INFERENCES FOR INTEGER-VALUED TIME SERIES MODELS

NURUL NAJIHAH MOHAMAD

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

INFERENCES FOR INTEGER-VALUED TIME SERIES MODELS

NURUL NAJIHAH MOHAMAD

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: NURUL NAJIHAH BINTI MOHAMAD

Registration/Matrix No.:SHB 120008

Name of Degree:DOCTOR OF PHILOSOPHY (MATHEMATICS & SCIENCE PHILOSOPHY)

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

INFERENCES FOR INTEGER-VALUED TIME SERIES MODELS

Field of Study:STATISTICAL TIME SERIES (STATISTICS)

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date

Subscribed and solemnly declared before,

Witness's Signature Date

Name:
Designation:

ABSTRACT

Recently there has been a growing interest in integer-valued volatility models. The need for such time series models arises in different areas including biomedicine, insurance and finance. Here, we look at a class of integer-valued GARCH time series models which are of interest to the practitioners. The models are assuming the form of GARCH model such that the conditional distribution of the process follows one of the following distributions; Poisson, negative binomial and zero-inflated Poisson.

In this study, a general theorem on the moment properties of the class of integer-valued volatility models is derived using martingale transformation with much simpler proofs. We show the first two moments obtained in the recent literature as special cases. In addition, we derive the closed form expressions of the kurtosis and skewness formula for these three models. The results are very useful in understanding the behaviour of the processes.

We then estimate the parameters of the class of integer-valued volatility models via the quadratic estimating functions theory. Specifically, the optimal estimating functions for each process are derived. Through a finite sample size investigation, we compare the performance of the quadratic estimating functions (QEF) method with the maximum likelihood and estimating functions (EF) methods. We show that the quadratic estimating functions method performs better in terms of unbiasness and mean square error. For illustration, we fit the models on real data sets.

ABSTRAK

Sejak kebelakangan ini, minat terhadap model-model volatiliti siri masa taburan data bilang telah bertambah. Keperluan kepada model siri masa tersebut telah meningkat di dalam pelbagai bidang termasuk biomedik, insurans dan kewangan. Di sini, kami menumpukan kajian terhadap satu kumpulan model GARCH yang bertaburan data bilang yang menjadi minat untuk pengguna. Model-model ini mengikuti bentuk model GARCH dimana taburan sandarannya mengikut salah satu daripada taburan berikut; Poisson, negatif binomial dan Poisson sifar melambung.

Dalam kajian ini, kami menerbitkan teorem am berkaitan dengan sifat-sifat momen untuk kelas model volatiliti siri masa taburan data bilang dengan pembuktian yang lebih mudah dengan menggunakan transformasi martingale. Dua sifat momen pertama yang boleh didapati di dalam kesusasteraan telah dibuktikan sebagai kes khas. Tambahan pula, kami menerbitkan formula bentuk tertutup untuk ukuran kurtosis dan kepencongan bagi ketiga-tiga model tersebut. Penemuan formula ini sangat berguna untuk memahami sifat-sifat proses tersebut.

Kami kemudiannya menganggar parameter untuk model volatiliti tersebut melalui teori fungsi anggaran kuadratik (QEF). Secara spesifiknya, kami menerbitkan fungsi anggaran optimal bagi setiap model. Melalui kajian saiz sampel terhingga, kami membandingkan prestasi kaedah QEF dengan kaedah kebolehjadian maksimum (MLE) dan kaedah fungsi anggaran linear (LEF). Untuk ilustrasi, kami memadankan model tersebut dengan data sebenar.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude and appreciation to my supervisors, Prof. Dr. Ibrahim bin Mohamed, Prof. Dr. Aerambamoothy Thavaneswaran and Prof. Dr. Mohd Sahar bin Yahya for their continuous encouragement and supervision throughout my study.

Secondly, special thank goes to my beloved husband, Mohd Zulkarnain bin Shamsuddin for his continuous support and encouragement at all stages of this study. To my parents and parents in law, Hj Mohamad bin Mahmud, Hjh Saudah binti Salleh, Shamsudin bin Hassan and Siti Hazarah binti Salleh, I would like to express my gratitude for their continuous love and prayers. Not forgotten my beloved daughters, Nur Aqilah Batrisya and Nur Aqilah Basyasya. With their blessing and prayers, I am very grateful to eventually complete this study.

Thirdly, I would like to express my appreciation to the International Islamic University Malaysia (IIUM) and Ministry of Higher Education (MOHE) for the opportunity to further my study in the field of statistics.

Finally, to all my friends (in particular, Huda Zuhrah Abd Halim, Adzhar Rambli, Lily Wong and Nadhirah Abd Halim) and staff members of the Institute of Mathematical Sciences, University of Malaya (paricularly Dr Ng Kok Haur and Puan Budiyah binti Yeop), I would like to say thank you so much for their assistance and support.

TABLE OF CONTENTS

Abstract iii

Abstrak iv

Acknowledgements v

Table of Contents vi

List of Figures x

List of Tables xi

List of Symbols and Abbreviations xiv

CHAPTER 1: INTRODUCTION 1

1.1 Background of the study 1

1.2 Statement of the problem 5

1.3 Objectives 5

1.4 Significance of research 5

1.5 Thesis outline 6

CHAPTER 2: LITERATURE REVIEW 7

2.1 Introduction 7

2.2 Development of INGARCH(p,q) Model 8

2.2.1 ARCH model 8

2.2.2 GARCH (p,q) Model 9

2.3 The Development of Quadratic Estimating Functions 10

2.3.1 Parameter Estimation 10

2.3.2 The Estimating Function 12

2.3.3 Derivation of Optimum EF 13

CHAPTER 3: MOMENT PROPERTIES OF SOME INTEGER-VALUED TIME SERIES MODEL 16

3.1 Introduction 16

3.2 The Class of Integer-Valued GARCH Models 16
3.3 First and Second Moments of The Model 17
3.4 Skewness and Kurtosis 19
3.4.1 Example on $p = 1$ and $q = 1$ 20
3.5 Feigin's Theorem on Stationary Distribution Of λ_t and X_t 23
3.6 Summary of The Chapter 24

CHAPTER 4: THE QUADRATIC ESTIMATING FUNCTIONS 25
4.1 Introduction 25
4.1.1 General Model and Method 26
4.2 Zero-inflated Model 28
4.2.1 Basic zero-Inflated Poisson Model 29
4.2.2 Zero-Inflated Poisson Regression Model 31
4.3 Summary of The Chapter 33

CHAPTER 5: INGARCH(p,q) MODEL 34
5.1 Introduction 34
5.2 The Moment Properties 34
5.2.1 The Moment Properties of Conditional Distribution of INGARCH(p,q) Model 35
5.2.2 The Moment Properties of Unconditional Distribution of INGARCH(p,q) Model 36
5.2.3 Empirical Study 38
5.3 Quadratic Estimating Functions on INGARCH(p,q) Model 38
5.4 Performance of The Estimation Method in INGARCH $(1,1)$ 42
5.4.1 Conditional LS Derivation of INGARCH $(1,1)$ 42
5.4.2 MLE Derivation of INGARCH $(1,1)$ 43
5.4.3 QEF of INGARCH $(1,1)$ 43
5.4.4 Simulation Study 44
5.4.5 The Result 45

5.5 Real Example ... 56

5.6 Summary of The Chapter ... 59

CHAPTER 6: NBINGARCH(p,q) MODEL ... 60

6.1 Introduction ... 60

6.2 The Moment Properties ... 60

6.2.1 The Moments of Conditional Distribution in NBINGARCH(p,q) Model ... 60

6.2.2 The Moments Properties of Unconditional Distribution in NBINGARCH(p,q) Model ... 63

6.2.3 Empirical Study ... 64

6.3 Quadratic Estimating Functions on NBINGARCH(p,q) Model ... 65

6.4 Performance of The Estimation Methods in NBINGARCH $(1,1)$ ... 69

6.4.1 MLE Derivation of NBINGARCH $(1,1)$ ... 70

6.4.2 EF Derivation of NBINGARCH $(1,1)$ ... 70

6.4.3 QEF Derivation of NBINGARCH $(1,1)$ ... 71

6.4.4 Simulation Study ... 71

6.4.5 The Result ... 71

6.5 Real Example ... 72

6.6 Summary of The Chapter ... 85

CHAPTER 7: ZIPINGARCH(p,q) MODEL ... 86

7.1 Introduction ... 86

7.2 The Moment Properties ... 87

7.2.1 The Moment Properties of Conditional Distribution in ZIPINGARCH(p,q) Model ... 87

7.2.2 The Moment Properties of Unconditional Distribution on ZIPINGARCH(p,q) Model ... 89

7.2.3 Empirical Study ... 91

7.3 Quadratic Estimating Functions on ZIPINGARCH(p,q) Model ... 92

7.4 Performance of The Estimation Methods on ZIPINGARCH $(1,1)$ ... 96

7.4.1 MLE Derivation of ZIPINGARCH $(1,1)$ 96

7.4.2 EF Derivation of ZIPINGARCH $(1,1)$ 96

7.4.3 QEF Derivation of ZIPINGARCH $(1,1)$ 97

7.4.4 Simulation Study 97

7.4.5 The Result 97

7.5 Real Example 109

7.6 Summary of The Chapter 111

CHAPTER 8: CONCLUSION **112**

8.1 Summary of The Study 112

8.2 Further Research 113

References **115**

Appendices 120

LIST OF FIGURES

Figure 5.1: The monthly strike data from January 1994 to December 2002 56

Figure 5.2: (a)The Pearson residual plot. (b)The periodagram plot 58

Figure 6.1: The Polio data in the United State from 1970 to 1983 83

Figure 6.2: (a)The Pearson residual plot. (b)The periodagram plot 85

Figure 7.1: The monthly counts of arson in the 13th police car beat plus in Pittsburgh from January 1990 until December 2001 109

Figure 7.2: (a)The Pearson residual plot. (b)The periodagram plot 111

LIST OF TABLES

Table 5.1: Generated data and true values for the moment structures with $\gamma = 0.1$ 38

Table 5.2: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.2$, and $\beta_1 = 0.3$ 46

Table 5.3: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.4$, and $\beta_1 = 0.1$ 47

Table 5.4: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1$, and $\beta_1 = 0.4$ 48

Table 5.5: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.4$, and $\beta_1 = 0.2$ 49

Table 5.6: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.5, \alpha_1 = 0.2$, and $\beta_1 = 0.3$ 50

Table 5.7: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.6$, and $\beta_1 = 0.3$ 51

Table 5.8: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.7$, and $\beta_1 = 0.2$ 52

Table 5.9: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.8$, and $\beta_1 = 0.1$ 53

Table 5.10: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1$, and $\beta_1 = 0.8$ 54

Table 5.11: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.4, \alpha_1 = 0.3$, and $\beta_1 = 0.6$ 55

Table 5.12: The estimated parameter of INGARCH$(1,1)$ model 57

Table 5.13: AIC and BIC for INARCH(1) and INGARCH $(1,1)$ models. 57

Table 5.14: Diagnostics for INGARCH(1,1) model 58

Table 6.1: Generated data and true values for the moment structures with $\gamma = 0.1$ 65

Table 6.2: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.2$, and $\beta_1 = 0.3$ 73

Table 6.3: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.4$, and $\beta_1 = 0.1$ 74

Table 6.4: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1,$ and $\beta_1 = 0.4$ 75

Table 6.5: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.4,$ and $\beta_1 = 0.2$ 76

Table 6.6: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.5, \alpha_1 = 0.2,$ and $\beta_1 = 0.3$ 77

Table 6.7: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.6,$ and $\beta_1 = 0.3$ 78

Table 6.8: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.7,$ and $\beta_1 = 0.2$ 79

Table 6.9: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.8,$ and $\beta_1 = 0.1$ 80

Table 6.10: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1,$ and $\beta_1 = 0.8$ 81

Table 6.11: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.4, \alpha_1 = 0.3,$ and $\beta_1 = 0.6$ 82

Table 6.12: AIC and BIC values for NBINGARCH$(1,1)$ 83

Table 6.13: The estimated parameter of NBINGARCH$(1,1)$ model 83

Table 6.14: AIC and BIC for INGARCH $(1,1)$ and NBINGARCH $(1,1)$ models. 84

Table 6.15: Diagnostics for NBINGARCH(1,1) model 85

Table 7.1: Generated data and true values for the moment structures with $\gamma = 0.1$ 91

Table 7.2: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.2,$ and $\beta_1 = 0.3$ 99

Table 7.3: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.4,$ and $\beta_1 = 0.1$ 100

Table 7.4: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1,$ and $\beta_1 = 0.4$ 101

Table 7.5: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.4,$ and $\beta_1 = 0.2$ 102

Table 7.6: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.5, \alpha_1 = 0.2,$ and $\beta_1 = 0.3$ 103

Table 7.7: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.6,$ and $\beta_1 = 0.3$ 104

Table 7.8: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.7,$ and $\beta_1 = 0.2$ 105

Table 7.9: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.8,$ and $\beta_1 = 0.1$ 106

Table 7.10: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1,$ and $\beta_1 = 0.8$ 107

Table 7.11: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.4, \alpha_1 = 0.3,$ and $\beta_1 = 0.6$ 108

Table 7.12: The estimated parameter of ZIPINGARCH$(1,1)$ model 110

Table 7.13: Diagnostics for ZIPINGARCH(1,1) model 110

LIST OF SYMBOLS AND ABBREVIATIONS

X_t The time series process

μ The mean

σ^2 The variance

γ Autocovariance

ρ Autocorellation

Γ Skewness

$Corr(X_t, X_s)$ The correlation

$Cov(X_t, X_s)$ The covariance

K Kurtosis

$\Im_{t-1}^{X}$ The $\sigma-$field generated by $X_{t-1}, X_{t-2}, \ldots, X_1$

$\gamma_0, \alpha_i, \beta_j$ Parameters in the model

λ_t Intensity parameter

ω Inflation parameter

$\mathbf{g}_Q^*(\boldsymbol{\theta})$ Optimal QEF

$\mathbf{I}_{\mathbf{g}_Q^*}(\boldsymbol{\theta})$ Information for the optimal QEF

a_{t-1}^* and b_{t-1}^* Coefficient in QEF method

m_t, s_t, h_t, u_t martingale differences

$\langle m_t \rangle$ variance of m_t

$\langle s_t \rangle$ variance of s_t

$\langle m, s \rangle_t$ covariance of m_t and s_t

$IVTS$ Integer-valued time series

$INAR$ Integer-valued Autoregressive

$ARCH$ Autoregressive conditional heteroskedasticity

$GARCH$ Generalized Autoregressive conditional heteroskedasticity

INGARCH Integer-valued GARCH

NBINGARCH Negative binomial integer-valued GARCH

ZIPINGARCH Zero-inflated Poisson integer-valued GARCH

EF Estimating functions

QEF Quadratic Estimating functions

LIK Likelihood

CHAPTER 1

INTRODUCTION

1.1 Background of the study

In the real world, we are bounded with time and space. In order to understand the events or incidents around us, observations are frequently made sequentially over time, for example, the yearly dengue cases, monthly unemployment figures, daily money exchange rates and share prices. These data are known as time series data where the observations $\{X_t\}$ are being recorded at specific times. To be specific, according to Brockwell & Davis (1991), a time series model can be defined as a specification of the joint distributions (or possibly only the means and covariances) of a sequence of random variables $\{X_t\}$ of which $\{x_t\}$ is postulated to be a realization.

In the past, various time series models have been proposed in the literature. The well-known time series model is the autoregressive integrated moving average model (ARIMA) which can be identified by the Box-Jenkins methodology. The model is developed by looking at several important concepts including stationary condition, the autocorrelation function and white noise process. The stationary properties depends on the first and second order moments of X_t. Generally, we have two types of stationary properties, which are called weakly stationarity and strictly stationarity. The time series $\{X_t\}$ is said to be weakly stationary if the mean function of X_t, say $\mu_X(t)$, is independent of t and the covariance function, denoted by $\gamma_X(t+h,t)$, is independent of t for each integer h. Therefore, we can say that a weakly stationary time series has the mean and variance constant over time. Meanwhile, strict stationarity of a time series is defined by the condition that $(X_1,\ldots,X_n)$ and $(X_{1+h},\ldots,X_{n+h})$ have the same joint distributions for all integers h and $n > 0$. For autocorrelation property, we are measuring the correlation of the observation at time t with the kth past observations where k=1,2,.... is called a lag k autocorrelation for the process. It is very useful in the determination of the order of ARIMA models and can also be used to indicate a white noice process. A time series process is said to be white noise if the correlation between $\{X_t\}$ and $\{X_s\}$ is zero for all $t \neq s$, that is,

$$Corr(X_t X_s) = 0 \text{ for all } t \neq s.$$

In real life applications, we frequently have time series data with nonconstant variance especially in finance and medicine. Examples include stock price in Baillie & Bollerslev (2002) and campylabacterosis cases in Ferland et al. (2006). Hence, there is a strong need to develop non-linear time series models for such data. Engle (1982) proposed an autoregressive conditionally heteroscedastic model namely ARCH(p), where $p \geq 1$ is the order of the model, to model the financial time series that exhibit time-varying volatility. The basic idea of ARCH(p) model is that the process $\{X_t\}$ is dependent although the series are uncorrelated and the dependence can be described in a simple quadratic function of its lagged values. Specifically, the model is assumed to have zero mean, serially uncorrelated processes with nonconstant variances conditional on the past, but constant unconditional variances. The ARCH(p) model has been shown to be useful not only in the dynamic volatility and correlation modeling, but also forecasting, risk management, market microstructure modeling, duration modeling and ultra-high-frequency data analysis (see Diebold (2004)). However, there were some shortcomings encountered with ARCH(p) model. Tsay (2014) highlighted four such weaknesses. Firstly, the positive and negative observations have the same effect since it depends on the square of previous observations. Secondly, some condition of the ARCH(p) model limits its ability to capture excess kurtosis of the process. Thirdly, the model may fail to provide insight on the cause of the variation observed. Finally, the model might also overpredict the volatility due to a slow response to large isolated shock.

Later, Bollerslev (1986) extended the ARCH(p) model process by proposing the generalized autoregressive conditional heteroscedasticity, GARCH(p,q) models to allow for a longer memory and more flexible lag structure. The GARCH(p,q) model is actually ARCH (∞) model with the error is conditionally normally distributed. The major advantage of this model is that the model allows time-varying volatility and leads to a fundamental change to the approaches used in finance. Tsay (2014) stated that this model enable us to construct volatility term structure for an observation and eventually improves the modelling and prediction of autoregressive moving average (ARMA) models. Later, the GARCH(p,q) model has been extended to other forms. For instance, Nelson (1991) proposed exponential GARCH model, namely EGARCH(p,q), by considering the weight innovation to allow for asymmetric effects between positive and negative assets returns. In addition, Engle & Ng (1993) introduced a nonsymmetric GARCH(p,q) model called as NGARCH(p,q) which can capture the asymmetric volatility response to the positive and negative observations while Gray (1996) introduced Markov-switching GARCH (MSW-GARCH) model to determine the short term interest rate by an unobservable Markov-process.

Time series data might involve count data, for example, in observing the changes of disease activity, the occurrence of virus infection, the count of price changes, the number of customers to be served and the incident in a city over a period of time (see Harvey & Fernandes (1989), Li et al. (2014), Ferland et al. (2006) and Davis et al. (2016)). This leads to the development of integer-valued time series models. For example, McKenzie (1985) introduced integer-valued autoregressive, INAR(p) model as the extension of AR(p) model. The model is essentially Markov chains, but is structurally autoregressions, and depends on only a few parameters. Then, Al-Osh & Alzaid (1987) investigated the properties of INAR(1) model and showed that the process can have negative value of autocorrelation. Later, various number of extension of INAR(p) model had being suggested, for instance, Kachour (2009) presented rounded INAR model namely RINAR(p) model, Pedeli & Karlis (2011) applied the classical integer-valued uutoregressive (INAR) model to the bivariate case known as bivariate Poisson INAR(p) model and Nastić et al. (2012) introduced the combined geometric integer-valued autoregressive with order $p > 0$, namely CGINAR(p) model using the negative binomial thinning. On the other hand, other types of count data time series also can be found in the literature. For example, Aly & Bouzar (1994) extended the ARMA model into integer-valued cases, Carvalho & Tanner (2007) proposed a model based on local mixtures of Poisson autoregressive models, Davis & Wu (2009) studied the generalized linear models for time series of counts where the latent observation is modeled by a negative binomial distribution and Houseman et al. (2006) proposed a nonstationarity negative binomial model for time series in modeling the enterococcus disease cases in Boston Harbor.

In addition, zero-inflated models are also frequently used when the time series data have the number of zeroes observed greater than what would be expected for the model. For example, Wang (2001) discussed the distribution changes in two-state of Markov chain for zero-inflated Poisson (ZIP) regression model, Yau et al. (2004) proposed a zero-inflated Poisson mixed autoregression model to analyze the incidence of workplace injuries in a hospital among cleaners, Lukusa et al. (2015) introduced a parameter estimation which is the inverse probability weighting (IPW) method to estimate the parameters of the ZIP regression model with missing covariates and Sellers & Raim (2016) suggested a model to model the relationship between explanatory and response variables namely zero-inflated Conway Maxwell Poisson (ZICMP) regression.

In our research, we are interested in the time series of count model introduced by Ferland et al. (2006). The authors suggested an integer-valued analogue of the classical generalized autoregressive conditional heteroskedastic, GARCH(p,q) model with conditional distribution follows Poisson namely INGARCH(p,q). Later, Zhu (2011) proposed error which following negative binomial instead of Poisson to overcome the overdispersion and zero-inflated model which are zero-inflated Poisson (ZIPINGARCH (p,q)) and zero-inflated negative binomial (ZINBINGARCH (p,q)). In this thesis, one of the contribution is to derive the higher order of moments via martingale difference.

Parameter estimation plays an important role in analyzing the time series analysis. Various estimation methods have been proposed to estimate the parameter of interests, for instance, the traditional methods including maximum likelihood estimation (MLE), least squares (LS) estimator and method of moments estimator (MME) and modern approaches to estimation methods including linear estimating function (EF) estimator and generalized method of moments (GMM) estimator. In estimation approaches for the INGARCH(p,q) model, Ferland et al. (2006) used MLE method to estimate the parameters of the process. But, the evaluation of the matrices in MLE is a cumbersome task. Meanwhile, for NBINGARCH(p,q) model, Zhu (2011) used three methods which are Yule-Walker (YW), MLE and LS estimators. MLE shows the best performance compared to LS and YW. However, the asymptotic theory of the MLE need geometric ergodicity of the process $\{X_t\}$. On the other hand, the common method of parameter estimation for ZIP models is MLE. However, as pointed by Nanjundan & Naika (2012), the MLE of

ZIP models have no closed form expressions. Therefore, in this research work, we use the quadratic estimating functions (QEF) approach proposed by Liang et al. (2011) as an alternative method in the parameter estimation of INGARCH(p,q), NBINGARCH(p,q) and ZIPINGARCH(p,q) models.

1.2 Statement of the problem

In real life, we may deal with count data which lead to the development of integer-valued time series (IVTS) model. The model of our interest is a class of INGARCH model assuming that the conditional distribution of the variable follows Poisson, negative binomial or zero-inflated Poisson distributions. The first two moments of such models have been discussed in literature. In this study, we derive the higher order moments for these models via martingale difference. On the other hand, a number of estimation methods have been used for the IVTS models but with shortcomings. Here, we will develop the estimation theory for the QEF method for the IVTS models and show its superiority compared to the other existing methods. .

1.3 Objectives

Based on the statement of problem above, the researcher has outlined the following objectives for this study:

1. Derive simpler form of the moments up to order four for a class of IVTS models via martingale representation.

2. Show the superiority of QEF method on zero-inflated models.

3. Develop new estimation theory based on QEF approach to estimate the parameters of the class of IVTS models.

4. Apply the methodology into real data sets.

1.4 Significance of research

The findings from this study will be advantageous in the following ways:

1. Contribute to the knowledge in statistics regarding the modeling of IVTS data and their higher order moments.

2. Optimize the estimation of parameters in IVTS model using QEF method.

3. Provide an alternative methods of estimations to be applied in real count data set.

1.5 Thesis outline

This research attempts to develop statistical methodologies for some problems in IVTS models. This research is outlined as follows:

Chapter 2 provides a literature review on the integer-valued time series models. We review the development of INGARCH (p,q) and the quadratic estimating function (QEF) method.

Chapter 3 discusses the moment properties of IVTS models. We present the class of INGARCH (p,q) models. Then, we derive the moment properties of the model up to order four.

Chapter 4 explains the QEF method. Here, we talk about the general method of QEF techniques and provide two examples of a command modeling for illustration.

Chapter 5, **Chapter 6** and **Chapter 7** consider INGARCH (p,q), NBINGARCH (p,q) and ZIPINGARCH (p,q) models respectively. In each chapter, we find the skewness and kurtosis formula of the model. We then derive the optimal QEF function and do a simulation study using R-cran programming to compare the performance of QEF method with other estimation methods. We provide the algorithm in solving the optimal QEF. We also apply such parameter estimation on a real data set and do some diagnostic checking using Pearson Residual to ascertain whether the data fit with the models.

Chapter 8 presents the summary of this research work. We also give some suggestions for extending the research work in the future.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Integer-valued time series or count data time series is widely used in many areas and field including epidemiology, criminal cases, insurance policy and public health surveillance. The importance of count data time series modeling have been discussed in the literature. For example, Aly & Bouzar (1994) underlined the theory of integer-valued autoregressive moving average model, Nastić et al. (2012) presented the combined geometric integer-valued autoregressive (CGINAR) model and Ferland et al. (2006) discussed on GARCH model for discrete data. The importance of discrete-valued time series also can be referred in MacDonald & Zucchini (1997) and Cameron & Trivedi (2001).

In practice, especially in medical field, many count data sets have high frequency of zeroes. For example, the rare diseases with low infection rates, the observed counts typically contain a high frequency of zeroes but the counts can also be very large during outbreak period. Therefore, Lambert (1992) introduced zero-inflated Poisson (ZIP) regression model in modeling such data sets. His study showed that ZIP model is not only easy to interpret, but also leads to more refined data analysis as the model can accommodate overdispersion. Following the findings of Lambert (1992), many studies and applications of ZIP model have been put forward. For instance, Mullahy (1997) proposed the hurdle model, Zhu (2011) studied the errors of GARCH model that follow the zero-inflated Poisson and negative binomial and Yang (2012) stressed the importance of zero-inflated model. On the other hand, the wide use of such models led to the creation of ZIM package in R software (see Yang et al. (2014)).

The growing interest in studying the various nature and origin of IVTS led to the development of new models. For example, Ferland et al. (2006) proposed INGARCH(p,q) model as analogue of GARCH (p,q) process with Poisson deviate instead of normal deviate. As an INGARCH(p,q) involves the Poisson variates, the conditional mean and conditional variance are the same, referred as equi-dispersion. But, in practice, frequently the data are overdispersed, with the variance is greater than the mean which may lead to poor performance in the existence of potential extreme observations. Therefore, to account for overdispersion and deal with potential extreme observations, Zhu (2011) introduced a new version of Ferland's model with negative binomial deviates namely NBINGARCH(p,q) model. On the other hand, Zhu (2012) used the same approach as INGARCH(p,q) model by modeling the process with high frequency of zeroes data sets known as ZIPINGARCH(p,q) model. These three models will be discussed and explained in detail for our research work.

The estimation approach is very important in time series analysis. For our models, namely, INGARCH(p,q), NBINGARCH(p,q) and ZIPINGARCH(p,q), the common estimation method used is MLE. However, the MLE approach has some drawbacks : (i) for INGARCH(p,q), Ferland et al. (2006) noted that the evaluation of the matrices is difficult to handle and needs to use the bootstrap technique which renders the estimation of the parameters of interest very complicated. (ii) for NBINGARCH(p,q) model, Zhu (2011) found that the geometric ergodicity of the process $\{X_t\}$ is needed in the theory of MLE asymptotics and (iii) for ZIPINGARCH(p,q), the conditional distribution follows ZIP model where the likelihood is divided into two parts which are the likelihood for zero and non-zero observations. Therefore, the MLE of this model have no closed form expressions (see Nanjundan & Naika (2012) and Zhu (2012)). In this research, we use the quadratic estimating functions (QEF) approach as the alternative method to estimate the parameters of interest. The theory of the QEF method will be discussed in section 2.3.

2.2 Development of INGARCH(p,q) Model

2.2.1 ARCH model

In real world application, the variance or volatility changes with time and this is typical of many classes of time series. Such changing known as heteroscedastic. It causes problems in predicting the future volatility pattern whether it is increasing or decreasing per time period.

Therefore, Engle (1982) introduced an autoregressive conditionally heteroscedastic model namely ARCH(p) originally to describe U.K. inflationary uncertainty as a function of past errors. It is called autoregressive since the model allows an AR (p) structure for the conditional heteroscedasticity with $\varepsilon_t \sim WN\left(0, \sigma_\varepsilon^2\right)$, with WN is white noise, $\alpha_0 > 0, \alpha_i \geq 0$ for $i = 1, 2, \ldots, p$ and σ_t^2 is the variance of process. The model is defined as

$$\begin{aligned} X_t &= \sigma_t \varepsilon_t; \ \forall t, \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2. \end{aligned} \tag{2.1}$$

Introduction of ARCH (p) models led to significant changes in financial modeling in which modeling of asset price volatility became more efficient.

2.2.2 GARCH (p,q) Model

Generalized Autoregressive Conditional Heteroskedasticity (GARCH(p,q)) was first introduced by Bollerslev (1986) to solve the problems encountered in ARCH (p) model with high order by allowing for more flexible lag structure. It is sometimes useful to consider the ARCH (∞) representation of a GARCH (p,q) process or one can say that the GARCH model is an infinite order ARCH model and often provides a highly parsimonious lag shape. Using the same approach corresponding to ARMA models, this extension of ARCH process is used to reduce the infinite number of estimated parameter. The GARCH (p,q) process uses values of the past squared observations and past variances to model the variance at time t whereby, in order word, the process $\{X_t\}$ conditioned on the past value, the σ-field generated from $X_{t-1}, X_{t-2}, \ldots, X_1$, $\Im_{t-1}^X$ follows normal distribution with mean zero and variance, h_t. It can be defined as

$$X_t | \Im_{t-1}^X \sim N(0, h_t), \tag{2.2}$$

$$h_t = \gamma_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}. \tag{2.3}$$

From Equation (2.2), it is clearly stated that the conditional variance of X_t given the past follows normal distribution with mean, $\mu = 0$ and variance, h_t. GARCH(p,q) models have been widely used in many areas especially in financial field.

Since the mid-1980s, these models have become actively discussed and studied among both academics researchers and practitioners and lead to several extensions of

GARCH (p,q) being developed. For example, integrated GARCH model (IGARCH) model was suggested by Engle & Bollerslev (1986) whereby the model is defined to be integrated in variance. Such model is argued to be both theoretically important for the asset pricing models and empirically relevant. On the other hand, Nelson (1991) proposed the exponential GARCH model known as EGARCH in order to capture the leverage effect of the stock market. In 1995, Sentana (1995) introduced quadratic GARCH model (QGARCH) to cope with asymmetric effects of shocks on volatility. In addition, the parameters in the GARCH (p,q) model change according to the sign or the size of shock ε_i which lead the model being interpreted as a regime-switching model known as Markov-Switching GARCH (MSW-GARCH) model. This model developed by Gray (1996) assumes that the regime is determined by an unobservable Markov-process.

Since the many different approaches have been proposed to model time series count data, Ferland et al. (2006) proposed an integer-valued GARCH (p,q) model known as INGARCH(p,q). This is inspired by GARCH (p,q) model using the integer-valued as the conditional distribution. The model use the identity link function for the conditional mean and assume the observations, conditionally on its past, to follow a Poisson distribution. Later, Zhu (2011) extended the INGARCH (p,q) model by assuming the conditional distribution follows negative binomial, denoted as NBINGARCH (p,q) model to overcome the shortcoming of the previous model. Then, the increasing number of data with high frequency of zeroes encourage Zhu (2012) to propose INGARCH(p,q) model with conditional distribution follows zero-infalted Poisson, ZIPINGARCH(p,q). These three models will be studied and explored in our research work. The details can be refer to Section 5, 6 and 7 respectively.

2.3 The Development of Quadratic Estimating Functions

2.3.1 Parameter Estimation

The estimation of parameters is very important in time series analysis. There are some desirable properties of estimators. Firstly, an estimator is said to be consistent if the estimated parameter converge to the true value as the sample size increases. Then, the estimator should be unbiased. The bias is defined as the deviation of the estimator from the true value. Here, we seek for the unbiased estimator whereby the deviation is close to zero. Lastly, the efficiency. It means that the estimator should give small variance.

The common and very well known estimator is the maximum likelihood estimation (MLE). In this method the estimator is obtained by maximizing the likelihood of the observed data. It is formally defined as

$$\hat{\boldsymbol{\theta}} = \arg_{\boldsymbol{\theta}} \ \max \mathrm{LIK}(\boldsymbol{\theta}, \ \boldsymbol{y}),$$

where $\boldsymbol{\theta}$ is the set of parameters, LIK is the likelihood and $\boldsymbol{y}$ is the vector of time series observations. But, MLE are not always the best under all circumstances. Bahadur (1958) claimed that not all MLE are consistent by giving two examples to show the inconsistency on MLE method in some cases and later, Le Cam (1990) add another six more examples. Using such examples, therefore, it is support that, for some cases, the MLE procedure cannot be recommended and need the alternative estimation method to estimate the parameters of interest.

In addition, Bera & Bilias (2002) reviewed important phases in the development of parameter estimation in both the statistical and econometric literature. He claimed that the optimality of MLE rests on the assumption that the true density function is known. But, in practise, the true distribution is seldom known which leads econometricians and statisticians to move away from applying MLE in estimating the estimated parameters. Besides, Vinod (1997) stated that MLE is sensitive to misspecification of the likelihood function. This is supported by Ng & Peiris (2013) by showing that incorrect likelihood function will affect the parameter estimates in terms of the mean and standard error. Moreover, according to Crowder (1987), in some cases, MLE may fail to give reasonable results by providing some examples. In his first example, he showed that the MLE fails to use the information on the parameters in the second moment of observations given. Then, in second example, the author illustrated that if the variance specification is not precisely correct, the consequences can be asymptotically disastrous. For the last example, he demonstrated that the MLE does not yield consistent estimator for some parametric function.

Therefore, various estimation methods have been proposed as alternative to MLE estimator. One of them is estimating functions (EF) (see details in Section 2.3.2). The work on estimating functions approach starts with the introduction of Godambe's information criterion in Godambe (1960) and later extended to discrete time stochastic processes, in Godambe (1985). This method have been studied and widely used by many authors including Bera & Bilias (2002), Merkouris et al. (2007) and Allen et al. (2013b).

There are many advantages of EF compared to MLE. Godambe & Heyde (2010) proved that the EF estimator yields asymptotically shortest confidence interval. Moreover, Ng & Peiris (2013) declared that the EF method is more computationally efficient and easy to apply in practise rather than MLE. In their research, they find that since the true distribution is rarely known, the EF is more useful and it gives a good approach and very reliable estimates. In addition, Bera & Bilias (2002) compared some estimation methods and noted that the EF approach is much easier to combine and it is invariant under one-to-one transformation of parameters. They discovered that the EF is well suited for semiparametric models because it only requires the assumption of a few moments. Furthermore, the result presented in Ng & Peiris (2013) proved that there is no significant difference in forecasting ability of EF compared to MLE in terms of computation. They also argued that the EF is more easier to evaluate and the estimates can be obtained without the knowledge of distribution of errors. Besides, Allen et al. (2013a) found that the computation time of EF is shorter than MLE (also can be see in Ng & Peiris (2013) and Ng et al. (2014)).

Liang et al. (2011) extended EF approach to quadratic estimating function (QEF) approach. They showed that the applications of QEF method are superior than EF method in some nonlinear time series models. The details of QEF estimator will be explained in Chapter 4. Thavaneswaran et al. (2012) and Thavaneswaran et al. (2015) used the QEF to jointly estimate the parameters of RCA models with GARCH errors and for generalized duration models respectively. In this research work, we will apply the QEF methods to integer-valued time series and compare the performance with MLE and EF estimators.

2.3.2 The Estimating Function

Fisher (1935) noted that the estimate of parameter $\boldsymbol{\theta}$ can be obtained by solving an equation,

$$g_h(X;\boldsymbol{\theta}) = 0, \tag{2.4}$$

where $g_h(X;\boldsymbol{\theta})$ is a function of the observation vector $\boldsymbol{X} = (X_t, t = 1,2,\ldots,n)$ and parameter $\boldsymbol{\theta}$. The classic approach of estimation required conditions on the resulting estimator $\hat{\boldsymbol{\theta}}$, such as unbiasedness, consistency, invariance and minimum variance. But, EF has a different approach. Instead of looking at the properties of estimator $\hat{\boldsymbol{\theta}}$, it is more concerned on the properties of EF itself. For instance, we will think about an unbiased EF rather than an unbiased $\hat{\boldsymbol{\theta}}$, i.e, we need

$$E\left[g_h(X;\boldsymbol{\theta})\right] = 0. \tag{2.5}$$

The significant role of unbiasedness and sufficiency were discussed in detail by Kimball et al. (1946). Another criterion of a good estimator is optimality. Durbin (1960) said that, "it seems reasonable to develop the idea of unbiased estimating equation with minimum variance". Therefore, such idea leads to the derivation of optimal linear unbiased EFs.

2.3.3 Derivation of Optimum EF

Godambe (1960) was the first person who introduced regular estimating function (EF) that satisfies certain conditions and came up with the procedure to choose an optimal EF. The required conditions for a function $g(X;\boldsymbol{\theta})$ to be a regular EF are:

(i) $E[g(X;\boldsymbol{\theta})] = \int g(X;\boldsymbol{\theta})f(X;\boldsymbol{\theta})dX = 0$,

(ii) $\frac{dg(X;\boldsymbol{\theta})}{d\boldsymbol{\theta}}$ exists for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, where $\boldsymbol{\Theta}$ is the parameter space,

(iii) $\int g(X;\boldsymbol{\theta})f(X;\boldsymbol{\theta})dX$ is differentiable under the sign of integration,

(iv) $E\left[\frac{dg(X;\boldsymbol{\theta})}{d\boldsymbol{\theta}}\right]^2 > 0$, for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$,

(v) $Var[g(X;\boldsymbol{\theta})] = E[g^2(X;\boldsymbol{\theta})] < \infty$, where $f(X;\boldsymbol{\theta})$ is the density function of extreme value distribution.

According to Godambe (1960), to find the optimal function of EF, denoted as $g^*(X;\boldsymbol{\theta})$, two criteria should be satisfied. First, the estimated parameter should be as close as possible to the true value. It means that the variance, $Var[g(X;\boldsymbol{\theta})] = E\left[g^2(X;\boldsymbol{\theta})\right]$ should be minimized, therefore $E\left[g^{*2}(X;\boldsymbol{\theta})\right] \leq E\left[g^2(X;\boldsymbol{\theta})\right]$ where $g^*(X;\boldsymbol{\theta})$ is the optimal estimating functions. The second criteria is the expected value of derivatives of function $g(X;\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$, $\left\{E\left[\frac{dg(X;\boldsymbol{\theta})}{d\boldsymbol{\theta}}\right]\right\}$ should be as large as possible i.e $\left\{E\left[\frac{dg^*(X;\boldsymbol{\theta})}{d\boldsymbol{\theta}}\right]\right\} \geq \left\{E\left[\frac{dg(X;\boldsymbol{\theta})}{d\boldsymbol{\theta}}\right]\right\}$. That measure of sensitivity requirement can be observed as an identification condition. By applying both principles, the $g^*(X;\boldsymbol{\theta})$ can be defined.

Definition 2.3.1. *A $g^* \in \mathscr{G}$ is said to be optimal if*

$$\frac{E\left[g^{*2}(X;\boldsymbol{\theta})\right]}{\{E\left[\frac{dg^*(X;\boldsymbol{\theta})}{d\boldsymbol{\theta}}\right]\}^2} \leq \frac{E\left[g^{2}(X;\boldsymbol{\theta})\right]}{\{E\left[\frac{dg(X;\boldsymbol{\theta})}{d\boldsymbol{\theta}}\right]\}^2}$$

for all $g^* \in \mathscr{G}$ and $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and where $\mathscr{G}$ is denoted as the class of all regular EFs.

Further, Godambe (1985) studied the inference of discrete stochastic processes using EF. He constrained the research to the class L of linear combination of martingale differences, say h_t's and the optimal function given as $\{g : g(\boldsymbol{\theta}) = \sum_{t=1}^{n} a_{t-1}h_t\}$ where $\boldsymbol{a}_{t-1}$ be $p \times q$ matrices depending on $\{X_1, X_2, \ldots, X_n\}$ and $\boldsymbol{\theta}$ with $E\left(h_t|\Im_{t-1}^X\right) = 0$ as $\Im_{t-1}^X$ is the $\sigma-$field generated by the process $\{X_1, X_2, \ldots, X_{t-1}\}$. Theorem (2.3.1) is the Godambe's (1985) result on optimal EF for the dependent case.

Theorem 2.3.1. *Define the EF as $\{g : g(\boldsymbol{\theta}) = \sum_{t=1}^{n} a_{t-1}h_t\}$ where h_t and a_{t-1} are assumed to be differentiable with respect to $\boldsymbol{\theta}$ for $t = 1, 2, \ldots, n$. Therefore, the optimal estimating function g^* that minimized $\frac{E\left[g^2(X;\boldsymbol{\theta})\right]}{\{E\left[\frac{dg(X;\boldsymbol{\theta})}{d\boldsymbol{\theta}}\right]\}^2}$ is $\left\{g^*(\boldsymbol{\theta}) = \sum_{t=1}^{n} a_{t-1}^* h_t\right\}$ where $a_{t-1}^* = \frac{E\left[\frac{dh_t}{d\boldsymbol{\theta}}|\Im_{t-1}^X\right]}{E\left[h_t^2|\Im_{t-1}^X\right]}$.*

The proof is available at Appendix A.

Furthermore, for multi-parameter case, the definition for the optimal EF is defined as Definition 2.3.2 below.

Definition 2.3.2. *A $g^* \in \mathscr{G}$ is said to be optimal if*

$$Var[g_s^*] \leq Var[g_s] \tag{2.6}$$

$$or\ g_s^* \leq g_s \tag{2.7}$$

$$or\ D_{g^*}^{-1}\sum_{g^*} D_{g^*}^{t-1} \leq D_g^{-1}\sum_{g} D_g^{t-1} \tag{2.8}$$

where $\sum_g = Var[g(X,\boldsymbol{\theta})] = E\left[g(X,\boldsymbol{\theta})g'(X,\boldsymbol{\theta})\right]$, $D_g = E\left[\frac{\partial g(X,\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]$ and $g_s(X,\boldsymbol{\theta}) = D_g^{-1}g(X,\boldsymbol{\theta})$ is the standardized vector EF i.e: the difference of the left hand side matrix from the right side matrix is nonnegative for all $g \in \mathscr{G}$.

The proof is available in Bera et al. (2006).

This method further developed onto QEF method. Such estimator was first introduced by Liang et al. (2011) whereby the authors studied the quadratic martingale estimating functions and showed that when the conditional mean and variance of the observed process depend on the same parameter of interest, then the QEF method shows better performance compared to EF method. Furthermore, from Liang et al. (2011), the QEF method is shown to be more informative compared to EF method by comparing their information. The methodology proposed later has been investigated by many authors including Ng et al. (2015) and Thavaneswaran et al. (2015). They showed that QEF method give the superior results either in simulation studies or real examples in ACD model compared to existing methods. However, the method has not been investigated to IVTS model. For this research, we attempt to apply QEF method to estimate the parameters of the IVTS models. The detail of QEF is given in Chapter 4.

CHAPTER 3

MOMENT PROPERTIES OF SOME INTEGER-VALUED TIME SERIES MODEL

3.1 Introduction

IVTS models are broadly applied in various fields especially biomedicine, epidemiology, economics and meteorology. For example, Zeger (1988) studied the monthly cases of Polio infection in US from 1970 to 1983, while Johansson (1996) looked at the effect of lowering the speed limits on the number of accidents and Li et al. (2014) studied the implication of crime cases over time. Hence, there is a strong need for IVTS models to be studied and consequently improved further, which is the focus of this study.

Ferland et al. (2006) proposed a new integer-valued time series model as an analogue of the generalized autoregressive conditional heteroskedastic (GARCH (p,q)) model with Poisson as conditional distribution. This model will be explained in detail in Chapter 5. It is later extended, for example, by Weiß (2013) in modelling time series of counts dealing with overdispersion. Zhu (2011) introduced NBINGARCH(p,q) while Zhu (2012) suggested ZIPINGARCH(p,q) model. Fokianos et al. (2009) considered geometric ergodicity and likelihood-based inference for linear and nonlinear Poisson autoregression.

3.2 The Class of Integer-Valued GARCH Models

In this thesis, we focus on three IVTS models, namely,

(a) INGARCH(p,q):

$$X_t|\Im_{t-1}^X \sim P(\lambda_{t,P})$$

$$\lambda_{t,P} = \gamma + \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=1}^{q} \beta_j \lambda_{t-j,P}$$

(b) NBINGARCH(p,q):

$$X_t|\Im_{t-1}^X \sim NB(r, \lambda_{t,NB})$$

$$\lambda_{t,NB} = \frac{1-p_t}{p_t} = \gamma + \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=1}^{q} \beta_j \lambda_{t-j,NB}$$

(c) ZIPINGARCH(p,q):

$$X_t|\Im_{t-1}^X \sim ZIP(\lambda_{t,ZIP},\omega)$$
$$\lambda_{t,ZIP} = \gamma + \sum_{i=1}^{p}\alpha_i X_{t-i} + \sum_{j=1}^{q}\beta_j\lambda_{t-j,ZIP}$$

where p_t is the probability of successes trials, ω is the inflation parameter lies between zeo and unity, r is the number of succeses trials, λ_t is the intensity parameter, $\Im_{t-1}^X$ is the $\sigma-$field generated by $X_{t-1}, X_{t-2}, \ldots, X_1$, $\gamma > 0$, $\alpha_i \geq 0$, $i = 1,2,\ldots,p$, $\beta_j \geq 0$ and $j = 1,2,\ldots,q$

The above models can be written as follow:

$$E\left(X_t|\Im_{t-1}^X\right) = a\lambda_{t,TP} \tag{3.1}$$
$$\lambda_{t,TP} = \gamma + \sum_{i=1}^{p}\alpha_i X_{t-i} + \sum_{j=1}^{q}\beta_j\lambda_{t-j,TP} \tag{3.2}$$

where $TP = P, NB$ or ZIP, a is the coefficient of the conditional mean with $a = 1$ and $a = r$ are for INGARCH and NBINGARCH models respectively while for ZIPINGARCH, it is given as $a = 1-\omega$ with $\gamma > 0$, $\alpha_i \geq 0$, $i = 1,2,\ldots,p$ and $\beta_j \geq 0$, $j = 1,2,\ldots,q$. Each of these models will be explained in detail in Chapters 5-7 respectively.

3.3 First and Second Moments of The Model

The new class of models can be written in standard ARMA representation. Using martingale transformation, $u_t = X_t - E(X_t|\Im_{t-1}^X) = X_t - a\lambda_{t,TP}$ with $E(u_t) = 0$ and $var(u_t) = \sigma_u^2$ and multiplying Equation (3.2) by a gives

$$a\lambda_{t,TP} = a\gamma + a\sum_{i=1}^{p}\alpha_i X_{t-i} + a\sum_{j=1}^{q}\beta_j\lambda_{t-j,TP}.$$

The Equation then can be rewritten as

$$X_t - u_t = a\gamma + a\sum_{i=1}^{p}\alpha_i X_{t-i} + \sum_{j=1}^{q}\beta_j\left(X_{t-j} - u_{t-j}\right)$$
$$\left(X_t - a\sum_{i=1}^{p}\alpha_i X_{t-i} - \sum_{j=1}^{q}\beta_j X_{t-j}\right) = a\gamma + u_t - \sum_{j=1}^{q}\beta_j u_{t-j}. \tag{3.3}$$

Since $\{u_t\}$ is martingale difference sequence and $\{X_t\}$ is a time series process, Equation

(3.3) can be written in backshift operator as

$$\left(1-a\sum_{i=1}^{p}\alpha_iB^i-\sum_{j=1}^{q}\beta_jB^j\right)X_t=a\gamma+\left(1-\sum_{j=1}^{q}\beta_jB^j\right)u_t. \tag{3.4}$$

From (3.4), the mean of the process is:

$$\begin{aligned}
E\left[\left(1-a\sum_{i=1}^{p}\alpha_iB^i-\sum_{j=1}^{q}\beta_jB^j\right)X_t\right] &= E\left[a\gamma+\left(1-\sum_{j=1}^{q}\beta_jB^j\right)u_t\right],\\
\left(1-a\sum_{i=1}^{p}\alpha_iB^i-\sum_{j=1}^{q}\beta_jB^j\right)E\left[X_t\right] &= a\gamma+\left(1-\sum_{j=1}^{q}\beta_jB^j\right)E\left[u_t\right],\\
\left(1-a\sum_{i=1}^{p}\alpha_i-\sum_{j=1}^{q}\beta_j\right)\mu &= a\gamma+\left(1-\sum_{j=1}^{q}\beta_j\right)(0),\\
\left(1-a\sum_{i=1}^{p}\alpha_i-\sum_{j=1}^{q}\beta_j\right)\mu &= a\gamma,\\
\mu &= \frac{a\gamma}{1-a\sum_{i=1}^{p}\alpha_i-\sum_{j=1}^{q}\beta_j}.
\end{aligned} \tag{3.5}$$

Now, from Equation (3.4), let $\phi(B)=1-\left(a\sum_{i=1}^{p}\alpha_iB^i+\sum_{j=1}^{q}\beta_jB^j\right)$ and $\theta(B)=1-\sum_{j=1}^{p}\beta_jB^j$. Therefore Equation (3.3) can be represented in the following form:

$$\phi(B)X_t=a\gamma+\theta(B)u_t. \tag{3.6}$$

We shall make the following stationarity assumptions for prosess $\{X_t\}$ having an ARMA(R,q) representation with $R=max(p,q)$:

- All zeroes of the polynomial $\phi(B)$ lie outside of the unit circle.
- $\sum_{j=0}^{\infty}\psi_j^2<\infty$ where the ψ's are obtained from the relation $\psi(B)\phi(B)=\theta(B)$.

These assumptions ensure that the u_t's are uncorrelated with zero mean and finite variance and that the process $\{X_t\}$ is weakly stationary. In this case, the autocorrelation function of $\{X_t\}$ will be exactly the same as that for a stationary ARMA(R,q) model. The Equation (3.6) can be written as $X_t-\mu=\psi(B)u_t$, i.e,

$$X_t-\mu=\sum_{j=0}^{\infty}\psi_ju_{t-j} \tag{3.7}$$

and the variance of X_t is given by

$$\sigma_X^2 \quad = \quad \sigma_u^2 \sum_{j=0}^{\infty} \psi_j^2. \tag{3.8}$$

The first interest here is to derive the general formula for the first two moments, autocovariance and autocorrelation of the integer-valued process $\{X_t\}$. The result is given in Proposition 1.

Proposition 1. *Under the stationarity assumption, the mean, variance, autocovariance and autocorrelation of the the integer-valued process $\{X_t\}$ are*

$$(a) \quad \mu_X = \frac{a\gamma}{1 - a\sum_{i=1}^{p} \alpha_i - \sum_{j=1}^{q} \beta_j},$$

$$(b) \quad \sigma_X^2 = \sigma_u^2 \sum_{j=0}^{\infty} \psi_j^2$$

$$(c) \quad \gamma_k^X = \sigma_u^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}$$

$$(d) \quad \rho_k^X = \frac{\sum_{j=0}^{\infty} \psi_j \psi_{j+k}}{\sum_{j=0}^{\infty} \psi_j^2}$$

The proof is available at Appendix C.

3.4 Skewness and Kurtosis

In the literature only the first two moments and the autovovariance are given for integer-valued volatility models considered. In this section, following Thavaneswaran et al. (2015), we obtain a general expression for the skewness and kurtosis for conditionally Poisson, negative binomial and zero-inflated Poisson distributions.

Proposition 2. *Under the assumption of stationarity and finite fourth moment, the process in the form $X_t - \mu = \sum_{j=0}^{\infty} \psi_j u_{t-j}$ where μ is the mean of the random process, $\{u_t\}$ is an uncorrelated noise process with mean zero, variance, σ_u^2, skewness, $\Gamma^{(u)}$ and kurtosis, $K^{(u)}$. Define $S_t = (X_t - \mu)^2$. Then,*

$$(a) \quad Var(S_t) = \left(K^{(u)} - 3\right) \sigma_u^4 \sum_{j=0}^{\infty} \psi_j^4 + 2\sigma_u^4 \left(\sum_{j=0}^{\infty} \psi_j^2\right)^2,$$

$$(b) \quad \Gamma^{(X)} = \frac{\sum_{j=0}^{\infty} \psi_j^3 \Gamma^{(u)}}{(\sum_{j=0}^{\infty} \psi_j^2)^{3/2}},$$

$$(c) \quad K^{(X)} = 3 + \frac{\left(K^{(u)} - 3\right) \sum_{j=0}^{\infty} \psi_j^4}{\left(\sum_{j=0}^{\infty} \psi_j^2\right)^2},$$

$$(d) \quad \rho_k^S = \frac{(K^{(u)} - 2) \sum_{j=0}^{\infty} \psi_j^2 \psi_{j+k}^2 + 2\left(\sum_{j=0}^{\infty} \psi_j \psi_{j+k}\right)^2}{(K^{(u)} - 3) \sum_{j=0}^{\infty} \psi_j^4 + 2\left(\sum_{j=0}^{\infty} \psi_j^2\right)^2}.$$

The proof is available at Appendix D.

3.4.1 Example on $p = 1$ and $q = 1$

From Section 3.3, the class of IVTS models considered in this study can be written in ARMA representation as Equation (3.7) with parameter ψ_j. For the case of $p = 1$ and $q = 1$, we can find the weight of ψ_j by using $\phi(B) = a\alpha_1 B + \beta_1 B$ and Equation (3.7), therefore, we have

$$\sum_{j=1}^{\infty} \psi_j B^j = \frac{1 - \beta_1 B}{1 - \phi(B)}. \tag{3.9}$$

Using geometric series, Equation (3.9) become

$$\sum_{j=0}^{\infty} \psi_j B^j = \left\{1 + \phi B + \phi^2 B^2 + \phi^3 B^3 + \ldots\right\}(1 - \beta_1 B).$$

Hence, the weight of ψ_j is given as

$$\begin{aligned}
\psi_0 &= 1, \\
\psi_1 &= \phi - \beta_1 = (a\alpha_1 + \beta_1) - \beta_1 = a\alpha_1, \\
\psi_2 &= \phi^2 - \phi\beta_1 = (a\alpha_1 + \beta_1)^2 - \beta_1 (a\alpha_1 + \beta_1) = (a\alpha_1 + \beta_1)(a\alpha_1 + \beta_1 - \beta_1) \\
&= \alpha_1 (a\alpha_1 + \beta_1), \\
\psi_3 &= \phi^3 - \phi^2\beta_1 = (a\alpha_1 + \beta_1)^3 - \beta_1 (a\alpha_1 + \beta_1)^2 = (a\alpha_1 + \beta_1)^2 (a\alpha_1 + \beta_1 - \beta_1) \\
&= \alpha_1 (a\alpha_1 + \beta_1)^2, \\
&\vdots \\
\psi_j &= \alpha_1 (a\alpha_1 + \beta_1)^{j-1}.
\end{aligned} \tag{3.10}$$

In a nutshell, it is shown that the weight $\psi_j = a\alpha_1 (a\alpha_1 + \beta_1)^{j-1}$ for $\psi_0 = 1$, for $j = 1, 2, \ldots$ where $a = 1$, for INGARCH $(1,1)$, $a = r$ for NBINGARCH $(1,1)$ and $a = 1 - \omega$ for ZIPINGARCH $(1,1)$.

Then, using the results obtained in Equation (3.10), power summation for the weight of ψ_j are given as:

$$\begin{aligned}
\sum_{j=0}^{\infty} \psi_j^2 &= 1 + (a\alpha_1)^2 + (a\alpha_1 (a\alpha_1 + \beta_1))^2 + \ldots + \left(a\alpha_1 (a\alpha_1 + \beta_1)^{j-1}\right)^2 + \ldots \\
&= 1 + (a\alpha_1)^2 \left\{1 + (a\alpha_1 + \beta_1)^2 + \ldots + (a\alpha_1 + \beta_1)^{2j-2} + \ldots\right\}.
\end{aligned} \tag{3.11}$$

Then, we summarize Equation (3.11) giving

$$\begin{aligned}
\sum_{j=0}^{\infty} \psi_j^2 &= 1 + (a\alpha_1)^2 \left[\frac{1}{1 - (a\alpha_1 + \beta_1)^2}\right] \\
&= \frac{1 - (a\alpha_1 + \beta_1)^2 + (a\alpha_1)^2}{1 - (a\alpha_1 + \beta_1)^2}.
\end{aligned} \tag{3.12}$$

By using the same approach finding $\sum_{j=0}^{\infty} \psi_j^2$, we have

$$\begin{aligned}
\sum_{j=0}^{\infty} \psi_j^3 &= \frac{1 - 3\alpha_1^2 a^2 \beta_1 - 3a\alpha_1 \beta_1^2 - \beta_1^3}{1 - (a\alpha_1 + \beta_1)^3}, \quad \text{and} \\
\sum_{j=0}^{\infty} \psi_j^4 &= \frac{1 - 4a^3\alpha_1^3\beta_1 - 6a^2\alpha_1^2\beta_1^2 - 4a\alpha_1\beta_1^3 - \beta_1^4}{1 - (a\alpha_1 + \beta_1)^4}.
\end{aligned}$$

On the other hand, we can find $\sum_{j=0}^{\infty} \psi_j \psi_{j+k}$, using the following steps.

$$\begin{aligned}
\sum_{j=0}^{\infty} \psi_j \psi_{j+k} &= \left[a\alpha_1 (a\alpha_1 + \beta_1)^{k-1}\right] + a\alpha_1 \left[a\alpha_1 (a\alpha_1 + \beta_1)^{k}\right] \\
&\quad + a\alpha_1 (a\alpha_1 + \beta_1) \left[a\alpha_1 (a\alpha_1 + \beta_1)^{k+1}\right] \\
&\quad + a\alpha_1 (a\alpha_1 + \beta_1)^2 \left[a\alpha_1 (a\alpha_1 + \beta_1)^{k+2}\right] + \ldots, \\
&= a\alpha_1 (a\alpha_1 + \beta_1)^{k-1} + a\alpha_1 \left[a\alpha_1 (a\alpha_1 + \beta_1)^{k}\right] \\
&\quad + a\alpha_1 \left[a\alpha_1 (a\alpha_1 + \beta_1)^{k+2}\right] \\
&\quad + a\alpha_1 \left[a\alpha_1 (a\alpha_1 + \beta_1)^{k+4}\right] + \ldots.
\end{aligned} \tag{3.13}$$

Equation (3.13) can be summarized as

$$\begin{aligned}\sum_{j=0}^{\infty}\psi_j\psi_{j+k} &= a\alpha_1\left(a\alpha_1+\beta_1\right)^{k-1}\left\{\begin{array}{c}1+a\alpha_1\left(a\alpha_1+\beta_1\right)+a\alpha_1\left(a\alpha_1+\beta_1\right)^3\\ +a\alpha_1\left(a\alpha_1+\beta_1\right)^5+\ldots\end{array}\right\},\\ &= a\alpha_1\left(a\alpha_1+\beta_1\right)^{k-1}\left\{1+a\alpha_1\left(a\alpha_1+\beta_1\right)\left[\begin{array}{c}1+\left(a\alpha_1+\beta_1\right)^2+\\ \left(a\alpha_1+\beta_1\right)^4+\ldots\end{array}\right]\right\}.\end{aligned}$$

Treating the Equation using geometric series, we have:

$$\begin{aligned}\sum_{j=0}^{\infty}\psi_j\psi_{j+k} &= a\alpha_1\left(a\alpha_1+\beta_1\right)^{k-1}\left\{1+a\alpha_1\left(a\alpha_1+\beta_1\right)\left[\frac{1}{1-\left(a\alpha_1+\beta_1\right)^2}\right]\right\},\\ &= a\alpha_1\left(a\alpha_1+\beta_1\right)^{k-1}\left\{\frac{1-a\alpha_1\beta_1-\beta_1^2}{1-\left(a\alpha_1+\beta_1\right)^2}\right\}.\end{aligned} \tag{3.14}$$

Meanwhile, we use the same technique for $\sum_{j=0}^{\infty}\psi_j^2\psi_{j+k}^2$ and we obtain $\sum_{j=0}^{\infty}\psi_j^2\psi_{j+k}^2$ as below:

$$\sum_{j=0}^{\infty}\psi_j^2\psi_{j+k}^2 = a^2\alpha_1^2\left(a\alpha_1+\beta_1\right)^{2k-2}\left\{\frac{1-\left(a\alpha_1+\beta_1\right)^4+a^2\alpha_1^2\left(a\alpha_1+\beta_1\right)^2}{1-\left(a\alpha_1+\beta_1\right)^4}\right\}. \tag{3.15}$$

Using Equation (3.11) and (3.14), the variance, autocovariance and autocorrelation for the models with order $p=1$ and $q=1$ are given by

$$var\left(X_t\right) = \sigma_u^2\sum_{j=0}^{\infty}\psi_j^2 = \sigma_u^2\left[\frac{1-2a\alpha_1\beta_1-\beta_1^2}{1-\left(a\alpha_1+\beta_1\right)^2}\right], \tag{3.16}$$

$$cov\left(X_tX_{t+k}\right) = \sigma_u^2\sum_{j=0}^{\infty}\psi_j\psi_{j+k} = \sigma_u^2 a\alpha_1\left(a\alpha_1+\beta_1\right)^{k-1}\left\{\frac{1-a\alpha_1\beta_1-\beta_1^2}{1-\left(a\alpha_1+\beta_1\right)^2}\right\}, \tag{3.17}$$

and

$$corr\left(X_tX_{t+k}\right) = \frac{\sigma_u^2\sum_{j=0}^{\infty}\psi_j\psi_{j+k}}{\sigma_u^2\sum_{j=0}^{\infty}\psi_j^2} = \frac{a\alpha_1\left(a\alpha_1+\beta_1\right)^{k-1}\left(1-a\alpha_1\beta_1-\beta_1^2\right)}{1-2a\alpha_1\beta_1-\beta_1^2} \tag{3.18}$$

respectively.

3.5 Feigin's Theorem on Stationary Distribution Of λ_t and X_t.

Let count time series, X_t at time t, when conditioned on λ_t is assumed to have a Poisson distribution, $P(.)$. Feigin et al. (2008) defined a single source of error (SSOE) model as

$$X_t|\lambda_t \sim P(\lambda_t) \tag{3.19}$$

$$\lambda_t = \lambda + \phi\lambda_{t-1} + \alpha(X_{t-1} - \lambda_{t-1}), \quad \text{for } t = 2,3,\ldots,T, \tag{3.20}$$

where the restrictions $\lambda > 0, \phi \geq \alpha \geq 0$ and $\phi \leq 1$. Here, for illustration, we consider only on first-order lags. Defining $\ell_{t-1} = (\lambda_1, X_1, X_2, \ldots, X_{t-1})$, from Equation (3.19) it follows that, conditional on ℓ_{t-1}, the mean and variance of X_t is λ_t.

In seeking to characterize the stationary distribution of λ_t when $\phi < 1$, Feigin et al. (2008) derived the expression for their Laplace transform. The Laplace transform of X_t and λ_t denoted as $L_{X_t}(.)$ and $L_{\lambda_t}(.)$ respectively. For conditional distribution of the process, $X_t|\ell_{t-1}$ follows Poisson distribution, then,

$$L_{X_t}(u) = E\left(e^{-uX_t}\right) = E\left[E\left(e^{-uX_t}\right)|\ell_{t-1}\right] = E\left[-\lambda_t\left(1-e^{-u}\right)\right] = L_{\lambda_t}\left(1-e^{-u}\right).$$

If the limit of $L_{\lambda_t}(.)$, say $L_\lambda(.)$, exists, hence the Laplace transform of the stationary distribution of X_t, say $L_X(.)$ will also exist and satisfy

$$L_X(u) = L_\lambda\left(1-e^{-u}\right). \tag{3.21}$$

Feigin et al. (2008) proposed the following Theorem 3.5.1 for stationarity distribution.

Theorem 3.5.1. *Given* $0 \leq \phi < 1$, $L_{\lambda_t}(\upsilon)$ *converges to*

$$L_\lambda(\upsilon) = exp\left\{-\lambda\sum_{k=0}^{\infty} g^{(k)}(\upsilon)\right\},$$

as $t \to \infty$ *where* $g(\upsilon) \equiv g(\upsilon;\delta,\alpha) = \upsilon\delta + (1-e^{-\upsilon\alpha})$ *with* $1 > \delta \equiv \phi - \alpha \geq 0$, $g^{(k)}(\upsilon)$ *is the function of* $g(\upsilon)$ *with power* k *and* $\sum_{k=0}^{\infty} g^{(k)}(\upsilon) < \infty$.

The proof is available at Appendix B. Hence, under this theorem, we can conclude that, our IVTS process are stationary as $t \to \infty$. Using the stationarity condition, $E(X_t) = \mu$ and for large t, $E(\lambda_t)$ converge to λ. Using this convergence in mean, we can apply that $lim_{t\to\infty} E\left[|\lambda_t|^k\right] = E\left[|\lambda|^k\right]$, where $\lambda = \dfrac{\mu}{a}$.

3.6 Summary of The Chapter

In this chapter, we introduced a class of IVTS models inspired by GARCH(p,q) model with unconditional distribution following Poisson, negative binomial and zero-inflated Poisson namely INGARCH(p,q), NBINGARCH(p,q) or ZIPINGARCH(p,q) models respectively. Then, we represented these models in ARMA representation and achieved our first objective by producing general close form expressions for the first four moments (mean, variance, skewness and kurtosis). Martingale differences were used to simplify the derivations and for the special case of $p = 1$ and $q = 1$, the ψ_j weights were obtained explicitly. However, for the existence of first four moments, only Ferland et al. (2006) showed the existence of moments for all order for INGARCH(p,q) model and for the other two models, NBINGARCH(p,q) and ZIPINGARCH(p,q), their existence of higher order of moments are still in discussion. Such existence for INGARCH(p,q) model can be found through the additive property of its conditional distribution which is Poisson distribution. Through such property, the process is built and its sequences are obtained using a cascade on a random variable via a sequence of i.i.d Poisson random variables, known as thinning operation. Unfortunately, such property does not work for the other two models, which are, NBINGARCH(p,q) and ZIPINGARCH(p,q) models. For NBINGARCH(p,q) model, where its conditional distribution follows negative binomial distribution, such property is applicable only on same p_i's while for ZIPINGARCH(p,q) model, nobody have looked at the problems. In fact, neither the thinning operation in Ferland et al. (2006) nor the coupling technique discussed in Franke (2010) in building the model can be applied on zero-inflated cases. Therefore, since the construction cannot be shown, its existence higher order moments also cannot be shown. To prove such existence, a new technique should be developed and implemented where it is still under discussion among the statisticians.

CHAPTER 4

THE QUADRATIC ESTIMATING FUNCTIONS

4.1 Introduction

Since the introduction by Godambe (1960), estimating functions (EF) had been applied in many areas including time series and related models (see Allen et al. (2013b), Thavaneswaran & Abraham (1988), Li et al. (2014), Chandra & Taniguchi (2001) and Thavaneswaran et al. (2015)). It is shown that in many studies, EF is computationally more efficient and easy to apply in real cases compared to the traditional parameter estimation, MLE.

Later, the EF method have been extended to quadratic estimating functions (QEF) by Liang et al. (2011). In their study, it is shown that QEF methods are more informative than the EF method when the first four conditional moments of the model are known (see Thavaneswaran et al. (2012) and Thavaneswaran et al. (2015)). In addition, Thavaneswaran et al. (2015), Liang et al. (2011), Merkouris et al. (2007) and Crowder (1987) showed that this extension leads to improvement in terms of the efficiencies of resulting estimates. At the same time, QEF method removes the problem of identifiability. Furthermore, according to Thavaneswaran et al. (2015), QEF method has standard asymptotic properties such as consistency and asymptotic normality compared to EF method. Moreover, the result of Monte Carlo simulation study presented in Ng et al. (2015) showed that the QEF estimators outperform the EF estimators in almost all cases in autoregressive conditional duration (ACD) model.

Since QEF estimator has not been discussed and applied on IVTS model, in our research work, we will focus on studying the performance of QEF method on these processes. We will compare the results with other estimation methods. This method will be explained in the next section.

4.1.1 General Model and Method

In order to use the quadratic estimating functions (QEF), it is necessary that the first four conditional moments are known. Consider a discrete time stochastic process, $\{X_t, t = 1, 2, \ldots\}$ conditonal on $\Im_{t-1}^X$ where $\Im_{t-1}^X$ is the σ-field generated from $X_{t-1}, X_{t-2}, \ldots, X_1$. The first four conditional moments are

$$\mu_t(\boldsymbol{\theta}) = \mathrm{E}\left[X_t | \Im_{t-1}^X\right], \tag{4.1}$$

$$\sigma_t^2(\boldsymbol{\theta}) = \mathrm{E}\left[(X_t - \mu_t(\boldsymbol{\theta}))^2 | \Im_{t-1}^X\right], \tag{4.2}$$

$$\Gamma_t(\boldsymbol{\theta}) = \frac{1}{\sigma_t^3(\boldsymbol{\theta})} \mathrm{E}\left[(X_t - \mu_t(\boldsymbol{\theta}))^3 | \Im_{t-1}^X\right], \tag{4.3}$$

and

$$\kappa_t(\boldsymbol{\theta}) = \frac{1}{\sigma_t^4(\boldsymbol{\theta})} \mathrm{E}\left[(X_t - \mu_t(\boldsymbol{\theta}))^4 | \Im_{t-1}^X\right] - 3, \tag{4.4}$$

For the skewness and the excess kurtosis of the process $\{X_t\}$, we assume that such moment properties do not contain any additional parameters. For QEF method in estimating the parameter of interest, $\boldsymbol{\theta}$ based on the observations $X_1, \ldots, X_n$, we consider two classes of martingale differences $\{m_t(\boldsymbol{\theta}) = X_t - \mu_t(\boldsymbol{\theta}), t = 1, \ldots, n\}$ and $\{s_t(\boldsymbol{\theta}) = m_t^2(\boldsymbol{\theta}) - \sigma_t^2(\boldsymbol{\theta}), t = 1, \ldots, n\}$. The variance and covariance of such martingale differences, m_t and s_t, can be described as:

$$\begin{aligned}
\langle m \rangle_t &= \mathrm{E}\left[m_t^2(\boldsymbol{\theta}) | \Im_{t-1}^X\right] = \mathrm{E}\left[(X_t - \mu_t(\boldsymbol{\theta}))^2 | \Im_{t-1}^X\right] = \sigma_t^2(\boldsymbol{\theta}), && (4.5)\\
\langle s \rangle_t &= \mathrm{E}\left[s_t^2(\boldsymbol{\theta}) | \Im_{t-1}^X\right] \\
&= \mathrm{E}\left[(X_t - \mu_t(\boldsymbol{\theta}))^4 + \sigma_t^4(\boldsymbol{\theta}) - 2\sigma_t^2(\boldsymbol{\theta})(X_t - \mu_t(\boldsymbol{\theta}))^2 | \Im_{t-1}^X\right], \\
&= \sigma_t^4(\boldsymbol{\theta})(\kappa_t(\boldsymbol{\theta}) + 2), && (4.6)\\
\langle m, s \rangle_t &= \mathrm{E}\left[m_t(\boldsymbol{\theta}) s_t(\boldsymbol{\theta}) | \Im_{t-1}^X\right] = \mathrm{E}\left[(X_t - \mu_t(\boldsymbol{\theta}))^3 - \sigma_t^2(\boldsymbol{\theta})(X_t - \mu_t(\boldsymbol{\theta})) | \Im_{t-1}^X\right], \\
&= \sigma_t^3(\boldsymbol{\theta}) \Gamma_t(\boldsymbol{\theta}). && (4.7)
\end{aligned}$$

Here, skewness and the excess kurtosis are assumed to have no additional parameters.

The optimal estimating functions also can be found for each martingale difference

m_t and s_t, that are

$$\begin{aligned}
\mathbf{g}_M^*(\boldsymbol{\theta}) &= \sum_{t=1}^{n} \frac{E\left(\frac{\partial m_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} | \Im_{t-1}^X\right)}{E\left(m_t^2 | \Im_{t-1}^X\right)} m_t(\boldsymbol{\theta}) = -\sum_{t=1}^{n} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{m_t(\boldsymbol{\theta})}{\langle m \rangle_t}, \text{ and} \\
\mathbf{g}_S^*(\boldsymbol{\theta}) &= \sum_{t=1}^{n} \frac{E\left(\frac{\partial s_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} | \Im_{t-1}^X\right)}{E\left(s_t^2 | \Im_{t-1}^X\right)} s_t(\boldsymbol{\theta}) = -\sum_{t=1}^{n} \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{s_t(\boldsymbol{\theta})}{\langle s \rangle_t}.
\end{aligned}$$

Then, the corresponding information for each component of $\mathbf{g}_M^*(\boldsymbol{\theta})$ and $\mathbf{g}_S^*(\boldsymbol{\theta})$ are

$$\begin{aligned}
\mathbf{I}_{\mathbf{g}_M}(\boldsymbol{\theta}) &= \sum_{t=1}^{n} \frac{\left(E[\frac{\partial m_t}{\partial \boldsymbol{\theta}} | \Im_{t-1}^X]\right)\left(E[\frac{\partial m_t}{\partial \boldsymbol{\theta}} | \Im_{t-1}^X]\right)'}{E\left[m_t m_t' | \Im_{t-1}^X\right]} = \sum_{t=1}^{n} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{1}{\langle m \rangle_t} \text{ and} \\
\mathbf{I}_{\mathbf{g}_S}(\boldsymbol{\theta}) &= \sum_{t=1}^{n} \frac{\left(E[\frac{\partial s_t}{\partial \boldsymbol{\theta}} | \Im_{t-1}^X]\right)\left(E[\frac{\partial s_t}{\partial \boldsymbol{\theta}} | \Im_{t-1}^X]\right)'}{E\left[s_t s_t' | \Im_{t-1}^X\right]} = \sum_{t=1}^{n} \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{1}{\langle s \rangle_t}.
\end{aligned}$$

For the discrete time stochastic process, $\{X_t\}$, the following Theorem 4.1.1 provides the optimal function and optimal information matrix of the QEF for the multiparameter case.

Theorem 4.1.1. *For the general model in (4.1) to (4.4), in the class of all quadratic estimating functions of the form* $\mathscr{G}_Q = \{\mathbf{g}_Q(\boldsymbol{\theta}) = \sum_{t=1}^n (\mathbf{a}_{t-1} m_t(\boldsymbol{\theta}) + \mathbf{b}_{t-1} s_t(\boldsymbol{\theta}))\}$,
(a) the optimal estimating functions is given by $\mathbf{g}_Q^*(\boldsymbol{\theta}) = \sum_{t=1}^n \left(\mathbf{a}_{t-1}^* m_t(\boldsymbol{\theta}) + \mathbf{b}_{t-1}^* s_t(\boldsymbol{\theta})\right)$,
where

$$\mathbf{a}_{t-1}^* = \left(1 - \frac{\langle m,s \rangle_t^2}{\langle m \rangle_t \langle s \rangle_t}\right)^{-1} \left(-\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{1}{\langle m \rangle_t} + \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t}\right)$$

and

$$\mathbf{b}_{t-1}^* = \left(1 - \frac{\langle m,s \rangle_t^2}{\langle m \rangle_t \langle s \rangle_t}\right)^{-1} \left(\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} - \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{1}{\langle s \rangle_t}\right);$$

(b) the information $\mathbf{I}_{g_Q^*}(\boldsymbol{\theta})$ *is given by*

$$\begin{aligned}
\mathbf{I}_{\mathbf{g}_Q^*}(\boldsymbol{\theta}) &= \sum_{t=1}^{n} \left(1 - \frac{\langle m,s \rangle_t^2}{\langle m \rangle_t \langle s \rangle_t}\right)^{-1} \left(\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{1}{\langle m \rangle_t} + \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{1}{\langle s \rangle_t} \right. \\
&\quad \left. - \left(\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} + \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right) \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t}\right);
\end{aligned}$$

(c) the gain in information $\mathbf{I}_{\mathbf{g}_Q^*}(\boldsymbol{\theta}) - \mathbf{I}_{\mathbf{g}_M^*}(\boldsymbol{\theta})$ *is given by*

$$\sum_{t=1}^{n} \left(1 - \frac{\langle m,s \rangle_t^2}{\langle m \rangle_t \langle s \rangle_t}\right)^{-1} \left(\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{\langle m,s \rangle_t^2}{\langle m \rangle_t^2 \langle s \rangle_t} + \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{1}{\langle s \rangle_t}\right.$$

$$-\left(\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}+\frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right)\frac{\langle m,s\rangle_t}{\langle m\rangle_t\langle s\rangle_t}\Bigg);$$

(d) the gain in information $\mathbf{I}_{\mathbf{g}_Q^*}(\boldsymbol{\theta})-\mathbf{I}_{\mathbf{g}_S^*}(\boldsymbol{\theta})$ *is given by*

$$\sum_{t=1}^{n}\left(1-\frac{\langle m,s\rangle_t^2}{\langle m\rangle_t\langle s\rangle_t}\right)^{-1}\Bigg(\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\frac{1}{\langle m\rangle_t}+\frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\frac{\langle m,s\rangle_t^2}{\langle m\rangle_t\langle s\rangle_t-\langle m,s\rangle_t^2}$$
$$-\left(\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}+\frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right)\frac{\langle m,s\rangle_t}{\langle m\rangle_t\langle s\rangle_t}\Bigg).$$

The proof is available at Appendix E

Corollary 4.1.2. *When the conditional skewness* γ *and kurtosis* κ *are constants, the optimal QEF and associated information, based on the martingale differences* $m_t(\boldsymbol{\theta})=X_t-\mu_t(\boldsymbol{\theta})$ *and* $s_t(\boldsymbol{\theta})=m_t(\boldsymbol{\theta})^2-\sigma_t(\boldsymbol{\theta})^2$*, are given by*

$$\mathbf{g}_Q^*(\boldsymbol{\theta})=\left(1-\frac{\gamma^2}{\kappa+2}\right)^{-1}\sum_{t=1}^{n}\frac{1}{\sigma_t^3(\boldsymbol{\theta})}\left(\begin{array}{c}\left\{-\sigma_t(\boldsymbol{\theta})\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}+\frac{\gamma}{\kappa+2}\frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right\}m_t(\boldsymbol{\theta})\\ +\frac{1}{\kappa+2}\left\{\gamma\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}-\frac{1}{\sigma_t(\boldsymbol{\theta})}\frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right\}s_t\end{array}\right)$$

and

$$\mathbf{I}_{\mathbf{g}_Q^*}(\boldsymbol{\theta})=\left(1-\frac{\gamma^2}{\kappa+2}\right)^{-1}\left(\mathbf{I}_{\mathbf{g}_M^*}(\boldsymbol{\theta})+\mathbf{I}_{\mathbf{g}_S^*}(\boldsymbol{\theta})-\frac{\gamma}{\kappa+2}\sum_{t=1}^{n}\frac{1}{\sigma_t^3(\boldsymbol{\theta})}\left\{\begin{array}{c}\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\\ +\\ \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\end{array}\right\}\right).$$

4.2 Zero-inflated Model

In this section, to investigate theoretically the performance of QEF method compared with EF method when applied on several count models, we derive the the optimal function and information gain using EF and QEF methods by considering two types of zero-inflated models, namely basic zero-inflated Poisson and basic zero-inflated Poisson regression models. We compare the results based on the information gain.

Zero-inflated count data models have probability function given by

$$f(y) = \begin{cases} \omega + (1-\omega)g(y) & \text{for } y = 0 \\ (1-\omega)g(y) & \text{for } y = 1,2,3,\ldots \end{cases} \tag{4.8}$$

where y is a count-valued random variable, $\omega \in [0,1]$ is a zero-inflation parameter (the probability of a strategic zero), and $g(\cdot)$ is the probability function of the *parent* count model. The mean of the zero-inflated count data model is

$$E(y) = \sum_{y=0}^{\infty} y f(y) = (1-\omega)E_g(y) \tag{4.9}$$

where $E_g(y)$ denotes the mean of the parent distribution. A fully parametric zero-inflated count data model is obtained once the probability function of the parent count model is specified.

4.2.1 Basic zero-Inflated Poisson Model

The simplest example would be the ZIP model obtained from (4.8) by taking

$$g(y_t;\lambda) = \frac{\exp(-\lambda)\lambda^{y_t}}{y_t!},\ \lambda > 0 \tag{4.10}$$

with mean $E_g(y_t) = \lambda$, $\mu_1 = E(y_t) = (1-\omega)\lambda$ and $\mu_2 = E(y_t^2) = \lambda(1-\omega)(\lambda+1)$. The parameter set is $\boldsymbol{\theta} = (\lambda,\omega)^{'}$. Following Kharrati-Kopaei & Faghih (2011), we define two martingale differences as $m_t = y_t - \mu_1$ and $\mathfrak{s}_t = y_t^2 - \mu_2$. The expression for $\langle m \rangle_t = E[m_t^2|\Im_{t-1}^X] = \sigma_{11}$, $\langle \mathfrak{s} \rangle_t = E[\mathfrak{s}_t^2|\Im_{t-1}^X] = \sigma_{22}$ and $\langle m,\mathfrak{s} \rangle_t = E[m_t\mathfrak{s}_t|\Im_{t-1}^X] = \sigma_{12}$ are given below. The derivative of μ_1 and μ_2 with respect to $\boldsymbol{\theta}$ are

$$\frac{\partial \mu_1}{\partial \boldsymbol{\theta}} = (1-\omega,-\lambda)^{'} \text{ and } \frac{\partial \mu_2}{\partial \boldsymbol{\theta}} = \left((1-\omega)(1+2\lambda),-(\lambda+\lambda^2)\right)^{'}.$$

Using results in Kharrati-Kopaei & Faghih (2011), the variances and covariance of the martingale difference are:

$$\sigma_{11} = Var(y_t) = \lambda\,(1-\omega)\,(1+\lambda) - \lambda^2\,(1-\omega)^2,$$

$$\sigma_{12} = Cov(y_t y_t^2) = \lambda\,(1-\omega)\left(\lambda^2 + 3\lambda + 1 - (1-\omega)\left(\lambda + \lambda^2\right)\right), \text{ and}$$

$$\sigma_{22} = Var(y_t^2) = \lambda\,(1-\omega)\left(\lambda^3 + 6\lambda^2 + 7\lambda + 1 - (1-\omega)\,\lambda\,(1+\lambda)^2\right).$$

Hence, using Theorem 4.1.1, the optimal QEF for each parameter are:

$$\begin{aligned}
\mathbf{g}_Q^*(\lambda) &= \left(1 - \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}}\right)^{-1} \left[\begin{array}{c} \left(-\dfrac{1-\omega}{\sigma_{11}} + \dfrac{(1-\omega)(1+2\lambda)\sigma_{11}}{\sigma_{11}\sigma_{22}}\right) m_t \\ + \left(\dfrac{(1-\omega)\sigma_{12}}{\sigma_{11}\sigma_{22}} + \dfrac{(1-\omega)(1+2\lambda)\sigma_{12}}{\sigma_{22}}\right) s_t \end{array} \right] \\
&= (1-\omega)\left(\frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}^2}\right) \sum_{t=1}^{n} \left[\begin{array}{c} \left(-\sigma_{22} + \sigma_{12}(1+2\lambda)\right) m_t + \\ \\ \left(\sigma_{12} - \sigma_{11}(1+2\lambda)\right) \mathfrak{s}_t \end{array} \right] \\
\mathbf{g}_Q^*(\omega) &= \left(1 - \frac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}}\right)^{-1} \left[\left(\frac{\lambda}{\sigma_{11}} - \frac{(\lambda+\lambda^2)\sigma_{11}}{\sigma_{11}\sigma_{22}}\right) m_t - \left(\frac{\lambda\sigma_{12}}{\sigma_{11}\sigma_{22}} + \frac{(\lambda+\lambda^2)\sigma_{12}}{\sigma_{22}}\right) s_t \right] \\
&= \lambda\left(\frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}^2}\right) \sum_{t=1}^{n} \left(\left(\sigma_{22} - (1+\lambda)\sigma_{12}\right) m_t + \left(\sigma_{11}(1+\lambda) - \sigma_{12}\right) \mathfrak{s}_t \right).
\end{aligned}$$

The information matrix of the optimal $\boldsymbol{\theta}$ is

$$\mathbf{I}_{\mathbf{g}_Q^*}(\boldsymbol{\theta}) = \begin{bmatrix} I_{\lambda\lambda}^Q & I_{\lambda\omega}^Q \\ \\ I_{\omega\lambda}^Q & I_{\omega\omega}^Q \end{bmatrix}$$

where

$$\begin{aligned}
I_{\lambda\lambda}^Q &= \frac{n\,(1-\omega)^2\left(\sigma_{22} + [1+2\lambda]\left[(1+2\lambda)\,\sigma_{11} - 2\sigma_{12}\right]\right)}{\sigma_{11}\sigma_{22} - \sigma_{12}^2}, \\
I_{\omega\omega}^Q &= \frac{n\lambda^2\left(\sigma_{22} + \sigma_{11}\,(1+\lambda)^2 - 2\,(1+\lambda)\,\sigma_{12}\right)}{\sigma_{11}\sigma_{22} - \sigma_{12}^2}, \\
\text{and } I_{\omega\lambda}^Q = I_{\lambda\omega}^Q &= \frac{n\lambda\,(1-\omega)\left(-\sigma_{22} - \sigma_{11}\left(1 + 3\lambda + 2\lambda^2\right) + \sigma_{12}\,(2+3\lambda)\right)}{\sigma_{11}\sigma_{22} - \sigma_{12}^2}.
\end{aligned}$$

For illustration, we compare only the information gain for parameter λ. The information for each martingale differences, m_t and $\mathfrak{s}_t$ are $I_{\lambda\lambda}^m = n(1-\omega)^2/\sigma_{11}$ and $I_{\lambda\lambda}^{\mathfrak{s}} = n(1-\omega)^2(1+2\lambda)/\sigma_{22}$ respectively. Therefore, we can clearly see that the value of denominator of $I_{\lambda\lambda}^Q$ is smaller than that $I_{\lambda\lambda}^m$ and value of numerator is larger than that of

$I^m_{\lambda\lambda}$ which leads to a greater gain information for $I^Q_{\lambda\lambda}$ compared to $I^m_{\lambda\lambda}$ and $I^{\mathfrak{s}}_{\lambda\lambda}$. That is $I^Q_{\lambda\lambda} > I^m_{\lambda\lambda}$ and $I^Q_{\lambda\lambda} > I^{\mathfrak{s}}_{\lambda\lambda}$. Similarly for $I^Q_{\omega\omega}$ as well. Hence, one can say that, the combined estimating functions are more informative than the information from each component.

4.2.2 Zero-Inflated Poisson Regression Model

In some cases, we may parameterize both λ and ω in terms of exogenous explanatory variables, say x and z where z can be identical to x, overlap with x or completely distinct from x. Following the definition given by Staub & Winkelmann (2013), we assume that

$$\lambda = \exp(\lambda_0 + \lambda_1 x) \text{ and } \omega = \frac{\exp(\delta_0 + \delta_1 z)}{1 + \exp(\delta_0 + \delta_1 z)}. \tag{4.11}$$

The parameter is $\boldsymbol{\theta} = (\lambda_0, \lambda_1, \delta_0, \delta_1)'$. The conditional expectation function of the corresponding ZIP model is given by

$$E(y \mid x, z) = (1 - \omega)\lambda = \frac{\exp(\lambda_0 + \lambda_1 x)}{1 + \exp(\delta_0 + \delta_1 z)}. \tag{4.12}$$

Here, we consider independent counts $y_t,\ t = 1, 2, \dots, n$, with parameters λ_t and ω_t arising from Equation (4.11). Hence, the mean, variance, skewness and kurtosis of the process are

$$\begin{aligned}
\mu_t(\boldsymbol{\theta}) &= \frac{\exp(\lambda_0 + \lambda_1 x_t)}{1 + \exp(\delta_0 + \delta_1 z_t)}, \\
\sigma_t^2(\boldsymbol{\theta}) &= \frac{\exp(\lambda_0 + \lambda_1 x_t)[1 + \exp(\delta_0 + \delta_1 z_t) + \exp(\lambda_0 + \lambda_1 x_t + \delta_0 + \delta_1 z_t)]}{[1 + \exp(\delta_0 + \delta_1 z_t)]^2}, \\
\Gamma_t(\boldsymbol{\theta}) &= \frac{1 + 3\lambda_t\omega_t + \lambda_t^2\omega_t + 2\lambda_t^2\omega_t^2}{\mu_t\,[1 + \lambda_t\omega_t]^{\frac{3}{2}}}, \text{ and} \\
\kappa_t(\boldsymbol{\theta}) &= \frac{\omega_t\lambda_t^3\left(6\omega_t^2 - 6\omega_t + 1\right) + 6\omega_t\lambda_t^2\,(2\omega_t - 1) + 7\omega_t\lambda_t + 1}{(1 - \omega_t)\,\lambda_t\,(1 + \omega_t\lambda_t)^2}.
\end{aligned}$$

respectively. We take the martingale differences to be $m_t(\boldsymbol{\theta}) = y_t - \mu_t(\boldsymbol{\theta})$ and $s_t(\boldsymbol{\theta}) = m_t^2(\boldsymbol{\theta}) - \sigma_t^2(\boldsymbol{\theta})$. The derivative of $\mu_t(\boldsymbol{\theta})$ with respect to each parameter is expressed as $\partial\mu_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = (B_{1,t}, B_{2,t}, B_{3,t}, B_{4,t})'$ where

$$B_{1,t} = \frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\lambda_0} = \frac{\exp(\lambda_0 + \lambda_1 x_t)}{1 + \exp(\delta_0 + \delta_1 z_t)},\ B_{2,t} = \frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\lambda_1} = \frac{\exp(\lambda_0 + \lambda_1 x_t)x_t}{1 + \exp(\delta_0 + \delta_1 z_t)},$$

$$B_{3,t} = \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \delta_0} = -\frac{\exp(\lambda_0 + \lambda_1 x_t + \delta_0 + \delta_1 z_t)}{[1+\exp(\delta_0 + \delta_1 z_t)]^2}, \text{ and}$$
$$B_{4,t} = \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \delta_1} = -\frac{\exp(\lambda_0 + \lambda_1 x_t + \delta_0 + \delta_1 z_t) z_t}{[1+\exp(\delta_0 + \delta_1 z_t)]^2}.$$

Let $$Y_t = [1/(\eta_t \sigma_t^4(\boldsymbol{\theta}))] \begin{pmatrix} -\sigma_t^2(\boldsymbol{\theta})\left(\kappa_t(\boldsymbol{\theta}) + 2 - (1+\lambda_t \omega_t)\right) \\ +\Gamma_t \sigma_t(\boldsymbol{\theta})(Y_t - \mu_t(\boldsymbol{\theta})) \begin{pmatrix} 1 + \lambda_t \omega_t + (Y_t - \mu_t(\boldsymbol{\theta})) \end{pmatrix} \\ -\sigma_t^3(\boldsymbol{\theta})\Gamma_t(\boldsymbol{\theta}) \end{pmatrix},$$

where $\eta_t = \kappa_t(\boldsymbol{\theta}) + 2 - \gamma_t^2(\boldsymbol{\theta}), \quad t = 1, 2, \cdots, n$, and therefore, the optimal QEF for λ_0, λ_1, δ_0 and δ_1 are

$$\begin{aligned} \mathbf{g}_Q^*(\boldsymbol{\theta}) &= \left(\mathbf{g}_Q^*(\lambda_0), \mathbf{g}_Q^*(\lambda_1), \mathbf{g}_Q^*(\delta_0), \mathbf{g}_{\mathbf{Q}}^*(\delta_1)\right)' \\ &= \left(\sum_{t=1}^n Y_t B_{1,t}, \sum_{t=1}^n Y_t B_{2,t}, \sum_{t=1}^n Y_t B_{3,t}, \sum_{t=1}^n Y_t B_{4,t}\right)'. \end{aligned}$$

On the other hand, the information matrix for QEF is

$$\mathbf{I}_{\mathbf{g}_Q^*}(\boldsymbol{\theta}) = \sum_{t=1}^n \frac{1}{\Re_t} \begin{pmatrix} \sigma_t(\boldsymbol{\theta})(\kappa_t(\boldsymbol{\theta}) + 2) W_t^2 \\ -2\Gamma_t(\boldsymbol{\theta}) W_t \end{pmatrix} \begin{bmatrix} B_{1,t}^2 & B_{1,t}B_{2,t} & B_{1,t}B_{3,t} & B_{1,t}B_{4,t} \\ B_{2,t}B_{1,t} & B_{2,t}^2 & B_{2,t}B_{3,t} & B_{2,t}B_{4,t} \\ B_{3,t}B_{1,t} & B_{3,t}B_{2,t} & B_{3,t}^2 & B_{3,t}B_{4,t} \\ B_{4,t}B_{1,t} & B_{4,t}B_{2,t} & B_{4,t}B_{3,t} & B_{4,t}^2 \end{bmatrix}$$

where $\Re_t = \sigma_t^3(\boldsymbol{\theta})\eta_t$ and $W_t = \exp(\lambda_0 + \lambda_1 x_t + \delta_0 + \delta_1 z_t) / \exp(\delta_0 + \delta_1 z_t)$. For comparison purpose, we focus only on the parameters λ_0 and λ_1 and the information gain by the estimating functions based on the single element of martingale differences $m_t(\boldsymbol{\theta})$ and $s_t(\boldsymbol{\theta})$ are derived as follows: For $m_t(\boldsymbol{\theta})$, the information matrix based on parameter λ are $I^m_{\lambda_0\lambda_0} = \sum_{t=1}^n B_{1,t}^2/\sigma_t^2(\boldsymbol{\theta})$ and $I^m_{\lambda_1\lambda_1} = \sum_{t=1}^n B_{2,t}^2/\sigma_t^2(\boldsymbol{\theta})$. On the other hand, for $s_t(\boldsymbol{\theta})$ we have $I^s_{\lambda_0\lambda_0} = \sum_{t=1}^n [1/\wp_t] W_t^2 B_{1,t}^2$ and $I^s_{\lambda_1\lambda_1} = \sum_{t=1}^n [1/\wp_t] W_t^2 B_{2,t}^2$ where $\wp_t = \sigma_t^4(\boldsymbol{\theta})(\kappa_t(\boldsymbol{\theta}) + 2)$. In a nutshell, it is obvious that $I^Q_{\lambda_0\lambda_0} > I^m_{\lambda_0\lambda_0}$ and $I^Q_{\lambda_0\lambda_0} > I^s_{\lambda_0\lambda_0}$ as well as $I^{Q}_{\lambda_1\lambda_1} > I^{m}_{\lambda_1\lambda_1}$ and $I^{Q}_{\lambda_1\lambda_1} > I^{m}_{\lambda_1\lambda_1}$. From the comparison, one can say that, following the determinant optimality discussed in Bera et al. (2006), the determinant of the combined estimating function is larger than the determinants of the information matrices associated with the components estimating functions. That is $|\mathbf{I}_{\mathbf{g}^*_Q}(\boldsymbol{\theta})| \geq |\mathbf{I}_{\mathbf{g}^*_m}(\boldsymbol{\theta})|$ and $|\mathbf{I}_{\mathbf{g}^*_Q}(\boldsymbol{\theta})| \geq |\mathbf{I}_{\mathbf{g}^*_s}(\boldsymbol{\theta})|$.

4.3 Summary of The Chapter

In this chapter, we discussed the theory on QEF method proposed by Liang et al. (2011) when applied on count data. Here, we derived the EF and QEF estimators for two types of ZIP models namely the basic ZIP and ZIP regression. In addition, we have also obtained closed form expressions of the information matrices for all the models and show that the QEF with combined estimating functions is more informative than EF with the component estimating functions. Thus, QEF method has improved the efficiency of the estimation of parameter $\boldsymbol{\theta}$ compared to the EF method. Hence, we can say that, the derivation of optimal QEF function and their information matrix have shown the superiority of QEF method when applied on selected zero-inflated models.

CHAPTER 5

INGARCH(p,q) MODEL

5.1 Introduction

Since GARCH models have given rise to new directions for research in probability and statistics and towards developing more parsimonious models for count time series data, Ferland et al. (2006) considered the model known as integer-valued GARCH model, namely INGARCH(p,q) where it is a GARCH(p,q) process with Poisson conditional distribution. He defined an INGARCH(p,q) process to be an integer-valued $\{X_t\}$ such that

$$X_t|\Im^X_{t-1} \sim P(\lambda_t) \tag{5.1}$$

$$\lambda_t = \gamma + \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=1}^{q} \beta_j \lambda_{t-j} \tag{5.2}$$

where $\Im^X_{t-1}$ is the $\sigma-$field generated by $X_{t-1}, X_{t-2}, \ldots, X_1$, $\gamma > 0$, $\alpha_i \geq 0$, $i = 1,2,\ldots,p$, and $\beta_j \geq 0$, $j = 1,2,\ldots,q$.

From the definition of INGARCH(p,q) model, it is clear that INGARCH $(p,0)$ model actually an INARCH(p) process. It should be noted that the second-order stationarity of the process is satisfied if and only if the summation of parameter lies within zero and one, whereby $0 < \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$.

This model is originally proposed to model the number of cases of campylobacterosis infections from January 1990 to the end of October 2000 in the north of the Province of Quèbec in Ferland et al. (2006). In this chapter, we will further study this model to explore its moment properties by using martingale transformation of the model.

5.2 The Moment Properties

The first objective of our study is finding the moment properties of unconditional distribution for all three models. The derivation of this unconditional moments need the conditional moment properties. Therefore, in this section, the moments for both unconditional and conditional distribution of INGARCH(p,q) model will be derived up to order four.

5.2.1 The Moment Properties of Conditional Distribution of INGARCH(p,q) Model

In INGARCH(p,q) set up, the conditional distribution of X_t given $\Im_{t-1}^X$ follows Poisson. To simplify, let $Y_t = X_t|\Im_{t-1}^X$ follows the Poisson distribution with parameter λ_t, denoted as $Y_t \sim P(\lambda_t)$. The moment properties of Poisson model can be obtained via its probability generating function (pgf) which is $G(z) = exp[\lambda_t(z-1)]$. We can get the moments of the model by finding its derivatives with evaluated at $z = 1$. The first three derivatives of such pgf are $G'(z) = \frac{d(exp(\lambda_t[z-1]))}{dz} = \lambda_t e^{(\lambda_t[z-1])}$, $G''(z) = \frac{d\lambda_t(e^{(\lambda_t[z-1])})}{dz} = \lambda_t^2 e^{(\lambda_t[z-1])}$ and $G'''(z) = \frac{d\lambda_t^2(e^{(\lambda_t[z-1])})}{dz} = \lambda_t^3 e^{(\lambda_t[z-1])}$ respectively. In general, we can show that, for the kth derivative, the pgf of Poisson distribution is in the form of $G^k(z) = \lambda_t^k e^{\lambda_t(z-1)}$.

The mean of Poisson is given as $\mu_t = G'(1) = E(y_t) = \lambda_t$. The second derivative is $G''(1) = E(y_t(y_t-1)) = \lambda_t^2$. By expanding that, we will get $E\left(y_t^2\right) - E\left(y_t\right) = \lambda_t^2$. Rearrange it, then, $E\left(y_t^2\right) = \lambda_t^2 + \lambda_t$. The variance of Poisson distribution is $\sigma_{y,t}^2 = E\left(y_t^2\right) - E\left(y_t\right)^2 = \lambda_t^2 + \lambda_t - \lambda_t^2 = \lambda_t$. For the third derivative is $G'''(1) = E(y_t(y_t-1)(y_t-2)) = \lambda_t^3$. This derivative can be expanded as $E\left(y_t^3 - 3y_t^2 + 2y_t\right) = \lambda_t^3$. Therefore, $E\left(y_t^3\right) = \lambda_t^3 + 3\lambda_t^2 + \lambda_t$. Using this result, the third moment is

$$
\begin{aligned}
E\left[(y_t-\lambda_t)^3\right] &= E\left(y_t^3 - 3y_t^2\lambda_t + 3y_t\lambda_t^2 - \lambda_t^3\right), \\
&= E\left(y_t^3\right) - 3E\left(y_t^2\right)\lambda_t + 3E\left(y_t\right)\lambda_t^2 - \lambda_t^3, \\
&= \lambda_t^3 + 3\lambda_t^2 + \lambda_t - 3\lambda_t^3 - 3\lambda_t^2 + 3\lambda_t^3 - \lambda_t^3, \\
&= \lambda_t. \qquad (5.3)
\end{aligned}
$$

Using Equation (D.5) in Appendix D, (5.3) and σ_y^2, the skewness is obtained in the following form:

$$
\begin{aligned}
\Gamma_t^{(P)} &= \frac{E\left[(y_t-\mu)^3\right]}{(var(y_t))^{3/2}}, \\
&= \frac{\lambda_t}{\lambda_t^{3/2}}, \\
&= \frac{1}{\sqrt{\lambda_t}}. \qquad (5.4)
\end{aligned}
$$

Now, we find the kurtosis of Poisson distribution. The fourth derivatives of the pgf is $G''''(1) = E\left(y_t\left(y_t-1\right)\left(y_t-2\right)\left(y_t-3\right)\right) = \lambda_t^4$. Again, we expand the factorization to obtain $E\left(y_t^4\right)$ whereby $E\left(y_t^4 - 6y_t^3 + 11y_t^2 - 6y_t\right) = \lambda_t^4$ and will lead to $E\left(y_t^4\right) = \lambda_t^4 + 6\lambda_t^3 + 7\lambda_t^2 + \lambda_t$. Therefore, the fourth moment is

$$\begin{aligned}
E\left[\left(y_t-\lambda_t\right)^4\right] &= E\left(y_t^4 - 4y_t^3\lambda_t + 6\lambda_t^2 y_t^2 - 4y_t\lambda_t^3 - \lambda_t^4\right), \\
&= E\left(y_t^4\right) - 4E\left(y_t^3\right)\lambda_t + 6\lambda_t^2 E\left(y_t^2\right) - 4E\left(y_t\right)\lambda_t^3 - \lambda_t^4, \\
&= \lambda_t^4 + 6\lambda_t^3 + 7\lambda_t^2 + \lambda_t - 4\lambda_t\left(\lambda_t^3 + 3\lambda_t^2 + \lambda_t\right) + 6\lambda_t^2\left(\lambda_t^2 + \lambda_t\right) - 4\lambda_t^3\lambda_t + \lambda_t^4, \\
&= \lambda_t^4 + 6\lambda_t^3 + 7\lambda_t^2 + \lambda_t - 4\lambda_t^4 - 12\lambda_t^3 - 4\lambda_t^2 + 6\lambda_t^4 + 6\lambda_t^3 - 4\lambda_t^4 + \lambda_t^4, \\
&= 3\lambda_t^2 + \lambda_t. \qquad (5.5)
\end{aligned}$$

Using (D.2) in Appendix D, σ_y^2 and Equation (5.5), the kurtosis of Poisson is

$$\begin{aligned}
K_t^{(P)} &= \frac{E\left[\left(y_t-\lambda_t\right)^4\right]}{\sigma_y^4}, \\
&= \frac{3\lambda_t^2 + \lambda_t}{\lambda_t^2}, \\
&= 3 + \frac{1}{\lambda_t}. \qquad (5.6)
\end{aligned}$$

5.2.2 The Moment Properties of Unconditional Distribution of INGARCH(p,q) Model

Using the martingale difference, u_t and the results in section 3.3, we obtain the mean as $\mu = \sigma_u^2 = E\left(\lambda_t\right)$. For large t, $\sigma_u^2 = \lambda = \mu$ and using Theorem 3.3.1 (b), the variance of INGARCH(p,q) model become

$$\sigma_X^2 = \mu \sum_{j=0}^{\infty} \psi_j^2.$$

Now, we find the skewness and kurtosis of INGARCH(p,q) model. We first have to find the skewness and kurtosis of the martingale difference, $\Gamma^{(u)}$ and $K^{(u)}$ respectively. For the skewness of u_t, we have

$$\Gamma^{(u)} = \frac{E\left(u_t^3\right)}{\left\{E\left(u_t^2\right)\right\}^{3/2}}$$

$$
\begin{aligned}
&= \frac{E\left(E\left[(\{X_t - \mu_t\}|\Im_{t-1})^3\right]\right)}{\sigma_u^3} \\
&= \frac{E\left(\lambda_t\right)}{\sigma_u^3}. \qquad (5.7)
\end{aligned}
$$

Similarly, using the result above on Equation (5.7), the excess kurtosis of the process (5.1) is given by

$$
\begin{aligned}
K^{(u)} &= \frac{E\left(u_t^4\right)}{\left\{E\left(u_t^2\right)\right\}^2} - 3 \\
&= \frac{E\left(E\left[(\{X_t - \mu_t\}|\Im_{t-1})^2\right]\right)}{\sigma_u^4} - 3 \\
&= \frac{E\left(3\lambda_t^2 + \lambda_t\right)}{\sigma_u^4} - 3. \qquad (5.8)
\end{aligned}
$$

Under large t, Equations (5.7) and (5.8) become

$$
\begin{aligned}
\Gamma^{(u)} &= \frac{\mu}{\mu^{3/2}} \\
&= \frac{1}{\sqrt{\mu}}, \qquad (5.9)
\end{aligned}
$$

and

$$
\begin{aligned}
K^{(u)} &= \frac{3\mu^2 + \mu}{\mu^2} - 3 \\
&= \frac{1}{\mu}. \qquad (5.10)
\end{aligned}
$$

Substituting Equation (5.9) into Equation (D.5) and Equation (5.10) into Equation (D.2), the skewness and kurtosis for INGARCH(p,q) are

$$
\Gamma^{(X)} = \frac{\sum_{j=0}^{\infty} \psi_j^3}{\sqrt{\mu}\left(\sum_{j=0}^{\infty} \psi_j^2\right)^{3/2}}
$$

and

$$
K^{(X)} = 3 + \frac{\left(\frac{1}{\mu} - 3\right)\sum_{j=0}^{\infty} \psi_j^4}{\left(\sum_{j=0}^{\infty} \psi_j^2\right)^2}
$$

respectively.

Ferland et al. (2006) showed the moment properties of unconditional INGARCH(p,q) distribution only for $p = 1$ and $q = 1$ case. It should also be noted that, with our finding, we are able to find this moments up to order four for all cases of p and q.

5.2.3 Empirical Study

To demonstrate the moments structure are correctly derived, we do the empirical study for the cases when the process is close and far from the boundary of stationarity. Here, we generate 500 samples of size 2000 and calculate the mean and the mean square error of the estimated moments as tabulated in Table 5.1. It can be seen that the estimated values are close to the true values.

Table 5.1: Generated data and true values for the moment structures with $\gamma = 0.1$

		$\alpha_1 = 0.1$			$\alpha_1 = 0.5$		
		Estimated	True	MSE	Estimated	True	MSE
$\beta = 0.2$							
	μ_X	0.9012	0.9018	8.85E-06	1.028	1.027	1.14E-05
	σ_2^X	0.9158	0.9162	3.58E-04	1.1158	1.1162	4.12E-04
	$\Gamma^{(X)}$	1.258	1.262	1.88E-04	1.052	1.059	9.48E-05
	$K^{(X)}$	4.125	4.108	5.22E-03	4.123	4.215	1.21E-04
$\beta = 0.4$							
	μ_X	0.7581	0.7588	1.26E-05	1.125	1.129	1.25E-05
	σ_2^X	0.9147	0.9158	1.28E-04	1.088	1.093	4.12E-04
	$\Gamma^{(X)}$	1.228	1.305	1.06E-03	1.235	1.244	1.21E-03
	$K^{(X)}$	4.113	4.212	4.12E-04	4.215	4.198	7.53E-04

5.3 Quadratic Estimating Functions on INGARCH(p,q) Model

In this section, we derive the optimal quadratic estimating functions, $\mathbf{g}_Q^*(\boldsymbol{\theta})$ for INGARCH(p,q) model as described in Section 4.1.1.

In order to extract the $\mathbf{g}_Q^*(\boldsymbol{\theta})$, the first four conditional moments of INGARCH(p,q) should be known. From Section 5.2.1, we obtained the first four conditional moments

$$\mu_t(\boldsymbol{\theta}) = \lambda_t(\boldsymbol{\theta}), \tag{5.11}$$

$$\sigma_t^2(\boldsymbol{\theta}) = \lambda_t(\boldsymbol{\theta}), \tag{5.12}$$

$$\Gamma_t(\boldsymbol{\theta}) = \frac{1}{\sqrt{\lambda_t(\boldsymbol{\theta})}}, \tag{5.13}$$

and

$$\kappa_t(\boldsymbol{\theta}) = \frac{1}{\lambda_t(\boldsymbol{\theta})} \tag{5.14}$$

where $\boldsymbol{\theta} = \gamma, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q$.

In QEF method, we use two martingale differences, namely m_t and s_t. They are defined as $m_t(\boldsymbol{\theta}) = X_t - \lambda_t(\boldsymbol{\theta})$ and for $s_t(\boldsymbol{\theta}) = m_t^2(\boldsymbol{\theta}) - \lambda_t(\boldsymbol{\theta})$ for $t = 1, 2, \ldots, n$. The variances and covariance of such martingale differences are

$$\begin{aligned} \langle m \rangle_t &= E\left[m_t^2(\boldsymbol{\theta}) \,|\Im_{t-1}^X\right] \\ &= \sigma_t^2(\boldsymbol{\theta}), \\ &= \lambda_t(\boldsymbol{\theta}), \end{aligned} \tag{5.15}$$

$$\begin{aligned} \langle s \rangle_t &= \sigma_t^4(\boldsymbol{\theta})(\kappa_t(\boldsymbol{\theta}) + 2), \\ &= \lambda_t^2(\boldsymbol{\theta})\left(\frac{1}{\lambda_t(\boldsymbol{\theta})} + 2\right), \\ &= \lambda_t^2(\boldsymbol{\theta})\left(\frac{1 + 2\lambda_t(\boldsymbol{\theta})}{\lambda_t(\boldsymbol{\theta})}\right), \\ &= \lambda_t(\boldsymbol{\theta})(1 + 2\lambda_t(\boldsymbol{\theta})), \end{aligned} \tag{5.16}$$

and

$$\begin{aligned} \langle m, s \rangle_t &= \sigma_t^4(\boldsymbol{\theta})\Gamma_t(\boldsymbol{\theta}), \\ &= \lambda_t^{3/2}(\boldsymbol{\theta})\frac{1}{\lambda_t(\boldsymbol{\theta})}, \\ &= \lambda_t^{1/2}(\boldsymbol{\theta}). \end{aligned} \tag{5.17}$$

Next, we find the partial derivatives $\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \theta}$ and $\frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \theta}$. Since the conditional mean, μ_t and conditional variance, σ_t^2 are the same, we have

$$\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \theta} = \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \theta} = \frac{\partial \lambda_t(\boldsymbol{\theta})}{\partial \theta}.$$

The partial derivative of $\lambda_t(\boldsymbol{\theta})$ with respect to parameter $\boldsymbol{\theta}$ are given as:

$$\begin{aligned}
\lambda_t'(\boldsymbol{\theta}) &= \left(\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\gamma}, \frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\alpha_1}, \ldots, \frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\alpha_p}, \frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\beta_1}, \ldots, \frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\beta_q}\right)' \\
&= \left(1+\sum_{j=1}^{q}\beta_j\frac{\partial\lambda_{t-j}(\boldsymbol{\theta})}{\partial\gamma}, X_{t-i}+\sum_{j=1}^{q}\beta_j\frac{\partial\lambda_{t-j}(\boldsymbol{\theta})}{\partial\alpha_i}, \lambda_{t-j}+\sum_{k=1}^{p}\beta_k\frac{\partial\lambda_{t-k}(\boldsymbol{\theta})}{\partial\beta_j},\right)',
\end{aligned}$$

where $i = 1,2,\ldots,p$ and $j = 1,2,\ldots,q$.

From Theorem 4.1.1 (a), the optimal estimation function can be achieved by finding $\mathbf{a}^*_{t-1}$ and $\mathbf{b}^*_{t-1}$. Since

$$\begin{aligned}
\left(1-\frac{\langle m,s\rangle_t^2}{\langle m\rangle_t\langle s\rangle_t}\right)^{-1} &= \left(1-\frac{\lambda_t^2(\boldsymbol{\theta})}{\lambda_t^2(\boldsymbol{\theta})\,(1+2\lambda_t(\boldsymbol{\theta}))}\right)^{-1}, \\
&= \frac{1+2\lambda_t(\boldsymbol{\theta})}{2\lambda_t(\boldsymbol{\theta})},
\end{aligned}$$

we have

$$\begin{aligned}
\mathbf{a}^*_{t-1} &= \left(1-\frac{\langle m,s\rangle_t^2}{\langle m\rangle_t\langle s\rangle_t}\right)^{-1}\left(-\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{1}{\langle m\rangle_t}+\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{\langle m,s\rangle_t}{\langle m\rangle_t\langle s\rangle_t}\right), \\
&= \frac{1+2\lambda_t(\boldsymbol{\theta})}{2\lambda_t(\boldsymbol{\theta})}\left[-\lambda_t\prime(\boldsymbol{\theta})\frac{1}{\lambda_t(\boldsymbol{\theta})}+\lambda_t\prime(\boldsymbol{\theta})\frac{\lambda_t(\boldsymbol{\theta})}{(1+2\lambda_t(\boldsymbol{\theta}))\lambda_t^2(\boldsymbol{\theta})}\right], \\
&= \frac{-\lambda_t\prime(\boldsymbol{\theta})}{\lambda_t(\boldsymbol{\theta})},
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{b}^*_{t-1} &= \left(1-\frac{\langle m,s\rangle_t^2}{\langle m\rangle_t\langle s\rangle_t}\right)^{-1}\left(\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{\langle m,s\rangle_t}{\langle m\rangle_t\langle s\rangle_t}-\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{1}{\langle s\rangle_t}\right), \\
&= \frac{1+2\lambda_t(\boldsymbol{\theta})}{2\lambda_t(\boldsymbol{\theta})}\left[\lambda_t\prime(\boldsymbol{\theta})\left(\frac{1}{\lambda_t(\boldsymbol{\theta})\,(1+2\lambda_t(\boldsymbol{\theta}))}\right)-\lambda_t\prime(\boldsymbol{\theta})\left(\frac{1}{\lambda_t(\boldsymbol{\theta})\,(1+2\lambda_t(\boldsymbol{\theta}))}\right)\right], \\
&= 0.
\end{aligned}$$

Since $\mathbf{b}^*_{t-1}$ is zero, we can say that, for INGARCH(p,q), the QEF estimator reduces to EF estimator. Hence, the optimal estimating functions for each parameter are:

$$\mathbf{g}^*_Q(\gamma) = \sum_{t=1}^{n}\frac{1}{\gamma+\sum_{i=1}^{p}\alpha_iX_{t-i}+\sum_{j=1}^{q}\beta_j\lambda_{t-j}(\boldsymbol{\theta})}$$

$$\left[1+\sum_{j=1}^{q}\beta_j\frac{\partial\lambda_{t-j}(\boldsymbol{\theta})}{\partial\gamma}\right]\left(X_t-\{\gamma+\sum_{i=1}^{p}\alpha_iX_{t-i}+\sum_{j=1}^{q}\beta_j\lambda_{t-j}(\boldsymbol{\theta})\}\right), \tag{5.18}$$

$$\mathbf{g}_Q^*(\alpha_i) = \sum_{t=1}^{n}\frac{1}{\gamma+\sum_{i=1}^{p}\alpha_iX_{t-i}+\sum_{j=1}^{q}\beta_j\lambda_{t-j}(\boldsymbol{\theta})}$$

$$\left[X_{t-i}+\sum_{j=1}^{q}\beta_j\frac{\partial\lambda_{t-j}(\boldsymbol{\theta})}{\partial\alpha_i}\right]\left(X_t-\{\gamma+\sum_{i=1}^{p}\alpha_iX_{t-i}+\sum_{j=1}^{q}\beta_j\lambda_{t-j}(\boldsymbol{\theta})\}\right), \text{and} \tag{5.19}$$

$$\mathbf{g}_Q^*(\beta_j) = \sum_{t=1}^{n}\frac{1}{\gamma+\sum_{i=1}^{p}\alpha_iX_{t-i}+\sum_{j=1}^{q}\beta_j\lambda_{t-j}(\boldsymbol{\theta})}$$

$$\left[\lambda_{t-j}+\sum_{k=1}^{p}\beta_k\frac{\partial\lambda_{t-k}(\boldsymbol{\theta})}{\partial\beta_j}\right]\left(X_t-\{\gamma+\sum_{i=1}^{p}\alpha_iX_{t-i}+\sum_{k=1}^{q}\beta_k\lambda_{t-k}(\boldsymbol{\theta})\}\right), \tag{5.20}$$

for $i = 1,2.\ldots,p$ and $j = 1,2,\ldots,q$.

On the other hand, using Theorem 4.1.1(b), the information for the optimal QEF can be obtained by

$$\begin{aligned}
\mathbf{I}_{\mathbf{g}_Q^*}(\boldsymbol{\theta}) &= \sum_{t=1}^{n}\left(1-\frac{\langle m,s\rangle_t^2}{\langle m\rangle_t\langle s\rangle_t}\right)^{-1}\left(\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\frac{1}{\langle m\rangle_t}+\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\frac{1}{\langle s\rangle_t}\right.\\
&\quad\left.-\left(\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}+\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right)\frac{\langle m,s\rangle_t}{\langle m\rangle_t\langle s\rangle_t}\right)\\
&= \frac{1+2\lambda_t}{2\lambda_t}\frac{\partial\lambda_t}{\partial\boldsymbol{\theta}}\frac{\partial\lambda_t}{\partial\boldsymbol{\theta}'}\left(\frac{1}{\lambda_t}+\frac{1}{\lambda_t(1+2\lambda_t)}-2\frac{\lambda_t}{\lambda_t^2(1+2\lambda_t)}\right)\\
&= \frac{1}{\lambda_t}\frac{\partial\lambda_t}{\partial\boldsymbol{\theta}}\frac{\partial\lambda_t}{\partial\boldsymbol{\theta}'}.
\end{aligned}$$

Let $M_{1,t} = \dfrac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\gamma}$, $M_{2,i,t} = \dfrac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\alpha_{i,t}}$ for $i = 1,2,\ldots,p$ and $M_{3,j,t} = \dfrac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\beta_{j,t}}$ for $j = 1,2,\ldots,q$ or can be written as:

$$\begin{aligned}
\frac{\partial\mu_t}{\partial\theta}=\frac{\partial\sigma_t^2}{\partial\theta} &= \left(\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\gamma},\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\alpha_{1,t}},\ldots,\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\alpha_p},\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\beta_j}\ldots,\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\beta_q}\right)'\\
&= \left(M_{1,t},M_{2,i,t},\ldots,M_{2,p,t},M_{3,j,t},\ldots,M_{3,q,t}\right)',
\end{aligned}$$

and $N_t = \dfrac{1}{\lambda_t}$, therefore, the information matrix of the optimal quadratic estimating functions for $\boldsymbol{\theta}$ is

$$\mathbf{I}_{\mathbf{g}^*_Q}(\boldsymbol{\theta}) = \begin{pmatrix} \sum_{t=1}^{n} N_t M_{1,t}^2 & \sum_{t=1}^{n} N_t M_{1,t} M_{2,i,t} & \sum_{t=1}^{n} N_t M_{1,t} M_{3,j,t} \\ \sum_{t=1}^{n} N_t M_{2,i,t} M_{1,t} & \sum_{t=1}^{n} N_t M_{2,i,t}^2 & \sum_{t=1}^{n} N_t M_{2,i,t} M_{3,j,t} \\ \sum_{t=1}^{n} N_t M_{3,j,t} M_{1,t} & \sum_{t=1}^{n} N_t M_{3,j,t} M_{2,i,t} & \sum_{t=1}^{n} N_t M_{3,j,t}^2 \end{pmatrix}.$$

5.4 Performance of The Estimation Method in INGARCH $(1,1)$

Here, we compare the performance of QEF in INGARCH $(1,1)$ model with EF and MLE estimators. The process of INGARCH $(1,1)$ is :

$$X_t|\Im_{t-1}^X \sim P(\lambda_t((\boldsymbol{\theta})),$$
$$\lambda_t(\boldsymbol{\theta}) = \gamma + \alpha_1 X_{t-1} + \beta_1 \lambda_{t-1}(\boldsymbol{\theta}). \tag{5.21}$$

Note that the QEF estimator and the EF estimator are the same for this model. To evaluate the performance of QEF, MLE and LS methods, a simulation study was carried out with $N = 500$ replications and two sample sizes, $n = 100$ and $n = 2000$. The simulation is carried out using R-cran software to obtain the estimated parameters of interest.

5.4.1 Conditional LS Derivation of INGARCH $(1,1)$

Lehmann & Casella (1991) defined the least square estimator as in Definition 5.4.1

Definition 5.4.1. *Let $Y = f(X) + noise$. Suppose f is known up to a finite number $p \leq n$ of parameters $\boldsymbol{\theta} = (\theta_1, ..., \theta_p)$, the least squares estimator, denoted by $\hat{\boldsymbol{\theta}}$ is that value of θ that minimizes $\sum_{i=1}^{n}(y_i - E(y_i))^2$, that is*

$$\hat{\boldsymbol{\theta}} = min \sum_{i=1}^{n} (y_i - E(y_i))^2.$$

The INGARCH $(1,1)$ is defined by Equation (5.1) and (5.2) with mean, $\mu_t = \lambda_t$. Therefore, using Definition 5.4.1, the LS of such model can be derived as

$$\hat{\boldsymbol{\theta}} = \min \sum_{i=1}^{n} (X_t - \mu)^2. \tag{5.22}$$

We minimize the Equation above to obtain the value of estimated parameters via simulation in R-cran software using *nlminb* command.

5.4.2 MLE Derivation of INGARCH $(1,1)$

The most well-known and traditional parameter estimation is MLE estimator. According to Casella & Berger (2002), MLE can be defines as in Definition (5.4.2)

Definition 5.4.2. *Let $X_1, X_2, \ldots, X_n$ be a random sample from distribution that depends on one or more unknown parameters $\theta_1, \theta_2, \ldots, \theta_n$ with probability mass or density function (pdf/pmf) $f(\theta_1, \theta_2, \ldots, \theta_n)$. Suppose that $\theta_1, \theta_2, \ldots, \theta_n$ is restricted to a given parameter space Ω. Then, the maximum likelihood estimator for θ_i for $i = 1, 2, \ldots, n$ is*

$$\hat{\boldsymbol{\theta}} = arg_{\theta}\ max\ L(\boldsymbol{\theta},\ \boldsymbol{x}) = arg_{\theta}\ \mathscr{L}(\boldsymbol{\theta},\ \boldsymbol{x}),$$

where $L(\theta_1, \theta_2, \ldots, \theta_n, \boldsymbol{x}) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, \ldots, \theta_n)$ and max $\{L(\boldsymbol{\theta},\ \boldsymbol{x})\} = \mathscr{L}(\boldsymbol{\theta},\ \boldsymbol{x})$.

Therefore, from Definition 5.4.2, to obtain the value of estimated parameter using MLE, we have to maximize the loglikelihood of the model. Using Equation (5.1) and (5.2), the loglikelihood of INGARCH $(1,1)$ is

$$maxLikelihood = \mathscr{L} = \sum_{t=1}^{n} \{-\lambda_t + X_t\ ln\ \lambda_t - ln\ (X_t!)\}. \tag{5.23}$$

We minimize the negative likelihood using *nlminb* command in R-cran software in order to estimate the parameters of interest.

5.4.3 QEF of INGARCH $(1,1)$

From Section 5.3, we derive the QEF method for INGARCH(p,q) model. Using such derivation via Equation (5.19)-(5.20), we can simply find the optimal functions for INGARCH $(1,1)$ as follow:

$$\mathbf{g}_Q^*(\gamma) \quad = \quad \sum_{t=1}^{n} \frac{1}{\gamma + \alpha_1 X_{t-1} + \beta_1 \lambda_{t-1}(\boldsymbol{\theta})}$$

$$\left[1+\beta_1\frac{\partial\lambda_{t-1}(\boldsymbol{\theta})}{\partial\gamma}\right](X_t-\{\gamma+\alpha_1X_{t-1}+\beta_1\lambda_{t-1}(\boldsymbol{\theta})\}),$$

(5.24)

$$\mathbf{g}_Q^*(\alpha_i) = \sum_{t=1}^{n}\frac{1}{\gamma+\alpha_1X_{t-1}+\beta_1\lambda_{t-1}(\boldsymbol{\theta})}$$

$$\left[X_{t-i}+\beta_1\frac{\partial\lambda_{t-1}(\boldsymbol{\theta})}{\partial\alpha_1}\right](X_t-\{\gamma+\alpha_1X_{t-1}+\beta_1\lambda_{t-1}(\boldsymbol{\theta})\}),\text{and}$$

(5.25)

$$\mathbf{g}_Q^*(\beta_j) = \sum_{t=1}^{n}\frac{1}{\gamma+\alpha_1X_{t-1}+\beta_1\lambda_{t-1}(\boldsymbol{\theta})}$$

$$\left[\lambda_{t-1}+\beta_1\frac{\partial\lambda_{t-1}(\boldsymbol{\theta})}{\partial\beta_1}\right](X_t-\{\gamma+\alpha_1X_{t-1}+\beta_1\lambda_{t-1}(\boldsymbol{\theta})\}).$$

(5.26)

By letting $\mathbf{g}_Q^*(\boldsymbol{\theta}) = 0$, the simultaneous Equations can be solved using R-cran software to give the QEF estimate of the parameters.

5.4.4 Simulation Study

Here, we discuss the steps to estimate parameters for INGARCH(1,1) model using LS, MLE and EF/QEF methods:

- *Step 1- Generate the data*: We first generate the data for a given parameter vector $(\gamma, \alpha_1, \beta_1)$ of size $n = 100$. We choose 10 parameter vectors which are closed to boundary and also not close to boundary. The parameter vectors are: $(0.2, 0.4, 0.1)$, $(0.1, 0.6, 0.3)$, $(0.3, 0.4, 0.2)$, $(0.1, 0.7, 0.2)$, $(0.4, 0.3, 0.6)$, $(0.2, 0.8, 0.1)$, $(0.3, 0.1, 0.8)$, $(0.1, 0.2, 0.3)$, $(0.3, 0.1, 0.4)$ and $(0.5, 0.2, 0.3)$.
- *Step 2: Initialize Parameters*: Following Wurtz et al. (2009), in the second step, we set the initial values for $\alpha_1 = 0.1$ and $\beta_1 = 0.8$ On the other hand, we take the value of γ to be the mean of generated data in step 1, $\mu_{X_t} = \dfrac{\gamma}{1-\alpha_1-\beta_1}$, namely, $\gamma = 0.1\mu_{X_t}$, (see Ferland et al. (2006)). The choice of initial values does not influence the final estimates. It only impact on the computational time.
- *Step 3- Parameter estimation*: The estimated values of γ, α_1 and β_1 can be obtained as follows:

◇ For LS and MLE methods, we use command in R-cran to minimize the Equation (5.22) and negative likelihood in Equation (5.23).

◇ For QEF method, we solve the simultaneous optimal Equations (5.25)-(5.26) using command in R-cran.

• *Step 4- Replication*: Step 1- Step 3 are repeated $N = 500$ times and compute the empirical mean, bias, standard error and mean square error for each parameter.

These steps are repeated for series of length $n = 1500$ and $n = 2000$.

5.4.5 The Result

The performance is measured based on mean, bias, standard error(SE) and mean squared error (MSE) for each parameters in LS, MLE and QEF estimators as shown in Table 5.2-5.11.

Discussion

The simulation results are shown in Tables 5.2-5.11. A number of interesting results can be highlighted. Firstly, for the set of parameters in stationarity cases, Table 5.2-5.6, the QEF method is comparable with MLE method and outperforms LS method such that smaller values of biasness, SE and MSE are observed for QEF. Secondly, when the data are close to non-stationarity cases, that is when $\alpha_1 + \beta_1$ approaches unity, we can see from Table 5.7-5.11, QEF is clearly superior than MLE and LS in some the measures especially their SE. Thirdly, as n increases from small sample size, $n = 100$ to large sample sizesm $n = 1500$ and $n = 2000$, the results show that there is a decrease in the values of the biasness, standard errors and mean square standard error for all parameters for QEF method as shown in all tables. Hence, we can conclude that QEF method perform well compared to LS and MLE methods in this IVTS model.

Table 5.2: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.2$, and $\beta_1 = 0.3$

		n=100			n=1500			n=2000		
		MLE	LS	QEF	MLE	LS	QEF	MLE	LS	QEF
$\hat{\gamma}$										
Mean		0.146	0.149	0.144	0.068	0.074	0.077	0.072	0.072	0.090
Bias		0.046	0.049	0.044	0.032	0.026	0.023	0.033	0.028	0.010
SE		0.058	0.062	0.055	0.021	0.022	0.019	0.017	0.017	0.013
MSE		9.12E-03	2.11E-02	8.58E-03	1.38E-03	1.95E-03	1.39E-03	9.06E-04	1.20E-03	9.08e-04
$\hat{\alpha}_1$										
Mean		0.102	0.102	0.108	0.147	0.143	0.148	0.136	0.142	0.186
Bias		0.098	0.098	0.092	0.053	0.057	0.052	0.064	0.058	0.014
SE		0.071	0.079	0.068	0.021	0.029	0.020	0.012	0.025	0.012
MSE		1.11E-02	1.52E-02	1.15E-02	5.98E-03	6.40E-03	5.99E-03	2.29E-03	4.76E-03	2.30e-03
$\hat{\beta}_1$										
Mean		0.342	0.348	0.343	0.322	0.319	0.319	0.328	0.320	0.307
Bias		0.042	0.048	0.043	0.022	0.019	0.019	0.023	0.020	0.007
SE		0.073	0.077	0.073	0.013	0.015	0.014	0.010	0.011	0.009
MSE		2.12E-02	2.18E-02	2.13E-02	5.64E-03	7.04E-03	5.86E-03	2.01E-03	3.94E-35	2.08e-03

Table 5.3: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.4$, and $\beta_1 = 0.1$

		n=100			n=1500			n=2000		
		MLE	LS	QEF	MLE	LS	QEF	MLE	LS	QEF
$\hat{\gamma}$										
Mean		0.138	0.135	0.139	0.155	0.155	0.156	0.159	0.156	0.161
Bias		0.062	0.065	0.061	0.045	0.045	0.044	0.041	0.044	0.039
SE		0.041	0.044	0.041	0.033	0.036	0.029	0.028	0.030	0.021
MSE		1.21E-02	1.25E-02	1.23E-02	3.03E-03	3.17E-03	2.99E-03	1.01E-03	2.12E-03	1.03e-03
$\hat{\alpha}_1$										
Mean		0.218	0.213	0.219	0.312	0.309	0.319	0.312	0.303	0.343
Bias		0.182	0.187	0.181	0.088	0.091	0.081	0.088	0.097	0.057
SE		0.155	0.160	0.156	0.047	0.052	0.043	0.041	0.048	0.037
MSE		3.11E-02	3.18E-02	3.12E-02	1.69E-03	2.46E-03	2.12E-03	1.04E-03	1.54E-03	1.05e-03
$\hat{\beta}_1$										
Mean		0.231	0.239	0.229	0.187	0.189	0.117	0.120	0.118	0.104
Bias		0.132	0.139	0.129	0.087	0.089	0.017	0.020	0.018	0.004
SE		0.145	0.148	0.142	0.022	0.049	0.023	0.010	0.023	0.009
MSE		3.12E-02	3.44E-02	3.18E-02	5.19E-03	5.46E-03	4.77E-03	2.07E-03	2.38E-03	2.07e-03

Table 5.4: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1,$ and $\beta_1 = 0.4$

	n=100			n=1500			n=2000		
	MLE	LS	QEF	MLE	LS	QEF	MLE	LS	QEF
$\hat{\gamma}$									
Mean	0.218	0.217	0.219	0.249	0.226	0.250	0.241	0.216	0.281
Bias	0.082	0.083	0.081	0.051	0.074	0.050	0.059	0.084	0.019
SE	0.106	0.112	0.107	0.081	0.090	0.082	0.072	0.086	0.076
MSE	9.46E-03	9.68E-03	9.49E-03	2.31E-03	4.22E-03	3.31E-03	1.55E-03	2.21E-03	1.59e-03
$\hat{\alpha}_1$									
Mean	0.059	0.057	0.061	0.074	0.068	0.076	0.074	0.068	0.092
Bias	0.041	0.043	0.039	0.026	0.032	0.024	0.026	0.032	0.008
SE	0.156	0.159	0.152	0.034	0.065	0.036	0.008	0.055	0.009
MSE	7.65E-03	8.12E-03	7.71E-03	1.25E-03	4.38E-03	1.53E-03	9.76E-04	1.23E-03	9.82e-04
$\hat{\beta}_1$									
Mean	0.452	0.468	0.447	0.410	0.455	0.407	0.405	0.424	0.398
Bias	0.052	0.068	0.047	0.010	0.055	0.007	0.005	0.024	0.002
SE	0.079	0.082	0.077	0.019	0.019	0.017	0.008	0.018	0.010
MSE	1.23E-02	2.13E-02	1.28E-02	7.75E-03	8.91E-03	7.94E-03	2.31E-03	6.05E-03	2.26e-03

Table 5.5: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.4,$ and $\beta_1 = 0.2$

	n=100			n=1500			n=2000		
	MLE	LS	QEF	MLE	LS	QEF	MLE	LS	QEF
$\hat{\gamma}$									
Mean	0.154	0.149	0.156	0.212	0.209	0.229	0.211	0.205	0.277
Bias	0.146	0.151	0.144	0.088	0.091	0.071	0.089	0.095	0.023
SE	0.193	0.201	0.192	0.037	0.041	0.036	0.031	0.036	0.021
MSE	5.33E-03	5.38E-03	5.31E-03	1.05E-03	1.42E-03	1.06E-03	9.56E-04	1.92E-03	9.56e-04
$\hat{\alpha}_1$									
Mean	0.315	0.311	0.318	0.389	0.411	0.391	0.389	0.379	0.398
Bias	0.085	0.089	0.082	0.011	0.011	0.009	0.011	0.021	0.002
SE	0.145	0.147	0.142	0.028	0.041	0.027	0.025	0.037	0.021
MSE	8.91E-03	9.21E-03	8.75E-03	2.58E-03	3.86E-03	2.27E-03	1.04E-03	2.21E-03	1.04e-03
$\hat{\beta}_1$									
Mean	0.257	0.259	0.251	0.223	0.225	0.210	0.224	0.228	0.197
Bias	0.057	0.059	0.051	0.023	0.025	0.010	0.024	0.028	0.003
SE	0.142	0.148	0.144	0.067	0.076	0.066	0.017	0.068	0.012
MSE	8.12E-03	9.05E-03	8.23E-03	1.06E-03	1.14E-03	1.01E-06	9.01E-04	1.19E-03	9.21e-04

Table 5.6: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.5, \alpha_1 = 0.2$, and $\beta_1 = 0.3$

		n=100			n=1500			n=2000		
		MLE	LS	QEF	MLE	LS	QEF	MLE	LS	QEF
$\hat{\gamma}$										
Mean		0.451	0.453	0.457	0.486	0.484	0.487	0.485	0.478	0.491
Bias		0.049	0.047	0.043	0.014	0.016	0.013	0.015	0.022	0.009
SE		0.119	0.119	0.116	0.09	0.094	0.09	0.069	0.079	0.069
MSE		8.98E-03	9.42E-03	9.01E-03	4.73E-03	4.75E-03	4.11E-03	1.38E-03	2.92E-03	1.39e-03
$\hat{\alpha}_1$										
Mean		0.113	0.110	0.115	0.143	0.137	0.145	0.214	0.141	0.182
Bias		0.087	0.090	0.085	0.057	0.063	0.055	0.014	0.059	0.018
SE		0.087	0.091	0.082	0.023	0.034	0.022	0.016	0.029	0.015
MSE		9.85E-03	1.26E-02	9.81E-03	6.59E-03	8.22E-03	6.54E-03	1.48E-03	2.85E-03	1.50e-03
$\hat{\beta}_1$										
Mean		0.351	0.358	0.347	0.322	0.319	0.315	0.323	0.320	0.309
Bias		0.051	0.058	0.047	0.022	0.019	0.015	0.023	0.020	0.009
SE		0.056	0.062	0.055	0.010	0.011	0.010	0.006	0.009	0.006
MSE		1.23E-03	1.38E-03	1.18E-03	5.36E-03	7.46E-03	5.92E-03	1.11E-03	3.59E-03	1.15e-03

Table 5.7: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.6,$ and $\beta_1 = 0.3$

		n=100			n=1500			n=2000		
		MLE	LS	QEF	MLE	LS	QEF	MLE	LS	QEF
$\hat{\gamma}$										
Mean		0.054	0.051	0.068	0.079	0.074	0.082	0.077	0.069	0.091
Bias		0.046	0.049	0.032	0.021	0.026	0.018	0.023	0.031	0.009
SE		0.714	0.715	0.511	0.509	0.514	0.01	0.508	0.513	0.005
MSE		2.61E-01	2.70E-01	5.55E-02	1.98E-02	2.44E-02	1.52E-03	9.68E-02	1.55E-02	9.71e-04
$\hat{\alpha}_1$										
Mean		0.645	0.651	0.631	0.582	0.572	0.584	0.582	0.581	0.590
Bias		0.045	0.051	0.031	0.018	0.028	0.016	0.018	0.019	0.010
SE		0.712	0.723	0.415	0.526	0.538	0.029	0.523	0.528	0.020
MSE		4.32E-02	4.85E-02	1.22E-02	6.49E-02	6.92E-02	5.34E-03	1.42E-02	1.85E-02	1.48e-03
$\hat{\beta}_1$										
Mean		0.219	0.216	0.231	0.221	0.220	0.225	0.221	0.218	0.290
Bias		0.081	0.084	0.069	0.079	0.080	0.075	0.079	0.082	0.010
SE		0.612	0.618	0.431	0.428	0.442	0.029	0.425	0.436	0.015
MSE		5.21E-02	5.38E-02	9.81E-03	4.76E-02	4.82E-02	3.45E-03	1.57E-02	2.61E-02	1.58e-03

Table 5.8: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.7,$ and $\beta_1 = 0.2$

		n=100			n=1500			n=2000		
		MLE	LS	QEF	MLE	LS	QEF	MLE	LS	QEF
$\hat{\gamma}$										
Mean		0.053	0.050	0.068	0.074	0.072	0.076	0.074	0.071	0.088
Bias		0.047	0.050	0.032	0.026	0.028	0.024	0.026	0.029	0.012
SE		0.811	0.819	0.528	0.609	0.617	0.009	0.508	0.515	0.002
MSE		8.92E-02	9.11E-02	9.48E-03	5.69E-02	6.58E-02	5.72E-03	1.62E-02	2.41E-02	1.68e-03
$\hat{\alpha}_1$										
Mean		0.568	0.565	0.581	0.682	0.681	0.682	0.681	0.608	0.692
Bias		0.132	0.135	0.119	0.018	0.019	0.018	0.019	0.092	0.008
SE		0.751	0.759	0.544	0.431	0.451	0.021	0.517	0.544	0.002
MSE		7.45E-01	9.86E-01	8.13E-02	7.04E-02	7.45E-02	7.03E-03	5.40E-02	7.06E-02	5.45e-03
$\hat{\beta}_1$										
Mean		0.258	0.261	0.235	0.220	0.227	0.216	0.226	0.221	0.207
Bias		0.058	0.061	0.035	0.020	0.027	0.016	0.026	0.021	0.007
SE		0.817	0.825	0.584	0.534	0.558	0.037	0.618	0.544	0.018
MSE		9.13E-01	9.58E-01	7.15E-02	8.47E-02	9.53E-02	8.45E-03	2.46E-02	5.77E-02	2.48e-03

Table 5.9: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.8,$ and $\beta_1 = 0.1$

		n=100			n=1500			n=2000		
		MLE	LS	QEF	MLE	LS	QEF	MLE	LS	QEF
$\hat{\gamma}$										
Mean		0.121	0.119	0.145	0.166	0.165	0.177	0.155	0.161	0.191
Bias		0.079	0.081	0.055	0.034	0.035	0.023	0.045	0.039	0.009
SE		0.916	0.921	0.527	0.621	0.636	0.020	0.618	0.631	0.008
MSE		8.61E-02	9.11E-02	6.17E-03	1.95E-02	2.97E-02	2.18E-03	1.06E-02	1.51E-02	1.07e-03
$\hat{\alpha}_1$										
Mean		0.618	0.613	0.625	0.722	0.720	0.727	0.727	0.727	0.779
Bias		0.182	0.187	0.175	0.078	0.080	0.073	0.073	0.073	0.021
SE		0.852	0.869	0.321	0.542	0.468	0.040	0.416	0.558	0.016
MSE		1.15E-01	2.02E-01	9.21E-02	6.35E-02	7.81E-02	6.41E-03	1.33E-02	4.51E-02	1.32e-03
$\hat{\beta}_1$										
Mean		0.216	0.217	0.172	0.178	0.186	0.173	0.190	0.192	0.129
Bias		0.116	0.117	0.072	0.078	0.086	0.073	0.090	0.092	0.029
SE		0.891	0.913	0.215	0.547	0.678	0.045	0.541	0.567	0.039
MSE		8.61E-02	9.21E-02	8.08E-03	7.88E-02	7.92E-02	7.49E-03	5.93E-02	6.24E-02	5.91e-03

Table 5.10: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1$, and $\beta_1 = 0.8$

		n=100			n=1500			n=2000		
		MLE	LS	QEF	MLE	LS	QEF	MLE	LS	QEF
$\hat{\gamma}$										
Mean		0.358	0.361	0.355	0.336	0.251	0.324	0.303	0.291	0.299
Bias		0.058	0.061	0.055	0.036	0.049	0.024	0.003	0.009	0.001
SE		0.517	0.531	0.257	0.398	0.398	0.084	0.355	0.390	0.057
MSE		8.62E-02	9.18E-02	7.11E-03	1.53E-02	1.72E-02	1.59E-03	1.05E-02	1.55E-02	1.02e-03
$\hat{\alpha}_1$										
Mean		0.062	0.059	0.068	0.079	0.074	0.088	0.078	0.073	0.091
Bias		0.038	0.041	0.032	0.021	0.026	0.011	0.022	0.027	0.009
SE		0.865	0.869	0.258	0.612	0.613	0.011	0.610	0.610	0.008
MSE		8.32E-02	9.18E-02	7.88E-03	2.72E-02	5.07E-02	2.81E-03	1.01E-02	3.95E-02	1.02e-03
$\hat{\beta}_1$										
Mean		0.845	0.848	0.845	0.809	0.819	0.804	0.822	0.827	0.801
Bias		0.045	0.048	0.045	0.009	0.019	0.004	0.022	0.027	0.001
SE		0.897	0.916	0.512	0.425	0.43	0.021	0.511	0.526	0.008
MSE		9.12E-02	9.88E-02	7.59E-03	1.72E-02	2.32E-02	1.83E-03	1.05E-02	1.42E-02	1.07e-03

Table 5.11: Simulation results for INGARCH $(1,1)$ with $\gamma = 0.4, \alpha_1 = 0.3$, and $\beta_1 = 0.6$

		n=100			n=1500			n=2000		
		MLE	LS	QEF	MLE	LS	QEF	MLE	LS	QEF
$\hat{\gamma}$										
Mean		0.353	0.351	0.379	0.388	0.384	0.389	0.382	0.378	0.394
Bias		0.047	0.049	0.021	0.012	0.016	0.011	0.018	0.022	0.006
SE		0.187	0.190	0.128	0.046	0.050	0.048	0.039	0.044	0.035
MSE		8.92E-02	9.25E-02	9.12E-03	3.04E-02	3.68E-02	2.33E-03	1.03E-02	1.98E-02	1.03e-03
$\hat{\alpha}_1$										
Mean		0.198	0.192	0.205	0.205	0.218	0.221	0.203	0.216	0.261
Bias		0.102	0.108	0.095	0.095	0.082	0.079	0.097	0.084	0.039
SE		0.516	0.528	0.214	0.014	0.026	0.018	0.002	0.015	0.007
MSE		9.38E-02	9.87E-02	8.26E-03	1.52E-02	1.85E-02	1.55E-03	1.16E-02	1.29E-02	1.17e-03
$\hat{\beta}_1$										
Mean		0.452	0.449	0.475	0.574	0.549	0.582	0.589	0.578	0.596
Bias		0.148	0.151	0.125	0.026	0.051	0.018	0.011	0.022	0.004
SE		0.058	0.062	0.035	0.019	0.021	0.019	0.008	0.012	0.008
MSE		7.32E-02	7.89E-02	1.02E-02	2.19E-02	2.47E-02	2.39E-03	1.05E-02	1.59E-02	1.08e-03

5.5 Real Example

We consider 108 monthly strike data from January 1994 to December 2002 found in Jung et al. (2005). The data are available at U.S. Bureau of Labor Statistics website (http://www.bls.gov/wsp/). It describes the number of work stoppages leading to 1000 workers or more being idle in effect in the period. The number of workers considered are those who participated in work stoppages that began in the calendar year and were counted more than once if they are involved in more than one stoppages during the given period. WeiB (2010) fitted the INARCH(1) model to the data and showed that the mean and variance of such model become closer to the empirical values indicating the INARCH(1) model is a good choice. Therefore, we now consider INGARCH(1,1) model which can be as the generalized model of INARCH(1) to investigate the adequacy of the model to the data. The line plot of the data is given in Figure 5.1 where the plot shows a possible change in mean. However, it does not effect on model estimation since the INGARCH model set up the conditional mean is presented as a function of variances at previous times and previous observations

Upon fitting the INGARCH$(1,1)$ model, we obtain the parameter estimates $\hat{\boldsymbol{\theta}}$ and standard errors in parentheses using two methods, MLE and QEF estimators as shown in Table 5.12. QEF methods show smaller AIC and BIC compared to MLE method. Besides, the standard errors in QEF method also give the smaller value compare that in MLE method.

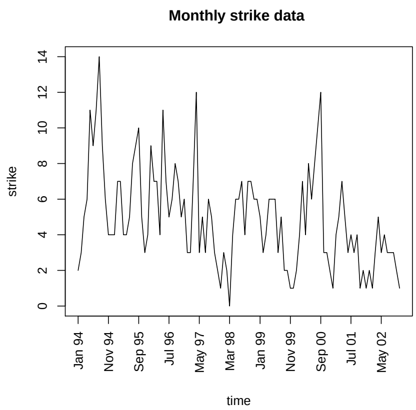


Figure 5.1: The monthly strike data from January 1994 to December 2002

Hence, from Table 5.12, our fitted model via QEF estimator for the data is

$$X_t|\Im_{t-1}^X \sim P(\lambda_t\left(\hat{\boldsymbol{\theta}}\right)); \ \lambda_t\left(\hat{\boldsymbol{\theta}}\right) = 1.56 + 0.546X_{t-1} + 0.139\lambda_{t-1} \tag{5.27}$$

All the parameter estimates $\hat{\boldsymbol{\theta}} = \left(\hat{\gamma}, \hat{\alpha}_1, \hat{\beta}_1\right)$ gained using QEF method are positive and the summation of $\hat{\alpha}_1 + \hat{\beta}_1$ lies between zero and one indicating the process is stationary.

Table 5.12: The estimated parameter of INGARCH$(1,1)$ model

Method	$\hat{\gamma}$	$\hat{\alpha}_1$	$\hat{\beta}_1$	AIC	BIC
MLE	1.53 (0.028)	0.535 (0.041)	0.157 (0.036)	385.09	377.07
QEF	1.56 (0.023)	0.546 (0.037)	0.139 (0.038)	383.45	375.43

The mean and variance obtained from our fitted model are 4.98 and 7.68 respectively. Meanwhile, the empirical mean and empirical variance are 4.94 and 7.92 accordingly. Weiß (2010) applied the data into INARCH(1) and obtained the variance as 8.37. Therefore, from the mean and variance for this model, we can see that the INGARCH $(1,1)$ have the smaller variance. Therefore, we can claim that the data fit better in INGARCH $(1,1)$ model compared to the INARCH(1). But, to strengthen the claim, next, we will investigate the Akaike information criterion (AIC), Bayesian information criterion (BIC) on INARCH(1) and INGARCH $(1,1)$ models.

Table 5.13: AIC and BIC for INARCH(1) and INGARCH $(1,1)$ models.

	AIC	BIC
INARCH(1)	463.97	468.54
INGARCH$(1,1)$	383.45	375.43

From Table 5.13, it is clear that the INGARCH$(1,1)$ model gives the smallest AIC and BIC compared to INARCH(1). It indicates that the INGARCH$(1,1)$ model fits the data better. To investigate the model fitting adequacy, we consider the Pearson residual defined as $z_t = \left(X_t - \lambda_t(\hat{\boldsymbol{\theta}})\right)/\sqrt{\lambda_t(\hat{\boldsymbol{\theta}})}$. According to Kedem & Fokianos (2005), under the specified model, there are two requirements should be satisfied. First, the sequence of z_t should have mean and variance close to 0 and 1 respectively and secondly, the sequence does not have serial correlation. For our data, we found that the mean and variance of Pearson residuals is close to zero and unity which are 0.032 and 1.009 respectively. Therefore, the first condition of specified model is satisfied. For the second condition, we use the Ljung-Box (LB) statistics to determine the existence of a serial correlation.

Table 5.14: Diagnostics for INGARCH(1,1) model

	$LB_{30}(z_t)$	$LB_{30}(z_t^2)$
χ^2	24.3	21.6
p-value	0.758	0.868

From Table 5.14, the p-value is larger that the significance level, $\alpha = 0.05$ indicating that there is no significant serial correlation in the residual. Therefore, the second requirement is fulfilled. We conclude the INGARCH $(1,1)$ model fits the data well.

On the other hand, the model fitting adequacy can also be investigated using the randomness of Pearson residual plot and the cumulative periodagram plot. For the latter, according to Brockwell & Davis (2013), the model is adequate if the plot do not cross the dotted line.

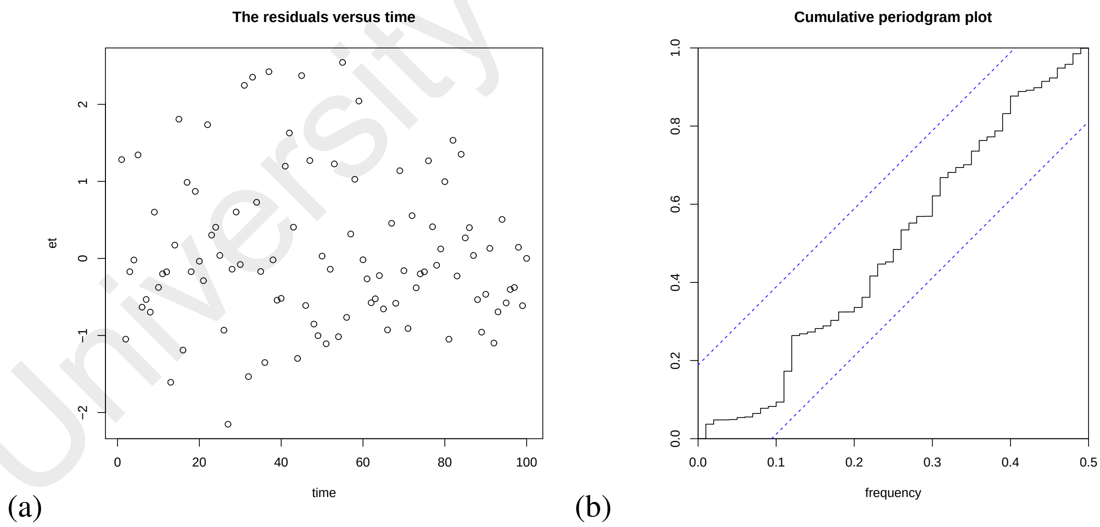


Figure 5.2: (a)The Pearson residual plot. (b)The periodagram plot

Figure 5.2(a) gives the plot of z_t given t. It is clearly shown that such residuals are randomly distributed and do not have specified trend. From Figure 5.2(b), we can obviously see that the plot does not exceed the dotted line. Thus, we conclude that the INGARCH$(1,1)$ models fit the data well.

5.6 Summary of The Chapter

In this chapter, we focussed only on INGARCH(p,q) model. We presented a general approach for any p and q instead of given by Ferland who gives the results only for $p = 1$ and $q = 1$. We derived the optimal QEF functions and their information matrix. To see the performance of this method, we carried out a simulation study in which the QEF was compared to MLE and LS. The results show that the QEF estimator is superior compared to MLE and LS methods. Lastly, we applied the methodology on a set of real data and showed that the INGARCH$(1,1)$ model was a good fit and that the QEF worked well in this case.

CHAPTER 6

NBINGARCH(p,q) MODEL

6.1 Introduction

Zhu (2011) highlighted that the INGARCH(p,q) model has few disadvantages. Firstly, it does not accommodate covariates. Secondly, the ACF is always positive and lastly the conditional distribution of INGARCH(p,q) model following Poisson, meaning that the conditional mean and conditional variance are the same which will lead to overdispersion

Therefore, to overcome these drawbacks, Zhu (2011) proposed the same model with negative binomial as conditional distribution which is denoted as NBINGARCH(p,q) . The model is defined as

$$
\begin{aligned}
\left(X_t|\Im_{t-1}^X\right) &\sim NB\left(r,p_t\right), \\
\frac{1-p_t}{p_t} = \lambda_t &= \gamma + \sum_{i=1}^{p}\alpha_i X_{t-i} + \sum_{j=1}^{q}\beta_j\lambda_{t-j}.
\end{aligned} \tag{6.1}
$$

where $\gamma > 0$, $\alpha_i \geq 0$, $i = 1,2,\ldots,p$, $j = 1,2,\ldots,q$, r is the number of successes trials, $\Im_{t-1}^X$ is the $\sigma-$field generated from $X_{t-1}, X_{t-2},\ldots,X_1$ and p_t is the probability of successes. Here, as in INGARCH(p,q) model, the model satisfies the stationarity if and only if $0 < r\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j < 1$.

6.2 The Moment Properties

In this Section, we will derive the moments of both unconditional and conditional distribution of NBINGARCH(p,q) model up to order four.

6.2.1 The Moments of Conditional Distribution in NBINGARCH(p,q) Model

For NBINGARCH(p,q) process, the conditional distribution follows negative binomial distribution. To simplify, let $Y_t = X_t|\Im_{t-1}^X$ follows the negative binomial distribution, $Y_t \sim NB(r,p_t)$ where p_t is the probability of successes. According to Johnson et al. (2005), probability generating function (pgf) of negative binomial distribution is $G(z) = \left(\frac{1-q_t}{1-q_t z}\right)^r$

with $q_t = 1 - p_t$. The first derivative can be obtained as below:

$$\begin{aligned}
G'(z) &= \frac{d}{dz}\left(\frac{1-q_t}{1-q_t z}\right)^r \\
&= \frac{d}{dz}(1-q_t)^r(1-q_t z)^{-r} \\
&= (1-q_t)^r(-r)(1-q_t z)^{-r-1}(-q_t) \\
&= rq_t(1-q_t)^r(1-q_t z)^{-r-1}. \qquad (6.2)
\end{aligned}$$

Using the same approach, the second and third derivatives are

$$G''(z) = rq_t^2(1-q_t)^r(r+1)(1-q_t z)^{-r-2},$$

and

$$G'''(z) = rq_t^3(1-q_t)^r(r+1)(r+2)(1-q_t z)^{-r-3},$$

respectively. Therefore, the k-th derivative is

$$G^k(z) = rq_t^k(1-q_t)^r(r+1)(r+2)\dots(r+k-1)(1-q_t z)^{-r-k}.$$

From Equation (6.2), the mean is:

$$\mu_t = E(y_t) = G'(1) = \frac{rq_t}{1-q_t}. \qquad (6.3)$$

The second derivative is

$$G''(1) = E\left(y_t^2\right) - E(y_t) = \frac{rq_t^2(r+1)}{(1-q_t)^2}.$$

Rearrange it to be $E\left(y^2\right)$, therefore,

$$E\left(y_t^2\right) = \frac{rq_t^2(r+1)}{(1-q_t)^2} + E(y_t) = \frac{rq_t^2(r+1)}{(1-q_t)^2} + \frac{rq_t}{1-q_t} = \frac{rq_t}{1-q_t}\left(\frac{rq_t+1}{1-q_t}\right),$$

then, the variance for negative binomial distribution is:

$$var(y_t) = E\left(y_t^2\right) - [E(y_t)]^2,$$

$$
\begin{aligned}
&= \frac{rq_t}{1-q_t}\left(\frac{rq_t+1}{1-q_t}\right)-\left(\frac{rq_t}{1-q_t}\right)^2, \\
&= \frac{rq_t}{(1-q_t)^2}\left(rq_t+1-rq_t\right), \\
&= \frac{rq_t}{(1-q_t)^2}. \qquad (6.4)
\end{aligned}
$$

Using the same approach, we obtain

$$E\left(y_t^3\right)=\frac{rq_t}{(1-q_t)^3}\left(r^2q_t^2+3rq_t+1+q_t\right)$$

and

$$E\left(y_t^4\right)=\frac{rq_t}{(1-q_t)^4}\left(r^3q_t^3+6r^2q_t^2+7rq_t+4rq_t^2+q_t^2+1+4q_t\right)$$

which lead to

$$E\left[(y-\mu_t)^3\right]=\frac{rq_t}{(1-q_t)^3}(1+q_t)$$

and

$$E\left[(y-\mu_t)^4\right]=\frac{rq_t}{(1-q_t)^4}\left(3rq_t+q_t^2+1+4q_t\right)$$

respectively. Therefore, the skewness of the negative binomial distribution is given by:

$$
\begin{aligned}
\Gamma_t^{(NB)} &= \frac{E\left[(y_t-\mu_t)^3\right]}{(var(y_t))^{3/2}}, \\
&= \frac{\frac{rq_t}{(1-q_t)^3}(1+q_t)}{\left(\frac{rq_t}{(1-q_t)^2}\right)^{3/2}}, \\
&= \frac{1+q_t}{\sqrt{rq_t}}, \qquad (6.5)
\end{aligned}
$$

and the kurtosis for negative binomial is:

$$
\begin{aligned}
K_t^{(NB)} &= \frac{E\left[(y_t-\mu_t)^4\right]}{(var(y_t))^2}, \\
&= \frac{\frac{rq_t}{(1-q_t)^4}\left(3rq_t+q_t^2+1+4q_t\right)}{\left(\frac{rq_t}{(1-q_t)^2}\right)^2}, \\
&= \frac{3rq_t+q_t^2+1+4q_t}{rq_t}. \qquad (6.6)
\end{aligned}
$$

We can represent the moment properties of negative binomial distribution in terms of λ_t where define

$$\lambda_t = \frac{1-p_t}{p_t}.$$

Therefore, the first four moments of negative binomial distribution are:

$$\begin{aligned} \mu_t &= r\lambda_t, \ \sigma_t^2 = r\lambda_t(1+\lambda_t), \\ \Gamma^{(NB)} &= \frac{1+2\lambda_t}{\sqrt{r\lambda_t(1+\lambda_t)}}, \ K^{(NB)} = \frac{\lambda_t(1+\lambda_t)(3r+4)+2\lambda_t^2-2\lambda_t+1}{r\lambda_t(1+\lambda_t)}. \end{aligned} \quad (6.7)$$

6.2.2 The Moments Properties of Unconditional Distribution in NBINGARCH(p,q) Model

The mean of the NBINGARCH(p,q) is $\mu = E(X_t) = E\left[E\left(X_t|\Im_{t-1}^X\right)\right] = E[r\lambda_t]$, therefore, under large t, we have $\mu = r\lambda$. The variance of the martingale difference u_t is

$$\sigma_u^2 = E\left(u_t^2\right) = rE(\lambda_t) + rE\left(\lambda_t^2\right). \quad (6.8)$$

For the skewness and kurtosis of the u_t, we have

$$\Gamma^{(u)} = \frac{E\left(u_t^3\right)}{\left\{E\left(u_t^2\right)\right\}^{3/2}} = \frac{E\left(r\lambda_t+3r\lambda_t^2+2r\lambda_t^3\right)}{\left[E\left(r\lambda_t+r\lambda_t^2\right)\right]^{3/2}}. \quad (6.9)$$

Moreover, for the excess kurtosis of the model, using the same approach as skewness of u_t, $\Gamma^{(u)}$,

$$\begin{aligned} K^{(u)} &= \frac{E\left(u_t^4\right)}{\left\{\left\{E\left(u_t^2\right)\right\}\right\}^2} - 3 \\ &= \frac{E\left(E\left[\left(\{X_t - r\lambda_t\}|\Im_{t-1}^X\right)^2\right]\right)}{\sigma_u^4} - 3 \\ &= \frac{E\left(r^2\lambda_t^2 - 8r\lambda_t^2 - 4r\lambda_t + r^2\lambda_t^3 - 4r\lambda_t^3\right)}{\left[E\left(r\lambda_t + r\lambda_t^2\right)\right]^2}. \end{aligned} \quad (6.10)$$

For large t and $\lambda = \frac{\mu}{r}$, Equations (6.8), (6.9) and (6.10) become

$$\sigma_u^2 = \mu(1+\frac{\mu}{r}), \quad (6.11)$$

$$\Gamma^{(u)} = \frac{\mu\left(r^2+3r\mu+2\mu^2\right)}{r^{1/2}\left(r\mu+\mu^2\right)^{3/2}}, \text{ and} \tag{6.12}$$

$$K^{(u)} = \frac{\mu^2r^2-8\mu^2r-8\mu r^2+\mu^3r-4\mu^3}{\left(r\mu+\mu\right)^2} \tag{6.13}$$

respectively.

Hence, by using Equations (6.11), (6.12) and (6.13), we can find the variance, skewness and kurtosis of NBINGARCH(p,q) process which are

$$\begin{aligned}
\sigma_X^2 &= \mu\left(1+\frac{\mu}{r}\right)\sum_{j=0}^{\infty}\psi_j^2, \\
\Gamma^{(X)} &= \frac{\mu\left(r^2+3r\mu+2\mu^2\right)\sum_{j=0}^{\infty}\psi_j^3}{r^{1/2}\left(r\mu+\mu^2\right)^{3/2}\left(\sum_{j=0}^{\infty}\psi_j^2\right)^{3/2}}, \text{ and} \\
K^{(X)} &= 3+\frac{\left(\mu^2r^2-8\mu^2r-8\mu r^2+\mu^3r-4\mu^3\right)\sum_{j=0}^{\infty}\psi_j^4}{\left(r\mu+\mu\right)^2\left(\sum_{j=0}^{\infty}\psi_j^2\right)^2}
\end{aligned}$$

accordingly.

Zhu (2011) derived only on the mean and variance of unconditional distribution in the NBINGARCH(p,q) model. The derivation for variance presented in Zhu's work is very complicated and with the order of q to be zero. Here we produced a closed form expression for the moments up to order four for all p and q cases and the derivation is easier to understand. This finding is very important in understanding the behavior of the model through its higher order moment properties.

6.2.3 Empirical Study

Again, similar to Section 5.2.3, to demonstrate the moments structure are correctly derived, we do the empirical study for the cases when the process is close and far from the boundary of stationarity. Here, we generate 500 samples of size 2000 and fixed the value $r = 2$, then, calculate the mean and the mean square error of the estimated moments as tabulated in Table 6.1. It can be seen that the estimated values are close to the true values.

Table 6.1: Generated data and true values for the moment structures with $\gamma = 0.1$

		$\alpha_1 = 0.1$			$\alpha_1 = 0.5$		
		Estimated	True	MSE	Estimated	True	MSE
$\beta = 0.2$							
	μ_X	0.8127	0.8213	7.21E-05	1.142	1.151	1.22E-04
	σ_2^X	0.8177	0.8156	1.23E-03	1.315	1.329	8.12E-03
	$\Gamma^{(X)}$	1.581	1.567	1.66E-04	1.079	1.082	5.58E-05
	$K^{(X)}$	4.032	4.108	4.11E-03	4.014	4.025	1.28E-05
$\beta = 0.4$							
	μ_X	0.8213	0.8236	1.44E-03	1.774	1.792	5.23E-03
	σ_2^X	0.8321	0.8344	9.12E-03	1.158	1.167	3.11E-03
	$\Gamma^{(X)}$	1.210	1.251	8.21E-03	1.582	1.596	4.53E-03
	$K^{(X)}$	4.234	4.247	5.14E-03	4.771	4.785	8.11E-03

6.3 Quadratic Estimating Functions on NBINGARCH(p,q) Model

The third contribution of this study is to find the optimal quadratic estimating functions for each of our models. To apply the QEF in estimating the parameters of interest, the first four conditional moments of NBINGARCH(p,q) are required and they are

$$\mu_t(\boldsymbol{\theta}) = r\lambda_t(\boldsymbol{\theta}), \tag{6.14}$$

$$\sigma_t^2(\boldsymbol{\theta}) = r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta})), \tag{6.15}$$

$$\Gamma_t(\boldsymbol{\theta}) = \frac{1+2\lambda_t(\boldsymbol{\theta})}{\sqrt{r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))}}, \tag{6.16}$$

and

$$\kappa_t(\boldsymbol{\theta}) = \frac{r\lambda_t^2(\boldsymbol{\theta}) - 4\lambda_t(\boldsymbol{\theta}) - 4}{r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))}. \tag{6.17}$$

The martingale differences of m_t and s_t for such model are defined to be $m_t(\boldsymbol{\theta}) = y_t - r\lambda_t(\boldsymbol{\theta})$ and $s_t(\boldsymbol{\theta}) = m_t^2(\boldsymbol{\theta}) - \sigma_t^2(\boldsymbol{\theta})$ for $t = 1,2,\ldots,n$. Then, their variances and covariance can be easily found as shown below:

$$\begin{aligned} \langle m \rangle_t &= \sigma_t^2(\boldsymbol{\theta}) = r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta})), \qquad (6.18) \\ \langle s \rangle_t &= \sigma_t^4(\boldsymbol{\theta})(\kappa_t(\boldsymbol{\theta})+2), \end{aligned}$$

$$= [r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))]^2 \left[\frac{r\lambda_t^2(\boldsymbol{\theta}) - 4\lambda_t(\boldsymbol{\theta}) - 4}{r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))} + 2\right],$$

$$= [r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))]^2 \left[\frac{r\lambda_t^2(\boldsymbol{\theta}) - 4\lambda_t(\boldsymbol{\theta}) - 4 + 2r\lambda_t(\boldsymbol{\theta}) + 2r\lambda_t^2(\boldsymbol{\theta})}{r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))}\right],$$

$$= r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))\left[\lambda_t^2(\boldsymbol{\theta})\{1+2r\} + 2\lambda_t(\boldsymbol{\theta})\{r-2\} - 4\right], \tag{6.19}$$

$$\text{and } \langle m,s\rangle_t = \sigma_t^4(\boldsymbol{\theta})\Gamma_t(\boldsymbol{\theta}),$$

$$= [r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))]^{3/2}\left[\frac{1+2\lambda_t(\boldsymbol{\theta})}{\sqrt{r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))}}\right],$$

$$= r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))[1+2\lambda_t(\boldsymbol{\theta})]. \tag{6.20}$$

On the other hand, in applying the QEF estimator, we have to find the derivatives of the mean and variance with respect to the parameter $\boldsymbol{\theta}$, $\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\theta}$ and $\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\theta}$ respectively

$$\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\theta} = \frac{\partial(r\lambda_t(\boldsymbol{\theta}))}{\partial\theta}$$

$$= r\lambda_t'(\boldsymbol{\theta}), \tag{6.21}$$

and

$$\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\theta} = \frac{\partial[r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))]}{\partial\theta}, \tag{6.22}$$

$$= r\lambda_t'(\boldsymbol{\theta})[1+\lambda_t(\boldsymbol{\theta})] + \lambda_t'(\boldsymbol{\theta})r\lambda_t(\boldsymbol{\theta}),$$

$$= r\lambda_t'(\boldsymbol{\theta})[1+2\lambda_t(\boldsymbol{\theta})]. \tag{6.23}$$

The parameters of interests are given as $\boldsymbol{\theta} = (\gamma, \alpha_1, \ldots, \alpha_p, \beta_1, \beta_2, \ldots, \beta_q)'$. Therefore,

$$\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \left(r\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\gamma}, r\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\alpha_1}\ldots, r\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\alpha_p}, r\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\beta_1}, \ldots, r\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\beta_q}\right)',$$

$$= \left(B_{(1,t)}, B_{(2,1,t)}, \ldots, B_{(2,p,t)}, \ldots, B_{(3,1,t)}, \ldots, B_{(3,q,t)}\right)'.$$

$$\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \left(D_t\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\gamma}, D_t\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\alpha_1}, \ldots, D_t\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\alpha_p}, D_t\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\beta_1}, \ldots, D_t\frac{\partial\lambda_t(\boldsymbol{\theta})}{\partial\beta_q},\right)',$$

$$= \left(H_{(1,t)}, H_{(2,1,t)}, \ldots, H_{(2,p,t)}, \ldots, H_{(3,1,t)}, \ldots, H_{(3,q,t)},\right)',$$

where $D_t = r(1+2\lambda_t(\boldsymbol{\theta}))$, $B_{(i,j,t)}$ is the partial derivative of the mean with respect to each parameter and $H_{(i,j,t)}$ is the partial derivative of the variance with respect to each parameter. In order to obtain the optimal function, we let $R_t = \left(1 - \frac{\langle m,s\rangle_t^2}{\langle m\rangle_t \langle s\rangle_t}\right)^{-1}$, $B_{k,t}^m = \frac{-B_{k,t}}{\langle m\rangle_t}$, $H_{k,t}^m = \frac{H_{k,t}\langle m,s\rangle_t}{\langle m\rangle_t \langle s\rangle_t}$, $T_{k,t}^v = \frac{B_{k,t}\langle m,s\rangle_t}{\langle m\rangle_t \langle s\rangle_t}$, and $Z_{k,t}^v = \frac{-H_{k,t}}{\langle s\rangle_t}$. Therefore, the optimal estimating function for NBINGARCH(p,q) is given by

$$g_Q^*(\boldsymbol{\theta}) = \sum_{t=1}^{n} \left(\mathbf{a}_{t-1}^* m_t + \mathbf{b}_{t-1}^* s_t\right),$$

where

$$\mathbf{a}_{t-1}^* = R_t \begin{pmatrix} B_{(1,t)}^m + H_{(1,t)}^m, B_{(2,1,t)}^m + H_{(2,1,t)}^m, \ldots, B_{(2,p,t)}^m + H_{(2,p,t)}^m, \\ B_{(3,1,t)}^m + H_{(3,1,t)}^m, \ldots, B_{(3,q,t)}^m + H_{(3,q,t)}^m \end{pmatrix},$$

$$\mathbf{b}_{t-1}^* = R_t \begin{pmatrix} T_{(1,t)}^v + Z_{(1,t)}^v, T_{(2,1,t)}^v + Z_{(2,1,t)}^v, \ldots, T_{(2,p,t)}^v + Z_{(2,p,t)}^v, \\ T_{(3,1,t)}^v + Z_{(3,1,t)}^v, \ldots, T_{(3,q,t)}^v + Z_{(3,q,t)}^v \end{pmatrix}.$$

Thus, the optimal quadratic estimating functions for each component of $\boldsymbol{\theta}$ are

$$\begin{aligned} g_Q^*(\gamma) &= \sum_{t=1}^{n} R_t \left[\left(B_{(1,t)}^m + H_{(1,t)}^m\right) m_t + \left(T_{(1,t)}^v + Z_{(1,t)}^v\right) s_t\right], \\ g_Q^*(\alpha_i) &= \sum_{t=1}^{n} R_t \left[\left(B_{(2,i,t)}^m + H_{(2,i,t)}^m\right) m_t + \left(T_{(2,i,t)}^v + Z_{(2,i,t)}^v\right) s_t\right], \; i = 1, \ldots, p, \end{aligned} \tag{6.24}$$

$$g_Q^*(\beta_j) = \sum_{t=1}^{n} R_t \left[\left(B_{(3,j,t)}^m + H_{(3,j,t)}^m\right) m_t + \left(T_{(3,j,t)}^v + Z_{(3,j,t)}^v\right) s_t\right], \; j = 1, \ldots, q \; . \tag{6.25}$$

To estimate the parameters of interest, we can solve $g_Q^*(\boldsymbol{\theta})$ using R-cran software. The information matrix of the optimal estimating function for $\boldsymbol{\theta}$ is given by

$$\mathbf{I}_{\mathbf{g}_Q^*}(\boldsymbol{\theta}) = \begin{pmatrix} I^Q_{\gamma\gamma} & I^Q_{\gamma\alpha_i} & I^Q_{\gamma\beta_j} \\ I^Q_{\alpha_i\gamma} & I^Q_{\alpha_i\alpha_i} & I^Q_{\alpha_i\beta_j} \\ I^Q_{\beta_j\gamma} & I^Q_{\beta_j\alpha_i} & I^Q_{\beta_j\beta_j} \end{pmatrix}$$

where the elements in the matrix are

$$I^Q_{\gamma\gamma} = \sum_{t=1}^{n} R_t \left[\frac{B^2_{(1,t)}}{\langle m \rangle_t} + \frac{H^2_{(1,t)}}{\langle s \rangle_t} - 2B_{(1,t)}H_{(1,t)} \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\alpha_i\alpha_i} = \sum_{t=1}^{n} R_t \left[\frac{B^2_{(2,i,t)}}{\langle m \rangle_t} + \frac{H^2_{(2,i,t)}}{\langle s \rangle_t} - 2B_{(2,i,t)}H_{(2,i,t)} \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\beta_j\beta_j} = \sum_{t=1}^{n} R_t \left[\frac{B^2_{(3,j,t)}}{\langle m \rangle_t} + \frac{H^2_{(3,j,t)}}{\langle s \rangle_t} - 2B_{(3,j,t)}H_{(3,j,t)} \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\gamma\alpha_i} = \sum_{t=1}^{n} R_t \left[\frac{B_{(2,i,t)}B_{(1,t)}}{\langle m \rangle_t} + \frac{H_{(2,i,t)}H_{(1,t)}}{\langle s \rangle_t} - (B_{(1,t)}H_{(2,i,t)} + H_{(1,t)}B_{(2,i,t)}) \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\gamma\beta_j} = \sum_{t=1}^{n} R_t \left[\frac{B_{(3,j,t)}B_{(1,t)}}{\langle m \rangle_t} + \frac{H_{(3,j,t)}H_{(1,t)}}{\langle s \rangle_t} - (B_{(1,t)}H_{(3,j,t)} + H_{(1,t)}B_{(3,j,t)}) \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right]$$

$$I^Q_{\alpha_i\beta_j} = \sum_{t=1}^{n} R_t \left[\frac{B_{(2,i,t)}B_{(3,j,t)}}{\langle m \rangle_t} + \frac{B_{(2,i,t)}H_{(3,j,t)}}{\langle s \rangle_t} - (B_{(2,i,t)}H_{(3,j,t)} + H_{(2,i,t)}B_{(3,j,t)}) \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right]$$

and $I_{\alpha_i\gamma} = I_{\gamma\alpha_i}$, $I_{\beta_j\gamma} = I_{\gamma\beta_j}$ and $I_{\beta_j\alpha_i} = I_{\alpha_i\beta_j}$.

The optimal estimating function and associated information based on m_t is given by

$$g^*_m(\theta) = -\sum_{t=1}^{n} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \left\{ \frac{X_t - r\lambda_t(\boldsymbol{\theta})}{r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))} \right\} \tag{6.26}$$

and the optimal estimating function and associated information based on s_t is given by

$$g^*_s(\theta) = -\sum_{t=1}^{n} \frac{\partial \sigma^2_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \left\{ \frac{m^2_t(\boldsymbol{\theta}) - \sigma^4_t(\boldsymbol{\theta})}{r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))\left[\lambda^2_t(\boldsymbol{\theta})\{1+2r\} + 2\lambda_t(\boldsymbol{\theta})\{r-2\} - 4\right]} \right\}. \tag{6.27}$$

We then obtain the corresponding information matrix using estimating functions based on $m_t(\boldsymbol{\theta})$ and $s_t(\boldsymbol{\theta})$ denoted as $\mathbf{I}_{\mathbf{g}^*_m}(\boldsymbol{\theta})$ and $\mathbf{I}_{\mathbf{g}^*_s}(\boldsymbol{\theta})$ respectively such that

$$\mathbf{I}_{\mathbf{g}^*_m}(\boldsymbol{\theta}) = \begin{pmatrix} I^m_{\gamma\gamma} & I^m_{\gamma\alpha_i} & I^m_{\gamma\beta_j} \\ I^m_{\alpha_i\gamma} & I^m_{\alpha_i\alpha_i} & I^m_{\alpha_i\beta_j} \\ I^m_{\beta_j\gamma} & I^m_{\beta_j\alpha_i} & I^m_{\beta_j\beta_j} \end{pmatrix} \; ; \; \mathbf{I}_{\mathbf{g}^*_s}(\boldsymbol{\theta}) = \begin{pmatrix} I^s_{\gamma\gamma} & I^s_{\gamma\alpha_i} & I^s_{\gamma\beta_j} \\ I^s_{\alpha_i\gamma} & I^s_{\alpha_i\alpha_i} & I^s_{\alpha_i\beta_j} \\ I^s_{\beta_j\gamma} & I^s_{\beta_j\alpha_i} & I^s_{\beta_j\beta_j} \end{pmatrix}$$

where the elements for $\mathbf{I}_{\mathbf{g}^*_m}$ in the matrix are

$$I^m_{\gamma\gamma} = \sum_{t=1}^{n} \left[\frac{B^2_{(1,t)}}{\langle m \rangle_t} \right]; \; I^m_{\alpha_i\alpha_i} = \sum_{t=1}^{n} \left[\frac{B^2_{(2,i,t)}}{\langle m \rangle_t} \right]; \; I^m_{\beta_j\beta_j} = \sum_{t=1}^{n} \left[\frac{B^2_{(3,j,t)}}{\langle m \rangle_t} \right];$$
$$I^m_{\gamma\alpha_i} = \sum_{t=1}^{n} \left[\frac{B_{(2,i,t)}B_{(1,t)}}{\langle m \rangle_t} \right]; \; I^m_{\gamma\beta_j} = \sum_{t=1}^{n} \left[\frac{B_{(3,j,t)}B_{(1,t)}}{\langle m \rangle_t} \right], \text{ and } I^m_{\alpha_i\beta_j} = \sum_{t=1}^{n} \left[\frac{B_{(2,i,t)}B_{(3,j,t)}}{\langle m \rangle_t} \right],$$

and the remaining elements are obtained by symmetry. The elements $\mathbf{I}_{\mathbf{g}^*_s}$ are

$$I^s_{\gamma\gamma} = \sum_{t=1}^{n} \left[\frac{H^2_{(1,t)}}{\langle s \rangle_t} \right]; \; I^s_{\alpha_i\alpha_i} = \sum_{t=1}^{n} \left[\frac{H^2_{(2,i,t)}}{\langle s \rangle_t} \right]; \; I^s_{\beta_j\beta_j} = \sum_{t=1}^{n} \left[\frac{H^2_{(3,j,t)}}{\langle s \rangle_t} \right];$$
$$I^s_{\gamma\alpha_i} = \sum_{t=1}^{n} \left[\frac{H_{(2,i,t)}H_{(1,t)}}{\langle s \rangle_t} \right]; \; I^s_{\gamma\beta_j} = \sum_{t=1}^{n} \left[\frac{H_{(3,j,t)}H_{(1,t)}}{\langle s \rangle_t} \right], \text{ and } I^s_{\alpha_i\beta_j} = \sum_{t=1}^{n} \left[\frac{H_{(2,i,t)}H_{(3,j,t)}}{\langle s \rangle_t} \right].$$

From the information obtained using QEF and information via its components, $m_t(\boldsymbol{\theta})$ and $s_t(\boldsymbol{\theta})$, it is clearly seen that $I^Q_{\gamma\gamma} > I^m_{\gamma\gamma}$, $I^Q_{\gamma\gamma} > I^s_{\gamma\gamma}$, $I^Q_{\alpha_i\alpha_i} > I^m_{\alpha_i\alpha_i}$, $I^Q_{\alpha_i\alpha_i} > I^s_{\alpha_i\alpha_i}$, $I^Q_{\beta_j\beta_j} > I^m_{\beta_j\beta_j}$, $I^Q_{\beta_j\beta_j} > I^s_{\beta_j\beta_j}$, $I^Q_{\gamma\alpha_i} > I^m_{\gamma\gamma}$, $I^Q_{\gamma\alpha_i} > I^s_{\gamma\gamma}$, $I^Q_{\alpha_i\beta_j} > I^m_{\alpha_i\alpha_i}$, $I^Q_{\alpha_i\beta_j} > I^s_{\alpha_i\alpha_i}$, $I^Q_{\alpha_i\beta_j} > I^m_{\beta_j\beta_j}$, and $I^Q_{\beta_j\beta_j} > I^s_{\beta_j\beta_j}$. Based on these, we conclude that the QEF is more informative than the component estimating functions (see Ghahramani & Thavaneswaran (2009)).

6.4 Performance of The Estimation Methods in NBINGARCH $(1,1)$

In this section, we compare the performance of QEF method with EF and MLE methods in NBINGARCH $(1,1)$ model by a simulation study with replication $N = 500$, and two sample size, $n = 100$ and $n = 2000$.

NBINGARCH $(1,1)$ process is defined as:

$$\left(X_t | \Im^X_{t-1}\right) \sim NB\left(r, p_t\right),$$

$$\frac{1-p_t}{p_t} = \lambda_t(\boldsymbol{\theta}) = \quad \gamma + \alpha_1 X_{t-1} + \beta_1 \lambda_{t-1}. \tag{6.28}$$

6.4.1 MLE Derivation of NBINGARCH $(1,1)$

In order to apply the MLE for the model, we first have to find loglikelihood of NBINGARCH $(1,1)$ which is

$$\mathscr{L} = \sum_{t=1}^{n} \left\{ ln \left\{ \frac{(X_t+r-1)!}{(r-1)!(X_t-r)!} \right\} \right\} + r\, ln \left(\frac{1}{1+\lambda_t} \right) + (X_t - r)\ ln \left(\frac{\lambda_t}{\lambda_t+1} \right).$$

To obtain the parameters of interest, we have to maximize the likelihood function using *nlminb* command in R-cran.

6.4.2 EF Derivation of NBINGARCH $(1,1)$

The derivation of EF for NBINGARCH $(1,1)$ can be obtained from Equation (6.26) whereby

$$\begin{aligned} g_E^*(\boldsymbol{\theta}) &= -\sum_{t=1}^{n} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \left\{ \frac{X_t - r\lambda_t(\boldsymbol{\theta})}{r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))} \right\}, \\ &= -\sum_{t=1}^{n} r\lambda_t'(\boldsymbol{\theta}) \left\{ \frac{X_t - r\lambda_t(\boldsymbol{\theta})}{r\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))} \right\}, \\ &= -\sum_{t=1}^{n} \lambda_t'(\boldsymbol{\theta}) \left\{ \frac{X_t - r\lambda_t(\boldsymbol{\theta})}{\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))} \right\}. \end{aligned}$$

Therefore, the optimal EF function for the parameters are:

$$g_E^*(\gamma) = -\sum_{t=1}^{n} \frac{\{X_t - r\lambda_t(\boldsymbol{\theta})\}\, r \left\{ 1 + \beta_1 \frac{\partial \lambda_{t-1}(\boldsymbol{\theta})}{\partial \gamma} \right\}}{\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))}, \tag{6.29}$$

$$g_E^*(\alpha_1) = -\sum_{t=1}^{n} \frac{\{X_t - r\lambda_t(\boldsymbol{\theta})\}\, r \left\{ rX_{t-1} + \beta_1 \frac{\partial \lambda_{t-1}(\boldsymbol{\theta})}{\partial \alpha_1} \right\}}{\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))}, \tag{6.30}$$

$$g_E^*(\beta_1) = -\sum_{t=1}^{n} \frac{\{X_t - r\lambda_t(\boldsymbol{\theta})\}\, r \left\{ \lambda_{t-1} + \beta_1 \frac{\partial \lambda_{t-1}(\boldsymbol{\theta})}{\partial \beta_1} \right\}}{\lambda_t(\boldsymbol{\theta})(1+\lambda_t(\boldsymbol{\theta}))}. \tag{6.31}$$

To obtain the estimates, we solve the simultaneous Equations from Equation (6.29) until Equation (6.31), by letting them equal to zero using *nleqslv* command in R-cran software.

6.4.3 QEF Derivation of NBINGARCH $(1,1)$

From Equation (6.24)-(6.25) in Section (6.3), we can simply find the QEF optimal functions for NBINGARCH $(1,1)$ process where

$$\begin{aligned} g_Q^*(\gamma) &= \sum_{t=1}^{n} \left[R_t \left[\left(B_{(1,t)}^m + H_{(1,t)}^m \right) m_t + (T_1^v + Z_1^v) s_t \right] \right], && (6.32) \\ g_Q^*(\alpha_1) &= \sum_{t=1}^{n} \left[R_t \left[\left(B_{(2,1,t)}^m + H_{(2,1,t)}^m \right) m_t + \left(T_{(2,1,t)}^v + Z_{(2,1,t)}^v \right) s_t \right] \right], && (6.33) \\ g_Q^*(\beta_1) &= \sum_{t=1}^{n} \left[R_t \left[\left(B_{(3,1,t)}^m + H_{(3,1,t)}^m \right) m_t + \left(T_{(3,1,t)}^v + Z_{(3,1,t)}^v \right) s_t \right] \right]. && (6.34) \end{aligned}$$

Using R-cran software through *nlesqv* command, we can solve the simultaneous Equations (6.33)-(6.34).

6.4.4 Simulation Study

We use R-cran to obtain the parameter estimates using MLE, EF and QEF methods for NBINGARCH $(1,1)$ model. The steps in the algorithm are the same as those in Section 5.4.3. In Step 3, we use *nlminb* command in R-cran to minimize the negative likelihood of NBINGARCH $(1,1)$ model and for both EF and QEF methods, the optimal Equation g^* can be solved using *nleqslv* in R-cran.

6.4.5 The Result

In this section, we investigate the performance of QEF method compared to MLE and EF methods in the NBINGARCH $(1,1)$ model based on mean, bias, standard error(SE) and mean squared error (MSE). The results are presented in Table 6.3 to Table 6.6. We perform the simulation with sample size $n = 100$, $n = 1500$ and $n = 2000$ and for different sets of values of the parameters for the cases $r = 2$ and $r = 3$.

Discussion

We assess the performance of MLE, EF and QEF methods through a simulation study for NBINGARCH(p,q) model using R-cran software. From all tables, which are, Table 6.2 - Table 6.11, the results show that, the QEF method gives the value of estimated parameters closest to the true values for all cases for $n = 100$, $n = 1500$ and $n = 2000$.

In comparing EF and QEF estimators, from all tables, we can obviously see that, the

QEF method outperforms such that smaller values of biasness, SE and MSE are observed for QEF indicating the QEF method is superior estimation than EF method. Besides, in comparing with MLE estimator, for stationarity cases, see Table 6.2-6.6, we can clearly see that the value of the standard errors and mean square errors of MLE estimator are fluctuate with QEF method for almost combination set of parameters indicating that QEF method is comparable with MLE method. However, when the values of parameters approach nonstationarity condition, see Table 6.7-6.11, the MLE gives a slightly bigger standard errors and mean square errors. The results indicate that the QEF methods are better estimators compared to MLE method.

Besides, when we increase the number of sample size from $n = 100$ to large sample size, $n = 1500$ and $n = 2000$, the results show that the value of standard errors, biasness and mean square error for EF and QEF methods are decrease for all combination sets of parameters but fluctuate in MLE method. Therefore, it indicate that the EF and QEF estimators are more consistent compared to MLE estimator.

Hence, in a nutshell, we can conclude that QEF method is superior estimation method compared to EF and MLE methods for this integer-valued time series model.

6.5 Real Example

Following Zhu (2011), we apply the NBINGARCH(p,q) model on Polio data found in Zeger (1988). The data represent counts of poliomyelitis cases in the United State from 1970 to 1983 as reported by the Centres of Disease Control and has shown a long-term decrease in the rate of U.S. polio infection. The data also are available in Morbidity and Mortality Weekly Report Annual Summary. The disease mainly affects children under 5 years of age and may lead to irreversible paralysis, paralysed and immobilized breathing muscles. Such disease is caused by a virus that invades the nervous system and it can be spread via faecal-oral route and multiplies in the intestine. Zhu (2011) applied the NBINGARCH $(1,1)$ model for this data and estimated the parameters using MLE. Here, instead of using MLE, we use QEF estimates to fit such model.

Table 6.2: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.2,$ and $\beta_1 = 0.3$

	r=2									r=3								
	n=100			n=1500			n=2000			n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$																		
Mean	0.048	0.050	0.052	0.069	0.075	0.102	0.087	0.104	1.01E-01	0.045	0.048	0.510	0.067	0.069	0.078	0.079	0.106	0.998
Bias	0.052	0.050	0.048	0.031	0.025	0.002	0.013	0.02	7.31E-04	0.055	0.052	0.049	0.033	0.031	0.022	0.021	0.006	0.002
SE	0.059	0.056	0.049	0.022	0.036	0.019	0.014	0.009	0.003	0.062	0.060	0.055	0.047	0.042	0.021	0.023	0.012	0.005
MSE	8.21E-03	9.11E-03	8.27E-03	1.87E-03	2.27E-03	1.21E-03	1.20E-03	1.21E-03	1.09E-03	8.99E-03	9.28E-03	9.05E-03	2.95E-03	3.24E-03	2.59E-03	2.21E-03	1.65E-03	1.18e-03
$\hat{\alpha}_1$																		
Mean	0.108	0.114	0.123	0.139	0.153	0.171	0.166	0.211	0.195	0.109	0.11	0.118	0.133	0.145	0.168	0.159	0.188	0.204
Bias	0.092	0.086	0.077	0.061	0.047	0.029	0.034	0.011	0.005	0.091	0.09	0.082	0.067	0.055	0.032	0.041	0.012	0.004
SE	0.081	0.075	0.069	0.037	0.043	0.037	0.015	0.015	0.006	0.09	0.078	0.071	0.045	0.051	0.035	0.047	0.023	0.017
MSE	1.21E-02	1.18E-02	9.88E-03	7.37E-03	7.57E-03	4.65E-03	2.31E-03	2.33E-03	1.87E-03	2.01E-02	2.13E-03	9.97E-03	8.13E-03	8.29E-03	7.32E-03	3.88E-03	2.98E-03	1.93e-03
$\hat{\beta}_1$																		
Mean	0.345	0.341	0.332	0.322	0.318	0.290	0.397	0.308	0.305	0.351	0.346	0.340	0.324	0.321	0.311	0.333	0.312	0.292
Bias	0.045	0.041	0.032	0.022	0.018	0.010	0.097	0.008	0.005	0.051	0.046	0.040	0.024	0.021	0.011	0.033	0.012	0.008
SE	0.082	0.076	0.058	0.033	0.031	0.011	0.015	0.015	0.002	0.088	0.071	0.062	0.054	0.055	0.043	0.021	0.018	0.009
MSE	1.39E-02	1.11E-02	9.03E-03	9.27E-03	6.45E-03	6.83E-03	1.86E-03	1.85E-03	1.21E-03	2.03E-02	1.98E-02	9.67E-03	8.21E-03	7.13E-03	6.98E-03	2.13E-03	2.66E-03	1.75e-03

Table 6.3: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.4,$ and $\beta_1 = 0.1$

	r=2									r=3								
	n=100			n=1500			n=2000			n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$																		
Mean	0.129	0.127	0.131	0.158	0.174	0.188	0.168	0.164	0.178	0.167	0.208	0.196	0.124	0.123	0.129	0.155	0.213	0.188
Bias	0.071	0.073	0.069	0.042	0.026	0.012	0.032	0.036	0.022	0.033	0.008	0.004	0.076	0.077	0.071	0.045	0.013	0.012
SE	0.081	0.075	0.070	0.035	0.031	0.030	0.034	0.033	0.029	0.01	0.010	0.009	0.083	0.083	0.079	0.021	0.025	0.023
MSE	7.11E-03	7.21E-03	7.18E-03	1.47E-03	1.30E-03	1.21E-03	1.55E-03	1.48E-03	1.01E-03	1.04E-03	1.05E-03	1.01E-03	7.38E-03	7.41E-03	7.25E-03	1.32E-03	1.21E-03	9.53e-04
$\hat{\alpha}_1$																		
Mean	0.286	0.292	0.307	0.315	0.394	0.398	0.311	0.383	0.389	0.332	0.401	0.399	0.281	0.288	0.298	0.351	0.403	0.401
Bias	0.114	0.108	0.093	0.085	0.006	0.001	0.089	0.017	0.011	0.007	0.001	8.01E-04	0.119	0.112	0.102	0.049	0.030	0.001
SE	0.081	0.086	0.075	0.045	0.052	0.050	0.051	0.052	0.049	0.01	0.010	0.005	0.089	0.091	0.081	0.021	0.017	0.009
MSE	4.31E-03	4.97E-03	4.22E-03	1.43E-03	1.71E-03	1.35E-03	1.61E-03	1.89E-03	1.72E-03	1.03E-03	1.10E-03	1.02E-03	4.87E-03	4.95E-03	4.92E-03	1.20E-03	1.33E-03	1.22e-03
$\hat{\beta}_1$																		
Mean	0.152	0.157	0.143	0.119	0.117	0.109	0.121	0.082	0.108	0.115	0.097	0.102	0.158	0.158	0.147	0.124	0.104	0.103
Bias	0.052	0.057	0.043	0.019	0.017	0.009	0.021	0.018	0.008	0.015	0.003	0.002	0.058	0.058	0.047	0.024	0.004	0.003
SE	0.061	0.062	0.058	0.022	0.035	0.01	0.032	0.033	0.032	0.019	0.019	0.009	0.072	0.078	0.067	0.022	0.024	0.019
MSE	8.13E-03	8.17E-03	7.92E-03	4.16E-03	5.61E-03	4.34E-03	5.41E-03	5.62E-03	5.01E-03	1.06E-03	1.08E-03	1.03E-03	8.27E-03	8.32E-03	8.11E-03	2.08E-03	2.81E-03	2.32e-03

Table 6.4: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1$, and $\beta_1 = 0.4$

	r=2									r=3								
	n=100			n=1500			n=2000			n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$																		
Mean	0.209	0.208	0.216	0.222	0.251	0.332	0.255	0.29	0.294	0.205	0.207	0.211	0.213	0.249	0.261	0.243	0.285	0.290
Bias	0.091	0.092	0.084	0.078	0.049	0.032	0.045	0.010	0.006	0.095	0.093	0.089	0.087	0.051	0.039	0.057	0.015	0.010
SE	0.085	0.080	0.072	0.021	0.043	0.018	0.019	0.018	0.01	0.092	0.087	0.079	0.034	0.034	0.021	0.042	0.025	0.016
MSE	5.68E-03	6.12E-03	5.24E-03	1.21E-03	2.73E-03	1.25E-03	1.06E-03	1.06E-03	1.01E-03	6.32E-03	6.81E-03	6.11E-03	1.34E-03	2.52E-03	2.11E-03	2.09E-03	1.19E-03	1.01e-03
$\hat{\alpha}_1$																		
Mean	0.155	0.149	0.138	0.069	0.078	0.085	0.094	0.103	0.099	0.161	0.154	0.141	0.055	0.072	0.079	0.055	0.085	0.092
Bias	0.055	0.049	0.038	0.031	0.022	0.015	0.006	0.003	0.001	0.061	0.054	0.041	0.045	0.028	0.021	0.045	0.015	0.008
SE	0.063	0.058	0.051	0.006	0.005	0.004	0.001	0.001	7.23E-04	0.071	0.062	0.051	0.023	0.033	0.023	0.011	0.015	0.002
MSE	7.13E-03	9.68E-03	6.88E-03	1.87E-03	9.50E-03	4.61E-03	1.02E-03	1.16E-03	1.03E-06	7.58E-03	9.89E-03	7.55E-03	2.21E-03	7.86E-03	5.87E-03	1.49E-03	3.77E-03	1.03e-03
$\hat{\beta}_1$																		
Mean	0.453	0.452	0.443	0.416	0.411	0.396	0.490	0.392	0.401	0.471	0.471	0.458	0.421	0.421	0.411	0.434	0.409	0.399
Bias	0.053	0.052	0.043	0.016	0.011	0.004	0.090	0.008	0.001	0.071	0.071	0.058	0.021	0.021	0.011	0.034	0.009	0.001
SE	0.073	0.071	0.059	0.020	0.022	0.015	0.017	0.015	0.011	0.081	0.076	0.061	0.032	0.024	0.023	0.02	0.017	0.015
MSE	7.44E-03	9.38E-03	7.51E-03	5.20E-03	5.56E-03	2.21E-03	1.62E-03	1.80E-03	1.33E-03	8.16E-03	1.03E-02	7.82E-03	6.78E-03	6.99E-03	5.34E-03	2.45E-03	2.83E-03	1.41e-03

Table 6.5: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.4,$ and $\beta_1 = 0.2$

	r=2									r=3								
	n=100			n=1500			n=2000			n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$																		
Mean	0.301	0.387	0.372	0.213	0.366	0.259	0.242	0.344	0.284	0.408	0.209	0.215	0.211	0.356	0.242	0.279	0.345	0.292
Bias	0.101	0.087	0.072	0.087	0.066	0.041	0.058	0.044	0.016	0.108	0.091	0.085	0.089	0.056	0.058	0.021	0.045	0.008
SE	0.087	0.075	0.062	0.036	0.088	0.035	0.018	0.021	0.017	0.092	0.082	0.071	0.044	0.052	0.031	0.019	0.026	0.018
MSE	6.32E-03	3.21E-03	2.11E-03	1.51E-03	1.82E-03	1.36E-03	1.09E-03	1.09E-03	1.01E-03	2.11E-03	1.98E-03	1.53E-03	2.11E-03	1.98E-03	1.53E-03	3.21E-03	1.10E-03	1.05e-04
$\hat{\alpha}_1$																		
Mean	0.339	0.338	0.344	0.389	0.378	0.390	0.324	0.389	0.402	0.334	0.338	0.345	0.381	0.381	0.389	0.332	0.391	0.392
Bias	0.061	0.062	0.056	0.011	0.022	0.010	0.076	0.011	0.002	0.066	0.062	0.055	0.019	0.019	0.011	0.068	0.009	0.008
SE	0.079	0.067	0.055	0.037	0.059	0.036	0.018	0.018	0.01	0.082	0.071	0.055	0.056	0.060	0.045	0.021	0.022	0.016
MSE	9.21E-03	5.84E-03	5.18E-03	2.56E-03	2.73E-03	2.04E-03	1.05E-03	1.18E-03	1.01E-03	9.66E-03	5.91E-03	5.01E-03	5.14E-03	4.76E-03	4.23E-03	3.09E-03	2.98E-03	1.09e-03
$\hat{\beta}_1$																		
Mean	0.301	0.297	0.285	0.223	0.119	0.221	0.216	0.189	0.208	0.301	0.299	0.292	0.231	0.220	0.217	0.223	0.187	0.211
Bias	0.101	0.097	0.085	0.023	0.081	0.021	0.016	0.011	0.008	0.101	0.099	0.092	0.031	0.020	0.017	0.023	0.013	0.011
SE	0.067	0.068	0.066	0.067	0.068	0.066	0.030	0.038	0.024	0.081	0.089	0.052	0.081	0.089	0.052	0.068	0.061	0.049
MSE	4.06E-03	4.31E-03	4.07E-03	4.06E-03	4.31E-03	4.07E-03	1.86E-03	1.86E-03	1.48E-03	3.99E-03	4.42E-03	3.87E-03	3.99E-03	4.42E-03	3.87E-03	4.02E-03	1.98E-03	1.55e-03

Table 6.6: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.5, \alpha_1 = 0.2,$ and $\beta_1 = 0.3$

	r=2									r=3								
	n=100			n=1500			n=2000			n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$																		
Mean	0.441	0.443	0.567	0.484	0.486	0.509	0.416	0.508	0.499	0.433	0.439	0.458	0.456	0.476	0.511	0.453	0.487	0.498
Bias	0.059	0.057	0.033	0.016	0.014	0.009	0.084	0.008	0.001	0.067	0.061	0.042	0.043	0.024	0.011	0.046	0.013	0.002
SE	0.056	0.062	0.043	0.021	0.019	0.009	0.016	0.013	0.003	0.067	0.069	0.052	0.052	0.034	0.027	0.025	0.028	0.013
MSE	7.13E-03	7.89E-03	7.05E-03	2.96E-03	3.65E-03	1.71E-03	1.40E-03	1.41E-03	1.06E-03	7.69E-03	8.15E-03	7.58E-03	3.17E-03	4.27E-03	2.88E-03	2.23E-03	1.98E-03	1.11e-03
$\hat{\alpha}_1$																		
Mean	0.103	0.109	0.134	0.137	0.147	0.181	0.164	0.222	0.195	0.102	0.110	0.129	0.121	0.135	0.166	0.158	0.231	0.187
Bias	0.097	0.091	0.066	0.063	0.053	0.019	0.036	0.022	0.005	0.098	0.090	0.071	0.079	0.065	0.034	0.042	0.031	0.013
SE	0.067	0.071	0.055	0.029	0.036	0.027	0.012	0.010	0.007	0.079	0.082	0.059	0.039	0.034	0.031	0.027	0.021	0.019
MSE	8.13E-03	8.03E-03	7.12E-03	4.85E-03	5.55E-03	5.05E-03	2.52E-03	2.51E-03	2.00E-03	8.92E-03	9.12E-03	8.05E-03	5.11E-03	5.79E-03	5.11E-03	3.11E-03	4.76E-03	3.24e-03
$\hat{\beta}_1$																		
Mean	0.358	0.355	0.339	0.322	0.319	0.293	0.317	0.29	0.302	0.361	0.352	0.346	0.324	0.331	0.325	0.319	0.321	0.319
Bias	0.058	0.055	0.039	0.022	0.019	0.007	0.017	0.010	0.002	0.061	0.052	0.046	0.024	0.031	0.025	0.019	0.021	0.019
SE	0.073	0.069	0.044	0.042	0.051	0.030	0.036	0.037	0.019	0.085	0.079	0.056	0.045	0.065	0.048	0.039	0.039	0.028
MSE	8.13E-03	9.92E-03	8.15E-05	5.73E-03	5.24E-03	4.09E-05	1.32E-03	1.29E-03	1.03E-03	9.66E-03	1.87E-02	9.58E-03	6.15E-03	6.33E-03	6.27E-03	2.16E-03	2.44E-03	2.09e-03

Table 6.7: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.6$, and $\beta_1 = 0.3$

	r=2									r=3								
	n=100			n=1500			n=2000			n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$																		
Mean	0.173	0.157	0.145	0.071	0.071	0.074	0.081	0.103	0.098	0.181	0.162	0.151	0.077	0.071	0.079	0.103	0.087	0.091
Bias	0.073	0.057	0.045	0.029	0.029	0.026	0.019	0.003	0.002	0.081	0.062	0.051	0.023	0.029	0.021	0.003	0.013	0.009
SE	0.511	0.099	0.073	0.410	0.009	0.003	0.405	0.005	0.001	0.528	0.103	0.089	0.433	0.016	0.009	0.441	0.005	9.21e-04
MSE	8.11E-02	8.63E-03	7.18E-03	1.74E-02	1.72E-03	1.34E-03	1.05E-02	1.10E-03	1.02E-03	8.59E-02	9.12E-03	5.21E-03	2.02E-02	1.88E-03	1.62E-03	1.41E-02	1.92E-03	1.03e-03
$\hat{\alpha}_1$																		
Mean	0.547	0.553	0.559	0.583	0.583	0.587	0.585	0.595	0.599	0.541	0.549	0.561	0.555	0.567	0.581	0.542	0.588	0.602
Bias	0.053	0.047	0.041	0.017	0.017	0.013	0.015	0.005	0.001	0.059	0.051	0.039	0.045	0.033	0.019	0.058	0.012	0.002
SE	0.712	0.089	0.056	0.531	0.034	0.026	0.518	0.019	0.011	0.738	0.092	0.088	0.532	0.052	0.041	0.556	0.022	0.016
MSE	9.21E-02	6.21E-03	5.28E-03	4.05E-02	5.98E-03	4.51E-03	1.44E-02	1.43E-03	1.21E-03	9.88E-02	6.88E-03	6.21E-03	5.78E-02	5.31E-03	5.20E-03	3.89E-02	1.53E-03	1.47e-03
$\hat{\beta}_1$																		
Mean	0.352	0.331	0.328	0.320	0.320	0.317	0.314	0.291	0.302	0.355	0.333	0.330	0.321	0.323	0.312	0.282	0.292	0.303
Bias	0.052	0.031	0.028	0.020	0.020	0.017	0.014	0.009	0.002	0.055	0.033	0.030	0.021	0.023	0.012	0.018	0.008	0.003
SE	0.635	0.079	0.065	0.435	0.034	0.033	0.425	0.024	0.019	0.641	0.083	0.059	0.444	0.043	0.039	0.446	0.027	0.018
MSE	9.12E-02	1.28E-02	1.01E-02	7.15E-02	8.09E-03	6.37E-03	2.93E-02	2.94E-03	1.97E-03	8.11E-02	8.11E-03	7.88E-03	8.11E-02	8.11E-03	7.88E-03	8.13E-02	5.99E-03	2.88e-03

Table 6.8: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.7,$ and $\beta_1 = 0.2$

	r=2									r=3								
	n=100			n=1500			n=2000			n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$																		
Mean	0.032	0.059	0.068	0.073	0.076	0.080	0.083	0.110	0.098	0.028	0.055	0.069	0.068	0.069	0.077	0.079	0.114	0.103
Bias	0.068	0.041	0.032	0.027	0.024	0.020	0.017	0.010	0.002	0.072	0.045	0.031	0.032	0.031	0.023	0.021	0.014	0.003
SE	0.713	0.083	0.071	0.407	0.003	0.002	0.405	0.002	8.11E-04	0.728	0.099	0.079	0.411	0.012	0.009	0.408	0.008	0.001
MSE	5.21E-02	7.11E-03	6.54E-03	1.16E-02	1.53E-03	1.41E-03	6.13E-02	6.14E-04	6.01E-04	5.93E-02	8.01E-03	7.52E-03	2.38E-02	2.88E-03	2.02E-03	1.32E-02	1.09E-03	9.11e-04
$\hat{\alpha}_1$																		
Mean	0.589	0.603	0.629	0.616	0.623	0.651	0.638	0.657	0.685	0.583	0.601	0.621	0.611	0.627	0.632	0.652	0.659	0.687
Bias	0.111	0.097	0.071	0.084	0.077	0.049	0.062	0.043	0.015	0.117	0.099	0.079	0.089	0.073	0.068	0.048	0.041	0.013
SE	0.556	0.063	0.052	0.556	0.063	0.052	0.521	0.022	0.011	0.577	0.07	0.07	0.577	0.07	0.07	0.581	0.063	0.051
MSE	9.22E-02	7.41E-03	6.88E-03	6.07E-02	6.25E-03	6.13E-03	2.81E-02	2.80E-03	2.34E-03	9.65E-02	5.89E-02	6.31E-03	6.21E-03	7.55E-03	6.93E-03	3.88E-02	3.77E-03	2.98e-03
$\hat{\beta}_1$																		
Mean	0.321	0.301	0.299	0.309	0.298	0.291	0.243	0.174	0.219	0.321	0.305	0.294	0.278	0.299	0.284	0.23	0.235	0.209
Bias	0.121	0.101	0.099	0.109	0.098	0.091	0.043	0.026	0.019	0.121	0.105	0.094	0.078	0.099	0.084	0.03	0.035	0.009
SE	0.753	0.095	0.087	0.542	0.072	0.042	0.526	0.061	0.033	0.755	0.097	0.091	0.556	0.077	0.067	0.572	0.073	0.053
MSE	8.51E-02	9.26E-03	8.11E-03	4.78E-02	7.86E-03	7.33E-03	3.06E-02	3.18E-03	3.04E-03	8.67E-02	9.33E-03	8.19E-03	5.02E-02	7.98E-03	6.89E-03	4.11E-02	3.89E-03	3.17e-03

Table 6.9: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.8,$ and $\beta_1 = 0.1$

	r=2									r=3								
	n=100			n=1500			n=2000			n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$																		
Mean	0.119	0.118	0.133	0.156	0.158	0.233	0.171	0.217	0.188	0.117	0.118	0.129	0.146	0.159	0.233	0.236	0.179	0.188
Bias	0.081	0.082	0.067	0.044	0.042	0.033	0.029	0.017	0.012	0.083	0.082	0.071	0.054	0.041	0.033	0.036	0.021	0.012
SE	0.687	0.087	0.062	0.418	0.043	0.018	0.412	0.013	0.01	0.691	0.092	0.065	0.428	0.034	0.019	0.444	0.023	0.009
MSE	8.33E-02	9.12E-03	7.52E-03	2.89E-02	3.18E-03	2.53E-03	1.69E-02	1.61E-03	1.41E-03	8.39E-02	9.31E-03	8.11E-03	3.11E-02	3.82E-03	3.21E-03	3.25E-02	2.67E-03	1.54e-03
$\hat{\alpha}_1$																		
Mean	0.712	0.728	0.732	0.725	0.742	0.783	0.746	0.784	0.791	0.703	0.725	0.739	0.711	0.741	0.782	0.765	0.788	0.808
Bias	0.088	0.072	0.068	0.075	0.058	0.017	0.054	0.016	0.009	0.097	0.075	0.061	0.089	0.059	0.018	0.035	0.012	0.008
SE	0.612	0.105	0.087	0.458	0.096	0.055	0.425	0.020	0.010	0.621	0.108	0.091	0.472	0.094	0.045	0.433	0.029	0.016
MSE	9.13E-02	9.21E-03	7.33E-03	6.11E-02	8.87E-03	5.01E-03	2.73E-02	2.71E-03	2.16E-03	9.08E-02	1.27E-02	8.23E-03	7.12E-02	9.87E-03	5.07E-03	3.51E-02	3.21E-03	2.08e-03
$\hat{\beta}_1$																		
Mean	0.232	0.191	0.175	0.196	0.187	0.137	0.135	0.127	0.105	0.235	0.197	0.139	0.199	0.191	0.144	0.132	0.127	0.111
Bias	0.132	0.091	0.075	0.096	0.087	0.037	0.035	0.027	0.005	0.135	0.097	0.039	0.099	0.091	0.044	0.032	0.027	0.011
SE	0.618	0.105	0.086	0.543	0.081	0.043	0.523	0.054	0.038	0.621	0.109	0.089	0.565	0.089	0.052	0.571	0.044	0.041
MSE	8.62E-02	5.33E-03	3.17E-03	1.68E-02	1.72E-03	1.50E-03	1.04E-02	1.07E-03	9.11E-04	8.91E-02	5.39E-03	3.28E-03	2.77E-02	2.45E-03	1.98E-03	2.01E-02	1.55E-03	1.07e-03

Table 6.10: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1$, and $\beta_1 = 0.8$

	r=2									r=3								
	n=100			n=1500			n=2000			n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$																		
Mean	0.254	0.272	0.283	0.285	0.287	0.291	0.315	0.308	0.302	0.255	0.269	0.279	0.275	0.281	0.288	0.268	0.311	0.305
Bias	0.046	0.028	0.017	0.015	0.013	0.009	0.015	0.008	0.002	0.045	0.031	0.021	0.025	0.019	0.012	0.032	0.011	0.005
SE	0.623	0.107	0.097	0.481	0.085	0.084	0.435	0.070	0.058	0.631	0.112	0.101	0.487	0.089	0.078	0.454	0.072	0.061
MSE	8.21E-02	9.16E-03	5.44E-03	4.22E-02	4.60E-03	1.71E-03	2.33E-02	2.39E-03	1.20E-03	8.65E-02	9.87E-03	5.46E-03	5.21E-02	5.02E-03	3.89E-03	2.67E-02	3.76E-03	1.66e-03
$\hat{\alpha}_1$																		
Mean	0.159	0.155	0.143	0.078	0.078	0.089	0.083	0.086	0.09	0.161	0.158	0.150	0.075	0.078	0.079	0.081	0.082	0.091
Bias	0.059	0.055	0.043	0.022	0.022	0.011	0.017	0.014	0.010	0.061	0.058	0.050	0.025	0.022	0.021	0.019	0.018	0.009
SE	0.753	0.085	0.057	0.516	0.015	0.014	0.511	0.008	0.004	0.761	0.091	0.062	0.523	0.023	0.021	0.514	0.011	0.005
MSE	1.01E-01	2.11E-02	9.26E-03	9.81E-02	9.43E-03	5.77E-03	4.56E-02	4.29E-03	2.52E-03	1.28E-01	2.38E-02	9.33E-03	9.92E-02	9.54E-03	6.34E-03	5.54E-02	5.31E-03	3.09e-03
$\hat{\beta}_1$																		
Mean	0.861	0.858	0.844	0.828	0.827	0.817	0.714	0.789	0.809	0.865	0.862	0.849	0.831	0.829	0.827	0.715	0.811	0.807
Bias	0.061	0.058	0.044	0.028	0.027	0.017	0.086	0.011	0.009	0.065	0.062	0.049	0.031	0.029	0.027	0.085	0.011	0.007
SE	0.719	0.087	0.066	0.534	0.035	0.031	0.521	0.022	0.013	0.725	0.092	0.071	0.553	0.069	0.561	0.032	0.033	0.018
MSE	8.12E-02	7.63E-03	4.11E-03	1.53E-02	1.95E-03	1.76E-03	1.07E-02	1.07E-03	9.51E-04	8.23E-02	7.58E-03	4.29E-03	2.87E-02	2.63E-03	1.98E-03	2.11E-02	1.88E-03	1.27e-03

Table 6.11: Simulation results for NBINGARCH $(1,1)$ with $\gamma = 0.4, \alpha_1 = 0.3$, and $\beta_1 = 0.6$

	r=2									r=3								
	n=100			n=1500			n=2000			n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$																		
Mean	0.348	0.367	0.372	0.388	0.389	0.39	0.394	0.403	0.399	0.344	0.363	0.369	0.376	0.382	0.388	0.372	0.392	0.398
Bias	0.052	0.033	0.028	0.012	0.011	0.010	0.006	0.003	0.001	0.056	0.037	0.031	0.024	0.018	0.012	0.028	0.008	0.002
SE	0.612	0.099	0.073	0.449	0.049	0.041	0.423	0.021	0.015	0.621	0.102	0.081	0.456	0.047	0.044	0.422	0.032	0.018
MSE	8.24E-02	6.13E-03	4.12E-03	2.31E-02	2.72E-03	2.06E-03	1.03E-02	1.08E-03	1.01E-03	9.32E-02	7.13E-03	4.58E-03	4.13E-02	4.52E-03	4.22E-03	2.89E-02	3.11E-03	2.09e-03
$\hat{\alpha}_1$																		
Mean	0.202	0.207	0.213	0.208	0.209	0.233	0.218	0.285	0.307	0.201	0.203	0.212	0.201	0.211	0.221	0.219	0.276	0.311
Bias	0.098	0.093	0.087	0.092	0.091	0.067	0.082	0.015	0.007	0.099	0.097	0.088	0.099	0.089	0.079	0.081	0.024	0.011
SE	0.715	0.089	0.055	0.62	0.024	0.021	0.611	0.019	0.012	0.781	0.092	0.061	0.732	0.031	0.023	0.742	0.016	0.011
MSE	9.38E-02	8.62E-03	7.14E-03	1.69E-02	2.66E-03	2.04E-03	1.16E-02	1.19E-03	1.08E-03	9.45E-02	8.73E-03	7.52E-03	2.92E-02	2.99E-03	2.11E-03	1.22E-02	1.28E-03	1.09e-03
$\hat{\beta}_1$																		
Mean	0.638	0.631	0.626	0.612	0.612	0.611	0.668	0.595	0.601	0.642	0.638	0.629	0.623	0.621	0.613	0.613	0.604	0.599
Bias	0.035	0.031	0.026	0.012	0.012	0.009	0.068	0.005	0.001	0.042	0.038	0.029	0.023	0.021	0.013	0.013	0.004	0.001
SE	0.613	0.065	0.051	0.424	0.025	0.020	0.418	0.02	0.013	0.616	0.068	0.059	0.445	0.038	0.033	0.418	0.023	0.014
MSE	4.15E-02	5.88E-03	2.87E-03	2.97E-02	2.91E-03	1.99E-03	1.31E-02	1.31E-03	1.05E-03	4.26E-02	5.97E-03	3.16E-03	3.12E-02	3.24E-03	2.89E-03	1.87E-02	1.78E-03	1.11e-03

The plot is given in Figure 6.1.

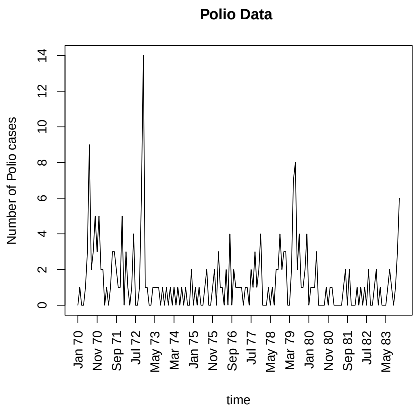


Figure 6.1: The Polio data in the United State from 1970 to 1983

We first identify the best value of r by comparing the AIC and BIC values for selected values of r as given in Table 6.12. It is clearly indicates the choice of $r = 2$.

Table 6.12: AIC and BIC values for NBINGARCH$(1,1)$

Initial value of r	1	2	3	4	5
AIC	521.325	519.321	522.369	526.128	530.965 ·
BIC	522.965	521.684	536.541	539.586	542.581

Then, we compare the performance of MLE, EF and QEF methods based on the information criteria AIC and BIC as shown in Table 6.13 with $r = 2$. The table presents the parameter estimates $\hat{\boldsymbol{\theta}}$, standard error in parenthesis, AIC and BIC values.

Table 6.13: The estimated parameter of NBINGARCH$(1,1)$ model

Method	$\hat{\gamma}$	$\hat{\alpha}_1$	$\hat{\beta}_1$	AIC	BIC
MLE	0.312 (0.008)	0.185 (0.003)	0.182 (0.018)	522.356	532.351
EF	0.325 (0.026)	0.190 (0.004)	0.192 (0.021)	520.833	532.168
QEF	0.321 (0.002)	0.183 (0.001)	0.185 (0.002)	519.321	521.684

We see that, the QEF method presents the smallest AIC and BIC compared to EF and MLE estimators indicating the fitted model via QEF method fit well to the data. The parameter estimates $(\gamma, \alpha_1, \beta_1)$ using QEF method are positive and the summation of $r\alpha_1 + \beta_1$ lies between zero and one, showing that the process satisfies the stationary requirement.

From Figure 6.1, it is clearly shown that, the data has an outlier at November 1972, which indicate that NBINGARCH$(1,1)$ model is preferable compared to INGARCH$(1,1)$ model. To prove the claim, we then compare the AIC and BIC based on QEF estimates for INGARCH$(1,1)$ and NBINGARCH$(1,1)$ models.

Table 6.14: AIC and BIC for INGARCH $(1,1)$ and NBINGARCH $(1,1)$ models.

	AIC	BIC
INGARCH $(1,1)$	550.23	563.77
NBINGARCH $(1,1)$	519.321	521.684

The results of AIC and BIC values in Table 6.14 clearly shown that the improvement of fit with using the NBINGARCH$(1,1)$ model is significant compared to INGARCH$(1,1)$ model. It indicating such model can deals with both overdispersion and potential extreme observations. Therefore, we can conclude that NBINGARCH$(1,1)$ model fits the data better than INGARCH$(1,1)$ model.

In order to check the model adequacy, we first define the residual of the fitted model. In this NBINGARCH $(1,1)$ process, we define the Pearson residual as

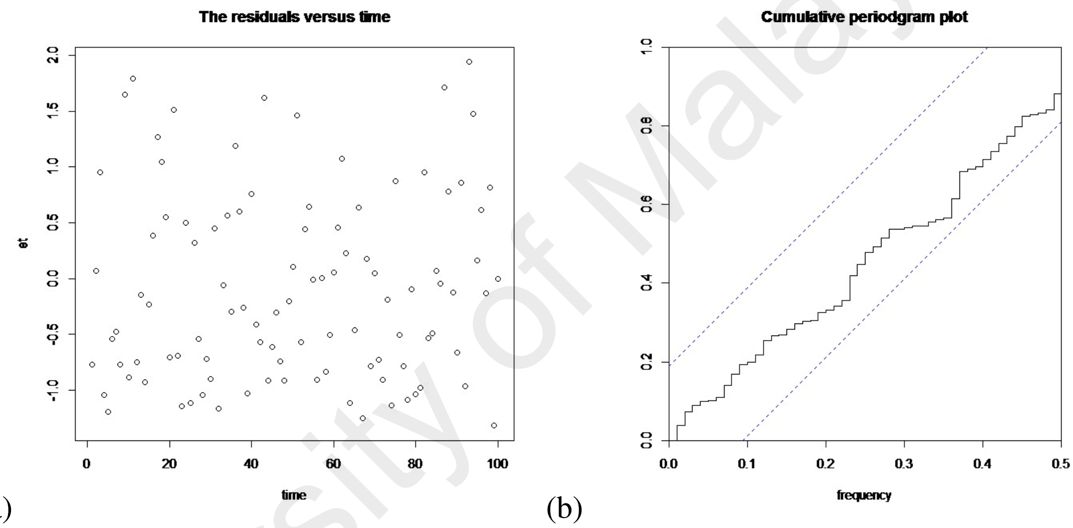
$$z_t = \left(X_t - r\lambda_t(\hat{\boldsymbol{\theta}})\right) / \sqrt{r\lambda_t(\hat{\boldsymbol{\theta}})\left(1+\lambda_t(\hat{\boldsymbol{\theta}})\right)}.$$

Using such residual z_t, we find the mean and variance as 0.0256 and 1.088 respectively whereby we can clearly see that the mean and variance of this Pearson residuals close to zero and unity which satisfy the first condition on specified model by Kedem & Fokianos (2005). To see whether the second condition is satisfied or not, we use the Ljung-Box (LB) statistics to determine the existence of serial correlation in the data. From Table 6.15, the p-value obtained is greater than the significance level $\alpha = 0.05$ for both z_t and z_t^2. Therefore, one can conclude that there is no significant serial correlation in the residual which satisfies the second requirement. Therefore, the process NBINGARCH $(1,1)$ is adequate for the data discussed.

Table 6.15: Diagnostics for NBINGARCH(1,1) model

	$LB_{30}(z_t)$	$LB_{30}(z_t^2)$
χ^2	29.1	14.1
p-value	0.513	0.994

In addition, the randomness of Pearson residual plot and the cumulative periodagram plot can be used in investigating the model fitting adequacy. Both plots shown in Figure 6.2 satisfy the specified model conditions by Brockwell & Davis (2013). Hence, we conclude that the model is adequate for the Polio data.



(a) (b)

Figure 6.2: (a)The Pearson residual plot. (b)The periodagram plot

6.6 Summary of The Chapter

In this chapter, we derived moments up to order four for NBINGARCH(p,q) model. We also derived the optimal QEF function and its information gain for this model. Theoretically, the information gain by QEF estimator is more informative compared to its individual components. Furthermore, we carried out the simulation studies to see the performance of QEF estimator compared to EF and MLE methods. The results show that the QEF method performed well compared to other two methods for this integer-valued model. Therefore, we can conclude that the performance of QEF estimator is superior compared to EF and MLE methods. Finally we applied the methodology on real data set found in Zeger (1988). We found that the NBINGARCH$(1,1)$ is fitted the data well.

CHAPTER 7

ZIPINGARCH(p,q) MODEL

7.1 Introduction

High occurrence of zeroes in data set may be observed especially in some areas of public health. For rare diseases with low infection rates, the observed counts typically contain a high frequency of zeroes (zero-inflation) but the counts can also be very large during outbreak period. Concerning with this excess zeroes issues, Neyman (1939) and Feller (1943) introduced the concept of zero-inflation to address the problem of excess of zeros. Following their findings, many studies and applications of zero-inflated models have been put forward, especially in regression context. For example, Lambert (1992) introduced zero-inflated Poisson (ZIP) regression model, Baksh et al. (2011) proposed the overdispersion test for the ZIP model and Lim et al. (2014) studied the ZIP mixture regression model.

Extensive studies related to the development of zero-inflated processes are also seen in time series analysis. A model introduced by Zhu (2012) is known as ZIPINGARCH(p,q) model. The model is inspired by GARCH (p,q) model with ZIP as conditional distribution. It is defined as

$$X_t|\Im_{t-1}^X \sim ZIP(\lambda_t, \omega), \tag{7.1}$$

$$\lambda_t = \gamma + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_j \lambda_{t-j} \tag{7.2}$$

where $\gamma > 0$, $\alpha_i \geq 0$, $i = 1,2,\ldots,p$, $\beta_j \geq 0$, $j = 1,2,\ldots,q$, $\Im_{t-1}$ is the σ-field generated from $X_{t-1}, X_{t-2}, \ldots, X_1$ and the inflation parameter, $0 < \omega < 1$. Therefore, in this chapter, we study the ZIP time series model namely ZIPINGARCH(p,q) process focusing on its moments up to order four and apply the QEF method as the estimation method.

7.2 The Moment Properties

Similar as INGARCH(p,q) and NBINGARCH(p,q) models, the moment properties for conditional and unconditional distributions of ZIPINGARCH(p,q) model are derived.

7.2.1 The Moment Properties of Conditional Distribution in ZIPINGARCH(p,q) Model

In ZIPINGARCH(p,q) model, the conditional distribution follows zero-inflated Poisson (ZIP). Let $Y_t = X_t|\Im_{t-1}^X$ follows the zero-inflated Poisson distribution with parameter λ_t and ω which is $Y_t \sim ZIP(\lambda_t, \omega)$. Zero-inflation models occur when zero counts is greater than expected for the Poisson distribution. According to Johnson et al. (2005) the distribution of Poisson known as zero-inflated Poisson where the zero class is misreported or over reported. The probability mass function (pmf) is

$$f(y_t) = \begin{cases} \omega + (1-\omega)e^{-\lambda_t} & \text{for } y_t = 0 \\ (1-\omega)\frac{e^{-\lambda_t}\lambda_t^{y_t}}{y_t!} & \text{for } y_t = 1,2,3,\ldots \end{cases} \tag{7.3}$$

It is a special case of finite mixture models, that is useful for count data containing many zeros. The probability generating function is $G(z) = \omega + (1-\omega)\,e^{\lambda_t(z-1)}$ (see Johnson et al. (2005) page 353). The mean of such distribution is the first derivatives of the pgf with $z = 1$

$$\begin{aligned} G'(z) &= \frac{d}{dz}\left[\omega + (1-\omega)\,e^{\lambda_t(z-1)}\right] \\ \mu_t &= E(y_t) = G'(1) = \lambda_t\,(1-\omega)\,e^{\lambda_t(z-1)} = (1-\omega)\,\lambda_t. \end{aligned}$$

Using the same step in INGARCH(p,q) and NBINGARCH(p,q), the kth derivative of ZIP's pgf is $G^k(z) = (1-\omega)\,e^{\lambda_t(z-1)}\lambda_t^k$. To find the variance of ZIP, we first have to obtain the value of $E\left(y_t^2\right)$ whereby from the second derivative of pgf, that is

$$G''(1) = E\left(y_t(y_t-1)\right) = E\left(y_t^2\right) - E(y_t) = (1-\omega)\,\lambda_t^2.$$

This implies that $E\left(y_t^2\right) = (1-\omega)\,\lambda_t\,(\lambda_t+1)$ and consequently, the variance is given by $\sigma_{y_t}^2 = E\left(y_t^2\right) - [E(y_t)]^2 = (1-\omega)\,\lambda_t\,(1+\lambda_t\omega)$.

By finding the third and fourth derivatives, we can find $E\left(y_t^3\right)$ and $E\left(y_t^4\right)$. We have $G'''(1) = E\left(y_t(y_t-1)(y_t-2)\right) = E\left(y_t^3\right) - 3E\left(y_t^2\right) + 2E(y_t) = (1-\omega)\lambda_t^3$ and $G''''(1) = E\left(y_t(y_t-1)(y_t-2)(y_t-3)\right) = E\left(y_t^4\right) - 6E\left(y_t^3\right) + 11E\left(y_t^2\right) - 6E(y_t) = (1-\omega)\lambda_t^4$ leads to $E\left(y_t^3\right) = (1-\omega)\lambda_t\left(\lambda_t^2+3\lambda_t+1\right)$ and $E\left(y_t^4\right) = (1-\omega)\lambda_t\left(\lambda_t^3+6\lambda_t^2+7\lambda_t+1\right)$. Using $E\left(y_t^3\right)$, we can find the third moment of ZIP model such that

$$E\left[\left(y_t-(1-\omega)\lambda_t\right)^3\right] = E\left\{y_t^3 - 3y_t^2(1-\omega)\lambda_t + 3y_t\left[(1-\omega)\lambda_t\right]^2 - (1-\omega)^3\lambda_t^3\right\}.$$

Then, using the expectation properties, we have

$$\begin{aligned}
E\left[\left(y_t-(1-\omega)\lambda_t\right)^3\right] &= E\left(y_t^3\right) - 3E\left(y_t^2\right)(1-\omega)\lambda_t + 3E(y_t)\left[(1-\omega)\lambda_t\right]^2 - (1-\omega)^3\lambda_t^3, \\
&= (1-\omega)\lambda_t\left\{\lambda_t^2+3\lambda_t+1\right\} - 3(1-\omega)^2\lambda_t^2\left\{\lambda_t+1\right\} \\
&\quad +3(1-\omega)^3\lambda_t^3 - (1-\omega)^3\lambda_t^3. \qquad (7.4)
\end{aligned}$$

Lastly, we factorize Equation (7.4) giving

$$\begin{aligned}
E\left[\left(y_t-(1-\omega)\lambda_t\right)^3\right] &= (1-\omega)\lambda_t\left\{\lambda_t^2+3\lambda_t+1-3(1-\omega)\lambda_t\left[\lambda_t+1\right]-(1-\omega)^2\lambda_t^2\right\}, \\
&= (1-\omega)\lambda_t\left\{\begin{array}{c} \lambda_t^2+3\lambda_t+1-3\lambda_t^2-3\lambda_t+3\omega\lambda_t^2-3\omega\lambda_t \\ +2\lambda_t^2-4\lambda_t^2\omega+2\lambda_t^2\omega^2 \end{array}\right\}, \\
&= (1-\omega)\lambda_t\left\{1+3\omega\lambda_t-\lambda_t^2\omega+2\lambda_t^2\omega^2\right\}. \qquad (7.5)
\end{aligned}$$

And, for the fourth moment, apply the same approach, giving

$$\begin{aligned}
E\left[\left(y_t-(1-\omega)\lambda_t\right)^4\right] &= E\left\{\begin{array}{c} y_t^4-4(1-\omega)\lambda_t y_t^3+6y_t^2(1-\omega)^2\lambda_t^2 \\ -4y_t(1-\omega)^3\lambda_t^3+(1-\omega)^4\lambda_t^4 \end{array}\right\}, \\
&= E\left(y_t^4\right)-4(1-\omega)\lambda_t E\left(y_t^3\right)+6E\left(y_t^2\right)(1-\omega)^2\lambda_t^2 \\
&\quad -4E(y_t)(1-\omega)^3\lambda_t^3+(1-\omega)^4\lambda_t^4. \qquad (7.6)
\end{aligned}$$

Substituting $E(y_t)$, $E\left(y_t^2\right)$,$E\left(y_t^3\right)$ and $E\left(y_t^4\right)$ into Equation (7.6), we have

$$E\left[\left(y_t-(1-\omega)\lambda_t\right)^4\right]=(1-\omega)\lambda_t\left\{\omega\lambda_t^3\left(3\omega^2-3\omega+1\right)+6\omega^2\lambda_t^2+\lambda_t(4\omega+3)+1\right\}. \tag{7.7}$$

Using (7.5) and the variance, the skewness for ZIP, $\Gamma_t^{(ZIP)}$ can be obtained

$$\begin{aligned}\Gamma_t^{(ZIP)} &= \frac{E\left(y_t-\mu_t\right)^3}{\left(var(y_t)\right)^{3/2}},\\ &= \frac{(1-\omega)\lambda_t\left\{1+3\omega\lambda_t-\lambda_t^2\omega+2\lambda_t^2\omega^2\right\}}{\left[(1-\omega)\lambda_t\left(1+\lambda_t\omega\right)\right]^{3/2}},\\ &= \frac{1+3\omega\lambda_t-\lambda_t^2\omega+2\lambda_t^2\omega^2}{\sqrt{(1-\omega)\lambda_t\left(1+\lambda_t\omega\right)^2}}.\end{aligned} \tag{7.8}$$

For kurtosis, $K^{(ZIP)}$, we substitute (7.7) and the variance into (D.2) :

$$\begin{aligned}K^{(ZIP)} &= \frac{E\left(y_t-\mu_t\right)^4}{\left(var(y_t)\right)^2},\\ &= \frac{(1-\omega)\lambda_t\left\{\omega\lambda_t^3\left(3\omega^2-3\omega+1\right)+6\lambda_t^2\omega^2+\lambda_t(4\omega+3)+1\right\}}{\left[(1-\omega)\lambda_t\left(1+\lambda_t\omega\right)\right]^2},\\ &= \frac{\omega\lambda_t^3\left(3\omega^2-3\omega+1\right)+6\lambda_t^2\omega^2+\lambda_t(4\omega+3)+1}{(1-\omega)\lambda_t\left(1+\lambda_t\omega\right)^2}.\end{aligned} \tag{7.9}$$

7.2.2 The Moment Properties of Unconditional Distribution on ZIPINGARCH(p,q) Model

Applying the same approach as in Section (5.2.2) and Section (6.2.2), we first find μ in terms of λ via mean of the model. The mean of ZIPINGARCH(p,q) is

$$\mu=E(X_t)=E\left[E\left(X_t|\Im_{t-1}^X\right)\right]=E\left[(1-\omega)\lambda_t\right]$$

and for large t, λ_t approaches λ, a constant. Hence,

$$\mu=(1-\omega)\lambda. \tag{7.10}$$

Therefore, we rewrite Equation (7.10) as $\lambda = \frac{\mu}{1-\omega}$. The variance of the martingale difference u_t is given as

$$\begin{aligned}\sigma_u^2 &= E\left(u_t^2\right) \\ &= (1-\omega)E\left(\lambda_t\right)+(1-\omega)\omega E\left(\lambda_t^2\right). \end{aligned} \tag{7.11}$$

On the other hand, the skewness of u_t is

$$\begin{aligned}\Gamma^{(u)} &= \frac{E\left(u_t^3\right)}{\left[E\left(u_t^2\right)\right]^{3/2}}, \\ &= \frac{\omega(2\omega+1)E\left(\lambda_t^2\right)+3\omega E\left(\lambda_t\right)+1}{E\left[(1-\omega)\lambda_t\left(1+\omega\left(\lambda_t\right)\right)\right]^{3/2}} \end{aligned} \tag{7.12}$$

Similarly, using the result in Equation (7.12), the excess kurtosis of the process is given by

$$\begin{aligned}K^{(u)} &= \frac{E\left(u_t^4\right)}{\left[E\left(u_t^2\right)\right]^2}-3 \\ &= \frac{\omega(6\omega^2-6\omega+1)E\left(\lambda_t^3\right)+6\omega(2\omega-1)E\left(\lambda_t^2\right)+7\omega E\left(\lambda_t\right)+1}{E\left\{\left((1-\omega)\lambda_t\right)\left(1+\omega\lambda_t\right)\right\}^2}. \end{aligned} \tag{7.13}$$

Thus, we can rewrite Equations (7.11), (7.12) and (7.13) , therefore, we have the variance, skewness and kurtosis of u_t as

$$\sigma_u^2 = \mu(1+\frac{\omega\mu}{1-\omega}), \tag{7.14}$$

$$\Gamma^{(u)} = \frac{(1-\omega)^2+3\mu\omega(1-\omega)-\mu^2\omega+2\mu^2\omega^2}{\sqrt{(1-\omega)\mu(1-\omega+\mu\omega)^3}}, \tag{7.15}$$

$$K^{(u)} = \frac{\omega\mu^3\left(3\omega^2-3\omega+1\right)+6\omega^2\mu^2(1-\omega)+\mu(1-\omega)^2(4\omega+3)+(1-\omega)^3}{(1-\omega)^2\mu(1-\omega+\mu\omega)}. \tag{7.16}$$

Therefore, the variance, skewness and kurtosis of ZIPINGARCH(p,q) can be obtained by substituting Equations (7.14), (7.15) and (7.16) into Theorem 1(b), Theorem 2(b) and

(c) respectively which are

$$\sigma_X^2 = \mu(1+\frac{\omega\mu}{1-\omega})\sum_{j=0}^{\infty}\psi_j^2,$$

$$\Gamma^{(X)} = \frac{(1-\omega)^2+3\mu\omega(1-\omega)-\mu^2\omega+2\mu^2\omega^2\sum_{j=0}^{\infty}\psi_j^3}{\sqrt{(1-\omega)\mu(1-\omega+\mu\omega)^3}\left(\sum_{j=0}^{\infty}\psi_j^2\right)^{3/2}},$$

$$K^{(X)} = \frac{\omega\mu^3\left(3\omega^2-3\omega+1\right)+6\omega^2\mu^2(1-\omega)+\mu(1-\omega)^2(4\omega+3)+(1-\omega)^3\sum_{j=0}^{\infty}\psi_j^4}{(1-\omega)^2\mu(1-\omega+\mu\omega)\left(\sum_{j=0}^{\infty}\psi_j^2\right)^2}.$$

Here, we compare our moment properties of unconditional ZIPINGARCH(p,q) model with Zhu (2012) who obtained only the first two moments of the model with the order $p=1$ and $q=1$. In the thesis, we form a closed form expression for the moments of the model up to order four for all p and q.

7.2.3 Empirical Study

To demonstrate the moments structure are correctly derived, we do the empirical study for the cases when the process is close and far from the boundary of stationarity. Here, we generate 500 samples of size 2000, fixed the inflation parameter, $\omega = 0.2$ and calculate the mean and the mean square error of the estimated moments as tabulated in Table 7.1. It can be seen that the estimated values are close to the true values.

Table 7.1: Generated data and true values for the moment structures with $\gamma = 0.1$

		$\alpha_1 = 0.1$			$\alpha_1 = 0.5$		
		Generated	True	MSE	Generated	True	MSE
$\beta = 0.2$							
	μ_X	0.9881	0.9892	1.23E-03	1.152	1.159	1.35E-04
	σ_2^X	0.8741	0.8753	7.12E-04	1.068	1.076	5.69E-04
	$\Gamma^{(X)}$	1.335	1.358	1.88E-03	1.368	1.438	7.25E-03
	$K^{(X)}$	4.125	4.137	5.21E-03	4.745	4.751	8.27E-04
$\beta = 0.4$							
	μ_X	0.7661	0.7682	1.25E-03	1.021	1.035	2.04E-03
	σ_2^X	0.8768	0.8753	5.11E-03	1.367	1.355	5.11E-03
	$\Gamma^{(X)}$	1.621	1.612	4.15E-03	1.812	1.827	8.01E-03
	$K^{(X)}$	4.147	4.159	9.11E-04	4.132	4.144	8.29E-03

7.3 Quadratic Estimating Functions on ZIPINGARCH(p,q) Model

In order to apply the QEF method, we first consider the first four conditional moments of the model (see Liang et al. (2011)). The mean, variance, skewness and kurtosis of y_t conditional on $\Im_{t-1}^{X}$ of ZIPINGARCH(p,q) model are

$$\mu_t(\boldsymbol{\theta}) = (1-\omega)\lambda_t(\boldsymbol{\theta}),$$

$$\sigma_t^2(\boldsymbol{\theta}) = (1-\omega)\lambda_t(\boldsymbol{\theta})\left(1+\omega\lambda_t(\boldsymbol{\theta})\right),$$

$$\Gamma(\boldsymbol{\theta}) = \frac{\omega(2\omega+1)\lambda_t^2(\boldsymbol{\theta})+3\omega\lambda_t(\boldsymbol{\theta})+1}{\left((1-\omega)\lambda_t(\boldsymbol{\theta})\right)^{1/2}\left(1+\omega\lambda_t(\boldsymbol{\theta})\right)^{3/2}},$$

$$\kappa(\boldsymbol{\theta}) = \frac{\omega(6\omega^2-6\omega+1)\lambda_t^3(\boldsymbol{\theta})+6\omega(2\omega-1)\lambda_t^2(\boldsymbol{\theta})+7\omega\lambda_t(\boldsymbol{\theta})+1}{\left((1-\omega)\lambda_t(\boldsymbol{\theta})\right)\left(1+\omega\lambda_t(\boldsymbol{\theta})\right)^2}.$$

Then, by taking the martingale differences as $m_t(\boldsymbol{\theta}) = y_t - \mu_t(\boldsymbol{\theta})$, $s_t(\boldsymbol{\theta}) = m_t^2(\boldsymbol{\theta}) - \sigma_t^2(\boldsymbol{\theta})$, we have $\langle m\rangle_t = \sigma_t^2(\boldsymbol{\theta})$, $\langle s\rangle_t = \sigma_t^4(\boldsymbol{\theta})(\kappa_t(\boldsymbol{\theta})+2)$ and $\langle m,s\rangle_t = \sigma_t^3(\boldsymbol{\theta})\gamma_t(\boldsymbol{\theta})$. Using arguments similar to those in the other models, the derivatives of mean and variance with respect to $\boldsymbol{\theta}$, $\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\theta}$ and $\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\theta}$ are

$$\begin{aligned}\frac{\partial\mu_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} &= \left(-\lambda_t,(1-\omega)\frac{\partial\lambda_t}{\partial\alpha_0},\ldots,(1-\omega)\frac{\partial\lambda_t}{\partial\alpha_p},(1-\omega)\frac{\partial\lambda_t}{\partial\beta_1},\ldots,(1-\omega)\frac{\partial\lambda_t}{\partial\beta_q}\right)',\\ &= \left(A_{1,t},A_{2,t},A_{(3,1,t)},\ldots,A_{(3,p,t)},\ldots,A_{(4,1,t)},\ldots,A_{(4,q,t)}\right)'.\end{aligned}$$

$$\begin{aligned}\frac{\partial\sigma_t^2(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} &= \left(-\lambda_t(1+2\omega\lambda_t),R_t\frac{\partial\lambda_t}{\partial\alpha_0},\ldots,R_t\frac{\partial\lambda_t}{\partial\alpha_p},R_t\frac{\partial\lambda_t}{\partial\beta_1},\ldots,R_t\frac{\partial\lambda_t}{\partial\beta_q},\right)',\\ &= \left(S_{1,t},S_{2,t},S_{(3,1,t)},\ldots,S_{(3,p,t)},\ldots,S_{(4,1,t)},\ldots,S_{(4,q,t)},\right)'.\end{aligned}$$

where define $R_t = (1-\omega)(1+2\omega\lambda_t(\boldsymbol{\theta}))$, $A_{(i,j,t)}$ is the partial derivative of the mean with respect to each parameter and $S_{(i,j,t)}$ is the partial derivative of the variance with respect to each parameter. Now, we let $A_{k,t}^m = \frac{-A_k}{\langle m\rangle_t}$, $S_{k,t}^m = \frac{S_k\langle m,s\rangle_t}{\langle m\rangle_t\langle s\rangle_t}$, $P_{k,t}^v = \frac{A_k\langle m,s\rangle_t}{\langle m\rangle_t\langle s\rangle_t}$, and $Q_{k,t}^v = \frac{-S_k}{\langle s\rangle_t}$.Hence, the optimal estimating function is given by

$$g_Q^*(\boldsymbol{\theta}) = \sum_{t=1}^{n}(\mathbf{a}_{t-1}^*m_t + \mathbf{b}_{t-1}^*s_t),$$

where

$$\mathbf{a}_{t-1}^* = R_t \left(A_{1,t}^m + S_{1,t}^m, \dots, A_{3,1,t}^m + S_{3,1,t}^m, \dots, A_{3,p,t}^m + S_{3,p,t}^m, A_{4,1,t}^m + S_{4,1,t}^m \dots, A_{4,q,t}^m + S_{4,q,t}^m\right),$$

$$\mathbf{b}_{t-1}^* = R_t \left(P_{1,t}^v + Q_{1,t}^v, \dots, P_{3,1,t}^v + Q_{3,1,t}^v, \dots, P_{3,p,t}^v + Q_{3,p,t}^v, P_{4,1,t}^v + Q_{4,1,t}^v \dots, P_{4,q,t}^v + Q_{4,q,t}^v\right).$$

Hence, for each component of $\boldsymbol{\theta}$, the optimal quadratic estimating functions can be formed as follow:

$$g_Q^*(\omega) = \sum_{t=1}^{n} R_t \left(\left(A_{1,t}^m + S_{1,t}^m\right) m_t + \left(P_{1,t}^v + Q_{1,t}^v\right) s_t\right),$$

$$g_Q^*(\gamma) = \sum_{t=1}^{n} R_t \left(\left(A_2^m + S_{2,t}^m\right) m_t + \left(P_{2,t}^v + Q_{2,t}^v\right) s_t\right),$$

$$g_Q^*(\alpha_i) = \sum_{t=1}^{n} R_t \left(\left(A_{(3,i,t)}^m + S_{(3,i,t)}^m\right) m_t + \left(P_{(3,i,t)}^v + Q_{(3,i,t)}^v\right) s_t\right) \qquad i = 1, \dots, p \quad \text{and}$$

$$g_Q^*(\beta_j) = \sum_{t=1}^{n} R_t \left(\left(A_{(4,j,t)}^m + S_{(4,j,t)}^m\right) m_t + \left(P_{(4,j,t)}^v + Q_{(4,j,t)}^v\right) s_t\right) \qquad j = 1, \dots, q.$$

The parameters of interest can be estimated by solving the optimal quadratic estimating functions, $g_Q^*(\boldsymbol{\theta})$. Here, we use R-cran via *nlminb* to obtain the estimates.

We also can find the information matrix of the optimal estimating function for $\boldsymbol{\theta}$ using Theorem 4.1.1(b) and this is given as

$$\mathbf{I}_{g_Q^*}(\boldsymbol{\theta}) = \begin{pmatrix} I_{\omega\omega}^Q & I_{\omega\alpha_0}^Q & I_{\omega\alpha_i}^Q & I_{\omega\beta_j} \\ I_{\alpha_0\omega}^Q & I_{\alpha_0\alpha_0}^Q & I_{\alpha_0\alpha_i}^Q & I_{\alpha_0\beta_j} \\ I_{\alpha_i\omega}^Q & I_{\alpha_i\alpha_0}^Q & I_{\alpha_i\alpha_i}^Q & I_{\alpha_i\beta_j} \\ I_{\beta_j\omega}^Q & I_{\beta_j\alpha_0}^Q & I_{\beta_j\alpha_i}^Q & I_{\beta_j\beta_j} \end{pmatrix}$$

where the elements in the matrix are

$$I_{\omega\omega}^Q = \sum_{t=1}^{n} R_t \left[\frac{A_{1,t}^2}{\langle m \rangle_t} + \frac{S_{1,t}^2}{\langle s \rangle_t} - 2A_{1,t}S_{1,t}\frac{\langle m, s \rangle_t}{\langle m \rangle_t \langle s \rangle_t}\right],$$

$$I^Q_{\alpha_0\alpha_0} = \sum_{t=1}^{n} R_t \left[\frac{A_{1,t}^2}{\langle m \rangle_t} + \frac{S_{2,t}^2}{\langle s \rangle_t} - 2A_{2,t}S_{2,t} \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\alpha_i\alpha_i} = \sum_{t=1}^{n} R_t \left[\frac{A_{(3,i,t)}^2}{\langle m \rangle_t} + \frac{S_{(3,i,t)}^2}{\langle s \rangle_t} - 2A_{(3,i,t)}S_{(3,i,t)} \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\beta_j\beta_j} = \sum_{t=1}^{n} R_t \left[\frac{A_{(4,j,t)}^2}{\langle m \rangle_t} + \frac{S_{(4,j,t)}^2}{\langle s \rangle_t} - 2A_{(4,j,t)}S_{(4,j,t)} \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\omega\alpha_0} = \sum_{t=1}^{n} R_t \left[\frac{A_{1,t}A_{2,t}}{\langle m \rangle_t} + \frac{S_{1,t}S_{2,t}}{\langle s \rangle_t} - (A_{1,t}S_{2,t} + S_{1,t}A_{2,t}) \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right]$$

$$I^Q_{\omega\alpha_i} = \sum_{t=1}^{n} R_t \left[\frac{A_{1,t}A_{(3,i,t)}}{\langle m \rangle_t} + \frac{S_{1,t}S_{(3,i,t)}}{\langle s \rangle_t} - (A_{1,t}S_{(3,i,t)} + S_{1,t}A_{(3,i,t)}) \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\omega\beta_j} = \sum_{t}^{n} R_t \left[\frac{A_{1,t}A_{(4,i)}}{\langle m \rangle_t} + \frac{S_{1,t}S_{(4,i)}}{\langle s \rangle_t} - (A_{1,t}S_{(4,i,t)} + S_{1,t}A_{(4,i,t)}) \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\alpha_0\alpha_i} = \sum_{t=1}^{n} R_t \left[\frac{A_{(3,i,t)}A_{2,t}}{\langle m \rangle_t} + \frac{S_{(3,i,t)}S_{2,t}}{\langle s \rangle_t} - (A_{2,t}S_{(3,i,t)} + S_{2,t}A_{(3,i,t)}) \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\alpha_0\beta_j} = \sum_{t=1}^{n} R_t \left[\frac{A_{(4,j,t)}A_{2,t}}{\langle m \rangle_t} + \frac{S_{(4,j,t)}S_{2,t}}{\langle s \rangle_t} - (A_{2,t}S_{(4,j,t)} + S_{2,t}A_{(4,j,t)}) \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right],$$

$$I^Q_{\alpha_i\beta_j} = \sum_{t=1}^{n} R_t \left[\frac{A_{(3,i,t)}A_{(4,j,t)}}{\langle m \rangle_t} + \frac{A_{(3,j)}S_{(4,j,t)}}{\langle s \rangle_t} - (A_{(3,i,t)}S_{(4,j,t)} + S_{(3,i,t)}A_{(4,j,t)}) \frac{\langle m,s \rangle_t}{\langle m \rangle_t \langle s \rangle_t} \right]$$

and remaining elements can be obtained by the symmetry of the matrix: $I_{\alpha_0\omega} = I_{\omega\alpha_0}$, $I_{\alpha_i\omega} = I_{\omega\alpha_i}$, $I_{\beta_j\omega} = I_{\omega\beta_j}$, $I_{\alpha_i\alpha_0} = I_{\alpha_0\alpha_i}$, $I_{\beta_j\alpha_0} = I_{\alpha_0\beta_j}$ and $I_{\beta_j\alpha_i} = I_{\alpha_i\beta_j}$.

In addition, we can also derive the optimal estimating function m_t which is given by

$$g_m^*(\theta) = -\sum_{t=1}^{n} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \left\{ \frac{X_t - (1-\omega)\lambda_t(\boldsymbol{\theta})}{(1-\omega)\lambda_t(\boldsymbol{\theta})(1+(1-\omega)\lambda_t(\boldsymbol{\theta}))} \right\} \tag{7.17}$$

and the optimal estimating function on s_t given by

$$g_s^*(\theta) = -\sum_{t=1}^{n} \frac{\partial \sigma_t^2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \left\{ \frac{m_t^2(\boldsymbol{\theta}) - \sigma_t^4(\boldsymbol{\theta})}{\sigma_t^4(\kappa_t + 2)} \right\}. \tag{7.18}$$

Then, the corresponding information matrices for $m_t(\boldsymbol{\theta})$ and $s_t(\boldsymbol{\theta})$ identified as $\mathbf{I}_{\mathbf{g}_m^*}(\boldsymbol{\theta})$ and $\mathbf{I}_{\mathbf{g}_s^*}(\boldsymbol{\theta})$ respectively are

$$\mathbf{I}_{\mathbf{g}_m^*}(\boldsymbol{\theta}) = \begin{pmatrix} I^m_{\gamma\gamma} & I^m_{\gamma\alpha_i} & I^m_{\gamma\beta_j} & I^m_{\gamma\omega} \\ I^m_{\alpha_i\gamma} & I^m_{\alpha_i\alpha_i} & I^m_{\alpha_i\beta_j} & I^m_{\gamma\omega} \\ I^m_{\beta_j\gamma} & I^m_{\beta_j\alpha_i} & I^m_{\beta_j\beta_j} & I^m_{\gamma\omega} \\ I^m_{\omega\gamma} & I^m_{\omega\alpha_i} & I^m_{\omega\beta_j} & I^m_{\omega\omega} \end{pmatrix} \; ; \; \mathbf{I}_{\mathbf{g}_m^*}(\boldsymbol{\theta}) = \begin{pmatrix} I^s_{\gamma\gamma} & I^s_{\gamma\alpha_i} & I^s_{\gamma\beta_j} & I^s_{\gamma\omega} \\ I^s_{\alpha_i\gamma} & I^s_{\alpha_i\alpha_i} & I^s_{\alpha_i\beta_j} & I^s_{\gamma\omega} \\ I^s_{\beta_j\gamma} & I^s_{\beta_j\alpha_i} & I^s_{\beta_j\beta_j} & I^s_{\gamma\omega} \\ I^s_{\omega\gamma} & I^s_{\omega\alpha_i} & I^s_{\omega\beta_j} & I^s_{\omega\omega} \end{pmatrix}$$

. The elements of $\mathbf{I}_{\mathbf{g}_m^*}$ matrix are

$$I^m_{\omega\omega} = \sum_{t=1}^{n}\left[\frac{A^2_{1,t}}{\langle m\rangle_t}\right];\; I^m_{\gamma\gamma} = \sum_{t=1}^{n}\left[\frac{A^2_{2,t}}{\langle m\rangle_t}\right];\; I^m_{\alpha_i\alpha_i} = \sum_{t=1}^{n}\left[\frac{A^2_{(3,i,t)}}{\langle m\rangle_t}\right];\; I^m_{\beta_j\beta_j} = \sum_{t=1}^{n}\left[\frac{A^2_{(4,j,t)}}{\langle m\rangle_t}\right];$$
$$I^m_{\omega\gamma} = \sum_{t=1}^{n}\left[\frac{A_{1,t}A_{2,t}}{\langle m\rangle_t}\right];\; I^m_{\omega\alpha_i} = \sum_{t=1}^{n}\left[\frac{A_{1,t}A_{(3,i,t)}}{\langle m\rangle_t}\right];\; I^m_{\omega\beta_j} = \sum_{t=1}^{n}\left[\frac{A_{1,t}A_{(4,j,t)}}{\langle m\rangle_t}\right];$$
$$I^m_{\gamma\alpha_i} = \sum_{t=1}^{n}\left[\frac{A_{2,t}A_{(3,i,t)}}{\langle m\rangle_t}\right];\; I^m_{\gamma\beta_j} = \sum_{t=1}^{n}\left[\frac{A_{2,t}A_{(4,j,t)}}{\langle m\rangle_t}\right];\; \text{and } I^m_{\alpha_i\beta_j} = \sum_{t=1}^{n}\left[\frac{A_{(3,i,t)}A_{(4,j,t)}}{\langle m\rangle_t}\right],$$

and the elements of $\mathbf{I}_{\mathbf{g}_s^*}$ are

$$I^m_{\omega\omega} = \sum_{t=1}^{n}\left[\frac{S^2_{1,t}}{\langle m\rangle_t}\right];\; I^m_{\gamma\gamma} = \sum_{t=1}^{n}\left[\frac{S^2_{2,t}}{\langle m\rangle_t}\right];\; I^m_{\alpha_i\alpha_i} = \sum_{t=1}^{n}\left[\frac{S^2_{(3,i,t)}}{\langle m\rangle_t}\right];\; I^m_{\beta_j\beta_j} = \sum_{t=1}^{n}\left[\frac{S^2_{(4,j,t)}}{\langle m\rangle_t}\right];$$
$$I^m_{\omega\gamma} = \sum_{t=1}^{n}\left[\frac{S_{1,t}S_{2,t}}{\langle m\rangle_t}\right];\; I^m_{\omega\alpha_i} = \sum_{t=1}^{n}\left[\frac{S_{1,t}S_{(3,i,t)}}{\langle m\rangle_t}\right];\; I^m_{\omega\beta_j} = \sum_{t=1}^{n}\left[\frac{S_{1,t}S_{(4,j,t)}}{\langle m\rangle_t}\right];$$
$$I^m_{\gamma\alpha_i} = \sum_{t=1}^{n}\left[\frac{S_{2,t}S_{(3,i,t)}}{\langle m\rangle_t}\right];\; I^m_{\gamma\beta_j} = \sum_{t=1}^{n}\left[\frac{S_{2,t}S_{(4,j,t)}}{\langle m\rangle_t}\right];\; \text{and } I^m_{\alpha_i\beta_j} = \sum_{t=1}^{n}\left[\frac{S_{(3,i,t)}S_{(4,j,t)}}{\langle m\rangle_t}\right].$$

From the information gain using QEF over its components, $m_t(\boldsymbol{\theta})$ and $s_t(\boldsymbol{\theta})$, it is clearly seen from $I^Q_{\omega\omega} > I^m_{\omega\omega}$, $I^Q_{\omega\omega} > I^s_{\omega\omega}$, $I^Q_{\gamma\gamma} > I^m_{\gamma\gamma}$, $I^Q_{\gamma\gamma} > I^s_{\gamma\gamma}$, $I^Q_{\alpha_i\alpha_i} > I^m_{\alpha_i\alpha_i}$, $I^Q_{\alpha_i\alpha_i} > I^s_{\alpha_i\alpha_i}$, $I^Q_{\beta_j\beta_j} > I^m_{\beta_j\beta_j}$, $I^Q_{\beta_j\beta_j} > I^s_{\beta_j\beta_j}$, $I^Q_{\omega\gamma} > I^m_{\omega\gamma}$, $I^Q_{\omega\gamma} > I^s_{\omega\gamma}$, $I^Q_{\omega\alpha_i} > I^m_{\omega\alpha_i}$, $I^Q_{\omega\alpha_i} > I^s_{\omega\alpha_i}$, $I^Q_{\omega\beta_j} > I^m_{\omega\beta_j}$, $I^Q_{\omega\beta_j} > I^s_{\omega\beta_j}$, $I^Q_{\gamma\alpha_i} > I^m_{\gamma\alpha_i}$, $I^Q_{\gamma\alpha_i} > I^s_{\gamma\alpha_i}$, $I^Q_{\gamma\beta_j} > I^m_{\gamma\beta_j}$, $I^Q_{\gamma\beta_j} > I^s_{\gamma\beta_j}$, $I^Q_{\alpha_i\beta_j} > I^m_{\alpha_i\beta_i}$, and $I^Q_{\alpha_i\beta_j} > I^s_{\alpha_i\beta_i}$. Since the information associated with QEF is larger than the information associated with the component estimating functions. Therefore, we can say that QEF estimator is more informative than the EF estimator.

7.4 Performance of The Estimation Methods on ZIPINGARCH $(1,1)$

To evaluate the performance of QEF, MLE and EF methods, a simulation study on ZIPINGARCH $(1,1)$ was carried out with the inflation parameter $\omega = 0.2$, replication $N = 500$, and two sample sizes, $n = 100$ and $n = 2000$. The process is defined as

$$X_t|\Im_{t-1}^{X} \sim ZIP(\lambda_t, \omega), \tag{7.19}$$

$$\lambda_t = \gamma + \alpha_1 X_{t-1} + \beta_1 \lambda_{t-1}. \tag{7.20}$$

7.4.1 MLE Derivation of ZIPINGARCH $(1,1)$

Since ZIPINGARCH $(1,1)$ has two separate probability mass functions (pmf), the loglikelihood for a model has two parts namely

$$\mathscr{L} = \begin{cases} \sum_{t=1}^{\infty} ln\left\{\omega + (1-\omega)\,e^{-\lambda_t}\right\} & \text{for } X_t = 0, \\ \sum_{t=1}^{\infty} \{ln\ (1-\omega) - \lambda_t + X_t\ ln\,\lambda_t - ln(X_t!)\} & \text{for } X_t = 1,2,3,\ldots \end{cases} \tag{7.21}$$

We then maximized the loglikelihood to extract the estimated parameter of interest in R-cran.

7.4.2 EF Derivation of ZIPINGARCH $(1,1)$

Using the same approach as Section 6.4.2, we derive the optimal EF function through Equation (7.17). The derivative of mean, μ_t with respect to $\boldsymbol{\theta}$ given in following form

$$\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{cases} \lambda_t & \text{for } \theta = \omega \\ -(1-\omega)\,\lambda_t' \ \text{ for } \theta = \gamma, \alpha_1, \beta_1. \end{cases} \tag{7.22}$$

Hence, from Theorem 1, the optimal for each component $g_E^*(\boldsymbol{\theta})$ of the model are:

$$g_E^*(\gamma) = \sum_{t=1}^{n}\left[\left\{\frac{-1-\beta_1\frac{d\lambda_t}{d\boldsymbol{\gamma}}}{\lambda_t(1+\omega\lambda_t)}\right\}(X_t-(1-\omega)\,\lambda_t)\right], \tag{7.23}$$

$$g_E^*(\alpha_i) = \sum_{t=1}^{n}\left[\left\{\frac{-X_{t-1}-\beta_1\frac{d\lambda_t}{d\boldsymbol{\alpha}_1}}{\lambda_t(1+\omega\lambda_t)}\right\}(X_t-(1-\omega)\,\lambda_t)\right], \tag{7.24}$$

$$g_E^*(\beta_j) = \sum_{t=1}^{n}\left[\left\{\frac{-\lambda_{t-1}-\beta_1\frac{d\lambda_t}{d\boldsymbol{\beta}_1}}{\lambda_t(1+\omega\lambda_t)}\right\}(X_t-(1-\omega)\,\lambda_t)\right], \tag{7.25}$$

$$g_E^*(\omega) \quad = \quad \sum_{t=1}^{n} \left[\frac{X_t - (1-\omega)\lambda_t}{(1-\omega)(1+\lambda_t\omega)} \right]. \tag{7.26}$$

We solve the simultaneous Equation of optimal $g_E^* = 0$ using *nleqslv* command in R-cran to get the estimates of interest.

7.4.3 QEF Derivation of ZIPINGARCH $(1,1)$

For the QEF method on the model, we use the derivation in Section (7.3) which leads to

$$g_Q^*(\omega) = \sum_{t=1}^{n} R_t \left(\left(A_{1,t}^m + S_{1,t}^m \right) m_t + \left(P_{1,t}^v + Q_{1,t}^v \right) s_t \right),$$

$$g_Q^*(\gamma) = \sum_{t=1}^{n} R_t \left(\left(A_{2,t}^m + S_{2,t}^m \right) m_t + \left(P_{2,t}^v + Q_{2,t}^v \right) s_t \right),$$

$$g_Q^*(\alpha_1) = \sum_{t=1}^{n} R_t \left(\left(A_{(3,1,t)}^m + S_{(3,1,t)}^m \right) m_t + \left(P_{(3,1,t)}^v + Q_{(3,1,t)}^v \right) s_t \right) \quad \text{and}$$

$$g_Q^*(\beta_1) = \sum_{t=1}^{n} R_t \left(\left(A_{(4,1,t)}^m + S_{(4,1,t)}^m \right) m_t + \left(P_{(4,1,t)}^v + Q_{(4,1,t)}^v \right) s_t \right).$$

Again, we solve such simultaneous Equation $g_E^* = 0$ using *nlesqv* command in R-cran software.

7.4.4 Simulation Study

The steps of QEF algorithm for ZIPINGARCH $(1,1)$ are almost the same as in Section 5.4.3. Since we have one extra parameter, ω, in this model, in Step 1, the data are generated for a given value of the vectors of parameters $(\omega, \gamma, \alpha_1, \beta_1)$. In Step 2, the INGARCH $(1,1)$ model will be changed to ZIPINGARCH $(1,1)$ model and for Step 3, we use the similar method as in Section 6.4.4.

7.4.5 The Result

The performance is measured based on the mean, bias, standard error(SE) and mean squared error (MSE) from the repetitions $(N = 500)$ for each parameter from EF, MLE and QEF estimators. These are shown in Tables 7.2-7.11 .

Discussion

Since the MLE for this model does not have a closed form, we need to search for the maximum value of the likelihood function. Note that in this model the likelihood is divided into two parts with the zero and nonzero observed data. Therefore, the EF and QEF methods are computationally more efficient and easy to apply in practice than MLE method.

From Tables 7.2-7.11, we can see that, in term of biasness, the QEF method produce the smallest bias compared to other two methods. Furthermore, the QEF method produce the better result in terms of estimated bias, standard error and mean square error compared to EF estimator in all tables, either small sample size, $n = 100$ or large sample sizes, $n = 1500$ and $n = 2000$. The simulation results obtained support theoretical results observed earlier in the study. Therefore, we can conclude that the QEF estimator is superior compared to EF estimator.

On the other hand, although MLE method gives slightly smaller value of standard errors mean square errors compared EF and QEF methods in some cases, specifically we can see from Table 7.2-7.6, which are satisfied the stationarity condition, but the values increase as the combination set of parameters approach the non-stationarity conditions as shown in Table 7.7-7.11. Compared with EF and QEF methods, the value remain small and do have no effect on these combination set of parameters for all sample sizes. Therefore, we can say that, the EF and QEF estimators are better compared to MLE method for ZIPINGARCH$(1,1)$ case.

Therefore, from the results obtained, we can say that the QEF method performs well compared to EF and MLE methods. Hence, we can conclude that, the QEF estimator is better estimation method in this IVTS model.

Table 7.2: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.2$, and $\beta_1 = 0.3$

	n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$									
Mean	0.049	0.051	0.055	0.07	0.109	0.108	0.069	0.106	0.104
Bias	0.051	0.049	0.045	0.030	0.009	0.008	0.031	0.006	0.004
SE	0.087	0.091	0.082	0.030	0.056	0.032	0.015	0.016	0.015
MSE	5.23E-03	5.85E-03	5.38E-03	1.30E-03	1.72E-03	1.49E-03	1.17E-03	1.19E-03	1.08e-03
$\hat{\alpha}_1$									
Mean	0.108	0.107	0.122	0.139	0.210	0.197	0.138	0.193	0.199
Bias	0.092	0.093	0.078	0.061	0.010	0.003	0.062	0.007	0.001
SE	0.075	0.079	0.069	0.047	0.064	0.054	0.041	0.042	0.033
MSE	8.54E-03	8.66E-03	8.50E-03	2.53E-03	2.82E-03	1.39E-03	1.02E-03	1.09E-03	1.02e-03
$\hat{\beta}_1$									
Mean	0.353	0.355	0.350	0.321	0.326	0.297	0.322	0.311	0.302
Bias	0.053	0.055	0.050	0.021	0.026	0.003	0.022	0.011	0.002
SE	0.152	0.15	0.144	0.076	0.082	0.074	0.075	0.075	0.058
MSE	6.31E-03	6.37E-03	6.28E-03	2.22E-03	2.84E-03	1.87E-03	1.19E-03	1.18E-03	1.05e-03
$\hat{\omega}$									
Mean	0.221	0.228	0.218	0.188	0.233	0.208	0.188	0.228	0.203
Bias	0.021	0.028	0.018	0.012	0.033	0.008	0.012	0.028	0.003
SE	0.099	0.114	0.095	0.024	0.096	0.042	0.016	0.029	0.013
MSE	5.37E-03	6.21E-03	5.44E-03	1.74E-03	2.42E-03	1.14E-03	1.04E-03	1.09E-03	1.02e-03

Table 7.3: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.4$, and $\beta_1 = 0.1$

	n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$									
Mean	0.139	0.142	0.148	0.157	0.182	0.193	0.158	0.196	0.198
Bias	0.061	0.058	0.052	0.043	0.018	0.007	0.042	0.004	0.002
SE	0.078	0.081	0.079	0.043	0.087	0.036	0.038	0.038	0.015
MSE	9.31E-03	9.52E-03	9.38E-03	3.62E-03	4.34E-03	2.42E-03	2.48E-03	2.42E-03	1.87e-03
$\hat{\alpha}_1$									
Mean	0.316	0.324	0.416	0.316	0.324	0.416	0.317	0.396	3.999e-01
Bias	0.084	0.076	0.016	0.084	0.076	0.016	0.083	0.004	1.02e-04
SE	0.065	0.075	0.055	0.065	0.075	0.055	0.060	0.061	0.013
MSE	1.14E-03	1.17E-03	1.15E-03	1.14E-03	1.17E-03	1.15E-03	1.07E-03	1.07E-03	1.02e-03
$\hat{\beta}_1$									
Mean	0.153	0.155	0.151	0.119	0.115	0.111	0.119	0.107	0.104
Bias	0.053	0.055	0.051	0.019	0.015	0.011	0.019	0.007	0.004
SE	0.096	0.092	0.083	0.015	0.021	0.013	0.013	0.014	0.004
MSE	1.23E-02	2.11E-02	9.88E-03	7.05E-03	7.47E-03	3.82E-03	2.01E-03	2.02E-03	1.39e-03
$\hat{\omega}$									
Mean	0.092	0.099	0.106	0.109	0.180	0.188	0.107	0.185	0.194
Bias	0.108	0.101	0.094	0.091	0.020	0.012	0.093	0.015	0.006
SE	0.095	0.091	0.085	0.070	0.080	0.007	0.042	0.039	0.002
MSE	7.28E-03	7.31E-03	7.26E-03	1.64E-03	4.23E-03	1.66E-03	1.05E-03	1.24E-03	1.07e-03

Table 7.4: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1$, and $\beta_1 = 0.4$

	n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$									
Mean	0.217	0.220	0.228	0.241	0.305	0.302	0.235	0.302	0.301
Bias	0.083	0.080	0.072	0.059	0.005	0.002	0.065	0.002	7.22e-04
SE	0.082	0.075	0.066	0.037	0.010	0.005	0.008	0.007	0.001
MSE	9.27E-03	9.31E-03	9.11E-03	3.93E-03	5.24E-03	2.39E-03	1.15E-03	1.16E-03	1.03e-03
$\hat{\alpha}_1$									
Mean	0.045	0.047	0.052	0.070	0.088	0.107	0.071	0.091	0.101
Bias	0.055	0.053	0.038	0.030	0.012	0.007	0.029	0.009	0.001
SE	0.087	0.082	0.075	0.029	0.039	0.022	0.015	0.019	0.017
MSE	5.11E-03	5.31E-03	5.18E-03	1.75E-03	2.75E-03	1.15E-03	1.02E-03	1.02E-03	9.05e-04
$\hat{\beta}_1$									
Mean	0.325	0.359	0.341	0.427	0.419	0.417	0.437	0.408	0.402
Bias	0.075	0.071	0.059	0.027	0.019	0.017	0.037	0.008	0.002
SE	0.075	0.071	0.064	0.025	0.063	0.050	0.013	0.015	0.009
MSE	6.31E-03	6.88E-03	6.32E-03	3.21E-03	3.87E-03	3.53E-03	1.13E-03	1.16E-03	1.08e-03
$\hat{\omega}$									
Mean	0.147	0.151	0.159	0.180	0.233	0.227	0.179	0.214	0.201
Bias	0.053	0.049	0.041	0.020	0.033	0.027	0.021	0.014	0.001
SE	0.076	0.073	0.071	0.051	0.040	0.029	0.044	0.031	0.015
MSE	7.39E-03	7.54E-03	7.41E-03	3.17E-03	3.77E-03	1.46E-03	1.21E-03	1.27E-03	1.01e-03

Table 7.5: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.4$, and $\beta_1 = 0.2$

	n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$									
Mean	0.181	0.182	0.191	0.211	0.272	0.279	0.211	0.287	0.294
Bias	0.119	0.118	0.109	0.089	0.028	0.021	0.089	0.013	0.006
SE	0.062	0.059	0.055	0.045	0.050	0.037	0.012	0.024	0.011
MSE	7.32E-03	7.85E-03	7.46E-03	1.61E-03	1.52E-03	1.47E-03	1.27E-03	1.25E-03	1.01e-03
$\hat{\alpha}_1$									
Mean	0.308	0.312	0.324	0.335	0.363	0.378	0.351	0.385	0.408
Bias	0.092	0.088	0.076	0.065	0.037	0.022	0.049	0.015	0.008
SE	0.089	0.085	0.076	0.041	0.051	0.045	0.037	0.041	0.032
MSE	8.13E-03	8.28E-03	8.09E-03	2.68E-03	2.75E-03	2.65E-03	1.09E-03	1.10E-03	1.04e-03
$\hat{\beta}_1$									
Mean	0.252	0.248	0.239	0.223	0.221	0.210	0.223	0.211	0.205
Bias	0.052	0.048	0.039	0.023	0.021	0.010	0.023	0.011	0.005
SE	0.121	0.128	0.118	0.091	0.097	0.012	0.046	0.050	0.007
MSE	8.36E-03	8.43E-03	8.15E-03	5.01E-03	6.20E-03	3.93E-03	1.11E-03	1.10E-36	1.02e-03
$\hat{\omega}$									
Mean	0.253	0.256	0.252	0.197	0.196	0.202	0.199	0.201	0.201
Bias	0.053	0.054	0.052	0.003	0.004	0.002	0.001	0.001	7.23E-04
SE	0.082	0.079	0.062	0.042	0.074	0.022	0.037	0.038	0.012
MSE	8.15E-03	8.21E-03	7.03E-03	2.14E-03	3.21E-03	1.82E-03	1.07E-03	1.08E-03	7.52e-04

Table 7.6: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.5, \alpha_1 = 0.2$, and $\beta_1 = 0.3$

	n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$									
Mean	0.432	0.437	0.442	0.455	0.452	0.474	0.443	0.486	0.490
Bias	0.068	0.063	0.058	0.045	0.048	0.026	0.056	0.014	0.010
SE	0.043	0.039	0.031	0.014	0.031	0.026	0.011	0.013	0.012
MSE	7.35E-03	7.52E-03	7.39E-03	4.22E-03	4.52E-03	4.24E-03	1.92E-03	1.92E-03	1.12e-03
$\hat{\alpha}_1$									
Mean	0.129	0.132	0.143	0.138	0.175	0.182	0.137	0.185	0.197
Bias	0.071	0.068	0.057	0.062	0.025	0.018	0.063	0.015	0.003
SE	0.066	0.062	0.049	0.031	0.056	0.038	0.027	0.047	0.026
MSE	5.19E-03	5.27E-03	5.03E-03	1.04E-03	1.23E-03	1.09E-03	7.94E-04	8.27E-04	7.99e-04
$\hat{\beta}_1$									
Mean	0.391	0.388	0.371	0.351	0.364	0.361	0.362	0.317	0.309
Bias	0.091	0.088	0.071	0.052	0.064	0.061	0.062	0.017	0.009
SE	0.062	0.056	0.044	0.017	0.032	0.013	0.013	0.02	0.010
MSE	7.12E-03	6.53E-03	6.19E-03	2.47E-03	2.40E-03	1.64E-03	9.57E-04	9.54E-04	8.11e-04
$\hat{\omega}$									
Mean	0.159	0.159	0.165	0.190	0.230	0.210	0.190	0.221	0.215
Bias	0.041	0.041	0.035	0.010	0.030	0.010	0.010	0.021	0.005
SE	0.076	0.073	0.059	0.030	0.039	0.018	0.018	0.014	0.008
MSE	9.01E-03	9.25E-03	9.08E-03	1.09E-03	1.18E-03	1.02E-03	8.07E-04	8.08E-04	7.39e-04

Table 7.7: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma=0.1, \alpha_1=0.6$, and $\beta_1=0.3$

	n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$									
Mean	0.017	0.041	0.048	0.069	0.115	0.089	0.069	0.108	0.101
Bias	0.083	0.059	0.052	0.031	0.015	0.011	0.031	0.008	0.001
SE	0.689	0.127	0.119	0.411	0.029	0.024	0.309	0.01	0.008
MSE	3.68E-01	8.32E-03	7.11E-03	1.95E-02	2.58E-03	2.02E-03	1.16E-02	1.18E-03	1.03e-03
$\hat{\alpha}_1$									
Mean	0.542	0.568	0.576	0.582	0.612	0.590	0.582	0.592	0.597
Bias	0.058	0.032	0.024	0.018	0.012	0.010	0.018	0.008	0.003
SE	0.612	0.091	0.082	0.443	0.081	0.050	0.438	0.039	0.014
MSE	9.88E-02	8.23E-03	8.01E-03	6.84E-02	7.96E-03	5.78E-03	1.77E-02	1.81E-03	1.11e-03
$\hat{\beta}_1$									
Mean	0.241	0.279	0.282	0.321	0.290	0.294	0.321	0.293	0.299
Bias	0.059	0.021	0.018	0.021	0.010	0.006	0.021	0.007	0.001
SE	0.581	0.102	0.085	0.441	0.048	0.040	0.435	0.035	0.016
MSE	1.23E-02	1.33E-03	1.08E-03	7.21E-02	7.20E-03	5.60E-03	2.23E-02	2.25E-03	1.26e-03
$\hat{\omega}$									
Mean	0.148	0.179	0.181	0.180	0.188	0.189	0.230	0.208	0.203
Bias	0.052	0.021	0.019	0.020	0.012	0.011	0.030	0.008	0.003
SE	0.613	0.113	0.089	0.414	0.019	0.013	0.411	0.010	0.005
MSE	9.78E-02	8.52E-03	6.32E-03	1.46E-02	2.50E-03	1.51E-03	1.01E-02	1.02E-03	9.68e-04

Table 7.8: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.1, \alpha_1 = 0.7,$ and $\beta_1 = 0.2$

	n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$									
Mean	0.018	0.152	0.069	0.072	0.105	9.95E-02	0.072	0.104	0.1001
Bias	0.082	0.052	0.031	0.028	0.005	0.0005	0.028	0.004	0.0001
SE	0.813	0.072	0.065	0.412	0.027	0.017	0.410	0.015	0.010
MSE	8.88E-02	9.35E-03	7.21E-03	1.59E-02	2.79E-03	2.38E-03	1.02E-02	1.19E-03	1.02e-03
$\hat{\alpha}_1$									
Mean	0.646	0.721	0.713	0.689	0.719	0.689	0.68	0.689	0.694
Bias	0.054	0.021	0.013	0.011	0.019	0.011	0.02	0.011	0.006
SE	0.612	0.213	0.201	0.454	0.089	0.028	0.426	0.029	0.019
MSE	2.31E-01	2.51E-02	1.08E-02	7.59E-02	6.93E-03	2.48E-03	2.66E-02	2.63E-03	1.26e-03
$\hat{\beta}_1$									
Mean	0.251	0.232	0.225	0.222	0.220	0.211	0.222	0.209	0.195
Bias	0.051	0.032	0.025	0.022	0.020	0.011	0.022	0.009	0.005
SE	0.852	0.147	0.099	0.449	0.055	0.041	0.422	0.032	0.007
MSE	7.53E-02	9.18E-03	8.37E-03	2.94E-02	3.77E-03	2.42E-03	2.03E-02	2.04E-03	1.84e-03
$\hat{\omega}$									
Mean	0.229	0.215	0.210	0.204	0.207	0.204	0.206	0.204	0.203
Bias	0.029	0.015	0.010	0.004	0.007	0.004	0.006	0.004	0.003
SE	0.642	0.123	0.083	0.442	0.088	0.078	0.438	0.062	0.039
MSE	9.33E-02	8.99E-03	7.08E-03	3.11E-02	5.81E-03	3.21E-03	1.05E-02	1.11E-03	1.07e-03

Table 7.9: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.2, \alpha_1 = 0.8,$ and $\beta_1 = 0.1$

	n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$									
Mean	0.119	0.147	0.152	0.154	0.167	0.23	0.153	0.182	0.189
Bias	0.081	0.053	0.048	0.046	0.033	0.03	0.047	0.018	0.011
SE	0.618	0.099	0.078	0.422	0.043	0.026	0.418	0.022	0.019
MSE	8.35E-02	9.28E-03	7.44E-03	4.28E-02	4.17E-03	2.22E-03	3.45E-02	3.46E-03	1.15e-03
$\hat{\alpha}_1$									
Mean	0.696	0.703	0.719	0.716	0.732	0.789	0.716	0.749	0.793
Bias	0.104	0.097	0.081	0.084	0.068	0.011	0.084	0.051	0.007
SE	0.712	0.123	0.107	0.45	0.052	0.041	0.444	0.034	0.015
MSE	8.22E-02	8.97E-03	8.26E-03	5.50E-02	5.88E-03	5.71E-03	5.26E-02	3.29E-03	3.27e-03
$\hat{\beta}_1$									
Mean	0.152	0.132	0.124	0.121	0.118	0.114	0.125	0.108	0.102
Bias	0.052	0.032	0.024	0.021	0.018	0.014	0.025	0.008	0.002
SE	0.654	0.149	0.093	0.456	0.04	0.031	0.446	0.019	0.007
MSE	9.22E-02	8.39E-03	7.89E-03	2.42E-02	2.93E-03	2.58E-03	1.07E-02	1.07E-03	1.07e-03
$\hat{\omega}$									
Mean	0.148	0.217	0.211	0.195	0.209	0.205	0.204	0.199	2.008e-01
Bias	0.052	0.017	0.011	0.005	0.009	0.005	0.004	0.001	7.51E-04
SE	0.817	0.078	0.065	0.509	0.003	0.001	0.506	0.001	8.22e-04
MSE	1.23E-01	2.11E-02	1.28E-02	3.98E-02	4.37E-03	4.01E-03	1.98E-02	1.97E-03	1.65e-03

Table 7.10: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.3, \alpha_1 = 0.1,$ and $\beta_1 = 0.8$

	n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$									
Mean	0.351	0.328	0.221	0.311	0.375	0.289	0.285	0.357	0.308
Bias	0.051	0.028	0.021	0.011	0.075	0.011	0.015	0.057	0.008
SE	0.913	0.088	0.075	0.511	0.060	0.021	0.508	0.045	0.007
MSE	9.31E-02	8.17E-03	7.69E-03	2.03E-02	2.53E-03	2.09E-03	1.22E-02	1.22E-03	1.91e-03
$\hat{\alpha}_1$									
Mean	0.142	0.129	0.121	0.076	0.116	0.115	0.076	0.108	0.102
Bias	0.042	0.029	0.021	0.024	0.016	0.015	0.024	0.008	0.002
SE	0.613	0.118	0.093	0.414	0.062	0.058	0.411	0.043	0.013
MSE	7.58E-02	8.27E-03	5.32E-03	1.13E-02	2.11E-03	1.15E-03	1.08E-02	1.08E-03	1.01e-03
$\hat{\beta}_1$									
Mean	0.753	0.769	0.782	0.815	0.815	0.793	0.826	0.806	0.799
Bias	0.047	0.031	0.018	0.015	0.015	0.007	0.026	0.006	0.001
SE	0.793	0.088	0.057	0.525	0.030	0.016	0.518	0.018	0.004
MSE	8.21E-02	7.52E-03	6.21E-03	4.56E-02	4.58E-03	3.02E-03	1.32E-02	1.28E-03	1.19e-03
$\hat{\omega}$									
Mean	0.122	0.148	0.165	0.140	0.153	0.224	0.260	0.162	0.191
Bias	0.078	0.052	0.035	0.060	0.047	0.024	0.060	0.038	0.009
SE	0.513	0.097	0.061	0.414	0.064	0.042	0.412	0.042	0.026
MSE	9.85E-02	8.27E-03	7.08E-03	3.18E-02	4.48E-03	1.89E-03	1.19E-02	1.17E-03	1.02e-03

Table 7.11: Simulation results for ZIPINGARCH $(1,1)$ with $\gamma = 0.4, \alpha_1 = 0.3,$ and $\beta_1 = 0.6$

	n=100			n=1500			n=2000		
	ML	EF	QEF	ML	EF	QEF	ML	EF	QEF
$\hat{\gamma}$									
Mean	0.466	0.448	0.435	0.364	0.488	0.448	0.361	0.415	0.401
Bias	0.066	0.048	0.035	0.036	0.088	0.048	0.039	0.015	0.001
SE	0.568	0.105	0.098	0.443	0.063	0.061	0.437	0.054	0.039
MSE	8.39E-02	9.33E-03	8.21E-03	3.70E-02	4.51E-03	3.75E-03	1.89E-02	1.64E-03	1.35e-03
$\hat{\alpha}_1$									
Mean	0.163	0.198	0.202	0.199	0.391	0.217	0.198	0.317	0.299
Bias	0.137	0.102	0.098	0.101	0.091	0.083	0.102	0.017	0.001
SE	0.613	0.215	0.201	0.417	0.099	0.079	0.415	0.051	0.024
MSE	9.99E-02	8.27E-03	7.58E-03	1.55E-02	1.64E-03	1.39E-03	1.04E-02	1.04E-03	1.01e-03
$\hat{\beta}_1$									
Mean	0.656	0.638	0.631	0.614	0.613	0.599	0.614	0.611	5.992e-01
Bias	0.056	0.038	0.031	0.014	0.013	0.001	0.014	0.011	8.23e-04
SE	0.712	0.128	0.097	0.425	0.030	0.013	0.422	0.013	0.006
MSE	8.13E-02	9.86E-03	7.65E-03	3.70E-02	3.75E-03	2.39E-03	2.07E-02	2.14E-03	1.76e-03
$\hat{\omega}$									
Mean	0.288	0.254	0.243	0.240	0.213	0.211	0.240	0.210	0.208
Bias	0.088	0.054	0.043	0.040	0.013	0.011	0.040	0.010	0.008
SE	0.718	0.108	0.087	0.413	0.021	0.020	0.410	0.013	0.009
MSE	6.31E-02	8.55E-03	6.38E-03	2.03E-02	2.48E-03	1.30E-03	1.23E-02	1.28E-03	8.39e-04

7.5 Real Example

Arson is the crime of intentionally burning any type of property where they are not only a home, building, structure, or a place with people inside, but also ones can commit arson by burning either personal property, buildings or land. Nowadays, such criminal widely occurs to show their protests and objections especially for the police. The people who commit arson are called as arsonist. The punishment for the arsonists depending on the degree of property damage and each country has their own law for such people.

For the model, we consider a count arson data from The Forecasting Principles site in the section Crime data. The data represent 144 monthly counts of arson in the 13th police car beat plus in Pittsburgh from January 1990 until December 2001. The data has 61 zeroes that is 42.4% of the series. Motivated from the high number of zeroes in the data, the ZIPINGARCH(p,q) model seems like a good candidates to be considered for this data set. The zero-inflation index for the data is 0.1815 where the index is defined as (see Puig & Valero (2006))

$$zi = 1 + \frac{log\,(p_0)}{\mu_p}$$

where p_0 is the proportion of zeroes and μ_p is the mean. The plot of the data is given in Figure 7.1.

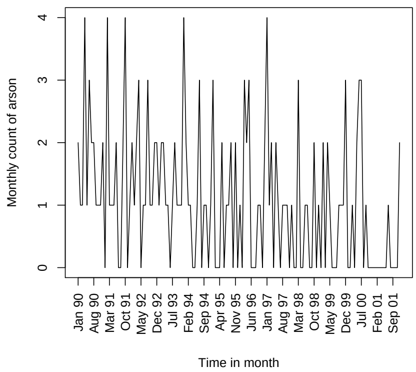


Figure 7.1: The monthly counts of arson in the 13th police car beat plus in Pittsburgh from January 1990 until December 2001

We fit the model and obtain the parameter estimated $\hat{\boldsymbol{\theta}}$ and standard error in parenthesis using MLE, EF and QEF methods as shown in Table 7.12.

Table 7.12: The estimated parameter of ZIPINGARCH$(1,1)$ model

Method	$\hat{\gamma}$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\omega}$	AIC	BIC
MLE	0.279 (0.007)	0.087 (0.014)	0.213 (0.051)	0.214 (0.017)	396.146	405.634
EF	0.281 (0.012)	0.084 (0.023)	0.233 (0.067)	0.208 (0.019)	397.231	408.372
QEF	0.277 (0.007)	0.089 (0.015)	0.217 (0.049)	0.214 (0.017)	395.146	405.355

Zhu (2011) fitted the data into ZIPINARCH(2) model to this data by estimating the parameter using the MLE method. The mean and variance obtained from this fitted model were 1.0369 and 1.3952 respectively while the empirical mean and empirical variance are 1.0417 and 1.3829 accordingly. But, using our fitted model based on QEF estimator as obtained in Table 7.12, the mean and variance are 1.0405 and 1.3821 respectively where they are closer to the empirical mean and empirical variance. From Table (7.12), it is clear that the QEF method gives the smallest AIC and BIC compared with EF and MLE methods. Therefore, we can say that, the estimated parameter obtained from QEF estimator produces the best fitted model compared to other two methods.

Now, we focus only on fitted model by QEF estimator. The parameter estimates $\left(\hat{\gamma}, \hat{\alpha}_1, \hat{\beta}_1\right)$ are positive and the summation of $(1-\hat{\omega})\hat{\alpha}_1 + \hat{\beta}_1$ lies between zero and one, indicating the process is stationary. To examine the model fitting adequacy, we define the Pearson residual for the process as

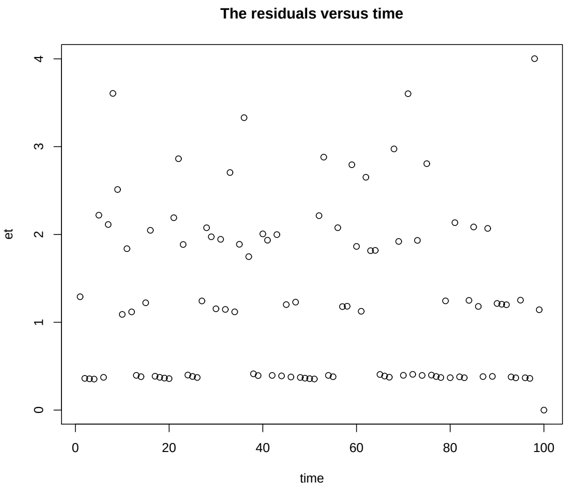
$$z_t = \left(X_t - (1-\hat{\omega})\lambda_t(\hat{\boldsymbol{\theta}})\right) / \sqrt{\lambda_t(\hat{\boldsymbol{\theta}})(1-\hat{\omega})(1+\lambda_t(\hat{\boldsymbol{\theta}}))}.$$

The mean and variance of Pearson residuals are close to zero and unity which are 0.039 and 0.992 respectively indicating adequacy of the model. To examine the serial in the sequence, we perform Ljung-Box (LB) test and the results are given in Table 7.13.

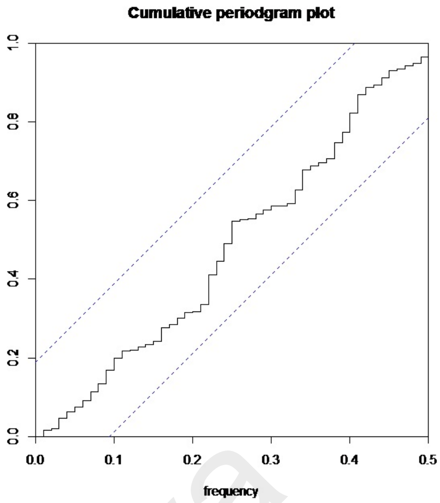
Table 7.13: Diagnostics for ZIPINGARCH(1,1) model

	$\text{LB}_{30}(z_t)$	$\text{LB}_{30}(z_t^2)$
χ^2	25.4	24.22
p-value	0.705	0.762

The results indicate that there is no significant serial correlation in the residuals showing that the ZIPINGARCH(p,q) model fit the data well.



(a)



(b)

Figure 7.2: (a)The Pearson residual plot. (b)The periodagram plot

Figures 7.2(a)and (b), a time series plot of the residuals and a cumulative periodogram plot, confirm that the model fits the data adequately.

7.6 Summary of The Chapter

In this chapter, we discussed the ZIPINGARCH(p,q) model based on the moment properties, QEF estimator and the associated information matrix. Regarding moment properties, we obtained the first four moments using a simpler method (martingale difference) compared with Zhu (2012) who derived only the mean and variance. We showed the superiority of QEF method on the information gain and simulation studies. In addition, we applied the method to a set of count time series data.

CHAPTER 8

CONCLUSION

This chapter provides a short summary of the main results presented in the thesis and discuss important challenges for future research.

8.1 Summary of The Study

This study focused on integer-valued time series models analogue of the generalized autoregressive conditional heteroscedastic, GARCH(p,q), with the conditional distribution following the integer-valued distributions, Poisson, negative binomial and zero-inflated Poisson. The models are known as INGARCH(p,q), NBINGARCH(p,q) and ZIPINGARCH(p,q) models respectively.

Firstly, we derived the moments of the models up to order four using martingale difference. The first and second moment properties are available in literatures but not for the third and fourth moment properties which are skewness and kurtosis. Zhu (2012) stated that the derivation of the unconditional distributions are complicated for these models and needs further investigation. Therefore, we found a general closed form expressions of skewness and kurtosis for these models in closed form expression. The theorem and derivation of general form of skewness and kurtosis of these models were shown in Chapter 3. In Chapter 5, 6 and 7, we obtained the skewness and kurtosis for each model by finding the skewness and kurtosis of each martingale difference of the model. This finding is very important in line with the opinion of Patton (2004) who stated that the knowledge of these higher moments, skewness and kurtosis will make a significant better forecast on assets, both statistically and economically.

Secondly, in time series analysis, one of important interest is to have a good parameter estimation method. Instead of the traditional MLE we consider the optimal QEF method proposed by Liang et al. (2011). The theory on QEF method was discussed in Chapter 4. We also gave some examples for simple integer-valued time series models based on zero-inflated poisson distribution namely basic zero-inflated Poisson and zero-inflated Poisson regression models. This method of parameter estimation have been studied and applied into continuous time series models, for example random coefficient autoregressive (RCA) model by Thavaneswaran et al. (2015) and autoregressive conditional duration (ACD) model (see Ng et al. (2015)). However, there were no study using QEF method on the integer-valued time series model. Therefore, in this research, we studied and discussed the methodology on integer-valued time series models to see its performance compared to other parameter estimations methods. Here, we derived the optimal QEF functions and their information matrices for each process. Theoretically, we compared the information gain by QEF method with each component of estimating functions. We conducted simulation studies to see the performance of QEF method compared to other estimation methods namely MLE and EF methods via R-cran programming. From the results obtained, we saw that the QEF estimators outperform MLE, LS and EF estimators in almost all cases especially for near non-stationary cases. Therefore, we can conclude that, the QEF estimator can be an efficient alternative method for parameter estimation on these models.

Lastly, we applied the methodology on real-world data set. We fitted these models (INGARCH(p,q), NBINGARCH(p,q) and ZIPINGARCH(p,q) models) in Chapter 5, 6 and 7 and estimated parameters using MLE, LS/EF and QEF. In each case, the model fit was examined by the AIC, BIC methods, Pearson residuals and their plots, Ljung-Box tests, and cumulative periodogram tests. In all cases we found that the models considered fitted well to the data sets.

From this study, we concluded that the findings are very important and will be beneficial to the statistical area. This research can adds more knowledge to the area of modeling integer-valued time series data and their properties at the higher order up to order four. The new methods of deriving moments using the martingale differences and other simpler methods will be useful for other statistical researchers.

8.2 Further Research

In this section, we discuss the possibilities for further research on the work presented in this thesis. Some suggestions are given as follows:

1. Extend the theory on zero-inflated negative binomial GARCH (ZINBINGARCH(p,q)) model.

From the research done, we can extend the moment properties to ZINBINGARCH(p,q) model. The moments properties for both conditional and unconditional for this model are quite complicate because of the complex expressions for the factorial cumulant. Due to this reason, it deserve a further investigation to simplify the derivation on the moments for conditional distribution of ZINBINGARCH(p,q) model and then plug in the result obtained into our Theorem 1 and Theorem 2 to have the unconditional moment properties of unconditional distribution of ZINBINGARCH(p,q) model.

2. Apply the QEF estimator on other integer-valued time series models.

In this research, we apply the QEF method on the class of integer-valued GARCH(p,q) models. We show that the estimator performs well compared to MLE and EF methods. Hence, the method should be extended for other class of integer-valued time series models, such as INAR(p) and etc.

3. Extend the use of higher moment properties of INGARCH(p,q) family especially in forecasting the volatility of the models.

Forecasting plays an important part in both linear and nonlinear time series cases. Thavaneswaran et al. (2005) mentioned that the kurtosis of a time series model can be used to forecast the error variance while Patton (2004) highlighted that significant of higher moments, namely skewness and kurtosis, can lead to better forecasts. Therefore, it is possible, we can extend the use of skewness and kurtosis of the models to forecast volatility.

REFERENCES

Abraham, . L. J., B. (2009). *Statistical methods for forecasting* (Vol. 234). New York, USA: John Wiley & Sons.

Allen, D., Ng, K., & Peiris, S. (2013a). The efficient modelling of high frequency transaction data: A new application of estimating functions in financial economics. *Economics Letters*, *120*(1), 117–122.

Allen, D., Ng, K., & Peiris, S. (2013b). Estimating and simulating weibull models of risk or price durations: An application to acd models. *The North American Journal of Economics and Finance*, *25*, 214–225.

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

Aly, E.-E. A., & Bouzar, N. (1994). On some integer-valued autoregressive moving average models. *Journal of multivariate analysis*, *50*(1), 132–151.

Bahadur, R. R. (1958). Examples of inconsistency of maximum likelihood estimates. *Sankhyā: The Indian Journal of Statistics*, *20*(3/4), 207–210.

Baillie, R. T., & Bollerslev, T. (2002). The message in daily exchange rates: A conditional-variance tale. *Journal of Business & Economic Statistics*, *20*(1), 60–68.

Baksh, M. F., Böhning, D., & Lerdsuwansri, R. (2011). An extension of an over-dispersion test for count data. *Computational Statistics & Data Analysis*, *55*(1), 466–474.

Bera, A. K., & Bilias, Y. (2002). The mm, me, ml, el, ef and gmm approaches to estimation: A synthesis. *Journal of Econometrics*, *107*(1), 51–86.

Bera, A. K., Bilias, Y., & Simlai, P. (2006). Estimating functions and equations: An essay on historical developments with applications to econometrics. *Palgrave Handbook of Econometrics*, *1*, 427–476.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Brockwell, P. J., & Davis, R. A. (1991). Time series: Theory and methods. *Springer Series in Statistics. Springer-V erlag,*, *1*(99), 1.

Brockwell, P. J., & Davis, R. A. (2013). *Time series: Theory and methods*. Colorada, USA: Springer Science & Business Media.

Cameron, A. C., & Trivedi, P. K. (2001). *Essentials of count data regression*. Blackwell, Oxford (U.K.): Blackwell Publishing Ltd.

Carvalho, A. X., & Tanner, M. A. (2007). Modelling nonlinear count time series with local mixtures of poisson autoregressions. *Computational statistics & data analysis*, *51*(11), 5266–5294.

Casella, G., & Berger, R. L. (2002). *Statistical inference* (Vol. 2). Florida, USA: Duxbury Pacific Grove, CA.

Chandra, S. A., & Taniguchi, M. (2001). Estimating functions for nonlinear time series models. *Annals of the Institute of Statistical Mathematics*, *53*(1), 125–141.

Crowder, M. (1987). On linear and quadratic estimating functions. *Biometrika*, *74*(3), 591–597.

Davis, R. A., Holan, S. H., Lund, R., & Ravishanker, N. (2016). *Handbook of discrete-valued time series*. New York: CRC Press.

Davis, R. A., & Wu, R. (2009). A negative binomial model for time series of counts. *Biometrika*, *96*(3), 735–749.

Diebold, F. X. (2004). The nobel memorial prize for robert f. engle. *The Scandinavian Journal of Economics*, *106*(2), 165–185.

Durbin, J. (1960). Estimation of parameters in time-series regression models. *Journal of the Royal Statistical Society. Series B (Methodological)*, *22*(1), 139–153.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, *50*(4), 987–1007.

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric reviews*, *5*(1), 1–50.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of finance*, *48*(5), 1749–1778.

Feigin, P. D., Gould, P., Martin, G. M., & Snyder, R. D. (2008). Feasible parameter regions for alternative discrete state space models. *Statistics & Probability Letters*, *78*(17), 2963–2970.

Feller, W. (1943). On a general class of" contagious" distributions. *The Annals of mathematical statistics*, *14*(4), 389–400.

Ferland, R., Latour, A., & Oraichi, D. (2006). Integer-valued garch process. *Journal of Time Series Analysis*, *27*(6), 923–942.

Fisher, R. A. (1935). The fiducial argument in statistical inference. *Annals of eugenics*, *6*(4), 391–398.

Fokianos, K., Rahbek, A., & Tjøstheim, D. (2009). Poisson autoregression. *Journal of the American Statistical Association*, *104*(488), 1430–1439.

Franke, J. (2010). *Weak dependence of functional ingarch processes*. Technische Universität Kaiserslautern, Fachbereich Mathematik.

Ghahramani, M., & Thavaneswaran, A. (2009). Combining estimating functions for volatility. *Journal of Statistical Planning and Inference*, *139*(4), 1449–1461.

Godambe, V. P. (1960). An optimum property of regular maximum likelihood estimation. *The Annals of Mathematical Statistics*, *31*(4), 1208–1211.

Godambe, V. P. (1985). The foundations of finite sample estimation in stochastic processes. *Biometrika*, *72*(2), 419–428.

Godambe, V. P., & Heyde, C. C. (2010). Quasi-likelihood and optimal estimation. In *Selected works of cc heyde* (Vol. 55, pp. 386–399). Springer.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, *42*(1), 27–62.

Harvey, A., & Fernandes, C. (1989). Time series models for insurance claims. *Journal of the Institute of Actuaries*, *116*(3), 513–528.

Houseman, E. A., Coull, B. A., & Shine, J. P. (2006). A nonstationary negative binomial time series with time-dependent covariates: Enterococcus counts in boston harbor. *Journal of the American Statistical Association*, *101*(476), 1365–1376.

Johansson, P. (1996). Speed limitation and motorway casualties: A time series count data regression approach. *Accident Analysis & Prevention*, *28*(1), 73–87.

Johnson, N. L., Kemp, A. W., & Kotz, S. (2005). *Univariate discrete distributions* (Vol. 444). Canada: John Wiley & Sons.

Jung, R. C., Ronning, G., & Tremayne, A. R. (2005). Estimation in conditional first order autoregression with discrete support. *Statistical papers*, *46*(2), 195–224.

Kachour, M. (2009). p-order rounded integer-valued autoregressive (rinar (p)) process. *Journal of Time Series Analysis*, *30*(4), 417-448.

Kedem, B., & Fokianos, K. (2005). *Regression models for time series analysis* (Vol. 488). USA: John Wiley & Sons.

Kharrati-Kopaei, M., & Faghih, H. (2011). Inferences for the inflation parameter in the zip distributions: The method of moments. *Statistical Methodology*, *8*(4), 377–388.

Kimball, B. F., et al. (1946). Sufficient statistical estimation functions for the parameters of the distribution of maximum values. *The Annals of Mathematical Statistics*, *17*(3), 299–309.

Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*, *34*(1), 1–14.

Le Cam, L. (1990). Maximum likelihood: An introduction. *International Statistical Review/Revue Internationale de Statistique*, *58*, 153–171.

Lehmann, E. L., & Casella, G. (1991). *Theory of point estimation*. USA: Wadsworth & Brooks/Cole Advanced Books & Software.

Li, G., Haining, R., Richardson, S., & Best, N. (2014). Space–time variability in burglary risk: A bayesian spatio-temporal modelling approach. *Spatial Statistics*, *9*, 180–191.

Liang, Y., Thavaneswaran, A., & Abraham, B. (2011). Joint estimation using quadratic estimating function. *Journal of Probability and Statistics*, *2011*.

Lim, H. K., Li, W. K., & Philip, L. (2014). Zero-inflated poisson regression mixture model. *Computational Statistics & Data Analysis*, *71*, 151–158.

Lukusa, T. M., Lee, S.-M., & Li, C.-S. (2015). Semiparametric estimation of a zero-inflated poisson regression model with missing covariates. *Metrika*, *79*(4), 457–483.

MacDonald, I. L., & Zucchini, W. (1997). *Hidden markov and other models for discrete-valued time series* (Vol. 110). USA: CRC Press.

McKenzie, E. (1985). Some simple models for discrete variate time series. *Water Resources Bulletins*, *21*, 645-650.

Merkouris, T., et al. (2007). Transform martingale estimating functions. *The Annals of Statistics*, *35*(5), 1975–2000.

Mullahy, J. (1997). Heterogeneity, excess zeros, and the structure of count data models. *Journal of Applied Econometrics*, *12*(3), 337–350.

Nanjundan, G., & Naika, T. R. (2012). Asymptotic comparison of method of moments estimators and maximum likelihood estimators of parameters in zero-inflated poisson model. *Applied mathematics*, *3*(06), 610.

Nastić, A. S., Ristić, M. M., & Bakouch, H. S. (2012). A combined geometric inar (p) model based on negative binomial thinning. *Mathematical and Computer Modelling*, *55*(5), 1665–1672.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, *59*(2), 347–370.

Neyman, J. (1939). On a new class of" contagious" distributions, applicable in entomology and bacteriology. *The Annals of Mathematical Statistics*, *10*(1), 35–57.

Ng, K. H., & Peiris, S. (2013). Modelling high frequency transaction data in financial economics: A comparative study based on simulations. *Economic Computation and Economic Cybernetics Studies and Research*, *47*(2), 189–202.

Ng, K. H., Peiris, S., & Gerlach, R. (2014). Estimation and forecasting with logarithmic autoregressive conditional duration models: A comparative study with an application. *Expert Systems with Applications*, *41*(7), 3323–3332.

Ng, K. H., Shelton, P., Thavaneswaran, A., & Ng, K.-H. (2015). Modelling the risk or price durations in financial markets: Quadratic estimating functions and applications. *Economic Computation & Economic Cybernetics Studies & Research*, *49*(1).

Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, *2*(1), 130–168.

Pedeli, X., & Karlis, D. (2011). A bivariate inar (1) process with application. *Statistical modelling*, *11*(4), 325–349.

Puig, P., & Valero, J. (2006). Count data distributions: Some characterizations with applications. *Journal of the American Statistical Association*, *101*(473), 332–340.

Sellers, K. F., & Raim, A. (2016). A flexible zero-inflated model to address data dispersion. *Computational Statistics & Data Analysis*, *99*, 68–80.

Sentana, E. (1995). Quadratic arch models. *The Review of Economic Studies*, *62*(4), 639–661.

Spiegel, M. (1999). *Schaum's mathematical handbook of formulas and tables*. USA: McGraw-Hill Osborne Media.

Staub, K. E., & Winkelmann, R. (2013). Consistent estimation of zero-inflated count models. *Health economics*, *22*(6), 673–686.

Thavaneswaran, A., & Abraham, B. (1988). Estimation for nonlinear time series models using estimating equations. *Journal of Time Series Analysis*, *9*(1), 99–108.

Thavaneswaran, A., Appadoo, S. S., & Samanta, M. (2005). Random coefficient garch models. *Mathematical and Computer Modelling*, *41*(6), 723–733.

Thavaneswaran, A., Liang, Y., & Frank, J. (2012). Inference for random coefficient volatility models. *Statistics & Probability Letters*, *82*(12), 2086–2090.

Thavaneswaran, A., Ravishanker, N., & Liang, Y. (2015). Generalized duration models and optimal estimation using estimating functions. *Annals of the Institute of Statistical Mathematics*, *67*(1), 129–156.

Tsay, R. S. (2014). *An introduction to analysis of financial data with r*. Chicago, USA: John Wiley & Sons.

Vinod, H. (1997). Using godambe-durbin estimating functions in econometrics. *Lecture Notes-Monograph Series*, *32*, 215–237.

Wang, P. (2001). Markov zero-inflated poisson regression models for a time series of counts with excess zeros. *Journal of Applied Statistics*, *28*(5), 623–632.

Weiß, C. H. (2010). The inarch (1) model for overdispersed time series of counts. *Communications in Statistics-Simulation and Computation*, *39*(6), 1269–1291.

Weiß, C. H. (2013). Integer-valued autoregressive models for counts showing underdispersion. *Journal of Applied Statistics*, *40*(9), 1931–1948.

Wurtz, D., Chalabi, Y., & Luksan, L. (2009). Parameter estimation of arma models with garch/aparch errors an r and splus software implementation. *Journal of Statistical Software, forthcoming*. Retrieved from http://www.jstatsoft.org/ doi: doi=10.1.1.693.9113

Yang, M. (2012). Statistical models for count time series with excess zeros.

Yang, M., Zamba, G. K., Cavanaugh, J. E., & Yang, M. M. (2014). Package zim.

Yau, K. K., Lee, A. H., & Carrivick, P. J. (2004). Modeling zero-inflated count series with application to occupational health. *Computer Methods and Programs in Biomedicine*, *74*(1), 47–52.

Zeger, S. L. (1988). A regression model for time series of counts. *Biometrika*, *75*(4), 621–629.

Zhu, F. (2011). A negative binomial integer-valued garch model. *Journal of Time Series Analysis*, *32*(1), 54–67.

Zhu, F. (2012). Zero-inflated poisson and negative binomial integer-valued garch models. *Journal of Statistical Planning and Inference*, *142*(4), 826–839.