

INTEGER-VALUED ARCH AND GARCH MODELS

by

Kezang Choden C

B.Sc., Sherubtse College, 2004

A Thesis

Submitted in Partial Fulfillment of the Requirements for the
Master of Science Degree

Department of Mathematics
in the Graduate School
Southern Illinois University Carbondale
August 2016

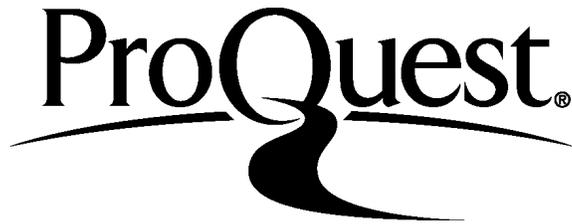
ProQuest Number: 10163458

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10163458

Published by ProQuest LLC (2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

THESIS APPROVAL

INTEGER-VALUED ARCH AND GARCH MODELS

By

Kezang Choden C

A Thesis Submitted in Partial

Fulfillment of the Requirements

for the Degree of

Master of Science

in the field of Mathematics

Approved by:

Dr. S. Yaser Samadi, Chair

Dr. Bhaskar Battacharya

Dr. David J. Olive

Graduate School
Southern Illinois University Carbondale
July 6, 2016

AN ABSTRACT OF THE THESIS OF

KEZANG CHODEN C, for the Master of Science degree in MATHEMATICS, presented on July 6, 2016, at Southern Illinois University Carbondale.

TITLE: INTEGER-VALUED ARCH AND GARCH MODELS

MAJOR PROFESSOR: Dr. S. Yaser Samadi

The models for volatility, autoregressive conditional heteroscedastic (ARCH) and generalized autoregressive conditional heteroscedastic (GARCH) are discussed. Stationarity condition and forecasting for simple ARCH(1) and GARCH(1,1) models are given. The model for discrete time series is proposed to be negative binomial integer-valued GARCH model, which is a generalization of the Poisson INGARCH model. The stationarity conditions and the autocorrelation function are given. For parameter estimation, three methodologies are presented with a focus on maximum likelihood approach. Simulation study on a sample size of 100 and 500 are carried out and the results are presented. An application of the model to a real time series is given indicating that the maximum likelihood estimation procedure performs better than the Poisson and double Poisson model-based methods.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my adviser Dr. S. Yaser Samadi for his support, guidance , understanding and encouragement in writing this paper. Without his help I would not have had the courage to complete my work.

I also like to acknowledge Dr. Bhaskar Bhattacharya and Dr. David J. Olive for being the member of my thesis committee and providing feedback in further improvement of my thesis.

I am also thankful to my fellow graduate students for rendering help in all possible ways and being there for me throughout my academic exploration.

Finally, I would like to thank my family for believing in me and supporting me in all my endeavors.

TABLE OF CONTENTS

Abstract	i
Acknowledgments	ii
List of Tables	v
List of Figures	vi
1 Introduction	1
2 Volatility	3
2.1 Characteristics of volatility	3
3 ARCH Models	9
3.1 The Model	9
3.2 Characteristics of the ARCH Model	12
3.3 Forecasting	14
4 GARCH Models	15
4.1 The Model	15
4.2 Conditions for GARCH model to be stationary	17
4.3 Forecasting	18
5 Negative Binomial Integer-valued GARCH (NBINGARCH) model	20
5.1 Conditions for stationary	21
5.2 Estimation Procedures	35
5.2.1 Yule Walker Estimation	35
5.2.2 CLS Estimation	39
5.2.3 Maximum likelihood Estimation	43
6 Data Analysis	46
6.1 Simulation Study	46
6.2 Real data Analysis	47
6.2.1 INARCH(1) Model	48

6.2.1.1	Negative Binomial	48
6.2.1.2	Poisson	49
6.2.1.3	Double Poisson	50
6.2.2	INGARCH(1,1) Model	52
6.2.2.1	Negative Binomial	54
6.2.2.2	Poisson	55
6.2.2.3	Double Poisson	55
	References	65
	Vita	66

LIST OF TABLES

4.1	EACF of absolute value of simulated GARCH(1,1) model	19
6.1	Simulated Results for NBINGARCH models	59
6.2	AIC and BIC for different values of r with NBINARCH(1) and NBINGARCH(1,1) models.	61
6.3	Parameter estimates with NB, P and DP INARCH(1) and INGARCH(1,1) models.	61

LIST OF FIGURES

2.1	Plot of Daily CREF Stock values.	4
2.2	Plot of Daily CREF Stock returns.	5
2.3	Daily CREF returns. (a) squared returns (b) absolute returns	6
2.4	Daily CREF returns. (a) sample ACF (b) sample PACF	7
2.5	Sample ACF and PACF of absolute daily CREF.	7
2.6	Sample ACF and PACF of square of daily CREF.	8
3.1	Simulated ARCH(1) process	10
3.2	ACF and PACF Simulated ARCH(1) process	12
4.1	Simulated GARCH(1,1) process	16
4.2	Sample ACF and PACF of Simulated GARCH(1,1) process	17
4.3	Sample ACF and PACF of absolute values of Simulated GARCH(1,1) process	18
4.4	Sample ACF and PACF of squared values Simulated GARCH(1,1) process	19
6.1	Polio data. (a) original series (b) ACF (c) PACF	60
6.2	Pearson residuals. (a) ACF (b) PACF (c) Q-Q plot (d) Kernel density	62
6.3	Randomized quantile residuals. (a) ACF (b) PACF (c) Q-Q plot (d) Kernel density	63
6.4	Normalized randomized quantile residual	64

CHAPTER 1

INTRODUCTION

The time series of counts are commonly observed and many authors have underlined the importance of discrete time series in real world. In medicine, especially in epidemiology, we observe such processes. Some of the examples of discrete time series are monthly reported cases of measles in New York City from 1928-1972, number of death in road accidents for past 10 years, monthly number of employed persons in US from 1990-2000, etc. Count time series are non-negative and are often overdispersed. To overcome the overdispersion, overdispersed Poisson and binomial regression models were considered. But since the mean and variance of Poisson distribution are equal, it was not a suitable choice. Negative binomial (NB) distribution whose variance is larger than the mean allows overdispersion. One of the models for overdispersed discrete time series is the integer-valued generalized autoregressive conditional heteroscedastic (INGARCH) model with Poisson deviates (Ferland et.al., 2006) and is defined as

$$\begin{cases} X_t | \mathcal{F}_{t-1} : \mathcal{P}(\lambda_t) : \forall t \in \mathbb{Z} \\ \lambda_t = \alpha_o + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j \lambda_{t-j}, \end{cases} \quad (1.1)$$

where $\alpha_o > 0$, $\alpha_i \geq 0$, and $\beta_j \geq 0$ and \mathcal{F}_{t-1} is the σ -field generated by $\{X_{t-1}, X_{t-2}, \dots\}$.

The conditional mean and conditional variance of Equation (1.1) are same, which brings us to negative binomial INGARCH (NBINGARCH) model that can simultaneously deal with both overdispersion and potential extreme observations. NBINGARCH model is main focus of the study here. But before we learn more about the proposed model, it is important to know what are autoregressive conditional heteroscedastic (ARCH) and generalized autoregressive conditional heteroscedastic(GARCH) models. So we will briefly looking at these two models in the beginning of the study. For more detailed information

on these two models, [1] and [2] can be referred.

In Chapter 2, we will familiarize ourselves with the word volatility, which is the key word in ARCH and GARCH models. We will learn some of its characteristics and look at some plots that will make the concept more clear. In Chapter 3, we will study about ARCH model, its characteristics and method to forecast for simple ARCH(1) model. Also, we will study simulated plots of size 500. Chapter 4 will help us with GARCH model, its conditions for stationarity, forecast for GARCH(1,1) model and simulated plots of GARCH(1,1) of size 500. In Chapter 5, NBINGARCH model will be described including the conditions for the process to be stationary. Three methods of estimation, i.e., the Yule-Walker (YW), conditional least square (CLS) and the maximum likelihood estimation (MLE) for particularly NBINARCH(1), NBINARCH(2) and NBINGARCH(1,1) models are adopted to estimate the value of the parameters. Some results of simulation study of the three models are presented in Table 6.1, which indicates MLE as a better approach to estimate parameters. In Chapter 6, analysis of data is studied by simulating the series of size 100 and 500, which results in a conclusion of Maximum likelihood estimation process as a better method to estimate the parameters. Maximum likelihood estimator methodology is applied to the polio data discussed in Zeger (1988). Parameter estimates with Poisson (P), Double Poisson (DP) and Negative Binomial (NB) for INARCH(1) and INGARCH(1,1) models are presented in Table 6.3 which indicates a better performance of NB model. Also model diagnostics study is done on residuals to test the goodness of the fitted model.

CHAPTER 2

VOLATILITY

Volatility in a time series is referred to as the phenomenon where conditional variance varies over time. The statistical methods for modeling the volatility of an asset return are referred to as conditional heteroscedasticity models. The study of dynamical pattern in the volatility of a time series constitutes the main subject of ARCH and GARCH models. But volatility is a feature that is not easily observed. To identify it, we will be looking at some features that are commonly seen in asset returns in the following section. To define asset return, let $\{p_t\}$ be the time series of daily price of some financial asset. The return on the t th day is given by

$$r_t = \log(p_t) - \log(p_{t-1}).$$

2.1 CHARACTERISTICS OF VOLATILITY

Some of the characteristics commonly seen in asset returns are (see Tsay, 2005)

- There exists volatility clusters. It is a pattern of alternating high and low volatile period of substantial duration.
- Volatility do not diverge to infinity - that is, volatility varies within some fixed period.
- Volatility reacts differently to big increase and big decrease in price, referred to as the leverage effect. Leverage effect is the relationship between stock returns and stock volatility. When volatility rises, expected returns tend to increase, leading to a drop in the stock price. So volatility and stock returns are negatively correlated.

In order to make ourselves clear with the above mentioned characteristics, let us consider the College Retirement Equities Fund (CREF) over a time period from August 26, 2004 to August 15, 2005 from [1]. CREF Stock Fund is an open ended variable annuity equity mutual fund launched by Teachers Insurance and Annuity Association College Retirement Equities Fund (TIAA-CREF). The fund invests in the public equity markets across the globe. The fund primarily invests in value stocks of companies across all market capitalization.

Figure 2.1 displays the time series plot of the CREF data. It shows a generally increasing trend.

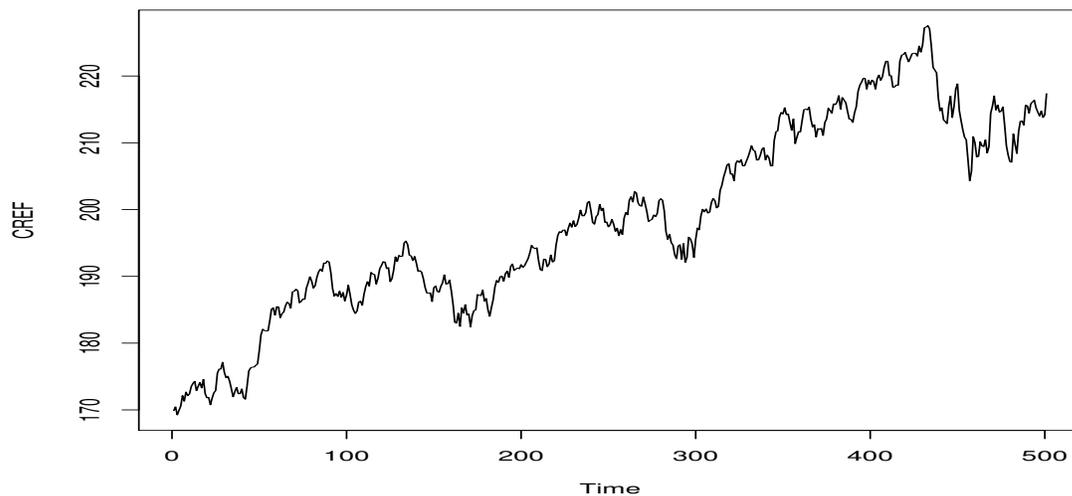


Figure 2.1: Plot of Daily CREF Stock values.

The CREF stock returns computed by $r_t = \log(p_t) - \log(p_{t-1})$ are very small. So the returns are multiplied by 100 and can be interpreted as percentage changes in the price. The multiplication also helps in reducing the numerical errors. The plot in Figure 2.2 shows that the returns were more volatile over some time periods and became very volatile toward the end. This observation is more clear in Figure 2.3, which are plots of absolute and squared returns.

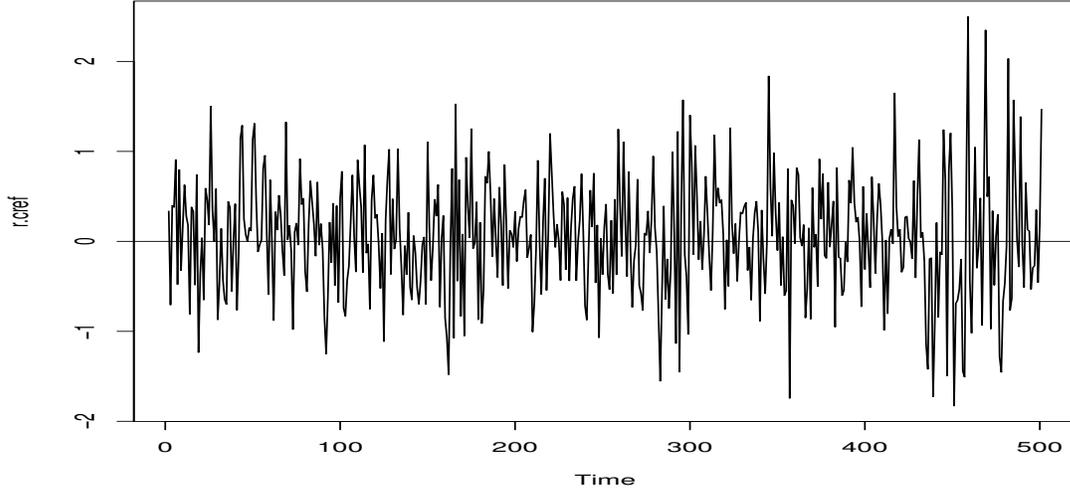


Figure 2.2: Plot of Daily CREF Stock returns.

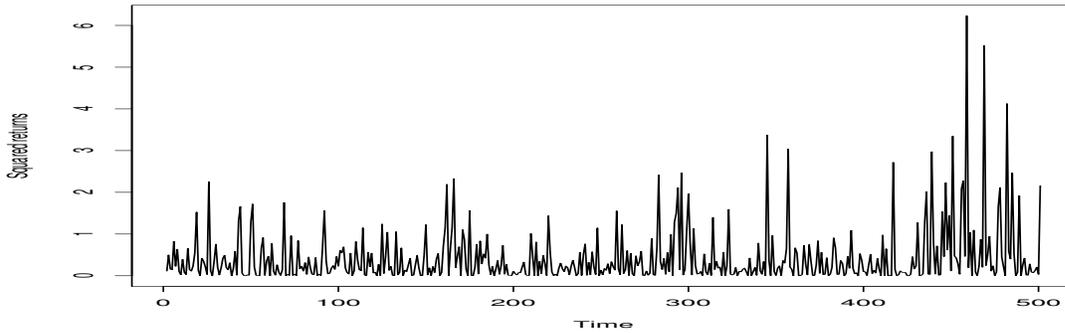
The sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF) of daily CREF returns in the plots of Figure 2.4 suggests that the returns have little serial correlation. Let $\{X_t : t = 0, \pm 1, \pm 2, \dots\}$, be a random process. Then autocorrelation function is defined as

$$\text{corr}(X_t, X_s) = \frac{\text{cov}(X_t, X_s)}{\sqrt{\text{var}(X_t)\text{var}(X_s)}} \quad \text{for } t, s = 0, \pm 1, \pm 2, \dots,$$

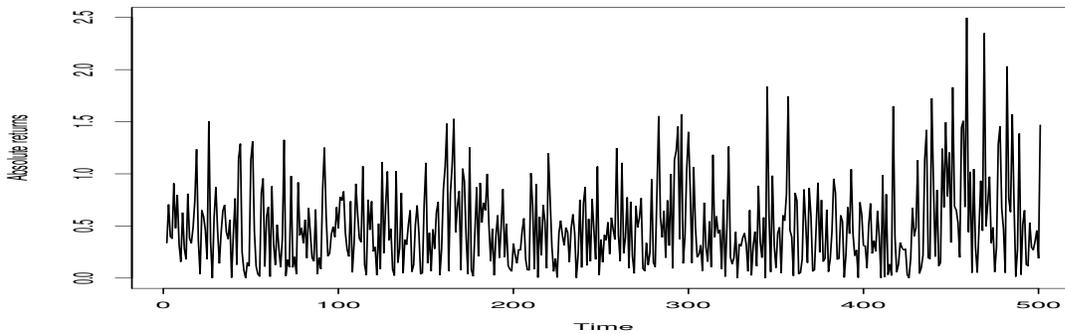
where $\text{cov}(X_t, X_s)$ is the autocovariance function. Partial autocorrelation is a conditional autocorrelation defined as the correlation between the series X_t and X_{t-k} after removing the effect of intermediate series $X_{t-1}, X_{t-2}, \dots, X_{t-k+1}$, where X_{t-k} is the k lag of X_t . So the partial autocorrelation between function X_t and X_{t-k} is written as

$$\begin{aligned} \phi_{kk} &= \text{corr}(X_t, X_{t-k} \mid X_{t-1}, \dots, X_{t-k+1}) \\ &= \frac{\text{cov}(X_t, X_{t-k} \mid X_{t-1}, \dots, X_{t-k+1})}{\sqrt{\text{var}(X_t \mid X_{t-1}, \dots, X_{t-k+1})\text{var}(X_{t-k} \mid X_{t-1}, \dots, X_{t-k+1})}}. \end{aligned}$$

The volatility clustering observed in daily CREF returns indicates that they may not be independently and identically distributed, otherwise the variance is constant over time.



(a)



(b)

Figure 2.3: Daily CREF returns. (a) squared returns (b) absolute returns

The dependency can be checked from the ACF, and PACF of squared and absolute value of the the return series. If series values are independent, then nonlinear instantaneous transformations such as taking logarithms, absolute values, or squaring preserves independence. The sample ACF and PACF of absolute and squared daily CREF returns are exhibited in Figure 2.5 and Figure 2.6. The plots of absolute and squared of daily CREF returns display some significant autocorrelations and hence provide some evidence that the daily CREF returns are not independently and identically distributed.

We can now say that financial asset return series, r_t is serially uncorrelated (even if the correlation exist, it is of small degree), but it is a dependent series.

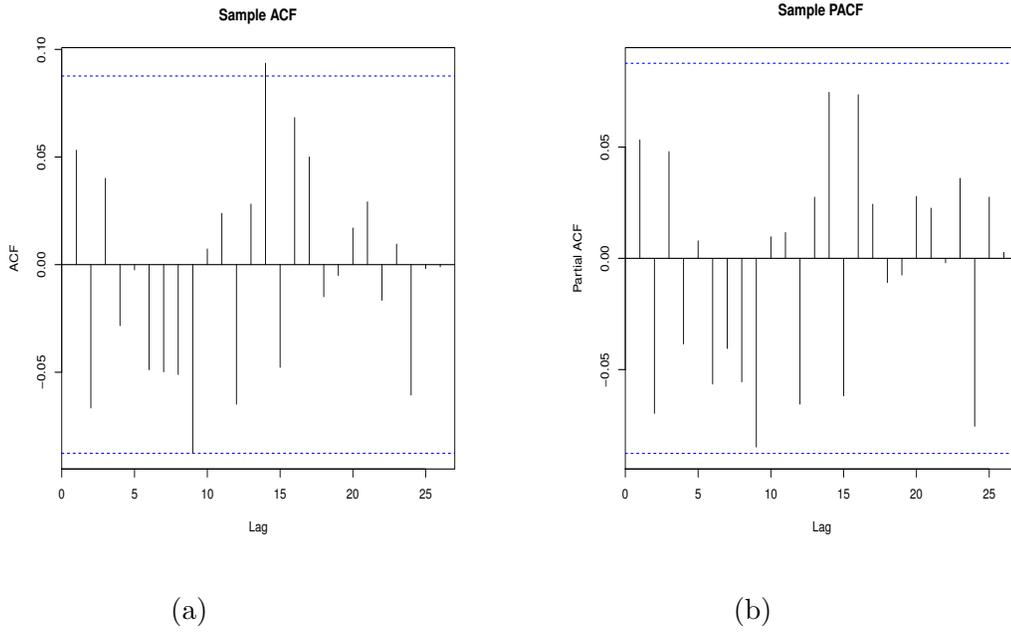


Figure 2.4: Daily CREF returns. (a) sample ACF (b) sample PACF

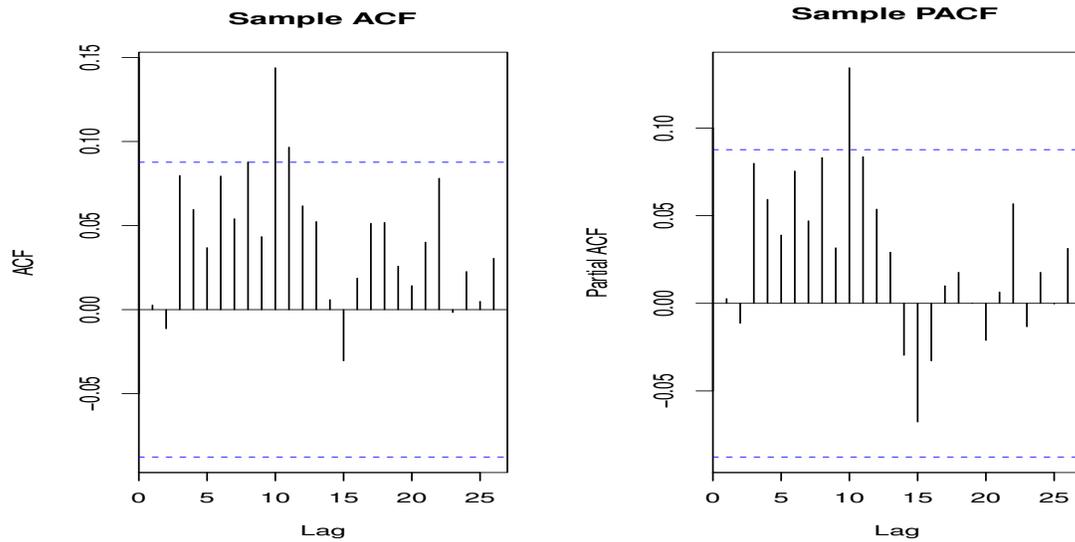


Figure 2.5: Sample ACF and PACF of absolute daily CREF.

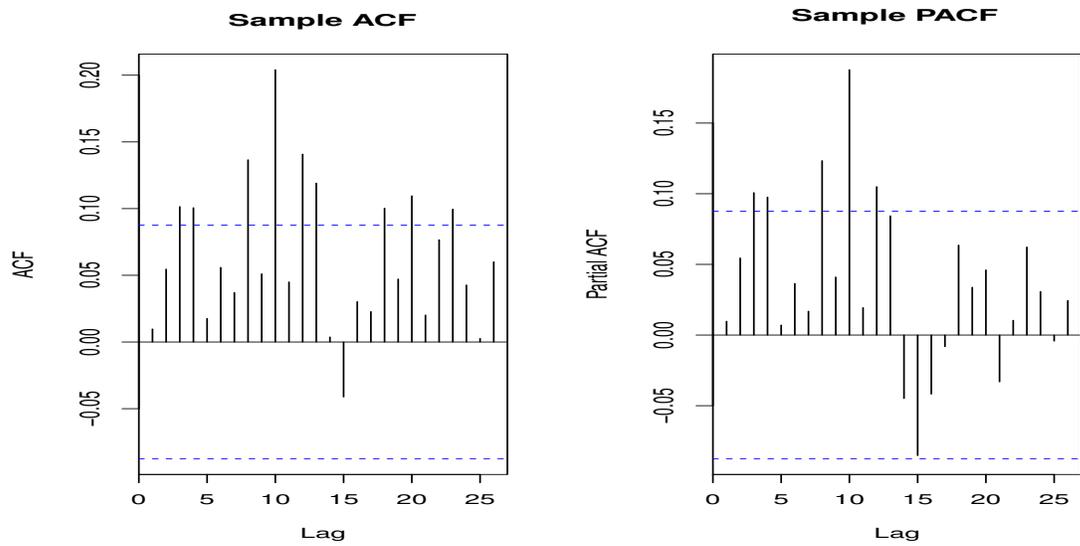


Figure 2.6: Sample ACF and PACF of square of daily CREF.

CHAPTER 3

ARCH MODELS

3.1 THE MODEL

An ARCH model was first proposed by Engle (1982). It is basically designed to model financial time series that displays volatility. As discussed in previous section, the financial asset return series $\{r_t\}$ is serially uncorrelated with zero mean, even when it exhibits volatility clustering. This indicates that conditional variance of $\{r_t\}$ given past returns is not constant. The conditional variance of $\{r_t\}$, will be denoted as $\sigma_{t|t-1}^2$, where the subscript $t - 1$ represents that the conditioning is upon returns through time $t - 1$. The ARCH model is a regression model with the conditional variance as response variable and the past lags of squared return as the covariates. Basically, ARCH(q) model assumes

$$r_t = \sigma_{t|t-1}\varepsilon_t, \quad \sigma_{t|t-1}^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \alpha_2 r_{t-2}^2 + \dots + \alpha_q r_{t-q}^2,$$

where $\{\varepsilon_t\}$ is a sequence of independently and identically distributed random variables with zero mean and unit variance and α_0, α_i are non negative numbers of unknown parameters. More over, ε_t is independent of $r_{t-j}, j = 1, 2, \dots$. So the conditional variance of r_t can be written as follows,

$$\begin{aligned} \text{var}(r_t | r_{t-j}, j = 1, 2, \dots) &= E(r_t^2 | r_{t-j}, j = 1, 2, \dots) - [E(r_t | r_{t-j}, j = 1, 2, \dots)]^2 \\ &= E(\sigma_{t|t-1}^2 \varepsilon_t^2 | r_{t-j}, j = 1, 2, \dots) - [E(\sigma_{t|t-1} \varepsilon_t | r_{t-j}, j = 1, 2, \dots)]^2 \\ &= \sigma_{t|t-1}^2 E(\varepsilon_t^2 | r_{t-j}, j = 1, 2, \dots) - \sigma_{t|t-1}^2 [E(\varepsilon_t | r_{t-j}, j = 1, 2, \dots)]^2 \\ &= \sigma_{t|t-1}^2 E(\varepsilon_t^2) - \sigma_{t|t-1}^2 [E(\varepsilon_t)]^2 - 0 \\ &= \sigma_{t|t-1}^2 [\text{var}(\varepsilon_t) + [E(\varepsilon_t)]^2] \\ &= \sigma_{t|t-1}^2 \end{aligned}$$

Since $[E(r_t | r_{t-j}, j = 1, 2, \dots)] = 0$, it is not wrong to say that

$$\text{var}(r_t | r_{t-j}, j = 1, 2, \dots) = E(r_t^2 | r_{t-j}, j = 1, 2, \dots) = \sigma_{t|t-1}^2$$

Example 3.1. ARCH(1) model is defined as

$$\begin{aligned} r_t &= \sigma_t \varepsilon_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2. \end{aligned} \tag{3.1}$$

If $\alpha_0 = 0.04$ and $\alpha_1 = 0.8$, the above model becomes

$$\sigma_t^2 = 0.04 + 0.8r_{t-1}^2.$$

Figure 3.1 shows the time series plot of a simulated series of size 500 from an ARCH(1) model with $\alpha_0 = 0.04$ and $\alpha_1 = 0.8$.

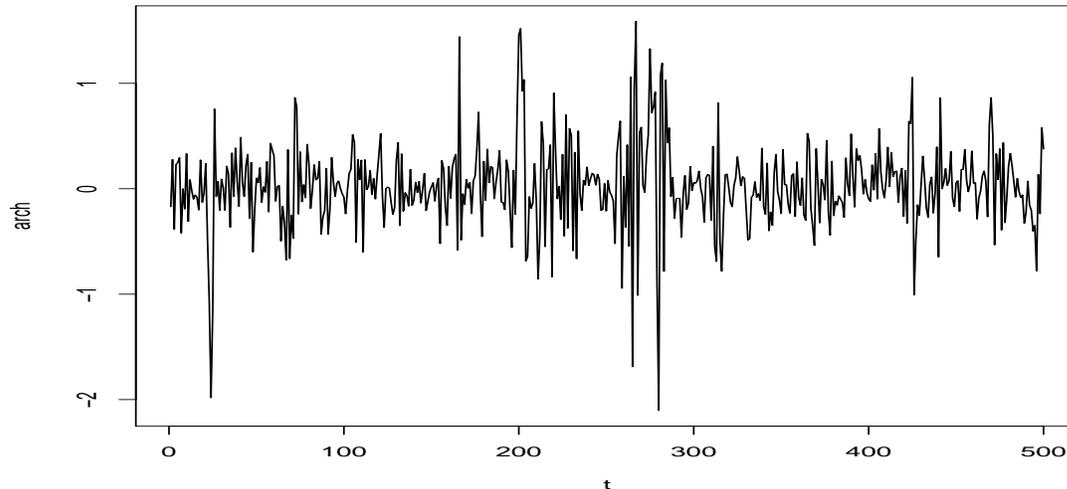


Figure 3.1: Simulated ARCH(1) process

As we can see that there exists volatility clustering and especially in the middle, the series is very volatile. ARCH(1) process is a good example of a white noise that has nonconstant conditional variance. In ARCH(1) model, r_t is an the ARCH(1) process. To

confirm r_t is a white noise, we will find the mean and variance of r_t .

$$\begin{aligned} E(r_t | r_{t-1}, r_{t-2}, \dots) &= E(\sigma_{t|t-1}\varepsilon_t | r_{t-1}, r_{t-2}, \dots) \\ &= \sigma_{t|t-1}E(\varepsilon_t | r_{t-1}, r_{t-2}, \dots) \\ &= 0. \end{aligned}$$

So,

$$E(r_t) = E(E(r_t | r_{t-1}, r_{t-2}, \dots)) = E(0) = 0,$$

and

$$\begin{aligned} \text{var}(r_t | r_{t-1}, r_{t-2}, \dots) &= E(r_t | r_{t-1}, r_{t-2}, \dots)^2 - [E(r_t | r_{t-1}, r_{t-2}, \dots)]^2 \\ &= E(\sigma_{t|t-1}^2\varepsilon_t^2 | r_{t-1}, r_{t-2}, \dots) - 0 \\ &= \sigma_{t|t-1}^2 E(\varepsilon_t^2 | r_{t-1}, r_{t-2}, \dots) \\ &= \sigma_{t|t-1}^2(1) \\ &= \sigma_{t|t-1}^2. \end{aligned}$$

So,

$$\begin{aligned} \text{var}(r_t) &= E[\text{var}(r_t | r_{t-1}, r_{t-2}, \dots)] + \text{var}[E(r_t | r_{t-1}, r_{t-2}, \dots)] \\ &= E(\sigma_{t|t-1}^2) + \text{var}(0) \\ &= \sigma_{t|t-1}^2. \end{aligned}$$

Hence ARCH(1) process is a white noise with mean 0 and finite variance. Now, to show r_t is serially uncorrelated, let us find the covariance of r_t . Multiplying both side of Equation (3.1) by r_{t-1} and finding expectation, we get

$$\begin{aligned} E(r_t r_{t-1}) &= E[E(r_t r_{t-1} | r_{t-1}, r_{t-2}, \dots)] \\ &= E[r_{t-1} E(r_t | r_{t-1}, r_{t-2}, \dots)] \\ &= E[r_{t-1}(0)] = 0. \end{aligned}$$

Covariance between r_t and r_{t-1} is

$$\text{cov}(r_t, r_{t-1}) = E(r_t r_{t-1}) - E(r_t)E(r_{t-1}) = 0.$$

So, the correlation between r_t and r_{t-1} given by $\text{corr}(r_t, r_{t-1}) = \frac{\text{cov}(r_t, r_{t-1})}{\sqrt{\text{var}(r_t)\text{var}(r_{t-1})}} = 0$. We can similarly show that $\text{cov}(r_t, r_{t-j}) = 0, \quad \forall j \geq 1$, and $\text{corr}(r_t, r_{t-j}) = 0, \quad \forall j \geq 1$. The claim that r_t is serially uncorrelated is also supported by the plot in Figure 3.2 of the simulated series. Hence ARCH(1) process is a white noise with mean 0 and finite variance.

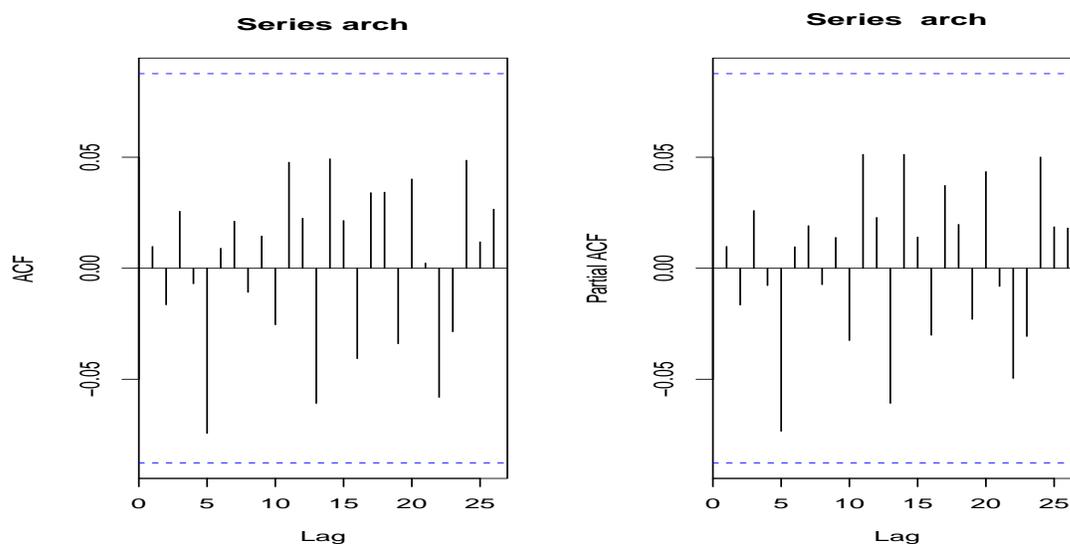


Figure 3.2: ACF and PACF Simulated ARCH(1) process

ARCH model is also applicable in other financial time series like changes in dollar exchange rate over a period of time, the amount that investments or stocks increase (or decrease) per time period, etc.

3.2 CHARACTERISTICS OF THE ARCH MODEL

The ARCH model is identical to a regression model. Since ARCH model models volatility and volatility is not easily observable, it provides a draw back in the use of the

model. For example, it is not easy to identify the regression relationship graphically. So, it is relevant to introduce some observable variables that will simplify the conditional variance. Let $\eta_t = r_t^2 - \sigma_{t|t-1}^2$, where η_t is a serially uncorrelated series with zero mean and also it is uncorrelated with past returns. That is

$$\begin{aligned} E(\eta_t) &= E(r_t^2) - E(\sigma_{t|t-1}^2) \\ &= E[E(r_t^2 | r_{t-j}, j = 1, 2, \dots)] - \sigma_{t|t-1}^2 \\ &= \sigma_{t|t-1}^2 - \sigma_{t|t-1}^2 = 0. \end{aligned}$$

Substituting $\sigma_{t|t-1}^2 = r_t^2 - \eta_t$ in Equation (3.1), we get

$$r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \eta_t \quad (3.2)$$

Since the squared return series must be non-negative, the value of parameters has to be non-negative. If the return series is stationary with variance σ^2 , then applying expectation on both sides of Equation (3.2), we have

$$\begin{aligned} E(r_t^2) &= \alpha_0 + \alpha_1 E(r_{t-1}^2) \\ E[E(r_t^2 | r_{t-j}, j = 1, 2, \dots)] &= \alpha_0 + \alpha_1 E[E(r_{t-1}^2 | r_{t-j}, j = 2, \dots)] \\ \sigma_{t|t-1}^2 &= \alpha_0 + \alpha_1 \sigma_{t-1|t-2}^2 \\ \sigma^2 &= \alpha_0 + \alpha_1 \sigma^2 \\ \sigma^2 &= \frac{\alpha_0}{1 - \alpha_1}, \end{aligned} \quad (3.3)$$

and hence $0 \leq \alpha_1 < 1$, is necessary and sufficient condition for stationarity of the ARCH(1) model.

3.3 FORECASTING

The ARCH model is used in predicting the future conditional variances. The h-step ahead forecast of conditional variance $\sigma_{t+h|t}^2$ is given by

$$\sigma_{t+h|t}^2 = E(r_{t+h}^2 | r_t, r_{t-1}, \dots). \quad (3.4)$$

In ARCH(1), if h=1: $\sigma_{t+1|t}^2 = \alpha_0 + \alpha_1 r_t^2 = (1 - \alpha_1)\sigma^2 + \alpha_1 r_t^2$.

We can use iterated expectation formula in Equation (3.4) to get an expression for h-step conditional variance as the following.

$$\begin{aligned} \sigma_{t+h|t}^2 &= E[E(r_{t+h}^2 | r_t, r_{t-1}, \dots) | r_{t+h-1}, r_{t+h-2}, \dots] \\ &= E[E(\sigma_{t+h|t+h-1}^2 \varepsilon_{t+h}^2 | r_{t+h-1}, r_{t+h-2}, \dots) | r_t, r_{t-1}, \dots] \\ &= E[\sigma_{t+h|t+h-1}^2 E(\varepsilon_{t+h}^2) | r_t, r_{t-1}, \dots] \\ &= E(\sigma_{t+h|t+h-1}^2 | r_t, r_{t-1}, \dots) \\ &= E(\alpha_0 + \alpha_1 r_{t+h-1}^2 | r_t, r_{t-1}, \dots) \\ &= \alpha_0 + \alpha_1 E(r_{t+h-1}^2 | r_t, r_{t-1}, \dots) \\ &= \alpha_0 + \alpha_1 \sigma_{t+h-1|t}^2. \end{aligned}$$

CHAPTER 4

GARCH MODELS

4.1 THE MODEL

Another approach to model volatility was proposed by Bollerslev(1986) and Taylor(1986), which introduces p lags of the conditional variance in the ARCH(q) model. It is called the GARCH(p, q) model, where p is the GARCH order. The GARCH(p, q) model is defined as

$$\sigma_{t|t-1}^2 = \alpha_0 + \beta_1 \sigma_{t-1|t-2}^2 + \dots + \beta_p \sigma_{t-p|t-p-1}^2 + \alpha_1 r_{t-1}^2 + \dots + \alpha_q r_{t-q}^2. \quad (4.1)$$

Since the conditional variance is nonnegative, the values of parameters in GARCH model is restricted to be nonnegative.

Example 4.1. Example 1. GARCH(1,1) is defined as

$$\sigma_{t|t-1}^2 = \alpha_0 + \beta_1 \sigma_{t-1|t-2}^2 + \alpha_1 r_{t-1}^2.$$

Figure 4.1 is the time series plot of simulated *GARCH*(1, 1) with the parameter values $\alpha_0 = 0.03$, $\alpha_1 = 0.05$, and $\beta_1 = 0.7$ of size 500.

It is evident in Figure 4.1 that there exists a volatility clustering. ACF and PACF of simulated process must be serially uncorrelated but dependent. To further confirm that it is a volatility model, we can look at ACF and PACF of absolute and squared simulated data which are exhibited in Figure 4.3 and Figure 4.4.

The fact that GARCH is a model for volatility, encourages us to again introduce the definition $\eta_t = r_t^2 - \sigma_{t|t-1}^2$, which is serially uncorrelated sequence and is uncorrelated with

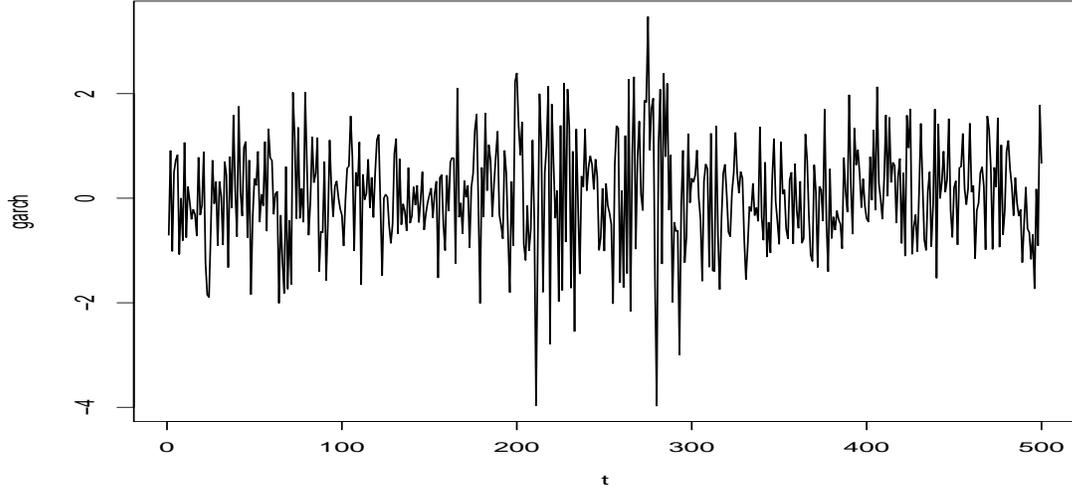


Figure 4.1: Simulated GARCH(1,1) process

past squared returns. Substituting $\sigma_{t|t-1}^2 = a_t^2 - \eta_t$ in Equation (4.1), we get

$$\begin{aligned}
 r_t^2 - \eta_t &= \alpha_0 + \beta_1(r_{t-1}^2 - \eta_{t-1}) + \dots + \beta_p(r_{t-p}^2 - \eta_{t-p}) + \alpha_1 r_{t-1}^2 + \dots + \alpha_q r_{t-q}^2 \\
 r_t^2 &= \alpha_0 + (\beta_1 + \alpha_1)r_{t-1}^2 + \dots + (\beta_{\max(p,q)} + \alpha_{\max(p,q)})r_{t-\max(p,q)}^2 + \eta_t - \beta_1\eta_{t-1} - \dots - \beta_p\eta_{t-p}.
 \end{aligned} \tag{4.2}$$

where $\beta_k = 0$, for all integers $k > p$ and $\alpha_k = 0$ for $k > q$. η_t is a serially uncorrelated series with mean zero.

The model for the squared returns resembles ARMA($\max(p, q), p$) model. To identify the values of ($\max(p, q), p$) we can use identification technique of ARMA model. Extended ACF (EACF) table of squared values can be used to identify the order of the ARMA model and hence the order of the GARCH model. More appropriately EACF of absolute returns can also be used to identify the order of GARCH model.

Table 4.1 displays the EACF for the absolute return of simulated GARCH(1,1) model with $\alpha_0 = 0.03$, $\alpha_1 = 0.05$, and $\beta_1 = 0.7$ of size 500.

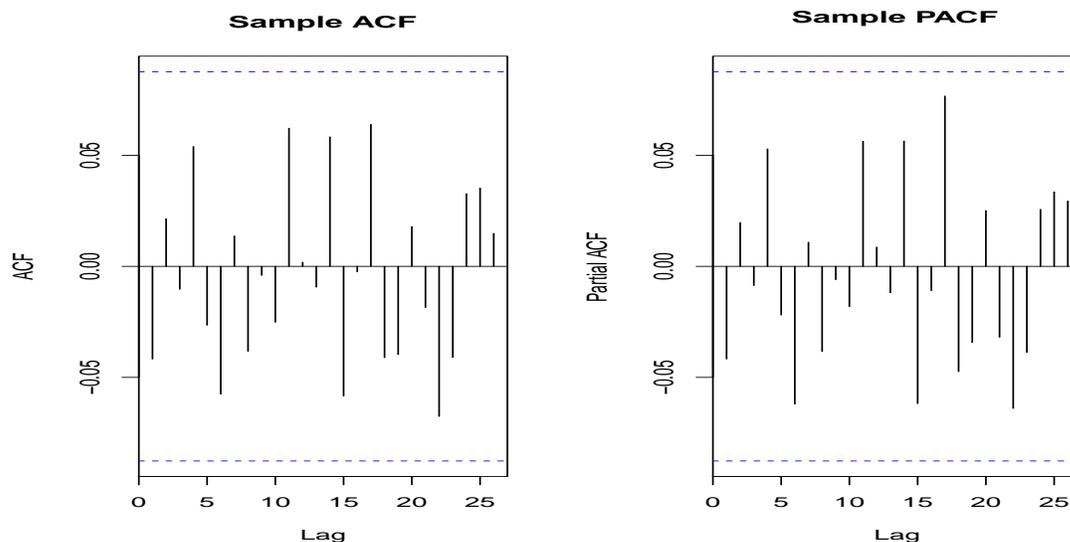


Figure 4.2: Sample ACF and PACF of Simulated GARCH(1,1) process

From Table 4.1, the suggested model is ARMA(1,1), and hence GARCH(1,1). It also suggests ARMA(2,2), but it's always better to adopt the model with less parameters.

4.2 CONDITIONS FOR GARCH MODEL TO BE STATIONARY

If the return series is stationary with variance σ^2 , then taking the expectation of (4.2) gives us

$$E(r_t^2) = \alpha_0 + (\beta_1 + \alpha_1)E(r_t^2) + \dots + (\beta_{\max(p,q)} + \alpha_{\max(p,q)})E(r_t^2)$$

$$\sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i)},$$

which is finite if $\sum_{i=1}^{\max(p,q)} (\beta_i + \alpha_i) < 1$, and is a necessary and sufficient condition for weak stationary of a GARCH(p, q) model. If $\alpha_1 + \beta_1 = 1$, then the GARCH(1,1) is no longer stationary and is called an Integrated GARCH(1,1) model.

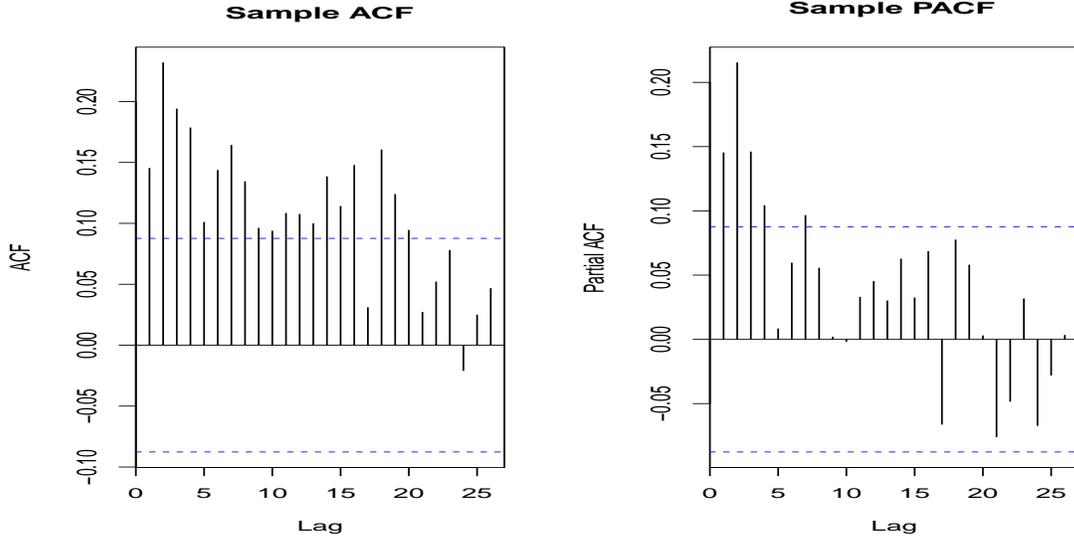


Figure 4.3: Sample ACF and PACF of absolute values of Simulated GARCH(1,1) process

4.3 FORECASTING

For simplicity, let us consider $p=q$. The forecast for h -step ahead conditional variance $\sigma_{t+h|t}^2$ is given by

$$\begin{aligned}\sigma_{t+h|t}^2 &= \alpha_0 + \alpha_1 r_{t+h-1}^2 + \cdots + \alpha_p r_{t+h-p}^2 + \beta_1 \sigma_{t-1|t-2}^2 + \cdots + \beta_p \sigma_{t+h-p|t+h-p-1}^2 \\ &= \alpha_0 + \sum_{i=1}^p \alpha_i r_{t+h-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t+h-i|t+h-i-1}^2.\end{aligned}$$

Consider the GARCH(1,1) model

$$\sigma_{t|t-1}^2 = \alpha_0 + \beta_1 \sigma_{t-1|t-2}^2 + \alpha_1 r_{t-1}^2.$$

Suppose there are n observations r_1, \dots, r_n . Let $\sigma_{1|0}^2 = \sigma^2 = \frac{\alpha_0}{1-\alpha_1-\beta_1}$, which is a condition for stationarity of GARCH(1,1). Then the GARCH(1,1) becomes

$$\sigma_{t|t-1}^2 = (1 - \alpha_1 - \beta_1)\sigma^2 + \beta_1 \sigma_{t-1|t-2}^2 + \alpha_1 r_{t-1}^2.$$

The 1-step ahead forecast $\sigma_{t+1|t}^2$ is written as

$$\sigma_{t+1|t}^2 = (1 - \alpha_1 - \beta_1)\sigma^2 + \beta_1 \sigma_{t|t-1}^2 + \alpha_1 r_t^2.$$

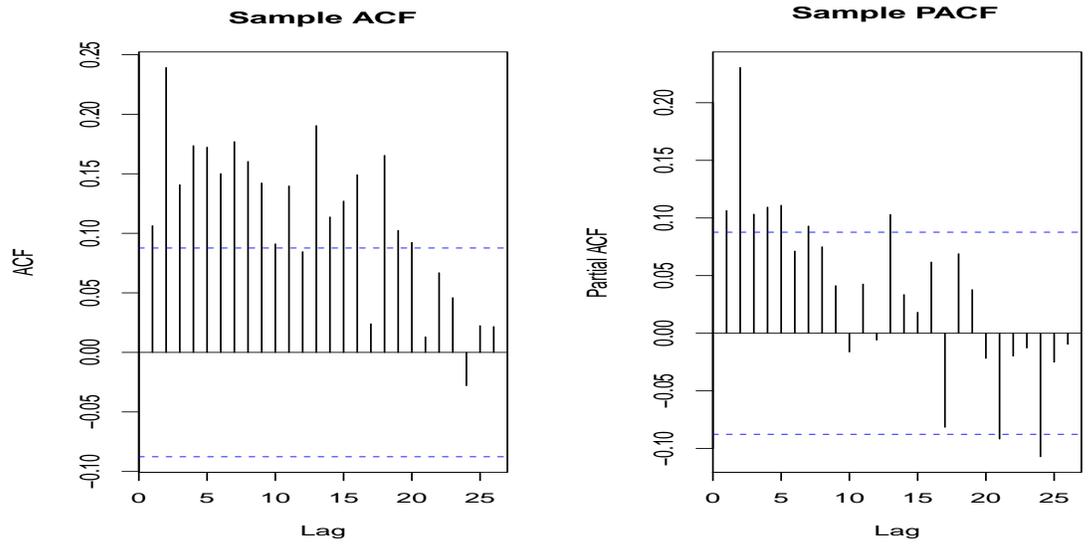


Figure 4.4: Sample ACF and PACF of squared values Simulated GARCH(1,1) process

Table 4.1: EACF of absolute value of simulated GARCH(1,1) model

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	o	o	o	o	o	o	o	o	o	o	o	o	o
3	x	o	x	o	o	o	o	o	o	o	o	o	o	o
4	o	x	x	x	o	o	o	o	o	o	o	o	o	o
5	x	x	x	x	x	o	o	o	o	o	o	o	o	o
6	x	x	x	o	x	o	o	o	o	o	o	o	o	o
7	x	x	x	o	o	o	o	o	o	o	o	o	o	o

CHAPTER 5

NEGATIVE BINOMIAL INTEGER-VALUED GARCH (NBINGARCH) MODEL

Let $\{X_t\}$ be a time series of discrete values. let \mathcal{F}_{t-1} be the σ -field generated by $\{X_{t-1}, X_{t-2}, \dots\}$ and conditional distribution of X_t be NB. That is,

$$X_t | \mathcal{F}_{t-1} : \mathcal{NB}(r, p_t), \quad (5.1)$$

where $r > 0$ and $0 \leq p_t \leq 1$. Note that p_t here represents the probability, not the price at time t as indicated in volatility section. Then p_t satisfies the model

$$\frac{1 - p_t}{p_t} = \lambda_t = \alpha_o + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j \lambda_{t-j}, \quad (5.2)$$

where $\alpha_o > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $p \geq 1$ and $q \geq 0$. The model in Equation (5.2) is denoted as NBINGARCH(p, q). Since the conditional distribution of X_t is NB, the probability mass function is given by

$$P(X_t = x_t | \mathcal{F}_{t-1}) = \binom{x_t + r - 1}{r - 1} p_t^r (1 - p_t)^{x_t},$$

where $x_t=0,1,2\dots$ with

$$p_t = \frac{1}{\lambda_t + 1}, \quad q_t = 1 - p_t = \frac{\lambda_t}{\lambda_t + 1}.$$

We can calculate conditional mean and variance of X_t as the following.

$$E(X_t | \mathcal{F}_{t-1}) = \frac{n(1 - p_t)}{p_t} = n\lambda_t, \quad Var(X_t | \mathcal{F}_{t-1}) = \frac{n(1 - p_t)}{p_t^2} = n\lambda_t(\lambda_t + 1). \quad (5.3)$$

Clearly, conditional variance is $(1 + \lambda_t)$ times that of conditional expectation. More over,

$$\begin{aligned} E(X_t) &= E(E(X_t | \mathcal{F}_{t-1})) \\ &= E(r\lambda_t) = rE(\lambda_t), \end{aligned} \quad (5.4)$$

$$\begin{aligned}
\text{Var}(X_t) &= \text{Var}(E(X_t | \mathcal{F}_{t-1})) + E(\text{Var}(X_t | \mathcal{F}_{t-1})) \\
&= \text{Var}(r\lambda_t) + E(r\lambda_t(\lambda_t + 1)) \\
&= r^2\text{Var}(\lambda_t) + rE(\lambda_t)^2 + rE(\lambda_t) \\
&= r^2\text{Var}(\lambda_t) + r\text{Var}(\lambda_t) + r(E(\lambda_t)^2 + E(\lambda_t)) \\
&= rE(\lambda_t) + r(E(\lambda_t))^2 + (r + r^2)\text{Var}(\lambda_t) \\
&> rE(\lambda_t) = E(X_t).
\end{aligned} \tag{5.5}$$

From the Equation (5.4) and Equation (5.5) , it is clear that $\text{Var}(X_t) > E(X_t)$. Since the expected value and variance of the process $\{X_t\}$ is not same unlike Poisson, Negative Binomial is a better distribution to be used for overdispersion.

5.1 CONDITIONS FOR STATIONARY

For simplicity, we assume $p \geq q$. The following theorems provide first and second order stationary conditions for models in Equation (5.1) and Equation (5.2).

Theorem 5.1. *A necessary and sufficient condition for NBINGARCH(p, q) to be first order stationary is that all roots of the equation*

$$1 - \sum_{i=1}^q (r\alpha_i\beta_i)Z^{-i} - \sum_{i=q+1}^p r\alpha_i Z^{-i} = 0 \tag{5.6}$$

lie inside the unit circle. (see Zhu, 2011)

Proof. let $\mu_t = E(X_t)$. Then

$$\begin{aligned}
E(X_t) &= E(E(X_t | \mathcal{F}_{t-1})) = rE(\lambda_t) \\
&= rE\left(\lambda_t = \alpha_o + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{i=1}^q \beta_i \lambda_{t-i}\right) \\
&= r\alpha_o + \sum_{i=1}^p r\alpha_i E(X_{t-i}) + \sum_{i=1}^q r\beta_i E(\lambda_{t-i}) \\
&= r\alpha_o + \sum_{i=1}^p r\alpha_i \mu_{t-i} + \sum_{i=1}^q \beta_i \mu_{t-i}.
\end{aligned} \tag{5.7}$$

From Goldberg(1958), the necessary and sufficient condition for a non-homogeneous difference equation to have a solution independent of t is that the roots of Equation (5.6) is inside the unit circle.

□

Remark 1. If the process $\{X_t\}$ follows a *NBINGARCH*(p, q) model and is a first order stationary, then

$$\begin{aligned}
E(X_t) = \mu &= r\alpha_o + \sum_{i=1}^p r\alpha_i \mu + \sum_{i=1}^q \beta_i \mu \\
&= \frac{r\alpha_o}{1 - \sum_{i=1}^p r\alpha_i - \sum_{i=1}^q \beta_i}.
\end{aligned}$$

For simplicity, in Theorem 5.2, we assume $q=0$ for second order stationarity condition.

Theorem 5.2. *Suppose the process $\{X_t\}$ is first order stationary. Then a necessary and sufficient condition for the process to be second order stationary is that all roots of $1 - C_1 z^{-1} - \dots - C_p z^{-p} = 0$ lie inside the unit circle, where for $u, l=1 \dots p-1$,*

$$C_u = (r + r^2) \left(\alpha_u^2 - \sum_{v=1}^{p-1} \sum_{|i-j|=v} \alpha_i \alpha_j b_{vu} \beta_{u0} \right), \quad C_p = (r + r^2) \alpha_p^2,$$

$$\beta_{l0} = r\alpha_l, \quad \beta_{lu} = r \sum_{|i-l|=l} \alpha_i - 1 \quad \text{and} \quad \beta_{lu} = r \sum_{|i-l|=u} \alpha_i, \quad u \neq l,$$

where B and B^{-1} are $(p-1) \times (p-1)$ matrices such that $B = (\beta_{ij})_{i,j}^{p-1}$ and $B^{-1} = (b_{ij})_{i,j}^{p-1}$. (See Zhu, 2011)

Proof. Let $\gamma_{it} = E(X_t X_{t-i})$ for $i=0,1,2,\dots,p$ and C be a constant independent of t . Suppose the process is second-order stationary, then let $\gamma_{st} = \gamma_{s,t-i}$ for $i=0,1,\dots,p$. To show that all roots lie inside the unit circle, consider the conditional second moment from which we get obtain second moment, which is a key for second-order stationary condition.

$$\begin{aligned}
E(X_t^2 | \mathcal{F}_{t-1}) &= \text{Var}(X_t | \mathcal{F}_{t-1}) + [E(X_t | \mathcal{F}_{t-1})]^2 \\
&= r\lambda_t(1 + \lambda_t) + (r\lambda_t)^2 \\
&= r\lambda_t + (r + r^2)\lambda_t^2 \\
&= r \left(\alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i} \right) + (r + r^2) \left(\alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i} \right)^2 \\
&= r\alpha_0 + r \sum_{i=1}^p \alpha_i X_{t-i} + (r + r^2) \left[\alpha_0^2 + 2\alpha_0 \sum_{i=1}^p \alpha_i X_{t-i} + \left(\sum_{i=1}^p \alpha_i X_{t-i} \right)^2 \right] \\
&= r\alpha_0 + (r + r^2)\alpha_0^2 + [r + 2\alpha_0(r + r^2)] \sum_{i=1}^p \alpha_i X_{t-i} + (r + r^2) \left(\sum_{i=1}^p \alpha_i X_{t-i} \right)^2 \\
&= r\alpha_0 + (r + r^2)\alpha_0^2 + [r + 2\alpha_0(r + r^2)] \sum_{i=1}^p \alpha_i X_{t-i} + \\
&\quad (r + r^2) \left(\sum_{i=1}^p \alpha_i^2 X_{t-i}^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^p \alpha_i \alpha_j X_{t-i} X_{t-j} \right)
\end{aligned}$$

For $l = 1, 2, \dots, p-1$, let the expected value of $(X_t X_{t-l})$ be denoted by γ_{lt} . So,

$$\begin{aligned}
\gamma_{lt} &= E[E(X_t X_{t-l} \mid \mathcal{F}_{t-1})] \\
&= E[E(X_t \mid \mathcal{F}_{t-1}) X_{t-l}] \\
&= E[(r\gamma_t) X_{t-l}] \\
&= rE \left[\left(\alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i} \right) X_{t-l} \right] \\
&= r\alpha_0\mu + rE \left(\sum_{i=l}^p \alpha_i X_{t-i} X_{t-l} + \sum_{\substack{i=1 \\ i \neq l}}^p \alpha_i X_{t-i} X_{t-l} \right) \\
&= r\alpha_0\mu + rE(\alpha_l X_{t-l} X_{t-l}) + \sum_{\substack{i=1 \\ i \neq l}}^p \alpha_i E(X_{t-i} X_{t-l}) \\
&= r\alpha_0\mu + r\alpha_l \gamma_{0,t-l} + r \sum_{\substack{i=1 \\ i \neq l}}^p \alpha_i \gamma_{|i-l|,t-i} \\
&= r\alpha_0\mu + r\alpha_l \gamma_{0,t-l} + r \sum_{\substack{i=1 \\ i \neq l}}^p \alpha_i \gamma_{|i-l|,t} \\
&= r\alpha_0\mu + r\alpha_l \gamma_{0,t-l} + r \left(\sum_{|i-l|=1} \alpha_i \gamma_{1t} + \dots + \sum_{|i-l|=l} \alpha_i \gamma_{lt} + \dots + \sum_{|i-l|=p-1} \alpha_i \gamma_{p-1,t} \right),
\end{aligned}$$

where $\gamma_{s,t-i}$ are being substituted by $\gamma_{s,t}$ for $i = 1, 2, \dots, p-1$ and for $l = 1, 2, \dots, p-1$,

$$r\alpha_0\mu + \beta_{l0}\gamma_{0,t-l} + \sum_{u=1}^{p-1} \beta_{lu}\gamma_{ut} = 0.$$

So,

$$\begin{aligned}
\sum_{u=1}^{p-1} \beta_{lu}\gamma_{ut} &= -(r\alpha_0\mu + \beta_{l0}\gamma_{0,t-l}) \\
&= -(r\alpha_0\mu + \beta_{10}\gamma_{0,t-1}, \dots, r\alpha_0\mu + \beta_{p-1,0}\gamma_{0,t-p+1}).
\end{aligned}$$

$$\text{when } l = 1, \quad \beta_{11}\gamma_{1t} + \beta_{12}\gamma_{2t} + \cdots + \beta_{1,p-1}\gamma_{p-1,t} = -(r\alpha_0\mu + \beta_{10}\gamma_{0,t-1})$$

$$\text{when } l = 2, \quad \beta_{21}\gamma_{1t} + \beta_{22}\gamma_{2t} + \cdots + \beta_{2,p-1}\gamma_{p-1,t} = -(r\alpha_0\mu + \beta_{20}\gamma_{0,t-2})$$

$$\vdots$$

$$\text{when } l = p - 1, \quad \beta_{p-1,1}\gamma_{1t} + \beta_{p-1,2}\gamma_{2t} + \cdots + \beta_{p-1,p-1}\gamma_{p-1,t} = -(r\alpha_0\mu + \beta_{p-1,0}\gamma_{0,t-p+1})$$

The system of equations can be expressed in the matrix form as

$$\begin{bmatrix} \beta_{11} & \cdots & \beta_{1,p-1} \\ \beta_{21} & \cdots & \beta_{2,p-1} \\ \vdots & \ddots & \vdots \\ \beta_{p-1,1} & \cdots & \beta_{p-1,p-1} \end{bmatrix} \times \begin{bmatrix} \gamma_{1t} \\ \gamma_{2t} \\ \vdots \\ \gamma_{p-1,t} \end{bmatrix} = - \begin{bmatrix} r\alpha_0\mu + \beta_{10}\gamma_{0,t-1} \\ r\alpha_0\mu + \beta_{20}\gamma_{0,t-2} \\ \vdots \\ r\alpha_0\mu + \beta_{p-1,0}\gamma_{0,t-p+1} \end{bmatrix}$$

where $B = (\beta_{ij})_{i,j}^{p-1}$. We can rewrite the matrix as

$$\begin{bmatrix} \gamma_{1t} \\ \gamma_{2t} \\ \vdots \\ \gamma_{p-1,t} \end{bmatrix} = - \begin{bmatrix} \beta_{11} & \cdots & \beta_{1,p-1} \\ \beta_{21} & \cdots & \beta_{2,p-1} \\ \vdots & \ddots & \vdots \\ \beta_{p-1,1} & \cdots & \beta_{p-1,p-1} \end{bmatrix}^{-1} \times \begin{bmatrix} r\alpha_0\mu + \beta_{10}\gamma_{0,t-1} \\ r\alpha_0\mu + \beta_{20}\gamma_{0,t-2} \\ \vdots \\ r\alpha_0\mu + \beta_{p-1,0}\gamma_{0,t-p+1} \end{bmatrix}$$

and since $B^{-1} = (b_{ij})_{i,j}^{p-1}$,

$$\begin{bmatrix} \gamma_{1t} \\ \gamma_{2t} \\ \vdots \\ \gamma_{p-1,t} \end{bmatrix} = - \begin{bmatrix} b_{11} & \cdots & b_{1,p-1} \\ b_{21} & \cdots & b_{2,p-1} \\ \vdots & \ddots & \vdots \\ b_{p-1,1} & \cdots & b_{p-1,p-1} \end{bmatrix} \times \begin{bmatrix} r\alpha_0\mu + \beta_{10}\gamma_{0,t-1} \\ r\alpha_0\mu + \beta_{20}\gamma_{0,t-2} \\ \vdots \\ r\alpha_0\mu + \beta_{p-1,0}\gamma_{0,t-p+1} \end{bmatrix}$$

Therefore,

$$(\gamma_{1t}, \gamma_{2t}, \dots, \gamma_{p-1,t})^T = -B^{-1} (r\alpha_0\mu + \beta_{10}\gamma_{0,t-1}, \dots, r\alpha_0\mu + \beta_{p-1,0}\gamma_{0,t-p+1})^T,$$

which means

$$\gamma_{1t} = -b_{11}(r\alpha_0\mu + \beta_{10}\gamma_{0,t-1}) - \dots - b_{1,p-1}(r\alpha_0\mu + \beta_{p-1,0}\gamma_{0,t-p+1}),$$

\vdots

$$\gamma_{p-1,t} = -b_{p-1,1}(r\alpha_0\mu + \beta_{p-1,0}\gamma_{0,t-p+1}) - \dots - b_{p-1,p-1}(r\alpha_0\mu + \beta_{p-1,0}\gamma_{0,t-p+1}).$$

So for $l = 1, 2, \dots, p-1$

$$\gamma_{lt} = -r\alpha_0\mu \sum_{u=1}^{p-1} b_{lu} - \sum_{u=1}^{p-1} b_{lu}\beta_{u0}\gamma_{0,t-u}. \quad (5.8)$$

The unconditional second moment is computed as the following,

$$\begin{aligned} E(X_t^2) &= \gamma_{0t} = E[E(X_t^2 \mid \mathcal{F}_{t-1})] \\ &= r\alpha_0 + (r + r^2)\alpha_0^2 + [r + 2\alpha_0(r + r^2)] \sum_{i=1}^p \alpha_i E(X_{t-i}) + \\ &\quad (r + r^2) \left(\sum_{i=1}^p \alpha_i^2 E(X_{t-i}^2) + \sum_{\substack{i,j=1 \\ i \neq j}}^p \alpha_i \alpha_j E(X_{t-i} X_{t-j}) \right) \\ &= r\alpha_0 + (r + r^2)\alpha_0^2 + [r + 2\alpha_0(r + r^2)] \sum_{i=1}^p \alpha_i \mu + \\ &\quad (r + r^2) \left(\sum_{i=1}^p \alpha_i^2 \gamma_{0,t-i} + \sum_{\substack{i,j=1 \\ i \neq j}}^p \alpha_i \alpha_j \gamma_{|i-j|,t} \right) \\ \text{let } C &= r\alpha_0 + (r + r^2)\alpha_0^2 + [r + 2\alpha_0(r + r^2)] \sum_{i=1}^p \alpha_i \mu \end{aligned}$$

Then,

$$\begin{aligned} E(X_t^2) &= C + (r + r^2) \left(\sum_{i=1}^p \alpha_i^2 \gamma_{0,t-i} + \sum_{\substack{i,j=1 \\ i \neq j}}^p \alpha_i \alpha_j \gamma_{|i-j|,t} \right) \\ &= C + (r + r^2) \left(\sum_{u=1}^p \alpha_u^2 \gamma_{0,t-u} + \sum_{v=1}^{p-1} \sum_{|i-j|=v} \alpha_i \alpha_j \gamma_{v,t} \right) \end{aligned}$$

Using Equation (5.8), we get

$$\begin{aligned} E(X_t^2) &= C + (r + r^2) \left[\sum_{u=1}^p \alpha_u^2 \gamma_{0,t-u} + \sum_{v=1}^{p-1} \sum_{|i-j|=v} \alpha_i \alpha_j \left(-r\alpha_0\mu \sum_{u=1}^{p-1} b_{vu} - \sum_{u=1}^{p-1} b_{vu}\beta_{u0}\gamma_{0,t-u} \right) \right] \\ &= C - (r + r^2)r\alpha_0\mu \sum_{v=1}^{p-1} \sum_{|i-j|=v} \alpha_i \alpha_j \sum_{u=1}^{p-1} b_{vu} + (r + r^2) \sum_{u=1}^p \alpha_u^2 \gamma_{0,t-u} - \\ &\quad (r + r^2) \sum_{v=1}^{p-1} \sum_{|i-j|=v} \alpha_i \alpha_j \sum_{u=1}^{p-1} b_{vu}\beta_{u0}\gamma_{0,t-u} \\ &= C_0 + (r + r^2) \left[\sum_{u=1}^p \alpha_u^2 \gamma_{0,t-u} - \sum_{u=1}^{p-1} \left(\sum_{v=1}^{p-1} \sum_{|i-j|=v} \alpha_i \alpha_j b_{vu}\beta_{u0} \right) \gamma_{0,t-u} \right] \\ &= C_0 + (r + r^2) \left[\sum_{u=1}^{p-1} \left(\alpha_u^2 - \sum_{v=1}^{p-1} \sum_{|i-j|=v} \alpha_i \alpha_j \beta_{v0} \right) \gamma_{0,t-u} + \alpha_p^2 \gamma_{0,t-p} \right] \\ &= C_0 + \sum_{u=1}^{p-1} (r + r^2) \left(\alpha_u^2 - \sum_{v=1}^{p-1} \sum_{|i-j|=v} \alpha_i \alpha_j b_{vu}\beta_{v0} \right) \gamma_{0,t-u} + (r + r^2) \alpha_p^2 \gamma_{0,t-p} \\ &= C_0 + \sum_{u=1}^{p-1} C_u \gamma_{0,t-u} + C_p \gamma_{0,t-p} \\ &= C_0 + \sum_{u=1}^p C_u \gamma_{0,t-u}, \end{aligned}$$

where $C_0 = C - (r + r^2)r\alpha_0\mu \sum_{v=1}^{p-1} \sum_{|i-j|=v} \alpha_i \alpha_j \sum_{u=1}^{p-1} b_{vu}$,

$C_u = (r + r^2) \left(\alpha_u^2 - \sum_{v=1}^{p-1} \sum_{|i-j|=v} \alpha_i \alpha_j b_{vu}\beta_{v0} \right)$ and $C_p = (r + r^2) \alpha_p^2$.

Again from Goldberg(1958), the non-homogeneous difference equation will have stable solution if all the roots of $1 - C_1 z^{-1} - \dots - C_p z^{-p} = 0$ lie inside the unit circle. \square

Following is a special case of Theorem 5.1 and Theorem 5.2.

Corollary 5.3. *Suppose the process $\{X_t\}$ follows a NBINARCH(p). The necessary and sufficient conditions that X_t is a first-order stationary for $p=1$ and $p=2$ are $r\alpha_1 < 1$ and $r(\alpha_1 + \alpha_2) < 1$ respectively.*

If X_t is first-order stationary, then, for $p=1$ and $p=2$, the second-order stationary conditions are $(r + r^2)\alpha_1^2 < 1$ and $\delta_2 + \delta_1 < 1$ respectively, where

$$\delta_1 = (r + r^2) \left(\alpha_1^2 + \frac{2r\alpha_1^2\alpha_2}{1 - r\alpha_2} \right), \quad \delta_2 = (r + r^2)\alpha_2^2. \quad (\text{See Zhu, 2011})$$

Proof. From Theorem 1, when $p=1$

$$1 - r\alpha_1 z^{-1} = 0$$

$$z = r\alpha_1 < 1.$$

When $p=2$, for stationary, we require the roots to be less than 1 in absolute value, and this is possible if and only if $r\alpha_1 + r\alpha_2 < 1$, $r\alpha_2 - r\alpha_1 < 1$, and $|r\alpha_2| < 1$. Now, suppose the process is first order stationary. Then, second-order stationary condition for $p=1$ is

$$1 - C_1 z^{-1} = 0$$

$$1 - (r + r^2)\alpha_1^2 z^{-1} = 0$$

$$z = (r + r^2)\alpha_1^2 < 1.$$

For $p=2$, from Theorem 2, the condition for second order stationarity is equivalent to the following condition

$$\delta_1 + \delta_2 < 1, \quad \delta_2 - \delta_1 < 1, \quad |\delta_2| < 1. \quad (5.9)$$

Then $\delta_1 > 0$ holds under the assumption of the first-order stationarity, thus the condition in (5.9) is equivalent to $\delta_1 + \delta_2 < 1$. \square

Theorem 5.4 provides us with set of equations from which we can find the variance and ACF.

Theorem 5.4. *Suppose that $\{X_t\}$ following the NBINGARCH(p, q) model is second-order stationary. Let $\gamma_X(k) = \text{cov}(X_t, X_{t-k})$, $\gamma_\lambda(k) = \text{cov}(\lambda_t, \lambda_{t-k})$, then they satisfy the following equations*

$$\begin{aligned}\gamma_X(k) &= \sum_{i=1}^p r\alpha_i \gamma_X(|k-i|) + \sum_{j=1}^{\min(k-1, q)} \beta_j \gamma_X(k-j) + \sum_{j=k}^k r^2 \beta_j \gamma_\lambda(j-k), \quad k \geq 1; \\ \gamma_\lambda(k) &= \sum_{i=1}^{\min(k, p)} r\alpha_i \gamma_\lambda(k-i) + \sum_{i=k+1}^p \frac{\alpha_i}{r} \gamma_X(i-k) + \sum_{j=1}^q \beta_j \gamma(|k-j|), \quad k \geq 0. \quad (\text{See Zhu, 2011})\end{aligned}$$

Proof. Let \mathcal{I}_t be the σ -field generated by $\{\lambda_t, \lambda_{t-1}, \dots\}$, then we have

$$E(X_t | \mathcal{F}_{t-1}, \mathcal{I}_t) = E(X_t | \mathcal{F}_{t-1}) = r\lambda_t \quad (5.10)$$

for $k \geq 0$, from Equation (5.3) and Equation (5.10), we have

$$\begin{aligned}\text{cov}(X_t - r\lambda_t, r\lambda_{t-k}) &= E[(X_t - r\lambda_t)r\lambda_{t-k}] - E(X_t - r\lambda_t)E(r\lambda_{t-k}) \\ &= E[(X_t - r\lambda_t)r\lambda_{t-k}] - 0\mu \\ &= E[(X_t - r\lambda_t)(r\lambda_{t-k} - \mu)] \\ &= E[E((X_t - r\lambda_t)(r\lambda_{t-k} - \mu) | \mathcal{I}_t)] \\ &= E[(r\lambda_{t-k} - \mu)E[(X_t - r\lambda_t) | \mathcal{I}_t]] \\ &= E[(r\lambda_{t-k} - \mu)E[E((X_t - r\lambda_t) | \mathcal{I}_t) | \mathcal{F}_{t-1}]] \\ &= E[(r\lambda_{t-k} - \mu)[E(E(X_t | \mathcal{F}_{t-1}, \mathcal{I}_t) | \mathcal{I}_t) - r\lambda_t]] \\ &= E[(r\lambda_{t-k} - \mu)[E(r\lambda_t | \mathcal{I}_t) - r\lambda_t]] = 0\end{aligned} \quad (5.11)$$

Similarly, for $k < 0$, from Equation (5.2), we have

$$\begin{aligned}
cov(X_t, X_{t-k} - r\lambda_{t-k}) &= E[X_t(X_{t-k} - r\lambda_{t-k})] - E(X_t)E(X_{t-k} - r\lambda_{t-k}) \\
&= E[X_t(X_{t-k} - r\lambda_{t-k})] - \mu 0 \\
&= E[(X_t - \mu)(X_{t-k} - r\lambda_{t-k})] \\
&= E[E[(X_t - \mu)(X_{t-k} - r\lambda_{t-k}) \mid \mathcal{F}_{t-k-1}]] \\
&= E[(X_t - \mu)E[(X_{t-k} - r\lambda_{t-k}) \mid \mathcal{F}_{t-k-1}]] \\
&= E[(X_t - \mu)(r\lambda_{t-k} - E(r\lambda_{t-k} \mid \mathcal{F}_{t-k-1}))]
\end{aligned} \tag{5.12}$$

Then from Equation (5.11) and Equation (5.12), we get

$$cov(X_t, r\lambda_{t-k}) = \begin{cases} cov(r\lambda_t, r\lambda_{t-k}) & k \geq 0, \\ cov(X_t, X_{t-k}) & k < 0. \end{cases} \tag{5.13}$$

For $k \geq 0$, from Equation (5.2) and Equation (5.13), we obtain

$$\begin{aligned}
\gamma_\lambda(k) &= cov(\lambda_t, \lambda_{t-k}) \\
&= cov\left(\alpha_o + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j \lambda_{t-j}, \lambda_{t-k}\right) \\
&= cov\left(\sum_{i=1}^p \alpha_i X_{t-i}, \lambda_{t-k}\right) + cov\left(\sum_{j=1}^q \beta_j \lambda_{t-j}, \lambda_{t-k}\right) \\
&= \sum_{i=1}^{\min(k,p)} \alpha_i cov(X_{t-i}, \lambda_{t-k}) + \sum_{i=k+1}^p \alpha_i cov(X_{t-i}, \lambda_{t-k}) + \sum_{j=1}^q \beta_j cov(\lambda_{t-j}, \lambda_{t-k}) \\
&= \sum_{i=1}^{\min(k,p)} r\alpha_i cov(\lambda_{t-i}, \lambda_{t-k}) + \sum_{i=k+1}^p \frac{\alpha_i}{r} cov(X_{t-i}, X_{t-k}) + \sum_{j=1}^q \beta_j cov(\lambda_{t-j}, \lambda_{t-k}) \\
&= \sum_{i=1}^{\min(k,p)} r\alpha_i \gamma_\lambda(k-i) + \sum_{i=k+1}^p \frac{\alpha_i}{r} \gamma_X(i-k) + \sum_{j=1}^q \beta_j \gamma_\lambda(|k-j|).
\end{aligned}$$

Similarly, for $k \geq 1$, we have

$$\begin{aligned}
\lambda_X(k) &= \text{cov}(X_t, X_{t-k}) \\
&= E[(X_t - \mu)(X_{t-k} - \mu)] \\
&= E[E[(X_t - \mu)(X_{t-k} - \mu) \mid \mathcal{F}_{t-1}]] \\
&= E[(X_{t-k} - \mu)(r\lambda_t - \mu)] \\
&= \text{cov}(r\lambda_t, X_{t-k}) \\
&= r\text{cov}(\lambda_t, X_{t-k}) \\
&= r\text{cov}\left(\alpha_o + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j \lambda_{t-j}, X_{t-k}\right) \\
&= \sum_{i=1}^p r\alpha_i \text{cov}(X_{t-i}, X_{t-k}) + \sum_{j=1}^q r\beta_j \text{cov}(\lambda_{t-j}, X_{t-k}) \\
&= \sum_{i=1}^p r\alpha_i \text{cov}(X_{t-i}, X_{t-k}) + \sum_{j=1}^{\min(k-1, q)} r\beta_j \text{cov}(\lambda_{t-j}, X_{t-k}) + \sum_{j=k}^q r\beta_j \text{cov}(\lambda_{t-j}, X_{t-k}) \\
&= \sum_{i=1}^p r\alpha_i \text{cov}(X_{t-i}, X_{t-k}) + \sum_{j=1}^{\min(k-1, q)} \beta_j \text{cov}(X_{t-j}, X_{t-k}) + \sum_{j=k}^q r^2 \beta_j \text{cov}(\lambda_{t-j}, \lambda_{t-k}) \\
&= \sum_{i=1}^p r\alpha_i \gamma_X(|k-i|) + \sum_{j=1}^{\min(k-1, p)} \beta_j \gamma_X(k-j) + \sum_{j=k}^q r^2 \beta_j \gamma_\lambda(j-k).
\end{aligned}$$

□

Example 5.1. Considering the NBINGARCH(1,1) model. Applying Theorem 5.4, we get

$$\begin{aligned}
\gamma_X(k) &= r\alpha_1 \gamma_X(k-1) + \beta_1 \gamma_X(k-1), \quad k \geq 2 \\
&= (r\alpha_1 + \beta_1) \gamma_X(k-1).
\end{aligned}$$

When $k = 2$: $\gamma_X(2) = (r\alpha_1 + \beta_1)\gamma_X(1)$

When $k = 3$: $\gamma_X(3) = (r\alpha_1 + \beta_1)\gamma_X(2) = (r\alpha_1 + \beta_1)^2\gamma_X(1)$

When $k = 4$: $\gamma_X(4) = (r\alpha_1 + \beta_1)^3\gamma_X(1)$

So, $\gamma_X(k) = (r\alpha_1 + \beta_1)^{k-1}\gamma_X(1)$, $k \geq 2$,

$$\begin{aligned}\gamma_\lambda(k) &= r\alpha_1\gamma_\lambda(k-1) + \beta_1\gamma_\lambda(k-1), \quad k \geq 1 \\ &= (r\alpha_1 + \beta_1)\gamma_\lambda(k-1)\end{aligned}$$

When $k = 1$: $\gamma_\lambda(1) = (r\alpha_1 + \beta_1)\gamma_\lambda(0)$

When $k = 2$: $\gamma_\lambda(2) = (r\alpha_1 + \beta_1)\gamma_\lambda(1) = (r\alpha_1 + \beta_1)^2\gamma_\lambda(0)$ (5.14)

So, $\gamma_\lambda(k) = (r\alpha_1 + \beta_1)^k\gamma_\lambda(0)$, $k \geq 1$.

Now from Theorem 5.4, and Equation (5.4) and Equation (5.5),

$$\begin{aligned}\gamma_X(1) &= r\alpha_1\gamma_X(0) + r^2\beta_1\gamma_\lambda(0) \\ &= r\alpha_1 \left(\mu + \frac{1}{r}\mu^2 + (r + r^2)\gamma_\lambda(0) \right) + r^2\beta_1\gamma_\lambda(0) \\ &= r^2(\alpha_1 + r\alpha_1 + \beta_1)\gamma_\lambda(0) + \alpha_1(r\mu + \mu^2).\end{aligned}$$
 (5.15)

From Theorem 5.4, and using Equation (5.14) and Equation (5.15), we have

$$\begin{aligned}
\gamma_\lambda(0) &= \frac{\alpha_1}{r} \gamma_X(1) + \beta \gamma_\lambda(1) \\
&= \frac{\alpha_1}{r} (r^2(\alpha_1 + r\alpha_1 + \beta_1)\gamma_\lambda(0) + \alpha_1(r\mu + \mu^2)) + \beta_1(r\alpha_1 + \beta_1)\gamma_\lambda(0) \\
&= (\alpha_1^2 r + \alpha_1^2 r^2 + \alpha_1 \beta_1)\gamma_\lambda(0) + \alpha_1^2 \left(\mu + \frac{\mu^2}{r} \right) + (r\alpha_1 \beta_1 + \beta_1^2)\gamma_\lambda(0) \\
&= [(r\alpha_1 + \beta_1)^2 + r\alpha_1^2]\gamma_\lambda(0) + \alpha_1^2 \left(\mu + \frac{\mu^2}{r} \right) \\
\text{var}(\lambda_t) = \gamma_\lambda(0) &= \frac{\alpha_1^2 \left(\mu + \frac{\mu^2}{r} \right)}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2},
\end{aligned}$$

where $\mu = \frac{r\alpha_0}{1-(r\alpha_1+\beta_1)}$. So substituting this value in Equation (5.5), we have

$$\begin{aligned}
\text{var}(X_t) &= \mu + \frac{\mu^2}{r} + (r + r^2) \left[\frac{\alpha_1^2 \left(\mu + \frac{\mu^2}{r} \right)}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \right] \\
&= \frac{(\mu + \frac{\mu^2}{r})(1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2) + (\mu + \frac{1}{r}\mu^2)(r + r^2)\alpha_1^2}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \\
&= \frac{1 - (r\alpha_1 + \beta_1)^2 + r^2\alpha_1^2}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \left(\mu + \frac{\mu^2}{r} \right).
\end{aligned}$$

We can now find autocovariances and autocorrelations.

$$\gamma_X(1) = r\alpha_1\gamma_X(0) + r^2\beta_1\gamma_\lambda(0)$$

$$\gamma_X(2) = r\alpha_1\gamma_X(1) + \beta_1\gamma_\lambda(1) = (r\alpha_1 + \beta_1)[r\alpha_1\gamma_X(0) + r^2\beta_1\gamma_\lambda(0)]$$

$$\gamma_X(3) = (r\alpha_1 + \beta_1)^2[r\alpha_1\gamma_X(0) + r^2\beta_1\gamma_\lambda(0)]$$

$$\Rightarrow \gamma_X(k) = (r\alpha_1 + \beta_1)^{k-1}[r\alpha_1\gamma_X(0) + r^2\beta_1\gamma_\lambda(0)]$$

$$= (r\alpha_1 + \beta_1)^{k-1} \left[r\alpha_1 \frac{(\mu + \frac{\mu^2}{r})[r^2\alpha_1^2 + 1 - (r\alpha_1 + \beta_1)^2]}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} + r^2\beta_1 \frac{\alpha_1^2 \left(\mu + \frac{\mu^2}{r}\right)}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \right]$$

$$= (r\alpha_1 + \beta_1)^{k-1} \frac{r\alpha_1(r^2\alpha_1^2 + 1 - (r\alpha_1 + \beta_1)^2 + r\alpha_1\beta_1)}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \left(\mu + \frac{\mu^2}{r}\right)$$

$$\gamma_X(k) = (r\alpha_1 + \beta_1)^{k-1} \frac{r\alpha_1[1 - \beta_1(\beta_1 + r\alpha_1)]}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \left(\mu + \frac{\mu^2}{r}\right), \quad k \geq 1;$$

and

$$\rho_X(k) = \frac{\gamma_X(k)}{\gamma_X(0)}$$

$$= (r\alpha_1 + \beta_1)^{k-1} \frac{r\alpha_1[1 - \beta_1(\beta_1 + r\alpha_1)]}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \left(\mu + \frac{\mu^2}{r}\right) \div \frac{r^2\alpha_1^2 + 1 - (r\alpha_1 + \beta_1)^2}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \left(\mu + \frac{\mu^2}{r}\right)$$

$$\rho_X(k) = \frac{(r\alpha_1 + \beta_1)^{k-1} r\alpha_1[1 - \beta_1(r\alpha_1 + \beta_1)]}{1 - (r\alpha_1 + \beta_1)^2 + r^2\alpha_1^2}, \quad k \geq 1.$$

(5.16)

Remark 2. From Example 1, we can see that the second-order stationary condition for NBINGARCH(1,1) model is $r\alpha_1^2 + (r\alpha_1 + \beta_1)^2 < 1$.

Corollary 5.5. Suppose that $\{X_t\}$ following NBINARCH(p) model is second-order stationary, then the autocovariance function $\gamma_X(k)$ satisfies the equation

$$\gamma_X(k) = \sum_{i=1}^p r\alpha_i\gamma_X(|k-i|), \quad k \geq 1. \quad (\text{See Zhu, 2011}) \quad (5.17)$$

5.2 ESTIMATION PROCEDURES

We will discuss three methods to estimate the parameters: Yule walker, Conditional least squares and Maximum likelihood estimations. We will bascially use these methods to estimate parameters for NBINGARCH(p,q) model, when p=1,2 and p=1,q=1.

5.2.1 Yule Walker Estimation

When p=1,q=0, $\lambda_t = \alpha_0 + \alpha_1 X_{t-1}$. From Corollary 5.5

$$\gamma_X(k) = \sum_{i=1}^k r \alpha_1 \gamma_X(|k-i|)$$

$$\gamma_X(k) = r \alpha_1 \gamma_X(k-1)$$

$$\alpha_1 = \frac{\gamma_X(k)}{r \gamma_X(k-1)}.$$

where $\gamma_X(k) = \frac{\sum_{t=k+1}^n (X_t - \bar{x})(X_{t-k} - \bar{x})}{n}$ is the sample autocovariance.

If k=1, then

$$\alpha_1 = \frac{\gamma_X(1)}{r \gamma_X(0)},$$

$$\text{where } \gamma_X(1) = \frac{\sum_{t=2}^n (X_t - \bar{x})(X_{t-1} - \bar{x})}{n} \quad \text{and} \quad \gamma_X(0) = \frac{\sum_{t=1}^n (X_t - \bar{x})^2}{n}.$$

$$\text{So, } \hat{\alpha}_1 = \frac{\sum_{t=2}^n (X_t - \bar{x})(X_{t-1} - \bar{x})}{r \sum_{t=1}^n (X_t - \bar{x})^2}$$

From Remark 1,

$$\begin{aligned} \mu &= \frac{r \alpha_0}{1 - r \alpha_1} \\ \Rightarrow \alpha_0 &= \frac{\mu(1 - r \alpha_1)}{r} \\ \hat{\alpha}_0 &= \frac{\bar{x}(1 - r \hat{\alpha}_1)}{r} \end{aligned}$$

When $p=2, q=0$,

$$\lambda_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2}$$

$$\gamma_X(k) = r\alpha_1\gamma_X(k-1) + r\alpha_2\gamma_X(k-2)$$

$$\text{If } k=1 : \gamma_X(1) = r\alpha_1\gamma_X(0) + r\alpha_2\gamma_X(1)$$

$$\text{If } k=2 : \gamma_X(2) = r\alpha_1\gamma_X(1) + r\alpha_2\gamma_X(0)$$

Solving by elimination method, we get

$$\begin{aligned}\hat{\alpha}_1 &= \frac{\gamma_X(1)[\gamma_X(0) - \gamma_X(2)]}{r[\gamma_X^2(0) - \gamma_X^2(1)]} \\ &= \frac{\rho_X(1)(1 - \rho_X(2))}{r(1 - \rho_X^2(1))}, \\ \hat{\alpha}_2 &= \frac{\rho_X(2) - r\hat{\alpha}_1\rho_X(1)}{r}, \quad \text{and} \\ \mu &= \frac{r\alpha_0}{1 - r\alpha_1 - r\alpha_2} \\ \Rightarrow \hat{\alpha}_0 &= \frac{\bar{x}(1 - r\hat{\alpha}_1 - r\hat{\alpha}_2)}{r},\end{aligned}$$

where $\rho_X(1)$ and $\rho_X(2)$ are sample autocorrelation.

When $p=q=1$,

$$\lambda_t = \alpha_0 + \alpha_1 X_{t-1} + \beta_1 \lambda_{t-1}.$$

Since the model has past lag of λ , it complicates the estimation process. In order to make the process simple, Lemma 2 from Ferland et al. (2006) is considered.

Lemma 2 : Suppose that $\{X_t\}$ and $\{Y_t\}$ are zero-mean stationary processes with autocovariance function $\gamma(\cdot)$ and that $\{Y_t\}$ is an ARMA(p,q) process. Then, $\{X_t\}$ is also an ARMA(p,q) process.

From the lemma and Equation (5.16), Equation (5.2) satisfies the ARMA(1,1) model

$$(X_t - \mu) - a(X_{t-1} - \mu) = e_t + be_{t-1}, \quad (5.18)$$

where $a = r\alpha_1 + \beta_1, b = -\beta_1,$ e_t is a white noise process with

$$\text{var}(e_t) = \frac{1 - (r\alpha_1 + \beta_1)^2}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \left(\mu + \frac{\mu^2}{r} \right).$$

To show that we can use ARMA(1,1) to estimate the parameters, we will show that covariance of ARMA(1,1) is same as the covariance of NBINGARCH(1,1). Multiplying both sides by X_{t-k} and finding expectation of Equation (5.18), we have

$$E(X_t X_{t-k}) = aE(X_{t-1} X_{t-k}) + E(e_t X_{t-k}) + bE(e_{t-1} X_{t-k})$$

$$\text{For } k = 0 : E(X_t X_t) = aE(X_{t-1} X_t) + E(e_t X_t) + bE(e_{t-1} X_t)$$

$$\gamma_X(0) = a\gamma_X(1) + \text{var}(e_t) + b[a \text{var}(e_t) + b \text{var}(e_t)] \quad (5.19)$$

$$= a\gamma_X(1) + \text{var}(e_t)(1 + ab + b^2)$$

$$\text{For } k = 1 : E(X_t X_{t-1}) = aE(X_{t-1} X_{t-1}) + E(e_t X_{t-1}) + bE(e_{t-1} X_{t-k}) \quad (5.20)$$

$$\gamma_X(1) = a\gamma_X(0) + b \text{var}(e_t).$$

Solving Equation (5.19) and Equation (5.20), we get

$$\begin{aligned} \gamma_X(0) &= a^2 \gamma_X(0) + ab \text{var}(e_t) + \text{var}(e_t)(1 + ab + b^2) \\ &= \frac{1 + 2ab + b^2}{1 - a^2} \text{var}(e_t). \end{aligned}$$

$$\text{For } k = 2 : \gamma_X(2) = a\gamma_X(1)$$

$$\text{For } k = 3 : \gamma_X(3) = a\gamma_X(2) = a^2\gamma_X(1)$$

$$\Rightarrow \gamma_X(k) = a^{k-1}\gamma_X(1)$$

$$= a^{k-1}[a\gamma_X(0) + b \text{var}(e_t)]$$

$$= a^{k-1} \left(a \frac{1 + 2ab + b^2}{1 - a^2} \text{var}(e_t) + b \text{var}(e_t) \right)$$

$$= a^{k-1} \left(\frac{a + b + a^2b + ab^2}{1 - a^2} \right) \text{var}(e_t)$$

$$= a^{k-1} \left(\frac{(a + b)(1 + ab)}{1 - a^2} \right) \text{var}(e_t).$$

If $a = r\alpha_1 + \beta_1$ and $b = -\beta_1$, then

$$\begin{aligned} \gamma_X(k) &= (r\alpha_1 + \beta_1)^{k-1} \frac{(r\alpha_1 + \beta_1 - \beta_1)(1 - \beta_1(r\alpha_1 + \beta_1))}{1 - (r\alpha_1 + \beta_1)^2} \text{var}(e_t) \\ &= (r\alpha_1 + \beta_1)^{k-1} \frac{(r\alpha_1 + \beta_1)(1 - \beta_1(r\alpha_1 + \beta_1))}{1 - (r\alpha_1 + \beta_1)^2} \text{var}(e_t). \end{aligned}$$

Now, if $\text{var}(e_t) = \frac{1 - (r\alpha_1 + \beta_1)^2}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \left(\mu + \frac{\mu^2}{r} \right)$, then

$$\gamma_X(k) = (r\alpha_1 + \beta_1)^{k-1} \frac{r\alpha_1[1 - \beta_1(\beta_1 + r\alpha_1)]}{1 - (r\alpha_1 + \beta_1)^2 - r\alpha_1^2} \left(\mu + \frac{\mu^2}{r} \right),$$

which is same as the covariance of NBINGARCH(1,1).

The autocorrelation of NBINGARCH(1,1) is given by

$$\begin{aligned}
\rho_X(k) &= a^{k-1} \left(\frac{a\gamma_X(0) + b \operatorname{var}(e_t)}{\gamma_X(0)} \right) \\
&= a^{k-1} \left(a + \frac{b \operatorname{var}(e_t)(1 - a^2)}{(1 + 2ab + b^2)\operatorname{var}(e_t)} \right) \\
&= a^{k-1} \left(\frac{a + b + a^2b + ab^2}{1 + 2ab + b^2} \right) \\
&= a^{k-1} \left(\frac{(a + b)(1 + ab)}{1 + 2ab + b^2} \right), \quad k \geq 1.
\end{aligned}$$

So, $\frac{\rho_X(2)}{\rho_X(1)} = a$.

Sample autocorrelation can be used to solve for a .

5.2.2 CLS Estimation

The CLS estimates mean to minimize

$$S(\alpha_0, \alpha_i) = \sum_{t=1}^n \left(X_t - r\alpha_0 - \sum_{i=1}^p r\alpha_i X_{t-i} \right)^2. \quad (5.21)$$

When $p=1, q=0$,

$$S(\alpha_0, \alpha_1) = \sum_{t=1}^n (X_t - r\alpha_0 - r\alpha_1 X_{t-1})^2.$$

$$\frac{\partial S(\alpha_0, \alpha_1)}{\partial \alpha_0} = -2r \left(\sum_{t=1}^n X_t - nr\alpha_0 - r\alpha_1 \sum_{t=1}^n X_{t-1} \right) = 0$$

$$\sum_{t=1}^n X_t = nr\alpha_0 + r\alpha_1 \sum_{t=1}^n X_{t-1} \quad \text{and}$$

$$\frac{\partial S(\alpha_0, \alpha_1)}{\partial \alpha_1} = -2r \left(\sum_{t=1}^n X_t X_{t-1} - r\alpha_0 \sum_{t=1}^n X_{t-1} - r\alpha_1 \sum_{t=1}^n X_{t-1} X_{t-1} \right) = 0$$

$$\sum_{t=1}^n X_t X_{t-1} = r\alpha_0 \sum_{t=1}^n X_{t-1} + r\alpha_1 \sum_{t=1}^n X_{t-1}^2$$

Expressing the above information in matrix form, we get

$$\begin{bmatrix} \sum_{t=1}^n X_t \\ \sum_{t=1}^n X_t X_{t-1} \end{bmatrix} = \begin{bmatrix} n-1 & \sum_{t=1}^n X_{t-1} \\ \sum_{t=1}^n X_{t-1} & \sum_{t=1}^n X_{t-1}^2 \end{bmatrix} \begin{bmatrix} r\alpha_0 \\ r\alpha_1 \end{bmatrix}$$

Value for α_0 and α_1 is obtained by solving the above matrix.

When $p=2, q=0$,

$$S(\alpha_0, \alpha_1, \alpha_2) = \sum_{t=1}^n (X_t - r\alpha_0 - r\alpha_1 X_{t-1} - r\alpha_2 X_{t-2})^2.$$

$$\frac{\partial S(\alpha_0, \alpha_1, \alpha_2)}{\partial \alpha_0} = -2r \sum_{t=1}^n (X_t - r\alpha_0 - r\alpha_1 X_{t-1} - r\alpha_2 X_{t-2}) = 0$$

$$\sum_{t=1}^n X_t = nr\alpha_0 + r\alpha_1 \sum_{t=1}^n X_{t-1} + r\alpha_2 \sum_{t=1}^n X_{t-2},$$

$$\frac{\partial S(\alpha_0, \alpha_1, \alpha_2)}{\partial \alpha_1} = -2r \sum_{t=1}^n (X_t X_{t-1} - r\alpha_0 X_{t-1} - r\alpha_1 X_{t-1}^2 - r\alpha_2 X_{t-1} X_{t-2}) = 0$$

$$\sum_{t=1}^n (X_t X_{t-1} = r\alpha_0 \sum_{t=1}^n X_{t-1} + r\alpha_1 \sum_{t=1}^n X_{t-1}^2 + r\alpha_2 \sum_{t=1}^n X_{t-1} X_{t-2})$$

$$\frac{\partial S(\alpha_0, \alpha_1, \alpha_2)}{\partial \alpha_2} = -2r \sum_{t=1}^n (X_t X_{t-2} - r\alpha_0 X_{t-2} - r\alpha_1 X_{t-1} X_{t-2} - r\alpha_2 X_{t-2}^2) = 0$$

$$\sum_{t=1}^n X_t X_{t-2} = r\alpha_0 \sum_{t=1}^n X_{t-2} + r\alpha_1 \sum_{t=1}^n X_{t-1} X_{t-2} + r\alpha_2 \sum_{t=1}^n X_{t-2}^2$$

So, the matrix form will be

$$\begin{bmatrix} \sum_{t=1}^n X_t \\ \sum_{t=1}^n (X_t X_{t-1}) \\ \sum_{t=1}^n X_t X_{t-2} \end{bmatrix} = \begin{bmatrix} n-2 & \sum_{t=1}^n X_{t-1} & \sum_{t=1}^n X_{t-2} \\ \sum_{t=1}^n X_{t-1} & \sum_{t=1}^n X_{t-1}^2 & \sum_{t=1}^n X_{t-1} X_{t-2} \\ \sum_{t=1}^n X_{t-2} & \sum_{t=1}^n X_{t-1} X_{t-2} & \sum_{t=1}^n X_{t-2}^2 \end{bmatrix} \begin{bmatrix} r\alpha_0 \\ r\alpha_1 \\ r\alpha_2 \end{bmatrix}$$

Solving the matrix, we will get the value for α_0 , α_1 , and α_2 .

When $p=q=1$, Equation (5.18) can also be written as

$$X_t = r\alpha_0 + aX_{t-1} + e_t + be_{t-1}$$

We need to find the value of e_t before we proceed with the estimation. So we will follow the following steps in doing so.

1. let $Y_t = X_t - \sum_{i=1}^n X_i/n$, fit the data by using AR(p) model, then obtain the CLS estimator for the autoregressive coefficient and define \hat{e}_t as the residual of process Y_t .

Remark 3. Smaller order of AP model provided better estimates for parameters unlike fitting higher order as indicated in [3].

2. Minimize

$$S(\alpha_0, a, b) = \sum_{t=1}^n (X_t - r\alpha_0 - aX_{t-1} - b\hat{e}_{t-1})^2$$

$$\frac{\partial S(\alpha_0, a, b)}{\partial \alpha_0} = -2r \sum_{t=1}^n (X_t - r\alpha_0 - aX_{t-1} - b\hat{e}_{t-1}) = 0$$

$$\hat{\alpha}_0 = \frac{\sum_{t=1}^n (X_t - aX_{t-1} - b\hat{e}_{t-1})}{r}$$

$$\frac{\partial S(\alpha_0, a, b)}{\partial a} = 2 \sum_{t=1}^n (X_t - r\alpha_0 - aX_{t-1} - b\hat{e}_{t-1})(X_{t-1}) = 0$$

$$\hat{a} = \frac{\sum_{t=1}^n (X_t - r\alpha_0 - b\hat{e}_{t-1})}{X_{t-1}}$$

$$\frac{\partial S(\alpha_0, a, b)}{\partial b} = 2 \sum_{t=1}^n (X_t - r\alpha_0 - aX_{t-1} - b\hat{e}_{t-1})(-\hat{e}_{t-1})$$

$$\hat{b} = \frac{-\sum_{t=1}^n (X_t - r\alpha_0 - aX_{t-1})}{\hat{e}_{t-1}}.$$

The value of α_0, a, b can also be obtained by converting the following system of equations to matrix form.

$$\begin{aligned}\sum X_t &= \sum r\alpha_0 + a \sum X_{t-1} + b \sum \hat{e}_{t-1} \\ \sum X_t X_{t-1} &= \sum r\alpha_0 X_{t-1} + a \sum X_{t-1}^2 + b \sum \hat{e}_{t-1} \\ \sum X_t \hat{e}_{t-1} &= \sum r\alpha_0 \hat{e}_{t-1} + a \sum X_{t-1} \hat{e}_{t-1} + b \sum \hat{e}_{t-1}^2,\end{aligned}$$

which is equivalent to the matrix

$$\begin{bmatrix} \sum X_t \\ \sum X_t X_{t-1} \\ \sum X_t \hat{e}_{t-1} \end{bmatrix} = \begin{bmatrix} n-1 & \sum X_{t-1} & \sum \hat{e}_{t-1} \\ \sum X_{t-1} & \sum X_{t-1}^2 & \sum \hat{e}_{t-1} X_{t-1} \\ \sum \hat{e}_{t-1} & \sum X_{t-1} \hat{e}_{t-1} & \sum \hat{e}_{t-1}^2 \end{bmatrix} \begin{bmatrix} r\alpha_0 \\ a \\ b \end{bmatrix}$$

So $\hat{\beta}_1 = -\hat{b}$ and $\hat{\alpha}_1 = \frac{\hat{a} + \hat{b}}{r}$.

5.2.3 Maximum likelihood Estimation

We know

$$\begin{aligned}
P(X_t = x_t | \mathcal{F}_{t-1}) &= \binom{x_t + r - 1}{r - 1} p_t^r (1 - p_t)^{x_t} \\
L(\theta) &= \prod_{t=1}^n \binom{x_t + r - 1}{r - 1} p_t^r (1 - p_t)^{x_t} \\
\log L(\theta) &= \sum_{t=1}^n r \log(p_t) + \sum_{t=1}^n X_t \log(1 - p_t) + \log \prod_{t=1}^n \binom{X_t + r - 1}{r - 1} \\
&= \sum_{t=1}^n \log \frac{1}{1 + \lambda_t} + \sum_{t=1}^n X_t \log \frac{\lambda_t}{1 + \lambda_t} + \log \prod_{t=1}^n \frac{(X_t + r - 1)!}{(r - 1)! X_t!} \\
&= - \sum_{t=1}^n r \log(1 + \lambda_t) + \sum_{t=1}^n X_t (\log \lambda_t - \log(1 + \lambda_t)) + \sum_{t=1}^n \log((X_t + r - 1)!) - \\
&\quad \sum_{t=1}^n \log((r - 1)!) - \sum_{t=1}^n \log(X_t!) \\
&= \sum_{t=1}^n X_t \log \lambda_t - \sum_{t=1}^n (r + X_t) \log(1 + \lambda_t) + \sum_{t=1}^n \log((X_t + r - 1)!) - \\
&\quad \sum_{t=1}^n \log((r - 1)!) - \sum_{t=1}^n \log(X_t!) \\
\log L(\theta) &= \sum_{t=1}^n [X_t \log \lambda_t - (r + X_t) \log(1 + \lambda_t) + \log((X_t + r - 1)!) - \log((r - 1)!) - \log(X_t!)] \\
&= \sum_{t=1}^n l_t(\theta).
\end{aligned} \tag{5.22}$$

The estimates cannot be obtained algebraically. So the numerical optimization method has to be used. To obtain asymptotic standard errors of MLE, we need the first derivative of $\log L(\theta)$ with respect to $\theta_i (i = 0, 1, \dots, p + q)$ which is as following

$$\frac{\partial l_t}{\partial \theta_i} = \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right) \frac{\partial \lambda_t}{\partial \theta_i}. \tag{5.23}$$

and the second derivative obtained as

$$\begin{aligned}
\frac{\partial^2 l_t}{\partial \theta_i \partial \theta_j} &= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right) \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} + \left(-\frac{X_t}{\lambda_t^2} + \frac{r + X_t}{(1 + \lambda_t)^2} \right) \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} \\
&= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right) \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} - \left(\frac{X_t}{\lambda_t^2} - \frac{r + X_t}{(1 + \lambda_t)^2} \right) \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j}, \quad \text{for } i, j = 0, 1, \dots, p + q.
\end{aligned} \tag{5.24}$$

Moreover, we know

$$\lambda_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j \lambda_{t-j}.$$

So,

$$\begin{aligned}
\frac{\partial \lambda_t}{\partial \alpha_0} &= 1 + \sum_{k=1}^q \beta_k \frac{\partial \lambda_{t-k}}{\partial \alpha_0} \\
\frac{\partial \lambda_t}{\partial \alpha_i} &= X_{t-i} + \sum_{k=1}^q \beta_k \frac{\partial \lambda_{t-k}}{\partial \alpha_i}, \quad i = 1, 2, \dots, p, \\
\frac{\partial \lambda_t}{\partial \beta_j} &= \lambda_{t-j} + \sum_{k=1}^q \beta_k \frac{\partial \lambda_{t-k}}{\partial \beta_j}, \quad j = 1, 2, \dots, q.
\end{aligned} \tag{5.25}$$

Taking expectation of Equation (5.23), we get

$$\begin{aligned}
E \left[\frac{\partial^2 l_t}{\partial \theta_i \partial \theta_j} \mid \mathcal{F}_{t-1} \right] &= \left(\frac{E(X_t \mid \mathcal{F}_{t-1})}{\lambda_t} - \frac{r + E(X_t \mid \mathcal{F}_{t-1})}{1 + \lambda_t} \right) \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} - \\
&\quad \left(\frac{E(X_t \mid \mathcal{F}_{t-1})}{\lambda_t^2} - \frac{r + E(X_t \mid \mathcal{F}_{t-1})}{(1 + \lambda_t)^2} \right) \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} \\
&= \left(\frac{r \lambda_t}{\lambda_t} - \frac{r + r \lambda_t}{1 + \lambda_t} \right) \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} - \left(\frac{r \lambda_t}{\lambda_t^2} - \frac{r + r \lambda_t}{(1 + \lambda_t)^2} \right) \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} \\
&= -r \left(\frac{1}{\lambda_t} - \frac{1}{1 + \lambda_t} \right) \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j}
\end{aligned} \tag{5.26}$$

Similarly, from Equation (5.3), we have

$$\begin{aligned}
E \left[\frac{\partial l_t}{\partial \theta_i} \frac{\partial l_t}{\partial \theta_j} \mid \mathcal{F}_{t-1} \right] &= E \left[\left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right)^2 \mid \mathcal{F}_{t-1} \right] \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} \\
&= \frac{E(X_t - r\lambda_t)^2 \mid \mathcal{F}_{t-1}}{\lambda_t^2(1 + \lambda_t)^2} \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} \\
&= \frac{r\lambda_t(1 + \lambda_t)}{\lambda_t^2(1 + \lambda_t)^2} \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} \\
&= r \left(\frac{1}{\lambda_t} - \frac{1}{1 + \lambda_t} \right) \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j}.
\end{aligned} \tag{5.27}$$

So, from Equation (5.26) and Equation (5.27), we obtain the information matrix equality

$$-E \left[\frac{\partial^2 l_t}{\partial \theta_i \partial \theta_j} \mid \mathcal{F}_{t-1} \right] = E \left[\frac{\partial l_t}{\partial \theta_i} \frac{\partial l_t}{\partial \theta_j} \mid \mathcal{F}_{t-1} \right], \quad i, j = 0, 1, \dots, p + q.$$

From Ferland et al. (2006), asymptotic standard errors of MLEs can be computed from the following matrix

$$\frac{1}{n} (\hat{D}_n \hat{S}_n^{-1} \hat{D}_n)^{-1}, \tag{5.28}$$

where

$$\hat{S}_n = \frac{1}{n} \sum_{t=1}^n \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta^T}, \quad \hat{D}_n = -\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \theta \partial \theta^T}.$$

CHAPTER 6

DATA ANALYSIS

6.1 SIMULATION STUDY

A simulation study is done using R software. The study is conducted for three NBINGARCH(p,q) models, where the value of r is assumed to be given. The value used here for r are 1 and 2. The three simulated models are NBINGARCH($p=1$), NBINGARCH($p=2$), and NBINGARCH($p=1,q=1$) with a set of parameters for each value of r . Sample sizes $n=100$ and $n=500$ are considered with each parameter set, and the number of replication used is 200. The simulation is being done using three methods discussed in the previous section.

The summary of the simulation results are provided in Table 6.1. It can be seen that as the sample size increases, the estimates approaches the exact value of the parameters. All the methods perform well but MLE performs better.

Note 1. Even with a small value of n , Maximum likelihood methodology performs better.

For each model, two sets of values of parameters are used.

1. $(\alpha_0, \alpha_1, r) = (2, 0.4, 1)$ and $(5, 0.3, 2)$, for A1= NBINGARCH(1), A2= NBINGARCH(1) respectively;
2. $(\alpha_0, \alpha_1, \alpha_2, r) = (2, 0.3, 0.2, 1)$ and $(3, 0.2, 0.2, 2)$ for B1= NBINGARCH(2), B2= NBINGARCH(2) respectively;
3. $(\alpha_0, \alpha_1, \beta_1, r) = (2, 0.5, 0.3, 1)$ and $(3, 0.2, 0.2, 2)$ for C1= NBINGARCH(1,1), C2= NBINGARCH(1,1) respectively.

The mean absolute errors are obtained by finding the absolute difference between the given value of the parameters and the simulated value of the parameters. The mean abso-

lute error are represented in parenthesis in Table 6.1.

6.2 REAL DATA ANALYSIS

As a practical demonstration of fitting INGARCH model, the polio data (168 observations) studied by Zeger (1988) consisting of monthly counts of poliomyelitis cases in the USA from the year 1970 to 1983 as reported by the Centers for Disease Control is considered. Poliomyelitis is an acute viral disease, usually affecting children and young adults, caused by polio viruses and can cause temporary or permanent paralysis. To provide insight to the data structure, Figure 6.1 presents the original series, the ACF and the PACF of the series.

The empirical mean and variance of the data are 1.333333 and 3.50499 respectively, indicating that the distribution is overdispersed. The data are fitted by NBINARCH(1) and NBINGARCH(1,1) models. The approach of Benjamin et.al(2003) and Davis and Wu (2009) is adopted, which maximizes the likelihood $\log L(\theta)$ with respect to θ for different values of r ($= 1, 2, 3, 4, 5$). The value of r is chosen taking into account the smallest Akaike information criterion (AIC) or Bayesian information criterion (BIC). So, $\hat{r} = 2$ provides the smallest AIC and BIC from the Table 6.2.

With the chosen value of r , NBINARCH(1) and NBINGARCH(1,1) models are fitted to the data. For comparison, Poisson (P) and Double Poisson(DP) models are considered and the results are displayed in Table 6.3 . In both INARCH(1) and INGARCH(1,1) models, we can see that NB performs better and between the two models NBINARCH(1) generates smaller AIC and BIC. Hence NBINARCH(1) model is preferred over NBINGARCH(1,1) model. The asymptotic standard errors (SE) are calculated using Equation (5.28) and are shown in parenthesis in Table 6.3. For each NBINARCH(1) and NBINGARCH(1,1) model, the calculation of asymptotic standard errors are presented in the following section.

6.2.1 INARCH(1) Model

From Equation 5.2,

$$\lambda_t = \alpha_0 + \alpha_1 X_{t-1}.$$

6.2.1.1 Negative Binomial

The conditional probability mass function of negative binomial model $X_t | \mathcal{F}_{t-1}$: $\mathcal{NB}(r, p_t)$ is given by

$$P(X_t = x_t | \mathcal{F}_{t-1}) = \binom{x_t + r - 1}{r - 1} p_t^r (1 - p_t)^{x_t}.$$

From Equation 5.22, the likelihood function is written as

$$\log L(\theta) = \sum_{t=1}^n [X_t \log \lambda_t - (r + X_t) \log(1 + \lambda_t) + \log((X_t + r - 1)!) - \log((r - 1)!) - \log(X_t!)] = \sum_{t=1}^n l_t(\theta).$$

Using Equation (5.23), we get

$$\begin{aligned} \frac{\partial l_t}{\partial \theta} &= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right) \begin{bmatrix} 1 \\ X_{t-1} \end{bmatrix} \\ \frac{\partial l_t}{\partial \theta^T} &= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right) \begin{bmatrix} 1 & X_{t-1} \end{bmatrix} \\ \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta^T} &= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right)^2 \begin{bmatrix} 1 & X_{t-1} \\ X_{t-1} & X_{t-1}^2 \end{bmatrix}. \end{aligned}$$

From Equation (5.24), we get the second derivatives as

$$\frac{\partial^2 l_t}{\partial \theta \partial \theta^T} = - \left(\frac{X_t}{\lambda_t^2} - \frac{r + X_t}{(1 + \lambda_t)^2} \right) \begin{bmatrix} 1 & X_{t-1} \\ X_{t-1} & X_{t-1}^2 \end{bmatrix}.$$

The matrix for second derivatives obtained above is also called the Hessian matrix (H). So,

$$\begin{aligned} \hat{S}_n &= \frac{1}{n} \sum_{t=1}^n \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta^T} \\ \hat{D}_n &= -\frac{1}{n} \sum_{t=1}^n \frac{\partial^2 l_t}{\partial \theta \partial \theta^T}, \end{aligned}$$

and the asymptotic standard error is found by solving

$$\frac{1}{n}(\hat{D}_n \hat{S}_n^{-1} \hat{D}_n)^{-1}.$$

6.2.1.2 Poisson

The conditional probability mass function of Poisson model $X_t | \mathcal{F}_{t-1} : \mathcal{P}(\lambda_t)$ is given by

$$f(x_t) = \frac{e^{-\lambda_t} \lambda_t^{X_t}}{X_t!}.$$

The likelihood function, $\log L(\theta)$ is given by

$$\log L(\theta) = \sum_{t=1}^n X_t \log \lambda_t - \sum_{t=1}^n \lambda_t - \sum_{t=1}^n \log(X_t!) = \sum_{t=1}^n l_t(\theta) = \sum_{t=1}^n l_t(\theta),$$

where $l_t(\theta) = X_t \log \lambda_t - \lambda_t - \log(X_t!)$. Now, for asymptotic standard errors, following expressions are required, which are obtained with the use of Equation (5.23) and Equation (5.24).

$$\begin{aligned}
\frac{\partial l_t}{\partial \theta} &= \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \theta} = \left(\frac{X_t}{\lambda_t} - 1 \right) \begin{bmatrix} 1 \\ X_{t-1} \end{bmatrix}, \\
\frac{\partial l_t}{\partial \theta^T} &= \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \theta} = \left(\frac{X_t}{\lambda_t} - 1 \right) \begin{bmatrix} 1 & X_{t-1} \end{bmatrix}, \\
\frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta^T} &= \left(\frac{X_t}{\lambda_t} - 1 \right)^2 \begin{bmatrix} 1 & X_{t-1} \\ X_{t-1} & X_{t-1}^2 \end{bmatrix}. \\
\frac{\partial^2 l_t}{\partial \theta \partial \theta^T} &= \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial^2 \lambda_t}{\partial \theta \partial \theta^T} - \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \theta} \frac{\partial \lambda_t}{\partial \theta^T} \\
&= - \left(\frac{X_t}{\lambda_t^2} \right) \begin{bmatrix} 1 & X_{t-1} \\ X_{t-1} & X_{t-1}^2 \end{bmatrix} = H
\end{aligned}$$

So making use of equation (5.28), the asymptotic standard errors can be calculated.

6.2.1.3 Double Poisson

The conditional probability mass function of Double Poisson model $X_t \mid \mathcal{F}_{t-1}$: $\mathcal{DP}(\lambda_t, \gamma)$ is

$$f(x_t) = (\gamma^{1/2} e^{-\gamma \lambda_t}) \left(\frac{e^{-X_t} X_t^{X_t}}{X_t!} \right) \left(\frac{e \lambda_t}{X_t} \right)^{\gamma X_t}, \quad \gamma > 0, \quad X_t = 0, 1, 2, \dots$$

The likelihood function, $\log L(\theta)$ is given by

$$\begin{aligned}
\log L(\theta) &= \frac{n}{2} \log \gamma - \gamma \sum_{t=1}^n \lambda_t - (1 - \gamma) \sum_{t=1}^n X_t + (1 - \gamma) \sum_{t=1}^n X_t \log X_t + \sum_{t=1}^n \gamma X_t \log \lambda_t - \sum_{t=1}^n \log(X_t!) \\
&= \sum_{t=1}^n l_t(\theta),
\end{aligned}$$

where $l_t(\theta) = \frac{1}{2} \log \gamma - \gamma \lambda_t - (1 - \gamma) X_t + (1 - \gamma) X_t \log X_t + \gamma X_t \log \lambda_t - \log(X_t!)$.

The fitted values of parameters are found slightly different here. First γ is assumed to be 1, which converts the model to Poisson. So the fitted value of α_0 and α_1 is same as the one obtained by fitting Poisson model. Then the fitted value of γ is found in a following way

$$\begin{aligned}
\frac{\partial \log L(\theta)}{\partial \gamma} &= \frac{n}{2\gamma} - \sum_{t=1}^n \lambda_t + \sum_{t=1}^n X_t - \sum_{t=1}^n X_t \log(X_t) + \sum_{t=1}^n X_t \log(\lambda_t) = 0 \\
\hat{\gamma} &= \frac{n}{2(\sum_{t=1}^n \lambda_t - \sum_{t=1}^n X_t + \sum_{t=1}^n X_t \log(X_t) - \sum_{t=1}^n X_t \log(\lambda_t))}.
\end{aligned}$$

For the asymptotic standard errors, differentiating l_t with respect to each parameter, we get

$$\begin{aligned}
\frac{\partial l_t}{\partial \alpha_0} &= -\gamma \frac{\partial \lambda_t}{\partial \alpha_0} + \frac{\gamma X_t}{\lambda_t} \frac{\partial \lambda_t}{\partial \alpha_0} = -\gamma + \frac{\gamma X_t}{\lambda_t}. \\
\frac{\partial l_t}{\partial \alpha_1} &= -\gamma \frac{\partial \lambda_t}{\partial \alpha_1} + \frac{\gamma X_t}{\lambda_t} \frac{\partial \lambda_t}{\partial \alpha_1} = \left(-\gamma + \frac{\gamma X_t}{\lambda_t}\right) X_{t-1}. \\
\frac{\partial l_t}{\partial \gamma} &= \frac{1}{2\gamma} - \lambda_t + X_t - X_t \log X_t + X_t \log \lambda_t
\end{aligned}$$

So

$$\frac{\partial l_t}{\partial \theta} = \begin{bmatrix} -\gamma + \frac{\gamma X_t}{\lambda_t} \\ -\gamma X_{t-1} + \frac{\gamma X_t X_{t-1}}{\lambda_t} \\ \frac{1}{2\gamma} - \lambda_t + X_t - X_t \log X_t + X_t \log \lambda_t \end{bmatrix},$$

$$\frac{\partial l_t}{\partial \theta^T} = \left[-\gamma + \gamma \frac{X_t}{\lambda_t} \quad -\gamma X_{t-1} + \gamma \frac{X_t X_{t-1}}{\lambda_t} \quad \frac{1}{2\gamma} - \lambda_t + X_t - X_t \log X_t + X_t \log \lambda_t \right].$$

The second derivatives are

$$\begin{aligned} \frac{\partial^2 l_t}{\partial \alpha_0^2} &= -\frac{\gamma X_t}{\lambda_t^2}, & \frac{\partial^2 l_t}{\partial \alpha_0 \partial \alpha_1} &= -\frac{\gamma X_t}{\lambda_t^2} X_{t-1}, & \frac{\partial^2 l_t}{\partial \alpha_0 \partial \gamma} &= -1 + \frac{X_t}{\lambda_t}, \\ \frac{\partial^2 l_t}{\partial \alpha_1 \alpha_0} &= \frac{-\gamma X_t X_{t-1}}{\lambda_t^2}, & \frac{\partial^2 l_t}{\partial \alpha_1^2} &= \frac{-\gamma X_t X_{t-1}^2}{\lambda_t^2}, & \frac{\partial^2 l_t}{\partial \alpha_1 \gamma} &= -X_{t-1} + \frac{X_t X_{t-1}}{\lambda_t}, \\ \frac{\partial^2 l_t}{\partial \gamma \alpha_0} &= -1 + \frac{X_t}{\lambda_t}, & \frac{\partial^2 l_t}{\partial \gamma \alpha_1} &= -X_{t-1} + \frac{X_t X_{t-1}}{\lambda_t}, & \frac{\partial^2 l_t}{\partial \gamma^2} &= -\frac{1}{2\gamma^2}. \end{aligned}$$

So the Hessian matrix (H) is

$$H = \frac{\partial^2 l_t}{\partial \theta \partial \theta^T} = \begin{bmatrix} -\frac{\gamma X_t}{\lambda_t^2} & -\frac{\gamma X_t}{\lambda_t^2} X_{t-1} & -1 + \frac{X_t}{\lambda_t} \\ -\frac{\gamma X_t X_{t-1}}{\lambda_t^2} & \frac{-\gamma X_t X_{t-1}^2}{\lambda_t^2} & -X_{t-1} + \frac{X_t X_{t-1}}{\lambda_t} \\ -1 + \frac{X_t}{\lambda_t} & -X_{t-1} + \frac{X_t X_{t-1}}{\lambda_t} & -\frac{1}{2\gamma^2} \end{bmatrix}.$$

We can now calculate asymptotic standard errors making use of Equation (5.28).

6.2.2 INGARCH(1,1) Model

From Equation 5.2,

$$\lambda_t = \alpha_0 + \alpha_1 X_{t-1} + \beta_1 \lambda_{t-1}.$$

Set $\lambda_0 = \bar{X}$ and $\frac{\partial \lambda_0}{\partial \theta_i} = 0$.

Using Equation (5.25), we get the first derivatives of λ_t as the following.

$$\begin{aligned}
a_0 &= \frac{\partial \lambda_t}{\partial \alpha_0} = 1 + \beta_1 \frac{\partial \lambda_{t-1}}{\partial \alpha_0} \\
a_1 &= \frac{\partial \lambda_t}{\partial \alpha_1} = X_{t-1} + \beta_1 \frac{\partial \lambda_{t-1}}{\partial \alpha_1} \\
b_1 &= \frac{\partial \lambda_t}{\partial \beta_1} = \lambda_{t-1} + \beta_1 \frac{\partial \lambda_{t-1}}{\partial \beta_1}.
\end{aligned} \tag{6.1}$$

The second derivatives of λ_t are

$$\begin{aligned}
\frac{\partial^2 \lambda_t}{\partial \alpha_0^2} &= \frac{\partial}{\partial \alpha_0} \left(\frac{\partial \lambda_t}{\partial \alpha_0} \right) = 0 \\
\frac{\partial^2 \lambda_t}{\partial \alpha_0 \partial \alpha_1} &= \frac{\partial}{\partial \alpha_1} \left(\frac{\partial \lambda_t}{\partial \alpha_0} \right) = 0 \\
\frac{\partial^2 \lambda_t}{\partial \alpha_0 \partial \beta_1} &= \frac{\partial}{\partial \beta_1} \left(\frac{\partial \lambda_t}{\partial \alpha_0} \right) = \beta_1 \frac{\partial^2 \lambda_{t-1}}{\partial \alpha_0 \partial \beta_1} + \frac{\partial \lambda_{t-1}}{\partial \alpha_0} \\
\frac{\partial^2 \lambda_t}{\partial \alpha_1 \partial \alpha_0} &= \frac{\partial}{\partial \alpha_0} \left(\frac{\partial \lambda_t}{\partial \alpha_1} \right) = 0 \\
\frac{\partial^2 \lambda_t}{\partial \alpha_1^2} &= \frac{\partial}{\partial \alpha_1} \left(\frac{\partial \lambda_t}{\partial \alpha_1} \right) = 0 \\
\frac{\partial^2 \lambda_t}{\partial \alpha_1 \partial \beta_1} &= \frac{\partial}{\partial \beta_1} \left(\frac{\partial \lambda_t}{\partial \alpha_1} \right) = \beta_1 \frac{\partial^2 \lambda_{t-1}}{\partial \alpha_1 \partial \beta_1} + \frac{\partial \lambda_{t-1}}{\partial \alpha_1} \\
\frac{\partial^2 \lambda_t}{\partial \beta_1 \partial \alpha_0} &= \frac{\partial}{\partial \alpha_0} \left(\frac{\partial \lambda_t}{\partial \beta_1} \right) = \beta_1 \frac{\partial^2 \lambda_{t-1}}{\partial \alpha_0 \partial \beta_1} + \frac{\partial \lambda_{t-1}}{\partial \alpha_0} \\
\frac{\partial^2 \lambda_t}{\partial \beta_1 \partial \alpha_1} &= \frac{\partial}{\partial \alpha_1} \left(\frac{\partial \lambda_t}{\partial \beta_1} \right) = \beta_1 \frac{\partial^2 \lambda_{t-1}}{\partial \alpha_1 \partial \beta_1} + \frac{\partial \lambda_{t-1}}{\partial \alpha_1} \\
\frac{\partial^2 \lambda_t}{\partial \beta_1^2} &= \frac{\partial}{\partial \beta_1} \left(\frac{\partial \lambda_t}{\partial \beta_1} \right) = \beta_1 \frac{\partial^2 \lambda_{t-1}}{\partial \beta_1^2} + 2 \frac{\partial \lambda_{t-1}}{\partial \beta_1}.
\end{aligned} \tag{6.2}$$

6.2.2.1 Negative Binomial

We know

$$l_t(\theta) = X_t \log \lambda_t - (r + X_t) \log(1 + \lambda_t) + \log((X_t + r - 1)!) - \log((r - 1)!) - \log(X_t!).$$

Making use of Equation (5.23) , we get

$$\begin{aligned} \frac{\partial l_t}{\partial \theta} &= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right) \frac{\partial \lambda_t}{\partial \theta} \\ &= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right) \begin{bmatrix} a_0 \\ a_1 \\ b_1 \end{bmatrix} \\ \frac{\partial l_t}{\partial \theta^T} &= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right) \begin{bmatrix} a_0 & a_1 & b_1 \end{bmatrix} \\ \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta^T} &= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right)^2 \begin{bmatrix} a_0^2 & a_0 a_1 & a_0 b_1 \\ a_0 a_1 & a_1^2 & a_1 b_1 \\ a_0 b_1 & a_1 b_1 & b_1^2 \end{bmatrix}, \end{aligned}$$

and using Equation (5.24), we get

$$\begin{aligned} \frac{\partial^2 l_t}{\partial \theta \partial \theta^T} &= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right) \frac{\partial \lambda_t}{\partial \theta \partial \theta^T} - \left(\frac{X_t}{\lambda_t^2} - \frac{r + X_t}{(1 + \lambda_t)^2} \right) \frac{\partial^2 \lambda_t}{\partial \theta} \frac{\partial \lambda_t}{\partial \theta^T} \\ &= \left(\frac{X_t}{\lambda_t} - \frac{r + X_t}{1 + \lambda_t} \right) \begin{bmatrix} \frac{\partial^2 \lambda_t}{\partial \alpha_0^2} & \frac{\partial^2 \lambda_t}{\partial \alpha_0 \partial \alpha_1} & \frac{\partial^2 \lambda_t}{\partial \alpha_0 \partial \beta_1} \\ \frac{\partial^2 \lambda_t}{\partial \alpha_1 \partial \alpha_0} & \frac{\partial^2 \lambda_t}{\partial \alpha_1^2} & \frac{\partial^2 \lambda_t}{\partial \alpha_1 \partial \beta_1} \\ \frac{\partial^2 \lambda_t}{\partial \beta_1 \partial \alpha_0} & \frac{\partial^2 \lambda_t}{\partial \beta_1 \partial \alpha_1} & \frac{\partial^2 \lambda_t}{\partial \beta_1^2} \end{bmatrix} - \left(\frac{X_t}{\lambda_t^2} - \frac{r + X_t}{(1 + \lambda_t)^2} \right) \begin{bmatrix} a_0^2 & a_0 a_1 & a_0 b_1 \\ a_0 a_1 & a_1^2 & a_1 b_1 \\ a_0 b_1 & a_1 b_1 & b_1^2 \end{bmatrix} \end{aligned}$$

So, now with the use of Equation (5.28), we can calculate the asymptotic standard errors.

6.2.2.2 Poisson

For Poisson INGARCH model, we know

$$l_t(\theta) = X_t \log \lambda_t - \lambda_t - \log(X_t!).$$

Again, using Equation (5.23) and Equation (5.24), we get the following expressions.

$$\frac{\partial l_t}{\partial \theta} = \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \theta} = \left(\frac{X_t}{\lambda_t} - 1 \right) \begin{bmatrix} a_0 \\ a_1 \\ b_1 \end{bmatrix}$$

$$\frac{\partial l_t}{\partial \theta^T} = \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \theta} = \left(\frac{X_t}{\lambda_t} - 1 \right) \begin{bmatrix} a_0 & a_1 & b_1 \end{bmatrix}$$

$$\frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta^T} = \left(\frac{X_t}{\lambda_t} - 1 \right)^2 \begin{bmatrix} a_0^2 & a_0 a_1 & a_0 b_1 \\ a_0 a_1 & a_1^2 & a_1 b_1 \\ a_0 b_1 & a_1 b_1 & b_1^2 \end{bmatrix}$$

$$\begin{aligned} \frac{\partial^2 l_t}{\partial \theta \partial \theta^T} &= \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial^2 \lambda_t}{\partial \theta \partial \theta^T} - \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \theta} \frac{\partial \lambda_t}{\partial \theta^T} \\ &= \left(\frac{X_t}{\lambda_t} - 1 \right) \begin{bmatrix} \frac{\partial^2 \lambda_t}{\partial \alpha_0^2} & \frac{\partial^2 \lambda_t}{\partial \alpha_0 \partial \alpha_1} & \frac{\partial^2 \lambda_t}{\partial \alpha_0 \partial \beta_1} \\ \frac{\partial^2 \lambda_t}{\partial \alpha_1 \partial \alpha_0} & \frac{\partial^2 \lambda_t}{\partial \alpha_1^2} & \frac{\partial^2 \lambda_t}{\partial \alpha_1 \partial \beta_1} \\ \frac{\partial^2 \lambda_t}{\partial \beta_1 \partial \alpha_0} & \frac{\partial^2 \lambda_t}{\partial \beta_1 \partial \alpha_1} & \frac{\partial^2 \lambda_t}{\partial \beta_1^2} \end{bmatrix} - \left(\frac{X_t}{\lambda_t^2} \right) \begin{bmatrix} a_0^2 & a_0 a_1 & a_0 b_1 \\ a_0 a_1 & a_1^2 & a_1 b_1 \\ a_0 b_1 & a_1 b_1 & b_1^2 \end{bmatrix}. \end{aligned}$$

So, using Equation (5.28), the asymptotic standard errors can be determined.

6.2.2.3 Double Poisson

For Double Poisson INGARCH model, we know

$$l_t(\theta) = \frac{1}{2} \log \gamma - \gamma \lambda_t - (1 - \gamma) X_t + (1 - \gamma) X_t \log X_t + \gamma X_t \log \lambda_t - \log(X_t!).$$

Differentiating l_t with respect to each parameter, we get

$$\begin{aligned}\frac{\partial l_t}{\partial \alpha_0} &= -\gamma \frac{\partial \lambda_t}{\partial \alpha_0} + \frac{\gamma X_t}{\lambda_t} \frac{\partial \lambda_t}{\partial \alpha_0} = \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \alpha_0} \\ \frac{\partial l_t}{\partial \alpha_1} &= -\gamma \frac{\partial \lambda_t}{\partial \alpha_1} + \frac{\gamma X_t}{\lambda_t} \frac{\partial \lambda_t}{\partial \alpha_1} = \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \alpha_1} \\ \frac{\partial l_t}{\partial \beta_1} &= -\gamma \frac{\partial \lambda_t}{\partial \beta_1} + \frac{\gamma X_t}{\lambda_t} \frac{\partial \lambda_t}{\partial \beta_1} = \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \beta_1} \\ \frac{\partial l_t}{\partial \gamma} &= \frac{1}{2\gamma} - \lambda_t + X_t - X_t \log X_t + X_t \log \lambda_t.\end{aligned}$$

So

$$\frac{\partial l_t}{\partial \theta} = \begin{bmatrix} \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \alpha_0} \\ \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \alpha_1} \\ \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \beta_1} \\ \frac{1}{2\gamma} - \lambda_t + X_t - X_t \log X_t + X_t \log \lambda_t \end{bmatrix},$$

$$\frac{\partial l_t}{\partial \theta^T} = \left[-\gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \alpha_0} \quad \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \alpha_1} \quad \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \beta_1} \quad \frac{1}{2\gamma} - \lambda_t + X_t - X_t \log X_t + X_t \log \lambda_t \right].$$

Multiplying the above two matrices, we obtain $\frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta^T}$.

Now, the second derivatives of l_t are

$$\begin{aligned}\frac{\partial^2 l_t}{\partial \alpha_0^2} &= \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial^2 \lambda_t}{\partial \alpha_0^2} - \gamma \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \alpha_0} \frac{\partial \lambda_t}{\partial \alpha_0} \\ &= -\gamma \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \alpha_0} \frac{\partial \lambda_t}{\partial \alpha_0} \\ \frac{\partial^2 l_t}{\partial \alpha_0 \partial \alpha_1} &= \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial^2 \lambda_t}{\partial \alpha_0 \partial \alpha_1} - \gamma \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \alpha_0} \frac{\partial \lambda_t}{\partial \alpha_1} \\ &= -\gamma \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \alpha_0} \frac{\partial \lambda_t}{\partial \alpha_1} \\ \frac{\partial^2 l_t}{\partial \alpha_0 \partial \beta_1} &= \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial^2 \lambda_t}{\partial \alpha_0 \partial \beta_1} - \gamma \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \alpha_0} \frac{\partial \lambda_t}{\partial \beta_1}\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l_t}{\partial \alpha_1^2} &= \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial^2 \lambda_t}{\partial \alpha_1^2} - \gamma \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \alpha_1} \frac{\partial \lambda_t}{\partial \alpha_1} \\
&= -\gamma \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \alpha_1} \frac{\partial \lambda_t}{\partial \alpha_1} \\
\frac{\partial^2 l_t}{\partial \alpha_1 \partial \beta_1} &= \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial^2 \lambda_t}{\partial \alpha_1 \partial \beta_1} - \gamma \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \alpha_1} \frac{\partial \lambda_t}{\partial \beta_1} \\
\frac{\partial^2 l_t}{\partial \beta_1^2} &= \gamma \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial^2 \lambda_t}{\partial \beta_1^2} - \gamma \left(\frac{X_t}{\lambda_t^2} \right) \frac{\partial \lambda_t}{\partial \beta_1} \frac{\partial \lambda_t}{\partial \beta_1} \\
\frac{\partial^2 l_t}{\partial \beta_1 \partial \gamma} &= \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \beta_1} \\
\frac{\partial^2 l_t}{\partial \gamma^2} &= -\frac{1}{2\gamma^2}.
\end{aligned}$$

So the Hessian matrix (H) is obtained as the following.

$$H = \frac{\partial^2 l_t}{\partial \theta \partial \theta^T} = \begin{bmatrix} \frac{\partial^2 l_t}{\partial \alpha_0^2} & \frac{\partial^2 l_t}{\partial \alpha_0 \partial \alpha_1} & \frac{\partial^2 l_t}{\partial \alpha_0 \partial \beta_1} & \frac{\partial^2 l_t}{\partial \alpha_0 \partial \gamma} \\ \frac{\partial^2 l_t}{\partial \alpha_0 \partial \alpha_1} & \frac{\partial^2 l_t}{\partial \alpha_1^2} & \frac{\partial^2 l_t}{\partial \alpha_1 \partial \beta_1} & \frac{\partial^2 l_t}{\partial \alpha_1 \partial \gamma} \\ \frac{\partial^2 l_t}{\partial \alpha_0 \partial \beta_1} & \frac{\partial^2 l_t}{\partial \alpha_1 \partial \beta_1} & \frac{\partial^2 l_t}{\partial \beta_1^2} & \frac{\partial^2 l_t}{\partial \beta_1 \partial \gamma} \\ \frac{\partial^2 l_t}{\partial \alpha_0 \partial \gamma} & \frac{\partial^2 l_t}{\partial \alpha_1 \partial \gamma} & \frac{\partial^2 l_t}{\partial \beta_1 \partial \gamma} & \frac{\partial^2 l_t}{\partial \gamma^2} \end{bmatrix}$$

Now, using Equation 5.28, asymptotic standard errors can be calculated.

To study the goodness of fit of the NBINARCH(1) model, analysis of residuals from the fitted model is considered. Pearson residual method defined as

$$r_{1t} = \frac{X_t - \hat{r}\hat{\lambda}_t}{[\hat{r}\hat{\lambda}_t(1 + \hat{\lambda}_t)]^{1/2}},$$

where $\hat{\lambda}_t = \hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}$, is used. The ACF, PACF, normal Q-Q and kernel density plots are shown in Figure 6.2. There is no correlation within the residuals from the plots of

ACF and PACF and is supported by the Ljung-Box statistics of 13.20877 based on $k=15$ lags. The p-value is 0.51 with degree of freedom of $k - 1 = 14$ of chi-square distribution. Ljung-Box statistics is calculated using the formula

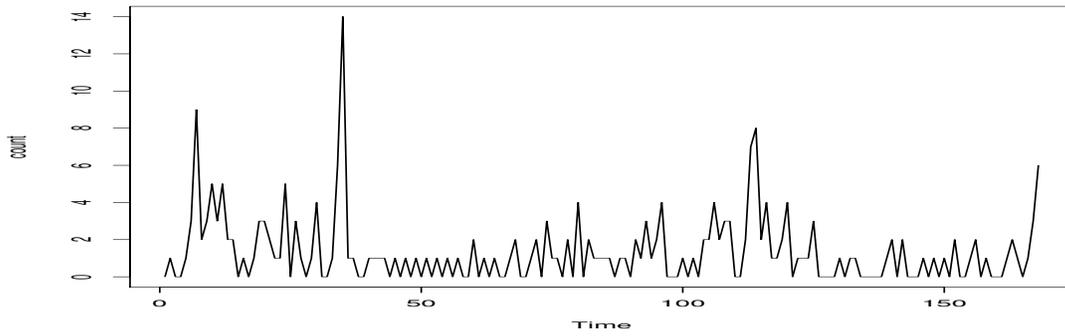
$$Q = n(n + 2) \sum_{i=1}^k \frac{\hat{r}_i^2}{n - i},$$

where \hat{r} is the residual ACF's. The Q-Q plot and kernel plot indicates that the residuals are highly non-normally distributed and is supported by Shapiro-Wilk normality test with a p-value of 4.861×10^{-12} .

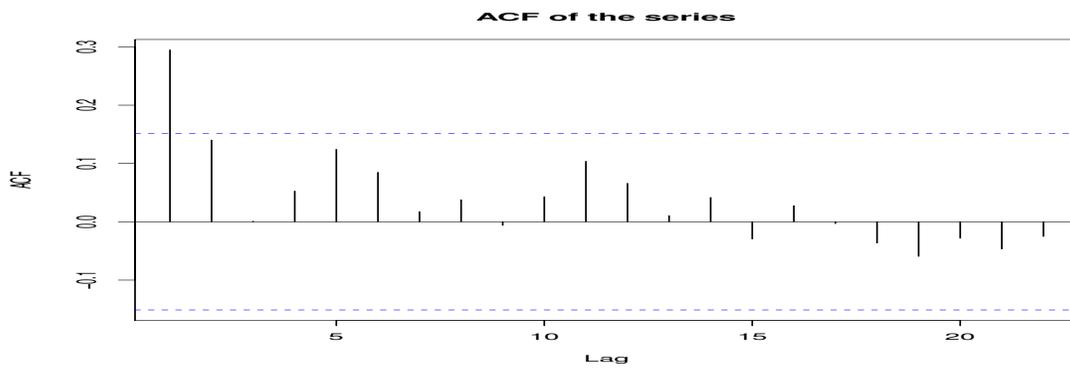
We also consider the normalized conditional randomized quantile residuals of Dunn and Smyth (1996) and Benjamin et.al (2003). The normalized randomized quantile residual is given by $r_{2t} = \Phi^{-1}(u_t)$, where Φ^{-1} is the inverse cumulative distribution function of a standard normal variable and u_t is a random value from the uniform distribution in the interval $[F(X_t - 1, \hat{\alpha}, \hat{r}), F(X_t, \hat{\alpha}, \hat{r})]$, where $F(X_t, \hat{\alpha}_0, \hat{r})$ is the fitted conditional NB cumulative distribution function. The corresponding residual analysis is given in Figure 6.3. We see that there is no correlation within the residuals and is supported by the Ljung-Box statistics of 15.72503 based on 15 lags. The p-value is 0.33 with degree of freedom of 14 of chi-square distribution. The Q-Q plot and kernel plots appears to be normally distributed and is supported by Shapiro-Wilk normality test with a p-value of 0.8972. Figure 6.4 shows a scatter plot of the normalized randomized quantile residuals against time. All of these indicates that the chosen model provides an adequate fit to the data.

Table 6.1: Simulated Results for NBINGARCH models

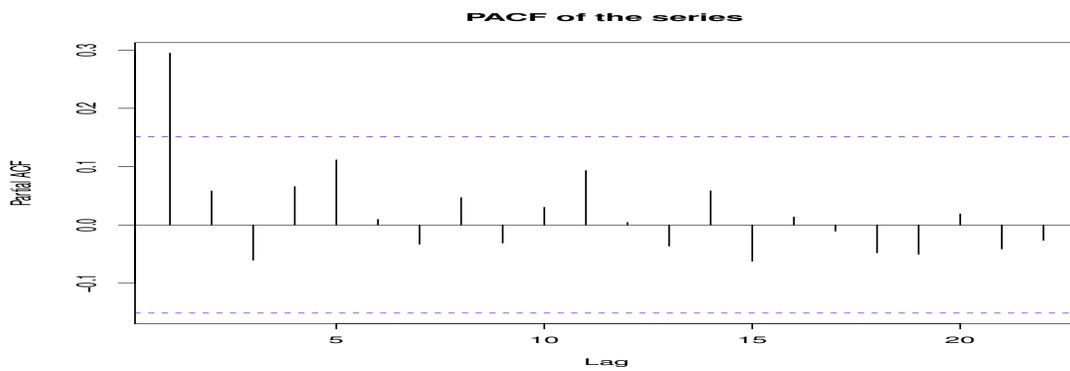
Model	n	Method	α_0 (Error)	α_1 (Error)	α_2 or β_1 (Error)
A1	100	YW	2.184435 (0.356641)	0.319169 (0.131690)	
		CLS	2.186407 (0.358673)	0.322384 (0.131318)	
		MLE	2.024442 (0.286056)	0.366937 (0.101582)	
	500	YW	2.070921 (0.193466)	0.373337 (0.071208)	
		CLS	2.071400 (0.194187)	0.374046 (0.071140)	
		MLE	1.996345 (0.118478)	0.397673 (0.049042)	
A2	100	YW	5.840418 (1.068042)	0.249448 (0.060471)	
		CLS	5.804475 (1.084485)	0.254096 (0.059203)	
		MLE	5.148558 (0.715489)	0.282398 (0.048495)	
	500	YW	5.334206 (0.700325)	0.283283 (0.037134)	
		CLS	5.336073 (0.703978)	0.283783 (0.036976)	
		MLE	5.068856 (0.355454)	0.294157 (0.020889)	
B1	100	YW	2.397071 (0.538210)	0.236929 (0.124910)	0.129244 (0.115791)
		CLS	2.406549 (0.554631)	0.235885 (0.127317)	0.134176 (0.113802)
		MLE	2.135391 (0.431835)	0.274392 (0.105936)	0.174367 (0.106789)
	500	YW	2.136410 (0.298498)	0.187440 (0.039690)	0.185613 (0.065613)
		CLS	2.138010 (0.298928)	0.276881 (0.067319)	0.186501 (0.065580)
		MLE	2.021851 (0.204806)	0.303536 (0.048611)	0.197870 (0.047137)
B2	100	YW	4.851662 (1.939932)	0.160993 (0.064587)	0.147203 (0.073094)
		CLS	4.851705 (1.957805)	0.159556 (0.066007)	0.155016 (0.069746)
		MLE	2.857774 (1.598617)	0.135756 (0.086554)	0.133762 (0.087672)
	500	YW	3.913638 (0.981200)	0.183339 (0.040213)	0.178599 (0.042745)
		CLS	3.918509 (0.989420)	0.183033 (0.040583)	0.179969 (0.042170)
		MLE	2.636327 (0.893688)	0.162735 (0.049617)	0.161340 (0.052444)
C1	100	CLS	3.846397 (2.068827)	0.370303 (0.182733)	0.175333 (0.245783)
		MLE	2.414842 (0.936882)	0.477183 (0.156486)	0.204068 (0.208017)
	500	CLS	2.830926 (1.001288)	0.427911 (0.110880)	0.266000 (0.135775)
		MLE	2.169787 (0.477650)	0.510534 (0.072072)	0.310405 (0.155190)
C2	100	CLS	3.931298 (1.355761)	0.173914 (0.055244)	0.101130 (0.219163)
		MLE	3.197880 (0.913585)	0.185840 (0.082177)	0.179380 (0.130755)
	500	CLS	3.293011 (0.625514)	0.195852 (0.029590)	0.166294 (0.097891)
		MLE	3.146980 (0.513980)	0.198687 (0.022328)	0.179760 (0.077436)



(a)



(b)



(c)

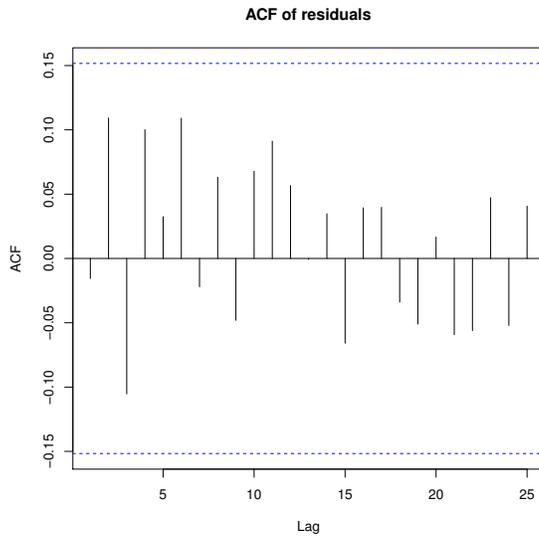
Figure 6.1: Polio data. (a) original series (b) ACF (c) PACF

Table 6.2: AIC and BIC for different values of r with NBINARCH(1) and NBINGARCH(1,1) models.

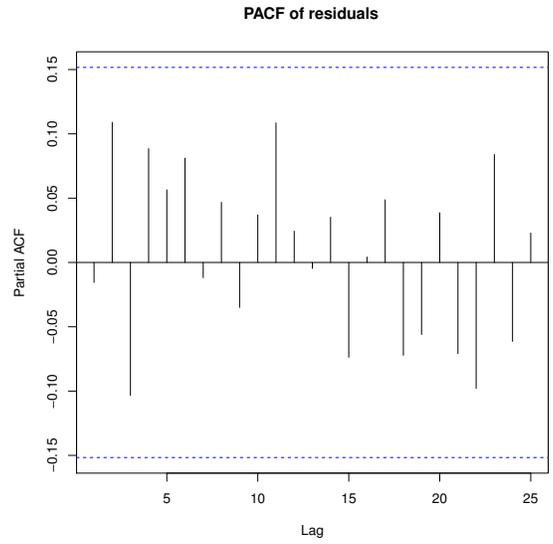
Model		r				
		1	2	3	4	5
NBINARCH(1)	AIC	523.5978	520.5613	524.4632	528.5592	532.1036
	BIC	532.9697	529.9332	533.8351	537.9311	541.4755
NBINGARCH(1,1)	AIC	529.8315	528.3122	532.8693	537.3188	615.9582
	BIC	539.2034	537.6841	542.2412	546.6907	625.3301

Table 6.3: Parameter estimates with NB, P and DP INARCH(1) and INGARCH(1,1) models.

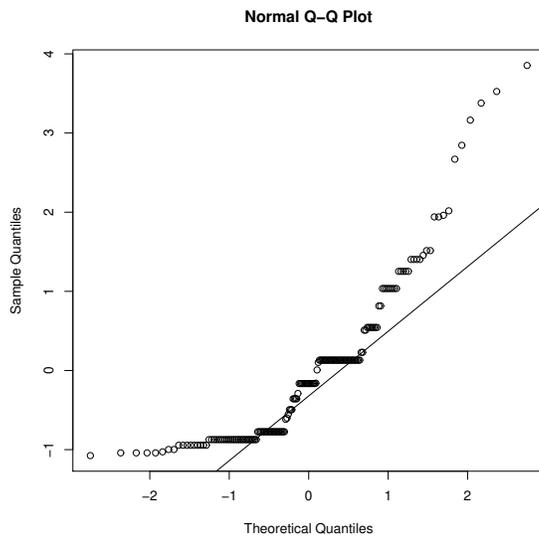
Model		$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	\hat{r} or $\hat{\gamma}$	AIC	BIC
		(SE)	(SE)	(SE)	(SE)		
INARCH(1)	NB	0.427734 (0.002645)	0.188481 (0.003417)		2	520.5613	529.9332
	P	0.865626 (0.013139)	0.364406 (0.016898)			562.2899	568.5379
	DP	0.865626 (0.013116)	0.364406 (0.016854)		0.554362 (0.003327)	528.5624	537.9343
INGARCH(1,1)	NB	0.311908 (0.006358)	0.184325 (0.003514)	0.181478 (0.017588)	2	521.0778	533.5737
	P	0.635683 (0.034461)	0.351473 (0.018544)	0.184559 (0.027449)		562.0793	571.4512
	DP	0.635683 (0.036136)	0.351473 (0.018577)	0.184559 (0.028609)	0.559420 (0.003724)	529.3326	541.8385



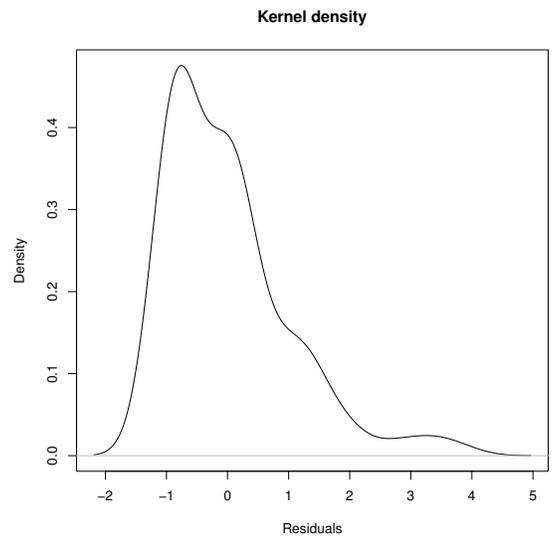
(a)



(b)

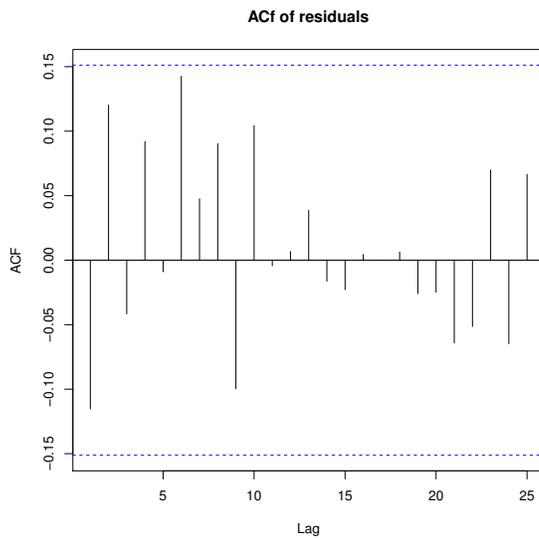


(c)

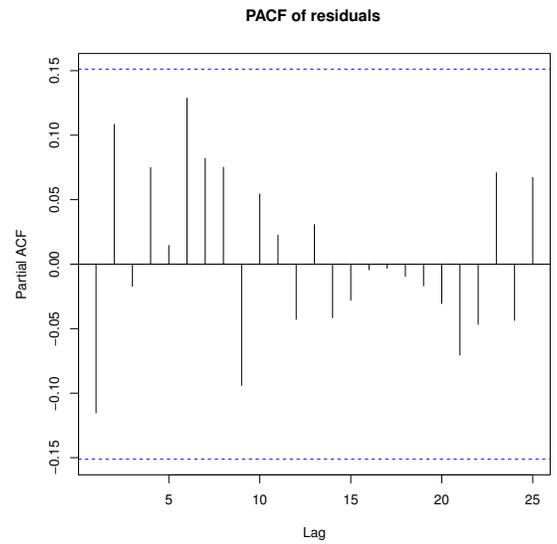


(d)

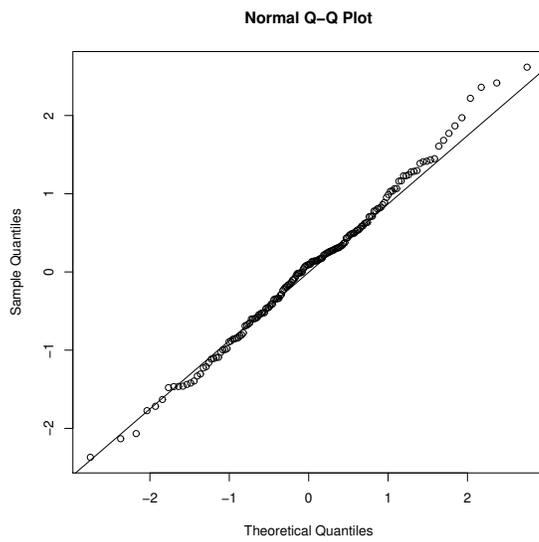
Figure 6.2: Pearson residuals. (a) ACF (b) PACF (c) Q-Q plot (d) Kernel density



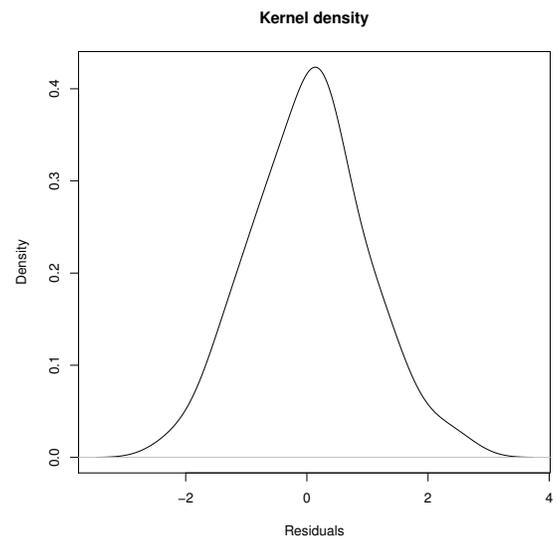
(a)



(b)



(c)



(d)

Figure 6.3: Randomized quantile residuals. (a) ACF (b) PACF (c) Q-Q plot (d) Kernel density

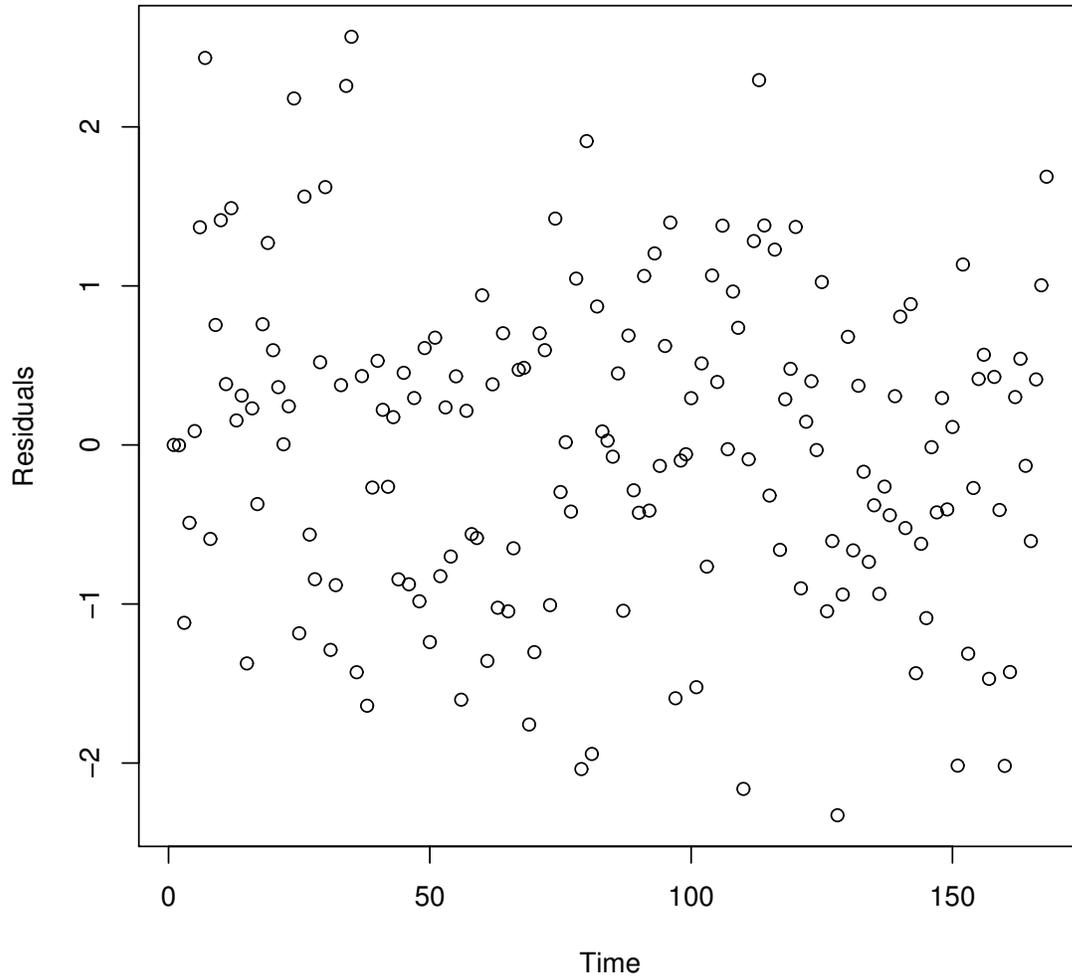


Figure 6.4: Normalized randomized quantile residual

REFERENCES

- [1] Chan, Kung-Sik and D. Cryer, Jonathan (2009). *Time series analysis with applications in R*, 2nd edition, Springer.
- [2] Tsay, Ruey S (2005). *Analysis of financial time series*, 2nd edition, John Wiley & Sons.
- [3] Zhu, Fukang (2011). *A negative binomial integer-valued GARCH model*, Journal of Time Series Analysis, vol. 32, p.(54-64), Wiley Online Library.
- [4] Ferland, Ren and Latour, Alain and Oraichi, Driss (2006). *Integer Valued GARCH Process*, Journal of Time Series Analysis, vol. 27, p.(923-942), Wiley Online Library.
- [5] Benjamin, Michael A and Rigby, Robert A and Stasinopoulos, D Mikis (2003). *Generalized autoregressive moving average models*, Journal of the American Statistical association, vol.98, p.(214-223), Taylor & Francis.
- [6] Davis, Richard A and Wu, Rongning (2009). *A negative binomial model for time series of counts*, Biometrika, vol. 96, p.(735-749), Biometrika Trust

VITA

Graduate School
Southern Illinois University

Kezang Choden C

email address (ksngcee@gmail.com)

Sherubtse College at Kanglung, Bhutan
Bachelor of Science, General, December 2004

Thesis Title:
INTEGER-VALUED ARCH AND GARCH MODELS

Major Professor: Dr. S. Yaser Samadi