POISSON INTEGER-VALUED GARCH MODEL

STRUCTURE, PARAMETER ESTIMATION
AND A REAL-DATA EXAMPLE

By

Twiza Aggie Namukwanya

A dissertation submitted in partial fulfillment of the requirements for the degree of Master of Science in Statistics

UNIVERSITY OF ZAMBIA
LUSAKA

2023

COPYRIGHT

DECLARATION

The work described in this Master of Science (MSc) dissertation was carried out under the supervision of Dr. John Musonda, Department of Mathematics and Statistics, School of Natural Sciences, University of Zambia.

The MSc dissertation represents original work by the author and has not otherwise been submitted in any form for any degree or diploma to any other University. Where use has been made of the work of others it is duly acknowledged in the text.

Name of Student: ....

Computer number ....

Signature ....

Date ....

APPROVAL

This dissertation of Twiza Namukwanya has been approved as fulfilling the requirements or partially fulfilling the requirements for the award of Master of Science in Mathematics by the University of Zambia.

Examiner 1: ....

Signature:..

Date:...

Examiner 2: ....

Signature:..

Date:...

Examiner 3: ....

Signature:..

Date:...

Chairperson of Board of Examiners: ....

Signature:..

Date:...

Supervisor: ....

Signature:...

Date...

ABSTRACT

In this dissertation, we present an integer-valued generalised autoregressive conditional heteroscedastic (INGARCH)(p, q) model based on the Poisson distribution. This model proves to be effective for modelling overdispersed integer-valued time series with conditional heteroscedasticity. The basic properties of the model are studied, and a condition for the existence of such a process is given. For the case $p = 1$, $q = 1$, it is explicitly shown that an INGARCH process is a standard autoregressive moving average (ARMA)$(1, 1)$ process. An approach for parameter estimation of the model is given using the method of conditional maximum likelihood estimation. In cases where analytical results from the likelihood function are not found, we suggest numerical optimisation methods, particularly using the R software package tscount.

The model is illustrated on an overdispersed integer-valued time series. Specifically, we analyse the daily number of shares traded on the Lusaka Securities Exchange (LuSE) from October 1, 2021, to May 10, 2022 (excluding weekends and public holidays), comprising a total of 150 observations. This data set is publicly available on the LuSE website at https://luse.co.zm/market-data. The empirical mean and variance of the data are 2.48287 and 5.76448 respectively, indicating overdispersion in the marginal distribution of the data.

In addition to fitting the INGARCH model, we assess its predictive performance using one-step-ahead forecasts and compute the corresponding confidence intervals. The model captures both the volatility clustering and the overdispersion present in the data. The forecasts closely match the actual data, providing reliable predictions for future observations.

Keywords. Integer-valued time series; INGARCH model; heteroscedasticity; overdispersion; one-step-ahead forecasts; predictive accuracy.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my supervisor, Dr. John Musonda, for his invaluable guidance, expertise, and continuous support throughout the research process. His insightful feedback and constructive criticism have played a crucial role in shaping the content and improving the quality of this dissertation.

I extend my heartfelt thanks to my colleagues and peers who have provided valuable input, shared their expertise, and engaged in fruitful discussions. Their diverse perspectives and thoughtful insights have broadened my understanding of the subject matter and enriched the overall quality of this work.

I also acknowledge the funding I received from the International Science Programmeme (ISP) through the Eastern African Universities Mathematics Programme (EAUMP) to enable me to do my studies. Special thanks go to the Zambian coordinator, Dr. Mervis Kikonko. I also thank Evelyn Hone College for the financial support that was provided for the research part of my studies.

I acknowledge the Department of Mathematics and Statistics, School of Natural Sciences at the University of Zambia for the study space, access to Maths Lab and internet facilities.

I am indebted to my friends and family for their unwavering support, encouragement, and belief in my abilities. Your love, understanding and patience have been a constant source of motivation and inspiration throughout this journey.

Although it is impossible to name everyone individually, I am sincerely grateful to all those who have played a part, big or small, in making this dissertation possible. Your support and encouragement have been invaluable and I am truly honoured to have had the opportunity to work on this dissertation.

Thank you for your support and belief in me.

Lusaka, December, 2023

Twiza Aggie Namukwanya

DEDICATION

This dissertation is dedicated to Dad and Mum, (Cletus Simukwanya and Rosemary Mwamba). The sky is indeed my playing ground.

TABLE OF CONTENTS

Copyright i

DECLARATION ii

APPROVAL iii

ABSTRACT iv

ACKNOWLEDGMENTS v

DEDICATION vi

TABLE OF CONTENTS viii

INDEX OF NOTATION ix

1 INTRODUCTION 1

- 1.1 Statement of the problem . . . 3
- 1.2 Aim of the study . . . 3
- 1.3 Research Objectives . . . 3
- 1.4 Research Questions . . . 3
- 1.5 Significance of the Study . . . 3
- 1.6 Literature Review . . . 3
- 1.7 Research Methodology . . . 5
- 1.8 Organisation of the dissertation . . . 7

2 PRELIMINARIES 8

- 2.1 Convergence in probability theory . . . 8
  - 2.1.1 Modes of convergence . . . 8
  - 2.1.2 Uniqueness of convergence . . . 11
  - 2.1.3 Relations between the modes of convergence . . . 11
- 2.2 Thinning operations . . . 12
  - 2.2.1 Binomial thinning operator . . . 12
  - 2.2.2 Alternative thinning concepts . . . 14

3 STRUCTURE OF THE PROCESS 16

- 3.1 INGARCH(p, q) process and a necessary existence condition . . . 16
- 3.2 Construction of the process . . . 19

3.3 Almost sure convergence of $X_t^{(n)}$ 21
3.4 Stationarity of the INGARCH process X_t 22
3.5 Mean-square convergence of $X_t^{(n)}$ 25
3.6 Conditional law of $\{X_t\}_{t\in\mathbb{Z}}$ given $\mathcal{F}_{t-1}$ 25
3.7 INGARCH(1,1) process 27

4 AN APPROACH FOR PARAMETER ESTIMATION **37**
4.1 Conditional maximum likelihood estimation 37
4.2 One-step-ahead forecast confidence interval 40

5 REAL-DATA EXAMPLE **41**

6 DISCUSSION **44**
6.1 On the structure of the process 44
6.2 On the approach for parameter estimation 47
6.3 On the real-data example 48

7 CONCLUSION AND RECOMMENDATIONS **51**
7.1 Conclusion 51
7.2 Recommendations 52

A Data Set for the Illustration **54**

B R Code for the Illustration **56**
B.1 Table 5.1: Summary Statistics 57
B.2 Figure 5.1: Plot of original data 58
B.3 Figure 5.2: SACF and SPACF 58
B.4 Table 5.2: Estimated parameter values 58
B.5 Figure 5.3: One-step-ahead forecasts 59

INDEX OF NOTATION

Below is a list of symbols that will be frequently used and a brief indication of their meaning.

$\mathbb{N}$	the set of positive integers
$\mathbb{N}_0$	the set of nonnegative integers
$\mathbb{R}$	the set of real numbers
$\mathbb{Q}$	the set of rational numbers
$\mathbb{E}$	the expectation of a variable
$GARCH$	generalised autoregressive conditional heteroscedastic model
$INGARCH$(p,q)	Integer-valued GARCH process of order p, q

CHAPTER 1

INTRODUCTION

Time-series analysis is a popular approach for analysing data that is collected over successive time intervals. In this approach, observations are arranged in chronological order, and the objective is to build mathematical models that can accurately represent the underlying stochastic process responsible for generating the data. These models not only help us understand the causal structure of the process, but also enable us to make reliable predictions about future behavior based on past patterns.

For a long time, time series modelling has predominantly focused on real-valued stochastic processes. Such models have been successful in many applications, ranging from economics to climate science. However, when dealing with data in the form of counts, known as integer-valued time series, traditional models may fail to address the peculiarities of this type of data.

In many real-world scenarios, data come in the form of counts, representing events or occurrences within specific time periods. Examples of integer-valued time series are prevalent in diverse domains, including website visits per hour, reported monthly insurance claims, or the number of defects in manufactured products. Due to the significant practical relevance and widespread occurrence of such data, there has been a growing interest in developing specialised methodologies to analyse integer-valued time series effectively.

Integer-valued time series exhibit distinct empirical characteristics that differentiate them from their real-valued counterparts. One of the key features of count data is overdispersion, which refers to the phenomenon in which the variance of the data is greater than the mean. Overdispersion is frequently observed in the count data, challenging traditional models that assume constant variance. Ignoring this characteristic can lead to unreliable results.

Another important characteristic of integer-valued time series is conditional heteroskedasticity, which refers to the non-constant conditional variance over time. In other words, the variance of the counts can change with the level of the count, introducing time-varying patterns in the data. These dynamic volatility patterns pose further challenges for modelling and require specialized techniques to be properly addressed. While the classical generalized autoregressive conditional heteroscedastic (GARCH) model has been successful in modelling real-valued time series with similar conditional heteroscedasticity, it may not be well suited for effectively capturing the volatility patterns in integer-valued time series. Due to its continuous nature (it is based on the normal distribution), the GARCH model might not fully account for the discrete nature of count data, and it may result in inappropriate parameter estimates and inaccurate predictions. In contrast, the integer-valued generalized autoregressive conditional heteroscedastic (INGARCH) model based on the Poisson distribution is designed to analyze time series of counts. Basically, this model is the generalised autoregressive conditional heteroscedastic (GARCH) model but with the normal distribution replaced with the Poisson distribution. The Poisson distribution is a discrete probability distribution that specifically models the probability of a given number of events occurring within a fixed interval of time or space. Its characteristics align well with count data, assuming independence between counts in non-overlapping intervals, making it a natural choice for modelling integer-valued time series. Thus, the INGARCH model can capture the unique characteristics of integer-valued time series data. Specifically, the INGARCH model considers the conditional variance of the time series to be influenced by the squared errors of previous observations, allowing it to account for the observed volatility clustering in integer-valued time series data. For a broader view of this active area of research, see [3, 14–27, 32–44].

The **Lusaka Securities Exchange (LuSE)** is Zambia's main stock exchange, established in 1993 to provide a platform for raising capital through the buying and selling of shares and other securities. It plays a vital role in the country's financial market by facilitating investment and supporting economic growth. LuSE is regulated by the *Securities and Exchange Commission (SEC)* of Zambia and operates under the *Securities Act*, adhering to international standards for transparency and investor protection. The exchange lists various financial instruments, including equities, bonds, and exchange-traded funds (ETFs), offering both domestic and international companies the opportunity to raise capital.

The LuSE All-Share Index (LASI) serves as a benchmark to evaluate the performance of listed companies and provides valuable information for investors and analysts. LuSE contributes significantly to the Zambian economy by mobilizing savings for investment, offering a platform for companies to access long-term financing, and promoting corporate governance and economic transparency. However, the exchange faces challenges such as low liquidity compared to other African stock exchanges and limited public awareness about securities trading.

Despite these challenges, LuSE remains a critical source of financial data, including stock prices, trading volumes, and market indices. This data is instrumental for testing financial and economic models, providing insights into market trends, risk assessments, and investor behaviour within Zambia's economy.

1.1. Statement of the problem

A commonly used model for overdispersed time series with conditional heteroscedasticity is the classical GARCH model. One limitation of this model is that it is based on a continuous distribution (the normal distribution), and this can lead to poor performance in the case where the data is integer-valued and scanty. In this dissertation, we address this limitation by proposing an integer-valued GARCH model based on a discrete distribution (the Poisson distribution) for this type of data.

1.2. Aim of the study

To present a Poisson INGARCH model for the effective modelling of overdispersed integer-valued time series data with conditional heteroscedasticity.

1.3. Research Objectives

(i) To formulate a Poisson INGARCH model and establish its basic properties/proofs behind the INGARCH model .

(ii) To give an approach for parameter estimation of a Poisson INGARCH model.

(iii) To illustrate the INGARCH model on an overdispersed integer-valued time series data.

1.4. Research Questions

(i) How can we formulate a Poisson INGARCH model and what are its basic properties?

(ii) How can estimate the parameters of a Poisson INGARCH model?

(iii) How can we illustrate the INGARCH model on an overdispersed count time series?

1.5. Significance of the Study

The study provides an effective tool for modelling overdispersed integer-valued time series data with conditional heteroscedasticity.

1.6. Literature Review

Time series of counts are commonly observed in real-world applications. This has led to a number of time series models being proposed to describe the marginal distribution and autocorrelation structure of this type of data, see [2, 19, 21, 24, 30].

The Poisson integer-valued GARCH (PIGARCH) model is a time series model that is used to model the conditional variance of a sequence of integer-valued random variables. The model was first proposed by Linton and Terrell [38] as an extension of the GARCH model to model count data. The PIGARCH model is based on the assumption that the conditional variance of the next observation in the sequence is a function of the squared errors of the previous observations. The Poisson distribution provides a standard framework for the analysis of count data, but the requirement that the variance should equal the mean is often too restrictive in practice. Frequently data are overdispersed, with the variance greater than the mean. Most literature on integer-valued models take as reference the modelling by the real-valued stochastic processes, namely the autoregressive moving average (ARMA) evolution. One of these approaches was proposed by Jacob and Lewis [9] who developed a discrete ARMA model using a mixture of a sequence of independent and identically distributed discrete random variables. Another way of obtaining models for integer-valued data involves replacing the usual multiplication in the standard ARMA models by a random operator, which preserves the discreteness of the process. This operator was introduced in [13] as the binomial thinning and leads to the family of integer-valued ARMA models. WeiB [14] provides a more recent review of a broad variety of such thinning operations. Alternatively to the thinning operator, Kachour et al. [10] used the rounding off operator to the nearest integer to introduce the rounded integer-valued autoregressive model. Some integer-valued bilinear models have been also introduced in [2, 3]. Taking into account of conditional heteroscedasticity, Heinen [8] defined an autoregressive conditional Poisson (ACP) model by adapting the autoregressive conditional duration model of Engle and Russell [4] to the integer-valued case, assuming a conditional Poisson distribution.

This model has already received considerable study in the literature. In particular, it has been presented by Heinen [40] a first-order stationarity condition of the model for any orders p and q, and the corresponding variance and autocorrelation function for the particular case $p = q = 1$. Ferland et al. [25] extended the studies of this model establishing a condition for the existence of a strictly stationary process which has finite first and second-order moments and deduced the maximum likelihood parameters estimators. They also stated a condition under which all moments of an INGARCH(1,1) model are finite. Fokianos et al. [26] considered likelihood-based inference when $p = q = 1$ using a perturbed version of the model. Weiß [43] derived a set of equations from which the variance and the autocorrelation function of the general model can be obtained. Neumann [56], Davis and Liu [16] and Christou and Fokianos [13] discussed some aspects related to the ergodicity. For the INARCH(p) model, Zhu and Wang [45] derived conditional weighted least squares estimators of the parameters and presented a test for conditional heteroscedasticity. Given the simple structure and the practical relevance of the INARCH(1) process, Weiß [43, 44] studied its properties in more detail. He Characterised the stationary marginal distribution in terms of its cumulants, showed how to approximate its marginal process distribution via the Poisson-Charlier expansion and calculated its higher-order moments and jumps. He also provided a conditional least squares approach for the estimation of its two parameters and constructed various simultaneous confidence regions. Because of its analogy with the standard GARCH model introduced by Bollerslev [1] in 1986, Ferland et al. [5] suggested to denominate these models as integer-valued GARCH.

Bagestani et al. [2] used the PIGARCH model to model the volatility of stock prices in the Iranian stock market. They found that the model was able to capture the volatility clustering and heteroscedasticity that is often observed in stock prices.

Zhang et al. [33] used the PIGARCH model to model the volatility of exchange rates in the Chinese foreign exchange market. They found that the model was able to capture the volatility clustering and heteroscedasticity that is often observed in exchange rates.

Chen et al. [4] used the PIGARCH model to model the volatility of insurance claims in the United States. They found that the model was able to capture the volatility clustering and heteroscedasticity that is often observed in insurance claims.

These studies have shown that the PIGARCH model can be a useful tool for modelling the volatility of a variety of time series data.

1.7. Research Methodology

The formulation of our model is inspired by the definition of the standard GARCH model. In the GARCH model, the conditional variance is expressed as a linear function of the squared past values of the series. This particular specification captures the main stylized facts characterizing integer-valued series. At the same time, it is simple enough to allow for a complete study of the solutions. To ensure the existence of a weakly stationary solution, we introduce two polynomials in the backshift operator and ensure that their roots lie outside the unit circle.

The model is constructed by successive approximations, as an almost sure limit of a sequence of independent and identically distributed Poisson random variables. The sequence itself is obtained through a cascade of thinning operations. Strict stationarity and other relevant probabilistic properties of the building sequence (and thus of the process) are shown using probability generating functions.

The estimation procedure for the parameters is analogous to the one used for the traditional GARCH model (conditional maximum likelihood estimation). In the case that analytical results from the likelihood function cannot be found, we use numerical optimisation methods to find out the optimal values of the parameters. In particular, we utilise a faster and more flexible method based on the R software package tscount to estimate the parameters.

The model is illustrated on an overdispersed integer-valued time series. In particular, we study the daily number of shares trading on the Lusaka Securities Exchange (LuSE) from October 1, 2021 to May 10, 2022 (except for weekends and public holidays), giving a total of 150 observations. This is part of the data set given on the LuSE website and is available for download at https://luse.co.zm/market-data. The empirical mean and variance of the data are 2.48287 and 5.76448 respectively, indicating that the true marginal distribution of the data is indeed overdispersed.

1.8. Organisation of the dissertation

The rest of the dissertation is organised as follows;

Chapter 2. In this chapter, we provide the necessary preliminary information required to comprehend the content discussed in this dissertation. The style here is deliberately tense, since this chapter is intended as a reference rather than a systematic exposition.

Chapter 3. In this chapter we present a Poisson integer-valued generalized autoregressive conditional heteroscedastic model (a Poisson INGARCH model) and its basic properties. In the first section, we define the process and give a necessary condition for its existence. The second section briefly discusses the construction of the process. In the third section, we check for almost sure convergence of the building sequence $X_t^{(n)}$ of the model. In the fourth section, we discuss stationarity of the process. The fifth section discusses the concept of mean-square convergence of $X_t^{(n)}$ while the sixth section looks at the conditional law of the actual process. The last section looks at a particular case of the model, the INGARCH(1,1) process.

Chapter 4. This chapter, looks at the estimation procedure for the parameters of the INGARCH model. In particular, it shows that analytical estimates from the likelihood function cannot be found. Therefore, the chapter suggests the use of numerical optimisation methods. In particular, the R software package tscount.

Chapter 5. In this chapter, we illustrate the INGARCH model on an overdispersed integer-valued time series real data with conditional heteroscedasticity. That is, we study the daily number of shares trading on the Lusaka Securities Exchange (LuSE) from October 1, 2021 to March 10, 2022 (except for weekends and public holidays), giving a total of 150 observations.

Chapter 6. In this chapter, we delve into the core of our research findings, meticulously analysing the integer-valued generalised autoregressive conditional heteroscedastic (INGARCH) process. The chapter serves as the focal point where we unravel the intricacies of our model. Through a comprehensive exploration of its theoretical underpinnings, estimation techniques, and practical applications, we aim to unravel the complexities of this model and shed light on its relevance in overdispersed integer-valued time series analysis.

Chapter 7 In this final chapter, we summarise the main findings of our research and underscore their implications. We offer concrete recommendations based on our insights to guide future actions and initiatives. Ultimately, this chapter serves as a blueprint for leveraging our findings to enact positive change in the relevant domain.

CHAPTER 2

PRELIMINARIES

In this chapter, we provide the necessary preliminary information required to comprehend the content discussed in this dissertation. The style here is deliberately terse, since this chapter is intended as a reference rather than a systematic exposition.

2.1. Convergence in probability theory

2.1.1 Modes of convergence

There are several modes of convergence in probability theory. We discuss four of them here.

Definition 2.1.1. *The sequence of random variables X_n converges almost surely (a.s.) to the random variable X as $n \to \infty$ if and only if*

$$P(\{\omega : X_n(\omega) \to X(\omega) \text{ as } n \to \infty\}) = 1.$$

Notation: $X_n \xrightarrow{a.s.} X$ *as* $n \to \infty$.

Definition 2.1.2. *The sequence of random variables X_n converges in probability to the random variable X as $n \to \infty$ if and only if, for all $\epsilon > 0$,*

$$P(|X_n - X| > \epsilon) \to 0 \text{ as } n \to \infty.$$

Notation: $X_n \xrightarrow{p} X$ *as* $n \to \infty$.

Definition 2.1.3. *The sequence of random variables X_n converges in r-mean to the random variable X as $n \to \infty$ if and only if*

$$E|X_n - X|^r \to 0 \text{ as } n \to \infty.$$

Notation: $X_n \xrightarrow{r} X$ *as* $n \to \infty$.

Definition 2.1.4. *The sequence of random variables X_n converges in distribution to the random variable X as $n \to \infty$ if and only if*

$$F_{X_n}(x) \to F_X(x) \text{ as } n \to \infty \text{ for all } x \in C(F_X),$$

where $C(F_X) = \{x : F_X(x) \text{ is continuous at } x\}$ is the continuity set of F_X.

Notation: $X_n \xrightarrow{d} X$ *as* $n \to \infty$.

Remark 2.1.1. When dealing with almost-sure convergence, we consider every $\omega \in \Omega$ and check whether or not the real numbers $X_n(\omega)$ converge to the real number $X(\omega)$ as $n \to \infty$. We have almost-sure convergence if the ω-set for which there is convergence has probability 1 or, equivalently, if the ω-set for which we do not have convergence has probability 0. Almost-sure convergence is also called convergence with probability 1. This convergence is equivalent to almost everywhere convergence in Lebesgue measure.

Remark 2.1.2. Convergence in 2-mean ($r = 2$ in Definition 2.1.3) is usually called convergence in square mean (or mean-square convergence).

Remark 2.1.3. Note that in Definition 2.1.4, the random variables are present only in terms of their distribution functions. Thus, they need not be defined on the same probability space.

Remark 2.1.4. We will permit ourselves the convenient abuse of notation such as $X_n \xrightarrow{d} N(0,1)$ or $X_n \xrightarrow{d} \text{Po}(\lambda)$ as $n \to \infty$ instead of the formally more correct, but lengthier $X_n \xrightarrow{d} X$ as $n \to 1$, where $X \sim N(0,1)$, and $X_n \xrightarrow{d} X$ as $n \to \infty$, where $X \sim \text{Po}(\lambda)$, respectively.

Remark 2.1.5. One can show that a distribution function has at most only a countable number of discontinuities. As a consequence, $C(F_X)$ equals the whole real line except, possibly, for at most a countable number of points.

We end this section with some examples.

Example 2.1.1. Let X_n be a random variable with support on the interval $\left(-\frac{1}{n}, 1\right)$, and assume that X_n has mean $E[X_n] = 1$ and variance $\text{Var}(X_n) = \frac{1}{n}$. Show that $X_n \xrightarrow{p} 1$ as $n \to \infty$.

Solution:

We want to show that X_n converges to 1 in probability, i.e., for all $\epsilon > 0$,

$$P(|X_n - 1| > \epsilon) \to 0 \quad \text{as} \quad n \to \infty.$$

First, by assumption, $E[X_n] = 1$ and $\text{Var}(X_n) = \frac{1}{n}$. Now, applying Chebyshev's inequality, we have:

$$P(|X_n - 1| > \epsilon) \leq \frac{\text{Var}(X_n)}{\epsilon^2} = \frac{1/n}{\epsilon^2} = \frac{1}{n\epsilon^2}.$$

As $n \to \infty$, we see that

$$P(|X_n - 1| > \epsilon) \leq \frac{1}{n\epsilon^2} \to 0.$$

Therefore, $X_n \xrightarrow{p} 1$ as $n \to \infty$.

Example 2.1.2. Let $X_1, X_2, \ldots$ be independent random variables with common distribution function

$$F(x) = \begin{cases} 1 - \left(\frac{1}{x}\right)^{\lambda}, & \text{for } x > 1, \lambda > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Define $Y_n = n^{-1/\lambda} \cdot \max_{1\le k\le n} X_k$, $n \ge 1$. Show that Y_n converges in distribution as $n \to \infty$, and determine the limit distribution.

Solution:

To solve the problem, let's first compute the common cumulative distribution function (CDF) for the random variables X_k. We are given that:

$$F(x) = 1 - \left(\frac{1}{x}\right)^{\lambda} \quad \text{for } x > 1.$$

The corresponding probability density function (PDF) is:

$$f(x) = \frac{d}{dx}F(x) = \frac{\lambda}{x^{\lambda+1}}, \quad \text{for } x > 1.$$

Now, we consider $Y_n = n^{-1/\lambda} \cdot \max_{1\le k\le n} X_k$. We want to find the limiting distribution of Y_n as $n \to \infty$.

The CDF of Y_n, denoted as $F_{Y_n}(x)$, is the probability that $Y_n \le x$, which is the same as:

$$F_{Y_n}(x) = P\left(Y_n \le x\right) = P\left(n^{-1/\lambda} \cdot \max_{1\le k\le n} X_k \le x\right).$$

This can be rewritten as:

$$F_{Y_n}(x) = P\left(\max_{1\le k\le n} X_k \le n^{1/\lambda} \cdot x\right).$$

Since the X_k's are independent, we have:

$$F_{Y_n}(x) = \prod_{k=1}^{n} P\left(X_k \le n^{1/\lambda} \cdot x\right) = \left[F\left(n^{1/\lambda} \cdot x\right)\right]^n.$$

Using the form of $F(x)$, for $n^{1/\lambda} \cdot x > 1$, we get:

$$F(n^{1/\lambda} \cdot x) = 1 - \left(\frac{1}{n^{1/\lambda} \cdot x}\right)^{\lambda} = 1 - \frac{1}{n \cdot x^{\lambda}}.$$

Thus, the CDF becomes:

$$F_{Y_n}(x) = \left(1 - \frac{1}{n \cdot x^{\lambda}}\right)^n.$$

Now, take the limit as $n \to \infty$. Using the approximation $\left(1 - \frac{1}{n \cdot x^\lambda}\right)^n \approx e^{-1/x^\lambda}$ for large n, we have:

$$F_{Y_n}(x) \to e^{-1/x^\lambda} \quad \text{as} \quad n \to \infty.$$

Therefore, Y_n converges in distribution to a random variable with CDF:

$$F_Y(x) = \begin{cases} e^{-1/x^\lambda}, & \text{for } x > 1, \\ 0, & \text{otherwise.} \end{cases}$$

Example 2.1.3. let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with mean μ and finite variance σ^2 and set $S_n = X_1 + X_2 + \ldots + X_n$, $n \geq 1$. The law of large numbers states that

$$P\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) \to 0 \text{ as } n \to \infty \text{ for all } \epsilon > 0,$$

that is,

$$\frac{S_n}{n} \xrightarrow{p} \mu \text{ as } n \to \infty.$$

The proof of this statement under the above assumptions follows from Chebyshev's inequality:

$$P\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2} \to 0 \text{ as } n \to \infty.$$

2.1.2 Uniqueness of convergence

We present a result which states that convergence is unique. In other words, that the limiting random variable is uniquely defined in the following sense: If $X_n \to X$ and $X_n \to Y$ almost surely, in probability, or in r-mean, then $X = Y$ almost surely, that is, $P(X = Y) = 1$ (or, equivalently, $P(\{\omega : X(\omega) \neq Y(\omega)\}) = 0$). For distributional convergence, uniqueness means $F_X(x) = F_Y(x)$ for all x, that is, $X \stackrel{d}{=} Y$.

Proposition 2.1.5. *Let $X_1, X_2, \ldots$ be a sequence of random variables. If X_n converges almost surely, in probability, in r-mean, or in distribution as $n \to \infty$, then the limiting random variable (distribution) is unique.*

2.1.3 Relations between the modes of convergence

The obvious question is whether or not the modes of convergence we have introduced really are different and if they are, whether or not they can be ordered in some sense. The following result summarises these notions.

Proposition 2.1.6. *Let X and $X_1, X_2, \ldots$ be random variables. The following implications hold as $n \to \infty$:*

$$
\begin{array}{ccccc}
X_n \xrightarrow{a.s.} X & \Longrightarrow & X_n \xrightarrow{p} X & \Longrightarrow & X_n \xrightarrow{d} X \\
& & \Uparrow & & \\
& & X_n \xrightarrow{r} X. & &
\end{array}
$$

Note that all implications are strict. However, convergence in probability and convergence in distribution are equivalent if the limiting random variable is degenerate. Finally, the concepts a.s. convergence and convergence in r-mean cannot be ordered; neither implies the other.

2.2. Thinning operations

Consider the first-order autoregressive (AR(1)) model given by

$$Y_t = \alpha \cdot Y_{t-1} + \epsilon_t,$$

where α is a parameter and ϵ_t are some random innovations. This model cannot be applied to count processes. Even if the innovations ϵ_t are assumed to be integer-valued with range $\mathbb{N}_0$, the observations Y_t would still not be integer-valued since the multiplication "$\alpha\cdot$" does not preserve the discrete range (the so-called multiplication problem).

2.2.1 Binomial thinning operator

A way of avoiding the multiplication problem, as described above, is to use the probabilistic operation of binomial thinning, sometimes also referred to as binomial sub-sampling. If X is a discrete random variable with range $\mathbb{N}_0$ and if $\alpha \in (0, 1)$, then the random variable

$$\alpha \circ X := \sum_{i=1}^{X} Z_i$$

is said to arise from X by binomial thinning, and the Z_i are referred to as the counting series. They are i.i.d. binary random variables with

$$P(Z_i = 1) = \alpha,$$

which are also independent of X. So by construction, $\alpha \circ X$ can only lead to integer values between 0 and X. The boundary values $\alpha = 0$ and $\alpha = 1$ might be included in this definition by setting $0 \circ X := 0$ and $1 \circ X := X$. Since each Z_i satisfies $Z_i \sim \text{Bin}(1, \alpha)$, and since the binomial distribution is additive, $\alpha \circ X$ has a conditional binomial distribution given the value of X; that is, $\alpha \circ X | X \sim \text{Bin}(X, \alpha)$. In particular, using the law of total expectation, it follows that

$$E[\alpha \circ X] = E[E[\alpha \circ X|X]] = E[\alpha \cdot X](= \alpha\mu).$$

So the binomial thinning $\alpha \circ X$ and the multiplication $\alpha \cdot X$ have the same mean, which motivates us to use binomial thinning within a modified AR(1) recursion. However, they differ in many other properties; in particular, the multiplication is not a random operation. As an example, the law of total variance implies that

$$\begin{aligned} V[\alpha \circ X] &= V[E[\alpha \circ X|X]] + E[V[\alpha \circ X|X]] \\ &= V[\alpha \cdot X] + E[\alpha(1-\alpha) \cdot X](= \alpha^2\sigma^2 + \alpha(1-\alpha)\mu), \end{aligned}$$

so we have $V[\alpha \circ X] \neq V[\alpha \cdot X]$. For the interpretation of the binomial thinning operation, we consider the following example:

Example 2.2.1. Consider a population of size X at a certain time t. If we observe the same population at a later time, $t+1$, then the population may have shrunk, because some of the individuals had died between times t and $t+1$. If the individuals survive independently of each other, and if the probability of surviving from t to $t+1$ is equal to α for all individuals, then the number of survivors is given by $\alpha \circ X$.

Using the random operator "$\circ$", the integer-valued AR(1) process can be defined as below:

Definition 2.2.1. *Let the innovations $(\varepsilon_t)_{t\in\mathbb{N}}$ be an i.i.d. process with range $\mathbb{N}_0$, denote $E[\varepsilon_t] = \mu_\varepsilon$, $V[\varepsilon_t] = \sigma_\varepsilon^2$. Let $\alpha \in (0,1)$. A process $(X_t)_{t\in\mathbb{N}_0}$ of observations satisfying*

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

is said to be an integer-valued AR(1) process (INAR(1) process) if all thinning operations are performed independently of each other and of $(\varepsilon_t)_{t\in\mathbb{N}}$, and if the thinning operations at each time t as well as ε_t are independent of $(X_s)_{s<t}$.

Note that it would be more correct to write "$\circ_t$" in the above recursion to emphasize the fact that the thinning is realised at each time t anew. However, for the sake of readability, the time index is avoided.

The INAR(1) recursion of Definition 2.2.1 can be interpreted as follows:

$$\underbrace{X_t}_{\text{Population at time } t} = \underbrace{\alpha \circ X_{t-1}}_{\text{Survivors from time } t-1} + \underbrace{\epsilon_t}_{\text{Immigration}}.$$

The INAR(1) process is a homogeneous Markov chain with 1-step transition probabilities:

$$p_{k|l} := P(X_t = k|X_{t-1} = l)$$

$$= \sum_{i=0}^{k} \binom{l}{i} \alpha^i (1-\alpha)^{l-i} \cdot P(\varepsilon_t = k - i)$$

For conditional mean and variance, we have:

$$E[X_t|X_{t-1}] = \alpha \cdot X_{t-1} + \mu_\varepsilon$$

$$V[X_t|X_{t-1}] = \alpha(1-\alpha) \cdot X_{t-1} + \sigma_\varepsilon^2$$

which are both linear functions of X_{t-1}.

2.2.2 Alternative thinning concepts

The idea behind a thinning operation (and the related time series models) can be modified in diverse ways. Such modified thinning concepts allow for different stochastic properties and alternative interpretation schemes. We shall pick out two of these alternative thinning concepts here, but many further approaches are described in Weiß [43].

The generalised thinning operation is given by

$$\alpha \cdot \beta \cdot X = \sum_{j=1}^{X} Z_j \quad \text{with} \quad \alpha \in (0,1) \quad \text{and} \quad \beta > 0.$$

where the random variables Z_j (counting series) are allowed to have the full range $\mathbb{N}_0$ instead of only $\{0,1\}$. Here, the Z_j are required to have mean α and variance β. Since now the Z_j may become larger than 1, the interpretation of INAR(1) as "survival indicators" is no longer appropriate, but they can be understood as describing a reproduction mechanism. Z_j might be the number of children being generated by the jth individual of the population behind X. Using their generalised thinning operation, the INAR(p) model can be extended as below:

$$X_t = \alpha_1 \cdot \beta_1 \cdot X_{t-1} + \cdots + \alpha_p \cdot \beta_p \cdot X_{t-p} + \varepsilon_t,$$

where again the condition $\alpha \cdot \beta = \sum_{j=1}^{p} \alpha_j < 1$ is assumed. In this model, a time index for the thinning operations is omitted for the sake of readability, but it is understood that all thinnings are performed independently.

Another family of thinning operations assumes the counting series to be Bernoulli distributed, but now the thinning probability α is allowed to be random itself. The resulting thinning operation is then called random coefficient (RC) thinning, and it has been applied in the context of count time series modelling. As a specific instance, we consider the case where $\alpha\phi$ follows the BETA $\left(1-\phi, \dfrac{1-\phi}{\phi}\right)$ distribution with $\alpha, \phi \in (0,1)$; that is, where $E[\alpha\phi] = \alpha$ and $\sigma_\alpha^2 := V[\alpha\phi] = \phi\alpha(1-\alpha)$. Then the conditional distribution of $\alpha\phi \circ X$ given

X is a beta-binomial distribution, and the thinning operation is referred to as beta-binomial thinning accordingly. A counterpart to the INAR(1) model using a general random coefficient thinning operation "$\alpha_t \circ$" (that is, where the distribution of α_t on $[0, 1)$ is only required to have mean α and a certain variance σ_α^2, but is not further specified). The RCINAR(1) model is defined by

$$X_t = \alpha_t \circ_t X_{t-1} + \varepsilon_t.$$

While the (conditional) mean and Autocorrelation Function (ACF) remain as in the INAR(1) case, the effect of the additional uncertainty manifests itself in the (conditional) variance:

$$V[X_t|X_{t-1}] = \sigma_\alpha^2(X_{t-1}^2 - X_{t-1}) + \alpha(1-\alpha)X_{t-1} + \sigma_\varepsilon^2,$$

$$\sigma^2 = V[X_t] = \mu(1-\alpha^2) + (\mu - 1)\sigma_\alpha^2 + (1-\alpha)\left(\frac{\sigma_\varepsilon^2}{\mu\varepsilon} - 1\right)(1-\alpha^2) - \sigma_\alpha^2.$$

Note that $V[X_t|X_{t-1}]$ is a quadratic function of X_{t-1}, and the observations are overdispersed even if the innovations are equidispersed.

CHAPTER 3

STRUCTURE OF THE PROCESS

In this chapter we present a Poisson integer-valued generalised autoregressive conditional heteroscedastic model (a Poisson INGARCH model) and its basic properties. In the first section, we define the process and give a necessary condition for its existence. The second section briefly discusses the construction of the process. In the third section, we check for almost sure convergence of the building sequence $X_t^{(n)}$ of the model. In the fourth section, we discuss stationarity of the process. The fifth section discusses the concept of mean-square convergence of $X_t^{(n)}$ while the sixth section looks at the conditional law of the actual process. The last section looks at a particular case of the model, the INGARCH(1,1) process.

3.1. INGARCH(p, q) process and a necessary existence condition

In this section we define a possion INGARCH(p, q) process and then give a necessary condition for the existence of such a process. A defintion of the classical GARCH(p, q) is also given and the two processes are contrasted in terms of how they are formulated.

Definition 3.1.1. *An integer-valued generalised autoregressive conditional heteroscedastic process of order p, q, abbreviated INGARCH(p, q), is an integer-valued process X_t such that:*

1. The conditional distribution of X_t given its past information up to time $t-1$ (denoted by $\mathcal{F}_{t-1}$ and commonly called the natural filtration of X_t) is given by the Poisson distribution with parameter λ_t, $\mathcal{P}(\lambda_t)$, for all $t \in \mathbb{Z}$.

2. The values of λ_t are determined recursively by the relation:

$$\lambda_t = \gamma_0 + \sum_{i=1}^{q} \gamma_i X_{t-i} + \sum_{j=1}^{p} \delta_j \lambda_{t-j}, \tag{3.1}$$

where $\gamma_0 > 0$, $\gamma_i \geqslant 0$ for $i = 1, \ldots, q$, and $\delta_j \geqslant 0$ for $j = 1, \ldots, p$.

Basically, the first part of this definition tells us that this type of an integer-valued process follows the Poisson distribution with parameter λ_t. The second part shows that the parameter λ_t is calculated using a linear combination of the previous q values of X_t and the previous p values of λ_t, with specific coefficients γ_i and δ_j.

An important point to highlight is that an INGARCH(p, q) process is an integer-valued counterpart to a GARCH(p, q) process. The GARCH(p, q) process is defined below.

Definition 3.1.2. *A generalised autoregressive conditional heteroscedastic process of order p, q, abbreviated GARCH(p, q), is a process X_t such that:*

1. *The conditional distribution of X_t given its natural filtration $\mathcal{F}_{t-1}$ follows the normal distribution with mean 0 and variance σ_t^2 for all $t \in \mathbb{Z}$*

2. *The values of the variance σ_t^2 are determined recursively by the relation:*

$$\sigma_t^2 = \gamma_0 + \sum_{i=1}^{q} \gamma_i X_{t-i} + \sum_{j=1}^{p} \delta_j \sigma_{t-j}^2, \tag{3.2}$$

where $\gamma_0 > 0$, $\gamma_i \geqslant 0$ for $i = 1, \ldots, q$, and $\delta_j \geqslant 0$ for $j = 1, \ldots, p$.

In contrast to a GARCH(p, q) process, an INGARCH(p, q) process involves Poisson random variables whose conditional mean (which is also the conditional variance) depends on both the past values of the series as well as on its own past values. This feature of an INGARCH(p, q) process addresses in a meaningful way the issue of heteroscedasticity for a given data set. Furthermore, we observe that an INGARCH($0, q$) process is essentially equivalent to an ARCH(q) process, which only considers the past values of X_t.

In the following proposition, we establish a necessary condition on the parameters of the model to ensure that the process is a second-order stationary (weakly stationary) process. To achieve this, we introduce two polynomials D and G given by

$$D(B) = 1 - \delta_1 B - \cdots - \delta_p B^p$$
$$G(B) = \gamma_1 B + \cdots + \gamma_q B^q,$$

where B is the backshift operator. Furthermore, we assume that the roots of $D(z) = 0$ lie outside the unit circle, which is equivalent to saying that $D(1) = \sum_{j=1}^{p} \delta_j < 1$ for a non-negative δ_j. Under this assumption, the operator $D(B)$ has an inverse denoted by $D^{-1}(B)$, and we can express the parameter λ_t as

$$\lambda_t = D^{-1}(B) \left(\gamma_0 + G(B) X_t \right) = \gamma_0 D^{-1}(1) + H(B) X_t, \tag{3.3}$$

where

$$H(B) = G(B) D^{-1}(B) = \sum_{j=1}^{\infty} \psi_j B^j.$$

The coefficients ψ_j are given by the power expansion of the rational function $G(z)/D(z)$ in the neighbourhood of 0. In the following, we denote the polynomial $D(B) - G(B)$ by $K(B)$.

Proposition 3.1.3. *For a second-order stationary process $\{X_t\}_{t \in \mathbb{Z}}$ to satisfy (3.1) it is necessary that $K(1) = D(1) - G(1) > 0$ or equivalently that $0 \leqslant \delta_1 + \cdots + \delta_p + \gamma_1 + \cdots + \gamma_q < 1$.*

Proof. Let ψ_j be the coefficient of z^j in the Taylor expansion of $G(z)/D(z)$, that is,

$$\begin{aligned}\frac{G(z)}{D(z)} &= \frac{\gamma_1 z + \cdots + \gamma_q z^q}{1 - \delta_1 z - \cdots - \delta_p z^p} \\ &= \psi_0 + \psi_1 z + \psi_2 z^2 + \cdots\end{aligned}$$

We can then calculate the mean μ of the process $\{X_t\}_{t\in\mathbb{Z}}$ as follows:

$$\begin{aligned}\mu &= \mathbb{E}[X_t] \\ &= E(E(X_t|\mathcal{F}_{t-1})), \quad \text{by the tower law} \\ &= E(\lambda_t), \quad \text{since } X_t|\mathcal{F}_{t-1} \text{ is Poisson distributed with parameter } \lambda_t \\ &= \mathbb{E}[\gamma_0 D^{-1}(1) + \sum_{j=1}^{\infty} \psi_j X_{t-j}], \quad \text{by equation 3.3} \\ &= \gamma_0 D^{-1}(1) + \sum_{j=1}^{\infty} \psi_j \mathbb{E}[X_{t-j}] \\ &= \gamma_0 D^{-1}(1) + \sum_{j=1}^{\infty} \psi_j \mu, \quad \text{since } X_t \text{ should be second-order stationary} \\ &= \gamma_0 D^{-1}(1) + \mu \sum_{j=1}^{\infty} \psi_j \\ &= \gamma_0 D^{-1}(1) + \mu D^{-1}(1) G(1), \quad \text{by assumption.}\end{aligned}$$

Now, solving for μ:

$$
\begin{aligned}
\mu &= \gamma_0 D^{-1}(1) + \mu D^{-1}(1)G(1)\\
\mu - \mu D^{-1}(1)G(1) &= \gamma_0 D^{-1}(1)\\
\mu\Big(1 - D^{-1}(1)G(1)\Big) &= \gamma_0 D^{-1}(1)\\
\mu &= \frac{\gamma_0 D^{-1}(1)}{1 - D^{-1}(1)G(1)}\\
\mu &= \frac{\gamma_0 D^{-1}(1)}{D(1)D^{-1}(1) - D^{-1}(1)G(1)}\\
\mu &= \frac{\gamma_0 D^{-1}(1)}{D^{-1}(1)\Big(D(1) - G(1)\Big)}\\
\mu &= \frac{\gamma_0}{D(1) - G(1)}\\
\mu &= \frac{\gamma_0}{1 - \delta_1 - \cdots - \delta_p - \gamma_1 - \cdots - \gamma_q}\\
\mu &= \frac{\gamma_0}{1 - \sum_{i=1}^{p}\delta_i - \sum_{i=1}^{q}\gamma_i}\\
\mu &= \frac{\gamma_0}{K(1)}\\
\mu &= \gamma_0 K^{-1}(1)
\end{aligned}
$$

In light of the above, the parameters $\delta_j, j = 1, \ldots, p$ and $\gamma_i, i = 1, \ldots, q$ of the non-negative integer-valued process $\{X_t\}_{t\in\mathbb{Z}}$ must satisfy necessarily the condition:

$$
1 - \sum_{i=1}^{p}\delta_i - \sum_{i=1}^{q}\gamma_i > 0, \quad \text{or equivalently,} \quad 0 \leq \sum_{i=1}^{p}\delta_i + \sum_{i=1}^{q}\gamma_i < 1, \tag{3.4}
$$

as required, and the proof is complete. □

3.2. Construction of the process

The construction of the INGARCH process involves iterative approximations. Consider a sequence $\{U_t\}_{t\in\mathbb{Z}}$ of independent Poisson variables with a common mean $\psi_0 = \gamma_0/D(1)$. For each $t \in \mathbb{Z}$ and $i \in \mathbb{N}$, let $\mathcal{Z}_{t,i} = \{Z_{t,i,j}\}_{j\in\mathbb{N}}$ represent a sequence of independent Poisson variables with a common mean ψ_i. It is assumed that all variables $U_s, Z_{t,i,j}, (s \in \mathbb{Z}, t \in \mathbb{Z}, i \in \mathbb{N}$, and $j \in \mathbb{N})$ are mutually independent. Using these variables, the building sequence $X_t^{(n)}$ of the INGARCH process is then defined as follows:

$$
X_t^{(n)} = \begin{cases} 0, & n < 0,\\ U_t, & n = 0,\\ U_t + \sum_{i=1}^{n}\sum_{j=1}^{X_{t-i}^{(n-i)}} Z_{t-i,i,j}, & n > 0. \end{cases} \tag{3.5}
$$

The aim here is to show that as $n \to \infty$, the building sequence $X_t^{(n)}$ converges almost surely to the INGARCH process X_t (this is Proposition 3.3.1), and the resulting process $\{X_t\}_{t\in\mathbb{Z}}$ satisfies (3.1).

Remark 3.2.1. It is essential to note that, for each $t \in \mathbb{Z}$, there exists a random variable U_t, and for each $(t, i) \in \mathbb{Z}\times\mathbb{N}$, a sequence of independent and identically distributed Poisson random variables with parameter ψ_i is available, denoted as $\mathcal{Z}_{t,i} = \{Z_{t,i,j}\}_{j\in\mathbb{N}}$.

The limiting process $\{X_t\}$ will prove to be a strictly stationary process as a consequence of Proposition 3.4.2, in which it will be established that for any fixed n, $\{X_t^{(n)}\}_{t\in\mathbb{Z}}$ is strictly stationary. To do so, we will first show that $X_t^{(n)}$ can be obtained through a cascade of so-called thinning operations on a random variable using a sequence of independent and identically distributed Poisson random variables, which is defined as follows: Given a non-negative integer-valued random variable W and an independent sequence of independent and identically distributed $\{V_j\}_{j\in\mathbb{N}}$ Poisson random variables of common mean ψ, the 'thinned random variable,' $\psi \circ W$, is defined as:

$$\psi \circ W = \begin{cases} \sum_{j=1}^{W} V_j, & \text{if } W > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Remark 3.2.2. Using the thinning operation, $X_t^{(n)}$ admits the representation

$$X_t^{(n)} = U_t + \sum_{i=1}^{n} \psi_i^{(t-i)} \circ X_{t-i}^{(n-i)}, \quad n > 0. \tag{3.6}$$

In this representation, $\psi_i^{(\tau)}\circ$ indicates that the Poisson random variables with common mean ψ_i involved in the thinning operation correspond to time τ. n the present case, the sequence $\mathcal{Z}_{\tau,i} = \{Z_{\tau,i,j}\}_{j\in\mathbb{N}}$). This representation shows that $X_t^{(n)}$ is obtained through a cascade of thinning operations along the sequence $\{U_t\}_{t\in\mathbb{Z}}$. For instance, we have

$$\begin{aligned} X_t^{(1)} =& U_t + \psi_1^{(t-1)} \circ U_{t-1} \\ X_t^{(2)} =& U_t + \psi_2^{(t-2)} \circ U_{t-2} + \psi_1^{(t-1)} \circ \left(U_{t-1} + \psi_1^{(t-2)} \circ U_{t-2}\right) \\ X_t^{(3)} =& U_t + \psi_3^{(t-3)} \circ U_{t-3} + \psi_2^{(t-2)} \circ \left(U_{t-2} + \psi_1^{(t-3)} \circ U_{t-3}\right) \\ & + \psi_1^{(t-1)} \circ \left[U_{t-1} + \psi_2^{(t-3)} \circ U_{t-3} + \psi_1^{(t-2)} \circ \left(U_{t-2} + \psi_1^{(t-3)} \circ U_{t-3}\right)\right]. \end{aligned}$$

For any value n, the sequence $X_t^{(n)}$ can be expanded in the above way, a fact that will be used in the proof of Proposition 3.3.1 in the next section.

3.3. Almost sure convergence of $X_t^{(n)}$

In this section we check for the almost sure convergence of the building sequence $X_t^{(n)}$ of the model. We first note that the expectation and the variance of $X_t^{(n)}$ are well defined because $X_t^{(n)}$ is a finite sum of independent Poisson variables. Furthermore the expectation $\mathbb{E}\left[X_t^{(n)}\right]$ does not depend on t but just depends on n and will thus be denoted by μ_n. Now using the fact that $\mu_k = 0$ if $k < 0$, we can express the expectation $\mathbb{E}\left[X_t^{(n)}\right]$ as

$$\begin{aligned} \mu_n &= \sum_{i=1}^{n} \mathbb{E}\left[\sum_{k=1}^{X_{t-i}^{(n-i)}} Z_{t-i,i,k}\right] + D^{-1}(1)\gamma_0 \\ &= \sum_{i=1}^{\infty} \psi_i \mu_{n-i} + D^{-1}(1)\gamma_0. \end{aligned} \tag{3.7}$$

The last equality can be written as:

$$\mu_n = H(B)\mu_n + D^{-1}(B)\gamma_0$$

and rearranging gives

$$D(B)\mu_n = G(B)\mu_n + \gamma_0$$

from which we deduce that

$$[D(B) - G(B)]\mu_n = K(B)\mu_n = \gamma_0.$$

The last equation indicates that the sequence $\{\mu_n\}$ satisfies a finite difference equation with constant coefficients. Note that the characteristic polynomial of this finite difference equation is $K(B)$ and all its roots are outside the unit circle if $K(1) > 0$ (see proposition 3.1.3). Finally, for a fixed value of t, the sequence $\left\{X_t^{(n)}\right\}_{n\in\mathbb{Z}}$ is a non-decreasing sequence of non-negative integer-valued random variables which can be shown by induction with respect to n as

$$X_t^{(n+1)} - X_t^{(n)} = \sum_{j=1}^{U_{t-n-1}} Z_{t-n-1,n+1,j} + \sum_{i=1}^{n} \sum_{j=X_{t-i}^{(n-i)}+1}^{X_{t-i}^{(n+1-i)}} Z_{t-i,i,j} \geqslant 0.$$

We are now ready to present our main results of this section in the form of a proposition.

Proposition 3.3.1. *If $K(1) > 0$ then the sequence $\left\{X_t^{(n)}\right\}_{n\in\mathbb{N}}$ has an almost sure limit.*

Proof. Proof. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the common probability space on which the relevant random variables are defined. Since $X_j^{(n)}$ is a non-decreasing sequence of nonnegative integers, for all $\omega \in \Omega$, we have

$$\lim_{n\to\infty} X_t^{(n)}(\omega) = X_t,$$

which is either finite or infinite. The set

$$\mathcal{A}_\infty = \{\omega : X_t(\omega) = \infty\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \limsup A_n$$

is of probability zero, where for $n > 1$,

$$\mathcal{A}_n = \left\{\omega : X_t^{(n)}(\omega) - X_t^{(n-1)}(\omega) > 0\right\}.$$

On the one hand, we have

$$\mathbb{E}\left[X_t^{(n)} - X_t^{(n-1)}\right] \geqslant \sum_{k=1}^{\infty} \Pr\left\{\omega : X_t^{(n)}(\omega) - X_t^{(n-1)}(\omega) = k\right\} = \Pr\left\{\mathcal{A}_n\right\},$$

while on the other hand, we have

$$\mathbb{E}\left[X_t^{(n)} - X_t^{(n-1)}\right] = \mu_n - \mu_{n-1} = \upsilon_n.$$

The sequence $\{\upsilon_n\}$ satisfies a homogeneous finite difference equation with a characteristic polynomial, namely $K(B)$, that has all its roots outside the unit circle. The sequence $\{\upsilon_n\}_n$ tends towards zero with a geometric rate as $n \to \infty$. In other words, there exist a constant $M \geqslant 0$ and a constant $0 < c < 1$ such that $\upsilon_n \leqslant Mc^n$. This implies that

$$\sum_{n=1}^{\infty} \Pr\left\{\mathcal{A}_n\right\} \leqslant \sum_{n=1}^{\infty} \upsilon_n \leqslant M \sum_{n=1}^{\infty} c^n < \infty.$$

It therefore follows from the Borel–Cantelli lemma that $\Pr\left\{\mathcal{A}_\infty\right\} = 0$, which shows that the sequence $\left\{X_t^{(n)}\right\}_{n\in\mathbb{N}}$ converges almost surely to the INGARCH process X_t □

3.4. Stationarity of the INGARCH process X_t

In this section, we prove that the INGARCH process X_t is strictly stationary. In view of Proposition 3.3.1, this will be proved if we can prove that for any fixed value of n the building sequence $\left\{X_t^{(n)}\right\}$ of X_t is a strictly stationary process. To do so, we use the concept of probability generating functions, whose definition is given below.

Definition 3.4.1. *Consider a random vector* $\mathbf{W}$ *with non-negative integer-valued entries, and let* $p(\mathbf{W})$ *denote its probability mass function. The probability generating function of*

$\mathbf{W}$ *is given by:*

$$g_{\mathbf{w}}(\mathbf{s}) = \mathbb{E}\left[s_1^{W_1} \ldots s_k^{W_k}\right] = \sum_{\mathbf{w} \in (\mathbb{Z}^+)^k} p(\mathbf{w}) \prod_{i=1}^{k} s_i^{w_i},$$

where $\mathbf{s} = (s_1, \ldots, s_k)' \in \mathbb{C}^k$.

Proposition 3.4.2. *For each* n*, the sequence* $\left\{X_t^{(n)}\right\}_{t \in \mathbb{Z}}$ *forms a strictly stationary process.*

Proof. To show strict stationarity, we need to demonstrate that the finite-dimensional joint distributions of $\mathbf{X}_{1..k}^{(n)} = \left(X_1^{(n)}, X_2^{(n)}, \ldots, X_k^{(n)}\right)$ and $\mathbf{X}_{1+h..k+h}^{(n)} = \left(X_{1+h}^{(n)}, X_{2+h}^{(n)}, \ldots, X_{k+h}^{(n)}\right)$ are identical for all $k \geq 1$ and $h \geq 0$.

Let us compare their probability generating functions, denoted $g_{\mathbf{X}}(\mathbf{s})$. The probability generating function for $\mathbf{X}_{1..k}^{(n)}$ is:

$$g_{\mathbf{X}_{1..k}^{(n)}}(\mathbf{s}) = \mathbb{E}\left[s_1^{X_1^{(n)}} s_2^{X_2^{(n)}} \cdots s_k^{X_k^{(n)}}\right].$$

Using the law of total expectation and conditioning on a latent sequence $\mathbf{U}_{t-n..t+k}$ (assumed to influence $X_t^{(n)}$):

$$g_{\mathbf{X}_{1..k}^{(n)}}(\mathbf{s}) = \mathbb{E}\left[\mathbb{E}\left[\prod_{i=1}^{k} s_i^{X_i^{(n)}} \mid \mathbf{U}_{t-n..t+k}\right]\right].$$

Substituting the explicit dependence of $X_t^{(n)}$ on $\mathbf{U}_{t-n..t+k}$, we write:

$$g_{\mathbf{X}_{1..k}^{(n)}}(\mathbf{s}) = \sum_{\mathbf{z} \in \mathbb{Z}_+^{k+n}} \mathbb{E}\left[\prod_{i=1}^{k} s_i^{X_i^{(n)}} \mid \mathbf{U}_{t-n..t+k} = \mathbf{z}\right] \mathbb{P}(\mathbf{U}_{t-n..t+k} = \mathbf{z}).$$

For the shifted process $\mathbf{X}_{1+h..k+h}^{(n)}$, a similar expression holds:

$$g_{\mathbf{X}_{1+h..k+h}^{(n)}}(\mathbf{s}) = \sum_{\mathbf{z} \in \mathbb{Z}_+^{k+n}} \mathbb{E}\left[\prod_{i=1}^{k} s_i^{X_{i+h}^{(n)}} \mid \mathbf{U}_{t-n+h..t+k+h} = \mathbf{z}\right] \mathbb{P}(\mathbf{U}_{t-n+h..t+k+h} = \mathbf{z}).$$

From the problem's assumptions, the latent sequence $\mathbf{U}_{t-n..t+k}$ determines $X_t^{(n)}$. If $\mathbf{U}_{t-n..t+k} = \mathbf{U}_{t-n+h..t+k+h}$, the joint conditional distributions of $\mathbf{X}_{1..k}^{(n)}$ and $\mathbf{X}_{1+h..k+h}^{(n)}$ are identical. Thus:

$$\mathbb{E}\left[\prod_{i=1}^{k} s_i^{X_i^{(n)}} \mid \mathbf{U}_{t-n..t+k} = \mathbf{z}\right] = \mathbb{E}\left[\prod_{i=1}^{k} s_i^{X_{i+h}^{(n)}} \mid \mathbf{U}_{t-n+h..t+k+h} = \mathbf{z}\right].$$

Since the marginal distributions $\mathbb{P}(\mathbf{U}_{t-n..t+k} = \mathbf{z})$ and $\mathbb{P}(\mathbf{U}_{t-n+h..t+k+h} = \mathbf{z})$ are the same, it follows that:

$$g_{\mathbf{X}^{(n)}_{1..k}}(\mathbf{s}) = g_{\mathbf{X}^{(n)}_{1+h..k+h}}(\mathbf{s}).$$

The joint distributions of $\mathbf{X}^{(n)}_{1..k}$ and $\mathbf{X}^{(n)}_{1+h..k+h}$ are therefore identical. Thus, $\{X_t^{(n)}\}_{t\in\mathbb{Z}}$ is strictly stationary.

□

Therefore, the main result (stationarity of the INGARCH process) for this section then follows, as presented in the following corollary.

Corollary 3.4.3. *The process $\{X_t\}_{t\in\mathbb{Z}}$ is a strictly stationary process.*

We end this section by establishing some results concerning the first two moments of the INGARCH process.

Proposition 3.4.4. *Assuming $K(1) > 0$, if $\{X_t\}_{t\in\mathbb{Z}}$ is the almost sure limit of Proposition 3.3.1, then its first moment is finite.*

Proof. As we have shown earlier, $\left\{X_t^{(n)}\right\}_{t\in\mathbb{Z}}$ is a strictly stationary process. Consequently, μ_n can be expressed as:

$$\mu_n = \sum_{i=1}^{n} \psi_i \mu_n + \gamma_0 D^{-1}(1) = \frac{\gamma_0 D^{-1}(1)}{1 - \sum_{i=1}^{n} \psi_i}.$$

Since $X_t^{(n)}$ increases with n towards X_t, applying Beppo Levi's theorem gives:

$$\begin{aligned}
\mathbb{E}\left[X_t\right] = \mu &= \frac{\gamma_0 D^{-1}(1)}{1 - \sum_{i=1}^{\infty} \psi_i} \\
&= \frac{\gamma_0 D^{-1}(1)}{1 - G(1)D^{-1}(1)} \\
&= \frac{\gamma_0}{D(1) - G(1)} \\
&= \frac{\gamma_0}{K(1)}.
\end{aligned}$$

From the above result, we can conclude that the first moment is finite. □

Proposition 3.4.5. *Assuming that $K(1) > 0$, if $\{X_t\}_{t\in\mathbb{Z}}$ is again the almost sure limit of Proposition 3.3.1, then its second moment is finite.*

Proof. For the second moment, we proceed as follows. By using the Cauchy–Schwarz inequality and noting that $\left\{X_t^{(n)}\right\}_{n\in\mathbb{N}}$ is a non-decreasing sequence in n, we obtain:

$$\mathbb{E}\left[\left(X_t^{(n)}\right)^2\right] \leqslant \sum_{i=1}^{n} \psi_i \mathbb{E}\left[X_t\right] + \left(\sum_{i=1}^{n} \psi_i\right)^2 \mathbb{E}\left[X_t^{(n)}\right]^2 + 2\gamma_0 D^{-1}(1) \sum_{i=1}^{n} \psi_i \mathbb{E}\left[X_t\right] + \mathbb{E}\left[U_t^2\right].$$

Thus, we have:

$$\begin{aligned}
\mathbb{E}\left[\left(X_t^{(n)}\right)^2\right] &\leqslant \frac{\mathbb{E}\left[X_t\right]\left(1 + 2\gamma_0 D^{-1}(1)\right) \sum_{i=1}^{n} \psi_i + \mathbb{E}\left[U_t^2\right]}{1 - \left(\sum_{i=1}^{n} \psi_i\right)^2} \\
&\leqslant \frac{\mathbb{E}\left[X_t\right]\left(1 + 2\gamma_0 D^{-1}(1)\right) \sum_{i=1}^{\infty} \psi_i + \mathbb{E}\left[U_t^2\right]}{1 - \left(\sum_{i=1}^{\infty} \psi_i\right)^2} = C.
\end{aligned}$$

By applying Lebesgue's dominated convergence theorem, we can conclude that $\mathbb{E}\left[X_t^2\right] \leqslant C$. Therefore, the second moment is finite. □

3.5. Mean-square convergence of $X_t^{(n)}$

The mean-square convergence of a Poisson INGARCH process $\{X_t\}_{t\in\mathbb{Z}}$ is the property that the mean-square error between the true value of the building sequence $\left\{X_t^{(n)}\right\}_{n\in\mathbb{N}}$ of $\{X_t\}_{t\in\mathbb{Z}}$ and its estimate converges to zero as the sample size n increases.

Proposition 3.5.1. *If $K(1) > 0$, then the sequence $\left\{X_t^{(n)}\right\}_{n\in\mathbb{N}}$ has a mean-square limit.*

Proof. We need to show that the first two moments of the building sequence $\left\{X_t^{(n)}\right\}_{n\in\mathbb{N}}$ of the INGARCH process $\{X_t\}_{t\in\mathbb{Z}}$ are finite, and consequently, the random variables $X_t^{(n)}$ are in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Let us define $V_t^{(n)} = \left(X_t^{(n)} - X_t\right)^2$. Then the sequence $\left\{V_t^{(n)}\right\}_{n\in\mathbb{N}}$ is decreasing, bounded below by 0, and $\mathbb{E}\left[V_t^{(0)}\right] = \mathbb{E}\left[(U_t - X_t)^2\right] < \infty$. Consequently,

$$\lim_{n\to\infty} \mathbb{E}\left[V_t^{(n)}\right] = \mathbb{E}\left[\lim_{n\to\infty} V_t^{(n)}\right] = 0,$$

because $\left\{X_t^{(n)}\right\}_{n\in\mathbb{N}}$ converges almost surely to X_t. Hence, we deduce that $\left\{X_t^{(n)}\right\}_{n\in\mathbb{N}}$ also converges to X_t in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. □

3.6. Conditional law of $\{X_t\}_{t\in\mathbb{Z}}$ given $\mathcal{F}_{t-1}$

In this section, we discuss the distributional properties of the INGARCH process $\{X_t\}_{t\in\mathbb{Z}}$. We have so far established that the building sequence $\left\{X_t^{(n)}\right\}$ converges almost surely and in mean-square, but we still have to prove that the distributional properties are satisfied. The main result of this section is given in the proposition below.

Definition 3.6.1. *A sequence of random variables $\{Y_n\}_{n\in\mathbb{N}}$ converges to a random variable*

Y in L^1 if:

$$\lim_{n\to\infty} \mathbb{E}\left[|Y_n - Y|\right] = 0.$$

Proposition 3.6.2. *The conditional distribution of the INGARCH process $\{X_t\}_{t\in\mathbb{Z}}$ given its natural filtration $\mathcal{F}_{t-1}$ follows the Poisson distribution with parameter λ_t for all $t \in \mathbb{Z}$.*

Proof. Given $\mathcal{F}_{t-1} = \sigma\left(X_{t-1}, X_{t-2}, \ldots\right)$, for $t \in \mathbb{Z}$, let

$$\lambda_t = \gamma_0 D^{-1}(1) + \sum_{j=1}^{\infty} \psi_j X_{t-j} \tag{3.8}$$

The sequence $\{\lambda_t\}$ satisfies:

$$\lambda_t = \gamma_0 + \sum_{i=1}^{q} \gamma_i X_{t-i} + \sum_{j=1}^{p} \delta_j \lambda_{t-j}.$$

For a fixed t, consider the sequence $\left\{r_t^{(n)}\right\}_{n\in\mathbb{N}}$ defined by

$$r_t^{(n)} = U_t + \sum_{i=1}^{n} \sum_{k=1}^{X_{t-i}} Z_{t-i,i,k}.$$

We claim there is a subsequence $\{n_k\}$ such that $r_t^{(n_k)}$ converges almost surely to X_t. Indeed, since

$$X_t - r_t^{(n)} = \left(X_t - X_t^{(n)}\right) + \left(X_t^{(n)} - r_t^{(n)}\right),$$

we just have to find a subsequence of

$$Y_t^{(n)} = r_t^{(n)} - X_t^{(n)} = \sum_{i=1}^{n} \sum_{j=X_{t-i}^{(n-i)}+1}^{X_{t-i}} Z_{t-i,i,j}$$

which converges almost surely to zero. But

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}\left[Y_t^{(n)}\right] &= \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}\left[\sum_{j=X_{t-i}^{(n-i)}+1}^{X_{t-i}} Z_{t-i,i,j}\right] \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}\left[X_{t-i} - X_{t-i}^{(n-i)}\right] \psi_i \\ &= \lim_{n\to\infty} \left(\mu \sum_{i=1}^{n} \psi_i - \sum_{i=1}^{n} \mu_{n-i}\psi_i\right) \\ &= \gamma_0 \frac{H(1)}{D(1)-G(1)} - \lim_{n\to\infty} \sum_{i=1}^{n} \mu_{n-i}\psi_i \\ &= \gamma_0 \frac{G(1)}{D(1)[D(1)-G(1)]} - \lim_{n\to\infty} \mu_n + \frac{\gamma_0}{D(1)} \\ &= 0.\end{aligned}$$

This means $Y_t^{(n)}$ converges to zero in L^1 because $Y_t^{(n)}$ in non-negative. Therefore, there is a subsequence $Y_t^{(n_k)}$ converging almost surely to the same limit. Now the required distributional property of X_t follows easily. Since $r_t^{(n_k)}$ converges almost surely to $X_t, r_t^{(n)} \mid \mathcal{F}_{t-1}$ converges almost surely to $X_t \mid \mathcal{F}_{t-1}$. However,

$$r_t^{(n)} \mid \mathcal{F}_{t-1} : \mathcal{P}\left(\gamma_0 D^{-1}(1) + \sum_{j=1}^{n} \psi_j X_{t-j}\right),$$

Thus, the INGARCH process $\{X_t\}_{t\in\mathbb{Z}}$ given its natural filtration $\mathcal{F}_{t-1}$ follows the Poisson distribution with parameter λ_t, $\mathcal{P}(\lambda_t)$, for all $t \in \mathbb{Z}$. □

3.7. INGARCH(1,1) process

In this section we look at a particular case of the INGARCH(p,q) process. It is well known that a standard GARCH process is also, from the weakly stationarity point of view, an ARMA process. This is still true in the case of an INGARCH(p,q) process giving an ARMA $(\max\{p,q\},p)$ model (see [33]). For simplicity, in this section we give some properties of the model in the special case where $p = q = 1$. For this case, we have for all $t \in \mathbb{Z}$ the model

$$\begin{cases} X_t \mid \mathcal{F}_{t-1} : \mathcal{P}(\lambda_t) \\ \lambda_t = \gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1} \end{cases} \tag{3.9}$$

In this model, the parameters are positive and less than 1. Under suitable conditions, we show that moments of all orders exist and thereafter, we give the mean and the autocovariance function of the INGARCH$(1,1)$ process.

The following lemma refers to Stirling numbers of the second kind, denoted by $\{ {n \atop k} \}$, count the number of ways to partition a set of n elements into k non-empty subsets. They arise naturally in combinatorial settings and are useful in expanding expressions involving powers of sums, such as λ^k, when written in terms of lower-order terms. .
For $n \geq 0$ and $0 \leq k \leq n$, these numbers satisfy the recurrence:

$$\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k \left\{ {n-1 \atop k} \right\}$$

where $\{ {n \atop n} \} = 1$ for all $n \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$, $\left\{ {n \atop 0} \right\} = 0$ for all $n \in \mathbb{N} = \{1, 2, \ldots\}$ and $\{ {n \atop k} \} = 0$ if $k > n$

Lemma 3.7.1. *Let X be a Poisson random variable with mean λ. The uncentred moments of X satisfy*

$$\mathbb{E}\left[X^k\right] = \sum_{i=0}^{k} \left\{ {k \atop i} \right\} \lambda^i$$

Proof. Let X be a Poisson random variable with mean λ. The moment-generating function (MGF) of X is given by:

$$M_X(t) = \mathbb{E}[e^{tX}] = e^{\lambda(e^t - 1)}$$

The kth moment of X is then given by:

$$\mathbb{E}[X^k] = \frac{1}{k!} \frac{d^k}{dt^k} M_X(t) \bigg|_{t=0}$$

Taking derivatives and evaluating at $t = 0$, we find:

$$\mathbb{E}[X^k] = \lambda^k$$

Now, let's express λ^k in terms of Stirling numbers of the second kind:

$$\lambda^k = \sum_{i=0}^{k} \left\{ {k \atop i} \right\} \lambda^{(i)}$$

where $\{ {k \atop i} \}$ denotes the Stirling number of the second kind, and the proof is complete. $\square$

Example 3.7.1. Understanding the Connection to λ^k with Iterations

Step 1: Base Case ($k = 1$)
For $k = 1$, we know:

$$\mathbb{E}[X] = \lambda.$$

This agrees with the expression:

$$\mathbb{E}[X^1] = \sum_{i=0}^{1} \left\{ \begin{array}{c} 1 \\ i \end{array} \right\} \lambda^i = \left\{ \begin{array}{c} 1 \\ 1 \end{array} \right\} \lambda^1 = \lambda.$$

Here, $\left\{ \begin{array}{c} 1 \\ 1 \end{array} \right\} = 1.$

Step 2: Second Iteration ($k = 2$)
For $k = 2$, we use:

$$\mathbb{E}[X^2] = \sum_{i=0}^{2} \left\{ \begin{array}{c} 2 \\ i \end{array} \right\} \lambda^i.$$

From the properties of Stirling numbers of the second kind:

$$\left\{ \begin{array}{c} 2 \\ 2 \end{array} \right\} = 1, \quad \left\{ \begin{array}{c} 2 \\ 1 \end{array} \right\} = 1.$$

Thus:

$$\mathbb{E}[X^2] = \left\{ \begin{array}{c} 2 \\ 2 \end{array} \right\} \lambda^2 + \left\{ \begin{array}{c} 2 \\ 1 \end{array} \right\} \lambda^1 = \lambda^2 + \lambda.$$

This matches the known result for the second raw moment of a Poisson random variable:

$$\mathbb{E}[X^2] = \lambda^2 + \lambda.$$

Proposition 3.7.2. *The moments of an INGARCH* $(1, 1)$ *are all finite if and only if*

$$\gamma_1 + \delta_1 < 1.$$

Proof. Since $X_t \mid \mathcal{F}_{t-1}$ is a Poisson random variable with mean $\lambda_t = \gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1}$, conditionally to time $t - 1$, the m th moment is

$$\mathbb{E}_{\mathcal{F}_{t-1}} [X_t^m] = \sum_{i=0}^{m} \left\{ \begin{array}{c} m \\ i \end{array} \right\} \lambda_t^i.$$

So,

$$\mathbb{E}\left[X_t^m\right] = \sum_{i=0}^{m} \left\{ \begin{matrix} m \\ i \end{matrix} \right\} \mathbb{E}\left[\lambda_t^i\right]$$

and

$$\lambda_t^i = (\gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1})^i = \sum_{n=0}^{i} \left(\begin{matrix} i \\ n \end{matrix} \right) \gamma_0^{i-n} \sum_{j=0}^{n} \left(\begin{matrix} n \\ j \end{matrix} \right) \gamma_1^j \delta_1^{n-j} X_{t-1}^j \lambda_{t-1}^{n-j}$$

We note that:

$$\mathbb{E}_{\mathcal{F}_{t-2}}\left[X_{t-1}^j \lambda_{t-1}^{n-j}\right] = \sum_{k=0}^{j} \left\{ \begin{matrix} j \\ k \end{matrix} \right\} \lambda_{t-1}^{k+n-j},$$

Consequently,

$$\mathbb{E}_{\mathcal{F}_{t-2}}\left[\lambda_t^i\right] = \sum_{n=0}^{i} \left(\begin{matrix} i \\ n \end{matrix} \right) \gamma_0^{i-n} \sum_{j=0}^{n} \sum_{k=0}^{j} \left\{ \begin{matrix} j \\ k \end{matrix} \right\} \left(\begin{matrix} n \\ j \end{matrix} \right) \gamma_1^j \delta_1^{n-j} \lambda_{t-1}^{n+k-j} \tag{3.10}$$

Let $\mathbf{\Lambda}_t = (\lambda_t^m, \ldots, \lambda_t)'$. In the algebraic expression of $\mathbb{E}_{\mathcal{F}_{t-2}}\left[\lambda_t^i\right], i = 1, \ldots, m$, all the powers of λ_{t-1} are $\leq i$. Therefore, a constant vector $\mathbf{d}$ and an upper triangular matrix $\mathbf{D}$ exist such that the following equation is satisfied:

$$\mathbb{E}_{\mathcal{F}_{t-2}}\left[\mathbf{\Lambda}_t\right] = \mathbf{d} + \mathbf{D}\mathbf{\Lambda}_{t-1}.$$

Iterating this recurrence ℓ times gives:

$$\mathbb{E}_{\mathcal{F}_{t-2-i}}\left[\mathbf{\Lambda}_t\right] = \left(\sum_{r=0}^{\ell} D^r\right) \mathbf{d} + \mathbf{D}^{\ell+1} \mathbf{\Lambda}_{t-(\ell+1)}.$$

Substituting $k = \ell + 2$ leads to:

$$\mathbb{E}_{\mathcal{F}_{t-k}}\left[\mathbf{\Lambda}_t\right] = \left(\sum_{r=0}^{k-2} D^r\right) \mathbf{d} + \mathbf{D}^{k-1} \mathbf{\Lambda}_{t-(k-1)}.$$

If the eigenvalues of $\mathbf{D}$ are inside the unit circle, one can write:

$$\mathbb{E}_{\mathcal{F}_{t-k}}\left[\mathbf{\Lambda}_t\right] = (\mathbf{I} - \mathbf{D})^{-1} \left(\mathbf{I} - \mathbf{D}^{k-1}\right) \mathbf{d} + \mathbf{D}^{k-1} \mathbf{\Lambda}_{t-(k-1)}.$$

D being a triangular matrix, the eigenvalues are the diagonal entries. The ℓ th diagonal entry corresponds to the case where in Equation (3.12), we let $i = m - \ell + 1$. To obtain the coefficient of $\lambda_{t-1}^{m-\ell+1}$, we look at the terms corresponding to $n = m - \ell + 1$ and $k = j$. They give:

$$d_{\ell\ell} = \sum_{j=0}^{m-\ell+1} \left(\begin{matrix} m - \ell + 1 \\ j \end{matrix} \right) \gamma_1^j \delta_1^{m-\ell+1-j} \lambda_{t-1}^{m-\ell+1} = (\gamma_1 + \delta_1)^{m-\ell+1}.$$

We conclude that the eigenvalues are inside the unit circle if, and only if, $\gamma_1 + \delta_1 < 1$. Consequently, we have

$$\mathbb{E}[\mathbf{\Lambda}_t] = \lim_{k\to\infty} \mathbb{E}_{\mathcal{F}_{t-k}}[\mathbf{\Lambda}_t] = (\mathbf{I} - \mathbf{D})^{-1}\mathbf{d}$$

and all the moments of X_t of order $\leqslant m$ are finite. □

It is quite interesting that all moments of the process exist. This is an unexpected result considering our experience with common GARCH models. In view of this, from now onwards, we will assume that $\gamma_1 + \delta_1 < 1$.

Proposition 3.7.3. *The expected value of the INGARCH* $(1,1)$ *process,* μ*, is given by*

$$\mu = \frac{\gamma_0}{1-(\gamma_1+\delta_1)}.$$

Proof. Recall that the INGARCH(1,1) process is defined for all $t \in \mathbb{Z}$ by

$$\begin{cases} X_t \mid \mathcal{F}_{t-1} : \mathcal{P}(\lambda_t) \\ \lambda_t = \gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1} \end{cases} \tag{3.11}$$

We take expectations on both sides of the equation $\lambda_t = \gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1}$, that is,

$$\mathbb{E}[\lambda_t] = \mathbb{E}[\gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1}]$$

Using linearity of expectation and the fact that X_{t-1} is measurable with respect to $\mathcal{F}_{t-1}$:

$$\mu = \gamma_0 + \gamma_1 \mu + \delta_1 \mu$$

Solving for μ, we have

$$\mu - \gamma_1\mu - \delta_1\mu = \gamma_0$$

$$\mu(1-\gamma_1-\delta_1) = \gamma_0$$

$$\mu = \frac{\gamma_0}{1-(\gamma_1+\delta_1)}$$

Hence, the expected value of the process is indeed $\mu = \frac{\gamma_0}{1-(\gamma_1+\delta_1)}$. □

Proposition 3.7.4. *The variance of the INGARCH* $(1,1)$ *process is given by*

$$\mathbb{V}[X_t] = \frac{\mu\left(1-(\gamma_1+\delta_1)^2+\gamma_1^2\right)}{1-(\gamma_1+\delta_1)^2}$$

Proof. To find the variance of the process described by the system of equations, denoted by

σ^2, we start by considering the expected value of X_t^2, denoted as $\mathbb{E}[X_t^2]$, which is equal to $\mathbb{E}[\lambda_t^2]$, where λ_t is defined as:

$$\lambda_t = \gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1}$$

Expanding $\mathbb{E}[\lambda_t^2]$, we get:

$$\mathbb{E}[X_t^2] = \mathbb{E}[(\gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1})^2]$$

Using linearity of expectation and properties of conditional expectations:

$$\mathbb{E}[X_t^2] = \gamma_0^2 + 2\gamma_0\gamma_1\mathbb{E}[X_{t-1}] + 2\gamma_0\delta_1\mathbb{E}[\lambda_{t-1}] + \gamma_1^2\mathbb{E}[X_{t-1}^2] + 2\gamma_1\delta_1\mathbb{E}[X_{t-1}\lambda_{t-1}] + \delta_1^2\mathbb{E}[\lambda_{t-1}^2]$$

Given that $\mathbb{E}[X_{t-1}] = \mathbb{E}[\lambda_{t-1}] = \mu$ and $\mathbb{E}[X_{t-1}\lambda_{t-1}] = \mu^2$, and using the expression for μ derived in the previous proposition, we simplify:

$$\mathbb{E}[X_t^2] = \gamma_0^2 + 2\gamma_0\gamma_1\mu + 2\gamma_0\delta_1\mu + \gamma_1^2(\mu^2 + \sigma^2) + 2\gamma_1\delta_1\mu^2 + \delta_1^2(\mu^2 + \sigma^2)$$

Now, we need to compute $(\mathbb{E}[X_t])^2 = \mu^2$. We already know that $\mathbb{E}[X_t] = \mu$. Putting it all together, we have:

$$\sigma^2 = \mathbb{E}[X_t^2] - (\mathbb{E}[X_t])^2 = \left(\gamma_0^2 + 2\gamma_0\gamma_1\mu + 2\gamma_0\delta_1\mu + \gamma_1^2(\mu^2 + \sigma^2) + 2\gamma_1\delta_1\mu^2 + \delta_1^2(\mu^2 + \sigma^2)\right) - \mu^2$$

After simplifying the expression, we get:

$$\sigma^2 = \frac{\mu(1 - (\gamma_1 + \delta_1)^2 + \gamma_1^2)}{1 - (\gamma_1 + \delta_1)^2}$$

Hence, the variance of the given process is $\mathbb{V}[X_t] = \frac{\mu\left(1-(\gamma_1+\delta_1)^2+\gamma_1^2\right)}{1-(\gamma_1+\delta_1)^2}$. □

Remark 3.7.1. As shown in the last two propositions, the mean and the variance are different. Hence, the marginal distribution of X_t is not Poisson.

Proposition 3.7.5. *The autocovariance function of a INGARCH* $(1,1)$ *process is given by*

$$\gamma(r) = \frac{\gamma_1\left(1 - \delta_1\left(\gamma_1 + \delta_1\right)\right)\left(\gamma_1 + \delta_1\right)^{r-1}\mu}{1 - \left(\gamma_1 + \delta_1\right)^2}, \quad \forall r \geqslant 1.$$

Proof. To find the autocovariance function $\gamma(r)$ of the process, we start by considering the covariance between X_t and X_{t-r}. Using the definition of covariance, we have:

$$\gamma(r) = \text{cov}(X_t, X_{t-r}) = \mathbb{E}[X_t X_{t-r}] - \mathbb{E}[X_t]\mathbb{E}[X_{t-r}]$$

Using the conditional expectation property, we can express $\mathbb{E}[X_t X_{t-r}]$ as

$$\mathbb{E}[X_t X_{t-r}] = \mathbb{E}[\lambda_t \lambda_{t-r}]$$

Expanding $\mathbb{E}[\lambda_t \lambda_{t-r}]$ using the given system of equations, we have

$$\mathbb{E}[X_t X_{t-r}] = \mathbb{E}[(\gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1})(\gamma_0 + \gamma_1 X_{t-r-1} + \delta_1 \lambda_{t-r-1})]$$

Using linearity of expectation and properties of conditional expectations, we have

$$\mathbb{E}[X_t X_{t-r}] = \gamma_0^2 + \gamma_0\gamma_1 \mathbb{E}[X_{t-1}] + \gamma_0\delta_1 \mathbb{E}[\lambda_{t-1}] + \gamma_1^2 \mathbb{E}[X_{t-1} X_{t-r-1}] +$$
$$\gamma_1\delta_1 \mathbb{E}[X_{t-1}\lambda_{t-r-1}] + \delta_1^2 \mathbb{E}[\lambda_{t-1}\lambda_{t-r-1}]$$

Given that $\mathbb{E}[X_{t-1}] = \mathbb{E}[\lambda_{t-1}] = \mu$, $\mathbb{E}[X_{t-1} X_{t-r-1}] = \mu^2$, and $\mathbb{E}[X_{t-1}\lambda_{t-r-1}] = \mu\mathbb{E}[\lambda_{t-r-1}] = \mu^2$, we simplify further:

$$\mathbb{E}[X_t X_{t-r}] = \gamma_0^2 + 2\gamma_0\gamma_1\mu + 2\gamma_0\delta_1\mu + \gamma_1^2\mu^2 + 2\gamma_1\delta_1\mu^2 + \delta_1^2 \mathbb{E}[\lambda_{t-1}\lambda_{t-r-1}]$$

Now, to compute $\mathbb{E}[\lambda_{t-1}\lambda_{t-r-1}]$, we expand it using the given system of equations:

$$\mathbb{E}[\lambda_{t-1}\lambda_{t-r-1}] = \mathbb{E}[(\gamma_0 + \gamma_1 X_{t-2} + \delta_1 \lambda_{t-2})(\gamma_0 + \gamma_1 X_{t-r-2} + \delta_1 \lambda_{t-r-2})]$$

Following similar steps as before, we eventually obtain:

$$\mathbb{E}[\lambda_{t-1}\lambda_{t-r-1}] = \mu^2 + \frac{\gamma_1 \left(1 - \delta_1 \left(\gamma_1 + \delta_1\right)\right) \left(\gamma_1 + \delta_1\right)^{r-2} \mu}{1 - \left(\gamma_1 + \delta_1\right)^2}$$

Substituting this back into the expression for $\mathbb{E}[X_t X_{t-r}]$, we get:

$$\mathbb{E}[X_t X_{t-r}] = \gamma_0^2 + 2\gamma_0\gamma_1\mu + 2\gamma_0\delta_1\mu + \gamma_1^2\mu^2 + 2\gamma_1\delta_1\mu^2 + \delta_1^2 \left(\mu^2 + \frac{\gamma_1 \left(1 - \delta_1 \left(\gamma_1 + \delta_1\right)\right) \left(\gamma_1 + \delta_1\right)^{r-2} \mu}{1 - \left(\gamma_1 + \delta_1\right)^2}\right)$$

Now, we need to compute $\mathbb{E}[X_t]\mathbb{E}[X_{t-r}] = \mu^2$. Hence, after simplification, we get:

$$\gamma(r) = \frac{\gamma_1 \left(1 - \delta_1 \left(\gamma_1 + \delta_1\right)\right) \left(\gamma_1 + \delta_1\right)^{r-1} \mu}{1 - \left(\gamma_1 + \delta_1\right)^2}$$

Which is the required expression for the autocovariance function. □

Remark 3.7.2. According to the following lemma (Lemma 3.7.6 below), the autocovariance function of the INGARCH(1,1) process indicates that the INGARCH(1,1) process is also an ARMA(1,1) process, as described by Corollary 3.7.7. Note that, Lemma 3.7.6 below was given as a remark by Brockwell and Davis [12] on page 90, and as such, we simply state it without proving it.

Lemma 3.7.6. *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.*

Corollary 3.7.7. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an INGARCH $(1,1)$ process satisfying*

$$\begin{cases} X_t \mid \mathcal{F}_{t-1} : \mathcal{P}(\lambda_t) \\ \lambda_t = \gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1} \end{cases}$$

Then $\{X_t\}_{t\in\mathbb{Z}}$ is an ARMA$(1,1)$ *that can be written as*

$$(X_t - \mu) - (\gamma_1 + \delta_1)(X_{t-1} - \mu) = e_t - \delta_1 e_{t-1}, \tag{3.12}$$

where $\{e_t\}_{t\in\mathbb{Z}}$ is a white noise process of variance $\sigma^2 = \mu = \gamma_0/(1 - \gamma_1 - \delta_1)$.

Consider $\{Y_t\}_{t\in\mathbb{Z}}$ a Gaussian ARMA$(1,1)$ process satisfying the equation

$$Y_t - \phi Y_{t-1} = \varepsilon_t + \theta \varepsilon_{t-1}, \tag{3.13}$$

where $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a sequence of i.i.d. $\mathcal{N}(0, \sigma^2)$. In Equation (3.13), we let

$$\theta = -\delta_1, \quad \phi = \gamma_1 + \delta_1, \quad \sigma^2 = \mu.$$

The the autocovariance function of a ARMA$(1,1)$ process is given by

$$\gamma(r) = \frac{\gamma_1 (1 - \delta_1 (\gamma_1 + \delta_1)) (\gamma_1 + \delta_1)^{r-1} \mu}{1 - (\gamma_1 + \delta_1)^2}, \quad \forall r \geqslant 1,$$

which is the same as the autocovariance of the INGARCH $(1,1)$ process.

To derive the autocovariance function of the ARMA(1,1) process described, we can start by expressing the process equation in terms of the innovations ε_t:

$$Y_t - (\gamma_1 + \delta_1) Y_{t-1} = \varepsilon_t - \delta_1 \varepsilon_{t-1}$$

Rearranging terms:

$$Y_t = (\gamma_1 + \delta_1) Y_{t-1} + \varepsilon_t - \delta_1 \varepsilon_{t-1}$$

Now, let's compute the autocovariance function $\gamma(r)$ for $r \geq 1$. The autocovariance $\gamma(r)$ is defined as:

$$\gamma(r) = \text{Cov}(Y_t, Y_{t-r})$$

Expanding the covariance:

$$\gamma(r) = E\left[(Y_t - E[Y_t])(Y_{t-r} - E[Y_{t-r}])\right]$$

Substituting the expression for Y_t:

$$\gamma(r) = E\left[((\gamma_1 + \delta_1)Y_{t-1} + \varepsilon_t - \delta_1\varepsilon_{t-1} - E[Y_t])((\gamma_1 + \delta_1)Y_{t-r-1} + \varepsilon_{t-r} - \delta_1\varepsilon_{t-r-1} - E[Y_{t-r}])\right]$$

Since ε_t are i.i.d. with mean 0 and variance μ, and Y_t is linear in ε_t, we have $E[\varepsilon_t] = 0$ and $E[Y_t] = 0$ for all t.

$$\gamma(r) = E\left[((\gamma_1 + \delta_1)Y_{t-1} + \varepsilon_t - \delta_1\varepsilon_{t-1})((\gamma_1 + \delta_1)Y_{t-r-1} + \varepsilon_{t-r} - \delta_1\varepsilon_{t-r-1})\right]$$

Expanding this expression and using the fact that the innovations are uncorrelated with past values of the process, we get:

$$\begin{aligned}\gamma(r) = (\gamma_1 + \delta_1)(\gamma_1 + \delta_1)^{r-1}E[Y_{t-1}Y_{t-r-1}] + E[\varepsilon_t Y_{t-r-1}]\\ - \delta_1 E[\varepsilon_{t-1}Y_{t-r-1}] - \delta_1(\gamma_1 + \delta_1)^{r-1}E[\varepsilon_{t-1}\varepsilon_{t-r-1}]\end{aligned}$$

Since Y_t follows an ARMA(1,1) process, $Y_t = \psi_0 + \psi_1 Y_{t-1} + \theta_1\varepsilon_{t-1} + \varepsilon_t$, where $\psi_0 = 0$ (for simplicity), $\psi_1 = \gamma_1 + \delta_1$, and $\theta_1 = -\delta_1$. Now, let's calculate the terms:

$$\begin{aligned}E[Y_{t-1}Y_{t-r-1}] &= E[(\psi_1 Y_{t-2} + \theta_1\varepsilon_{t-2} + \varepsilon_{t-1})(\psi_1 Y_{t-r-2} + \theta_1\varepsilon_{t-r-2} + \varepsilon_{t-r-1})]\\ &= \psi_1^2 E[Y_{t-2}Y_{t-r-2}] + \theta_1^2 E[\varepsilon_{t-2}\varepsilon_{t-r-2}]\\ &= (\gamma_1 + \delta_1)^2\gamma(r-1) + \mu\delta_1^2\delta_{r-1}\end{aligned}$$

Where $\delta_r = \begin{cases} 1 & \text{if } r = 0 \\ 0 & \text{otherwise} \end{cases}$. Also,

$$\begin{aligned}E[\varepsilon_t Y_{t-r-1}] &= E[\varepsilon_t(\psi_1 Y_{t-r-2} + \theta_1\varepsilon_{t-r-2} + \varepsilon_{t-r-1})]\\ &= \theta_1 E[\varepsilon_t\varepsilon_{t-r-2}]\\ &= \mu\delta_{r-1}\end{aligned}$$

And,

$$\begin{aligned}E[\varepsilon_{t-1}Y_{t-r-1}] &= E[\varepsilon_{t-1}(\psi_1 Y_{t-r-2} + \theta_1\varepsilon_{t-r-2} + \varepsilon_{t-r-1})]\\ &= \psi_1 E[\varepsilon_{t-1}Y_{t-r-2}]\\ &= \mu(\gamma_1 + \delta_1)\delta_{r-1}\end{aligned}$$

Finally,

$$E[\varepsilon_{t-1}\varepsilon_{t-r-1}] = \mu\delta_{r-1}$$

Substituting these expressions back into the equation for $\gamma(r)$, we get:

$$\gamma(r) = (\gamma_1 + \delta_1)^{r-1}\gamma(1) + \mu\delta_1\delta_{r-1} - \mu(\gamma_1 + \delta_1)\delta_{r-1} - \delta_1(\gamma_1 + \delta_1)^{r-1}\mu$$

Simplifying,

$$\gamma(r) = (\gamma_1 + \delta_1)^{r-1}\gamma(1) - \mu(\gamma_1 + \delta_1)\delta_{r-1}$$

Given $\gamma(1)$ as the autocovariance at lag 1, we can express it as:

$$\gamma(1) = \frac{\gamma_1(1 - \delta_1(\gamma_1 + \delta_1))\mu}{1 - (\gamma_1 + \delta_1)^2}$$

Thus,

$$\gamma(r) = \frac{\gamma_1(1 - \delta_1(\gamma_1 + \delta_1))(\gamma_1 + \delta_1)^{r-1}\mu}{1 - (\gamma_1 + \delta_1)^2}, \quad r \geqslant 1.$$

The last result indicates that the classical least squares approach can be used to estimate the parameters of such a process. It should be noted that inference can also be made in this framework. In Section 4, we consider the conditional maximum likelihood estimation of the parameters.

CHAPTER 4

AN APPROACH FOR PARAMETER ESTIMATION

Recall from Definition 3.1.1 that an integer-valued generalised autoregressive conditional heteroscedastic process of order p, q, abbreviated as INGARCH(p, q), is an integer-valued process X_t satisfying the following two conditions:

1. The conditional distribution of X_t given its natural filtration $\mathcal{F}_{t-1}$ follows the Poisson distribution with parameter λ_t, $\mathcal{P}(\lambda_t)$, for all $t \in \mathbb{Z}$.
2. The values of λ_t are determined recursively by the relation:

$$\lambda_t = \gamma_0 + \sum_{i=1}^{q} \gamma_i X_{t-i} + \sum_{j=1}^{p} \delta_j \lambda_{t-j},$$

   where $\gamma_0 > 0$, $\gamma_i \geqslant 0$ for $i = 1, \ldots, q$, and $\delta_j \geqslant 0$ for $j = 1, \ldots, p$.

Now given the time series data $x_1, ...x_n$, the task is to estimate the value of the parameters $\Theta = (\gamma_0, \gamma_1, \ldots, \gamma_q, \delta_1, \ldots, \delta_p)' = (\theta_0, \ldots, \theta_{p+q})'$. This can be done using the method of conditional maximum likelihood estimation.

4.1. Conditional maximum likelihood estimation

The estimation procedure for the parameters of the INGARCH model is analogous to the one used for traditional GARCH model. For simplicity, we condition on the pre-sample values. As mentioned by Bollerslev [10], this does not affect the asymptotic results. The conditional likelihood function of the n observations $x_1, \ldots, x_n$, conditionally on the pre-sample values, is given by

$$L(\Theta) = \prod_{t=1}^{n} \frac{\mathrm{e}^{-\lambda_t} \lambda_t^{x_t}}{x_t!}$$

where

$$\Theta = (\gamma_0, \gamma_1, \ldots, \gamma_q, \delta_1, \ldots, \delta_p)' = (\theta_0, \ldots, \theta_{p+q})'$$
$$\lambda_t = \gamma_0 + \sum_{i=1}^{q} \gamma_i x_{t-i} + \sum_{j=1}^{p} \delta_j \lambda_{t-j}.$$

Analytical estimates from this likelihood function cannot be found, even if we compute and set to zero the derivatives. In fact, we obtain a system of non-linear equations. For this reason, we can only suggest the use of numerical optimisation methods to find out the optimal value of Θ. As usual, we work with the log-likelihood function

$$\mathcal{L}(\Theta) = \ln L(\Theta) = \sum_{t=1}^{n} [x_t \ln \lambda_t - \lambda_t - \ln (x_t!)] = \sum_{t=1}^{n} \ell_t(\Theta),$$

where $\ell_t(\Theta) = x_t \ln \lambda_t - \lambda_t - \ln (x_t!)$. Obviously, we can neglect the term $\ln (x_t!)$. The first derivatives of $\ell_t(\Theta)$ with respect to $\theta_i, i = 0, \dots, p+q$ are

$$\frac{\partial \ell_t}{\partial \theta_i} = \left(\frac{\partial \lambda_t}{\partial \theta_i}\right)\left(\frac{x_t}{\lambda_t} - 1\right) \tag{4.1}$$

while the second derivatives are

$$\frac{\partial^2 \ell_t}{\partial \theta_i \partial \theta_j} = \frac{\partial^2 \lambda_t}{\partial \theta_i \partial \theta_j}\left(\frac{x_t}{\lambda_t} - 1\right) - \frac{x_t}{\lambda_t^2}\frac{\partial \lambda_t}{\partial \theta_i}\frac{\partial \lambda_t}{\partial \theta_j} \tag{4.2}$$

for $0 \leqslant i, j \leqslant p+q$. Moreover, we have:

$$\begin{aligned}
\lambda_t &= \gamma_0 + \sum_{i=1}^{q} \gamma_i x_{t-i} + \sum_{j=1}^{p} \delta_j \lambda_{t-j}; \\
\frac{\partial \lambda_t}{\partial \gamma_0} &= 1 + \sum_{j=1}^{p} \delta_j \frac{\partial \lambda_{t-j}}{\partial \gamma_0}; \\
\frac{\partial \lambda_t}{\partial \gamma_i} &= x_{t-i} + \sum_{j=1}^{p} \delta_j \frac{\partial \lambda_{t-j}}{\partial \gamma_i}, \quad i = 1, \dots, q; \\
\frac{\partial \lambda_t}{\partial \delta_j} &= \lambda_{t-j} + \sum_{k=1}^{p} \delta_k \frac{\partial \lambda_{t-k}}{\partial \delta_j}, \quad j = 1, \dots, p.
\end{aligned}$$

As in Bollerslev [10], if n is large enough, the distribution of the maximum likelihood estimator Θ can be approximated by the following distribution:

$$\hat{\Theta} \sim \mathcal{N}\left(\Theta_0, n^{-1}\Im\left(\Theta_0\right)^{-1}\right),$$

where $\Im(\Theta_0)$ is the information matrix evaluated at Θ_0, the actual value of Θ. Let $\mathbf{W}_t' = (1, x_{t-1}, \dots, x_{t-q}, \lambda_{t-1}, \dots, \lambda_{t-p})$ and denote by ∇g the gradient of any function g. We have

$$\nabla \lambda_t = \mathbf{W}_t + \sum_{j=1}^{p} \delta_j \nabla \lambda_{t-j}$$

Equation (4.1) can now be written as

$$\nabla \ell_t = \left(\frac{x_t}{\lambda_t} - 1 \right) \nabla \lambda_t \tag{4.3}$$

and Equation (4.2) becomes

$$\mathbf{H}_t = \left(\frac{x_t}{\lambda_t} - 1 \right) \nabla \left[\nabla' \lambda_t \right] - \frac{x_t}{\lambda_t^2} \nabla \lambda_t \nabla' \lambda_t. \tag{4.4}$$

Taking the expectation on both sides of equation (4.2), we find:

$$\begin{aligned} \mathbb{E}_{\mathcal{F}_{t-1}} \left[\frac{\partial^2 \ell_t}{\partial \theta_i \partial \theta_j} \right] &= \frac{\partial^2 \lambda_t}{\partial \theta_i \partial \theta_j} \mathbb{E}_{\mathcal{F}_{t-1}} \left[\frac{x_t}{\lambda_t} - 1 \right] - \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} \mathbb{E}_{\mathcal{F}_{t-1}} \left[\frac{x_t}{\lambda_t^2} \right] \\ &= -\frac{1}{\lambda_t} \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} \end{aligned}$$

and

$$-\mathbb{E} \left[\frac{\partial^2 \ell_t}{\partial \theta_i \partial \theta_j} \right] = \mathbb{E} \left[\frac{1}{\lambda_t} \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} \right].$$

In a similar way, we obtain

$$\mathbb{E} \left[\frac{\partial \ell_t}{\partial \theta_i} \frac{\partial \ell_t}{\partial \theta_j} \right] = \mathbb{E} \left[\left(\frac{x_t}{\lambda_t} - 1 \right)^2 \left(\frac{\partial \lambda_t}{\partial \theta_j} \right) \left(\frac{\partial \lambda_t}{\partial \theta_i} \right) \right] = \mathbb{E} \left[\frac{1}{\lambda_t} \frac{\partial \lambda_t}{\partial \theta_i} \frac{\partial \lambda_t}{\partial \theta_j} \right]$$

The model satisfies the information matrix equality:

$$-\mathbb{E}[[\frac{\partial^2 \ell_t}{\partial \theta_i \partial \theta_j}]] = \mathbb{E}[[\frac{\partial \ell_t}{\partial \theta_i} \frac{\partial \ell_t}{\partial \theta_j}]].$$

The matrices

$$\hat{\mathbf{S}}_n = \frac{1}{n} \sum_{t=1}^{n} \nabla \ell_t \nabla' \ell_t$$

and

$$\hat{\mathbf{D}}_n = -\frac{1}{n}\sum_{t=1}^{n} \nabla\left[\nabla' \ell_t\right]$$

would be consistent estimates of the information matrix. Thus, both could be used to estimate the asymptotic covariance matrix of the maximum likehood estimator. The first one is the outer-product estimate and the second one is the second-derivative estimate. White [43] suggested using

$$\mathbb{E}\left[\left(\hat{\mathbf{\Theta}} - \mathbf{\Theta}_0\right)\left(\hat{\mathbf{\Theta}} - \mathbf{\Theta}_0\right)'\right] \equiv \frac{1}{n}\left(\hat{\mathbf{D}}_n \hat{\mathbf{S}}_n^{-1} \hat{\mathbf{D}}_n\right)^{-1} \tag{4.5}$$

in a quasi-maximum likelihood estimation approach.

4.2. One-step-ahead forecast confidence interval

A one-step-ahead forecast confidence interval is a range of values that is likely to contain the true value of the next observation in a time series. It is calculated using the historical data and the estimated parameters of a time series model.

The width of the confidence interval depends on the accuracy of the model and the amount of historical data. A wider confidence interval indicates that the model is less accurate or that there is less historical data available.

One-step-ahead forecast confidence intervals can be used to make decisions about future events. However, it is important to remember that they are only estimates and that the true value of the next observation may fall outside the confidence interval.

To construct a confidence interval for the one step ahead forecasts, we used the R statistical programming language to help us obtain the parameters.

CHAPTER 5

REAL-DATA EXAMPLE

The aim of this study is to effectively model overdispersed integer-valued time series data with conditional heteroscedasticity. In light of this, we illustrate the INGARCH model on an overdispersed integer-valued time series real data with conditional heteroscedasticity. That is, we study the daily number of shares trading on the Lusaka Securities Exchange (LuSE) from October 1, 2021 to May 10, 2022 (except for weekends and public holidays), giving a total of 150 observations. This is a part of data set given on the LuSE website and is available for download at https://luse.co.zm/market-data.

The number of shares traded on a stock exchange within a given period is typically measured on a daily basis. Volume reflects the level of activity and liquidity in the market. Trend indicates the direction in which trading volume is moving over time. Increasing volume may indicate growing interest or momentum in a particular stock or the market as a whole, while decreasing volume could signal a loss of interest or a potential reversal. Seasonality describes patterns or trends in trading volume that occur at specific times of the day, week, month, or year. For example, trading volume often spikes at the opening and closing hours of the trading day and may vary depending on the day of the week or month.

Some descriptive statistics of our share time series data are shown in Table 5.1 below:

Table 5.1: Summary Statistics of the daily number of shares, measured in 100,000 thousands

Statistic	Value
Sample Size	150
Maximum	8.96
Minimum	0.065
Mean	2.48287
Median	3.855
Variance	5.76448
Skewness	-0.450693
Kurtosis	-1.00845

These summary statistics show that the mean is 2.48 and the variance is 5.76, which indicates that the data is sufficiently overdispersed, and thus suitable to be used in an example of an application of the INGARCH model.

A plot of the data is given in Figure 5.1. Apart from a random variation somewhere in the middle of the time series, there seems to be no trend, indicating the stationarity of the data.

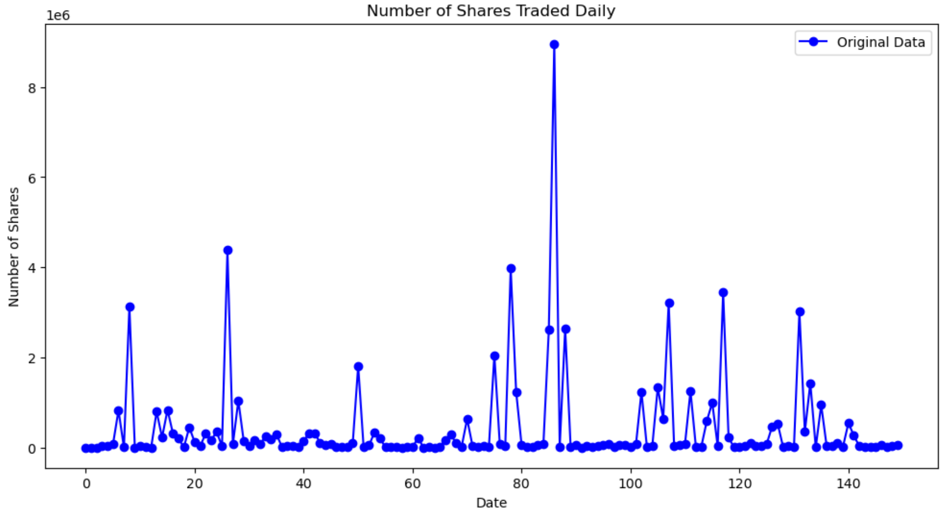

Figure 5.1: Daily number of shares trading on LuSE: October 1, 2021 to March 10, 2022

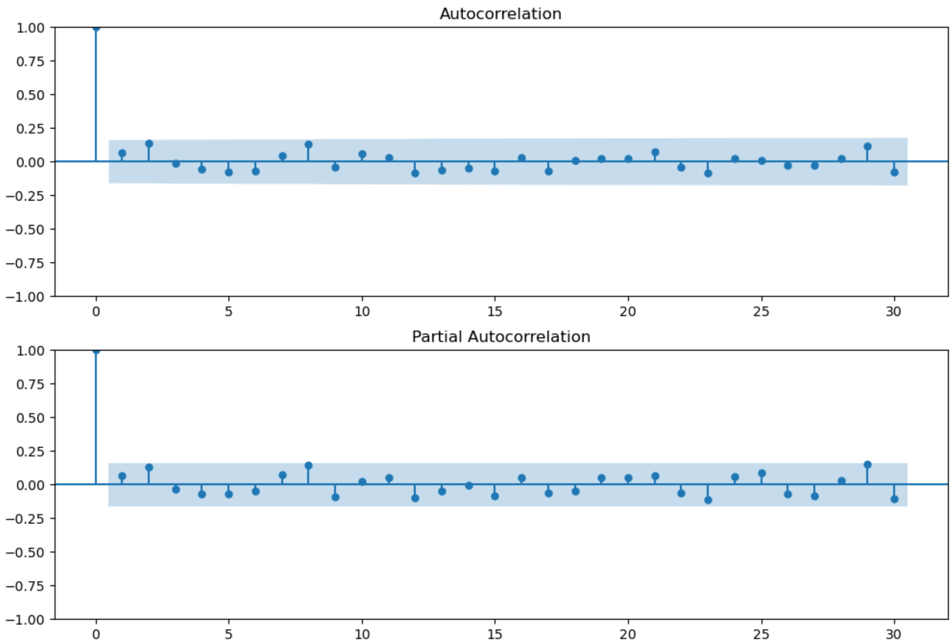

Figure 5.2: Sample autocorrelation function (SACF) and Sample partial autocorrelation function (SPACF) of the share time series data

From the sample autocorrelation function (SACF) and the sample partial autocorrelation function (SPACF) in Figure 5.2 above, we were led to opt for the following model:

$$\begin{cases} X_t \mid \mathcal{F}_{t-1} : \mathcal{P}\left(\lambda_t\right) \\ \lambda_t = \gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1} \end{cases} \tag{5.1}$$

We used the R statistical programming language to compute the maximum likelihood estimates of the parameters: γ_0, γ_1 and δ_1. The estimated parameters for this model are displayed in Table 5.2 below.

Table 5.2: Estimated parameter values

Parameter	γ_0	γ_1	δ_1
Estimated value	1.7384	0.3271	0.2844
Standard error	0.7721	0.0628	0.1012

In Figure 5.3 below, we represent the one step ahead forecasts using a 95% confidence interval. These intervals are quite tightly spaced.

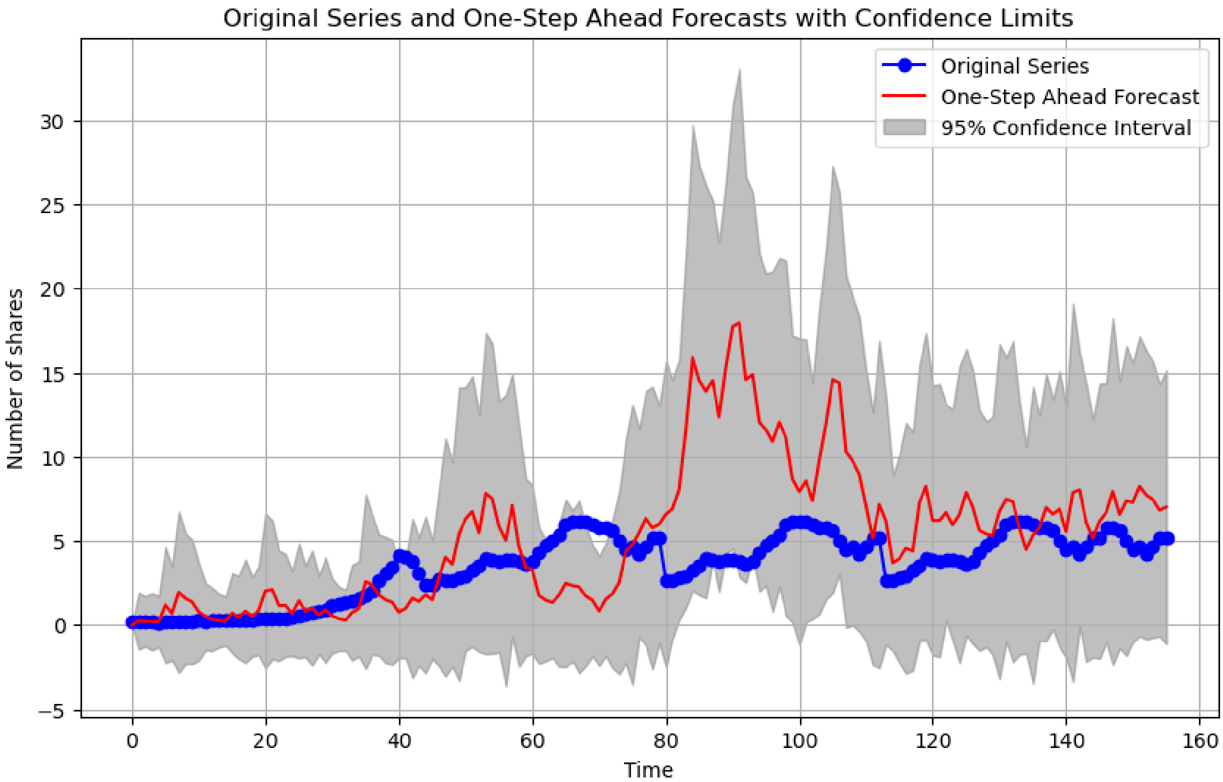

Figure 5.3: The daily number of shares trading on LuSE. Original series and one-step-ahead forecasts with confidence limits.

CHAPTER 6

DISCUSSION

In this chapter, we delve into the core of our research findings, meticulously analysing the Integer-Valued generalised Autoregressive Conditional heteroscedasticity (INGARCH) process. The chapter serves as the focal point where we unravel the intricacies of our model. Through a comprehensive exploration of its theoretical underpinnings, estimation techniques, and practical applications, we aim to unravel the complexities of this model and shed light on its relevance in overdispersed integer-valued time series analysis.

6.1. On the structure of the process

We start by delving into the definition and necessary conditions for the existence of an INGARCH(p,q) process, contrasting it with the classical GARCH(p,q) process.

Firstly, the definition of an INGARCH(p,q) process X_t outlines its unique characteristics, notably its reliance on the Poisson distribution, $\mathcal{P}(\lambda_t)$, for conditional distribution and the sequential determination of the parameter λ_t based on past values of X_t and λ_t. This definition underscores the modelling flexibility offered by an INGARCH(p,q) process, particularly in capturing the heteroscedasticity present in count data sets. A crucial point made here is the distinction between an INGARCH(p,q) and GARCH(p,q) process. While both aim to model volatility clustering, an INGARCH(p,q) process utilises discrete distributions and incorporates both past values of the series and its own past values in determining the conditional mean. This sets it apart from GARCH(p,q) processes, which rely solely on past values of the series for modelling conditional variance. Furthermore, Proposition 3.1.3 establishes a necessary condition for the second-order stationarity of an INGARCH(p,q) process. This condition, expressed in terms of the coefficients δ_j and γ_i, provides insight into the parameter space where the process behaves stably over time. Specifically, it highlights the importance of the relationship between the parameters in maintaining stationarity, emphasising the need for careful parameter selection in model specification.

The construction of an INGARCH process involves iterative approximations. We start with a sequence $\{U_t\}_{t\in\mathbb{Z}}$ of independent Poisson variables with a common mean $\psi_0 = \gamma_0/D(1)$. Additionally, for each $t \in \mathbb{Z}$ and $i \in \mathbb{N}$, we have a sequence $\mathcal{Z}_{t,i} = \{Z_{t,i,j}\}_{j\in\mathbb{N}}$ of independent Poisson variables with a common mean ψ_i. These variables are assumed to be mutually independent. Utilizing these variables, the building sequence $X_t^{(n)}$ of the INGARCH process is then defined recursively. It starts with $X_t^{(n)} = 0$ for $n < 0$, $X_t^{(n)} = U_t$ for $n = 0$, and for

$n > 0$, it is defined as $X_t^{(n)} = U_t + \sum_{i=1}^{n} \sum_{j=1}^{X_{t-i}^{(n-i)}} Z_{t-i,i,j}$. The aim here is to demonstrate that as $n \to \infty$, the building sequence $X_t^{(n)}$ converges almost surely to the INGARCH process X_t, as stated in Proposition 3.3.1. Moreover, the resulting process $\{X_t\}_{t\in\mathbb{Z}}$ satisfies the desired conditions. An important remark is that for each $t \in \mathbb{Z}$, there exists a random variable U_t, and for each $(t,i) \in \mathbb{Z}\times\mathbb{N}$, a sequence of independent and identically distributed Poisson random variables with parameter ψ_i is available, denoted as $\mathcal{Z}_{t,i} = \{Z_{t,i,j}\}_{j\in\mathbb{N}}$. The limiting process $\{X_t\}$ is proved to be a strictly stationary process as a consequence of Proposition 3.4.2, where it is be established that for any fixed n, the sequence $\{X_t^{(n)}\}_{t\in\mathbb{Z}}$ is strictly stationary. This is achieved by showing that $X_t^{(n)}$ can be obtained through a cascade of thinning operations on a random variable using a sequence of independent and identically distributed Poisson random variables. Using the thinning operation, $X_t^{(n)}$ admits the representation

$$X_t^{(n)} = U_t + \sum_{i=1}^{n} \psi_i^{(t-i)} \circ X_{t-i}^{(n-i)}, \quad n > 0. \tag{6.1}$$

This representation illustrates that $X_t^{(n)}$ is obtained through a cascade of thinning operations along the sequence $\{U_t\}_{t\in\mathbb{Z}}$. For any value of n, the sequence $X_t^{(n)}$ can be expanded iteratively. This recursive construction provides insight into the convergence of the building sequence $X_t^{(n)}$ to the desired INGARCH process X_t, laying the groundwork for further analysis and characterization of the resulting process.

In Section 3.3, we examined the almost sure convergence of the building sequence $\{X_t^{(n)}\}$ of the model. We first noted that the expectation and variance of $X_t^{(n)}$ are well-defined because $X_t^{(n)}$ is a finite sum of independent Poisson variables. Additionally, the expectation $\mathbb{E}[X_t^{(n)}]$ does not depend on t but only on n, denoted as μ_n.

Expressing the expectation $\mathbb{E}[X_t^{(n)}]$ as

$$\begin{aligned}\mu_n &= \sum_{i=1}^{n} \mathbb{E}\left[\sum_{k=1}^{X_{t-i}^{(n-i)}} Z_{t-i,i,k}\right] + D^{-1}(1)\gamma_0 \\ &= \sum_{i=1}^{\infty} \psi_i \mu_{n-i} + D^{-1}(1)\gamma_0, \end{aligned} \tag{6.2}$$

we derived the finite difference equation $[D(B) - G(B)]\mu_n = K(B)\mu_n = \gamma_0$, where $K(B)$ is the characteristic polynomial. The key proposition of Section 3.3 states that if $K(1) > 0$, then the sequence $\{X_t^{(n)}\}_{n\in\mathbb{N}}$ has an almost sure limit. The proof of this proposition relied on the Borel-Cantelli lemma and showed that the sequence $\{X_t^{(n)}\}$ converges almost surely to the INGARCH process X_t. This result is significant as it establishes the convergence of the building sequence to the desired process, providing confidence in the model's behavior over time.

In Section 3.4, we established the stationarity of the INGARCH process X_t by proving that its building sequence $\{X_t^{(n)}\}$ is a strictly stationary process for any fixed value of n. To achieve this, we utilised probability generating functions and showed that the joint distribution of $\mathbf{X}_{1..k}^{(n)}$ and $\mathbf{X}_{1+h..k+h}^{(n)}$ remains unchanged, thereby establishing strict stationarity. The main result is summarised in the following corollary:

Corollary 6.1.1. *The process $\{X_t\}_{t\in\mathbb{Z}}$ is a strictly stationary process.*

We also investigated the first two moments of the INGARCH process. Specifically, we showed that under the assumption $K(1) > 0$, both the first and second moments of X_t are finite. This is crucial for understanding the behaviour of the process and provides insights into its statistical properties. Overall, the stationarity of the INGARCH process, along with the finiteness of its moments, underscores its stability and suitability for modelling various phenomena in finance and economics involving count data.

We established the mean-square convergence property in Proposition 3.5.1, which states that if $K(1) > 0$, then the sequence $\{X_t^{(n)}\}_{n\in\mathbb{N}}$ has a mean-square limit. The proof hinges on demonstrating that the first two moments of the building sequence $\{X_t^{(n)}\}_{n\in\mathbb{N}}$ are finite, implying that the random variables $X_t^{(n)}$ are in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. By defining $V_t^{(n)} = (X_t^{(n)} - X_t)^2$, we showed that the sequence $\{V_t^{(n)}\}_{n\in\mathbb{N}}$ is decreasing, bounded below by 0, and its expected value tends to zero as n approaches infinity. This convergence in expectation implies that $\{X_t^{(n)}\}_{n\in\mathbb{N}}$ converges to X_t in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. The mean-square convergence property is fundamental in assessing the accuracy and reliability of the INGARCH process as a model for various phenomena involving count data, particularly in finance and economics. It ensures that as the sample size grows, the estimates provided by the building sequence approach the true values, enhancing the predictive power and usefulness of the model.

In Section 3.6, we delved into the distributional properties of the INGARCH process $\{X_t\}_{t\in\mathbb{Z}}$. While we have already established the almost sure and mean-square convergence of the building sequence $\left\{X_t^{(n)}\right\}$, we still need to verify the satisfaction of distributional properties.

The main result of Section 3.6, presented in Proposition 3.6.1, asserts that the conditional distribution of the INGARCH process $\{X_t\}_{t\in\mathbb{Z}}$ given its natural filtration $\mathcal{F}_{t-1}$ follows the Poisson distribution with parameter λ_t for all $t \in \mathbb{Z}$. The proof of Proposition 3.6.1 relies on establishing the parameter λ_t defined by Equation (3.8), which depends on the past values of the process up to lag $t-1$. By showing that the sequence $\{\lambda_t\}$ satisfies a recursive equation, we can conclude that the conditional distribution of X_t given $\mathcal{F}_{t-1}$ is Poisson with parameter λ_t. This result provides insight into the stochastic behavior of the process, revealing that each future value of X_t is determined probabilistically based on past information up to time $t-1$. The proof further establishes convergence properties, ensuring that the building sequence $X_t^{(n)}$ converges almost surely to X_t, thus confirming the desired distributional property. The distributional characterization of the INGARCH process is essential for understanding its probabilistic behaviour and forms the basis for various statistical analyses and modelling applications.

In Section 3.7, we looked at a particular case of the INGARCH(p, q) process. It is well known that a standard GARCH process is also, from the weakly stationarity point of view, an ARMA process. This is still true in the case of an INGARCH(p, q) process giving an ARMA $(\max\{p, q\}, p)$ model. Moving on to the main results, Proposition 3.7.2 establishes a condition for the finiteness of moments in the INGARCH(1,1) process. Specifically, it states that the moments are all finite if and only if the sum of the autoregressive and moving average parameters is less than 1.

The intuition behind this result is that the moments of the process depend on the moments of the latent variable λ_t, which, in turn, depend on the parameters γ_1 and δ_1. If the sum of these parameters exceeds 1, the moments of λ_t may not be finite, leading to infinite moments of X_t.

Propositions 3.7.3 and 3.7.5 provide expressions for the mean and auto-covariance function of the INGARCH(1,1) process under the condition $\gamma_1 + \delta_1 < 1$. It demonstrates that when this condition is satisfied, the process is stationary, and its mean and auto-covariance function can be explicitly computed.

The mean of the process is proportional to γ_0 and inversely proportional to $1 - \gamma_1 - \delta_1$. This reflects the impact of the parameters on the long-term average behaviour of the process.

Similarly, the auto-covariance function involves a combination of the parameters γ_1 and δ_1, along with their powers and a scaling factor σ^2. This function describes the dependency structure of the process over different lags.

Overall, these results provide valuable insights into the behaviour of the INGARCH(1,1) process, shedding light on its stationarity and moment properties under certain parameter conditions. They offer a foundation for further analysis and modelling of integer-valued time series data

6.2. On the approach for parameter estimation

In real-world applications, the results obtained from the maximum likelihood estimation procedure provide valuable insights into the behaviour of the time series under consideration. For the Poisson INGARCH model, the conditional likelihood function

$$L(\Theta) = \prod_{t=1}^{n} \frac{\mathrm{e}^{-\lambda_t} \lambda_t^{x_t}}{x_t!}, \quad \text{where } \lambda_t = \gamma_0 + \sum_{i=1}^{q} \gamma_i x_{t-i} + \sum_{j=1}^{p} \delta_j \lambda_{t-j},$$

derived from the observations $x_1, \ldots, x_n$, conditioned on pre-sample values, allows us to model the dynamic behaviour of the volatility through the parameters

$$\Theta = (\gamma_0, \gamma_1, \ldots, \gamma_q, \delta_1, \ldots, \delta_p)'.$$

This likelihood function is essential for estimating the parameters that govern the INGARCH model. However, due to the non-linear nature of the likelihood function, analytical solutions for the parameter estimates cannot be derived. This necessitates the use of numerical optimisation methods to find the optimal values of Θ. The log-likelihood function $\mathcal{L}(\Theta)$ provides a convenient form for optimisation, facilitating the estimation process. The gradient and Hessian of the log-likelihood function with respect to the parameters Θ are crucial for optimisation algorithms. The gradient gives the direction of steepest ascent, while the Hessian provides information about the curvature of the log-likelihood surface. These quantities are instrumental in iteratively improving parameter estimates until convergence is achieved. Furthermore, the asymptotic properties of the maximum likelihood estimator $\hat{\Theta}$ provide insights into its statistical properties. The asymptotic distribution of $\hat{\Theta}$ can be approximated by a normal distribution, allowing for inference on the parameters' significance and confidence intervals. The information matrix equality, as demonstrated in equations (4.3) and (4.4), plays a pivotal role in understanding the efficiency of the maximum likelihood estimator. Consistent estimates of the information matrix, such as $\hat{\mathbf{S}}_n$ and $\hat{\mathbf{D}}_n$, enable the estimation of the asymptotic covariance matrix of $\hat{\Theta}$.

In Section 4.2, a one-step-ahead forecast confidence interval is a range of values that is likely to contain the true value of the next observation in a time series. It is calculated using the historical data and the estimated parameters of a time series model. The width of the confidence interval depends on the accuracy of the model and the amount of historical data. A wider confidence interval indicates that the model is less accurate or that there is less historical data available. One-step-ahead forecast confidence intervals can be used to make decisions about future events. However, it is important to remember that they are only estimates and that the true value of the next observation may fall outside the confidence interval.

6.3. On the real-data example

In Chapter 5, the model is illustrated on the daily number of shares trading on the Lusaka Securities Exchange (LuSE) from October 1, 2021, to May 10, 2022 (except for weekends and public holidays), giving a total of 150 observations.

The summary statistics in Table 5.1 offer a detailed glimpse into the characteristics of share data, shedding light on the distribution properties and variability. With a mean of 2.48287 and a variance of 5.76448, the data appears to be sufficiently overdispersed, indicating its potential suitability for modelling using the INGARCH model. This overdispersion, coupled with the range spanning from 0.0065 to 89.58 shares and a slightly higher median of 3.855 compared to the mean, suggests a negatively skewed distribution with occasional spikes in trading volume. The negative skewness value of -0.450693 further supports this notion. Additionally, the negative kurtosis value of -1.0084 indicates lighter tails compared to a normal distribution, implying a lower likelihood of extreme events. These insights underscore the importance of robust modelling approaches like the INGARCH model in capturing the persistent volatility patterns inherent in the data, offering valuable insights for risk management, investment strategies, and market analysis.

A plot of the data is given in Figure 5.1. Apart from a random variation somewhere in the middle of the time series, there seems to be no trend, indicating the stationarity of the data.

The analysis of the Sample Autocorrelation Function (SACF) and Sample Partial Autocorrelation Function (SPACF) in Figure 5.2 guided our model selection (Equation 6.3). The SACF plots the correlation between the time series (representing shares over time) and its lagged values, identifying significant correlations exceeding the significance bounds. The plot shows a gradual decline in correlation values, indicating some autocorrelation, though not particularly strong. Insights from the SACF plot include a gradual decay in correlation with lag, significant early lags indicating recent past values' stronger influence, and most correlation values lying within significance bounds. However, the lag 1 autocorrelation value exceeds the bounds, suggesting the presence of a moving average (MA) term in the model

The Sample Partial Autocorrelation Function (SPACF) plot illustrates the partial correlation between the time series, representing the number of shares over time, and its lagged values, controlling for the effects of other lags. In the SPACF plot, significant spikes are observed at certain lags, notably lag 1. This spike suggests a direct influence of the first lagged value on the current observation, hinting at the possible presence of an autoregressive term. The significant spikes in the SPACF plot, particularly at lag 1, imply the potential presence of an autoregressive (AR) term in the model. This indicates that the current observation is influenced by its past values, particularly those at lag 1.

Combining our observations from SACF and SPACF plots, we concluded that the required model is an ARMA(1,1) model. We then used Corollary 3.7.7, to opt for the INGARCH(1,1) model:

$$\begin{cases} X_t \mid \mathcal{F}_{t-1} : \mathcal{P}(\lambda_t), \\ \lambda_t = \gamma_0 + \gamma_1 X_{t-1} + \delta_1 \lambda_{t-1}, \end{cases} \tag{6.3}$$

where γ_0, γ_1, and δ_1 are some parameters to be estimated. From the estimated parameter values obtained through the R statistical programming language, it was observed that $\gamma_0 = 1.7384$, representing the baseline level of the process. The parameter γ_1 was estimated to be 0.3271, indicating a moderately positive influence of the previous observation on the current one. Similarly, the parameter δ_1 was estimated at 0.2844, suggesting a moderate positive influence of the previous value of the latent variable on the current one. These estimates, along with their associated standard errors, offer valuable insights into the conditional distribution and dynamics of the process.

Figure 5.3 shows the daily number of shares trading on LuSE: original series and one-step-ahead forecasts with confidence limits. The blue line represents the original series, which depicts the actual number of shares over time. This serves as a reference for evaluating the accuracy of the forecast. The red line represents the one-step-ahead forecast generated by the model. This line indicates the predicted number of shares at each time point based on the model's parameters and past observations. The shaded grey region surrounding the forecasted line represents the 95% confidence interval. This interval provides a range within which we can reasonably expect the actual number of shares to fall, given the uncertainty associated with the forecast. From the plot, we observe that the one-step-ahead forecast closely follows the trajectory of the original series. This suggests that the model captures the underlying patterns and trends in the data effectively. The confidence interval provides valuable insights into the uncertainty associated with the forecast. As the forecast progresses further into the future, the width of the confidence interval may widen, reflecting increased uncertainty about future observations. Overall, the plot demonstrates the model's ability to generate accurate forecasts of the number of shares over time while accounting for uncertainty. However, further analysis, including model diagnostics and validation, may be necessary to assess the model's performance comprehensively.

CHAPTER 7

CONCLUSION AND RECOMMENDATIONS

In this final chapter, we summarise the main findings of our research and underscore their implications. We offer concrete recommendations based on our insights to guide future actions and initiatives. Ultimately, this chapter serves as a blueprint for leveraging our findings to enact positive change in the relevant domain.

7.1. Conclusion

Our investigation into the structure of the INGARCH(p, q) process has provided valuable insights into its dynamic behavior and applicability in modelling count data. By contrasting it with the classical GARCH(p, q) process, we emphasized the unique characteristics of the INGARCH(p, q) model, particularly its reliance on the Poisson distribution for conditional distribution and the sequential determination of parameters based on past values.

Through rigorous theoretical analysis and mathematical proofs, we have established crucial properties of the INGARCH process, including its stationarity, convergence properties, and distributional characteristics. Notably, the almost sure convergence, mean-square convergence, and strict stationarity of the building sequence highlight the stability and reliability of the INGARCH process over time. Furthermore, our examination of parameter estimation methods illuminated the challenges posed by the non-linear likelihood function, necessitating the use of numerical optimisation techniques.

In practical applications, our analysis of a real-data example involving the daily number of shares traded on the Lusaka Securities Exchange showcased the INGARCH model's effectiveness in capturing volatility patterns and generating accurate forecasts. Through detailed analysis of summary statistics and autocorrelations, we identified the model's suitability for the data and demonstrated its forecasting capabilities.

Overall, our findings contribute to a comprehensive understanding of the INGARCH(p, q) process, its structural properties, estimation techniques, and forecasting capabilities. These insights pave the way for further research and applications in modelling and analysing count data phenomena.

7.2. Recommendations

Based on the comprehensive analysis of the integer-valued generalised autoregressive conditional heteroscedasticity (INGARCH) process presented in the discussion chapter, the following concrete recommendations are proposed to guide future actions and initiatives:

1. **Model Understanding and Application**: Given the intricate nature of the INGARCH process and its relevance in overdispersed integer-valued time series analysis, it is imperative for researchers and practitioners to deepen their understanding of its theoretical underpinnings, estimation techniques, and practical applications. This involves thorough exploration of the model's definition, conditions for existence, and construction process as outlined in Sections 3.1 to 3.4, enabling better utilization of the model in diverse domains such as finance, economics, and epidemiology.

2. **Parameter Estimation and Inference**: Future research efforts should focus on refining parameter estimation techniques for the INGARCH model, particularly in real-world applications where accurate estimation is crucial for model reliability and predictive performance. Section 4.1 provides insights into the maximum likelihood estimation procedure and the significance of asymptotic properties, offering a roadmap for improving estimation accuracy and robustness. Additionally, leveraging techniques such as numerical optimisation and asymptotic analysis can enhance the efficiency and reliability of parameter estimation, facilitating more accurate inference and decision-making.

3. **Model Evaluation and Validation**: To ensure the effectiveness and applicability of the INGARCH model in real-world scenarios, rigorous model evaluation and validation procedures are essential. Researchers should undertake comprehensive diagnostic checks, sensitivity analyses, and validation tests to assess the model's goodness-of-fit, predictive accuracy, and stability. This includes examining the distributional properties, stationarity, and moment properties of the INGARCH process, as discussed in Sections 3.6 and 3.7, to validate its suitability for capturing the underlying dynamics of count data sets.

4. **Practical Implementation and Software Development**: To facilitate wider adoption and utilization of the INGARCH model, efforts should be directed towards developing user-friendly software packages and computational tools for model estimation, inference, and forecasting. Collaborative initiatives between researchers, software developers, and industry practitioners can lead to the development of intuitive software solutions that streamline the application of the INGARCH model in various fields. Furthermore, the provision of comprehensive documentation, tutorials, and case studies can enhance users' understanding and proficiency in utilizing the model effectively.

5. **Empirical Studies and Real-World Applications**: Empirical studies and real-world applications of the INGARCH model across different domains are essential for validating

its effectiveness, identifying potential limitations, and exploring new avenues for model refinement and extension. Researchers are encouraged to conduct empirical studies using diverse datasets, ranging from financial time series to epidemiological data, to demonstrate the model's versatility and robustness across various contexts. Additionally, case studies and practical applications can provide valuable insights into the model's performance in capturing complex phenomena and informing decision-making processes.

By adhering to these recommendations, researchers, practitioners, and policymakers can harness the full potential of the INGARCH model in analysing and forecasting overdispersed integer-valued time series data, thereby advancing knowledge, facilitating informed decision-making, and addressing real-world challenges effectively.

APPENDIX A

Data Set for the Illustration

In Chapter 5, the INGARCH model is illustrated on the daily number of shares trading on the Lusaka Securities Exchange (LuSE) from October 1, 2021 to May 10, 2022 (except for weekends and public holidays), giving a total of 150 observations.

Table A.1: Daily Volume on the Lusaka Securities Exchange

Month	Day	Volume	Month	Day	Volume	Month	Day	Volume
oct	1	650	nov	30	1500331	jan	25	52589
oct	4	2451	dec	1	320286	jan	26	12540
oct	5	900	dec	2	324365	jan	27	21756
oct	6	37816	dec	3	109610	jan	28	49189
oct	7	27100	dec	6	48900	jan	31	73454
oct	8	82000	dec	7	72201	feb	1	2614778
oct	11	826285	dec	8	6602	feb	2	8958498
oct	12	6353	dec	9	11560	feb	3	20163
oct	13	3127647	dec	10	21214	feb	4	2634287
oct	14	1910	dec	13	91453	feb	7	12554
oct	15	37846	dec	14	1810923	feb	8	51352
oct	19	18561	dec	15	10122	feb	9	3851
oct	20	4968	dec	16	48305	feb	10	40423
oct	21	801070	dec	17	345564	feb	11	9658
oct	22	224605	dec	20	200026	feb	14	40027
oct	26	836781	dec	21	13128	feb	15	66251
oct	27	311931	dec	22	5570	feb	16	89793
oct	28	206207	dec	23	13851	feb	17	14000
oct	29	5734	dec	24	2253	feb	18	68362
nov	1	448519	dec	27	16203	feb	21	63572
nov	2	117117	dec	28	17667	feb	22	16246
nov	3	31081	dec	29	212257	feb	23	75054
nov	4	312807	dec	30	900	feb	24	1238882
nov	5	156394	dec	31	24480	feb	25	18005

Table A.2: Daily Volume on the Lusaka Securities Exchange (continued)

Month	Day	Volume	Month	Day	Volume	Month	Day	Volume
nov	8	354601	jan	3	3206	feb	28	42951
nov	9	31808	jan	4	7300	mar	1	1334289
nov	10	4383470	jan	5	171928	mar	2	641900
nov	11	74364	jan	6	295978	mar	3	3210493
nov	12	1041636	jan	7	91466	mar	4	35027
nov	15	138172	jan	10	23867	mar	7	49442
nov	16	33061	jan	11	633538	mar	9	83060
nov	17	157937	jan	12	33083	mar	10	1254701
nov	18	77213	jan	13	25329	mar	11	13507
nov	19	248665	jan	14	34971	mar	14	25776
nov	22	195012	jan	17	13323	mar	15	589953
nov	23	299801	jan	18	2051803	mar	16	996227
nov	24	15581	jan	19	69386	mar	17	29499
nov	25	31370	jan	20	27558	mar	21	3443195
nov	26	43929	jan	21	3972528	mar	22	220694
nov	29	8385	jan	24	1238882	mar	23	7930
mar	24	20562	apr	7	11097	apr	25	547008
mar	25	46954	apr	8	3017330	apr	26	263077
mar	28	101160	apr	11	349124	apr	27	33254
mar	29	43575	apr	12	1433410	apr	29	14199
mar	30	34823	apr	13	5497	may	3	12023
mar	31	82831	apr	14	945176	may	4	10500
apr	1	471427	apr	19	36342	may	5	65768
apr	4	537177	apr	20	28277	may	6	17950
apr	5	17742	apr	21	97087	may	9	36728
apr	6	42574	apr	22	20818	may	10	65311

APPENDIX B

R Code for the Illustration

In Chapter 5, the INGARCH model is illustrated on the daily number of shares trading on the Lusaka Securities Exchange (LuSE) from October 1, 2021 to May 10, 2022 (except for weekends and public holidays), giving a total of 150 observations. We now give the R code for the computational aspect of the illustration.

```
# Define the data set
data <- c(650, 2451, 900, 37816, 27100, 82000, 826285, 6353,
3127647, 1910, 37846, 18561, 4968, 801070, 224605, 836781,
311931, 206207, 5734, 448519, 117117, 31081, 312807, 156394,
354601, 31808, 4383470, 74364, 1041636, 138172, 33061, 157937,
77213, 248665, 195012, 299801, 15581, 31370, 43929, 8385, 150033,
320286, 324365, 109610, 48900, 72201, 6602, 11560, 21214, 91453,
1810923, 10122, 48305, 345564, 200026, 13128, 5570, 13851, 2253,
16203, 17667, 212257, 900, 24480, 3206, 7300, 171928, 295978,
91466, 23867, 633538, 33083, 25329, 34971, 13323, 2051803, 69386,
27558, 3972528, 1238882, 52589, 12540, 21756, 49189, 73454,
2614778, 8958498, 20163, 2634287, 12554, 51352, 3851, 40423,
9658, 40027, 66251, 89793, 14000, 68362, 63572, 16246, 75054,
1238882, 18005, 42951, 1334289, 641900, 3210493, 35027, 49442,
83060, 1254701, 13507, 25776, 589953, 996227, 29499, 3443195,
220694, 7930, 20562, 46954, 101160, 43575, 34823, 82831, 471427,
537177, 17742, 42574, 11097, 3017330, 349124, 1433410, 5497,
945176, 36342, 28277, 97087, 20818, 547008, 263077, 33254, 14199,
12023, 10500, 65768, 17950, 36728, 65311)
```

R Code Implementation:

```
# Load necessary libraries
library(tscount)
library(forecast)

# Define the data set
data <- c(650, 2451, 900, ..., 65311)

# Fit Poisson INGARCH(1,1) model
model <- ingarchFit(data, p = 1, q = 1, dist = "poisson")
summary(model)

# One-step ahead forecasting
forecast <- predict(model, n.ahead = 1)
forecast_values <- forecast$pred
forecast_ci <- forecast$interval
```

B.1. Table 5.1: Summary Statistics

```
# Summary statistics
sample_size <- length(data)
maximum <- max(data)
minimum <- min(data)
mean <- mean(data)
median <- median(data)
variance <- var(data)
skewness <- skewness(data)
kurt <- kurtosis(data)

# Displaying summary statistics as a table
headers <- c("Statistic", "Value")
table <- data.frame(
  Statistic = c("Sample_Size", "Maximum", "Minimum", "Mean",
  "Median", "Variance", "Skewness", "Kurtosis"),
  Value = c(sample_size, maximum, minimum, mean,
  median, variance, skewness, kurt)
)

print(table)
```

B.2. Figure 5.1: Plot of original data

```
# Load necessary libraries
library(fINGARCH)
library(ggplot2)


# Fit INGARCH model
model <- ingarchFit(formula = ~ingarch(1, 1), data = data)

# Forecast one-step ahead
forecast <- predict(model, n.ahead = 1)

# Extract forecast values and confidence interval
forecast_values <- forecast$pred
forecast_ci <- forecast$interval

# Plot the time series and forecasts
plot.ts(data, main = "Number of Shares Traded Daily", xlab = "Date",
ylab = "Number of Shares", col = "blue")
lines(forecast_values, col = "red")
lines(forecast_ci[,1], col = "pink", lty = 2)
lines(forecast_ci[,2], col = "pink", lty = 2)
legend("topright", legend = c("Original Data", "Forecast", "95% CI"),
col = c("blue", "red", "pink"), lty = 1)
```

B.3. Figure 5.2: SACF and SPACF

```
# Load necessary libraries
library(stats)

# Plot the ACF and PACF of the series
par(mfrow=c(2,1), mar=c(4, 4, 2, 1))
acf(data, lag.max=20, main="ACF of the series")
pacf(data, lag.max=20, main="PACF of the series")
```

B.4. Table 5.2: Estimated parameter values

```
# Load necessary libraries
library(knitr)

# Create the data frame
data <- data.frame(
  Participant = c(1, 2, 3, 4),
 Estimated value = c(1.7383, 0.3271, 0.2844)
Standard error = c( 0.7721, 0.0628, 0.101285)

# Print the table
knitr::kable(data, caption = "Pretest and Posttest Scores")
```

B.5. Figure 5.3: One-step-ahead forecasts

```
# Model parameters
gamma0 <- 0.1
gamma1 <- 0.5
delta1 <- 0.5
# Function to calculate one-step ahead forecast and confidence interval
forecast <- function(data, gamma0, gamma1, delta1, alpha=0.05) {
  n <- length(data)
  X <- rep(0, n+1)
  lambda <- rep(0, n+1)
  forecast_values <- rep(0, n)
  upper_limits <- rep(0, n)
  lower_limits <- rep(0, n)

  X[1] <- data[1]
  lambda[1] <- gamma0

  for (t in 2:(n+1)) {
    lambda[t] <- gamma0 + gamma1 * X[t-1] + delta1 * lambda[t-1]
    X[t] <- rpois(1, lambda[t])
    if (t <= n) {
      forecast_values[t] <- lambda[t]
      std_error <- sqrt(lambda[t])
      z_score <- abs(qnorm(alpha/2))
      upper_limits[t] <- forecast_values[t] + z_score * std_error
      lower_limits[t] <- forecast_values[t] - z_score * std_error
```

```
    }
  }

  list(forecast_values, upper_limits, lower_limits)
}

# Generating forecast
forecast_result <- forecast(data, gamma0, gamma1, delta1)
forecast_values <- forecast_result[[1]]
upper_limits <- forecast_result[[2]]
lower_limits <- forecast_result[[3]]

# Plotting
plot(1:length(data), data, type='o', col='blue', xlab='Time',
ylab='Number of shares', main='Original Series and
One-Step Ahead Forecasts with Confidence Limits')
lines(forecast_values, col='red')
polygon(c(1:length(data), rev(1:length(data))),
c(upper_limits, rev(lower_limits)), col='grey', border=NA, alpha=0.5)
legend('topright', legend=c('Original Series',
'One-Step Ahead Forecast', '95% Confidence Interval'),
col=c('blue', 'red', 'grey'), lty=c(1, 1, 0), pch=c(1, NA, NA))
```

REFERENCES

[1] Abramowitz, M. and Stegun, I. (1965a). Handbook of Mathematical Functions. Dover Publications, Inc., New York.

[2] Abramowitz, M. and Stegun, I. (1965b). Handbook of Mathematical Functions. Dover Publications, Inc., New York.

[3] Ahmad, A. and Francq, C. (2015). Poisson QMLE of count time series models. Journal of Time Series Analysis, page DOI: 10.1111/jtsa.12167.

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

[5] Al-Osh, M. and Alzaid, A. (1988). Integer-valued moving average (INMA) process. Statist. Papers, 29:281–300.

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

[7] Bakouch, H. (2010). Higher-order moments, cumulants and spectral densities of the NGINAR(1) process. Statistical Methodology, 7:1–21.

[8] Bentarzi, M. and Bentarzi, W. (2015). Periodic Integer-Valued GARCH(1,1) Model. Communications in Statistics - Simulation and Computation, DOI:10.1080/03610918.2014.994780.

[9] Bollerslev, T. (1986). generalised autoregressive conditional heteroscedasticity. Journal of Econometrics, 37:307–327.

[10] Brännäs, K. and Hall, A. (2001). Estimation in integer-valued moving average models. Appl. Stochastic Models Bus. Ind., 17:277–291.

[11] Brockwell, P. and Davis, R. (1991). Time series: Theory and Methods. Springer-Verlag, New York, 2nd edition.

[12] Christou, V. and Fokianos, K. (2014). Quasi-likelihood inference for negative binomial time series models. Journal of Statistical Planning and Inference, 147:139–151.

[13] Consul, P. C. and Famoye, F. (2006). Lagrangian Probability Distributions. Birkhäuser, Boston.

[14] Cramér, H. (1946). Mathematical Methods of Statistics. Princeton University Press, Princeton.

[15] Davis, R. and Liu, H. (2015). Theory and inference for a class of observation-driven time

series models with application to time series of counts. Statistica Sinica, 25:409–431.

[16] Doukhan, P., Latour, A., and Oraichi, D. (2006). A simple Integer-Valued bilinear time series model. Adv. Appl. Prob., 38:559–578.

[17] Drost, F., van den Akker, R., and Werker, B. (2008). Note on Integer-Valued bilinear time series models. Stat. and Prob. Letters, 78:992–996.

[18] Du, J. and Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model. Journal of Time Series Analysis, 12:129–142.

[19] Efron, B. (1986). Double exponential families and their use in generalised linear regression. Journal of the American Statistical Association, 81:709–721.

[20] Elaydi, S. (2005). An Introduction to difference equations. Springer Science+Business Media, Inc., New York, 3rd edition.

[21] Engle, R. and Russell, J. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. Econometrica, 66(5):1127–1162.

[22] Feller, W. (1968). An Introduction to Probability Theory and Its Applications, volume 1. Wiley, New York, 3rd edition.

[23] Ferland, R., Latour, A., and Oraichi, D. (2006). Integer-valued GARCH process. Journal of Time Series Analysis, 27:923–942.

[24] Fokianos, K., Rahbek, A., and Tjøthiem, D. (2009). Poisson autoregression. Journal of the American Statistical Association, 104:1430–1439.

[25] Freeland, R. and McCabe, B. (2004). Forecasting discrete valued low count time series. International Journal of Forecasting, 20:427–434.

[26] Fristedt, B. and Gray, L. (1997). A Modern Approach to Probability Theory. Birkhauser, Boston.

[27] Gantmacher, F. (1959). The theory of matrices, volume I. Chelsea, New York.

[28] Gauthier, G. and Latour, A. (1994). Convergence forte des estimateurs des paramètres d'un processus GENAR(p). Annales des Sciences Mathématiques du Québec, 18:49–71.

[29] Gnedenko, B. and Kolmogorov, A. (1954). Limit Distributions for Sums of Independent Random Variables. Addison-Wesley, Reading, Mass.

[30] Gonçalves, E., Leite, J., and Mendes Lopes, N. (2009). A mathematical approach to detect the Taylor property in TARCH processes. Stat. and Prob. Letters, 72:602–610.

[31] Gourieroux, C. (1997) ARCH Models and Financial Applications. Springer Series in Statistics. New York: Springer.

[32] Graham, R. L., Knuth, D. E. and Patashnik, O. (1989) Concrete Mathematics. Reading, MA: Addison Wesley

[33] Hamilton, J. D. (1994) Time Series Analysis. Princeton, NJ: Princeton University Press.

[34] Hardy, E. (1996) Modelisation de type ARCH pour séries chronologiques a valeurs entieres. Master thesis, Montreal, QC: Département de Mathématiques, Université du Québec à Montréal.

[35] Heinen, A. (2001) modelling Time Series Count Data: The Autoregressive Conditional Poisson Model. PhD thesis, San Diego, CA: Department of Economics, University of California.

[36] Kedem, B. and Fokanios, K. (2002) Regression Models for Time Series Analysis. Wiley series in Probability and Statistics. Hoboken, NJ: Wiley-Interscience.

[37] Linton, O. and Terrell, D. (1998) Local linear regression with measurement error. Journal of Econometrics, 87(2):239–259.

[38] Macdonald, I. L. and Zucchini, W. (1997). Hidden Markov and Other Models for Discrete-valued Time Series. Boca Raton, LA: Chapman & Hall.

[39] Rydbery, T. H. and Shephard, N. (2000) BIN Models for Trade-by-Trade Data. Modelling the Number of Trades in a Fixed Interval of Time. Technical report 0740 , Econometric Society. Available at http://ideas.repec.org/p/ecm/wc2000/0740.html (accessed on 25th July 2006).

[40] Streett, S. B. (2000) Some observation-driven models for time series. PhD thesis, Fort Collins, CO: Department of Statistics, Colorado State University.

[41] Weiß, C. H. (2008). Thinning operations for modelling time series of counts – a survey. *Advances in Statistical Analysis*, 92, 319–341.

[42] Weiß, C. H. (2009). Modelling time series of counts with overdispersion. *Statistical Methods and Applications*, 18, 507–519.

[43] White, H. (1982) Maximum likelihood estimation of misspecified models. Econometrica $50, 1 - 25$.