approx 1/\sqrt{T}$.

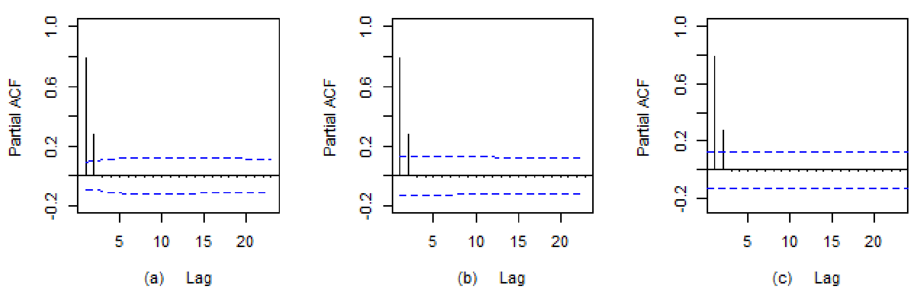

Figure 2.52: A plot of average SPACF values obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.6$ and $\rho_2 = 0.3$ using (a) Jackknife s.e of SPACF (b) simulated s.e of SPACF (c) $s.e(\phi_{kk})$ of AR(1) Gaussian process $\approx 1/\sqrt{T}$.

Table 2.9: Standard error of average SPACF obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho = 0.3$ and $\rho_2 = 0.2$ obtained from two different methods.

Lag	**Jackknife s.e**	**Simulated s.e**	**Lag**	**Jackknife s.e**	**Simulated s.e**
1	0.06123	0.07376	13	0.06102	0.06133
2	0.06225	0.06414	14	0.06090	0.06046
3	0.06311	0.06245	15	0.06069	0.06114
4	0.06289	0.06311	16	0.06058	0.06001
5	0.06262	0.06224	17	0.06047	0.06114
6	0.06227	0.06241	18	0.06037	0.06032
7	0.06207	0.06142	19	0.06019	0.06008
8	0.06184	0.06156	20	0.06006	0.05953
9	0.06159	0.06089	21	0.05993	0.05980
10	0.06147	0.06168	22	0.05978	0.05949
11	0.06135	0.06138	23	0.05967	0.05884
12	0.06112	0.06074			

The standard error values of ϕ_{kk} at each lag are shown in Tables 2.9 and 2.10 for both Jackknife and simulation methods. For example if we compare the Jackknife s.e of (0.06225) and simulated s.e of (0.06414) at lag 2 from Table 2.9 with the $s.e(r_k)$ of AR(2) Gaussian process $= \dfrac{1}{\sqrt{T}} = 0.06325$, it is obvious that these three values are approximately the same. Since these three methods give approximately equal results at each lag, leading to the same conclusions. Therefore, it is recommended to use Gaussian process approach to estimate the standard error which is $\approx \pm\dfrac{1}{\sqrt{T}}$.

Table 2.10: Standard error of average SPACF obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho = 0.6$ and $\rho_2 = 0.3$ obtained from two different methods.

Lag	Jackknife s.e	Simulated s.e	Lag	Jackknife s.e	Simulated s.e
1	0.04480	0.06719	13	0.05958	0.06178
2	0.04934	0.06700	14	0.05928	0.06118
3	0.05412	0.06696	15	0.05894	0.06131
4	0.05673	0.06679	16	0.05867	0.06115
5	0.05844	0.06595	17	0.05833	0.06069
6	0.05945	0.06569	18	0.05804	0.05994
7	0.06002	0.06432	19	0.05773	0.05988
8	0.06027	0.06363	20	0.05741	0.06044
9	0.06032	0.06344	21	0.05707	0.05960
10	0.06027	0.06263	22	0.05678	0.06036
11	0.06013	0.06256	23	0.05651	0.05953
12	0.05989	0.06306			

Now we will discuss the patterns in SACF and SPACF plots we obtained for Stationary Poisson AR(2) process from our simulation study and propose a method for identifying the order of AR term for Time series of count data. From equation (2.20) we can see that theoretical PACF of a Stationary Poisson AR(2) process is non zero at lags 1 , 2 and zero at lag 3 and high order.

From the simulation procedure we discussed earlier, we have obtained the following ACF and PACF plots of Stationary Poisson AR(2) process for different ρ_1 and ρ_2

values satisfying the conditions $0 \leq \rho_1 \leq (1 - \rho_2)$ and $\rho_1 \geq \rho_2$.

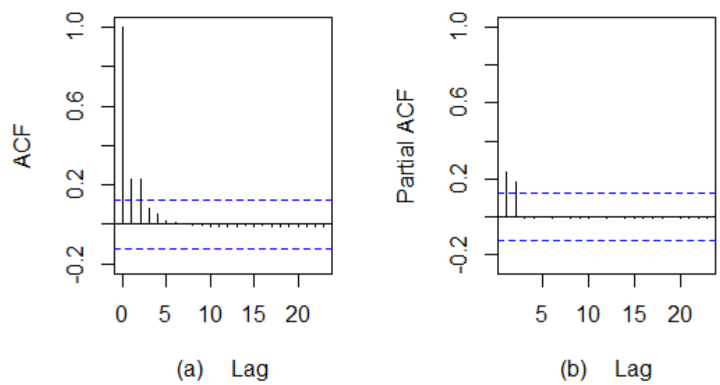

Figure 2.53: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.2$ and $\rho_2 = 0.2$.

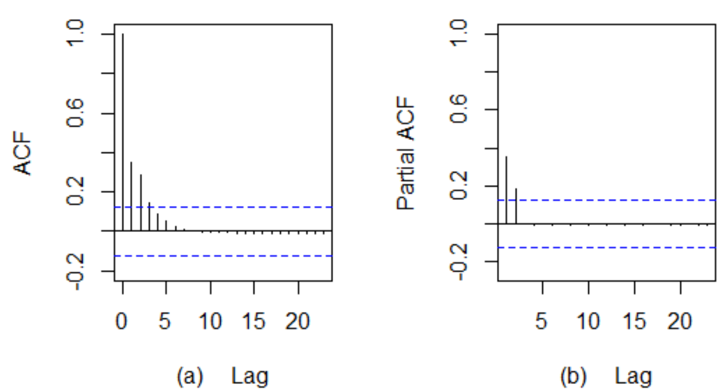

Figure 2.54: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.3$ and $\rho_2 = 0.2$.

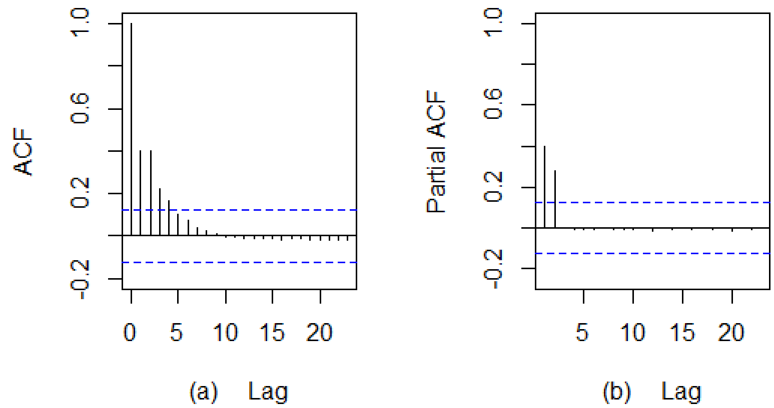

Figure 2.55: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.3$ and $\rho_2 = 0.3$.

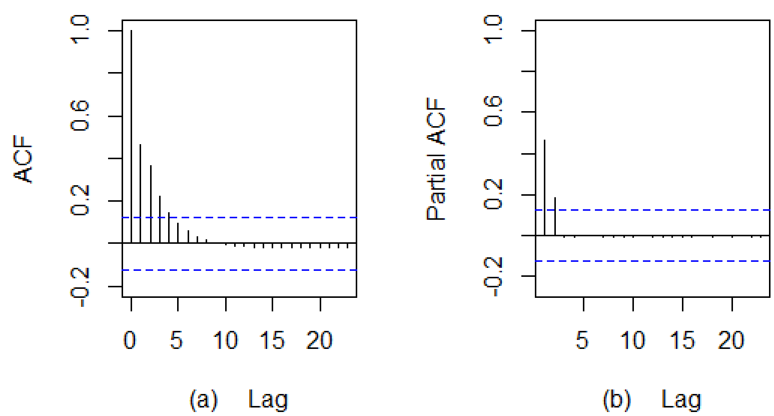

Figure 2.56: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.4$ and $\rho_2 = 0.2$.

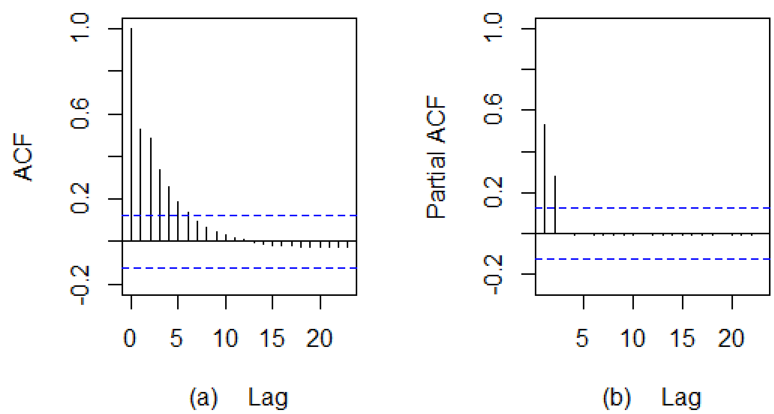

Figure 2.57: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.4$ and $\rho_2 = 0.3$.

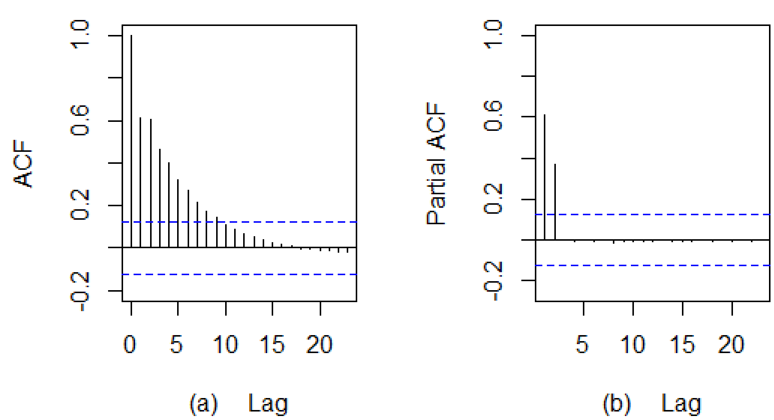

Figure 2.58: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.4$ and $\rho_2 = 0.4$.

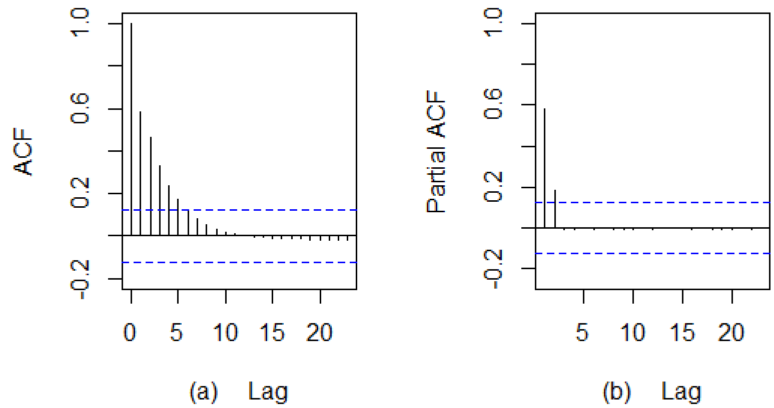

Figure 2.59: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.5$ and $\rho_2 = 0.2$.

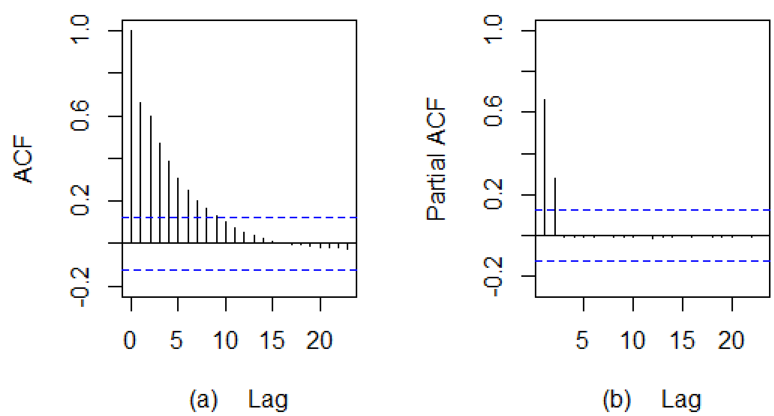

Figure 2.60: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.5$ and $\rho_2 = 0.3$.

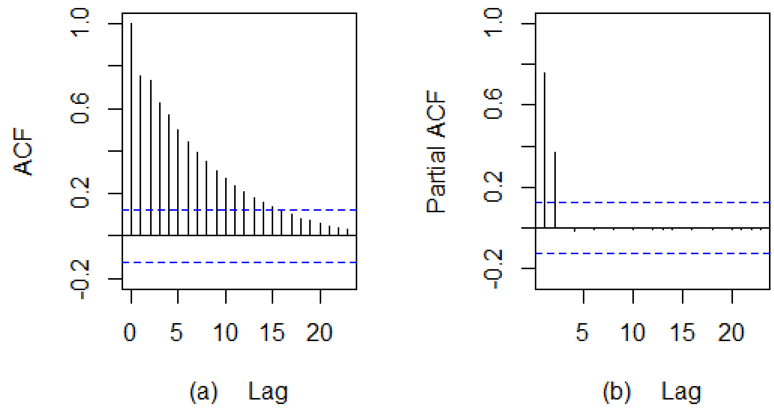

Figure 2.61: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.5$ and $\rho_2 = 0.4$.

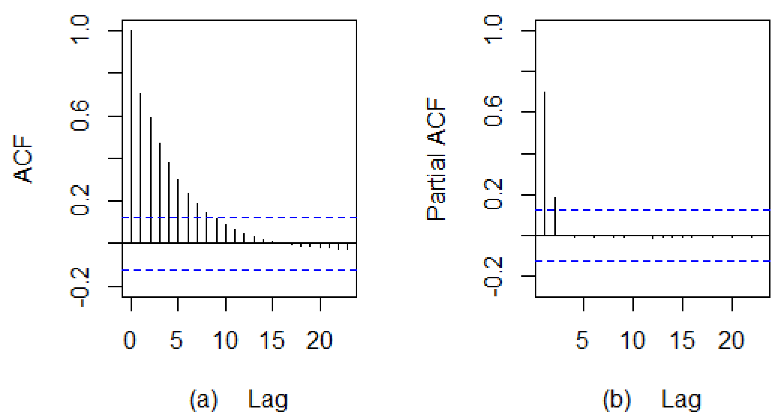

Figure 2.62: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.6$ and $\rho_2 = 0.2$.

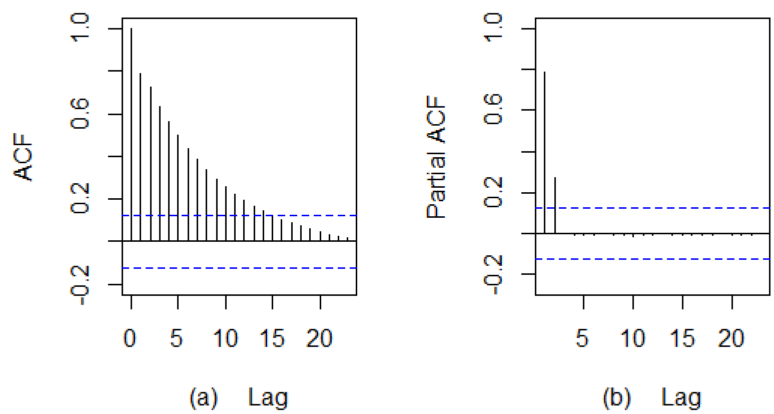

Figure 2.63: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.6$ and $\rho_2 = 0.3$.

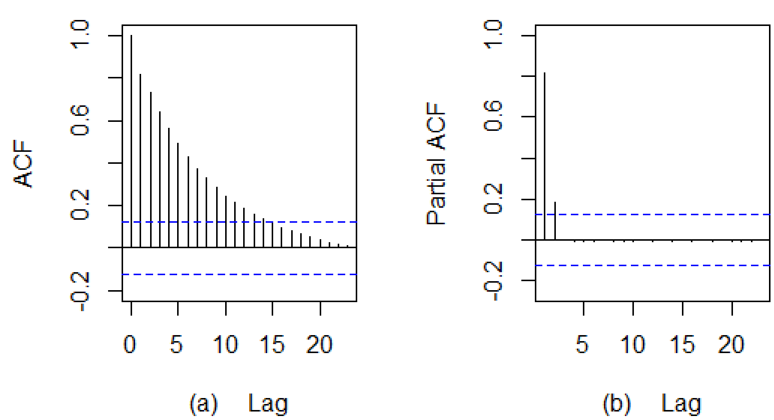

Figure 2.64: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.7$ and $\rho_2 = 0.2$.

The SPACF plots in Figures 2.53 to 2.64 show that the SPACF at lag 1 and lag 2 are significant and those increase when ρ_1 and ρ_2 increase respectively. From all of the SACF plots in Figures 2.53 to 2.64, we can clearly see that SACF are decaying exponentially.

We also considered few combinations of ρ_1 and ρ_2 values where $\rho_1 < \rho_2$. SPACF plots for those combinations in Figures 2.65 to 2.67 also show that the SPACF at lag 1 and lag 2 are significant and SPACF at lag 1 is greater than SPACF at lag 2 when there is no big difference between ρ_1 and ρ_2 even if $\rho_1 < \rho_2$. All of the SACF plots in Figures 2.65 to 2.67 clearly show that SACF are exponentially decaying.

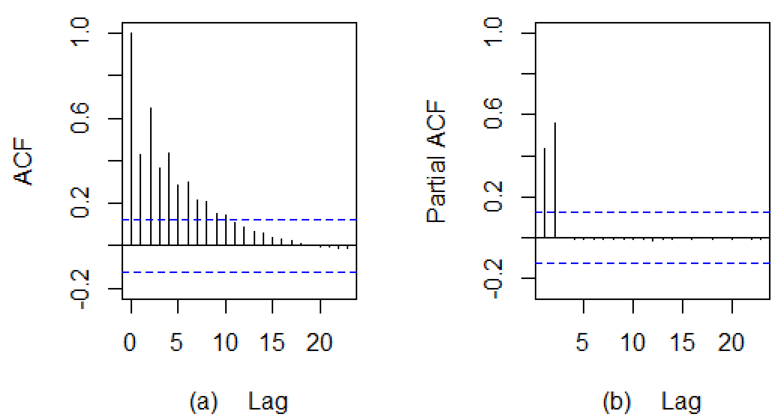

Figure 2.65: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.2$ and $\rho_2 = 0.6$.

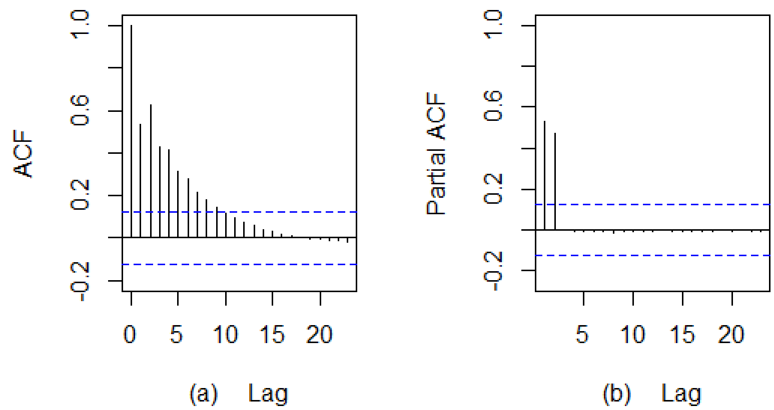

Figure 2.66: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.3$ and $\rho_2 = 0.5$.

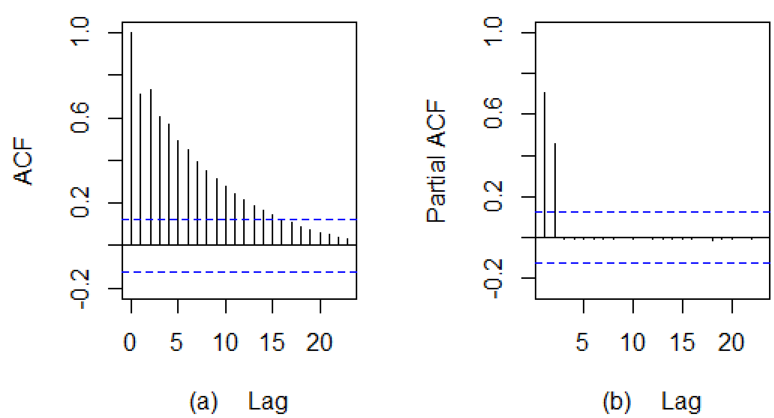

Figure 2.67: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(2) process with $\rho_1 = 0.4$ and $\rho_2 = 0.5$.

Therefore, we can come to a conclusion that the patterns in the ACF and PACF of Stationary Poisson AR(2) processes of count data are similar to that of Stationary models for Gaussian AR(2) processes.

2.3.8 Simulation Study of Stationary Poisson AR(3) Process

First we choose ρ_1, ρ_2 and ρ_3 values satisfying the conditions $0 \leq \rho_1 \leq (1 - \rho_2 - \rho_3)$ as we discussed in Section 2.2.2 and $\rho_1 \geq \rho_2 \geq \rho_3$. It is reasonable to assume that $\rho_1 \geq \rho_2 \geq \rho_3$, since it is more likely that the correlation between responses 1 distance apart will be higher than the correlation between the responses 2 distances apart which is more than the correlation between the responses 3 distances apart.

Now using the computed mean μ as discussed in the Section 2.3.1, we generate d_t from Poisson distribution with mean $\big(\mu(1 - \rho_1 - \rho_2 - \rho_3)\big)$. Then we generated y_t for $t = 1, 2, ..., 250$ from model (2.27).

Then we get SACF plot and SPACF plot using the simulated mean values of SACF and SPACF that we found as per the procedure discussed in Section 2.3.1. The above procedure was repeated to check for different combination of ρ_1, ρ_2 and ρ_3 values which are (0.2,0.2,0.2), (0.3,0.2,0.2), (0.3,0.3,0.2), (0.3,0.3,0.3), (0.4,0.2,0.2), (0.4,0.3,0.2), (0.5,0.2,0.2) and (0.35,0.32,0.29) respectively.

In order to test the significance of the SACF values, we first compare the jackknife standard error of SACF to the simulated standard error of SACF values obtained from 10000 simulations for some specific ρ_1, ρ_2 and ρ_3 values. Figures 2.68 and 2.69 show that Jackknife estimator of $s.e(\phi_{kk})$ and simulated $s.e(\phi_{kk})$ for Stationary Poisson AR(3) process are approximately equal to $s.e(\phi_{kk})$ of AR(3) Gaussian process $\approx \pm 1/\sqrt{T}$, which is the method generally use for continuous time series process.

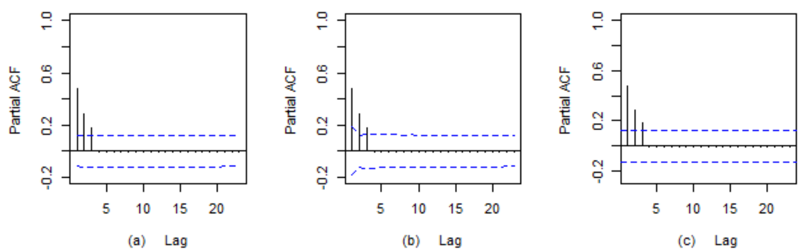

Figure 2.68: A plot of average SPACF values obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.25$ and $\rho_3 = 0.2$ using (a) Jackknife s.e of SPACF (b) simulated s.e of SPACF (c) $s.e(\phi_{kk})$ of AR(1) Gaussian process $\approx 1/\sqrt{T}$.

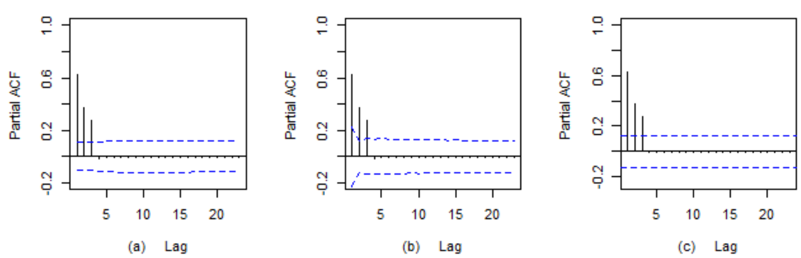

Figure 2.69: A plot of average SPACF values obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.3$ and $\rho_3 = 0.3$ using (a) Jackknife s.e of SPACF (b) simulated s.e of SPACF (c) $s.e(\phi_{kk})$ of AR(1) Gaussian process $\approx 1/\sqrt{T}$.

Table 2.11: Standard error of SPACF obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.25$ and $\rho_3 = 0.2$ obtained from two different methods.

Lag	**Jackknife s.e**	**Simulated s.e**	**Lag**	**Jackknife s.e**	**Simulated s.e**
1	0.05862	0.09217	13	0.06074	0.06127
2	0.05919	0.06015	14	0.06046	0.06116
3	0.05998	0.06412	15	0.06019	0.06066
4	0.06172	0.06457	16	0.05996	0.05993
5	0.06203	0.06374	17	0.05976	0.05956
6	0.06215	0.06348	18	0.05953	0.05971
7	0.06215	0.06341	19	0.05936	0.05948
8	0.06194	0.06222	20	0.05914	0.05997
9	0.06178	0.06262	21	0.05899	0.05924
10	0.06147	0.06170	22	0.05879	0.05866
11	0.06123	0.06084	23	0.05863	0.05902
12	0.06099	0.06108			

The standard error values of ϕ_{kk} at each lag are shown in Tables 2.11 and 2.12 for both Jackknife and simulation methods. For example if we compare the Jackknife s.e of (0.05998) and simulated s.e of (0.06412) at lag 3 from Table 2.11 with the $s.e(r_k)$ of AR(3) Gaussian process $= \dfrac{1}{\sqrt{T}} = 0.06325$, it is obvious that these three values are approximately the same. Since these three methods give approximately equal results at each lag, leading to same conclusions, it is recommended to use Gaussian process approach to estimate the standard error which is $\approx \pm\dfrac{1}{\sqrt{T}}$.

Table 2.12: Standard error of SPACF obtained from 10,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.3$ and $\rho_3 = 0.3$ obtained from two different methods.

Lag	Jackknife s.e	Simulated s.e	Lag	Jackknife s.e	Simulated s.e
1	0.05355	0.11282	13	0.05958	0.06302
2	0.05374	0.05995	14	0.05948	0.06214
3	0.05368	0.06768	15	0.05921	0.06249
4	0.05719	0.06633	16	0.05906	0.06121
5	0.05783	0.06735	17	0.05879	0.06090
6	0.05851	0.06557	18	0.05853	0.06096
7	0.05921	0.06522	19	0.05831	0.06122
8	0.05950	0.06455	20	0.05806	0.05987
9	0.05974	0.06359	21	0.05782	0.06077
10	0.05979	0.06425	22	0.05753	0.05938
11	0.05978	0.06372	23	0.05729	0.05931
12	0.05971	0.06290			

Now we will discuss the patterns in SACF and SPACF plots we obtained for Stationary Poisson AR(3) process from our simulation study and propose a method for identifying the order of AR term for Time series of count data. From equation (2.32) we can see that theoretical PACF of a Stationary Poisson AR(3) process is non zero at lags 1, 2, 3 and zero at lag 4 and high order.

From the simulation procedure we discussed earlier, we have obtained the following SACF and SPACF plots of Stationary Poisson AR(3) process for different ρ_1, ρ_2 and

ρ_3 values satisfying the conditions $0 \leq \rho_1 \leq (1 - \rho_2 - \rho_3)$ and $\rho_1 \geq \rho_2 \geq \rho_3$.

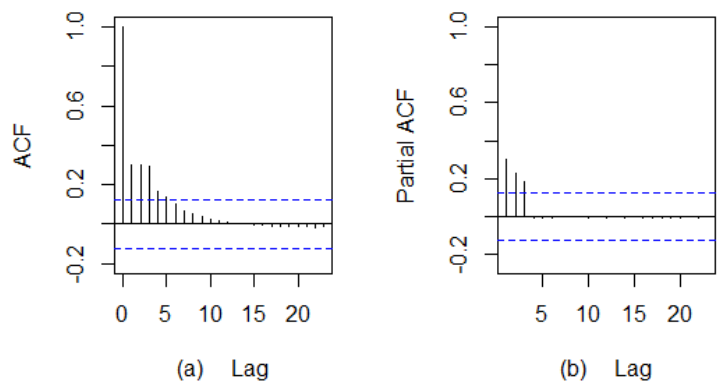

Figure 2.70: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.2$, $\rho_2 = 0.2$ and $\rho_3 = 0.2$.

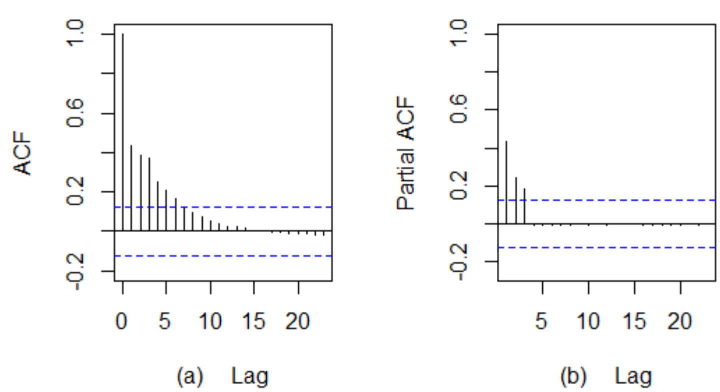

Figure 2.71: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.2$ and $\rho_3 = 0.2$.

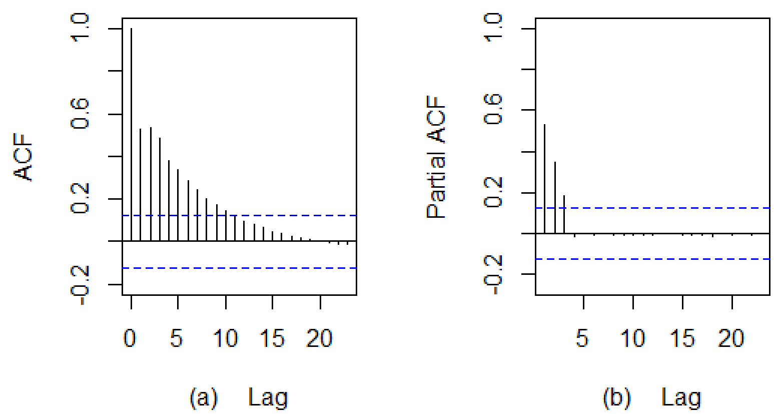

Figure 2.72: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.3$ and $\rho_3 = 0.2$.

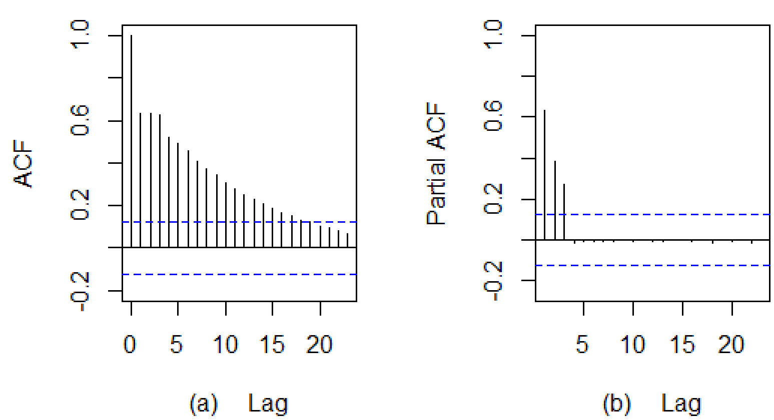

Figure 2.73: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.3$ and $\rho_3 = 0.3$.

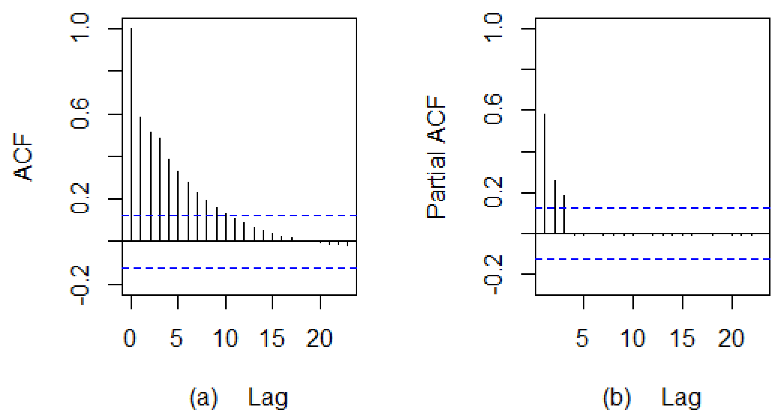

Figure 2.74: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.4$, $\rho_2 = 0.2$ and $\rho_3 = 0.2$.

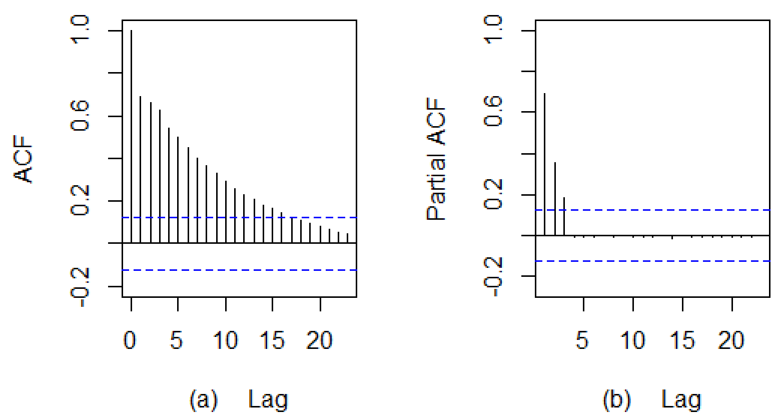

Figure 2.75: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.4$, $\rho_2 = 0.3$ and $\rho_3 = 0.2$.

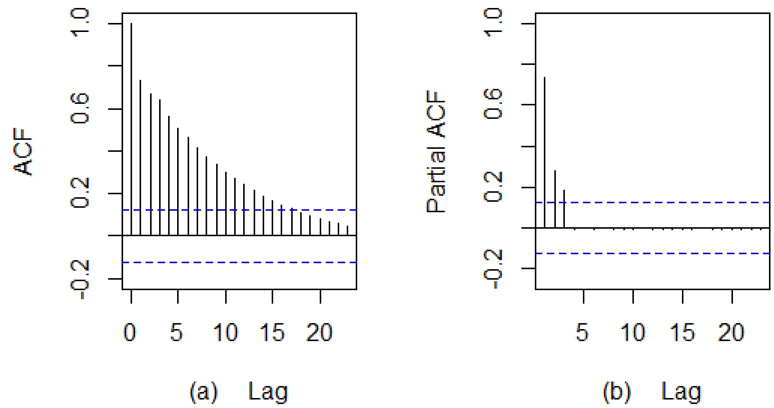

Figure 2.76: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.5$, $\rho_2 = 0.2$ and $\rho_3 = 0.2$.

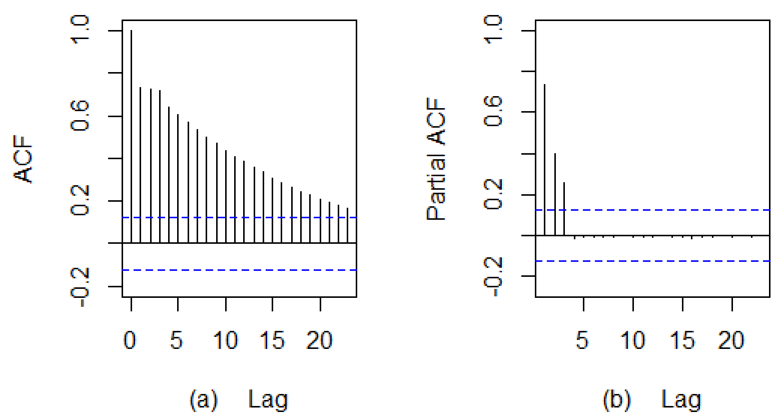

Figure 2.77: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.35$, $\rho_2 = 0.32$ and $\rho_3 = 0.29$.

The SACF plots in Figures 2.70 to 2.77 show that the SPACF at lag 1, lag 2 and lag 3 are significant and those increase when ρ_1, ρ_2 and ρ_3 increase respectively. From all of the SACF plots in Figures 2.70 to 2.77, we can clearly see that SACF are decaying exponentially.

We also considered few combinations of ρ_1, ρ_2 and ρ_3 values which do not satisfy the condition $\rho_1 \geq \rho_2 \geq \rho_3$. SPACF plots for those combinations in Figures 2.78 to 2.81 also show that the SPACF at lag 1, lag 2 and lag 3 are significant and SPACF at lag 1 is greater than SPACF at lag 2 when there is no big difference between ρ_1, ρ_2 and ρ_3 even if $\rho_1 < \rho_2 < \rho_3$. All of the SACF plots in Figures 2.78 to 2.81 show that SACF are exponentially decaying in overall.

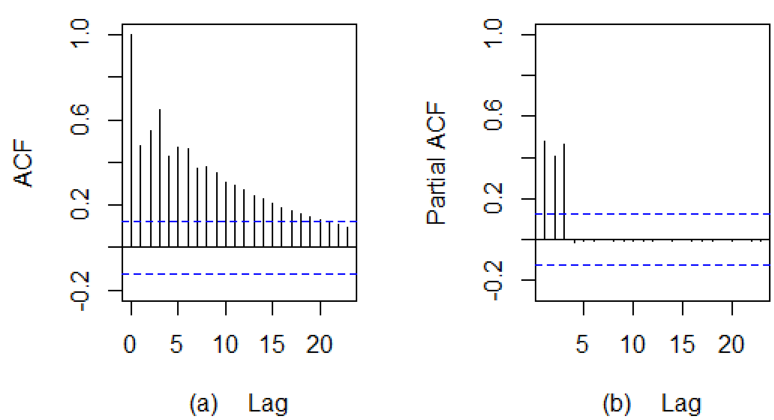

Figure 2.78: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.1$, $\rho_2 = 0.3$ and $\rho_3 = 0.5$.

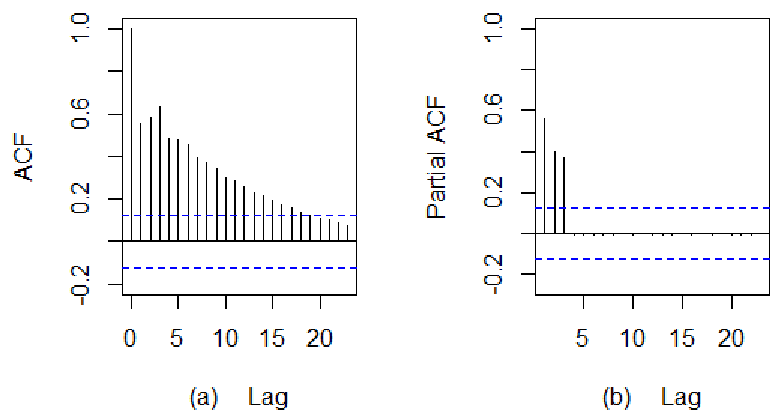

Figure 2.79: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.2$, $\rho_2 = 0.3$ and $\rho_3 = 0.4$.

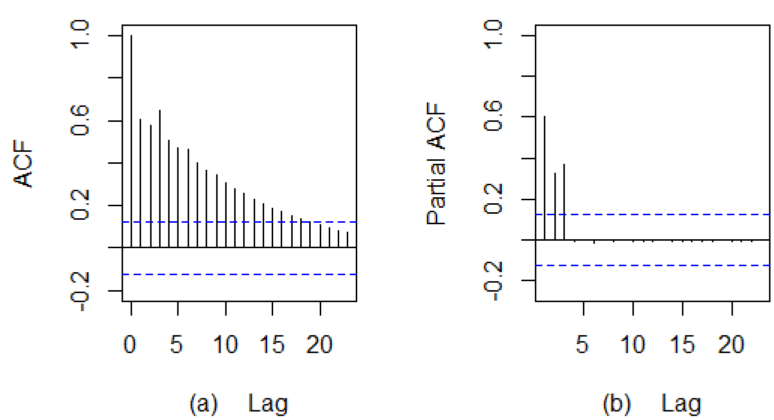

Figure 2.80: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.3$, $\rho_2 = 0.2$ and $\rho_3 = 0.4$.

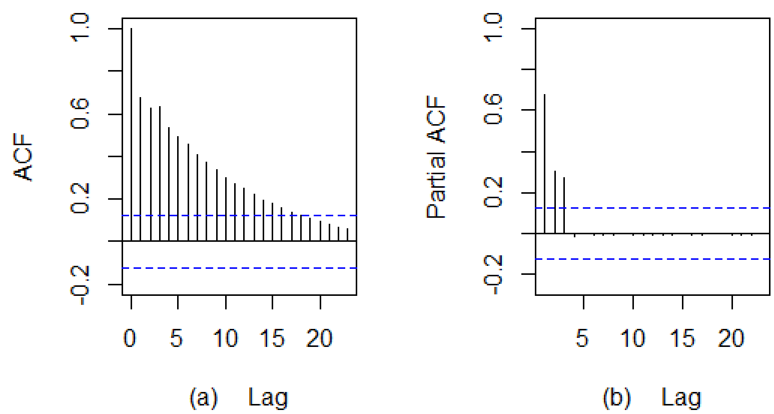

Figure 2.81: A plot of (a) average SACF (b) average SPACF; obtained from 1,000 simulations of 250 observations of Stationary Poisson AR(3) process with $\rho_1 = 0.4$, $\rho_2 = 0.2$ and $\rho_3 = 0.3$.

Therefore, we can come to a conclusion that the patterns in the ACF and PACF of Stationary Poisson AR(3) processes of count data are similar to that of Stationary models for Gaussian AR(3) processes.

Summary of our simulation study of stationary Poisson AR processes:

In each of the three Stationary Poisson AR processes, we have explained the patterns of ACF and PACF and we also have verified that the Jackknife estimator of ϕ_{kk} for the Stationary Poisson AR process is approximately equal to the $s.e_{(}\phi_{kk})$ for continuous time series process through our simulation study. Thus, we can expect that the sample PACF of Stationary Poisson AR(p) process will cut off after lag p and the sample PACF will decay exponentially. Hence, patterns in the ACF and PACF of Stationary Poisson AR processes are similar to that of AR models for Gaussian processes.

Chapter 3

The Non Stationary Poisson AR(3) Process

In Chapter 2, we discussed the order identification of stationary Poisson AR process of order up to three. The focus of our discussion throughout this chapter will be on the third order autoregressive dynamic model for non stationary count time series.

We mentioned that the autoregressive dynamic model of order one and two have been discussed by McKenzie (1988) and Zhang and Oyet (2014), respectively. We propose the model for AR(3) non stationary process and derive the basic properties of mean, variance, covariance and correlation. We also verify that the properties satisfy non stationary conditions. Then, we derive the estimators for model parameters using GQL and GMM. Finally, we estimate the model parameters and examine the accuracy of estimates through simulation study.

3.1 Proposed Non Stationary Poisson AR(3) Model

Let y_t represent the response at time t and μ_t the mean at time t. We assume that the mean μ_t is a function of covariates $\boldsymbol{x}_t$ such that $\mu_t = exp(\boldsymbol{x}_t'\boldsymbol{\beta})$ where $\boldsymbol{\beta}$ is a measure of the covariate effects.

Now the Poisson AR(3) non stationary model can be defined by,

$$\begin{aligned} & y_1 \sim Poi(\mu_1) \quad with \quad \mu_1 = exp(x_1'\boldsymbol{\beta}), \\ & y_2 = \sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2, \\ & y_3 = \sum_{j=1}^{y_2} b_{1j}(\rho_1) + \sum_{j=1}^{y_1} b_{2j}(\rho_2) + d_3, \\ & y_t = \sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1) + \sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2) + \sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3) + d_t, \quad for \quad t = 4, 5, ..., T \end{aligned} \tag{3.1}$$

where $\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1) = \rho_1 * y_{t-1}$ is the binomial thinning operation (Ed McKenzie, 1988). That is, $Pr[b_{ij}(\rho_i) = 1] = \rho_i$ and $Pr[b_{ij}(\rho_i) = 0] = 1 - \rho_i$ for $i = 1, 2, 3$. Furthermore, we make the following assumptions:

1. $d_2 \sim Poi(\mu_2 - \rho_1\mu_1)$
2. $d_3 \sim Poi(\mu_3 - \rho_1\mu_2 - \rho_2\mu_1)$
3. $d_t \sim Poi(\mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3})$, for $t = 4, 5, ..., T$
4. d_t and y_{t-1} are independent for $t = 2, 3, ..., T$
5. d_t and y_{t-2} are independent for $t = 3, 4, ..., T$
6. d_t and y_{t-3} are independent for $t = 4, 5, ..., T$

It is reasonable to assume that $\rho_1 \geq \rho_2 \geq \rho_3 \geq 0$, since it is more likely the correlation between the responses 1 distance apart is more than the correlation between the responses 2 distances apart which is more than the correlation between the responses 3 distances apart and so on. In addition, non-negativity of the mean of a Poisson random variable leads to the condition on the parameter ρ_1 which is for a fixed value of $\rho_2 \geq \rho_3 \geq 0$ and $t \geq 4$

$$0 \leq \rho_1 \leq min\Big(\frac{\mu_2}{\mu_1}, \frac{\mu_3 - \rho_2\mu_1}{\mu_2}, \frac{\mu_t - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}}{\mu_{t-1}}, 1\Big) \tag{3.2}$$

in model (3.1).

3.1.1 Basic properties

One of the main focus of our study was to derive the basic properties of the Poisson AR(3) non stationary process.

Mean

From the AR(3) model (3.1), we can see that $E(y_1) = \mu_1$.

Now, at $t = 2$, we find that

$$\begin{aligned}
E(y_2) &= E\Big[\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big] \\
&= E_{y_1} E\Big[\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2 \Big| y_1\Big] \\
&= E_{y_1}\Big[y_1\rho_1 + \mu_2 - \rho_1\mu_1\Big] \\
&= \rho_1\mu_1 + \mu_2 - \rho_1\mu_1 \\
&= \mu_2
\end{aligned}$$

Next at $t = 3$, we find that

$$\begin{aligned}
E(y_3) =& E\Big[\sum_{j=1}^{y_2} b_{1j}(\rho_1) + \sum_{j=1}^{y_1} b_{2j}(\rho_2) + d_3\Big] \\
&= E\Big[\sum_{j=1}^{y_2} b_{1j}(\rho_1)\Big] + E\Big[\sum_{j=1}^{y_1} b_{2j}(\rho_2)\Big] + E\Big[d_3\Big] \\
&= E_{y_2} E\Big[\sum_{j=1}^{y_2} b_{1j}(\rho_1)\Big| y_2\Big] + E_{y_1} E\Big[\sum_{j=1}^{y_1} b_{2j}(\rho_2)\Big| y_1\Big] + E\Big[d_3\Big] \\
&= E_{y_2}\Big[y_2\rho_1\Big] + E_{y_1}\Big[y_1\rho_2\Big] + E\Big[d_3\Big] \\
&= \rho_1\mu_2 + \rho_2\mu_1 + \mu_3 - \rho_1\mu_2 - \rho_2\mu_1 \\
&= \mu_3
\end{aligned}$$

In general, we can show that the mean of y_t for $t = 4, 5, ..., T$ is

$$
\begin{aligned}
E(y_t) &= E\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1) + \sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2) + \sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3) + d_t\Big] \\
&= E\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1)\Big] + E\Big[\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big] + E\Big[\sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3)\Big] + E\Big[d_t\Big] \\
&= E_{y_{t-1}} E\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1)\Big| y_{t-1}\Big] + E_{y_{t-2}} E\Big[\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big| y_{t-2}\Big] \\
&+ E_{y_{t-3}} E\Big[\sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3)\Big| y_{t-3}\Big] + E\Big[d_t\Big] \\
&= \rho_1 E[y_{t-1}] + \rho_2 E[y_{t-2}] + \rho_3 E[y_{t-3}] + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}
\end{aligned}
$$

Now, by taking successive expectation we find that $E(y_t) = \mu_t = exp(\boldsymbol{x}_t'\boldsymbol{\beta})$ for all $t = 1, 2, ..., T$.

Variance

As we discussed in Sections 2.2.2 and 2.2.3, since $b_{ij}(\rho_i)$ is a sequence of independent identically distributed (i.i.d) Binary random variables with $Pr[b_{ij}(\rho_i) = 1] = \rho_i$ and $Pr[b_{ij}(\rho_i) = 0] = 1 - \rho_i \quad \text{for all} \quad i = 1, 2, 3$, the following covariance terms will vanish when we take variance to the response y_t.

$$
Cov\Big(\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1), \sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big) = Cov\Big(\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1), \sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3)\Big)
$$
$$
= Cov\Big(\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2), \sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3)\Big) = 0
$$

We also note that from assumptions 4,5 and 6 of model (3.1),

$$
Cov\Big(\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1), d_t\Big) = Cov\Big(\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2), d_t\Big) = Cov\Big(\sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3), d_t\Big) = 0.
$$

From the AR(3) model (3.1), we see that $\sigma_{1,1} = Var(y_1) = \mu_1$. Now at $t = 2$, we find

$$\begin{aligned}
\sigma_{2,2} &= Var(y_2) \\
&= Var\Big[\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big] \\
&= E_{y_1}\Big[Var\Big(\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big|y_1\Big)\Big] + Var_{y_1}\Big[E\Big(\sum_{j=1}^{y_1} b_{1j}(\rho_1) + d_2\Big|y_1\Big)\Big] \\
&= E_{y_1}\Big[y_1\rho_1(1-\rho_1) + \mu_2 - \rho_1\mu_1\Big] + Var_{y_1}\Big[y_1\rho_1 + \mu_2 - \rho_1\mu_1\Big] \\
&= \rho_1(1-\rho_1)E_{y_1}\Big[y_1\Big] + \mu_2 - \rho_1\mu_1 + \rho_1^2 Var_{y_1}\Big[y_1\Big] \\
&= \rho_1(1-\rho_1)\mu_1 + \mu_2 - \rho_1\mu_1 + \rho_1^2\mu_1 = \mu_2.
\end{aligned}$$

Similarly we find the $Var(y_t)$ at $t = 3$,

$$\begin{aligned}
\sigma_{3,3} &= Var(y_3) \\
&= Var\Big[\sum_{j=1}^{y_2} b_{1j}(\rho_1) + \sum_{j=1}^{y_1} b_{2j}(\rho_2) + d_3\Big] \\
&= Var\Big[\sum_{j=1}^{y_2} b_{1j}(\rho_1)\Big] + Var\Big[\sum_{j=1}^{y_1} b_{2j}(\rho_2)\Big] + Var\Big[d_3\Big] \\
&= E_{y_2}\Big[Var\Big(\sum_{j=1}^{y_2} b_{1j}(\rho_1)\Big|y_2\Big)\Big] + Var_{y_2}\Big[E\Big(\sum_{j=1}^{y_2} b_{1j}(\rho_1)\Big|y_2\Big)\Big] \\
&\quad + E_{y_1}\Big[Var\Big(\sum_{j=1}^{y_1} b_{2j}(\rho_2)\Big|y_1\Big)\Big] + Var_{y_1}\Big[E\Big(\sum_{j=1}^{y_1} b_{2j}(\rho_2)\Big|y_1\Big)\Big] + Var\Big[d_3\Big] \\
&= E_{y_2}\Big[y_2\rho_1(1-\rho_1)\Big] + Var_{y_2}\Big[y_2\rho_1\Big] + E_{y_1}\Big[y_1\rho_2(1-\rho_2)\Big] + Var_{y_1}\Big[y_1\rho_2\Big] + Var\Big[d_3\Big] \\
&= \rho_1(1-\rho_1)E_{y_1}\Big[y_1\Big] + \mu_2 - \rho_1\mu_1 + \rho_1^2 Var_{y_1}\Big[y_1\Big] \\
&= \rho_1(1-\rho_1)\mu_2 + \rho_1^2\mu_2 + \rho_2(1-\rho_2)\mu_1 + \rho_2^2\mu_1 + \mu_3 - \rho_1\mu_2 - \rho_2\mu_1 \\
&= \mu_3
\end{aligned}$$

In general, to obtain the $Var(y_t)$ for $t = 4, 5, ..., T$, we first find

$$Var\Big(y_t\Big|y_{t-1}, y_{t-2}, y_{t-3}\Big) = Var\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1) + \sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2) + \sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3) + d_t\Big|y_{t-1}, y_{t-2}, y_{t-3}\Big]$$

$$= Var\Big[\sum_{j=1}^{y_{t-1}} b_{1j}(\rho_1)\Big|y_{t-1}\Big] + Var\Big[\sum_{j=1}^{y_{t-2}} b_{2j}(\rho_2)\Big|y_{t-2}\Big]$$

$$+ Var\Big[\sum_{j=1}^{y_{t-3}} b_{3j}(\rho_3)\Big|y_{t-3}\Big] + Var\Big[d_t\Big]$$

$$= y_{t-1}\rho_1(1-\rho_1) + y_{t-2}\rho_2(1-\rho_2) + y_{t-3}\rho_3(1-\rho_3)$$

$$+ \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}$$

Then, using the conditional variance property we find

$$Var\Big(y_t\Big|y_{t-2}, y_{t-3}\Big)$$

$$= E_{y_{t-1}}\Big[Var\Big(y_t\Big|y_{t-1}, y_{t-2}, y_{t-3}\Big)\Big|y_{t-2}, y_{t-3}\Big] + Var_{y_{t-1}}\Big[E\Big(y_t\Big|y_{t-1}, y_{t-2}, y_{t-3}\Big)\Big|y_{t-2}, y_{t-3}\Big]$$

$$= E_{y_{t-1}}\Big[y_{t-1}\rho_1(1-\rho_1) + y_{t-2}\rho_2(1-\rho_2) + y_{t-3}\rho_3(1-\rho_3) + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}\Big|y_{t-2}, y_{t-3}\Big]$$

$$+ Var_{y_{t-1}}\Big[y_{t-1}\rho_1 + y_{t-2}\rho_2 + y_{t-3}\rho_3 + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}\Big|y_{t-2}, y_{t-3}\Big]$$

$$= \mu_{t-1}\rho_1(1-\rho_1) + y_{t-2}\rho_2(1-\rho_2) + y_{t-3}\rho_3(1-\rho_3) + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}$$

$$+ Var_{y_{t-1}}\Big[y_{t-1}\rho_1\Big] + 0$$

$$= \rho_1(1-\rho_1)\mu_{t-1} + y_{t-2}\rho_2(1-\rho_2) + y_{t-3}\rho_3(1-\rho_3) + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3} + \rho_1^2\sigma_{t-1,t-1}$$

Again, using the conditional variance property we find

$$Var\Big(y_t\Big|y_{t-3}\Big)$$

$$= E_{y_{t-2}}\Bigg[Var\Big(y_t\Big|y_{t-2}, y_{t-3}\Big)\Bigg|y_{t-3}\Bigg] + Var_{y_{t-2}}\Bigg[E\Big(y_t\Big|y_{t-2}, y_{t-3}\Big)\Bigg|y_{t-3}\Bigg]$$

$$= E_{y_{t-2}}\Bigg[\rho_1(1-\rho_1)\mu_{t-1} + y_{t-2}\rho_2(1-\rho_2) + y_{t-3}\rho_3(1-\rho_3) + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}$$

$$+ \rho_1^2\sigma_{t-1,t-1}\Bigg|y_{t-3}\Bigg] + Var_{y_{t-2}}\Bigg[\rho_1\mu_{t-1} + y_{t-2}\rho_2 + y_{t-3}\rho_3 + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}\Bigg|y_{t-3}\Bigg]$$

$$= \rho_1(1-\rho_1)\mu_{t-1} + \rho_2(1-\rho_2)\mu_{t-2} + y_{t-3}\rho_3(1-\rho_3) + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3} + \rho_1^2\sigma_{t-1,t-1}$$

$$+ 0 + Var_{y_{t-2}}\Bigg[y_{t-2}\rho_2\Bigg] + 0$$

$$= \rho_1(1-\rho_1)\mu_{t-1} + \rho_2(1-\rho_2)\mu_{t-2} + y_{t-3}\rho_3(1-\rho_3) + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}$$

$$+ \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2}$$

Next, by using the property that

$$Var(y_t) = E_{y_{t-3}}\Big[Var\Big(y_t\Big|y_{t-3}\Big)\Big] + Var_{y_{t-3}}\Big[E\Big(y_t\Big|y_{t-3}\Big)\Big]$$

we can obtain a recursive relation variance formula for t = 4,5,...,T.

$$
\begin{aligned}
\sigma_{t,t} &= Var(y_t) \\
&= E_{y_{t-3}}\Big[Var\Big(y_t\Big|y_{t-3}\Big)\Big] + Var_{y_{t-3}}\Big[E\Big(y_t\Big|y_{t-3}\Big)\Big] \\
&= E_{y_{t-3}}\Big[\rho_1(1-\rho_1)\mu_{t-1} + \rho_2(1-\rho_2)\mu_{t-2} + y_{t-3}\rho_3(1-\rho_3) + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3} \\
&\quad + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2}\Big] \\
&\quad + Var_{y_{t-3}}\Big[\rho_1\mu_{t-1} + \rho_2\mu_{t-2} + y_{t-3}\rho_3 + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}\Big] \\
&= \rho_1(1-\rho_1)\mu_{t-1} + \rho_2(1-\rho_2)\mu_{t-2} + \rho_3(1-\rho_3)\mu_{t-3} + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3} \\
&\quad + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2} + 0 + 0 + Var_{y_{t-3}}\Bigg[y_{t-3}\rho_3\Bigg] + 0 \\
&= \rho_1(1-\rho_1)\mu_{t-1} + \rho_2(1-\rho_2)\mu_{t-2} + \rho_3(1-\rho_3)\mu_{t-3} + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3} \\
&\quad + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2} + \rho_3^2\sigma_{t-3,t-3} \\
&= (\mu_t - \rho_1^2\mu_{t-1} - \rho_2^2\mu_{t-2} - \rho_3^2\mu_{t-3}) + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2} + \rho_3^2\sigma_{t-3,t-3}
\end{aligned}
$$

Hence, the recursive relation formula is

$Var(y_t) = (\mu_t - \rho_1^2\mu_{t-1} - \rho_2^2\mu_{t-2} - \rho_3^2\mu_{t-3}) + \rho_1^2\sigma_{t-1,t-1} + \rho_2^2\sigma_{t-2,t-2} + \rho_3^2\sigma_{t-3,t-3} \quad for \quad t = 4, 5, ..., T.$

Now, by taking variance successively we show that $\sigma_{t,t} = Var(y_t) = \mu_t = exp(\boldsymbol{x}_t'\boldsymbol{\beta})$ for all $t = 1, 2, ..., T$.

Co-variance and Correlation

To obtain the correlation structure of the response, we first compute the covariance between observations k distances apart.

At $t = 2$, from the properties of AR(1) Poisson process(Ed McKenzie, 1988),

$$\sigma_{2,1} = Cov(y_2, y_1) = \rho_1\mu_1 \quad and \quad Corr(y_2, y_1) = \frac{Cov(y_2, y_1)}{\sqrt{Var(y_2)}\sqrt{Var(y_1)}}$$

$$= \frac{\rho_1\mu_1}{\sqrt{\mu_2}\sqrt{\mu_1}}$$

$$= \rho_1\sqrt{\frac{\mu_1}{\mu_2}}.$$

At $t = 3$, from the properties of AR(2) Poisson process(C. Zhang and A. J. Oyet, 2014),

$$\sigma_{3,2} = Cov(y_3, y_2) = \rho_1\mu_2 + \rho_1\rho_2\mu_1 \quad and \quad Corr(y_3, y_2) = \frac{Cov(y_3, y_2)}{\sqrt{Var(y_3)}\sqrt{Var(y_2)}}$$

$$= \frac{\rho_1\mu_2 + \rho_1\rho_2\mu_1}{\sqrt{\mu_3}\sqrt{\mu_2}}$$

$$= \rho_1\sqrt{\frac{\mu_2}{\mu_3}} + \rho_1\rho_2\sqrt{\frac{\mu_1}{\mu_2}}.$$

$$\sigma_{3,1} = Cov(y_3, y_1) = \rho_1^2\mu_1 + \rho_2\mu_1 \quad and \quad Corr(y_3, y_1) = \frac{Cov(y_3, y_1)}{\sqrt{Var(y_3)}\sqrt{Var(y_1)}}$$

$$= \frac{(\rho_1^2 + \rho_2)\mu_1}{\sqrt{\mu_3}\sqrt{\mu_1}}$$

$$= (\rho_1^2 + \rho_2)\sqrt{\frac{\mu_1}{\mu_3}}.$$

Next, for $t = 4, 5, ..., T$, one can derive

$$\sigma_{t,t-k} = Cov(y_t, y_{t-k})$$

$$= E\Big(y_t, y_{t-k}\Big) - E\Big(y_t\Big)E\Big(y_{t-k}\Big)$$

$$= E_{y_{t-k}}\Big[E\Big(y_t y_{t-k}\Big|y_{t-k}\Big)\Big] - \mu_t\mu_{t-k}$$

$$= E_{y_{t-k}}\Big[y_{t-k}\Big(y_{t-1}\rho_1 + y_{t-2}\rho_2 + y_{t-3}\rho_3 + \mu_t - \rho_1\mu_{t-1} - \rho_2\mu_{t-2} - \rho_3\mu_{t-3}\Big)\Big] - \mu_t\mu_{t-k}$$

$$= \rho_1\Big[E\Big(y_{t-1}y_{t-k}\Big) - \mu_{t-1}\mu_{t-k}\Big] + \rho_2\Big[E\Big(y_{t-2}y_{t-k}\Big) - \mu_{t-2}\mu_{t-k}\Big]$$

$$+ \rho_3\Big[E\Big(y_{t-3}y_{t-k}\Big) - \mu_{t-3}\mu_{t-k}\Big]$$

$$= \rho_1 Cov(y_{t-1}, y_{t-k}) + \rho_2 Cov(y_{t-2}, y_{t-k}) + \rho_3 Cov(y_{t-3}, y_{t-k})$$

$$= \rho_1\sigma_{t-1,t-k} + \rho_2\sigma_{t-2,t-k} + \rho_3\sigma_{t-3,t-k}$$

Hence, for $t = 4, 5, ..., T$ we found that the covariances can be obtained from the recursive relation as follows.

$\sigma_{t,t-k} = Cov(Y_t, Y_{t-k}) = \rho_1\sigma_{t-1,t-k} + \rho_2\sigma_{t-2,t-k} + \rho_3\sigma_{t-3,t-k}$.

and

$$
\begin{aligned}
Corr(y_t, y_{t-k}) &= \frac{Cov(y_t, y_{t-k})}{\sqrt{Var(y_t)}\sqrt{Var(y_{t-k})}} \\
&= \frac{Cov(y_t, y_{t-k})}{\sqrt{\mu_t}\sqrt{\mu_{t-k}}} \\
&= \frac{\rho_1\sigma_{t-1,t-k} + \rho_2\sigma_{t-2,t-k} + \rho_3\sigma_{t-3,t-k}}{\sqrt{\mu_t\mu_{t-k}}} \\
&= \rho_1\sqrt{\frac{\mu_{t-1}}{\mu_t}}Corr(y_{t-1}, y_{t-k}) + \rho_2\sqrt{\frac{\mu_{t-2}}{\mu_t}}Corr(y_{t-2}, y_{t-k}) \\
&\quad + \rho_3\sqrt{\frac{\mu_{t-3}}{\mu_t}}Corr(y_{t-3}, y_{t-k}).
\end{aligned}
$$

3.2 Estimation of Parameters of the Non Stationary Poisson AR(3) Model

3.2.1 GQL Estimation of $\boldsymbol{\beta}$

In this section, we develop a GQL estimating equation for $\boldsymbol{\beta}$. Let $\boldsymbol{\mu} = (\mu_1, \mu_2, ...\mu_t, ..., \mu_T)'$ be the $T \times 1$ dimensional mean vector of $\boldsymbol{y} = (y_1, y_2, ..., y_t, ..., y_T)'$. It has been shown that $E(y_t) = \boldsymbol{\mu}_t = exp(\boldsymbol{x}_t'\boldsymbol{\beta})$. So the partial derivative of μ_t becomes $\partial\mu_t/\partial\boldsymbol{\beta} = \boldsymbol{x}_t\mu_t$. Define the covariate matrix as $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_t, ..., \boldsymbol{x}_T)'$ and the diagonal matrix of means as $\boldsymbol{U} = diag(\mu_1, \mu_2, ...\mu_t, ..., \mu_T)$. Now, the so-called marginal generalized quasi-likelihood (GQL) estimating equation for the parameter $\boldsymbol{\beta}$ can be written as

$$
\frac{\partial\mu'}{\partial\boldsymbol{\beta}}\boldsymbol{\Sigma}^{-1}(\rho)(\boldsymbol{y} - \mu) = \boldsymbol{X}'\boldsymbol{U}\boldsymbol{\Sigma}^{-1}(\rho)(\boldsymbol{y} - \mu) = 0. \tag{3.3}
$$

Then we solve the equation (3.2) by using the Newton-Raphson iterative approach produces the following expression.

$$\hat{\boldsymbol{\beta}}_{(r+1)} = \hat{\boldsymbol{\beta}}_{(r)} + \left[\boldsymbol{X}'\boldsymbol{U}\boldsymbol{\Sigma}^{-1}(\rho)\boldsymbol{U}'\boldsymbol{X}\right]^{-1}\left[\boldsymbol{X}'\boldsymbol{U}\boldsymbol{\Sigma}^{-1}(\rho)(\boldsymbol{y}-\mu)\right]\Bigg|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{(r)}} \tag{3.4}$$

where $\hat{\boldsymbol{\beta}}_{(r)}$ is the value of $\boldsymbol{\beta}$ at r^{th} iteration. Now we iterate the above expression (3.3) to get the convergence of $\boldsymbol{\beta}$ for fixed values of ρ_1, ρ_2 and ρ_3.

3.2.2 GMM estimation of ρ_1, ρ_2 and ρ_3

In order to obtain the estimate for $\boldsymbol{\beta}$ using above expression (3.3), ρ_1, ρ_2 and ρ_3 values must be known. Thus, we need to find good estimates for the values of ρ_1, ρ_2 and ρ_3. In this section, we will develop GMM estimating equations for the ρ values. We first define the standardized variance and the standardized lag 1, lag 2 and lag 3 auto-covariances of the responses as

$$\begin{aligned}
S_{tt} &= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{y_t-\mu_t}{\sigma_t}\right)^2, \\
S_{t,t+1} &= \frac{1}{T-1}\sum_{t=1}^{T-1}\left(\frac{y_t-\mu_t}{\sigma_t}\right)\left(\frac{y_{t+1}-\mu_{t+1}}{\sigma_{t+1}}\right), \\
S_{t,t+2} &= \frac{1}{T-2}\sum_{t=1}^{T-2}\left(\frac{y_t-\mu_t}{\sigma_t}\right)\left(\frac{y_{t+2}-\mu_{t+2}}{\sigma_{t+2}}\right), \\
S_{t,t+3} &= \frac{1}{T-3}\sum_{t=1}^{T-3}\left(\frac{y_t-\mu_t}{\sigma_t}\right)\left(\frac{y_{t+3}-\mu_{t+3}}{\sigma_{t+3}}\right).
\end{aligned} \tag{3.5}$$

Then we obtain the expectation of above covariances as follows.

$$\begin{aligned} E\big[S_{tt}\big] &= 1, \\ E\big[S_{t,t+1}\big] &= \frac{1}{T-1}\sum_{t=1}^{T-1}\rho_{t,t+1}, \\ E\big[S_{t,t+2}\big] &= \frac{1}{T-2}\sum_{t=1}^{T-2}\rho_{t,t+2}, \\ E\big[S_{t,t+3}\big] &= \frac{1}{T-3}\sum_{t=1}^{T-3}\rho_{t,t+3}, \end{aligned} \tag{3.6}$$

where

$$\rho_{t,t+k} = \frac{\sigma_{t,t+k}}{\sqrt{\sigma_{t,t}}\sqrt{\sigma_{t+k,t+k}}} \quad for \quad k = 1, 2, 3.$$

Thus, from the covariance of y_t we find

$$\begin{aligned} \rho_{t,t+1} &= \frac{\sigma_{t,t+1}}{\sqrt{\sigma_{t,t}}\sqrt{\sigma_{t+1,t+1}}} \\ &= \frac{\rho_1\sigma_{t,t} + \rho_2\sigma_{t-1,t} + \rho_3\sigma_{t-2,t}}{\sqrt{\sigma_{t,t}}\sqrt{\sigma_{t+1,t+1}}} \\ &= \rho_1\sqrt{\frac{\sigma_{t,t}}{\sigma_{t+1,t+1}}} + \rho_2\sqrt{\frac{\sigma_{t-1,t}}{\sigma_{t,t}}\frac{\sigma_{t-1,t}}{\sigma_{t+1,t+1}}} + \rho_3\sqrt{\frac{\sigma_{t-2,t}}{\sigma_{t,t}}\frac{\sigma_{t-2,t}}{\sigma_{t+1,t+1}}}, \end{aligned} \tag{3.7}$$

$$\begin{aligned} \rho_{t,t+2} &= \frac{\sigma_{t,t+2}}{\sqrt{\sigma_{t,t}}\sqrt{\sigma_{t+2,t+2}}} \\ &= \frac{\rho_1\sigma_{t,t+1} + \rho_2\sigma_{t,t} + \rho_3\sigma_{t-1,t}}{\sqrt{\sigma_{t,t}}\sqrt{\sigma_{t+2,t+2}}} \\ &= \rho_1\sqrt{\frac{\sigma_{t,t+1}}{\sigma_{t,t}}\frac{\sigma_{t,t+1}}{\sigma_{t+2,t+2}}} + \rho_2\sqrt{\frac{\sigma_{t,t}}{\sigma_{t+2,t+2}}} + \rho_3\sqrt{\frac{\sigma_{t-1,t}}{\sigma_{t,t}}\frac{\sigma_{t-1,t}}{\sigma_{t+2,t+2}}}, \end{aligned} \tag{3.8}$$

$$\rho_{t,t+3} = \frac{\sigma_{t,t+3}}{\sqrt{\sigma_{t,t}}\sqrt{\sigma_{t+3,t+3}}}$$
$$= \frac{\rho_1\sigma_{t,t+2} + \rho_2\sigma_{t,t+1} + \rho_3\sigma_{t,t}}{\sqrt{\sigma_{t,t}}\sqrt{\sigma_{t+3,t+3}}} \tag{3.9}$$
$$= \rho_1\sqrt{\frac{\sigma_{t,t+2}}{\sigma_{t,t}}\frac{\sigma_{t,t+2}}{\sigma_{t+3,t+3}}} + \rho_2\sqrt{\frac{\sigma_{t,t+1}}{\sigma_{t,t}}\frac{\sigma_{t,t+1}}{\sigma_{t+3,t+3}}} + \rho_3\sqrt{\frac{\sigma_{t,t}}{\sigma_{t+3,t+3}}}.$$

We also obtain the moment equations using first order approximation and assuming higher order terms are negligible.

$$\begin{aligned}
\frac{S_{t,t+1}}{S_{t,t}} &= E\left[\frac{S_{t,t+1}}{S_{t,t}}\right] \approx \frac{E[S_{t,t+1}]}{E[S_{t,t}]} = E[S_{t,t+1}],\\
\frac{S_{t,t+2}}{S_{t,t}} &= E\left[\frac{S_{t,t+2}}{S_{t,t}}\right] \approx \frac{E[S_{t,t+2}]}{E[S_{t,t}]} = E[S_{t,t+2}],\\
\frac{S_{t,t+3}}{S_{t,t}} &= E\left[\frac{S_{t,t+3}}{S_{t,t}}\right] \approx \frac{E[S_{t,t+3}]}{E[S_{t,t}]} = E[S_{t,t+3}].
\end{aligned} \tag{3.10}$$

Now, we get three simultaneous equations by substituting (3.5) into (3.9)

$$\begin{aligned}
\frac{S_{t,t+1}}{S_{t,t}} &= \frac{\rho_1}{T-1}\sum_{t=1}^{T-1}\sqrt{\frac{\sigma_{t,t}}{\sigma_{t+1,t+1}}} + \frac{\rho_2}{T-2}\sum_{t=2}^{T-1}\sqrt{\frac{\sigma_{t-1,t}}{\sigma_{t,t}}\frac{\sigma_{t-1,t}}{\sigma_{t+1,t+1}}} + \frac{\rho_3}{T-3}\sum_{t=3}^{T-1}\sqrt{\frac{\sigma_{t-2,t}}{\sigma_{t,t}}\frac{\sigma_{t-2,t}}{\sigma_{t+1,t+1}}},\\
\frac{S_{t,t+2}}{S_{t,t}} &= \frac{\rho_1}{T-2}\sum_{t=1}^{T-2}\sqrt{\frac{\sigma_{t,t+1}}{\sigma_{t,t}}\frac{\sigma_{t,t+1}}{\sigma_{t+2,t+2}}} + \frac{\rho_2}{T-2}\sum_{t=1}^{T-2}\sqrt{\frac{\sigma_{t,t}}{\sigma_{t+2,t+2}}} + \frac{\rho_3}{T-3}\sum_{t=2}^{T-2}\sqrt{\frac{\sigma_{t-1,t}}{\sigma_{t,t}}\frac{\sigma_{t-1,t}}{\sigma_{t+2,t+2}}},\\
\frac{S_{t,t+3}}{S_{t,t}} &= \frac{\rho_1}{T-1}\sum_{t=1}^{T-1}\sqrt{\frac{\sigma_{t,t+2}}{\sigma_{t,t}}\frac{\sigma_{t,t+2}}{\sigma_{t+3,t+3}}} + \frac{\rho_2}{T-2}\sum_{t=2}^{T-1}\sqrt{\frac{\sigma_{t,t+1}}{\sigma_{t,t}}\frac{\sigma_{t,t+1}}{\sigma_{t+3,t+3}}} + \frac{\rho_3}{T-3}\sum_{t=3}^{T-1}\sqrt{\frac{\sigma_{t,t}}{\sigma_{t+3,t+3}}}
\end{aligned} \tag{3.11}$$

which are solved simultaneously to obtain GMM estimates for ρ_1, ρ_2 and ρ_3. The procedure is then iterated to reach convergence.

3.3 Simulation Study

In this section, we carry out simulation studies to examine the performance of the proposed GQL and GMM estimation methods.

3.3.1 Data Generation

We first set the number of responses $T = 250$ in model (3.1). We considered the number of time dependent covariates $p = 2$ which are defined as $\boldsymbol{x}_t' = (x_{t1}, x_{t2})$ with x_{t1} and x_{t2}) given by

$$x_{t1} = \begin{cases} 0.1, & t = 1, 2, ..., 50 \\ 0.5, & t = 51, 52, ..., 150 \\ 1, & t = 151, 152, ..., 250 \end{cases} \tag{3.12}$$

and

$$x_{t2} = \left(0.5 + \frac{(t-1)0.5}{T}\right), \quad t = 1, 2, ..., 250 \tag{3.13}$$

It is clear that, the mean μ_t of the process (3.1) will be non-stationary since the covariates are time dependent and the mean is in the form of $\mu_t = exp(\boldsymbol{x}_t'\boldsymbol{\beta})$.

Using the above defined covariate vectors and fixed combination of parameter $\boldsymbol{\beta}$ values, we first compute the mean μ_t for $t = 1, 2, ..., 250$ and then choose ρ_1, ρ_2 and ρ_3 which satisfying the condition (3.2) and $\rho_1 \geq \rho_2 \geq \rho_3 \geq 0$. Then we generate y_1 from the Poisson distribution $Poi(\mu_1)$ and using that we can generate all the y_t for $t = 2, 3, ..., 250$ from model (3.1). After we generate all the y_t values, we assign suitable initial values for

$\boldsymbol{\beta}$, ρ_1, ρ_2 and ρ_3. Then, we iterate (3.4) and the solutions we got by solving simultaneous equations set (3.11) to convergence of $\boldsymbol{\beta}$, ρ_1, ρ_2 and ρ_3. We repeated the same procedure 1,000 times for fixed values of $\boldsymbol{\beta}$, ρ_1, ρ_2 and ρ_3. Then, the above procedure was repeated to test for different combination of $\boldsymbol{\beta}$ values and ρ values.

3.3.2 Discussion of the results

For each of the above cases, the average (Simulated Mean-SM) and the standard error (Simulated Standard Error-SSE) of the estimates $\boldsymbol{\beta}$, ρ_1, ρ_2 and ρ_3 from 1,000 simulations are shown in Tables 3.1. In each different combinations, we also found true value of standard error (TSE) of $\boldsymbol{\beta}$ and estimated standard error (ESE) of $\boldsymbol{\beta}$. The results seem, in general, the proposed GQL and GMM methods have performed well in estimating the parameters of the non-stationary AR(3) Poisson process in each cases. However, the moment estimates for the correlation parameters seems less accurate.

Furthermore, bias of estimation of $\boldsymbol{\beta}$ values seems to be high for small values of $\boldsymbol{\beta}$ and bias of estimation of ρ values seems to be low for small values of ρ.

,

,

Table 3.1: GQL estimates of $\boldsymbol{\beta}$ and GMM estimates of ρ_1, ρ_2 and ρ_3 of the process (3.1) and their standard errors obtained from 1,000 Simulations based on time dependent covariates $\boldsymbol{x}_t' = (x_{t1}, x_{t2})$ in (3.12) and (3.13).

β_1	β_2	ρ_1	ρ_2	ρ_3		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$
0.3	0.5	0.2	0.15	0.1	SM	0.362	0.573	0.151	0.116	0.097
					SSE	0.628	1.506	0.062	0.060	0.057
					ESE	0.634	1.472			
					TSE	0.629	1.537			
0.5	1	0.2	0.15	0.1	SM	0.556	0.969	0.146	0.117	0.099
					SSE	0.523	1.271	0.062	0.063	0.055
					ESE	0.489	1.189			
					TSE	0.545	1.326			
1	1	0.2	0.15	0.1	SM	1.019	0.989	0.155	0.118	0.106
					SSE	0.456	1.111	0.066	0.064	0.060
					ESE	0.414	1.010			
					TSE	0.450	1.096			
0.3	0.5	0.25	0.2	0.1	SM	0.378	0.546	0.199	0.192	0.099
					SSE	0.667	1.672	0.067	0.062	0.052
					ESE	0.603	1.619			
					TSE	0.696	1.701			
0.5	1	0.25	0.2	0.1	SM	0.577	0.956	0.189	0.152	0.102
					SSE	0.552	1.346	0.069	0.070	0.056
					ESE	0.528	1.287			
					TSE	0.605	1.473			
1	1	0.25	0.2	0.1	SM	1.026	0.967	0.195	0.155	0.109
					SSE	0.510	1.244	0.068	0.070	0.060
					ESE	0.453	1.104			
					TSE	0.503	1.227			
0.3	0.5	0.3	0.2	0.1	SM	0.387	0.537	0.247	0.163	0.102
					SSE	0.755	1.748	0.066	0.068	0.062
					ESE	0.665	1.699			
					TSE	0.730	1.786			
0.5	1	0.3	0.2	0.1	SM	0.582	0.964	0.200	0.142	0.109
					SSE	0.591	1.432	0.060	0.061	0.053
					ESE	0.532	1.295			
					TSE	0.637	1.551			
1	1	0.3	0.2	0.1	SM	0.985	1.027	0.202	0.140	0.120
					SSE	0.554	1.342	0.062	0.064	0.057
					ESE	0.455	1.108			
					TSE	0.532	1.298			

Chapter 4

Concluding Remarks

,

In this chapter a summary of the main findings is presented in here, with the suggestions for future research. This thesis has focused on the Poisson AR and MA models for time series of count data. The patterns and order identification of continuous time series has been studied extensively by several authors. To the best of our knowledge, there are no studies have been done regarding this for count time series models. In Chapter 2, we have dealt with the patterns and order identification of stationary Poisson AR and MA Models for Count Time Series. We have considered stationary Poisson AR and MA models upto order 3 with Poisson marginal distribution.

In order to find the theoretical patterns of the models that we considered, we first discussed about the model structure and the basic properties (mean, variance, covariance and correlation) of stationary Poisson MA models up to order q and stationary Poisson AR models up to order 2 which were already proposed in literature. We proposed a model for stationary Poisson AR(3) process and derived it's basic properties. We also proved that the stationarity conditions of the proposed model are satisfied.

It is often very useful to plot the ACF and PACF against consecutive time lags to determine the order of AR and MA terms. For this purpose, we discussed ACF of stationary Poisson MA and AR processes up to order 3 and we derived PACF of stationary Poisson MA and AR processes up to order 2 as the form of PACF of higher order models are complicated. The Table 4.1 summarizes the theoretical patterns we found of ACF and PACF of stationary Poisson AR and MA processes.

Table 4.1: Patterns of theoretical ACF and PACF of stationary Poisson AR and MA models.

Process	**Theoretical ACF**	**Theoretical PACF**
MA(1)	Cuts off after lag 1.	Decays exponentially depending on the magnitude of the correlation parameter ρ.
MA(2)	Cuts off after lag 2.	Decays exponentially depending on the magnitude of the correlation parameters ρ_1 and ρ_2.
MA(3)	Cuts off after lag 3.	Decays exponentially depending on the magnitude of the correlation parameters ρ_1, ρ_2 and ρ_3.
AR(1)	Decays exponentially depending on the magnitude of the correlation parameter ρ.	Cuts off after lag 1.
AR(2)	Decays exponentially depending on the magnitude of the correlation parameters ρ_1 and ρ_2.	Cuts off after lag 2.
AR(3)	Decays exponentially depending on the magnitude of the correlation parameters ρ_1, ρ_2 and ρ_3.	Cuts off after lag 3.

We carried out simulation studies to examine the accuracy of the above patterns of ACF and PACF in each of the cases. To test whether the SACF and SPACF is significantly different from zero, we used Jackknife method to find the standard error of SACF and SPACF. For some different specific values of correlation parameters, we compared the Jackknife standard errors with the standard errors of SACF and SPACF for the continuous case and also with the simulated standard errors from 10,000 simulations and we found that the standard errors from all three methods are approximately same. The ACF and PACF plots obtained from simulation studies revealed that the patterns in the ACF and PACF for time series of Poisson counts that are similar to that of AR and MA models for Gaussian processes.

Chapter 3 dealt with the third order auto-regressive dynamic model for non stationary count time series. We proposed a model for non-stationary Poisson AR(3) process. By deriving the basic properties mean, variance, covariance and correlation we have verified that the properties satisfy non stationary conditions. We estimated the model parameters using GQL and the correlation parameters using GMM estimation methods. Finally, we examined the accuracy of estimates through a simulation study. Results from the simulation studies have shown that the proposed GQL and GMM methods have performed well in estimating the parameters. However, the moment estimates for the correlation parameters were less accurate.

In future, we wish to apply the non-stationary Poisson AR(3) estimation theory to a real-life data set and generalize the Poisson model identification for AR(p) process. In this thesis we derived patterns for model identification of any stationary Poisson AR(3) process. To the best of our knowledge, we note that these patterns have not been derived for AR(p), where $p > 3$ process. This is an open problem for future study. Specifically, the stationary Poisson AR(p) model can be written as (Alzaid and Al-Osh, 1990),

$$y_t = \sum_{i=1}^{p} \rho_i * y_{t-i} + d_t \quad \text{for} \quad t = 1, 2, ..., T. \tag{4.1}$$

with some assumptions similar to the ones given in section 2.2.3. It can be shown that the mean and variance of model (4.1) are $E(y_t) = \mu$ and $Var(y_t) = \mu$ respectively. Proceeding as before, we can show that the lag k auto covariance of model (4.1) is as follows.

$$\sigma_{t,t-k} = Cov(Y_t, Y_{t-k}) = \sum_{i=1}^{p} \rho_i \sigma_{t-i,t-k}.$$

The challenge will be in deriving a closed form expression for the lag k auto covariance.

Bibliography

[1] Al-Osh, M. and Alzaid, A. (1988). Integer-valued moving average (inma) process. *Statistical Papers*, 29(1):281–300.

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

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

[4] Alzaid, A. and Al-Osh, M. (1990). An integer-valued pth-order autoregressive structure (inar (p)) process. *Journal of Applied Probability*, 27(2):314–324.

[5] Alzaid, A. and Al-Osh, M. (1993). Some autoregressive moving average processes with generalized poisson marginal distributions. *Annals of the Institute of Statistical Mathematics*, 45(2):223–232.

[6] Bakouch, H. S. and Ristić, M. M. (2010). Zero truncated poisson integer-valued ar (1) model. *Metrika*, 72(2):265–280.

[7] Barron, D. N. (1992). The analysis of count data: Overdispersion and autocorrelation. *Sociological methodology*, pages 179–220.

[8] Böckenholt, U. (1998). Mixed inar (1) poisson regression models: Analyzing heterogeneity and serial dependencies in longitudinal count data. *Journal of Econometrics*, 89(1-2):317–338.

[9] Boslaugh, S. (2012). *Statistics in a nutshell: A desktop quick reference.* " O'Reilly Media, Inc.".

[10] Box, G. E. and Jenkins, G. M. (1976). Time series analysis, control, and forecasting. *San Francisco, CA: Holden Day*, 3226(3228):10.

[11] Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control.* John Wiley & Sons.

[12] Bozdogan, H. (1987). Model selection and akaike's information criterion (aic): The general theory and its analytical extensions. *Psychometrika*, 52(3):345–370.

[13] Crowder, M. (1995). On the use of a working correlation matrix in using generalised linear models for repeated measures. *Biometrika*, 82(2):407–410.

[14] de Gooijer, J. G., Abraham, B., Gould, A., and Robinson, L. (1985). Methods for determining the order of an autoregressive-moving average process: A survey. *International Statistical Review/Revue Internationale de Statistique*, pages 301–329.

[15] Freeland, R. K. and McCabe, B. (2005). Asymptotic properties of cls estimators in the poisson ar (1) model. *Statistics & probability letters*, 73(2):147–153.

[16] Garel, B. and Hallin, M. (1999). Rank-based autoregressive order identification. *Journal of the American Statistical Association*, 94(448):1357–1371.

[17] Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054.

[18] Hoff, J. C. (1983). *A practical guide to Box-Jenkins forecasting*. Lifetime Learning Publications.

[19] Ispány, M., Pap, G., and van Zuijlen, M. C. (2003). Asymptotic inference for nearly unstable inar (1) models. *Journal of Applied Probability*, 40(3):750–765.

[20] Jung, R. C. and Tremayne, A. (2006). Binomial thinning models for integer time series. *Statistical Modelling*, 6(2):81–96.

[21] Koreisha, S. G. and Pukkila, T. (1995). A comparison between different order-determination criteria for identification of arima models. *Journal of Business & Economic Statistics*, 13(1):127–131.

[22] Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22.

[23] McKenzie, E. (1985). Some simple models for discrete variate time series 1. *JAWRA Journal of the American Water Resources Association*, 21(4):645–650.

[24] McKenzie, E. (1988). Some arma models for dependent sequences of poisson counts. *Advances in Applied Probability*, 20(4):822–835.

[25] Oyet, A. J. and Sutradhar, B. C. (2013). Longitudinal modeling of infectious disease. *Sankhya B*, 75(2):319–342.

[26] Ozaki, T. (1977). On the order determination of arima models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 26(3):290–301.

[27] Pankratz, A. (2009). *Forecasting with univariate Box-Jenkins models: Concepts and cases*, volume 224. John Wiley & Sons.

[28] Park, Y. and Oh, C. W. (1997). Some asymptotic properties in inar (1) processes with poisson marginals. *Statistical Papers*, 38(3):287.

[29] Pukkila, T., Koreisha, S., and Kallinen, A. (1990). The identification of arma models. *Biometrika*, 77(3):537–548.

[30] Steutel, F. and Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability*, pages 893–899.

[31] Sutradhar, B. C. and Bari, W. (2007). On generalized quasilikelihood inference in longitudinal mixed model for count data. *Sankhyā: The Indian Journal of Statistics*, pages 671–699.

[32] Sutradhar, B. C. and Das, K. (1999). Miscellanea. on the efficiency of regression estimators in generalised linear models for longitudinal data. *Biometrika*, 86(2):459–465.

[33] Sutradhar, B. C., Jowaheer, V., and Sneddon, G. (2008). On a unified generalized quasi–likelihood approach for familial–longitudinal non-stationary count data. *Scandinavian Journal of Statistics*, 35(4):597–612.

[34] Vandaele, W. (1983). Applied time series and box-jenkins models. Technical report.

[35] Wasserman, L. (2006). *All of nonparametric statistics*. Springer Science & Business Media.

[36] Wedderburn, R. W. (1974). Quasi-likelihood functions, generalized linear models, and the gaussnewton method. *Biometrika*, 61(3):439–447.

[37] Wei, W. W. (2006). Time series analysis. In *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2*.

[38] Weiß, C. H. (2007). Controlling correlated processes of poisson counts. *Quality and reliability engineering international*, 23(6):741–754.

[39] Zhang, C. and Oyet, A. J. (2014). Second order longitudinal dynamic models with covariates: estimation and forecasting. *Metrika*, 77(7):837–859.