이학박사 학위논문

Parameter Change Test for Time Series of Counts

이산형 시계열 자료에서의 모수 변화 검정

2017년 8월

서울대학교 대학원

통계학과

이 영 미

Parameter Change Test for Time Series of Counts

지도교수 이 상 열

이 논문을 이학박사 학위논문으로 제출함
2017년 4월

서울대학교 대학원
통계학과
이 영 미

이영미의 이학박사 학위논문을 인준함
2017년 6월

위 원 장 ＿＿＿＿ 오 희 석 ＿＿＿＿ (인)

부위원장 ＿＿＿＿ 이 상 열 ＿＿＿＿ (인)

위 원 ＿＿＿＿ 임 채 영 ＿＿＿＿ (인)

위 원 ＿＿＿＿ 김 영 원 ＿＿＿＿ (인)

위 원 ＿＿＿＿ 고 방 원 ＿＿＿＿ (인)

Parameter Change Test for Time Series of Counts

by

Youngmi Lee

A Thesis
submitted in fulfillment of the requirement
for the degree of
Doctor of Philosophy
in Statistics

The Department of Statistics
College of Natural Sciences
Seoul National University
August, 2017

Abstract

Parameter Change Test for Time Series of Counts

Youngmi Lee
The Department of Statistics
The Graduate School
Seoul National University

In this thesis, we consider parameter change test for time series of counts. First we consider the problem of testing for parameter change in zero-inflated generalized Poisson (ZIGP) autoregressive models. We verify that the ZIGP process is stationary and ergodic and that the conditional maximum likelihood estimator (CMLE) is strongly consistent and asymptotically normal. Then, based on these results, we construct CMLE- and residual-based cumulative sum tests and show that their limiting null distributions are a function of independent Brownian bridges. Simulation results are provided for illustration and a real data analysis is performed on data of crimes in Australia. Second we consider bivariate Poisson integer-valued GARCH(1,1) models, constructed via a trivariate reduction method of independent Poisson variables. We verify that the CMLE of the model parameters is asymptotically normal. Then, based on these results, we construct CMLE- and residual-based CUSUM tests and derive that their limiting null distributions are a function of independent Brownian bridges. A simulation study are conducted for illustration. We analyze two daily data sets of car accidents that occurred in Sungdong and Seocho counties in Seoul, Korea. Finally, we consider the problem of testing for a parameter change in general nonlinear integer-valued time series models where the conditional distribution of current observations is assumed to follow a one-parameter exponential family. We consider score-, (standardized) residual-, and estimate-based CUSUM tests, and show that their limiting null distributions take the form of the functions of Brownian bridges. Based on the obtained results, we then conduct a comparison study of the performance of CUSUM tests, through the use of Monte Carlo simulations. Our findings demonstrate that the standardized residual-based CUSUM test largely outperforms the others.

Keywords : Time series of counts; zero-inflated generalized Poisson autoregressive model; integer-valued GARCH model; test for parameter change; CUSUM test; weak convergence to a Brownian bridge; bivariate Poisson INGARCH model; exponential family; comparison of tests.

Student number : 2003-20356

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In recent years, integer-valued time series have attracted much attention from researchers because count data sets are frequently encountered in practice. Time series models of counts can be classified into two categories, that is, one using a thinning approach and the other using a generalized linear model (GLM) approach. The former includes the autoregressive moving average (ARMA)-type models based on a binomial thinning, referred to as integer-valued ARMA models; for example, see Alzaid and Al-Osh (1990), Jin-Guan and Yuan (1991),Al-Osh and Aly (1992), McKenzie (2003),and Weiß (2008). The latter is considered by Zeger and Qaqish (1988), Li (1994), Davis et al. (2000), Fahrmeir and Tutz (2001) and Jung et al. (2006). Further, regression models with an intensity process are considered by Ferland et al. (2006) and Fokianos et al. (2009). Sequences of counts appear in many other application fields such as statistical quality control (Weiß (2009)) and insurance (Zhu and Joe (2006)). See also Winkelmann (2008), who provides a survey of statistical and econometric techniques for count data based on conditional distribution models, and Jung and Tremayne (2011), who provide an overview of some recent developments in the analysis of time series of counts.

Time series of counts data are often overdispersed; that is, the variance is bigger than the mean of data. In this case, apparently, the Poisson distribution is not suitable for practical purposes. An alternative approach to modeling overdispersed count data is to employ integer-valued GARCH (INGARCH) models (cf. Ferland et al. (2006)). On the other hand, the generalized Poisson (GP) distribution introduced by Consul and Jain (1973) is a natural extension of the Poisson distribution when the data are overdispersed or even underdispersed. It is this flexibility that led many authors to study GP regression models (cf. Consul and Famoye (1992), Famoye (1993), and Famoye et al. (2004)). Recently, Zhu (2012a) studied the stationarity and ergodicity of the GP-INGARCH process and demonstrated the consistency and asymptotic normality of the conditional maximum likelihood estimator (CMLE); see also Jung and Tremayne (2011). On the other hand, the zero-inflated Poisson distribution is considered suitable for data with excess zeros: see Lambert (1992) and Gupta et al. (1995, 1996). Later, Zhu (2012b) studied zero-inflated (ZI) INGARCH models and investigated their model properties. In this study, we combine the GP-INGARCH and ZI-INGARCH models into one model.

For marginal distributions, some researchers also consider the use of distributions other than the Poisson distribution. For example, Davis and Wu (2009), Zhu (2011), and Christou and Fokianos (2014) consider negative binomial INGARCH (NB-INGARCH) models, Zhu (2012c) considers ConwayMaxwell Poisson distribution and Hudecová (2013) and Fokianos et al. (2014) consider the binary time series model. Davis and Liu (2016) recently extended the Poisson AR model to one-parameter exponential distribution AR models—called general nonlinear INGARCH (GN-INGARCH) models—thus establishing its stationarity and ergodicity, as well as the asymptotic properties of the conditional maximum likelihood estimator (CMLE) under some regularity conditions.

Compared to the univariate model, only a few consider bivariate (multivariate) integer-valued time series models. We can refer to Quoreshi (2006) and Pedeli and Karlis (2011, 2013a,b) who introduce the INAR type models, and Heinen (2003), Liu (2012) and Andreassen (2013) who investigate the INGARCH type models. Heinen and Rengifo (2003) suggest the multivariate AR conditional Poisson models based on a double Poisson distribution of Efron (1986). Liu (2012) consider bivariate Poisson INGARCH(p, q) models constructed via the trivariate reduction and prove the stationarity and ergodicity under certain conditions. Andreassen (2013) verifies the consistency of the conditional maximum-likelihood estimation (CMLE) of bivariate Poisson INGARCH(1,1) models. The Poisson INGARCH-type model has a limitation since it can only accommodate the non-negative dependence between two time series. To overcome this drawback, Heinen and Rengifo (2007) and Andreassen (2013) consider a copula approach. Heinen and Rengifo (2007) use the continued extension argument proposed by Denuit and Lambert (2005) to guarantee the uniqueness of the copula distribution. We focus on the bivariate Poisson INGARCH model of Liu (2012) in Chapter because the model is much more tractable in developing the CUSUM test.

Integer-valued time series in epidemiology are well-known to often undergo a change as the result of variations in quality of health care and state of patients' health. In general, inferences that ignore a parameter change can lead to a false conclusion, and thus, the detection of a parameter change is an important issue in practice. The problem of change point detection has been investigated by many authors: see Csörgö and Horváth (1997) for a general review. Among the existing change point tests, the cumulative sum (CUSUM) test has long been popular since it is easy to understand and implement in practice. The change point test for integer-valued time series has been studied by several authors: see Fokianos and Fried (2010, 2012), Hudecová (2013), and Fokianos et al. (2014). Further, Kang and Lee (2009) proposed a CUSUM test for detecting change points in random coefficient integer-valued autoregressive models with Poisson innovations and used it to analyze polio data. Franke et al. (2012) investigated a CUSUM test based on estimated residuals from Poisson autoregressive models with intensity $\lambda_t = f(Y_{t-1})$ for some real-valued function f. Doukhan et al. (2013) proposed the Poisson autoregressive models with intensity $\lambda_t = f(Y_{t-1}, Y_{t-2}, \ldots)$ and a change point test based on the likelihood of observations. See also Liu (2012) for a relevant reference. Recently, Kang and Lee (2014) investigated the change point test for Poisson autoregressive models with $\lambda_t = f_\theta(Y_{t-1}, \lambda_{t-1})$ (cf. Fokianos et al. (2009)) that include INGARCH(1,1) models. They suggested two types of CUSUM tests: an estimates-based test using the CMLE, and a residual-based test. In this study, we aim to extend their method to zero-inflated generalized Poisson autoregressive (ZIGP AR) models, the bivariate Poisson AR models and GN-INGARCH models. Compared to the previous study of Kang and Lee (2014), a more careful analysis is needed to obtain the asymptotic results owing to model complexity.

Although the estimates-based CUSUM test generally performs well, the estimates-based test occasionally suffers from severe size distortions; for this reason, it cannot be completely trusted (Kang and Lee (2014), Lee et al. (2016a,b)). In contrast, the residual-based test performs much more stably and produces reasonably good powers (Lee et al. (2004), Lee and Lee (2015)). However, its performance power is not always satisfactory, and a great power loss can occur, particularly when dealing with a parameter change in conditional locations (Oh and Lee (2017b)). As an alternative, one can use the score vector-based CUSUM test (Berkes et al. (2004), Oh and Lee (2017a)), because it might outperform the residual-based CUSUM test in terms of power. This study, inspired by the work of Oh and Lee (2017b), additionally considers the residual-based CUSUM test using the standardized residuals, as doing so can to a great extent enhance test performance in terms of power; this is seen in the results of their and our simulation studies.

This thesis is organized as followed. In Chapter 2, we review the results relevant to our subject handled in this thesis. In Chapter 3, we consider the zero-inflated generalized Poisson AR models. We establish the asymptotic results for CMLE, introduce the CUSUM tests based on estimates and residuals and derive their limiting null distributions. A simulation study and a real data analysis are presented fo illustration. In Chap 4, we introduces the bivariate Poisson INGARCH model and shows the asymptotic normality of CMLE. Further, we introduce the CUSUM test based on the estimates and residuals and derives their limiting null distributions. A simulation study and a real data analysis are conducted for illustration. In Chapter 5, we introduces the one-parameter exponential family AR models and establishes the asymptotic results for the CMLE. Further we introduces the CUSUM tests based on score vectors, (standardized) residuals, and estimates. After verifying their limiting null distributions, we implement a simulation study for comparison and illustrate a real data example.

Chapter 2

Literature Review

2.1 CUSUM test

Lee et al. (2003) proposed the Cusum test for detecting a parameter change. Their Cusum test turned out to be widely applicable to various time series models. Let $\{x_t : t \in Z\}$ be the stationary time series, and let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_J)^T$ be the parameter vector. We wish to test the following hypotheses based on the estimators $\hat{\boldsymbol{\theta}}_n$:

H_0 : $\boldsymbol{\theta}$ does not change over $x_1, \ldots, x_n$ vs.

H_1 : not H_0.

Let $\hat{\boldsymbol{\theta}}_k$ be the estimator of $\boldsymbol{\theta}$ based on $x_1, \ldots, x_k$. They investigate the differences $\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n$, $k = 1, \ldots, n$ for constructing a Cusum test. Suppose that

$$\sqrt{k}(\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}) = \frac{1}{\sqrt{k}} \sum_{t=1}^{k} \boldsymbol{l}_t + \boldsymbol{\Delta}_k,$$

where $\boldsymbol{l}_t := \boldsymbol{l}_t(\boldsymbol{\theta})$ forms stationary martingale differences and $\boldsymbol{\Delta}_k = (\Delta_{1,k}, \ldots, \Delta_{J,k})^T$. Let $\Gamma = Var(\boldsymbol{l}_t)$ be the covariance matrix of $\boldsymbol{l}_t$. Assuming that Γ is nonsingular, we get

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[ns]}\Gamma^{-1/2}\boldsymbol{l}_t \xrightarrow{w} \mathbf{W}_J(s)$$

in the $D^J[0,1]$ space, where $\mathbf{W}_J(s) = (W_1(s), \dots W_J(s))^T$ denotes a $J-$dimensional standard Brownian motion. Now, suppose that for each $j = 1, \dots, J$,

$$\max_{1\le k\le n}\frac{\sqrt{k}}{\sqrt{n}}|\Delta_{j,k}| = o_P(1).$$

Then,

$$\begin{aligned}&\frac{[ns]}{\sqrt{n}}\Gamma^{-1/2}(\hat{\boldsymbol{\theta}}_{ns} - \boldsymbol{\theta})\\ =\ &\frac{1}{\sqrt{n}}\sum_{t=1}^{[ns]}\Gamma^{-1/2}\boldsymbol{l}_t + \Gamma^{-1/2}\frac{\sqrt{[ns]}}{\sqrt{n}}\boldsymbol{\Delta}_{[ns]} \xrightarrow{w} \mathbf{W}_J(s),\end{aligned}$$

and consequently,

$$\frac{[ns]}{\sqrt{n}}\Gamma^{-1/2}(\hat{\boldsymbol{\theta}}_{[ns]} - \hat{\boldsymbol{\theta}}_n) \xrightarrow{w} \mathbf{W}_J^o(s),$$

where $\mathbf{W}_J^o(s) = (W_1^o, \dots W_J^o(s))^T$ is a $J-$dimensional Brownian bridge. Therefore, under H_0, the test statistic

$$T_n := \max_{1\le k\le n}\frac{k^2}{n}(\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n)^T\Gamma^{-1}(\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n).$$

satisfies

$$T_n \xrightarrow{w} \sup_{0\le s\le 1}\sum_{j=1}^{J}(W_j^o(s))^2.$$

We can obtain the critical region $\{T_n \ge c_\alpha\}$ given a nominal level α, where c_α is the empirical $(1-\alpha)$ quantile values for $\sup_{0\le s\le 1}||\mathbf{W}_J^0(s)||^2$. The critical values are provided in Table 1 of Lee et al. (2003). When H_0 is rejected, we estimate the location of the change point as $\hat{k} = \arg\max_{1\le k\le n}\frac{k^2}{n}(\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n)^T\Gamma^{-1}(\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n)$.

2.2 The Poisson AR models

The Poisson autoregressive model is defined by

$$Y_t|\mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t), \quad \lambda_t = f_\theta(\lambda_{t-1}, Y_{t-1}), \quad \forall t \in \mathbb{Z} \tag{2.1}$$

where f_θ is some known positive function depending on the parameter $\theta \in \Theta \subset \mathbb{R}^d$. The contraction condition is as follows : for all $\theta \in \Theta$,

$$|f_\theta(\lambda, y) - f_\theta(\lambda', y')| \leq \omega_1|\lambda - \lambda'| + \omega_2|y - y'|,$$

for all $\lambda, \lambda' \geq 0$ and $y, y' \in \mathbb{N}_0$, where $\omega_1, \omega_2 > 0$ and $\omega_1 + \omega_2 < 1$. Under the contraction condition, there exists a unique strictly stationary ergodic solution for (2.1) (cf. Neumann (2011)) which has finite moments of any order (cf. Fokianos et al. (2009)). Fokianos and Tjøstheim (2012) proved that the weak consistency and the asymptotic normality of the CMLE for Poisson AR model. The strong consistency of the CMLE is given by Kang and Lee (2014).

In particular, Poisson INGARCH (1,1) model is defined by

$$Y_t|\mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t), \quad \lambda_t = \omega + \alpha\lambda_{t-1} + \beta Y_{t-1}, \quad \forall t \in \mathbb{Z}$$

where $\omega > 0$, $\alpha \geq 0, \beta \geq 0$ and $\mathcal{F}_{t-1} = \sigma(Y_{t-1}, Y_{t-2}, \ldots)$. This models well describes the overdispersion, since

$$E(Y_t) := \mu = \frac{\omega}{1-(\alpha+\beta)} < \mu\left(1 + \frac{\beta^2}{1-(\alpha+\beta)^2}\right) = Var(Y_t).$$

Further, it has a strictly stationary ergodic solution and all moments of Y_t and λ_t are finite when $\alpha + \beta < 1$ which satisfies the contraction condition.

Chapter 3

Parameter Change Test for Zero-Inflated Generalized Poisson Autoregressive Models

3.1 Introduction

In this Chapter, we consider zero-inflated generalized Poisson autoregressive (ZIGP AR) models. We show that the ZIGP AR model is ergodic and stationary and that the CMLE is strongly consistent and asymptotically normal. Further, based on these, we drive the limiting distribution of the CUSUM test.

This Chapter is organized as follows. In Section 3.2, we introduce the ZIGP AR model and establish the asymptotic results for CMLE. In Section 3.3, we introduce the CUSUM tests based on estimates and residuals and derive their limiting null distributions. In Section 3.4, we present a simulation study for illustration. In Section 3.3.2, we apply our tests to a real data set and demonstrate the existence of a parameter change. In Section 3.6, concluding remarks are provided. All the related proofs are provided in the Appendix and the supplementary material.

3.2 Zero-inflated generalized Poisson AR model

A random variable Y has a ZIGP distribution with parameter λ, κ and ρ if

$$P(Y=y)=\begin{cases}\rho+(1-\rho)e^{-\lambda} & x=0\\ (1-\rho)\lambda(\lambda+\kappa y)^{x-1}e^{-(\lambda+\kappa y)}/y! & y=1,2,\cdots\\ 0 & \text{for } y>m \text{ if } \kappa<0\end{cases}$$

(cf. Gupta et al. (1995, 1996)), where $\lambda > 0, 0 \le \rho < 1, \max(-1, -\lambda/m) < \kappa < 1$, and $m(\ge 4)$ is the largest positive integer for which $\lambda + \kappa m > 0$. The above distribution reduces to a generalized Poisson distribution when $\rho = 0$ and to the ordinary Poisson distribution with mean λ when $\rho = 0$ and $\kappa = 0$.

Using (2.5) of Zhu (2012b), we obtain the moments of Y as

$$E(Y^s)=\sum_{i=0}^{s} a_{si}\lambda^i, \quad s=1,2,\ldots, \tag{3.1}$$

where a_{si} is not related to λ and $a_{ss} = (1-\rho)/(1-\kappa)^s$. In particular, we have

$$E(Y)=\frac{1-\rho}{1-\kappa}\lambda \text{ and } Var(Y)=(1-\rho)\left\{\frac{\rho\lambda^2}{(1-\kappa)^2}+\frac{\lambda}{(1-\kappa)^3}\right\}.$$

The variance of Y is greater than or equal to the mean depending on whether $0 < \kappa < 1$ or $\kappa = 0$, respectively. When $\kappa < 0$, the variance is less than the mean, provided $\rho = 0$.

Let $\{Y_t\}$ be a time series of counts with the conditional distribution following a ZIGP distribution; that is,

$$Y_t|\mathcal{F}_{t-1} \sim \text{ZIGP}(\lambda_t^*, \kappa, \rho), \quad (1-\rho)\frac{\lambda_t^*}{1-\kappa}=\lambda_t(\theta^*)=f_{\theta^*}(\lambda_{t-1}, Y_{t-1}), \tag{3.2}$$

where $0 \leq \rho < 1, \max(-1, -\lambda_t^*/m) < \kappa < 1, \mathcal{F}_{t-1}$ is the $\sigma-$field generated by $Y_{t-1}, \ldots, Y_0, \lambda_0$ and f_{θ^*} is a positive function on $[0, \infty) \times \mathbb{N}_0, \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, depending on the parameter $\theta^* \in \Theta^* \subset \mathbb{R}^d$ and irrespective of ρ and κ. Let

$$\theta = (\rho, \phi, \theta^{*T})^T$$

where $\phi = 1/(1-\kappa)$. Note that the poisson parameter in the conditional mean equation is expressed as $\lambda_t^*(\theta) = \frac{1-\kappa}{1-\rho}\lambda_t(\theta^*)$ wherein the ρ and κ are assumed to be constant in the same spirit as in Zhu (2012a,b). The true value of θ is denoted by $\theta_0 = (\rho_0, \phi_0, \theta^{*T}_0)^T$.

In what follows, we assume that

(**A1**) For all $\lambda, \lambda' \geq 0$ and $y, y' \in \mathbb{N}_0$,

$$\sup_{\theta^* \in \Theta^*} |f_{\theta^*}(\lambda, y) - f_{\theta^*}(\lambda', y')| \leq \omega_1|\lambda - \lambda'| + \omega_2|y - y'|,$$

where $\omega_1, \omega_2 \geq 0$ satisfying $\omega_1 + \omega_2 < 1$.

Based on the results of Neumann (2011), one can show the stationarity and ergodicity of the process under assumption (**A1**). More precisely, we can obtain the following (see the Appendix for its proof).

Theorem 3.1. *Suppose that the bivariate chain $((Y_t, \lambda_t))_{t\in\mathbb{N}}$ in model (3.2) satisfies* (**A1**). *Then, it holds that*

(i) *There exists a unique stationary distribution.*

(ii) *The process $((Y_t, \lambda_t))_{t\in\mathbb{N}}$ is ergodic.*

(iii) *The process $((Y_t, \lambda_t))_{t\in\mathbb{N}}$ belongs to $\mathcal{L}^s$ for each $s > 0$.*

The result of Theorem 3.1 plays an important role in establishing the asymptotic results addressed below.

The conditional likelihood function of model (3.2) is given by

$$\tilde{L}(\theta) = \prod_{t=1}^{n} \left[\tilde{L}_{00}(\theta) I(Y_t = 0) + \tilde{L}_{11}(\theta) I(Y_t \geq 1) \right],$$

where

$$\begin{aligned}
\tilde{L}_{00}(\theta) &= \rho + (1-\rho) e^{-\tilde{\lambda}_t^*} = \rho + (1-\rho) \exp\left\{ -\frac{\tilde{\lambda}_t(\theta^*)}{\phi(1-\rho)} \right\}, \\
\tilde{L}_{11}(\theta) &= (1-\rho) \frac{\tilde{\lambda}_t^* (\tilde{\lambda}_t^* + \kappa Y_t)^{Y_t - 1} e^{-(\tilde{\lambda}_t^* + \kappa Y_t)}}{Y_t!} \\
&= \frac{\tilde{\lambda}_t(\theta^*) \left\{ \tilde{\lambda}_t(\theta^*) + (1-\rho)(\phi - 1) Y_t \right\}^{Y_t - 1} (1-\rho)^{-(Y_t - 1)} \phi^{-Y_t}}{Y_t!} \\
&\qquad \times \exp\left[-\frac{1}{(1-\rho)\phi} \left\{ \tilde{\lambda}_t(\theta^*) + (1-\rho)(\phi - 1) Y_t \right\} \right],
\end{aligned}$$

and the $\tilde{\lambda}_t$ are defined recursively by

$$\tilde{\lambda}_t(\theta^*) = f_{\theta^*}(\tilde{\lambda}_{t-1}(\theta^*), Y_{t-1}), \quad t \geq 2,$$

with an arbitrarily chosen initial random variable $\tilde{\lambda}_1$. The CMLE of θ_0 is defined as

$$\hat{\theta}_n = \arg\max_{\theta \in \Theta} \tilde{L}(\theta) = \arg\max_{\theta \in \Theta} \tilde{L}_n(\theta),$$

where

$$\tilde{L}_n(\theta) = \log \tilde{L}(\theta) = \sum_{t=1}^{n} \tilde{\ell}_t(\theta) = \sum_{t=1}^{n} \left[\tilde{\ell}_{t0}(\theta) I(Y_t = 0) + \tilde{\ell}_{t1}(\theta) I(Y_t \geq 1) \right]$$

and

$$\tilde{\ell}_{t0}(\theta) = \log \left[\rho + (1-\rho) \exp\left\{ -\frac{\tilde{\lambda}_t(\theta^*)}{\phi(1-\rho)} \right\} \right],$$

$$\begin{aligned}\tilde{\ell}_{t1}(\theta) &= \log \tilde{\lambda}_t(\theta^*) + (Y_t - 1)\log\left\{\tilde{\lambda}_t(\theta^*) + (1-\rho)(\phi-1)Y_t\right\} - Y_t \log\phi \\ &\quad -(Y_t-1)\log(1-\rho) - \frac{1}{\phi(1-\rho)}\left\{\tilde{\lambda}_t(\theta^*) + (1-\rho)(\phi-1)Y_t\right\} - \log Y_t!.\end{aligned}$$

To ensure the strong consistency and asymptotic normality of $\hat{\theta}_n$, we assume the following conditions:

(**A2**) $\theta_0 \in \Theta$ and Θ is compact. In addition, there exist $0 < \delta_L < \delta_U < \infty$ such that

$$0 < \delta_L < \rho < 1 - \delta_L < 1, \;\; 0 < \delta_L < \phi < \delta_U < \infty,$$
$$f_{\theta^*}(\lambda, x) \geq \delta_L > 0, \quad \forall \theta^* \in \Theta^*, \; \lambda \geq 0, \; x \in \mathbf{N}_0.$$

(**A3**) $E\left(\sup_{\theta^*\in\Theta^*} \lambda_1(\theta^*)\right) < \infty$ and $E\left(\sup_{\theta^*\in\Theta^*} \tilde{\lambda}_1(\theta^*)\right) < \infty$.

(**A4**) $\lambda_t(\theta^*) = \lambda_t(\theta_0^*)$ a.s. implies $\theta^* = \theta_0^*$.

(**A5**) θ_0^* is an interior point of Θ^*.

(**A6**) $\lambda_t(\theta^*)$ is twice continuously differentiable with respect to θ and satisfies

$$E\left(\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\|\right)^4 < \infty \;\text{ and }\; E\left(\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial^2\lambda_t(\theta^*)}{\partial\theta^*\partial\theta^{*T}}\right\|\right)^2 < \infty.$$

(**A7**) Let V stand for a generic integrable random variable and $0 < \eta < 1$ be a generic constant. Then, for all t, a.s.

$$\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^*} - \frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\| \leq V\eta^t \;\text{ and }\; \sup_{\theta^*\in\Theta^*}\left\|\frac{\partial^2\tilde{\lambda}_t(\theta^*)}{\partial\theta^*\partial\theta^*T} - \frac{\partial^2\lambda_t(\theta^*)}{\partial\theta^*\partial\theta^{*\top}}\right\| \leq V\eta^t.$$

(**A8**) $\nu^T \frac{\partial\lambda_t(\theta_0^*)}{\partial\theta^*} = 0$ implies $\nu = 0$.

(**A9**) There exists $\delta_L^* > 0$ such that for all Y_t, a.s.,

$$\lambda_t(\theta^*) + (1-\rho)(\phi-1)Y_t \geq \delta_L^* > 0, \;\; \tilde{\lambda}_t(\theta^*) + (1-\rho)(\phi-1)Y_t \geq \delta_L^* > 0.$$

Remark 3.1. *(**A9**) is necessary to define the log-likelihood function $\ell_{t1}(\theta)$. If $\phi > 1$ (over-dispersion) and (**A2**) holds, (**A9**) is automatically satisfied. In the case of $0 < \phi < 1$ (under-dispersion), owing to the parameter restriction in ZIGP distribution, that is, $\max(-1, -\lambda/m) < \kappa < 1$, where m is the largest positive integer for which $\lambda + \kappa m > 0$, $\lambda_t(\theta^*) + (1-\rho)(\phi-1)m > 0$ is implicitly assumed, which also guarantees $\zeta_t := \lambda_t(\theta^*) + (1-\rho)(\phi-1)Y_t > 0$ for all t. In implementation, given $Y_1, \ldots, Y_n$, one may take $\delta_L^* = \min_{1 \le t \le n} \zeta_t > 0$. Concerning these assumptions, though, a careful treatment might be needed when dealing with the under-dispersion case in practice.*

Then, we can obtain Theorems 3.2 and 3.3 below the proofs of which are presented in the Appendix.

Theorem 3.2. *Under (**A1**)-(**A4**) and (**A9**), as $n \to \infty$*

$$\hat{\theta}_n \to \theta_0 \quad a.s..$$

Theorem 3.3. *Under (**A1**)-(**A9**), as $n \to \infty$*

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{w} N\left(0, I(\theta_0)^{-1}\right),$$

where

$$I(\theta_0) := E\left(\frac{\partial \ell_t(\theta_0)}{\partial \theta} \frac{\partial \ell_t(\theta_0)}{\partial \theta^T}\right) = -E\left(\frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta^T}\right)$$

and

$$\begin{aligned}
\ell_t(\theta) &= \ell_{t0}(\theta) I(Y_t = 0) + \ell_{t1}(\theta) I(Y_t \geq 1), \\
\ell_{t0}(\theta) &= \log\left[\rho + (1-\rho)\exp\left\{-\frac{\lambda_t(\theta^*)}{\phi(1-\rho)}\right\}\right], \\
\ell_{t1}(\theta) &= \log \lambda_t(\theta^*) + (Y_t - 1)\log\left\{\lambda_t(\theta^*) + (1-\rho)(\phi-1)Y_t\right\} - (Y_t - 1)\log(1-\rho) \\
&\quad - Y_t \log \phi - \frac{1}{\phi(1-\rho)}\left\{\lambda_t(\theta^*) + (1-\rho)(\phi-1)Y_t\right\} - \log Y_t!.
\end{aligned}$$

Remark 3.2. *When $f_{\theta^*}(\lambda, x) = \omega + \alpha\lambda + \beta x$, model (3.2) becomes an integer-valued GARCH (1,1) model. In this case, the results in Theorems 3.2 and 3.3 hold when $\alpha + \beta < 1$. The detailed proof is omitted for brevity.*

Remark 3.3. *Ferland et al. (2006) investigated the properties of Poisson-INGARCH (p, q) models; Doukhan and Kengne (2015) considered the change point test problem for the model with $\lambda_t = f_\theta(Y_{t-1}, \ldots)$; Zhu (2012a,b) verified the stationarity of ZIP-INGARCH(p, q) and GP-INGARCH(p, q) models. In view of these articles, naturally, one may consider extending our method to model (3.2) with*

$$\lambda_t = f_\theta(\lambda_{t-1}, \ldots, \lambda_{t-p}, Y_{t-1}, \ldots, Y_{t-q}).$$

Since this issue is somewhat beyond the scope of this paper, it is a topic of a different research project.

3.3 Change point test

In this section, we propose estimates- and residual-based CUSUM tests for detecting a parameter change in ZIGP AR models. We would also like to test the null and alternative hypotheses:

$$H_0 : \theta \text{ does not change over } Y_1, \ldots, Y_n \ \ vs. \ \ H_1 : \text{not } H_0.$$

3.3.1 Estimates-based CUSUM test

To implement our test, we employ the test statistic

$$T_n^{est} = \max_{1 \le k \le n} \frac{k^2}{n} (\hat{\theta}_k - \hat{\theta}_n)^T \hat{I}_n (\hat{\theta}_k - \hat{\theta}_n), \tag{3.3}$$

where $\hat{\theta}_k$ is the CMLE of θ_0 based on $Y_1, \ldots, Y_k$ and

$$\hat{I}_n = -\frac{1}{n}\sum_{t=1}^{n} \frac{\partial \tilde{\ell}_t(\hat{\theta}_n)}{\partial\theta\partial\theta^T}.$$

By using Taylor's theorem, provided $\frac{1}{n}\frac{\partial^2 \tilde{L}_n(\theta_n')}{\partial\theta^2}$ is nonsingular, we can show that

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = -\left(\frac{1}{n}\frac{\partial^2 \tilde{L}_n(\theta_n')}{\partial\theta\partial\theta^T}\right)^{-1} \frac{1}{\sqrt{n}}\frac{\partial \tilde{L}_n(\theta_0)}{\partial\theta},$$

where the θ_n' is between $\hat{\theta}_n$ and θ_0 and subsequently,

$$I(\theta_0)\cdot\sqrt{n}(\hat{\theta}_n - \theta_0) = \frac{1}{\sqrt{n}}\frac{\partial \tilde{L}_n(\theta_0)}{\partial\theta} - \left\{I(\theta_0) + \frac{1}{n}\frac{\partial^2 \tilde{L}_n(\theta_n')}{\partial\theta\partial\theta^T}\right\}\left(\frac{1}{n}\frac{\partial^2 \tilde{L}_n(\theta_n')}{\partial\theta\partial\theta^T}\right)^{-1}\frac{1}{\sqrt{n}}\frac{\partial \tilde{L}_n(\theta_0)}{\partial\theta}.$$

Therefore, for $0 < s < 1$,

$$I(\theta_0)\frac{[ns]}{\sqrt{n}}(\hat{\theta}_{[ns]} - \theta_0) = \frac{1}{n}\frac{\partial \tilde{L}_{[ns]}(\theta_0)}{\partial\theta} + \sqrt{\frac{[ns]}{n}}\tilde{\Delta}_{[ns]},$$

where

$$\tilde{\Delta}_k = \begin{cases} -\left\{I(\theta_0) + \frac{1}{k}\frac{\partial^2 \tilde{L}_k(\theta_k')}{\partial\theta\partial\theta^T}\right\}\left(\frac{1}{k}\frac{\partial^2 \tilde{L}_k(\theta_n')}{\partial\theta\partial\theta^T}\right)^{-1}\frac{1}{\sqrt{k}}\frac{\partial \tilde{L}_k(\theta_0)}{\partial\theta}, & \text{if } \left(\frac{1}{k}\frac{\partial^2 \tilde{L}_k(\theta_n')}{\partial\theta\partial\theta^T}\right)^{-1} \text{ exists} \\ \left\{I(\theta_0) + \frac{1}{n}\frac{\partial^2 \tilde{L}_n(\theta_n')}{\partial\theta\partial\theta^T}\right\}\sqrt{k}(\hat{\theta}_k - \theta_0), & \text{otherwise.} \end{cases}$$

According to the Proposition 3.6 in the Appendix , $\{\partial\ell_t(\theta_0)/\partial\theta; \mathcal{F}_t\}$ forms a martingale difference sequence under H_0. Thus, using the functional central limit theorem for martingales, we can show that

$$I(\theta_0)^{-1/2}\frac{1}{\sqrt{n}}\frac{\partial L_{[ns]}(\theta_0)}{\partial\theta} \xrightarrow{w} \mathbf{B}_{d+2}(s),$$

where $\{\mathbf{B}_{d+2}(s), 0 < s < 1\}$ is a $(d+2)$-dimensional standard Brownian motion. From Proposition 3.4 in the Appendix, we can see that

$$I(\theta_0)^{-1/2}\frac{1}{\sqrt{n}}\frac{\partial \tilde{L}_{[ns]}(\theta_0)}{\partial\theta} \xrightarrow{w} \mathbf{B}_{d+2}(s),$$

Further, using Egorov's theorem and Proposition 5 in the Appendix (cf. Lemma 9 of Kang and Lee (2014)), we obtain

$$\max_{1\le k\le n}\sqrt{\frac{k}{n}}\|\tilde{\Delta}_k\| = o_P(1).$$

Then, we have

$$I(\theta_0)\frac{[ns]}{\sqrt{n}}(\hat{\theta}_{[ns]}-\theta_0) = \frac{1}{n}\frac{\partial \tilde{L}_{[ns]}(\theta_0)}{\partial\theta} + \sqrt{\frac{[ns]}{n}}\tilde{\Delta}_{[ns]} \xrightarrow{w} \mathbf{B}_{d+2}(s),$$

and, subsequently,

$$I(\theta_0)^{1/2}\frac{[ns]}{\sqrt{n}}(\hat{\theta}_{[ns]}-\hat{\theta}_n) \xrightarrow{w} \mathbf{B}^{\circ}_{d+2}(s).$$

Therefore, we obtain the following.

Theorem 3.4. *Suppose that H_0 and* (**A1**)-(**A9**) hold. Then,

$$I(\theta_0)^{1/2}\frac{[ns]}{\sqrt{n}}(\hat{\theta}_{[ns]}-\hat{\theta}_n) \xrightarrow{w} \mathbf{B}^{\circ}_{d+2}(s),$$

where $\left\{\mathbf{B}^{\circ}_{d+2}(s), 0 < s < 1\right\}$ is a $(d+2)$-dimensional Brownian bridge. Further, $\hat{I}_n$ in (3.3) is a consistent estimator of $I(\theta_0)$, and thus,

$$T_n^{est} \xrightarrow{w} \sup_{0\le s\le 1}\|\mathbf{B}^{\circ}_{d+2}(s)\|^2.$$

Remark 3.4. *Recently, Doukhan and Kengne (2015) suggested a cusum test, say C_n, that measures the discrepancy between the parameter estimates based on the first k and remaining $n-k$ observations. This approach has merit in that their estimator of $I(\theta_0)$ is easily proven to converge to a positive definite matrix both under the null and alternative hypotheses, and as such, the CUSUM test can be shown to be consistent. The test C_n is also applicable to our model and its limiting distribution can be obtained similarly to T_n^{est}. To compare C_n and T_n^{est}, some simulation study* *is conducted in Section 3.4: see Tables 3.5 and 3.6. As seen therein,* C_n *does not outperform* T_n^{est} *in our set-up. However, it could interesting to compare its performance with the estimates- and residual-based tests in other situations based on different models.*

3.3.2 Residual-based CUSUM test

One can test for a change based on the residuals defined by $\epsilon_t = Y_t - \lambda_t(\theta_0)$, as in Franke et al. (2012) and Kang and Lee (2014). In this case, we use a test such as

$$T_n^{res} = \max_{1\le k\le n} \frac{1}{\sqrt{n}\hat{\tau}_n} \left| \sum_{t=1}^{k} \hat{\epsilon}_t - \left(\frac{k}{n}\right) \sum_{t=1}^{n} \hat{\epsilon}_t \right|,$$

where $\hat{\epsilon}_t = Y_t - \hat{\lambda}_t$ with $\hat{\lambda}_t = f_{\hat{\theta}_n^*}(\hat{\lambda}_{t-1}, Y_{t-1})$, an arbitrarily chosen initial random variable $\hat{\lambda}_1$ and $\hat{\tau}_n^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{\epsilon}_t^2$. Then, one can see that the following result holds (cf. the proof of Theorem 6 of Kang and Lee (2014)).

Theorem 3.5. *Under* **(A1)**-**(A9)** *and* H_0*, we have*

$$T_n^{res} \to \sup_{0\le s\le 1} |\mathbf{B}_1^{\circ}(s)|.$$

Remark 3.5. *As seen in the simulation study below, the residual-based test tends to be more stable than the estimates-based test. However, the latter merits to produce better powers than the former in many situations. For the residual-based CUSUM test for GARCH type models, see De Pooter and van Dijk (2004) and Lee et al. (2016a). Recently, Fokianos and Fried (2012), Hudecová (2013) and Kirch and Tadjuidje Kamgaing (2014) proposed residual-based score type tests for their own purposes. All these tests are worth further investigation in our set-up as well for a comparison study. Due to its importance, this issue is left as our future research project.*

3.4 Simulation results

In this section, we report simulation results to evaluate the performance of T_n^{est} and T_n^{res}. We consider the INGARCH(1,1) model:

$$Y_t|\mathcal{F}_{t-1} \sim \text{ZIGP}(\lambda_t^*, \kappa, \rho), \quad (1-\rho)\frac{\lambda_t^*}{1-\kappa} = \lambda_t(\theta^*) = \omega + \alpha\lambda_{t-1}(\theta^*) + \beta Y_{t-1},$$

where λ_1 is assumed be 0. In this simulation study, we set the nominal level as 0.05, the repetition number as 1,000, and the sample size as $n = 300, 500, 1000$. The critical values for T_n^{est} and T_n^{res} are 3.899 and 1.353, respectively, which are the ones obtained through a Monte Carlo simulation of the limiting null distribution, $\sup_{0\le s\le 1} \|\mathbf{B}_J^{\circ}\|^2$ (cf. Lee *et al.* (2003)). Since $\hat{\theta}_k$ for small k could be inaccurate, instead of (3.3), we use the test statistic:

$$T_{m,n}^{est} = \max_{k_L \le k \le n} \frac{k^2}{n}(\hat{\theta}_k - \hat{\theta}_n)^T \hat{I}_n(\hat{\theta}_k - \hat{\theta}_n),$$

which has the same asymptotic properties as (3.3). In our simulation study, we use $k_L = 20$.

To calculate the empirical size, we consider the INGARCH(1,1) model with $\rho =$ 0, 0.1, 0.3, $\phi = 1.2$, 1.5 (i.e., $\kappa = 1/6, 1/3$, respectively), and

$$\theta^* = (\omega, \alpha, \beta) = (1, 0.1, 0.2), (1, 0.1, 0.5), (1, 0.1, 0.8).$$

The empirical sizes are illustrated in Table 3.1. One can see that T_n^{res} has no severe size distortions. On the other hand, T_n^{est} exhibits some size distortions either when $\alpha+\beta \approx 1$ or $(\rho, \phi) \neq (0, 1)$, namely, the case other than the pure Poisson AR model. However, the size gets closer to the nominal level as the sample size increases. This shows that a fairly large sample size is needed to achieve the stability of the test. In our past experience (cf. Kang and Lee (2014), Na et al. (2012) and Song (2008)), this phenomenon has been frequently observed when performing the CUSUM test for GARCH-type models.

In order to examine the power, we consider the case in which parameter θ changes to θ' at the middle point:

- Case 1: θ^* changes to $\theta^{*\prime}$ and $(\rho, \phi) = (0, 1), (0.1, 1), (0, 1.2), (0.1, 1.2)$ does not change.
- Case 2: (ρ, ϕ) changes to (ρ', ϕ') and $(\omega, \alpha, \beta) = (1, 0.1, 0.2)$ does not change.

Tables 3.2-3.4 exhibit the empirical powers, wherein we can see that T_n^{est} produces good powers in many cases while T_n^{res} produces small powers in Case 1 when n= 300, 500 (see Tables 3.2, 3.3) and in Case 2 (see Table 3.4). In Case 1, the power of T_n^{res} gets closer to 1 as the sample size increases. Meanwhile, in Case 2, the power is small since the estimated residuals are less affected by the change of (ρ, ϕ). Figure 3.1 shows the time plots of the estimated studentized residuals of simulated data when $\theta = (0, 1, 1, 0.1, 0.2)$ changes to $(0.3, 1.2, 1, 0.1, 0.2), (0, 1, 1, 0.5, 0.2)$ and $(0.3, 1.2, 1, 0.5, 0.2)$ at the middle point, respectively. This indicates that the residuals tend to have more stable movements only when (ρ, ϕ) changes, which subsequently results in producing small values of T_n^{res}, and thus T_n^{res} is not favored when we conduct a test for a change in (ρ, ϕ).

Tables 3.5 and 3.6 show the sizes and powers of the test C_n with $\nu_n = u_n = (\log n)^{5/2}$ and $q(\cdot) \equiv 1$ (see Doukhan and Kengne (2015) for ν_n, u_n and q) and compare its performance with T_n^{est}. Here, we only report a few cases since the other cases show a similar pattern: the symbol * in Tables 3.1-3.4 denotes the cases chosen for the comparison. As mentioned in Remark 3.4, it is seen that C_n has severer size distortions and produces no better powers than the estimates-based test in our set-up. In fact, it can be checked that C_n does not completely outperform T_n^{res} in power as well excepting the case that the change only occurs in ρ and ϕ.

3.5 Real data analysis

In this section, we illustrate our method through a real data analysis. We analyze two monthly data of counts of robbery with firearms and assault police in Inner Sydney during the period January, 1995 to December, 2013 (the sample size = 228), recorded by the New South Wales (NSW) Police Force.

3.5.1 Number of robbery with firearms in Inner Sydney

First, we consider the data of counts of robbery with firearm. The empirical mean and variance of the data are 1.013 and 1.493, respectively. The time plot and histogram of the data are given in Figure 3.2. There are 99 zeros (43.42%) in series. The zero-inflates index defined in Puig and Valero (2006) is 0.2682, indicating that the series is zero inflated. Moreover, the data exhibit serial dependency; see the autocorrelation and partial autocorrelation samples shown in Figure 3.3. Thus, we fit a ZIGP-INGARCH(1,1) model to the data. Applying the change point test in Section 3.3, we obtain T_n^{est} =17.388 (see the vertical line in Figure 3.4), which suggests the rejection of the null hypothesis at the nominal level 0.05, that is, a parameter change occurs. On the other hand, we have $T_n^{res} = 1.242$, which is less than the critical value 1.353 at the nominal level 0.05 but is greater than the critical value 1.219 at the nominal level 0.1. Hence, it can be reasoned that a parameter change exists with a high possibility. The estimated parameters are summarized in Table 3.7. It can be seen that the number of zeros is 33(34.02%) in the first period and 66(50.38%) in the second period. Further, the zero-inflates index is -0.0162 for the first period and 0.2983 for the second period, indicating that only the data in the second period are zero-inflated. This evidence strongly advocates the existence of a parameter change.

3.5.2 Number of assault police in Inner Sydney

Next, we analyze the data of counts of assault police. The empirical mean and variance of the data are 22.877 and 46.751, respectively, which indicates that the time series is over-dispersed. The time plot of the data, the sample autocorrelation and the partial autocorrelation functions are shown in Figures 3.5 and 3.6. Since there are no zero observations, we only fit the GP-INGARCH(1,1) model to the data. In fact, the likelihood ratio test as in Zhu (2012a) indicates that a GP-INGARCH(1,1) model is favored, that is, $\phi > 1$ (overdispersed).

Applying the change point test in Section 3.3, we obtain $T_n^{est} = 5.004$ and $T_n^{res} = 1.566$, and thus conclude that a parameter change exists. Further, T_n^{est} is maximized at $k = 106$ (see the vertical line in Figure 3.7), which suggests that the change occurs in October 2003. The estimated parameters for the subseries before/after the change point are summarized in Table 3.8.

3.6 Concluding remarks

In this study, we carried out estimates- and residual-based CUSUM tests for ZIGP AR models and derived their limiting null distribution under regularity conditions. Compared to ordinary Poisson AR models, the ZIGP AR model has greater flexibility and is thus more suitable for analyzing a wider class of time series of counts.

We showed the stationarity and ergodicity of the ZIGP AR model and verified the strong consistency and asymptotic normality of the CMLE. Through a simulation study and real data analysis, we demonstrated that our test performs adequately and provides a functional tool to analyze a crime data set. The simulation result shows that a fairly large sample size is required to ensure the stability of the test, which is rather conventional according to our past experience in analyzing GARCH-type models. This shortcoming may be overcome by using a bootstrap method, but a careful analysis is required to justify its usage. Due to its importance, we leave this topic as a task of our future study.

3.7 Appendix

In this Appendix, we prove the theorems in the previous sections. The proofs of the propositions below are found in the supplementary material.

Proposition 3.1. *For arbitrary $\lambda_1, \lambda_2 \geq 0$, we can construct on an appropriate probability space $Y_i \sim ZIGP(\lambda_i^*, \kappa, \rho), \lambda_i = (1-\rho)\lambda_i^*/(1-\kappa)$ for $i = 1, 2$, such that*

$$E|Y_1 - Y_2| = |\lambda_1 - \lambda_2| \ \text{ and } \ P(Y_1 \neq Y_2) \leq |\lambda_1 - \lambda_2|.$$

Proof of Theorem 1 Points (i) and (ii) in Theorem 3.1 can be easily proved by Proposition 3.1 in the Appendix and Theorems 2.1 and 3.1 of Neumann (2011). Meanwhile, point (iii) can be verified by (3.1) and Theorem 2.1 of Doukhan et al. (2012). We omit the details for brevity. □

Proposition 3.2. *Under* **(A1)**-**(A3)** *and* **(A9)**, *we have*

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^{n}\tilde{\ell}_t(\theta) - \frac{1}{n}\sum_{t=1}^{n}\ell_t(\theta)\right| \to 0 \quad a.s..$$

Proposition 3.3. *Under* (**A1**)-(**A3**) *and* (**A9**), *we have*

$$(a)\ E\left(\sup_{\theta\in\Theta}\ell_t(\theta)\right)<\infty \quad \text{and} \quad (b)\ \theta\neq\theta_0 \Rightarrow E\ell_t(\theta)<E\ell_t(\theta_0).$$

Proof of Theorem 2. We express

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^{n}\tilde{\ell}_t(\theta)-\frac{1}{n}\sum_{t=1}^{n}\ell_t(\theta)\right|\leq\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^{n}\tilde{\ell}_t(\theta)-E\ell_t(\theta)\right|+\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^{n}\ell_t(\theta)-E\ell_t(\theta)\right|.$$

By Proposition 3.2, the first term of the RHS of the above inequality converges to 0 a.s.. Since $\ell_t(\theta)$ is the stationary and ergodic and $E(\sup_{\theta\in\Theta}\ell_t(\theta))<\infty$, the second term also converges to 0 a.s.. Therefore, the strong consistency can be asserted by Proposition 3.3. □

Note that the first derivatives are as follows:

$$\frac{\partial\ell_t(\theta)}{\partial\theta}=\frac{\partial\ell_{t0}(\theta)}{\partial\theta}I(Y_t=0)+\frac{\partial\ell_{t1}(\theta)}{\partial\theta}I(Y_t\geq 1),$$

with

$$\begin{aligned}
\frac{\partial\ell_{t0}(\theta)}{\partial\rho} &= \frac{1-\exp\{-A_{t0}(\theta)\}-\frac{\lambda_t(\theta^*)}{\phi(1-\rho)}\exp\{-A_{t0}(\theta)\}}{\rho+(1-\rho)\exp\{-A_{t0}\}},\\
\frac{\partial\ell_{t0}(\theta)}{\partial\phi} &= \frac{\lambda_t(\theta^*)\exp\{-A_{t0}(\theta)\}}{\phi^2\left[\rho+(1-\rho)\exp\{-A_{t0}(\theta)\}\right]},\\
\frac{\partial\ell_{t0}(\theta)}{\partial\theta^*} &= -\frac{\exp\{-A_{t0}\}}{\phi\left[\rho+(1-\rho)\exp\{-A_{t0}(\theta)\}\right]}\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*},
\end{aligned}$$

where $A_{t0}(\theta)=\lambda_t(\theta^*)/\{\phi(1-\rho)\}$, and

$$\frac{\partial\ell_{t1}(\theta)}{\partial\rho} = -\frac{(\phi-1)(Y_t-1)Y_t}{\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t}+\frac{Y_t-1}{1-\rho}-\frac{\lambda_t(\theta^*)}{\phi(1-\rho)^2},$$

$$\frac{\partial\ell_{t1}(\theta)}{\partial\phi} = \frac{(1-\rho)(Y_t-1)Y_t}{\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t}-\frac{Y_t}{\phi}+\frac{\lambda_t(\theta^*)}{\phi^2(1-\rho)}-\frac{Y_t}{\phi^2},$$

$$\frac{\partial \ell_{t1}(\theta)}{\partial \theta^*} = \left\{ \frac{1}{\lambda_t(\theta^*)} + \frac{Y_t - 1}{\lambda_t(\theta^*) + (1-\rho)(\phi-1)Y_t} - \frac{1}{\phi(1-\rho)} \right\} \frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}.$$

The second derivatives are also obtained as follows:

$$\frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^T} = \frac{\partial^2 \ell_{t0}(\theta)}{\partial \theta \partial \theta^T} I(Y_t = 0) + \frac{\partial^2 \ell_{t1}(\theta)}{\partial \theta \partial \theta^T} I(Y_t \geq 1)$$

with

$$\frac{\partial^2 \ell_{t0}(\theta)}{\partial \rho^2} = \frac{-1 + \left\{ 2 + 2A_{t0}(\theta) + \frac{\rho A_{t0}^2(\theta)}{(1-\rho)} \right\} e^{-A_{t0}(\theta)} - \{1 + 2A_{t0}(\theta)\} e^{-2A_{t0}(\theta)}}{\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2},$$

$$\frac{\partial^2 \ell_{t0}(\theta)}{\partial \phi^2} = \frac{1-\rho}{\phi^2} \cdot \frac{\{A_{t0}^2(\theta) - 2A_{t0}(\theta)\} \rho e^{-A_{t0}(\theta)} - 2(1-\rho)A_{t0}(\theta)e^{-2A_{t0}(\theta)}}{\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2},$$

$$\begin{aligned} \frac{\partial^2 \ell_{t0}(\theta)}{\partial \theta^* \partial \theta^{*T}} &= \frac{\rho e^{-A_{t0}(\theta)}}{\phi^2(1-\rho)\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2} \cdot \frac{\partial \lambda_t(\theta^*)}{\partial \theta^*} \frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}} \\ &\quad - \frac{e^{-A_{t0}(\theta)}}{\phi\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]} \cdot \frac{\partial^2 \lambda_t(\theta^*)}{\partial \theta^* \partial \theta^{*T}}, \end{aligned}$$

$$\frac{\partial^2 \ell_{t0}(\theta)}{\partial \rho \partial \phi} = \frac{-\rho A_{t0}^2(\theta)e^{-A_{t0}(\theta)} - (1-\rho)A_{t0}(\theta)\left\{e^{-A_{t0}(\theta)} - e^{-2A_{t0}(\theta)}\right\}}{\phi\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2},$$

$$\frac{\partial^2 \ell_{t0}(\theta)}{\partial \rho \partial \theta^{*T}} = \frac{\rho A_{t0}(\theta)e^{-A_{t0}(\theta)} + (1-\rho)\left\{e^{-A_{t0}(\theta)} - e^{-2A_{t0}(\theta)}\right\}}{\phi(1-\rho)\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2} \cdot \frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}},$$

$$\frac{\partial^2 \ell_{t0}(\theta)}{\partial \phi \partial \theta^{*T}} = \frac{\{1 - A_{t0}(\theta)\} \rho e^{-A_{t0}(\theta)} + (1-\rho)e^{-2A_{t0}(\theta)}}{\phi^2\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2} \cdot \frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}}$$

and

$$\frac{\partial^2 \ell_{t1}(\theta)}{\partial \rho^2} = -\frac{(\phi-1)^2(Y_t-1)Y_t^2}{\{\lambda_t(\theta^*) + (1-\rho)(\phi-1)Y_t\}^2} + \frac{Y_t - 1}{(1-\rho)^2} - \frac{2\lambda_t(\theta^*)}{\phi(1-\rho)^3},$$

$$\frac{\partial^2 \ell_{t1}(\theta)}{\partial \phi^2} = -\frac{(1-\rho)^2(Y_t-1)Y_t^2}{\{\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t\}^2} + \frac{Y_t}{\phi^2} - \frac{2\lambda_t(\theta^*)}{\phi^3(1-\rho)} + \frac{2Y_t}{\phi^3},$$

$$\frac{\partial^2 \ell_{t1}(\theta)}{\partial \rho \partial \phi} = -\frac{\lambda_t(\theta^*)Y_t(Y_t-1)}{\{\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t\}^2} + \frac{\lambda_t(\theta^*)}{\phi^2(1-\rho)^2},$$

$$\frac{\partial^2 \ell_{t1}(\theta)}{\partial \rho \partial \theta^{*T}} = \left[\frac{(\phi-1)Y_t(Y_t-1)}{\{\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t\}^2} - \frac{1}{\phi(1-\rho)^2}\right]\frac{\partial \lambda_t(\theta^*)}{\partial \theta^*},$$

$$\frac{\partial^2 \ell_{t1}(\theta)}{\partial \phi \partial \theta^{*T}} = \left[-\frac{(1-\rho)Y_t(Y_t-1)}{\{\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t\}^2} + \frac{1}{\phi^2(1-\rho)}\right]\frac{\partial \lambda_t(\theta^*)}{\partial \theta^*},$$

$$\begin{aligned}\frac{\partial^2 \ell_{t1}(\theta)}{\partial \theta^* \partial \theta^{*T}} &= -\left[\frac{Y_t-1}{\{\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t\}^2} + \frac{1}{\lambda_t^2(\theta^*)}\right]\frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}\frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}} \\ &\quad + \left\{\frac{1}{\lambda_t(\theta^*)} + \frac{Y_t-1}{\lambda_t+(1-\rho)(\phi-1)Y_t} - \frac{1}{\phi(1-\rho)}\right\}\frac{\partial^2 \lambda_t(\theta^*)}{\partial \theta^* \partial \theta^{*T}}.\end{aligned}$$

Proposition 3.4. *Under* **(A1)**-**(A7)** *and* **(A9)**, *for* $i = 0, 1$,

$$\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{\partial \tilde{\ell}_{ti}(\theta_0)}{\partial \theta} - \frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{\partial \ell_{ti}(\theta_0)}{\partial \theta}\right\| = o_P(1).$$

Proposition 3.5. *Under* **(A1)**-**(A7)** *and* **(A9)**, *we have*

$$-\frac{1}{n}\frac{\partial^2 \tilde{L}_n(\theta'_n)}{\partial \theta \partial \theta^T} \to I(\theta_0) \quad a.s.$$

where θ'_n *is an intermediate point between* θ_0 *and* $\hat{\theta}_n$.

Proposition 3.6. *Assume that* **(A1)**-**(A3)** *and* **(A6)** *hold. Then,* $\{\partial \ell_t(\theta_0)/\partial \theta\,;\, \mathcal{F}_t\}$ *forms a stationary ergodic martingale difference sequence.*

Proof of Theorem 3. Since $\{\partial \ell_t(\theta_0)/\partial \theta\,;\, \mathcal{F}_t\}$ forms a martingale difference sequence according to Proposition 3.6, we can show that $\frac{1}{\sqrt{n}}\sum_{t=1}^n \partial \ell_t(\theta_0)/\partial \theta$ converges weakly to $N(0, I(\theta_0))$ by using a martingale central limit theorem and the Cramér-Wold device. Then, using Taylor's theorem and Propositions 3.4 and 3.5, we can assert the theorem. □

3.8 Supplementary Material

In this supplementary material, we provide the proofs of the propositions in the Appendix.

Proof of Propositon 3.1 Let

$$X_1 = BY + (1-B)N_1, \;\; X_2 = BY + (1-B)(N_1 + Z),$$

where $B \sim Bin(1,\rho), P(Y=0) = 1, N_1 \sim GP(\lambda_1^*, \kappa), Z \sim GP(\lambda_2^* - \lambda_1^*, \kappa)$ and $B, Y, Z,$ and N_1 are independent r.v.s. Then, X_1 and X_2 have a $ZIGP$ distribution, and thus, we have

$$\begin{aligned} &E|X_2 - X_1| = E(1-B)Z = |\lambda_2 - \lambda_1|, \\ &P(X_1 \neq X_2) = P\left((1-B)Z \neq 0\right) = P(1-B \neq 0, Z \neq 0) \leq |\lambda_2 - \lambda_1|. \end{aligned}$$

This completes the proof. □

Lemma 3.1. *Under* **(A1)**-**(A3)**, *we have*

$$\sup_{\theta^* \in \Theta^*} \left| \tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*) \right| \leq V \eta^t \quad a.s..$$

Proof. From **(A1)**, we have

$$\left| \tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*) \right| \leq \omega_1 \left| \tilde{\lambda}_{t-1}(\theta^*) - \lambda_{t-1}(\theta^*) \right| \leq \omega_1^{t-1} \left| \tilde{\lambda}_1(\theta^*) - \lambda_1(\theta^*) \right|.$$

This completes the proof. □

Proof of Propositon 3.2 It suffices to show that for $i = 0, 1$,

$$\sup_{\theta \in \Theta} \left| \tilde{\ell}_{ti}(\theta) - \ell_{ti}(\theta) \right| \to 0 \;\; a.s., \;\; \text{as} \;\; t \to \infty.$$

Since $|\tilde{\ell}_{ti}(\theta) - \ell_{ti}(\theta)| = (\tilde{\ell}_{ti}(\theta) - \ell_{ti}(\theta))^+ + (\ell_{ti}(\theta) - \tilde{\ell}_{ti}(\theta))^+$ for $i = 0, 1$, we first show that for $i = 0, 1$, as $t \to \infty$,

$$\sup_{\theta \in \Theta} \left\{ \tilde{\ell}_{ti}(\theta) - \ell_{ti}(\theta) \right\}^+ \to 0 \;\; \text{a.s.}. \tag{3.4}$$

By the mean value theorem and using the fact that $\log x \leq x - 1$, we have

$$\begin{aligned}
\left\{ \tilde{\ell}_{t0}(\theta) - \ell_{t0}(\theta) \right\}^+ &\leq \left\{ \tilde{\ell}_{t0}(\theta) - \ell_{t0}(\theta) \right\} \vee 0 \\
&\leq \frac{1-\rho}{\rho} \left| \exp\left\{ -\frac{\tilde{\lambda}_t(\theta^*)}{\phi(1-\rho)} \right\} - \exp\left\{ -\frac{\lambda_t(\theta^*)}{\phi(1-\rho)} \right\} \right| \\
&\leq \frac{1-\rho}{\rho} \cdot \frac{\left| \tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*) \right|}{\phi(1-\rho)} \cdot \exp\left\{ -\frac{\lambda_t^*(\theta^*)}{\phi(1-\rho)} \right\} \leq \frac{1}{\phi\rho} \left| \tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*) \right|
\end{aligned}$$

for some intermediate points $\lambda_t^*(\theta^*)$ between $\tilde{\lambda}_t(\theta^*)$ and $\lambda_t(\theta^*)$. Thus, by Lemma 3.1,

$$\sup_{\theta \in \Theta} \left\{ \tilde{\ell}_{t0}(\theta) - \ell_{t0}(\theta) \right\}^+ \to 0 \;\; \text{a.s.} \;\; \text{as} \;\; t \to \infty$$

Meanwhile, by (**A2**) and Lemma 1, we have

$$\begin{aligned}
\left\{ \tilde{\ell}_{t1}(\theta) - \ell_{t1}(\theta) \right\}^+ &\leq \left\{ \tilde{\ell}_{t1}(\theta) - \ell_{t1}(\theta) \right\} \vee 0 \\
&\leq \left| \tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*) \right| \left| \frac{1}{\lambda_t(\theta^*)} + \frac{Y_t - 1}{\lambda_t(\theta^*) + (1-\rho)(\phi - 1)Y_t} - \frac{1}{\phi(1-\rho)} \right| \\
&\leq \left\{ \frac{Y_t + 1}{\delta_L^*} + \frac{2}{\delta_L^2} \right\} V\eta^t.
\end{aligned}$$

To show (3.4), it suffices to show that $Y_t\eta^t \to 0$ $a.s.$ as $t \to \infty$. By using the Markov inequality and the stationarity of $\{Y_t\}$, we have

$$\sum_{t=1}^{\infty} P\left(Y_t\eta^t > \epsilon\right) \leq \sum_{t=1}^{\infty} \frac{E(Y_t\eta^t)^s}{\epsilon^s} < \infty$$

for some s in Theorem 3.1. Hence, by the Borel-Cantelli lemma, we obtain $Y_t\eta^t \to 0$ $a.s.$. Similarly, it can be seen that for $i = 0, 1$,

$$\sup_{\theta\in\Theta}\left\{\tilde{\ell}_{ti}(\theta) - \ell_{ti}(\theta)\right\}^- \to 0 \text{ a.s. as } t \to \infty.$$

This validates the lemma. □

Proof of Propositon 3.3 To prove (a), it suffices to show that for $i = 0, 1$,

$$E\left(\sup_{\theta\in\Theta}|\ell_{ti}(\theta)|\right) < \infty.$$

From **(A3)** and the stationarity of $\lambda_t(\theta^*)$, we have

$$E\left(\sup_{\theta^*\in\Theta^*}\lambda_t(\theta^*)\right) < \infty. \tag{3.5}$$

Further, from **(A2)**, we have

$$E\left(\sup_{\theta\in\Theta}|\ell_{t0}(\theta)|\right) \le -\log\delta_L < \infty.$$

Note that

$$\begin{aligned} E\left(\sup_{\theta\in\Theta}|\ell_{t1}(\theta)|\right) &\le E\left(\sup_{\theta^*\in\Theta^*}|\log\lambda_t(\theta^*)|\right) \\ &+E\left[|Y_t-1|\sup_{\theta\in\Theta}|\log\{\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t\}|\right] \\ &+E|Y_t-1|\sup|\log(1-\rho)| + E|Y_t|\sup|\log\phi| \\ &+\sup\left\{\frac{1}{\phi(1-\rho)}\right\}E\left[\sup_{\theta\in\Theta}\{\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t\}\right]+E(\log Y_t!). \end{aligned} \tag{3.6}$$

By using the fact that $\log x \le x - 1$, it can be seen that

$$\sup_{\theta\in\Theta}|\log\lambda_t(\theta^*)| \le -\log\delta_L I(\delta_L \le 1) + \sup\lambda_t(\theta^*). \tag{3.7}$$

Therefore, from **(A2)**, (3.5), (3.7) and (iii) in Theorem 1, we can verify that the LHS of (3.6) is finite, except for the second term. We can show that the second term of the LHS (3.6) is finite by using the same method as in Zhu and Wang (2011). Note that

$$\begin{aligned} &E\left[|Y_t-1|\sup_{\theta\in\Theta}|\log\{\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t\}|\right] \\ &\qquad\le E\left[|Y_t-1|\sup\{-I(A_{t1}(\theta)\le 1)\log A_{t1}(\theta)\}\right] \\ &\qquad+E\left[|Y_t-1|\sup\{I(1<A_{t1}(\theta)\le e)\log A_{t1}(\theta)\}\right] \\ &\qquad+E\left[|Y_t-1|\sup\{I(e<A_{t1}(\theta))\log A_{t1}(\theta)\}\right], \end{aligned} \tag{3.8}$$

where $A_{t1}(\theta)=\lambda_t(\theta^*)+(1-\rho)(\phi-1)Y_t$. Since $\log\delta_L^*<\log A_{t1}(\theta)<1$ when $A_{t1}(\theta)<e$, the first and second terms of LHS in (3.8) is finite. Meanwhile, using the Cauchy-Schwarz inequality, Jenson's inequality and (3.5), we get

$$\begin{aligned} &E\left[|Y_t-1|\sup_{\theta\in\Theta}\{I(e<A_{t1}(\theta))\log A_{t1}(\theta)\}\right] \\ &\quad\le\left\{E(Y_t-1)^2\right\}^{1/2}\left[E\left\{\sup_{\theta\in\Theta}I(A_{t1}(\theta)>e)\log A_{t1}(\theta)\right\}^2\right]^{1/2} \\ &\quad\le\left\{E(Y_t-1)^2\right\}^{1/2}\left[E\left\{I\left(\sup_{\theta\in\Theta}A_{t1}(\theta)>e\right)\log\left(\sup_{\theta\in\Theta}A_{t1}(\theta)\right)\right\}^2\right]^{1/2} \\ &\quad\le\left\{E(Y_t-1)^2\right\}^{1/2}\left[\log^2E\left\{\sup_{\theta\in\Theta}A_{t1}(\theta)I\left(\sup_{\theta\in\Theta}A_{t1}(\theta)>e\right)\right\}\right]^{1/2} \\ &\quad\le\left\{E(Y_t-1)^2\right\}^{1/2}\left[\log^2E\left\{\sup_{\theta\in\Theta}A_{t1}(\theta)\right\}\right]^{1/2}<\infty. \end{aligned}$$

This asserts (a).

For (b), we use $\log x\le x-1$ to get

$$\begin{aligned} E_{\theta_0}\{\ell_t(\theta)-\ell_t(\theta_0)\} &= E_{\theta_0}\{\log g(Y_1:\theta)-\log g(Y_1:\theta_0)\} \\ &= E_{\theta_0}\left\{\log\frac{g(Y_1:\theta)}{g(Y:\theta_0)}\right\} \\ &\le E_{\theta_0}\left\{\frac{g(Y_1:\theta)}{g(Y_1:\theta_0)}\right\}-1 \end{aligned}$$

$$
\begin{aligned}
&= \sum_y \frac{g(y:\theta)}{g(y:\theta_0)} g(y:\theta_0)dy - 1 \\
&= \sum_y g(y:\theta)dy - 1 = 0
\end{aligned}
$$

where $g(y:\theta)$ is the probability density function of Y_1. This completes the proof.□

Note that the first derivatives are as follows:

$$
\frac{\partial \ell_t(\theta)}{\partial \theta} = \frac{\partial \ell_{t0}(\theta)}{\partial \theta} I(Y_t = 0) + \frac{\partial \ell_{t1}(\theta)}{\partial \theta} I(Y_t \geq 1),
$$

with

$$
\begin{aligned}
\frac{\partial \ell_{t0}(\theta)}{\partial \rho} &= \frac{1 - \exp\{-A_{t0}(\theta)\} - \frac{\lambda_t(\theta^*)}{\phi(1-\rho)} \exp\{-A_{t0}(\theta)\}}{\rho + (1-\rho)\exp\{-A_{t0}\}}, \\
\frac{\partial \ell_{t0}(\theta)}{\partial \phi} &= \frac{\lambda_t(\theta^*)\exp\{-A_{t0}(\theta)\}}{\phi^2\left[\rho + (1-\rho)\exp\{-A_{t0}(\theta)\}\right]}, \\
\frac{\partial \ell_{t0}(\theta)}{\partial \theta^*} &= -\frac{\exp\{-A_{t0}\}}{\phi\left[\rho + (1-\rho)\exp\{-A_{t0}(\theta)\}\right]} \frac{\partial \lambda_t(\theta^*)}{\partial \theta^*},
\end{aligned}
$$

where $A_{t0}(\theta) = \lambda_t(\theta^*)/\{\phi(1-\rho)\}$, and

$$
\begin{aligned}
\frac{\partial \ell_{t1}(\theta)}{\partial \rho} &= -\frac{(\phi-1)(Y_t-1)Y_t}{A_{t1}(\theta)} + \frac{Y_t - 1}{1-\rho} - \frac{\lambda_t(\theta^*)}{\phi(1-\rho)^2}, \\
\frac{\partial \ell_{t1}(\theta)}{\partial \phi} &= \frac{(1-\rho)(Y_t-1)Y_t}{A_{t1}(\theta)} - \frac{Y_t}{\phi} + \frac{\lambda_t(\theta^*)}{\phi^2(1-\rho)} - \frac{Y_t}{\phi^2}, \\
\frac{\partial \ell_{t1}(\theta)}{\partial \theta^*} &= \left\{\frac{1}{\lambda_t(\theta^*)} + \frac{Y_t - 1}{A_{t1}(\theta)} - \frac{1}{\phi(1-\rho)}\right\} \frac{\partial \lambda_t(\theta^*)}{\partial \theta^*},
\end{aligned}
$$

where $A_{t1}(\theta) = \lambda_t(\theta^*) + (\phi - 1)(1-\rho)Y_t$.

The second derivatives are also obtained as follows:

$$
\frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^T} = \frac{\partial^2 \ell_{t0}(\theta)}{\partial \theta \partial \theta^T} I(Y_t = 0) + \frac{\partial^2 \ell_{t1}(\theta)}{\partial \theta \partial \theta^T} I(Y_t \geq 1)
$$

with

$$
\frac{\partial^2 \ell_{t0}(\theta)}{\partial \rho^2} = \frac{-1 + \left\{2 + 2A_{t0}(\theta) + \frac{\rho A_{t0}^2(\theta)}{(1-\rho)}\right\} e^{-A_{t0}(\theta)} - \{1 + 2A_{t0}(\theta)\} e^{-2A_{t0}(\theta)}}{\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2},
$$

$$\frac{\partial^2 \ell_{t0}(\theta)}{\partial \phi^2} = \frac{1-\rho}{\phi^2} \cdot \frac{\left\{A_{t0}^2(\theta) - 2A_{t0}(\theta)\right\} \rho e^{-A_{t0}(\theta)} - 2(1-\rho)A_{t0}(\theta)e^{-2A_{t0}(\theta)}}{\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2},$$

$$\begin{aligned}\frac{\partial^2 \ell_{t0}(\theta)}{\partial \theta^* \partial \theta^{*T}} &= \frac{\rho e^{-A_{t0}(\theta)}}{\phi^2(1-\rho)\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2} \cdot \frac{\partial \lambda_t(\theta^*)}{\partial \theta^*} \frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}} \\ &\quad - \frac{e^{-A_{t0}(\theta)}}{\phi\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]} \cdot \frac{\partial^2 \lambda_t(\theta^*)}{\partial \theta^* \partial \theta^{*T}},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 \ell_{t0}(\theta)}{\partial \rho \partial \phi} &= \frac{-\rho A_{t0}^2(\theta)e^{-A_{t0}(\theta)} - (1-\rho)A_{t0}(\theta)\left\{e^{-A_{t0}(\theta)} - e^{-2A_{t0}(\theta)}\right\}}{\phi\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2}, \\ \frac{\partial^2 \ell_{t0}(\theta)}{\partial \rho \partial \theta^{*T}} &= \frac{\rho A_{t0}(\theta)e^{-A_{t0}(\theta)} + (1-\rho)\left\{e^{-A_{t0}(\theta)} - e^{-2A_{t0}(\theta)}\right\}}{\phi(1-\rho)\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2} \cdot \frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}},\end{aligned}$$

$$\frac{\partial^2 \ell_{t0}(\theta)}{\partial \phi \partial \theta^{*T}} = \frac{\left\{1 - A_{t0}(\theta)\right\} \rho e^{-A_{t0}(\theta)} + (1-\rho)e^{-2A_{t0}(\theta)}}{\phi^2\left[\rho + (1-\rho)e^{-A_{t0}(\theta)}\right]^2} \cdot \frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}}$$

and

$$\begin{aligned}\frac{\partial^2 \ell_{t1}(\theta)}{\partial \rho^2} &= -\frac{(\phi-1)^2(Y_t-1)Y_t^2}{A_{t1}(\theta)^2} + \frac{Y_t-1}{(1-\rho)^2} - \frac{2\lambda_t(\theta^*)}{\phi(1-\rho)^3}, \\ \frac{\partial^2 \ell_{t1}(\theta)}{\partial \phi^2} &= -\frac{(1-\rho)^2(Y_t-1)Y_t^2}{A_{t1}(\theta)^2} + \frac{Y_t}{\phi^2} - \frac{2\lambda_t(\theta^*)}{\phi^3(1-\rho)} + \frac{2Y_t}{\phi^3}, \\ \frac{\partial^2 \ell_{t1}(\theta)}{\partial \rho \partial \phi} &= -\frac{\lambda_t(\theta^*)Y_t(Y_t-1)}{A_{t1}(\theta)^2} + \frac{\lambda_t(\theta^*)}{\phi^2(1-\rho)^2}, \\ \frac{\partial^2 \ell_{t1}(\theta)}{\partial \rho \partial \theta^{*T}} &= \left[\frac{(\phi-1)Y_t(Y_t-1)}{A_{t1}(\theta)^2} - \frac{1}{\phi(1-\rho)^2}\right] \frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}, \\ \frac{\partial^2 \ell_{t1}(\theta)}{\partial \phi \partial \theta^{*T}} &= \left[-\frac{(1-\rho)Y_t(Y_t-1)}{A_{t1}(\theta)^2} + \frac{1}{\phi^2(1-\rho)}\right] \frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}, \\ \frac{\partial^2 \ell_{t1}(\theta)}{\partial \theta^* \partial \theta^{*T}} &= -\left[\frac{Y_t-1}{A_{t1}(\theta)^2} + \frac{1}{\lambda_t^2(\theta^*)}\right] \frac{\partial \lambda_t(\theta^*)}{\partial \theta^*} \frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}} \\ &\quad + \left\{\frac{1}{\lambda_t(\theta^*)} + \frac{Y_t-1}{A_{t1}(\theta)} - \frac{1}{\phi(1-\rho)}\right\} \frac{\partial^2 \lambda_t(\theta^*)}{\partial \theta^* \partial \theta^{*T}}.\end{aligned}$$

Lemma 3.2. *Under (**A2**), we have the followings :*

$$
\begin{aligned}
&|e^{-\tilde{A}_{t0}(\theta)} - e^{-A_{t0}(\theta)}| \le C|\tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*)|,\\
&|e^{-2\tilde{A}_{t0}(\theta)} - e^{-2A_{t0}(\theta)}| \le 2C|\tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*)|,\\
&|\tilde{A}_{t0}(\theta)e^{-\tilde{A}_{t0}(\theta)} - A_{t0}(\theta)e^{-A_{t0}(\theta)}| \le \left\{C + C^2\lambda_t(\theta^*)\right\}|\tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*)|,\\
&|\tilde{A}_{t0}(\theta)e^{-A_{t0}(\theta)} - A_{t0}(\theta)e^{-\tilde{A}_{t0}(\theta)}| \le \left\{C + C^2\tilde{\lambda}_t(\theta^*)\right\}|\tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*)|,\\
&|\tilde{A}_{t0}(\theta)e^{-2\tilde{A}_{t0}(\theta)} - A_{t0}(\theta)e^{-2A_{t0}(\theta)}| \le \left\{C + 2C^2\lambda_t(\theta^*)\right\}|\tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*)|,\\
&|\tilde{A}^2_{t0}(\theta)e^{-\tilde{A}_{t0}(\theta)} - A^2_{t0}(\theta)e^{-A_{t0}(\theta)}|\\
&\qquad\qquad \le \left[C^2\left\{\lambda_t(\theta^*) + \tilde{\lambda}_t(\theta^*)\right\} + C^3\lambda_t^2(\theta)\right]|\tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*)|,\\
&|\tilde{A}^2_{t0}(\theta)e^{-A_{t0}(\theta)} - A^2_{t0}(\theta)e^{-\tilde{A}_{t0}(\theta)}|\\
&\qquad\qquad \le \left[C^2\left\{\lambda_t(\theta^*) + \tilde{\lambda}_t(\theta^*)\right\} + C^3\tilde{\lambda}_t^2(\theta)\right]|\tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*)|
\end{aligned}
$$

where $\tilde{A}_{t0}(\theta) = \tilde{\lambda}_t(\theta^*)/\phi(1-\rho)$ *and* C *is a sufficiently large constant.*

Proof. Use the mean value theorem. □

Lemma 3.3. *Let*

$$
\begin{aligned}
Z_{t0}(\theta) &= \frac{\exp\{-A_{t0}(\theta)\}}{\phi\left[\rho + (1-\rho)\exp\{-A_{t0}(\theta)\}\right]},\\
Z_{t1}(\theta) &= \frac{1}{\lambda_t(\theta^*)} + \frac{Y_t - 1}{\lambda_t(\theta^*) + (1-\rho)(\phi-1)Y_t} - \frac{1}{\phi(1-\rho)}.
\end{aligned}
$$

*Then, under (**A1**)-(**A3**) and (**A9**), a.s.,*

$$
\begin{aligned}
&\sup_{\theta\in\Theta}|Z_{t0}(\theta)| \le \frac{1}{\delta_L^2}, \quad \sup_{\theta\in\Theta}\left|\tilde{Z}_{t0}(\theta)\right| \le \frac{1}{\delta_L^2},\\
&\sup_{\theta\in\Theta}|Z_{t1}(\theta)| \le \frac{Y_t}{\delta_L^*} + C, \quad \sup_{\theta\in\Theta}\left|\tilde{Z}_{t1}(\theta)\right| \le \frac{Y_t}{\delta_L^*} + C,\\
&\sup_{\theta\in\Theta}\left|\tilde{Z}_{t0}(\theta) - Z_{t0}(\theta)\right| \le CV\eta^t,
\end{aligned}
$$

$$\sup_{\theta\in\Theta}\left|\tilde{Z}_{t1}(\theta)-Z_{t1}(\theta)\right|\leq\left(\frac{Y_t}{\delta_L^{*2}}+C\right)V\eta^t.$$

for some constant C and

$$E\left(\sup_{\theta^*\in\Theta^*}\tilde{\lambda}_t(\theta^*)\right)<\infty.$$

Proof. By (**A2**) and Lemma 3.1, we have

$$\sup_{\theta\in\Theta}|Z_{t0}(\theta)|\leq\sup_{\theta\in\Theta}\frac{1}{\phi\rho}\leq\frac{1}{\delta_L^2},\quad \sup_{\theta\in\Theta}\left|\tilde{Z}_{t0}(\theta)\right|\leq\frac{1}{\delta_L^2}$$

and

$$\begin{aligned}\sup_{\theta\in\Theta}\left|\tilde{Z}_{t1}(\theta)\right| &\leq \frac{1}{\delta_L}+\sup_{\theta\in\Theta}\left|\frac{Y_t-1}{\tilde{A}_{t1}(\theta)}\right|+\sup_{\theta\in\Theta}\left|\frac{1}{\phi(1-\rho)}\right|\leq\frac{1}{\delta_L}+\frac{Y_t+1}{\delta_L^*}+\frac{1}{\delta_L^2},\\ \sup_{\theta\in\Theta}|Z_{t1}(\theta)| &\leq \frac{1}{\delta_L}+\frac{Y_t+1}{\delta_L^*}+\frac{1}{\delta_L^2},\end{aligned}$$

where $\tilde{A}_{t1}(\theta)=\tilde{\lambda}-t(\theta^*)+(\phi-1)(1-\rho)Y_t$. Due to Lemma 3.2 and (**A2**), we get

$$\begin{aligned}\sup_{\theta\in\Theta}\left|\tilde{Z}_{t0}(\theta)-Z_{t0}(\theta)\right| &\leq\frac{1}{\delta_L^3}\sup_{\theta\in\Theta}\rho\left[e^{-\tilde{A}_{t0}(\theta)}-e^{-A_{t0}(\theta)}\right]\leq\frac{C}{\delta_L^3}V\eta^t,\\ \sup_{\theta\in\Theta}\left|\tilde{Z}_{t1}(\theta)-Z_{t1}(\theta)\right| &\leq\frac{1}{\delta_L^2}\sup_{\theta^*\in\Theta^*}\left|\tilde{\lambda}_t(\theta^*)-\lambda_t(\theta^*)\right|+\frac{|Y_t-1|}{\delta_L^{*2}}\sup_{\theta^*\in\Theta^*}\left|\lambda_t(\theta^*)-\tilde{\lambda}_t(\theta^*)\right|\\ &\leq\left(\frac{Y_t}{\delta_L^{*2}}+C\right)V\eta^t.\end{aligned}$$

This completes the proof. □

Lemma 3.4. *Under* (**A1**)-(**A6**) *and* (**A9**)*, we have*

$$E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial^2\ell_t(\theta)}{\partial\theta\partial\theta^T}\right\|\right)<\infty \text{ and } E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial\ell_t(\theta)}{\partial\theta}\frac{\partial\ell_t(\theta)}{\partial\theta^T}\right\|\right)<\infty.$$

Proof. It's sufficient to show that for $i=0,1$,

$$E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial^2\ell_{ti}(\theta)}{\partial\theta\partial\theta^T}\right\|\right)<\infty \text{ and } E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial\ell_{ti}(\theta)}{\partial\theta}\frac{\partial\ell_{ti}(\theta)}{\partial\theta^T}\right\|\right)<\infty. \tag{3.9}$$

First,by using (**A2**), Lemma 3.2 and Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\sup_{\theta\in\Theta}\left|\frac{\partial^2\ell_{t0}(\theta)}{\partial\rho^2}\right| &\le \frac{1}{\delta_L^2}\sup_{\theta\in\Theta}\left\{A_{t0}^2(\theta)/\delta_L+4A_{t0}(\theta)+4\right\},\\
\sup_{\theta\in\Theta}\left|\frac{\partial^2\ell_{t0}(\theta)}{\partial\phi^2}\right| &\le \frac{1}{\delta_L^4}\sup_{\theta\in\Theta}\left\{A_{t0}^2(\theta)+4A_{t0}(\theta)\right\},\\
\sup_{\theta\in\Theta}\left\|\frac{\partial^2\ell_{t0}(\theta)}{\partial\theta^*\partial\theta^{*T}}\right\| &\le \frac{4}{\delta_L^2}\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial^2\lambda_t(\theta^*)}{\partial\theta^*}\frac{\partial\lambda_t(\theta^*)}{\partial\theta^{*T}}\right\|+\frac{1}{\delta_L^2}\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial^2\lambda_t(\theta^*)}{\partial\theta^*\partial\theta^{*T}}\right\|,\\
\sup_{\theta\in\Theta}\left|\frac{\partial^2\ell_{t0}(\theta)}{\partial\rho\partial\phi}\right| &\le \frac{1}{\delta_L^3}\sup_{\theta\in\Theta}\left\{A_{t0}^2(\theta)+2A_{t0}(\theta)\right\},\\
\sup_{\theta\in\Theta}\left\|\frac{\partial^2\ell_{t0}(\theta)}{\partial\rho\partial\theta^{*T}}\right\| &\le \frac{1}{\delta_L^4}\sup_{\theta\in\Theta}\left\{A_{t0}(\theta)+2\right\}\cdot\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\|,\\
\sup_{\theta\in\Theta}\left\|\frac{\partial^2\ell_{t0}(\theta)}{\partial\phi\partial\theta^{*T}}\right\| &\le \frac{1}{\delta_L^5}\sup_{\theta\in\Theta}\left\{A_{t0}(\theta)+2\right\}\cdot\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\|,
\end{aligned}$$

and

$$\begin{aligned}
\sup_{\theta\in\Theta}\left|\frac{\partial^2\ell_{t1}(\theta)}{\partial\rho^2}\right| &\le \frac{1}{\delta_L^2}+\frac{(\phi_U-1)^2(Y_t+1)Y_t}{\delta_L^{*2}}+\frac{Y_t+1}{\delta_L^2}+\frac{2\sup_{\theta^*\in\Theta^*}\lambda_t(\theta^*)}{\delta_L^4},\\
\sup_{\theta\in\Theta}\left|\frac{\partial^2\ell_{t1}(\theta)}{\partial\phi^2}\right| &\le \frac{(Y_t+1)Y_t^2}{\delta_L^{*2}}+\frac{Y_t}{\delta_L^2}+\frac{2\sup_{\theta^*\in\Theta^*}\lambda_t(\theta^*)}{\delta_L^4}+\frac{2Y_t}{\delta_L^3},\\
\sup_{\theta\in\Theta}\left|\frac{\partial^2\ell_{t1}(\theta)}{\partial\rho\partial\phi}\right| &\le \left\{\frac{(Y_t+1)Y_t}{\delta_L^{*2}}+\frac{1}{\delta_L^5}\right\}\sup_{\theta^*\in\Theta^*}\lambda_t(\theta^*),\\
\sup_{\theta\in\Theta}\left\|\frac{\partial^2\ell_{t1}(\theta)}{\partial\rho\partial\theta^{*T}}\right\| &\le \left\{\frac{(\phi_U-1)Y_t(Y_t+1)}{\delta_L^{*2}}+\frac{1}{\delta_L^3}\right\}\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\|,\\
\sup_{\theta\in\Theta}\left\|\frac{\partial^2\ell_{t1}(\theta)}{\partial\phi\partial\theta^{*T}}\right\| &\le \left\{\frac{Y_t(Y_t+1)}{\delta_L^{*2}}+\frac{1}{\delta_L^3}\right\}\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\|,\\
\sup_{\theta\in\Theta}\left\|\frac{\partial^2\ell_{t1}(\theta)}{\partial\theta^*\partial\theta^{*T}}\right\| &\le \left(\frac{Y_t+1}{\delta_L^{*2}}+\frac{1}{\delta_L}\right)\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\frac{\partial\lambda_t(\theta^*)}{\partial\theta^{*T}}\right\|\\
&\quad+\left(\frac{Y_t+1}{\delta_L^*}+C\right)\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial^2\lambda_t(\theta)}{\partial\theta^*\partial\theta^{*T}}\right\|.
\end{aligned}$$

Therefore, the first part in (3.9) can be seen by using (**A3**) and (**A6**). Similarly, we can show the second part in (3.9) from the followings.

$$\sup_{\theta\in\Theta}\left\{\frac{\partial\ell_{t0}(\theta)}{\partial\rho}\right\}^2 \le \frac{3}{\delta_L^2}\sup_{\theta\in\Theta}A_{t0}^2(\theta),$$

$$
\begin{aligned}
&\sup_{\theta\in\Theta}\left\{\frac{\partial \ell_{t0}(\theta)}{\partial \phi}\right\}^2 \le \frac{1}{\delta_L^4}\sup_{\theta\in\Theta} A_{t0}^2(\theta),\\
&\sup_{\theta\in\Theta}\left\|\frac{\partial \ell_{t0}(\theta)}{\partial \theta^*}\cdot\frac{\partial \ell_{t0}(\theta)}{\partial \theta^{*T}}\right\| \le \frac{1}{\delta_L^4}\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}\cdot\frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}}\right\|,\\
&\sup_{\theta\in\Theta}\left|\frac{\partial \ell_{t0}(\theta)}{\partial \rho}\cdot\frac{\partial \ell_{t0}(\theta)}{\partial \phi}\right| \le \frac{1}{\delta_L^3}\sup_{\theta\in\Theta}\left\{2A_{t0}(\theta)+A_{t0}^2(\theta)\right\},\\
&\sup_{\theta\in\Theta}\left\|\frac{\partial \ell_{t0}(\theta)}{\partial \rho}\cdot\frac{\partial \ell_{t0}(\theta)}{\partial \theta^*}\right\| \le \frac{1}{\delta_L^3}\sup_{\theta\in\Theta}\left\{2+A_{t0}(\theta)\right\}\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}\right\|,\\
&\sup_{\theta\in\Theta}\left\|\frac{\partial \ell_{t0}(\theta)}{\partial \phi}\cdot\frac{\partial \ell_{t0}(\theta)}{\partial \theta^*}\right\| \le \frac{1}{\delta_L^4}\sup_{\theta\in\Theta} A_{t0}(\theta)\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}\right\|,
\end{aligned}
$$

and

$$
\begin{aligned}
&\sup_{\theta\in\Theta}\left\{\frac{\partial \ell_{t1}}{\partial \rho}\right\}^2 \le 3\left[\frac{(\phi_U-1)^2(Y_t-1)^2Y_t^2}{\delta_L^{*2}}+\frac{(Y_t-1)^2}{\delta_L^2}+\frac{\sup_{\theta^*\in\Theta^*}\lambda_t(\theta^*)}{\delta_L^4}\right],\\
&\sup_{\theta\in\Theta}\left\{\frac{\partial \ell_{t1}}{\partial \phi}\right\}^2 \le 4\left[\frac{(Y_t-1)^2Y_t^2}{\delta_L^{*2}}+\frac{Y_t^2}{\delta_L^2}+\frac{Y_t^2}{\delta_L^4}+\frac{\sup_{\theta^*\in\Theta^*}\lambda_t(\theta^*)}{\delta_L^6}\right],\\
&\sup_{\theta\in\Theta}\left|\frac{\partial \ell_{t1}}{\partial \rho}\cdot\frac{\partial \ell_{t1}}{\partial \phi}\right| \le \sup_{\theta\in\Theta}\left|\frac{\partial \ell_{t1}}{\partial \rho}\right|\cdot\sup_{\theta\in\Theta}\left|\frac{\partial \ell_{t1}}{\partial \phi}\right|,\\
&\sup_{\theta\in\Theta}\left|\frac{\partial \ell_{t1}(\theta)}{\partial \rho}\right|\cdot\left\|\frac{\partial \ell_{t1}(\theta^*)}{\partial \theta^*}\right\| \le \sup_{\theta\in\Theta}\left|\frac{\partial \ell_{t1}(\theta)}{\partial \rho}\right|\cdot\sup_{\theta^*}\left\|\frac{\partial \ell_{t1}(\theta^*)}{\partial \theta^*}\right\|\\
&\qquad\le\left\{\frac{1}{\delta_L}+\frac{(\phi_U-1)(Y_t+1)Y_t}{\delta_L^*}+\frac{Y_t+1}{\delta_L}+\frac{\sup_{\theta^*\in\Theta^*}\lambda_t(\theta^*)}{\delta_L^3}\right\}\\
&\qquad\qquad\times\sup_{\theta\in\Theta}|Z_{t1}(\theta)|\cdot\left\|\frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}\right\|,\\
&\sup_{\theta\in\Theta}\left|\frac{\partial \ell_{t1}(\theta)}{\partial \phi}\right|\cdot\left\|\frac{\partial \ell_{t1}(\theta^*)}{\partial \theta^*}\right\|\\
&\qquad\le\left\{\frac{(Y_t+1)Y_t}{\delta_L^*}+\frac{Y_t}{\delta_L}+\frac{\sup_{\theta^*\in\Theta^*}\lambda_t(\theta^*)}{\delta_L^3}+\frac{Y_t}{\delta_L^2}\right\}\cdot\sup_{\theta\in\Theta}|Z_{t1}(\theta)|\cdot\left\|\frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}\right\|,\\
&\sup_{\theta\in\Theta}\left\|\frac{\partial \ell_{t1}(\theta)}{\partial \theta^*}\cdot\frac{\partial \ell_{t1}(\theta)}{\partial \theta^{*T}}\right\| \le \sup_{\theta\in\Theta} Z_{t1}(\theta)^2\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}\frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}}\right\|\\
&\qquad\le\left\{\frac{Y_t}{\delta_L^*}+C\right\}^2\cdot\sup_{\theta^*\in\Theta^*}\left\|\frac{\partial \lambda_t(\theta^*)}{\partial \theta^*}\frac{\partial \lambda_t(\theta^*)}{\partial \theta^{*T}}\right\|.
\end{aligned}
$$

The proof is completed. □

Proof of Propositon 3.4 First, we show that

$$\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \tilde{\ell}_{t0}(\theta_0)}{\partial \theta} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_{t0}(\theta_0)}{\partial \theta} \right\| = o_P(1).$$

By Lemma 3.2 and **(A2)**, we get

$$\begin{aligned}
\left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \tilde{\ell}_{t0}(\theta_0)}{\partial \rho} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_{t0}(\theta_0)}{\partial \rho} \right| &\leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left| \frac{\partial \tilde{\ell}_{t0}(\theta_0)}{\partial \rho} - \frac{\partial \ell_{t0}(\theta_0)}{\partial \rho} \right| \\
&\leq \frac{1}{\sqrt{n}} \cdot \frac{1}{\delta_L^2} \sum_{t=1}^{n} \left[\left| e^{-A_{t0}(\theta)} - e^{-\tilde{A}_{t0}(\theta)} \right| + \rho \left| A_{t0}(\theta) e^{-A_{t0}(\theta)} - \tilde{A}_{t0}(\theta) e^{-\tilde{A}_{t0}(\theta)} \right| \right. \\
&\qquad \left. + (1-\rho) e^{-A_{t0}(\theta) - \tilde{A}_{t0}(\theta)} \left| A_{t0}(\theta) - \tilde{A}_{t0}(\theta) \right| \right] \\
&\leq \frac{1}{\sqrt{n}} \cdot \frac{1}{\delta_L^2} \sum_{t=1}^{n} V \eta^t \left\{ C + C + C^2 \lambda_t(\theta_0^*)) + C \right\} = o_P(1).
\end{aligned}$$

Similarly, we can have

$$\left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \tilde{\ell}_{t0}(\theta_0)}{\partial \phi} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_{t0}(\theta_0)}{\partial \phi} \right| \leq \frac{1}{\sqrt{n}} \cdot \frac{1}{\delta_L^2} \sum_{t=1}^{n} V \eta^t \left\{ 1 + C \lambda_t(\theta_0^*) \right\} = o_P(1).$$

Note that owing to Lemma 3.3 and **(A7)**,

$$\begin{aligned}
&\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \tilde{\ell}_{t0}(\theta_0)}{\partial \theta^*} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_{t0}(\theta_0)}{\partial \theta^*} \right\| \\
&\qquad \leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\| \tilde{Z}_{t0}(\theta_0) \frac{\partial \tilde{\lambda}_t(\theta_0^*)}{\partial \theta^*} - Z_{t0}(\theta_0^*) \frac{\partial \lambda_t(\theta_0^*)}{\partial \theta^*} \right\| \\
&\qquad \leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\| \tilde{Z}_{t0}(\theta_0) \left\{ \frac{\partial \tilde{\lambda}_t(\theta_0^*)}{\partial \theta^*} - \frac{\partial \lambda_t(\theta_0^*)}{\partial \theta^*} \right\} \right\| \\
&\qquad\qquad + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\| \left\{ \tilde{Z}_{t0}(\theta_0) - Z_{t0}(\theta_0) \right\} \frac{\partial \lambda_t(\theta_0^*)}{\partial \theta^*} \right\| \\
&\qquad \leq \frac{1}{\sqrt{n}} \cdot \frac{C_2}{\delta_L^3} \sum_{t=1}^{n} V \eta^t \frac{\partial \lambda_t(\theta^*)}{\partial \theta^*} + \frac{1}{\sqrt{n}} \cdot \frac{1}{\delta_L^2} \sum_{t=1}^{n} V \eta^t = o_P(1).
\end{aligned}$$

Next, note that

$$\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial\tilde{\ell}_{t1}(\theta_0)}{\partial\theta}-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial\ell_{t1}(\theta_0)}{\partial\theta}\right\|=o_P(1).$$

Further,

$$\begin{aligned}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial\tilde{\ell}_{t1}(\theta_0)}{\partial\rho}-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial\ell_{t1}(\theta_0)}{\partial\rho}\right| &\leq \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left|\frac{\partial\tilde{\ell}_{t1}(\theta_0)}{\partial\rho}-\frac{\partial\ell_{t1}(\theta_0)}{\partial\rho}\right| \\ &\leq \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left|\frac{(1-\rho_0)(\phi_0-1)Y_t}{\tilde{\lambda}_t(\theta_0^*)+(1-\rho_0)(\phi_0-1)Y_t}-\frac{(1-\rho_0)(\phi_0-1)Y_t}{\lambda_t(\theta_0^*)+(1-\rho_0)(\phi_0-1)Y_t}\right| \\ &\qquad +\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\left|\tilde{\lambda}_t(\theta_0^*)-\lambda_t(\theta_0^*)\right|}{\phi_0(1-\rho_0)^2} \\ &\leq \frac{1}{\sqrt{n}}\frac{|\phi-1|}{\delta_L^2}\sum_{t=1}^{n}Y_t\cdot\left|\tilde{\lambda}_t(\theta_0^*)-\lambda_t(\theta_0^*)\right|+\frac{1}{\phi_0(1-\rho_0)^2}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left|\tilde{\lambda}_t(\theta_0^*)-\lambda_t(\theta_0^*)\right| \\ &\leq \frac{\phi_U+1}{\delta_L^2}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}Y_tV\eta^t+\frac{1}{\delta_L^3}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}V\eta^t=o_P(1).\end{aligned}$$

Similarly,

$$\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial\tilde{\ell}_{t1}(\theta_0)}{\partial\phi}-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial\ell_{t1}(\theta_0)}{\partial\phi}\right|=o_P(1).$$

Since due to Lemma 3.3 and **(A7)**,

$$\begin{aligned}&\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial\tilde{\ell}_{t1}(\theta_0)}{\partial\theta^*}-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial\ell_{t1}(\theta_0)}{\partial\theta^*}\right\| \\ &\qquad\leq \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left\|\tilde{Z}_{t1}(\theta_0)\frac{\partial\tilde{\lambda}_t(\theta_0^*)}{\partial\theta^*}-Z_{t1}(\theta_0^*)\frac{\partial\lambda_t(\theta_0^*)}{\partial\theta^*}\right\| \\ &\qquad\leq \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left\|\tilde{Z}_{t1}(\theta_0)\left\{\frac{\partial\tilde{\lambda}_t(\theta_0^*)}{\partial\theta^*}-\frac{\partial\lambda_t(\theta_0^*)}{\partial\theta^*}\right\}\right\| \\ &\qquad\qquad +\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left\|\left\{\tilde{Z}_{t1}(\theta_0)-Z_{t1}(\theta_0)\right\}\frac{\partial\lambda_t(\theta_0^*)}{\partial\theta^*}\right\|\end{aligned}$$

$$\leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(\frac{Y_t}{\delta_L^*} + C \right) V \eta^t + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(\frac{Y_t}{\delta_L^{*2}} + C \right) V \eta^t \cdot \frac{\partial \lambda_t(\theta_0^*)}{\partial \theta^*} = o_P(1),$$

the lemma is established. □

Lemma 3.5. *Under (A1)-(A7) and (A9), we have, as $n \to 0$,*

$$\sup_{\theta \in \Theta} \left\| \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2 \tilde{\ell}_t(\theta)}{\partial \theta \partial \theta^T} - \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta^T} \right\| \to 0 \quad a.s.$$

Proof. It's sufficient to show that for $i = 0, 1$, as $t \to \infty$,

$$\sup_{\theta \in \Theta} \left\| \frac{\partial^2 \tilde{\ell}_{ti}(\theta)}{\partial \theta \partial \theta^T} - \frac{\partial^2 \ell_{ti}(\theta)}{\partial \theta \partial \theta^T} \right\| \to 0 \quad a.s.. \tag{3.10}$$

By Lemma 3.2, we have,

$$\begin{aligned}
&\left| \frac{\partial^2 \tilde{\ell}_{t0}(\theta)}{\partial \rho^2} - \frac{\partial^2 \ell_{t0}(\theta)}{\partial \rho^2} \right| \\
&\leq \frac{1}{\rho^2} \Big| 2\rho \left\{ e^{-\tilde{A}_t(\theta)} - e^{-A_{t0}(\theta)} \right\} + 2\rho^2 \left\{ \tilde{A}_{t0}(\theta) e^{-\tilde{A}_{t0}(\theta)} - A_{t0}(\theta) e^{-A_{t0}(\theta)} \right\} \\
&\quad + \frac{\rho^3}{1-\rho} \left\{ \tilde{A}_{t0}^2(\theta) e^{-\tilde{A}_{t0}(\theta)} - A_{t0}^2(\theta) e^{-A_{t0}(\theta)} \right\} + (1-2\rho) \left\{ e^{-2\tilde{A}_{t0}(\theta)} - e^{-2A_{t0}(\theta)} \right\} \\
&\quad - 2\rho^2 \left\{ \tilde{A}_{t0}(\theta) e^{-2\tilde{A}_{t0}(\theta)} - A_{t0}(\theta) e^{-2A_{t0}(\theta)} \right\} \\
&\quad + e^{-\tilde{A}_{t0}(\theta) - A_{t0}(\theta)} \Big[4\rho(1-\rho) \left\{ \tilde{A}_{t0}(\theta) - A_{t0}(\theta) \right\} \\
&\qquad + 2\rho^2(1-\rho) \left\{ \tilde{A}_{t0}^2(\theta) - A_{t0}^2(\theta) \right\} \\
&\qquad - 2(1-\rho) \left\{ e^{-\tilde{A}_{t0}(\theta)} - e^{-A_{t0}(\theta)} \right\} \\
&\qquad + 2(1-\rho)^2 \left\{ \tilde{A}_{t0}(\theta) e^{-A_{t0}(\theta)} - A_{t0}(\theta) e^{-\tilde{A}_{t0}(\theta)} \right\} \\
&\qquad + \rho(1-\rho) \left\{ \tilde{A}_{t0}^2(\theta) e^{-A_{t0}(\theta)} - A_{t0}^2(\theta) e^{-\tilde{A}_{t0}(\theta)} \right\} \\
&\qquad - 4\rho(1-\rho) \left\{ \tilde{A}_{t0}(\theta) e^{-\tilde{A}_{t0}(\theta)} - A_{t0}(\theta) e^{-A_{t0}(\theta)} \right\} \Big] \Big| \\
&\leq C \left\{ 1 + \lambda_t(\theta^*) + \tilde{\lambda}_t(\theta^*) + \lambda_t^2(\theta^*) + \tilde{\lambda}_t^2(\theta^*) \right\} \left| \tilde{\lambda}_t(\theta^*) - \lambda_t(\theta^*) \right|.
\end{aligned}$$

Similarly, we obtain

$$\left|\frac{\partial^2\tilde{\ell}_{t0}(\theta)}{\partial\phi^2}-\frac{\partial^2\ell_{t0}(\theta)}{\partial\phi^2}\right|$$
$$\leq C\left\{1+\lambda_t(\theta^*)+\tilde{\lambda}_t(\theta^*)+\lambda_t^2(\theta^*)+\tilde{\lambda}_t^2(\theta^*)\right\}\left|\tilde{\lambda}_t(\theta^*)-\lambda_t(\theta^*)\right|,$$
$$\left|\frac{\partial^2\tilde{\ell}_{t0}(\theta)}{\partial\rho\partial\phi}-\frac{\partial^2\ell_{t0}(\theta)}{\partial\rho\partial\phi}\right|$$
$$\leq C\left\{1+\lambda_t(\theta^*)+\tilde{\lambda}_t(\theta^*)+\lambda_t^2(\theta^*)+\tilde{\lambda}_t^2(\theta^*)\right\}\left|\tilde{\lambda}_t(\theta^*)-\lambda_t(\theta^*)\right|,$$
$$\left\|\frac{\partial^2\tilde{\ell}_{t0}(\theta)}{\partial\rho\partial\theta^{*T}}-\frac{\partial^2\ell_{t0}(\theta)}{\partial\rho\partial\theta^{*T}}\right\|\leq\frac{A_{t0}(\theta)+2}{\phi\rho^2(1-\rho)}\left\|\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^{*T}}-\frac{\partial\lambda_t(\theta^*)}{\partial\theta^{*T}}\right\|$$
$$+C\left\{1+\lambda_t(\theta^*)+\tilde{\lambda}_t(\theta^*)+\lambda_t^2(\theta^*)+\tilde{\lambda}_t^2(\theta^*)\right\}\left|\tilde{\lambda}_t(\theta^*)-\lambda_t(\theta^*)\right|\left\|\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^{*T}}\right\|,$$
$$\left\|\frac{\partial^2\tilde{\ell}_{t0}(\theta)}{\partial\phi\partial\theta^{*T}}-\frac{\partial^2\ell_{t0}(\theta)}{\partial\phi\partial\theta^{*T}}\right\|\leq\frac{A_{t0}(\theta)+2}{\phi^2\rho^2}\left\|\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^{*T}}-\frac{\partial\lambda_t(\theta^*)}{\partial\theta^{*T}}\right\|$$
$$+C\left\{1+\lambda_t(\theta^*)+\tilde{\lambda}_t(\theta^*)+\lambda_t^2(\theta^*)+\tilde{\lambda}_t^2(\theta^*)\right\}\left|\tilde{\lambda}_t(\theta^*)-\lambda_t(\theta^*)\right|\left\|\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^{*T}}\right\|.$$

And note that

$$\left\|\frac{\partial^2\tilde{\ell}_{t0}(\theta)}{\partial\theta^*\partial\theta^{*T}}-\frac{\partial^2\ell_{t0}(\theta)}{\partial\theta^*\partial\theta^{*T}}\right\|=\left\|\tilde{Y}_{t0}(\theta)\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^*}\left\{\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^{*T}}-\frac{\partial\lambda_t(\theta^*)}{\partial\theta^{*T}}\right\}\right.$$
$$+\tilde{Y}_{t0}(\theta)\left\{\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^*}-\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\}\frac{\partial\lambda_t(\theta^*)}{\partial\theta^{*T}}+\left\{\tilde{Y}_{t0}(\theta)-Y_{t0}(\theta)\right\}\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\frac{\partial\lambda_t(\theta^*)}{\partial\theta^{*T}}$$
$$\left.-\tilde{Z}_{t0}(\theta)\left\{\frac{\partial^2\tilde{\lambda}_t(\theta^*)}{\partial\theta^*\partial\theta^{*T}}-\frac{\partial^2\lambda_t(\theta^*)}{\partial\theta^*\partial\theta^{*T}}\right\}-\left\{\tilde{Z}_{t0}(\theta)-Z_{t0}(\theta)\right\}\frac{\partial^2\tilde{\lambda}_t(\theta^*)}{\partial\theta^*\partial\theta^{*T}}\right\|,$$

where

$$Y_{t0}=\frac{\rho e^{-A_{t0}(\theta)}}{\phi^2(1-\rho)\left[\rho+(1-\rho)e^{-A_{t0}(\theta)}\right]^2}.$$

Then for Y_{t0}, we have

$$\sup_{\theta\in\Theta}|Y_{t0}(\theta)|\leq\frac{1}{\delta_L^5},\quad\sup_{\theta\in\Theta}\left|\tilde{Y}_{t0}(\theta)\right|\leq\frac{1}{\delta_L^5}$$

and

$$\begin{aligned}
&\sup_{\theta\in\Theta}\left|\tilde{Y}_{t0}(\theta)-Y_{t0}(\theta)\right| \\
&\qquad\le \sup_{\theta\in\Theta}\left[\frac{1}{\phi^2\rho(1-\rho)}\left\{\rho^2+(1-\rho)e^{-\tilde{A}_{t0}(\theta)-A_{t0}(\theta)}\right\}\left|e^{-\tilde{A}_{t0}(\theta)}-e^{-A_{t0}(\theta)}\right|\right] \\
&\qquad\le C\sup_{\theta^*\in\Theta^*}|\tilde{\lambda}_t(\theta^*)-\lambda_t(\theta^*)|.
\end{aligned}$$

Then by **(A2)**, **(A3)**, **(A6)**, **(A7)** and Lemma 3.2, we can show that

$$\sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\ell}_{t0}(\theta)}{\partial\theta\partial\theta^T}-\frac{\partial^2\ell_{t0}(\theta)}{\partial\theta\partial\theta^T}\right\|\to 0 \quad \text{a.s.} \quad \text{as } t\to\infty.$$

By similar way, it can be seen that (3.10) holds when $i=1$ with followings:

$$\begin{aligned}
\left|\frac{\partial^2\tilde{\ell}_{t1}(\theta)}{\partial\rho^2}-\frac{\partial^2\ell_{t1}(\theta)}{\partial\rho^2}\right| &\le \left[\frac{(\phi-1)^2Y_t^3\{A_{t1}(\theta)+\tilde{A}_{t1}(\theta)\}}{\delta_L^{*4}}+\frac{2}{\delta_L^4}\right]\cdot V\eta^t \\
\left|\frac{\partial^2\tilde{\ell}_{t1}(\theta)}{\partial\phi^2}-\frac{\partial^2\ell_{t1}(\theta)}{\partial\phi^2}\right| &\le \left[\frac{Y_t^3\{A_{t1}(\theta)+\tilde{A}_{t1}(\theta)\}}{\delta_L^{*4}}+\frac{2}{\delta_L^4}\right]\cdot V\eta^t, \\
\left|\frac{\partial^2\tilde{\ell}_{t1}(\theta)}{\partial\rho\partial\phi}-\frac{\partial^2\ell_{t1}(\theta)}{\partial\rho\partial\phi}\right| &\le \left[\frac{A_{t1}(\theta)^2+\lambda_t(\theta^*)\{A_{t1}(\theta)+\tilde{A}_{t1}(\theta)\}}{\delta_L^{*4}}+\frac{1}{\delta_L^4}\right]\cdot V\eta^t, \\
\left\|\frac{\partial^2\tilde{\ell}_{t1}(\theta)}{\partial\rho\partial\theta^{*T}}-\frac{\partial^2\ell_{t1}(\theta)}{\partial\rho\partial\theta^{*T}}\right\| &\le \left|\frac{(\phi-1)Y_t(Y_t-1)}{\tilde{A}_{t1}(\theta)^2}+\frac{1}{\phi(1-\rho)^2}\right|\left\|\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^*}-\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\| \\
&+\left\|\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\|\left|\frac{(\phi-1)Y_t(Y_t-1)}{\tilde{A}_{t1}(\theta)^2}-\frac{(\phi-1)Y_t(Y_t-1)}{A_{t1}(\theta)^2}\right|, \\
\left\|\frac{\partial^2\tilde{\ell}_{t1}(\theta)}{\partial\phi\partial\theta^{*T}}-\frac{\partial^2\ell_{t1}(\theta)}{\partial\phi\partial\theta^{*T}}\right\| &\le \left|\frac{(1-\rho)Y_t(Y_t-1)}{\tilde{A}_{t1}(\theta)^2}+\frac{1}{\phi^2(1-\rho)}\right|\left\|\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^*}-\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\| \\
&+\left\|\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\right\|\left|\frac{(1-\rho)Y_t(Y_t-1)}{\tilde{A}_{t1}(\theta)^2}-\frac{(1-\rho)Y_t(Y_t-1)}{A_{t1}(\theta)^2}\right|.
\end{aligned}$$

And note that

$$\left\|\frac{\partial^2\tilde{\ell}_{t1}(\theta)}{\partial\theta^*\partial\theta^{*T}}-\frac{\partial^2\ell_{t1}(\theta)}{\partial\theta^*\partial\theta^{*T}}\right\| \le \left\|\left(\frac{Y_t-1}{\tilde{A}_{t1}(\theta)^2}+\frac{1}{\tilde{\lambda}_t^2(\theta^*)}\right)\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\left(\frac{\partial\lambda_t(\theta^*)}{\partial\theta^{*T}}-\frac{\partial\tilde{\lambda}_t}{\partial\theta^{*T}}\right)\right\|$$

$$
\begin{aligned}
&+\left\|\left(\frac{Y_t-1}{\tilde{A}_{t1}(\theta)^2}+\frac{1}{\tilde{\lambda}_t^2(\theta^*)}\right)\left(\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}-\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^*}\right)\frac{\partial\tilde{\lambda}_t(\theta^*)}{\partial\theta^{*T}}\right\| \\
&+\left\|\left(\frac{Y_t-1}{A_{t1}(\theta)^2}+\frac{1}{\lambda_t^2(\theta^*)}-\frac{Y_t-1}{\tilde{A}_{t1}(\theta)^2}-\frac{1}{\tilde{\lambda}_t^2(\theta^*)}\right)\frac{\partial\lambda_t(\theta^*)}{\partial\theta^*}\frac{\partial\lambda_t(\theta^*)}{\partial\theta^{*T}}\right\| \\
&+\left\|\tilde{Z}_{t1}\left(\frac{\partial^2\tilde{\lambda}_t}{\partial\theta^*\partial\theta^{*T}}-\frac{\partial^2\lambda_t(\theta^*)}{\partial\theta^*\partial\theta^{*T}}\right)+\left(\tilde{Z}_{t1}-Z_{t1}\right)\frac{\partial^2\lambda_t(\theta^*)}{\partial\theta^*\partial\theta^{*T}}\right\|
\end{aligned}
$$

and

$$
\begin{aligned}
&\sup_{\theta\in\Theta}\left|\frac{Y_t-1}{A_{t1}(\theta)^2}+\frac{1}{\lambda_t^2(\theta^*)}-\frac{Y_t-1}{\tilde{A}_{t1}(\theta)^2}-\frac{1}{\tilde{\lambda}_t^2(\theta^*)}\right| \\
&\qquad\le \sup_{\theta\in\Theta}\left|\frac{Y_t-1}{\tilde{A}_{t1}(\theta)^2}-\frac{Y_t-1}{A_{t1}(\theta)^2}\right|+\sup_{\theta\in\Theta}\left|\frac{1}{\tilde{\lambda}_t^2(\theta^*)}-\frac{1}{\lambda_t^2(\theta^*)}\right| \\
&\qquad\le \frac{Y_t+1}{\delta_L^4}\sup_{\theta\in\Theta}\left|\tilde{\lambda}_t(\theta^*)-\lambda_t(\theta^*)\right|\cdot\left|A_{t1}(\theta)+\tilde{A}_{t1}(\theta)\right| \\
&\qquad\qquad+\sup_{\theta\in\Theta}\frac{\tilde{\lambda}_t(\theta^*)+\lambda_t(\theta^*)}{\delta_L^4}\cdot\left|\tilde{\lambda}_t(\theta^*)-\lambda_t(\theta^*)\right|.
\end{aligned}
$$

The proof is completed. □

Proof of Propositon 3.5 According to Lemma 3.5, it suffices to show that

$$
\left\|\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2\ell_t(\theta_n')}{\partial\theta\partial\theta^T}-E\left(\frac{\partial^2\ell_t(\theta_0)}{\partial\theta\partial\theta^T}\right)\right\|\to 0 \quad a.s..
$$

Note that

$$
\begin{aligned}
&\left\|\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2\ell_t(\theta_n')}{\partial\theta\partial\theta^T}-E\left(\frac{\partial^2\ell_t(\theta_0)}{\partial\theta\partial\theta^T}\right)\right\| \qquad (3.11)\\
&\le\left\|\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2\ell_t(\theta_n')}{\partial\theta\partial\theta^T}-\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2\ell_t(\theta_0)}{\partial\theta\partial\theta^T}\right\|+\left\|\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2\ell_t(\theta_0)}{\partial\theta\partial\theta^T}-E\left(\frac{\partial^2\ell_t(\theta_0)}{\partial\theta\partial\theta^T}\right)\right\|.
\end{aligned}
$$

Since $\{\partial^2\ell_t(\theta_0)/\partial\theta\partial\theta^T\}$ is a stationary and ergodic process with

$$
E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial^2\ell_t(\theta)}{\partial\theta\partial\theta^T}\right\|\right)<\infty,
$$

the second term on the right hand side of (3.11) converges to 0 a.s.. To show that the first term of the RHS of (3.11) converges to 0 a.s., we put $h_t(\theta) = \partial^2 \ell_t(\theta)/\partial\theta\partial\theta^T$ and $\eta_n = \|\hat{\theta}_n - \theta_0\|$. Then, we have

$$\left\| \frac{1}{n}\sum_{t=1}^{n} \frac{\partial^2 \ell_t(\theta_n')}{\partial\theta\partial\theta^T} - \frac{1}{n}\sum_{t=1}^{n} \frac{\partial^2 \ell_t(\theta_0)}{\partial\theta\partial\theta^T} \right\| \leq \frac{1}{n}\sum_{t=1}^{n} \sup_{\|\theta'-\theta_0\|\leq\eta_n} \|h_t(\theta') - h_t(\theta_0)\|.$$

Since $\eta_n \to 0$ a.s., there exists an event E with $P(E) = 1$ such that $\eta_n(\omega) \to 0$ for all $\omega \in E$. Let $A_m = (\eta_n \leq 1/m$ for sufficiently large $n)$ and $A = \cap_{m=1}^{\infty} A_m$. Assume that $\|\theta' - \theta_0\| \leq \eta_n$. For any $\omega \in A$ and m, there exists $N(\omega, m)$ such that $\|\theta' - \theta_0\| < 1/m$ for $n \geq N(\omega, m)$. Thus,

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \sup_{\|\theta'-\theta_0\|\leq\eta_n(\omega)} \|h_t(\theta',\omega) - h_t(\theta_0,\omega)\| \\
&= \limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{N(\omega,m)} \sup_{\|\theta'-\theta_0\|\leq\eta_n(\omega)} \|h_t(\theta',\omega) - h_t(\theta_0,\omega)\| \\
&\quad + \limsup_{n\to\infty} \frac{1}{n}\sum_{t=N(\omega,m)+1}^{n} \sup_{\|\theta'-\theta_0\|\leq\eta_n(\omega)} \|h_t(\theta',\omega) - h_t(\theta_0,\omega)\| \\
&\leq \limsup_{n\to\infty} \frac{1}{n}\sum_{t=N(\omega,m)+1}^{n} \sup_{\|\theta'-\theta_0\|<1/m} \|h_t(\theta',\omega) - h_t(\theta_0,\omega)\| \\
&= \limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \sup_{\|\theta'-\theta_0\|<1/m} \|h_t(\theta',\omega) - h_t(\theta_0,\omega)\|.
\end{aligned}$$

Because $P(A) = 1$ and $\{h_t(\theta)\}$ is stationary and ergodic, we have that for all $m \geq 1$, almost surely,

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \sup_{\|\theta'-\theta_0\|\leq\eta_n} \|h_t(\theta') - h_t(\theta_0)\| \leq E \sup_{\|\theta'-\theta_0\|<1/m} \|h_t(\theta') - h_t(\theta_0)\|,$$

which converges to 0 as $m \to \infty$ by the dominated converge theorem. This indicates that the first term of the RHS of (3.11) converges to 0 a.s. and the lemma is established. □

Proof of Propositon 3.6 Let $\{Y_t\}$ be a time series with the conditional distribution following the GP distribution, i.e. $\rho = 0$, in (3.2). Then we have, almost surely,

$$
\begin{aligned}
E\left[\frac{\partial \ell_t(\theta_0)}{\partial \theta} \,|\, \mathcal{F}_{t-1}\right] &= E\left[\frac{\partial \ell_{t0}(\theta_0)}{\partial \theta} I(Y_t = 0) + \frac{\partial \ell_{t1}(\theta_0)}{\partial \theta} I(Y_t \geq 1) \,|\, \mathcal{F}_{t-1}\right] \\
&= \frac{\partial \ell_{t0}(\theta_0)}{\partial \theta} P(Y_t = 0 \,|\, \mathcal{F}_{t-1}) \\
&\quad + (1-\rho_0)\left[E\left\{\frac{\partial \ell_{t1}(\theta_0; Y_t)}{\partial \theta} \,|\, \mathcal{F}_{t-1}\right\} - \frac{\partial \ell_{t1}(\theta_0; 0)}{\partial \theta} P(Y_t = 0 \,|\, \mathcal{F}_{t-1})\right].
\end{aligned}
$$

Since according to Zhu (2012a),

$$
\begin{aligned}
E\left[\frac{(Y_t - 1)Y_t}{\lambda_t(\theta_0^*) + (1-\rho_0)(\phi_0 - 1)Y_t} \,|\, \mathcal{F}_{t-1}\right] &= \frac{\lambda_t(\theta_0^*)}{\phi_0(1-\rho_0)^2}\ , \\
E\left[\frac{Y_t - 1}{\lambda_t(\theta_0^*) + (1-\rho_0)(\phi_0 - 1)Y_t} \,|\, \mathcal{F}_{t-1}\right] &= \frac{1}{\phi_0(1-\rho_0)} - \frac{1}{\lambda_t(\theta_0^*)}\ ,
\end{aligned}
$$

the lemma is established. □

Table 3.1: Empirical sizes of T_n^{est} and T_n^{res} at the nominal level 0.05.

	T_n^{est}			T_n^{res}		
$(\rho, \phi, \omega, \alpha, \beta)$	$n = 300$	$n = 500$	$n = 1000$	$n = 300$	$n = 500$	$n = 1000$
(0,1,1,0.1,0.2)*	0.072	0.052	0.042	0.032	0.037	0.043
(0,1,1,0.1,0.5)	0.091	0.063	0.040	0.044	0.048	0.047
(0,1,1,0.1,0.8)	0.377	0.288	0.170	0.027	0.037	0.039
(0.1,1,1,0.1,0.2)*	0.097	0.085	0.066	0.027	0.030	0.038
(0.1,1,1,0.1,0.5)	0.104	0.070	0.056	0.037	0.030	0.035
(0.3,1,1,0.1,0.2)	0.136	0.088	0.072	0.021	0.034	0.044
(0.3,1,1,0.1,0.5)	0.118	0.072	0.058	0.040	0.038	0.045
(0,1.2,1,0.1,0.2)*	0.111	0.095	0.072	0.030	0.040	0.041
(0,1.2,1,0.1,0.5)	0.106	0.075	0.073	0.037	0.026	0.043
(0,1.5,1,0.1,0.2)	0.124	0.109	0.088	0.029	0.041	0.037
(0,1.5,1,0.1,0.5)	0.125	0.092	0.071	0.035	0.036	0.045
(0.1,1.2,1,0.1,0.2)	0.126	0.115	0.100	0.033	0.036	0.045
(0.1,1.2,1,0.1,0.5)	0.167	0.138	0.095	0.033	0.044	0.037
(0.1,1.5,1,0.1,0.2)*	0.129	0.121	0.091	0.031	0.031	0.037
(0.1,1.5,1,0.1,0.5)	0.172	0.160	0.103	0.026	0.033	0.051

Table 3.2: Empirical powers of T_n^{est} and T_n^{res} when $\omega = 1$ changes to $\omega' = 0.3$ and $(\rho, \phi, \alpha, \beta)$ does not change.

	T_n^{est}			T_n^{res}		
$(\rho, \phi, \alpha, \beta)$	$n = 300$	$n = 500$	$n = 1000$	$n = 300$	$n = 500$	$n = 1000$
(0,1,0.1,0.2)*	1	1	1	0.989	1	1
(0,1,0.1,0.5)	0.998	1	1	0.849	0.998	1
(0,1,0.1,0.8)	0.980	0.998	1	0.116	0.245	0.717
(0.1,1,0.1,0.2)*	1	1	1	0.984	1	1
(0.1,1,0.1,0.5)	1	1	1	0.747	0.985	1
(0,1.2,0.1,0.2)*	1	1	1	0.984	1	1
(0,1.2,0.1,0.5)	0.998	1	1	0.747	0.985	1
(0,1.2,0.1,0.8)	0.986	0.996	1	0.066	0.143	0.519
(0.1,1.2,0.1,0.2)*	1	1	1	0.927	0.998	1
(0.1,1.2,0.1,0.5)	1	1	1	0.664	0.966	1

Table 3.3: Empirical powers of T_n^{est} and T_n^{res} when $(\alpha, \beta) = (0.1, 0.2)$ changes to (α', β') and (ρ, ϕ, ω) does not change

	T_n^{est}			T_n^{res}		
$(\rho, \phi, \omega, \alpha, \beta)$	$n = 300$	$n = 500$	$n = 1000$	$n = 300$	$n = 500$	$n = 1000$
(0,1,1,0.3,0.2)*	0.636	0.844	0.98	0.479	0.753	0.987
(0,1,1,0.5,0.2)	1	1	1	0.607	0.996	1
(0,1,1,0.1,0.4)	0.596	0.840	0.978	0.470	0.771	0.980
(0,1,1,0.1,0.6)	1	1	1	0.984	1	1
(0.1,1,1,0.3,0.2)*	0.662	0.854	0.978	0.466	0.763	0.983
(0.1,1,1,0.5,0.2)	1	1	1	0.529	0.933	1
(0.1,1,1,0.1,0.4)	0.682	0.854	1	0.402	0.626	0.967
(0.1,1,1,0.1,0.6)	0.998	1	1	0.943	0.998	1
(0,1.2,1,0.3,0.2)*	0.574	0.768	0.954	0.341	0.688	0.960
(0,1.2,1,0.5,0.2)	1	1	1	0.545	0.972	1
(0,1.2,1,0.1,0.4)	0.570	0.766	0.988	0.306	0.583	0.928
(0,1.2,1,0.1,0.6)	0.998	1	1	0.931	0.997	1
(0.1,1.2,1,0.3,0.2)*	0.596	0.744	0.926	0.313	0.604	0.944
(0.1,1.2,1,0.5,0.2)	0.998	1	1	0.543	0.912	1
(0.1,1.2,1,0.1,0.4)	0.622	0.780	0.958	0.306	0.563	0.861
(0.1,1.2,1,0.1,0.6)	0.998	1	1	0.931	0.991	1

Table 3.4: Empirical powers of T_n^{est} and T_n^{res} when $(\rho, \phi) = (0, 1)$ changes to (ρ', ϕ') and $(\omega, \alpha, \beta) = (1, 0.1, 0.2)$ does not change.

	T_n^{est}			T_n^{res}		
$(\rho, \phi, \omega, \alpha, \beta)$	$n = 300$	$n = 500$	$n = 1000$	$n = 300$	$n = 500$	$n = 1000$
(0.3,1,1,0.1,0.2)*	0.794	0.986	1	0.032	0.046	0.043
(0,1.5,1,0.1,0.2)*	0.852	0.982	1	0.031	0.039	0.041
(0.1,1.2,1,0.1,0.2)*	0.708	0.870	0.992	0.027	0.031	0.044

Table 3.5: Empirical sizes of T_n^{est} and C_n at the nominal level 0.05.

	T_n^{est}		C_n	
$(\rho, \phi, \omega, \alpha, \beta)$	$n = 500$	$n = 1000$	$n = 500$	$n = 1000$
(0,1,1,0.1,0.2)	0.052	0.042	0.094	0.084
(0.1,1,1,0.1,0.2)	0.085	0.066	0.118	0.102
(0,1.2,1,0.1,0.2)	0.095	0.072	0.134	0.116
(1,1.2,1,0.1,0.2)	0.115	0.100	0.134	0.150

Table 3.6: Empirical powers of T_n^{est} and C_n when θ changes to θ'.

		T_n^{est}		C_n	
$\theta = (\rho, \phi, \omega, \alpha, \beta)$	$\theta' = (\rho', \phi', \omega', \alpha', \beta')$	$n = 500$	$n = 1000$	$n = 500$	$n = 1000$
(0,1,1,0.1,0.2)	(0,1,0.3,0.1,0.2)	1	1	1	1
	(0,1,1,0.3,0.2)	0.844	0.98	0.750	0.968
	(0.3,1,1,0.1,0.2)	0.986	1	0.952	1
	(0,1.5,1,0.1,0.2)	0.982	1	0.988	1
(0.1,1,1,0.1,0.2)	(0.1,1,0.3,0.1,0.2)	1	1	1	1
	(0.1,1,1,0.3,0.2)	0.854	0.978	0.768	0.936
(0,1.2,1,0.1,0.2)	(0,1.2,0.3,0.1,0.2)	1	1	1	1
	(0,1.2,1,0.3,0.2)	0.768	0.954	0.684	0.912
(0.1,1.2,1,0.1,0.2)	(0.1,1.2,0.3,0.1,0.2)	1	1	1	1
	(0.1,1.2,1,0.3,0.2)	0.744	0.926	0.684	0.904

Table 3.7: Estimated parameters for the robbery with a firearm data in Inner Sydney based on a ZIGP-INGARH(1,1) model. Standard errors are shown in parentheses.

Model	mean	variance	$\hat{\rho}$	$\hat{\phi}$	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
Full data	1.013	1.493	0.182	1.000	0.256	0.483	0.267
			(0.068)	(0.071)	(0.150)	(0.199)	(0.081)
First period	1.062	1.288	0.001	1.065	0.992	0.001	0.055
(Jan.1995-Jan.2003)			(0.193)	(0.126)	(0.473)	(0.509)	(0.165)
Second period	0.977	1.653	0.239	1.000	0.100	0.642	0.242
(Feb.2003-Dec.2012)			(0.101)	(0.104)	(0.072)	(0.138)	(0.085)

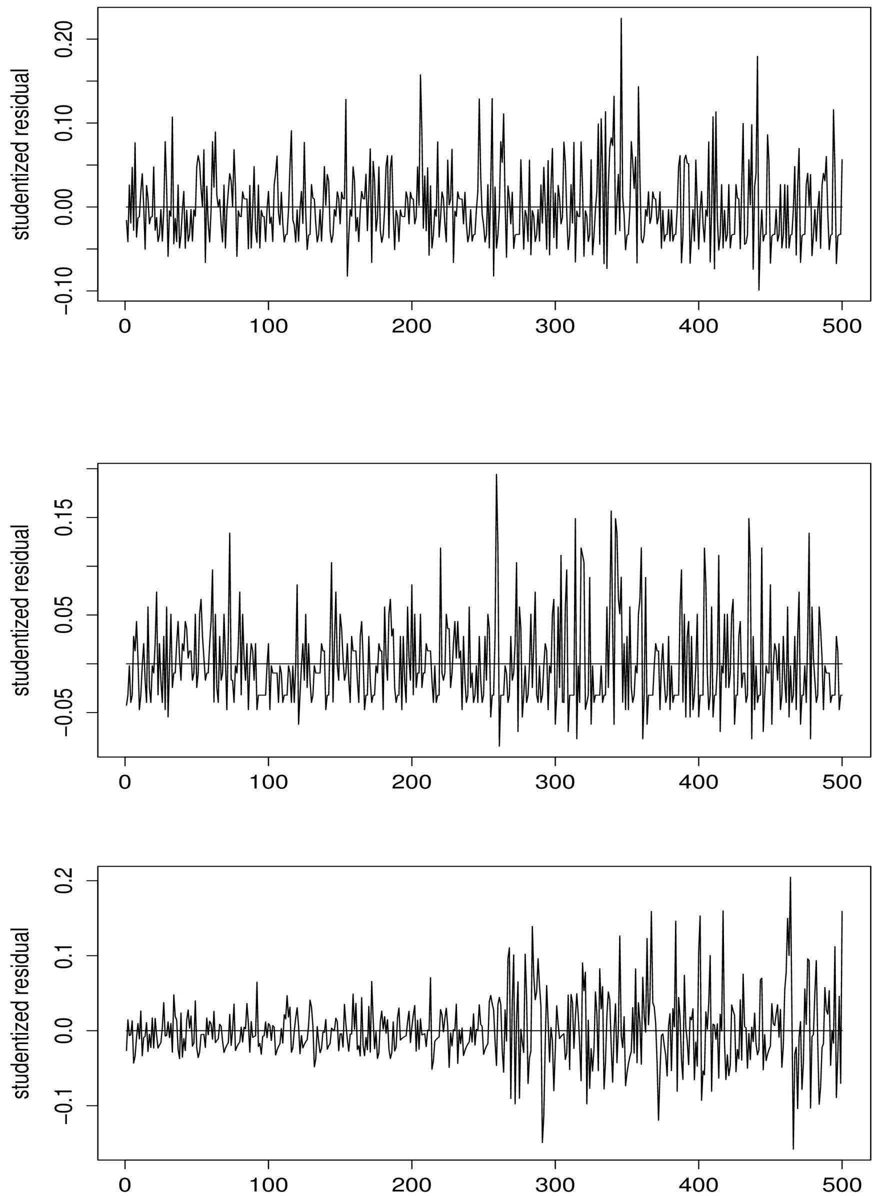

Figure 3.1: Plots of the estimated studentized residuals of simulated data when $\theta = (0, 1, 1, 0.1, 0.2)$ changes to $\theta' = (0.3, 1.2, 1, 0.1, 0.2), (0, 1, 1, 0.5, 0.2)$ and $(0.3, 1.2, 1, 0.5, 0.2)$.

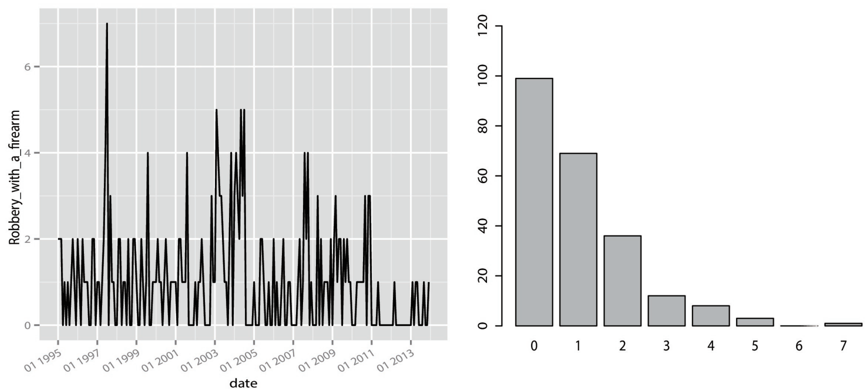

Figure 3.2: Plot of counts series and the histogram of the robbery with a firearm in Inner Sydney

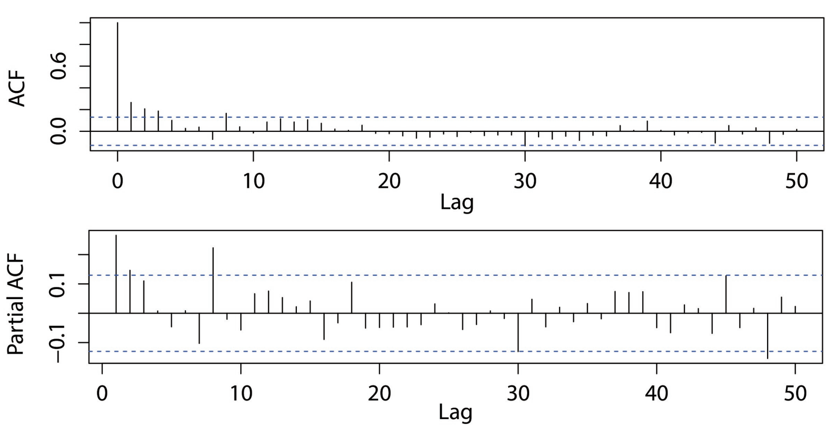

Figure 3.3: Plot of the sample autocorrealation and the sample partial autocorrealation from the robbery with a firearm data in Inner Sydney

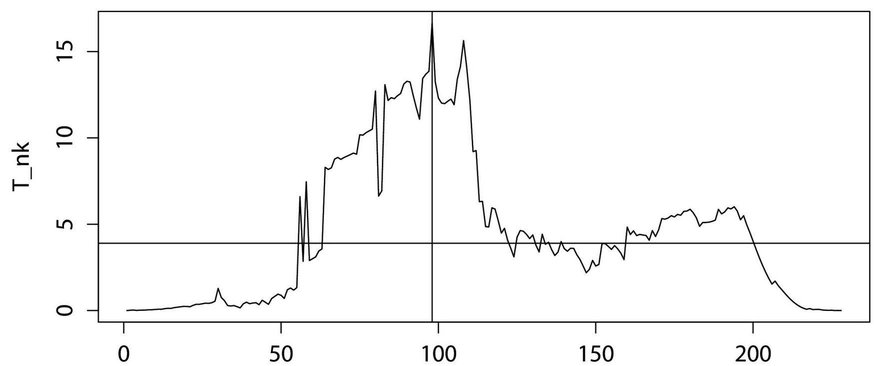

Figure 3.4: Plot of T_n^{est} from the robbery with a firearm data in Inner Sydney with ZIGP-GARCH(1,1)

Table 3.8: Estimated parameters for the assault police data in Sydney based on a GP-INGARH(1,1) model. Standard errors are shown in parentheses.

Model	mean	variance	$\hat{\phi}$	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
Full data	22.877	46.751	1.341	9.572	0.238	0.334
			(0.060)	(0.828)	(0.024)	(0.028)
First period	20.409	44.456	1.349	11.763	0.001	0.420
(Jan.1995-Sep.2003)			(0.088)	(1.220)	(0.036)	(0.042)
Second period	24.984	39.377	1.248	0.100	0.994	0.001
(Oct.2003-Dec.2012)			(0.082)	(1.127)	(0.033)	(0.038)

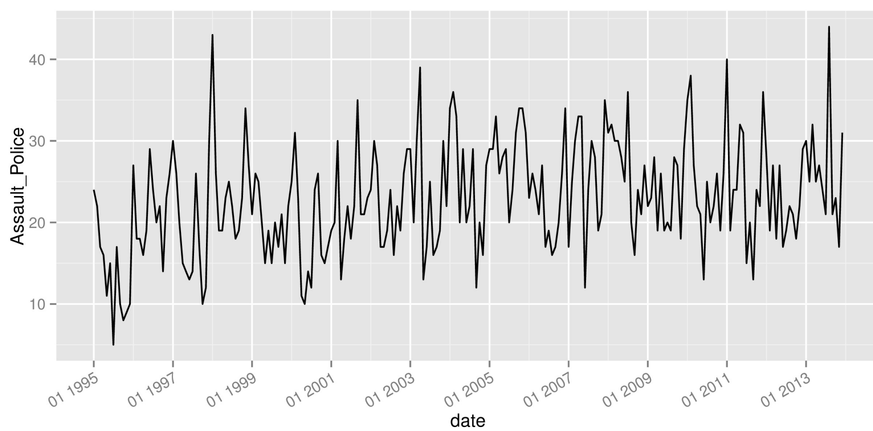

Figure 3.5: Plot of counts series from the assault police data in Inner Sydney

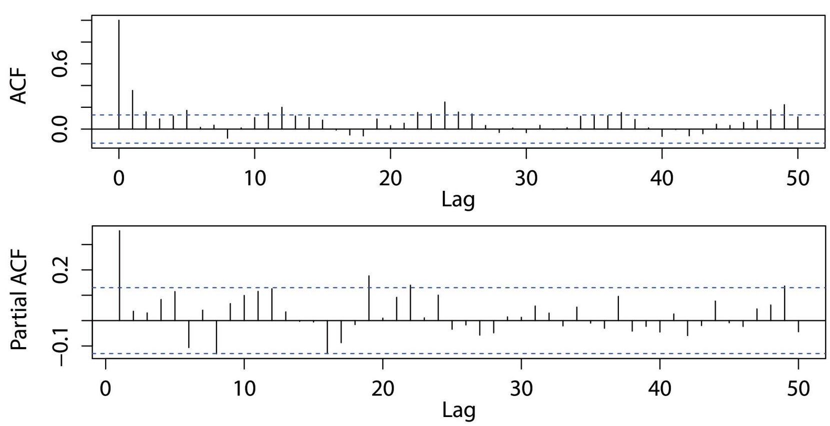

Figure 3.6: Plot of the sample autocorrealation and the sample partial autocorrealation form the assault police data in Inner Sydney

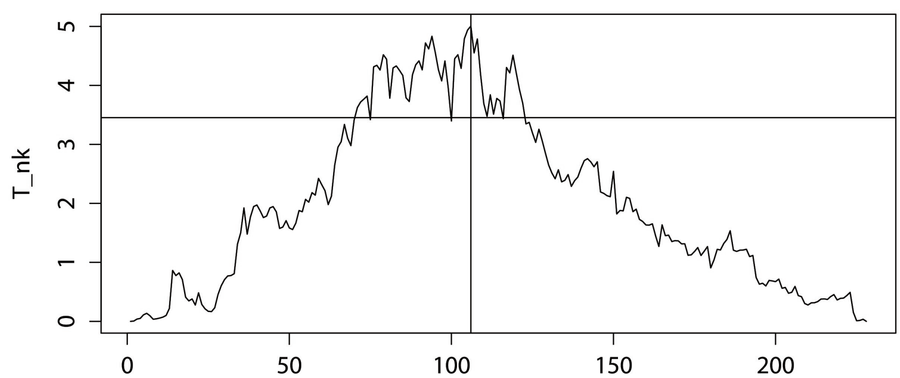

Figure 3.7: Plot of T_n^{est} from the assault police data in Inner Sydney with GP-GARCH(1,1)

Chapter 4

Asymptotic Normality and Parameter Change Test for Bivariate Poisson INGARCH Models

4.1 Introduction

In this Chapter, we consider the problem of testing for a parameter change in bivariate Poisson INGARCH model of Liu (2012).

This paper is organized as follow. Section 4.2 introduces the bivariate Poisson INGARCH model and shows the asymptotic normality of CMLE. Section 4.3 introduces the CUSUM test based on the estimates and residuals and derives their limiting null distributions. Sections 4.4 and 4.5 conduct a simulation study and real data analysis for illustration. Section 4.6 provides concluding remarks. Lastly, the proofs of the theorems in Sections 4.2 and 4.3 are provided in Section 4.7 and in Supplementary material.

4.2 Bivariate Poisson INGARCH model

Let $\mathbf{Y}_t = (Y_{t,1}, Y_{t,2})^T$ be a two dimensional vector of counts at time t, where $\{Y_{t,1}, t \geq 1\}$ and $\{Y_{t,2}, t \geq 1\}$ are the two time series of counts with the conditional distribution following a Poisson distribution with conditional mean $\lambda_{t,1}$ and $\lambda_{t,2}$, respectively. Suppose that $\{\mathbf{Y}_t\}$ follows a bivariate Poisson INGARCH(1,1) model:

$$\mathbf{Y}_t|\mathcal{F}_{t-1} \sim BP(\lambda_{t,1}, \lambda_{t,2}, \varphi), \;\; \boldsymbol{\lambda}_t = (\lambda_{t,1}, \lambda_{t,2})^T = \boldsymbol{\delta} + \boldsymbol{A}\boldsymbol{\lambda}_{t-1} + \boldsymbol{B}\mathbf{Y}_{t-1}, \tag{4.1}$$

where $\mathcal{F}_t$ is the σ-field generated by $\boldsymbol{\lambda}_1, \mathbf{Y}_1, \ldots, \mathbf{Y}_t, \varphi \geq 0, \boldsymbol{\delta} = (\delta_1, \delta_2)^T \in \mathbb{R}_+^2$ and $\boldsymbol{A} = \{\alpha_{ij}\}_{i,j=1,2}$ and $\boldsymbol{B} = \{\beta_{ij}\}_{i,j=1,2}$ are 2×2 matrices with nonnegative entries. Further, $\{\mathbf{Y}_t\}$ has the conditional joint probability mass function (pmf) of the form:

$$\begin{aligned} P(Y_{t,1} = n_1, Y_{t,2} = n_2|\mathcal{F}_{t-1}) &= e^{-(\lambda_{t,1}+\lambda_{t,2}-\varphi)} \frac{(\lambda_{t,1} - \varphi)_1^n}{n_1!} \frac{(\lambda_{t,2} - \varphi)_2^n}{n_2!} \\ &\times \sum_{s=0}^{m\wedge n} \binom{m}{s} \binom{n}{s} s! \left\{ \frac{\varphi}{(\lambda_{t,1} - \varphi)(\lambda_{t,2} - \varphi)} \right\}^s \end{aligned} \tag{4.2}$$

with $m \wedge n = \min\{n_1, n_2\}$ and $\varphi = Cov(Y_{t,1}, Y_{t,2}|\mathcal{F}_{t-1}) \in [0, \lambda_{t,1} \wedge \lambda_{t,2})$ deterministic and independent of t, obtained through a trivariate reduction method.

Let $\theta = (\delta_1, \delta_2, \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \varphi)^T$. For estimating the true parameter θ_0, we recursively define $\tilde{\boldsymbol{\lambda}}_t$, $t \geq 2$, by using an arbitrarily chosen initial value $\tilde{\boldsymbol{\lambda}}_1$ and the equations:

$$\tilde{\boldsymbol{\lambda}}_t = \boldsymbol{\delta} + \boldsymbol{A}\tilde{\boldsymbol{\lambda}}_{t-1} + \boldsymbol{B}\mathbf{Y}_{t-1}, \tag{4.3}$$

where $\boldsymbol{\delta}$, $\boldsymbol{A}$ and $\boldsymbol{B}$ are sometimes written as $\boldsymbol{\delta}(\theta)$, $\boldsymbol{A}(\theta)$ and $\boldsymbol{B}(\theta)$ when the role of θ is emphasized. Then, constructing the conditional likelihood function based on the observation $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$:

$$\tilde{L}(\theta) = \prod_{t=1}^{n} p_\theta(\mathbf{Y}_t \,|\, \tilde{\boldsymbol{\lambda}}_t),$$

where $p_\theta(\mathbf{Y}_t \,|\, \boldsymbol{\lambda}_t)$ is the conditional probability mass function in (4.2), we obtain the CMLE of θ_0 by

$$\hat{\theta}_n = \underset{\theta\in\Theta}{\arg\max}\, \tilde{L}(\theta) = \underset{\theta\in\Theta}{\arg\max}\, \tilde{L}_n(\theta) = \underset{\theta\in\Theta}{\arg\max} \sum_{t=1}^{n} \tilde{\ell}_t(\theta),$$

where $\tilde{\ell}_t(\theta) = \log p_\theta(\mathbf{Y}_t \,|\, \tilde{\boldsymbol{\lambda}}_t)$. According to Andreassen (2013), the CMLE is strongly consistent under the following regularity conditions:

(**B1**) $\theta_0 \in \Theta$ and Θ is compact.

(**B2**) $\delta(\theta), \boldsymbol{A}(\theta)$ and $\mathbf{B}(\theta)$ have non-negative entries and $\boldsymbol{B}(\theta)$ is full rank for all $\theta \in \Theta$.

(**B3**) $\varphi(\theta) < \min(a_1, a_2)$ where $(a_1, a_2)^T = (\mathbf{I} - \boldsymbol{A}(\theta))^{-1}\delta(\theta)$ for all $\theta \in \Theta$.

(**B4**) There exists a $p \in [1, \infty]$ such that $\|\boldsymbol{A}(\theta)\|_p + 2^{1-(1/p)}\|\boldsymbol{B}(\theta)\|_p < 1$ for all $\theta \in \Theta$,

where $\|\boldsymbol{A}\|_p = \max_{\mathbf{x}\neq 0}\{\|\boldsymbol{A}\mathbf{x}\|_p / \|\mathbf{x}\|_p : \mathbf{x} \in \mathbb{C}^n\}$ denotes the p-induced norm of matrix $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ for $1 \leq p \leq \infty$, and $\|\mathbf{x}\|_p$ is the p-norm. When $p = 1$ and $p = \infty$, the norms of $\boldsymbol{A} = \{a_{ij}\}_{1\leq i\leq m, 1\leq j\leq n}$ become $\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$ and $\|\boldsymbol{A}\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$, respectively.

According to Proposition 4.3.1 in Liu (2012), $\{(\mathbf{Y}_t, \boldsymbol{\lambda}_t)\}$ is ergodic and strictly stationary under the assumption (**B2**) and (**B4**): for example, $\|A\|_1 + \|B\|_1 < 1$, corresponding to $p = 1$ in (**B4**), is used for the models in our simulation study. For the univariate process, Doukhan and Kengne (2015) provided the ergodic and stationary conditions in their Assumption A_F.

Heinen (2003) suggest that either $\boldsymbol{A}$ or $\boldsymbol{B}$ is diagonal, since in practice, the diagonal set-up of $\boldsymbol{A}$ is a useful device to reduce the number of model parameters. Because this simplification makes the situation a lot more tractable and eases the verification of the asymptotic normality of CMLE, we also focus on the situation that $\boldsymbol{A}$ is diagonal.

In what follows, we set $\theta = (\theta_1^T, \theta_2^T, \varphi)^T$, where $\theta_1 = (\delta_1, \alpha_1, \beta_{11}, \beta_{12})^T$ and $\theta_2 = (\delta_2, \alpha_2, \beta_{21}, \beta_{22})^T$, $\boldsymbol{\delta} = (\delta_1, \delta_2)^T$, $\boldsymbol{A} = diag(\alpha_1, \alpha_2)^T$, and $\boldsymbol{B} = \{\beta_{ij}\}_{i,j=1,2}$. In this case, (**B1**) and (**B3**) are restated as follows:

(**B1**') $\theta_{01} \in \Theta_1, \theta_{02} \in \Theta_2$, $\varphi_0 \in \Theta_3$, where θ_{01}, θ_{02} and φ_0 are the true parameters of θ_1, θ_2 and φ, respectively, and $\Theta_1, \Theta_2, \Theta_3$ are compact sets; $\Theta = \Theta_1 \times \Theta_2 \times \Theta_3$.

(**B3**') $\varphi < \delta_1/(1-\alpha_1) \wedge \delta_2/(1-\alpha_2)$ for all $\theta \in \Theta$.

Below, we present the asymptotic normality of the CMLE.

Theorem 4.1. *Under* (**B1**)-(**B4**), *as* $n \to \infty$,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{w} N(0, I(\theta_0)^{-1}),$$

where

$$I(\theta_0) := E\left(\frac{\partial \ell_t(\theta_0)}{\partial \theta}\frac{\partial \ell_t(\theta_0)}{\partial \theta^T}\right) = -E\left(\frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta^T}\right)$$

and

$$\begin{aligned}\ell_t(\theta) \quad = \quad & -\{\lambda_{t,1}(\theta) + \lambda_{t,2}(\theta) - \varphi\} + Y_{t,1}\log\{\lambda_{t,1}(\theta) - \varphi\} \\ & + Y_{t,2}\log\{\lambda_{t,2}(\theta) - \varphi\} - \log Y_{t,1}! - \log Y_{t,2}! \\ & + \log\left[\sum_{s=0}^{Y_{t,1}\wedge Y_{t,2}} \binom{Y_{t,1}}{s}\binom{Y_{t,2}}{s} s! \left\{\frac{\varphi}{(\lambda_{t,1}(\theta) - \varphi)(\lambda_{t,2}(\theta) - \varphi)}\right\}^s\right].\end{aligned}$$

4.3 Change point test

In this section, we propose CUSUM tests for detecting a parameter change in bivariate Poisson AR models. We want to test the null and alternative hypotheses

$$H_0 : \theta \text{ does not change over } \mathbf{Y}_1, \cdots, \mathbf{Y}_n \ \ vs. \ \ H_1 : \text{not } H_0.$$

4.3.1 Estimate-based CUSUM test

The CUSUM test based on the estimates is given by

$$T_n^{est,1} = \max_{1\leq k\leq n} \frac{k^2}{n}(\hat{\theta}_k - \hat{\theta}_n)^T \hat{I}_n(\hat{\theta}_k - \hat{\theta}_n),$$

where $\hat{\theta}_k$ is the CMLE of θ_0 based on $\mathbf{Y}_1, \ldots, \mathbf{Y}_k$, and

$$\hat{I}_n = -\frac{1}{n}\sum_{t=1}^{n} \frac{\partial^2 \tilde{\ell}_t(\hat{\theta}_n)}{\partial\theta\partial\theta^T}.$$

It can be seen that $\hat{I}_n$ is a consistent estimator of $I(\theta_0)$ under H_0. The following shows that the CUSUM test has the supremum of independent Brownian bridges as its limiting null distribution, the proof of which is presented in Section 4.7.

Theorem 4.2. *Under the assumption* **(B1)**-**(B4)** *and* H_0*, we have*

$$T_n^{est,1} \xrightarrow{w} \sup_{0\leq s\leq 1} \|\mathbf{B}_9^\circ(s)\|^2,$$

where $\{\mathbf{B}_9^\circ(s), 0 < s < 1\}$ *is a 9-dimensional Brownian bridge. Here,* $\mathbf{B}_d^\circ$ *denotes a d-dimensional vector process the components of which are independent Brownian bridges.*

As an alternative of $T_n^{est,1}$, one can consider

$$T_n^{est,2} = \max_{v_n\leq k\leq n-v_n} \frac{k^2(n-k)^2}{n^3}(\hat{\theta}_k - \tilde{\theta}_k)^T \hat{I}_n'(\hat{\theta}_k - \tilde{\theta}_k), \tag{4.4}$$

(cf. Doukhan and Kengne (2015), where $\tilde{\theta}_k$ are the CMLE of θ_0 based on $\mathbf{Y}_{k+1}, \ldots, \mathbf{Y}_n$,

$$\hat{I}'_n = -\frac{1}{2}\left[\frac{1}{u_n}\sum_{t=1}^{u_n}\frac{\partial^2\tilde{\ell}_t(\hat{\theta}_{u_n})}{\partial\theta\partial\theta^T} + \frac{1}{n-u_n}\sum_{t=u_n+1}^{n}\frac{\partial^2\tilde{\ell}_t(\tilde{\theta}_{u_n})}{\partial\theta\partial\theta^T}\right],$$

and $\{u_n : n \geq 1\}$ and $\{v_n : n \geq 1\}$ are two integer valued sequences satisfying $u_n, v_n \to \infty, u_n/n, v_n/n \to 0$ as $n \to \infty$. Lee et al. (2016a) study (4.4) in univariate zero-inflated generalized Poisson AR models. It can be shown that under **(B1)**-**(B4)** and H_0,

$$T_n^{est,2} \xrightarrow{w} \sup_{0\leq s\leq 1} \|\mathbf{B}_9^{\circ}(s)\|^2, \tag{4.5}$$

the proof of which is provided in Section 6.

4.3.2 Residual-based CUSUM test

In this subsection, we consider the CUSUM test based on the residuals. Let $\boldsymbol{\epsilon}_t = (\epsilon_{t,1}, \epsilon_{t,2})^T = \mathbf{Y}_t - \boldsymbol{\lambda}_t(\theta_0)$ with $\epsilon_{t,i} = Y_{t,i} - \lambda_{t,i}(\theta_{i0})$ for $i = 1, 2$. Since $\boldsymbol{\epsilon}_t$ are not observable, we use the estimated residuals $\hat{\boldsymbol{\epsilon}}_t$:

$$\hat{\boldsymbol{\epsilon}}_t = (\hat{\epsilon}_{t,1}, \hat{\epsilon}_{t,2})^T = \mathbf{Y}_t - \hat{\boldsymbol{\lambda}}_t, \;\; \hat{\boldsymbol{\lambda}}_t = \hat{\boldsymbol{\delta}}_n + \hat{\boldsymbol{A}}_n\hat{\boldsymbol{\lambda}}_{t-1} + \hat{\boldsymbol{B}}_n\mathbf{Y}_{t-1}, \;\; t \geq 2,$$

where $\hat{\boldsymbol{\delta}}_n = (\hat{\delta}_{n,1}, \hat{\delta}_{n,2})^T, \hat{\boldsymbol{A}}_n = diag(\hat{\alpha}_{n,1}, \hat{\alpha}_{n,2}), \hat{\boldsymbol{B}}_n = \{\hat{\beta}_{n,ij}\}_{i,j=1,2}$ and $\hat{\boldsymbol{\lambda}}_1$ is an arbitrarily chosen initial random variable. Then, we employ the test statistic:

$$T_n^{res} = \max_{1\leq k\leq n}\frac{1}{\sqrt{n}}\left\|\hat{\boldsymbol{\Gamma}}^{-1/2}\left(\sum_{t=1}^{k}\hat{\boldsymbol{\epsilon}}_t - \frac{k}{n}\sum_{t=1}^{n}\hat{\boldsymbol{\epsilon}}_t\right)\right\|$$

where $\hat{\boldsymbol{\Gamma}}$ is a consistent estimator of $\boldsymbol{\Gamma} = Var(\boldsymbol{\epsilon}_t) = E\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1^T$: for example,

$$\hat{\boldsymbol{\Gamma}} = \begin{pmatrix} \widehat{Var(\epsilon_{1,1})} & \hat{\varphi}_n \\ \hat{\varphi}_n & \widehat{Var(\epsilon_{1,2})} \end{pmatrix}$$

with $\widehat{Var(\epsilon_{1,i})} = \sum_{t=1}^{n} \hat{\epsilon}_{t,i}^2/n$, $i = 1, 2$, is a consistent estimator of $\mathbf{\Gamma}$ owing to Lemma 11 of Kang and Lee (2014).

Since $\{\boldsymbol{\epsilon}_t\}$ is a stationary and ergodic martingale difference sequence with respect to a filtration $\{\mathcal{F}_t\}$, by a central limit theorem for martingales, we have, as $n \to \infty$,

$$\frac{1}{\sqrt{n}}\mathbf{\Gamma}^{-1/2}\sum_{t=1}^{[ns]}\boldsymbol{\epsilon}_t \xrightarrow{w} \mathbf{B}_2(s),$$

where $\{\mathbf{B}_2(s), 0 < s < 1\}$ is a 2-dimensional standard Brownian motion. Subsequently,

$$\frac{1}{\sqrt{n}}\mathbf{\Gamma}^{-1/2}\left(\sum_{t=1}^{[ns]}\boldsymbol{\epsilon}_t - \frac{[ns]}{n}\sum_{t=1}^{n}\boldsymbol{\epsilon}_t\right) \xrightarrow{w} \mathbf{B}_2^0(s),$$

where $\{\mathbf{B}_2^\circ(s), 0 < s < 1\}$ is a 2-dimensional Brownian bridge. From Lemma 10 of Kang and Lee (2014), we can have

$$\max_{1\le k\le n}\frac{1}{\sqrt{n}}\left|\left\{\sum_{t=1}^{k}\hat{\epsilon}_{t,i} - \frac{k}{n}\sum_{t=1}^{n}\hat{\epsilon}_{t,i}\right\} - \frac{1}{\sqrt{n}}\left\{\sum_{t=1}^{k}\epsilon_{t,i} - \frac{k}{n}\sum_{t=1}^{n}\epsilon_{t,i}\right\}\right| = o_P(1).$$

Thus,

$$\frac{1}{\sqrt{n}}\mathbf{\Gamma}^{-1/2}\left(\sum_{t=1}^{[ns]}\hat{\boldsymbol{\epsilon}}_t - \frac{[ns]}{n}\sum_{t=1}^{n}\hat{\boldsymbol{\epsilon}}_t\right) \xrightarrow{w} \mathbf{B}_2^0(s),$$

which implies the following.

Theorem 4.3. *Under the assumption* **(B1)**-**(B4)** *and* H_0, *we have*

$$T_n^{res} \xrightarrow{w} \sup_{0\le s\le 1}\|\mathbf{B}_2^0(s)\|.$$

4.4 Simulation results

In this section, we report simulation results to evaluate the performance of the proposed test statistics, $T_n^{est,1}, T_n^{est,2}$ and T_n^{res}. We consider model (4.1) with the initial value of $\boldsymbol{\lambda}_1$ equal to $(0,0)^T$. We employ the nominal level $\alpha = 0.05$, sample size $n = 300, 500, 1000$, and the number of realizations 500. The critical values for $T_n^{est,1}, T_n^{est,2}$ and T_n^{res} are obtained as 5.632, 5.632 and 2.408, respectively, through a Monte Carlo simulation as in Lee et al. (2003).

We consider the models:

Model 1 : $\boldsymbol{B}$ is a diagonal matrix : $(\beta_1, \beta_2) = (0.1, 0.2)$,

Model 2 : $\boldsymbol{B}$ is a non-diagonal matrix : $(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (0.1, 0.2, 0.1, 0.2)$

with $(\delta_1, \alpha_1, \delta_2, \alpha_2) = (3, 0.2, 1, 0.1), (3, 0.4, 1, 0.3)$ and $\varphi = 0, 0.3, 07$. These settings satisfy **(B1)**-**(B4)** (with $p = 1$) and particularly guarantee the ergodicity and stationarity of the bivariate Poisson INGARCH(1,1) model. From Table 4.1, we can see that the sizes of T_n^{res} are close the nominal level, whereas $T_n^{est,1}$ and $T_n^{est,2}$ have some size distortions: that is, the size has a tendency to increase as either $\boldsymbol{\alpha}$ or φ increases.

To examine the power, we consider the parameter change from θ to θ' at $[n\tau]$ with τ =1/3, 1/2, 2/3:

Case 1 : $(\delta_1, \delta_2) = (3, 1)$ changes to $(\delta_1', \delta_2') = (2.7, 1.5)$,

Case 2 : $(\alpha_1, \alpha_2) = (0.2, 0.1)$ changes to $(\alpha_1', \alpha_2') = (0.3, 0.2)$,

Case 3 : $\varphi = 0$ changes to $\varphi = 0.3, 0.7$,

wherein the other parameters remain constant. From Tables 4.2 and 4.5, we can see that $T_n^{est,1}$ and $T_n^{est,2}$ produce similar powers and T_n^{res} produces slightly better powers than these two. Also, the power gets closer to 1 as the sample size increase in Cases 1 and 2, but becomes lessened in Case 3: particularly, T_n^{res} performs poorly. As anticipated, all the tests appear to have the largest powers at $\tau = 1/2$.

Overall, our finding show that the CUSUM test is a functional tool to detect a parameter change for bivariate Poisson INGARCH models. Among the three tests, T_n^{res} seems to be the most recommendable, although not powerful at detecting a change of φ. For this, one can still use the other two tests but needs to develop a new method, which we leave as our future project.

4.5 Real data analysis

In this section, we give a real data example. We analyze two daily data sets of car accidents that occurred in Seongdong (Y_1) and Seocho (Y_2) counties in Seoul, Korea during the period from January 1, 2011 to December 31, 2012 (the sample size is 731). The time plots for the two data are given in Figure 4.1. The mean and variance are 2.927 and 3.574 for Y_1, and 5.661 and 7.164 for Y_2, indicating over-dispersion. The autocorrelation and partial autocorrelation functions of each data are given in Figure 4.2, indicating serial dependence. Because the covariance and correlation of the two data sets are obtained as 0.6987 and 0.1381, respectively, a bivariate Poisson INGARCH (1,1) model is fitted to the data.

The CUSUM test shows that $T_n^{est,1} = 5.2481, T_n^{est,2} = 15.236$ and $T_n^{res} = 2.410$. Based on this result, the null hypothesis of no changes is rejected by $T_n^{est,2}$ and T_n^{res} at the nominal level 0.05, and $T_n^{est,1}$ at the nominal level 0.1. Since $T_n^{est,2}$ and T_n^{res} are maximized at $t = 343$ (see Figure 4.3), the change point can be estimated as December 9, 2011. The parameter estimates for the full series and two subseries before/after the change point are presented in Table 4.6. Particularly, it shows that $\hat{\varphi}$ increases from 0.489 to 0.638.

4.6 Concluding remarks

In this study, we considered the problem of testing for a parameter change in bivariate Poisson INGARCH(1,1) models, constructed via a trivariate reduction method of independent Poisson variables. We verified that the conditional maximum likelihood estimator of the models parameters is asymptotically normal, and based on this, we constructed the CMLE- and residual-based CUSUM tests and derived their limiting null distributions. To evaluate the performance of the tests, we conducted a simulation study and real data analysis using two daily data sets of car accidents in Seoul, Korea during the period from January 1, 2011 to December 31, 2012. The results demonstrated the validity of the CUSUM tests. Although this work yields satisfactory results, there are some aspects that should be considered for its extension. First, the proposed bivariate Poisson INGARCH(1,1) model has a shortcoming that it can only cover the process with a positive correlation. Second, there is a demand to develop more general INGARCH type models, such as higher order Poisson INGARCH(p, q) models and multivariate models, to improve the applicability of the INGARCH models. At this moment, these issues are somewhat beyond the scope of the current study, so are left as our future project.

4.7 Appendix

In this section, we verify the theorems in the previous sections. The proofs of the lemmas below are provided in Supplementary material.

Lemma 4.1. *Let V stand for a generic positive integrable random variable and $0 < \rho < 1$ be a generic constant. Under* **(B1)**-**(B4)***, we have that for $i = 1, 2$,*

(i) $E\left(\sup_{\theta\in\Theta}\|\boldsymbol{\lambda}_t(\theta)\|\right) < \infty$ *and* $E\left(\sup_{\theta\in\Theta}\|\tilde{\boldsymbol{\lambda}}_t(\theta)\|\right) < \infty$.

(ii) $\boldsymbol{\lambda}_t(\theta) = \boldsymbol{\lambda}_t(\theta_0)$ *a.s. implies* $\theta_i = \theta_{i0}$.

(iii) $\lambda_{t,i}(\theta_i)$ *is twice continuously differentiable with respect to* θ_i *and satisfies*

$$E\left(\sup_{\theta_i\in\Theta_i}\left\|\frac{\partial\lambda_{t,i}(\theta_i)}{\partial\theta_i}\right\|_p\right)^4 < \infty \ \ and\ \ E\left(\sup_{\theta_i\in\Theta_i}\left\|\frac{\partial^2\lambda_{t,i}(\theta_i)}{\partial\theta_i\partial\theta_i^T}\right\|_p\right)^2 < \infty.$$

(iv) *For all* t, *a.s.,*

$$\sup_{\theta_i\in\Theta_i}\left\|\frac{\partial\tilde{\lambda}_{t,i}(\theta_i)}{\partial\theta_i} - \frac{\partial\lambda_{t,i}(\theta_i)}{\partial\theta_i}\right\|_p \le V\rho^t \ \ and\ \ \sup_{\theta_i\in\Theta_i}\left\|\frac{\partial^2\tilde{\lambda}_{t,i}(\theta_i)}{\partial\theta_i\partial\theta_i^T} - \frac{\partial^2\lambda_{t,i}(\theta_i)}{\partial\theta_i\partial\theta_i^T}\right\|_p \le V\rho^t.$$

(v) $\nu^T\frac{\partial\lambda_{t,i}(\theta_{i0})}{\partial\theta_i} = 0$ *implies* $\nu = 0$.

(vi) $\sup_{\theta\in\Theta}\left\|\tilde{\boldsymbol{\lambda}}_t(\theta) - \boldsymbol{\lambda}_t(\theta)\right\| \le V\rho^t$ *a.s. for all* t.

Note that the first derivatives of $\ell_t(\theta)$ are expressed as

$$\begin{aligned}\frac{\partial\ell_t(\theta)}{\partial\theta} &= \left(D_{t,1}(\theta)s_1(\theta_1)^T, D_{t,2}(\theta)s_2(\theta_2)^T, D_{t,3}(\theta)\right)^T \\ &= \begin{pmatrix} D_{t,1}(\theta)\mathbf{I}_4 & \mathbf{0}_{4\times4} & \mathbf{0}_{4\times1} \\ \mathbf{0}_{4\times4} & D_{t,2}(\theta)\mathbf{I}_4 & \mathbf{0}_{4\times1} \\ \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} & D_{t,3}(\theta)\end{pmatrix}\begin{pmatrix} s_1(\theta_1) \\ s_2(\theta_2) \\ 1\end{pmatrix} := \boldsymbol{D}_t(\theta)\Lambda_t(\theta),\end{aligned}$$

where I_4 denotes the 4×4 identity matrix and $\mathbf{0}_{m\times n}$ denotes the $m\times n$ matrix with all elements equal to 0,

$$\begin{aligned}s_i(\theta_i) &= \frac{\partial\lambda_{t,i}(\theta_i)}{\partial\theta_i},\ i = 1,2, \\ D_{t,i}(\theta) &= \frac{Y_{t,i}}{\lambda_{t,i}(\theta_i)-\varphi} - 1 - \frac{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,1)}{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,0)}\cdot\frac{1}{\lambda_{t,i}(\theta_i)-\varphi},\quad i=1,2, \\ D_{t,3}(\theta) &= \frac{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,1)}{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,0)}\cdot\frac{1}{\varphi} - 1 - D_{t,1}(\theta) - D_{t,2}(\theta),\end{aligned}$$

$$g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, r) = \sum_{s=0}^{Y_{t,1}\wedge Y_{t,2}} \binom{Y_{t,1}}{s}\binom{Y_{t,2}}{s} s! s^r f(\boldsymbol{\lambda}_t, \varphi)^s, \quad r = 0, 1, 2,$$
$$f(\boldsymbol{\lambda}_t, \varphi) = \frac{\varphi}{(\lambda_{t,1}(\theta_1) - \varphi)(\lambda_{t,2}(\theta_2) - \varphi)}.$$

Further, the second derivatives are given by

$$\begin{aligned}
\frac{\partial^2 \ell_t(\theta)}{\partial\theta\partial\theta^T} &= \begin{pmatrix} F_{t,11}(\theta)s_1(\theta_1)s_1(\theta_1)^T & F_{t,12}(\theta)s_1(\theta_1)s_2(\theta_2)^T & F_{t,13}(\theta)s_1(\theta_1) \\ F_{t,21}(\theta)s_2(\theta_2)s_1(\theta_1)^T & F_{t,22}(\theta)s_2(\theta_2)s_2(\theta_2)^T & F_{t,23}(\theta)s_2(\theta_2) \\ F_{t,31}(\theta)s_1(\theta_1)^T & F_{t,32}(\theta)s_2(\theta_2)^T & F_{t,33}(\theta) \end{pmatrix} \\
&\quad + \begin{pmatrix} D_{t,1}(\theta)s_{11}(\theta_1) & 0 & 0 \\ 0 & D_{t,2}(\theta)s_{22}(\theta_2) & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&:= \boldsymbol{F}_t(\theta) + \boldsymbol{D}_t(\theta)\frac{\partial\Lambda_t(\theta)}{\partial\theta},
\end{aligned}$$

where

$$\begin{aligned}
s_{ii}(\theta_i) &= \frac{\partial^2\lambda_{t,i}(\theta_i)}{\partial\theta_i\partial\theta_i^T}, \ i = 1, 2, \\
F_{t,ii}(\theta) &= -\left\{Y_{t,i} - h_t(\theta) - \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)}\right\}\frac{1}{(\lambda_{t,i}(\theta_i) - \varphi)^2}, \ i = 1, 2, \\
F_{t,ij}(\theta) &= \frac{h_t(\theta)}{(\lambda_{t,1}(\theta_1) - \varphi)(\lambda_{t,2}(\theta_2) - \varphi)}, \ 1 \le i \ne j \le 2, \\
F_{t,i3}(\theta) &= F_{t,3i}(\theta) = -\frac{h_t(\theta)}{\varphi(\lambda_{t,i}(\theta_i) - \varphi)} - F_{t,ii}(\theta) - F_{t,12}(\theta), \ i = 1, 2, \\
F_{t,33}(\theta) &= \frac{h_t(\theta)}{\varphi}\left\{\frac{1}{\varphi} + \frac{1}{\lambda_{t,1}(\theta_1) - \varphi} + \frac{1}{\lambda_{t,2}(\theta_2) - \varphi}\right\} \\
&\qquad - \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)} \cdot \frac{1}{\varphi^2} - F_{t,31}(\theta) - F_{t,32}(\theta), \\
h_t(\theta) &= \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 2)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)} - \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)^2}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)^2}.
\end{aligned}$$

Lemma 4.2. *For $i, j = 1, 2, 3$, let $\tilde{D}_{t,i}(\theta)$ and $\tilde{F}_{t,ij}(\theta)$ be the same as $D_{t,i}(\theta)$ and $F_{t,ij}(\theta)$ with $\boldsymbol{\lambda}_t$ replaced by $\tilde{\boldsymbol{\lambda}}_t$. Then, under (B1)-(B4),*

$$\sup_{\theta\in\Theta}|D_{t,i}(\theta)| \le C\|\mathbf{Y}_t\| + 1, \quad \sup_{\theta\in\Theta}|\tilde{D}_{t,i}(\theta)| \le C\|\mathbf{Y}_t\| + 1,$$
$$\sup_{\theta\in\Theta}|F_{t,ii}(\theta)| \le C\|\mathbf{Y}_t\|^2, \quad \sup_{\theta\in\Theta}|\tilde{F}_{t,ii}(\theta)| \le C\|\mathbf{Y}_t\|^2$$

for some positive constant C. Further, as $t \to \infty$,

$$\sup_{\theta\in\Theta}\left|\tilde{D}_{t,i}(\theta) - D_{t,i}(\theta)\right| \to 0 \quad a.s. \quad \text{and} \qquad \sup_{\theta\in\Theta}\left|\tilde{F}_{t,ij}(\theta) - F_{t,ij}(\theta)\right| \to 0 \quad \text{a.s..}$$

Lemma 4.3. *Under the assumption (B1)-(B4),*

$$E\sup_{\theta\in\Theta}\left\|\frac{\partial^2 \ell_t(\theta)}{\partial\theta\partial\theta^T}\right\|_p < \infty \;\; and \;\; E\sup_{\theta\in\Theta}\left\|\frac{\partial \ell_t(\theta)}{\partial\theta}\frac{\partial \ell_t(\theta)}{\partial\theta^T}\right\|_p < \infty.$$

Lemma 4.4. *Under assumption (B1)-(B4),*

$$\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{\partial \tilde{\ell}_t(\theta_0)}{\partial\theta} - \frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{\partial \ell_t(\theta_0)}{\partial\theta}\right\|_p = o_P(1).$$

Lemma 4.5. *Under assumption (B1)-(B4),*

$$\sup_{\theta\in\Theta}\left\|\frac{1}{n}\sum_{t=1}^n \frac{\partial^2 \tilde{\ell}_t(\theta)}{\partial\theta\partial\theta^T} - \frac{1}{n}\sum_{t=1}^n \frac{\partial^2 \ell_t(\theta)}{\partial\theta\partial\theta^T}\right\|_p \to 0 \quad \text{a.s..}$$

Lemma 4.6. *Under (B1)-(B4),*

$$-\frac{1}{n}\frac{\partial^2 \tilde{L}_n(\theta_n^*)}{\partial\theta\partial\theta^T} \to I(\theta_0) \quad a.s.,$$

where θ_n^* is any intermediate point between $\hat{\theta}_n$ and θ_0.

Lemma 4.7. *Under (B1)-(B4), $\{\partial\ell_t(\theta_0)/\partial\theta; \mathcal{F}_t\}$ forms a stationary ergodic martingale difference sequence.*

Proof of Theorem 4.1. In view of Lemma 4.1 and 4.3, we have that $I(\theta_0)$ exists and is positive definite. According to Lemma 4.7, by using a martingale central limit theorem and the Cramér-Wold device, we can show that $\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\partial\ell_t(\theta_0)/\partial\theta$ converges weakly to $N(0, I(\theta_0))$. By using Taylor' theorem, we have

$$0 = \frac{1}{\sqrt{n}}\frac{\partial\tilde{L}_n(\hat{\theta}_n)}{\partial\theta} = \frac{1}{\sqrt{n}}\frac{\partial\tilde{L}_n(\theta_0)}{\partial\theta} - \left(\frac{1}{n}\frac{\partial^2\tilde{L}_n(\theta_n')}{\partial\theta\partial\theta^T}\right)\sqrt{n}(\hat{\theta}_n - \theta_0), \tag{4.6}$$

where θ_n' is an intermediate point between θ_0 and $\hat{\theta}_n$. Then using by Lemma 4.4 and 4.6, we can assert the theorem. □

Proof of Theorem 4.2. From (4.6), we have that for $0 < s < 1$,

$$I(\theta_0)\frac{[ns]}{\sqrt{n}}(\hat{\theta}_{[ns]} - \theta_0) = \frac{1}{n}\frac{\partial\tilde{L}_{[ns]}(\theta_0)}{\partial\theta} + \sqrt{\frac{[ns]}{n}}\tilde{\Delta}_{[ns]}, \tag{4.7}$$

where

$$\tilde{\Delta}_k = \begin{cases} -\left\{I(\theta_0) + \frac{1}{k}\frac{\partial^2\tilde{L}_k(\theta_k')}{\partial\theta\partial\theta^T}\right\}\left(\frac{1}{k}\frac{\partial^2\tilde{L}_k(\theta_k')}{\partial\theta\partial\theta^T}\right)^{-1}\frac{1}{\sqrt{k}}\frac{\partial\tilde{L}_k(\theta_0)}{\partial\theta}, & \text{if } \left(\frac{1}{k}\frac{\partial^2\tilde{L}_k(\theta_k')}{\partial\theta\partial\theta^T}\right)^{-1} \text{ exists} \\ \left\{I(\theta_0) + \frac{1}{n}\frac{\partial^2\tilde{L}_n(\theta_n')}{\partial\theta\partial\theta^T}\right\}\sqrt{k}(\hat{\theta}_k - \theta_0), & \text{otherwise.} \end{cases}$$

By Lemma 4.7, using the functional central limit theorem for martingales, we can get

$$I(\theta_0)^{-1/2}\frac{1}{\sqrt{n}}\frac{\partial L_{[ns]}(\theta_0)}{\partial\theta} \xrightarrow{w} \mathbf{B}_9(s),$$

where $\{\mathbf{B}_9(s), 0 < s < 1\}$ is a 9-dimensional standard Brownian motion. Then, from Lemma 4.4, we have

$$I(\theta_0)^{-1/2}\frac{1}{\sqrt{n}}\frac{\partial\tilde{L}_{[ns]}(\theta_0)}{\partial\theta} \xrightarrow{w} \mathbf{B}_9(s).$$

Further, in a similar way to prove Lemma 9 of Kang and Lee (2014), we can show that

$$\max_{1\le k\le n}\sqrt{\frac{k}{n}}\|\tilde{\Delta}_k\| = o_P(1).$$

Thus, (4.7) converges weakly to $\mathbf{B}_9(s)$, which establishes the theorem. □

Proof of (4.5). From (4.7), we can express

$$I(\theta_0)\sqrt{n}(\hat{\theta}_k - \theta_0) = \frac{1}{k}\frac{\partial \tilde{L}_k(\theta_0)}{\partial \theta} + \frac{1}{\sqrt{k}}\tilde{\Delta}_k,$$
$$I(\theta_0)\sqrt{n}(\tilde{\theta}_k - \theta_0) = \frac{1}{n-k}\frac{\partial \tilde{L}^*_k(\theta_0)}{\partial \theta} + \frac{1}{\sqrt{n-k}}\tilde{\Delta}^*_k,$$

where $\tilde{L}^*_k(\theta) = \sum_{t=k+1}^{n} \tilde{\ell}_t(\theta)$ and

$$\tilde{\Delta}^*_k = \begin{cases} -\left\{I(\theta_0) + \frac{1}{k}\frac{\partial^2 \tilde{L}^*_k(\theta'_k)}{\partial\theta\partial\theta^T}\right\}\left(\frac{1}{k}\frac{\partial^2 \tilde{L}^*_k(\theta'_k)}{\partial\theta\partial\theta^T}\right)^{-1}\frac{1}{\sqrt{k}}\frac{\partial \tilde{L}^*_k(\theta_0)}{\partial\theta}, & \text{if } \left(\frac{1}{k}\frac{\partial^2 \tilde{L}^*_k(\theta'_k)}{\partial\theta\partial\theta^T}\right)^{-1} \text{ exists} \\ \left\{I(\theta_0) + \frac{1}{n}\frac{\partial^2 \tilde{L}_n(\theta'_n)}{\partial\theta\partial\theta^T}\right\}\sqrt{k}(\hat{\theta}_k - \theta_0), & \text{otherwise.} \end{cases}$$

Subsequently, we have

$$\begin{aligned} &\frac{k(n-k)}{n^{3/2}} I(\theta_0)(\hat{\theta}_k - \tilde{\theta}_k) \\ &\quad = \frac{1}{\sqrt{n}}\left\{\frac{\partial \tilde{L}_k(\theta_0)}{\partial\theta} - \frac{k}{n}\frac{\partial \tilde{L}_n(\theta_0)}{\partial\theta}\right\} + \frac{\sqrt{k(n-k)}}{n}\left(\sqrt{\frac{k}{n}}\tilde{\Delta}_k + \sqrt{\frac{n-k}{n}}\tilde{\Delta}^*_k\right), \end{aligned}$$

since $\tilde{L}^*_k(\theta) = \tilde{L}_n(\theta) - \tilde{L}_k(\theta)$. Then, (4.5) can be verified in a similar fashion to the proof of Theorem 4.2. □

4.8 Supplementary Material

In this Supplement we provide the proofs of Lemmas 4.1-4.7.

Proof of Lemma 4.1 By iterating (4.3) in Section 4.2, we have

$$\boldsymbol{\lambda}_t = (\mathbf{I} + \boldsymbol{A} + \cdots + \boldsymbol{A}^{t-2})\boldsymbol{\delta} + \boldsymbol{A}^{t-1}\boldsymbol{\lambda}_1 + \sum_{k=1}^{t-1} \boldsymbol{A}^{k-1}\boldsymbol{B}\mathbf{Y}_{t-k}. \tag{4.8}$$

Therefore, for $i = 1, 2$,

$$\lambda_{t,i}(\theta_i) = \frac{1-\alpha_i^{t-1}}{1-\alpha_i}\delta_i + \alpha_i^{t-1}\lambda_{1,i} + \sum_{k=1}^{t-1}\alpha_i^{k-1}\left(\beta_{i1}Y_{t-k-1,1} + \beta_{i2}Y_{t-k-1,2}\right).$$

Then, (i)-(v) can be obtained by using the arguments in the proof of Theorem 3 of Kang and Lee (2014). Meanwhile, from (4.8), we have

$$\sup_{\theta\in\Theta}\|\tilde{\boldsymbol{\lambda}}_t(\theta) - \boldsymbol{\lambda}_t(\theta)\| = \sup_{\theta\in\Theta}\|\boldsymbol{A}^{t-1}(\tilde{\boldsymbol{\lambda}}_1 - \boldsymbol{\lambda}_1)\| \le V\rho^t,$$

where $\rho = \sup_{\theta\in\Theta}\|\boldsymbol{A}\|_p$ and $V = \|\tilde{\boldsymbol{\lambda}}_1 - \boldsymbol{\lambda}_1\|/\rho$. This establishes (vi). □

Proof of Lemma 4.2. By our assumption, we have

$$\sup_{\theta\in\Theta}\frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, r)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)} \le (Y_{t,1}\wedge Y_{t,2})^r \cdot \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)} \le \|\mathbf{Y}_t\|^r, \tag{4.9}$$

$$\sup_{\theta\in\Theta}|h_t(\theta)| \le \|\mathbf{Y}_t\| + \|\mathbf{Y}_t\|^2 \le 2\|\mathbf{Y}_t\|^2. \tag{4.10}$$

Furthermore, according to **(B3')**, we can take ϵ such that

$$0 < \epsilon < \inf_{\theta\in\Theta}\left\{\min(\delta_1/(1-\alpha_1), \delta_2/(1-\alpha_2)) - \varphi\right\}, \tag{4.11}$$

which can lead to $\sup_{\theta\in\Theta}\{\lambda_{t,i}(\theta_i) - \varphi\} > \epsilon$. Then, using (4.9)-(4.11), we have that for $i, j = 1, 2, 3$,

$$\sup_{\theta\in\Theta}|D_{t,i}(\theta)| \le C\|\mathbf{Y}_t\| + 1, \quad \sup_{\theta\in\Theta}|\tilde{D}_{t,i}(\theta)| \le C\|\mathbf{Y}_t\| + 1,$$
$$\sup_{\theta\in\Theta}|F_{t,ii}(\theta)| \le C\|\mathbf{Y}_t\|^2, \quad \sup_{\theta\in\Theta}|\tilde{F}_{t,ii}(\theta)| \le C\|\mathbf{Y}_t\|^2.$$

Next, we show that

$$\sup_{\theta\in\Theta}\left|\tilde{D}_{t,1}(\theta) - D_{t,1}(\theta)\right| \to 0 \quad a.s. \tag{4.12}$$

Note that

$$\sup_{\theta\in\Theta}|\tilde{D}_{t,1}(\theta) - D_{t,1}(\theta)| \le Y_{t,1}\cdot\sup_{\theta\in\Theta}\left|\frac{1}{\tilde{\lambda}_{t,1}(\theta)-\varphi} - \frac{1}{\lambda_{t,1}(\theta)-\varphi}\right| \tag{4.13}$$

$$+\sup_{\theta\in\Theta}\left|\frac{g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,1)}{g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,0)}\cdot\frac{1}{\tilde{\lambda}_{t,1}(\theta)-\varphi}-\frac{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,1)}{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,0)}\cdot\frac{1}{\lambda_{t,1}(\theta)-\varphi}\right|.$$

By Lemma 4.1, we have

$$\sup_{\theta\in\Theta}\left|\frac{1}{\tilde{\lambda}_{t,1}(\theta_1)-\varphi}-\frac{1}{\lambda_{t,1}(\theta_1)-\varphi}\right|\le\sup_{\theta\in\Theta}\frac{\|\tilde{\boldsymbol{\lambda}}_t(\theta)-\boldsymbol{\lambda}_t(\theta)\|}{\epsilon^2}\le\frac{V}{\epsilon^2}\rho^t. \tag{4.14}$$

The second term of of the RHS of (4.13) is bounded by $I+II$ with

$$I=\sup_{\theta\in\Theta}\left|\frac{g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,1)}{g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,0)}\left\{\frac{1}{\tilde{\lambda}_{t,1}(\theta_1)-\varphi}-\frac{1}{\lambda_{t,1}(\theta_1)-\varphi}\right\}\right|,$$
$$II=\sup_{\theta\in\Theta}\left|\left\{\frac{g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,1)}{g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,0)}-\frac{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,1)}{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,0)}\right\}\frac{1}{\lambda_{t,1}(\theta_1)-\varphi}\right|.$$

Note that $I\le\|\mathbf{Y}_t\|V\rho^t/\epsilon^2$ owing to (4.9) and (4.14). Further, provided $f(\tilde{\boldsymbol{\lambda}}_t,\varphi)\ge f(\boldsymbol{\lambda}_t,\varphi)$, we have

$$\begin{aligned}
&|g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,r)-g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,r)|\\
\le&\sum_{s=0}^{Y_{t,1}\wedge Y_{t,2}}\binom{Y_{t,1}}{s}\binom{Y_{t,2}}{s}s^r s!|f(\tilde{\boldsymbol{\lambda}}_t,\varphi)^s-f(\boldsymbol{\lambda}_t,\varphi)^s|\\
=&\sum_{s=0}^{Y_{t,1}\wedge Y_{t,2}}\binom{Y_{t,1}}{s}\binom{Y_{t,2}}{s}s^r s!f(\boldsymbol{\lambda}_t,\varphi)^s\left(\frac{f(\tilde{\boldsymbol{\lambda}}_t,\varphi)^s}{f(\boldsymbol{\lambda}_t,\varphi)^s}-1\right)\\
\le&\sum_{s=0}^{Y_{t,1}\wedge Y_{t,2}}\binom{Y_{t,1}}{s}\binom{Y_{t,2}}{s}s^r s!f(\boldsymbol{\lambda}_t,\varphi)^s\left(\exp\left(2s\frac{V\rho^t}{\epsilon}\right)-1\right)\\
\le&\ g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,r)\left|\exp\left(\frac{2\|\mathbf{Y}_t\|V}{\epsilon}\rho^t\right)-1\right|.
\end{aligned}\tag{4.15}$$

Then, due to (4.9) and (4.15), we get

$$\begin{aligned}
&\left|\frac{g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,r)}{g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,0)}-\frac{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,r)}{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,0)}\right|\\
&\qquad\le\frac{g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,r)\left|g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,0)-g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,0)\right|}{g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,0)\cdot g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,0)}
\end{aligned}$$

$$+\frac{\left|g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,r)-g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,r)\right|}{g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,0)}$$

$$\leq 2\|\mathbf{Y}_t\|^r\left|\exp\left(\frac{2\|\mathbf{Y}_t\|V}{\epsilon}\rho^t\right)-1\right|. \tag{4.16}$$

If $f(\tilde{\boldsymbol{\lambda}}_t,\varphi)\leq f(\boldsymbol{\lambda}_t,\varphi)$, the inequality in (4.15) holds with $g(\mathbf{Y}_t,\boldsymbol{\lambda}_t,\varphi,r)$ replaced by $g(\mathbf{Y}_t,\tilde{\boldsymbol{\lambda}}_t,\varphi,r)$, and henceforth, (4.16) still holds for this case. Therefore,

$$II\leq\frac{2\|\mathbf{Y}_t\|}{\epsilon}\left|\exp\left(\frac{2\|\mathbf{Y}_t\|}{\epsilon}V\rho^t\right)-1\right|,$$

and as such, the RHS of (4.13) is bounded by

$$\frac{2\|\mathbf{Y}_t\|V}{\epsilon^2}\rho^t+\frac{2\|\mathbf{Y}_t\|}{\epsilon}\left|\exp\left(\frac{2\|\mathbf{Y}_t\|}{\epsilon}V\rho^t\right)-1\right|. \tag{4.17}$$

Since $\|\mathbf{Y}_t\|V\rho^t\to 0$ a.s. as $t\to\infty$, the first term of the above equation goes to 0 a.s. as $t\to\infty$. The mean value theorem shows that

$$\sum_{t=1}^{\infty}\|\mathbf{Y}_t\|\left|\exp\left(\frac{2\|\mathbf{Y}_t\|}{\epsilon}V\rho^t\right)-1\right|\leq\sum_{t=1}^{\infty}\left\{\frac{2\|\mathbf{Y}_t\|^2V\rho^t}{\epsilon^2}\exp\left(\frac{2\|\mathbf{Y}_t\|}{\epsilon}V\rho^t\right)\right\}$$

$$\leq\exp\left(2V\sup_{t\geq 1}\|\mathbf{Y}_t\|\rho^t/\epsilon\right)\sum_{t=1}^{\infty}2\|\mathbf{Y}_t\|^2V\rho^t/\epsilon^2. \tag{4.18}$$

Further, $E\left(\sup_{t\geq 1}\|\mathbf{Y}_t\|\rho^t\right)\leq\sum_{t=1}^{\infty}\rho^tE\|\mathbf{Y}_t\|<\infty$, which implies $\sup_{t\geq 1}\|\mathbf{Y}_t\|\rho^t<\infty$ a.s.. Therefore, (4.18) is a.s. finite, so that the second term of (4.17) also goes to 0 a.s. as $t\to\infty$. This implies (4.12). Similarly, it can be shown that for $i=2,3$,

$$\sup_{\theta\in\Theta}|\tilde{D}_{t,i}(\theta)-D_{t,i}(\theta)|\to 0\ \text{ a.s. as }\ t\to\infty.$$

Now, we show that as $t\to\infty$,

$$\sup_{\theta\in\Theta}|\tilde{F}_{t,11}(\theta)-F_{t,11}(\theta)|\to 0\quad\text{a.s.}. \tag{4.19}$$

We express

$$\begin{aligned}
&\tilde{F}_{t,11}(\theta) - F_{t,11}(\theta)\\
&\quad = \left\{ Y_{t,1} - \tilde{h}_t(\theta) - \frac{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 1)}{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 0)} \right\} \cdot \left\{ \frac{1}{(\tilde{\lambda}_{t,1}(\theta_1) - \varphi)^2} - \frac{1}{(\lambda_{t,1}(\theta_1) - \varphi)^2} \right\}\\
&\qquad + \left\{ \tilde{h}_t(\theta) - h_t(\theta) + \frac{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 1)}{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 0)} - \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)} \right\} \cdot \frac{1}{(\lambda_{t,1}(\theta_1) - \varphi)^2} .
\end{aligned}$$

From (4.9) and (4.16), we have

$$\begin{aligned}
&\sup_{\theta\in\Theta} \left| \tilde{h}_t(\theta) - h_t(\theta) \right|\\
&\leq \sup_{\theta\in\Theta} \left| \frac{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 2)}{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 0)} - \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 2)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)} \right| + \sup_{\theta\in\Theta} \left| \frac{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 1)^2}{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 0)^2} - \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)^2}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)^2} \right|\\
&\leq \sup_{\theta\in\Theta} \left| \frac{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 2)}{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 0)} - \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 2)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)} \right|\\
&\qquad + \sup_{\theta\in\Theta} \left| \frac{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 1)}{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 0)} - \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)} \right| \left\{ \frac{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 1)}{g(\mathbf{Y}_t, \tilde{\boldsymbol{\lambda}}_t, \varphi, 0)} + \frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)} \right\}\\
&\leq \left(2\|\mathbf{Y}_t\|^2 + 4\|\mathbf{Y}_t\|^2 \right) \left| \exp\left(\frac{2\|\mathbf{Y}_t\| V}{\epsilon} \rho^t \right) - 1 \right| .
\end{aligned}$$

Thus, using (4.9), (4.10) and Lemma 4.1, we can have

$$\begin{aligned}
&\sup_{\theta\in\theta} \left| \tilde{F}_{t,11}(\theta) - F_{t,11}(\theta) \right|\\
&\qquad \leq \frac{\|\mathbf{Y}_t\| + 2\|\mathbf{Y}_t\|^2 + \|\mathbf{Y}_t\|}{\epsilon^4} \sup_{\theta\in\Theta} \left\{ \lambda_{t,1}(\theta_1) + \tilde{\lambda}_{t,1}(\theta_1) - 2\varphi \right\} V \rho^t\\
&\qquad + \frac{6\|\mathbf{Y}_t\|^2 + 2\|\mathbf{Y}_t\|}{\epsilon^2} \left| \exp\left(\frac{2\|\mathbf{Y}_t\| V}{\epsilon} \rho^t \right) - 1 \right| .
\end{aligned}$$

Then, Lemma 4.1 and (4.18) asserts (4.19).

Next, we show that

$$\sup_{\theta\in\theta} \left| \tilde{F}_{t,12}(\theta) - F_{t,12}(\theta) \right| \to 0 \quad \text{a.s. as } t \to \infty. \tag{4.20}$$

Due to (4.14), we can see that

$$\sup_{\theta\in\Theta}\left|\frac{1}{(\tilde{\lambda}_{t,1}(\theta_1)-\varphi)(\tilde{\lambda}_{t,2}(\theta_2)-\varphi)}-\frac{1}{(\lambda_{t,1}(\theta_1)-\varphi)(\lambda_{t,2}(\theta_2)-\varphi)}\right|$$
$$\le \sup_{\theta\in\Theta}\left[\frac{1}{\tilde{\lambda}_{t,1}(\theta_1)-\varphi}\left|\frac{1}{\tilde{\lambda}_{t,2}(\theta)-\varphi}-\frac{1}{\lambda_{t,2}(\theta_2)-\varphi}\right|\right]$$
$$+\sup_{\theta\in\Theta}\left[\frac{1}{\lambda_{t,2}(\theta_2)-\varphi}\left|\frac{1}{\tilde{\lambda}_{t,1}(\theta_1)-\varphi}-\frac{1}{\lambda_{t,1}(\theta_1)-\varphi}\right|\right]\le\frac{2V}{\epsilon^3}\rho^t,$$

which in turn implies

$$\sup_{\theta\in\theta}\left|\tilde{F}_{t,12}(\theta)-F_{t,12}(\theta)\right|$$
$$=\sup_{\theta\in\theta}\left|\frac{\tilde{h}_t(\theta)}{(\tilde{\lambda}_{t,1}(\theta_1)-\varphi)(\tilde{\lambda}_{t,2}(\theta)-\varphi)}-\frac{h_t(\theta)}{(\lambda_{t,1}(\theta_1)-\varphi)(\lambda_{t,2}(\theta)-\varphi)}\right|$$
$$=\sup_{\theta\in\Theta}|\tilde{h}_t(\theta)|\left|\frac{1}{(\tilde{\lambda}_{t,1}(\theta_1)-\varphi)(\tilde{\lambda}_{t,2}(\theta))}-\frac{1}{(\lambda_{t,1}(\theta_1)-\varphi)(\lambda_{t,2}(\theta))}\right|$$
$$+\sup_{\theta\in\Theta}\left|\tilde{h}_t(\theta)-h_t(\theta)\right|\frac{1}{(\lambda_{t,1}(\theta_1)-\varphi)(\lambda_{t,2}(\theta_2)-\varphi)}$$
$$\le\frac{4\|\mathbf{Y}_t\|^2V}{\epsilon^3}\rho^t+\frac{6\|\mathbf{Y}_t\|^2}{\epsilon^2}\left|\exp\left(\frac{2\|\mathbf{Y}_t\|V}{\epsilon}\rho^t\right)-1\right|.$$

This asserts (4.20). Since it can be similarly shown that as $t\to\infty$,

$$\sup_{\theta\in\Theta}|\tilde{F}_{t,ij}(\theta)-F_{t,ij}(\theta)|\to 0\quad\text{a.s.}\quad\text{for } i,j=1,2,3,$$

the lemma is validated. □

Proof of Lemma 4.3 Since

$$E\sup_{\theta\in\Theta}\left\|\frac{\partial^2\ell_t(\theta)}{\partial\theta\partial\theta^T}\right\|_p\le E\sup_{\theta\in\Theta}\|\boldsymbol{F}_t(\theta)\|_p+E\sup_{\theta\in\Theta}\|\boldsymbol{D}_t(\theta)\|_p\cdot\left\|\frac{\partial\Lambda_t(\theta)}{\partial\theta}\right\|_p,$$

to prove the first inequality in the lemma, it suffices to show that for $i,j=1,2$,

$$E\sup_{\theta\in\Theta}\left\|F_{t,ij}(\theta)\frac{\partial\lambda_{t,i}(\theta_i)}{\partial\theta_i}\frac{\partial\lambda_{t,j}(\theta_j)}{\partial\theta_j^T}\right\|_p<\infty\quad\text{and}\quad E\sup_{\theta\in\Theta}\left\|D_{t,i}(\theta)\frac{\partial^2\lambda_{t,i}(\theta_i)}{\partial\theta_i\partial\theta_i^T}\right\|_p<\infty.$$

which, however, can be readily shown by using the Cauchy-Schwarz inequality and Lemmas 4.1 and 4.2, that is,

$$\begin{aligned} E\sup_{\theta\in\Theta}\left\|F_{t,ij}(\theta)\frac{\partial\lambda_{t,i}(\theta_i)}{\partial\theta_i}\frac{\partial\lambda_{t,j}(\theta_j)}{\partial\theta_j^T}\right\|_p &\le E\sup_{\theta\in\Theta}|F_{t,ij}|\cdot\sup_{\theta\in\Theta}\left\|\frac{\partial\lambda_{t,i}(\theta_i)}{\partial\theta_i}\frac{\partial\lambda_{t,j}(\theta_j)}{\partial\theta_j^T}\right\|_p \\ &\le C\left(E\|\mathbf{Y}_t\|^4\right)^{1/2}\cdot\left(E\sup_{\theta\in\Theta}\left\|\frac{\partial\lambda_{t,i}(\theta_i)}{\partial\theta_i}\frac{\partial\lambda_{t,j}(\theta_j)}{\partial\theta_j^T}\right\|_p^2\right)^{1/2}<\infty \end{aligned}$$

and

$$\begin{aligned} E\sup_{\theta\in\Theta}\left\|D_{t,i}(\theta)\frac{\partial^2\lambda_{t,i}(\theta_i)}{\partial\theta_i\partial\theta_i^T}\right\|_p &\le E\sup_{\theta\in\Theta}|D_{t,i}(\theta)|\cdot\sup_{\theta\in\Theta}\left\|\frac{\partial^2\lambda_{t,i}(\theta_i)}{\partial\theta_i\partial\theta_i^T}\right\|_p \\ &\le \left\{E(C\|\mathbf{Y}_t\|+1)^2\right\}^{1/2}\cdot\left(E\sup_{\theta\in\Theta}\left\|\frac{\partial^2\lambda_{t,i}(\theta_i)}{\partial\theta_i\partial\theta_i^T}\right\|_p^2\right)^{1/2}<\infty. \end{aligned}$$

The second inequality of the lemma is similarly proved. □

Proof of Lemma 4.4 Note that

$$\begin{aligned} \left\|\frac{\partial\tilde{\ell}_t(\theta_0)}{\partial\theta}-\frac{\partial\ell_t(\theta_0)}{\partial\theta}\right\|_p &= \|\tilde{\boldsymbol{D}}_t(\theta_0)\tilde{\Lambda}_t(\theta_0)-\boldsymbol{D}_t(\theta_0)\Lambda_t(\theta_0)\|_p \\ &\le \|\tilde{\boldsymbol{D}}_t(\theta_0)\|_p\|\tilde{\Lambda}_t(\theta_0)-\Lambda_t(\theta_0)\|_p+\|\Lambda_t(\theta_0)\|_p\|\tilde{\boldsymbol{D}}_t(\theta_0)-\boldsymbol{D}_t(\theta_0)\|_p. \end{aligned}$$

Further, from Lemma 4.1 and (4.9),

$$\begin{aligned} \|\tilde{\Lambda}_t(\theta_0)-\Lambda_t(\theta_0)\|_p &\le \left\|\frac{\partial\tilde{\lambda}_{t,1}(\theta_1^0)}{\partial\theta_1}-\frac{\partial\lambda_{t,1}(\theta_1^0)}{\partial\theta_1}\right\|_p+\left\|\frac{\partial\tilde{\lambda}_{t,2}(\theta_2^0)}{\partial\theta_2}-\frac{\partial\lambda_{t,2}(\theta_2^0)}{\partial\theta_2}\right\|_p\le 2V\rho^t, \\ \|\tilde{\boldsymbol{D}}_t(\theta_0)-\mathbf{D}_t(\theta_0)\|_p &\le \sum_{i=1}^3|\tilde{D}_{t,i}(\theta_0)-D_{t,i}(\theta_0)| \\ &\le \frac{C\|\mathbf{Y}_t\|}{\epsilon^2}\left|\exp\left(\frac{2\|\mathbf{Y}_t\|V}{\epsilon}\rho^t\right)-1\right|+\frac{C(1+\|\mathbf{Y}_t\|)}{\epsilon^2}V\rho^t \qquad (4.21) \end{aligned}$$

for some positive constant C. Then, we have

$$\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^n\frac{\partial\tilde{\ell}_t(\theta_0)}{\partial\theta}-\frac{1}{\sqrt{n}}\sum_{t=1}^n\frac{\partial\ell_t(\theta_0)}{\partial\theta}\right\|_p\le\frac{1}{\sqrt{n}}\sum_{t=1}^n\left\|\frac{\partial\tilde{\ell}_t(\theta_0)}{\partial\theta}-\frac{\partial\ell_t(\theta_0)}{\partial\theta}\right\|_p$$

$$\leq \frac{1}{\sqrt{n}}\sum_{t=1}^{n} 2\|\tilde{\boldsymbol{D}}_t(\theta_0)\|_p V\rho^t + \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\|\Lambda_t(\theta_0)\|_p \frac{C(1+\|\mathbf{Y}_t\|)}{\epsilon^2} V\rho^t$$
$$+\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\|\Lambda_t(\theta_0)\|_p \frac{C\|\mathbf{Y}_t\|}{\epsilon^2}\left|\exp\left(\frac{2\|\mathbf{Y}_t\|V}{\epsilon}\rho^t\right)-1\right|.$$

The first and second terms of the above second inequality are negligible owing to Lemma 4.1 since $\|\tilde{\mathbf{D}}_t(\theta_0)\|_p \leq C(\|\mathbf{Y}_t\|+1)$. On the other hand, using the mean value theorem, we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\|\Lambda_t(\theta_0)\|_p \frac{C\|\mathbf{Y}_t\|}{\epsilon^2}\left|\exp\left(\frac{2\|\mathbf{Y}_t\|V}{\epsilon}\rho^t\right)-1\right|$$
$$\leq \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\|\Lambda_t(\theta_0)\|_p \frac{C\|\mathbf{Y}_t\|}{\epsilon^2}\exp\left(\frac{2\|\mathbf{Y}_t\|V}{\epsilon}\rho^t\right)\frac{2\|\mathbf{Y}_t\|V}{\epsilon}\rho^t$$
$$\leq \frac{2CV}{\epsilon^3\sqrt{n}}\exp\left(\frac{2V}{\epsilon}\sup_{t\geq 1}\|\mathbf{Y}_t\|\rho^t\right)\sum_{t=1}^{\infty}\|\Lambda_t(\theta_0)\|_p\|\mathbf{Y}_t\|^2\rho^t.$$

Since $E(\sum_{t=1}^{\infty}\|\Lambda_t(\theta_0)\|\|\mathbf{Y}_t\|^2\rho^t) < \infty$, we have $\sum_{t=1}^{\infty}\|\Lambda_t(\theta_0)\|\|\mathbf{Y}_t\|^2\rho^t < \infty$ *a.s.*. Furthermore, $\exp(2V\sup_{t\geq 1}\|\mathbf{Y}_t\|\rho^t/\epsilon)$ is a.s. finite (See (4.18).). Therefore, the lemma is validated. □

Proof of Lemma 4.5. It suffices to show that as $t\to\infty$,

$$\sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\ell}_t(\theta)}{\partial\theta\partial\theta^T}-\frac{\partial^2\ell_t(\theta)}{\partial\theta\partial\theta^T}\right\|_p \to 0 \quad \text{a.s..}$$

Note that

$$\sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\ell}_t(\theta)}{\partial\theta\partial\theta^T}-\frac{\partial^2\ell_t(\theta)}{\partial\theta\partial\theta^T}\right\|_p \leq \sup_{\theta\in\Theta}\left\|\tilde{\boldsymbol{F}}_t(\theta)-\boldsymbol{F}_t(\theta)\right\|_p + \sup_{\theta\in\Theta}\left\|\tilde{\boldsymbol{D}}_t(\theta)\frac{\partial\tilde{\boldsymbol{\Lambda}}_t(\theta)}{\partial\theta}-\boldsymbol{D}_t(\theta)\frac{\partial\boldsymbol{\Lambda}_t(\theta)}{\partial\theta}\right\|_p \tag{4.22}$$

Since

$$\left\|\tilde{F}_{t,11}(\theta)\frac{\partial\tilde{\lambda}_{t,1}(\theta_1)}{\partial\theta_1}\frac{\partial\tilde{\lambda}_{t,1}(\theta_1)}{\partial\theta_1^T}-F_{t,11}(\theta)\frac{\partial\lambda_{t,1}(\theta_1)}{\partial\theta_1}\frac{\partial\lambda_{t,1}(\theta_1)}{\partial\theta_1^T}\right\|_p \tag{4.23}$$

$$\le |\tilde{F}_{t,11}(\theta) - F_{t,11}(\theta)| \left\| \frac{\partial \tilde{\lambda}_{t,1}(\theta_1)}{\partial \theta_1} \frac{\partial \tilde{\lambda}_{t,1}(\theta_1)}{\partial \theta_1^T} \right\|_p$$
$$+ |F_{t,11}(\theta)| \left\| \frac{\partial \tilde{\lambda}_{t,1}(\theta_1)}{\partial \theta_1} \frac{\partial \tilde{\lambda}_{t,1}(\theta_1)}{\partial \theta_1^T} - \frac{\partial \lambda_{t,1}(\theta_1)}{\partial \theta_1} \frac{\partial \lambda_{t,1}(\theta_1)}{\partial \theta_1^T} \right\|_p,$$

using Lemmas 4.1 and 4.2, we can show that the supremum over θ of (4.23) goes to 0 a.s. as $t \to \infty$. Similarly, we can show that all components of $\sup_{\theta\in\Theta} \|\tilde{\boldsymbol{F}}_t(\theta) - \boldsymbol{F}_t(\theta)\|_p$ go to 0 a.s. as $t \to \infty$. This implies $\sup_{\theta\in\Theta} \|\tilde{\mathbf{F}}_t(\theta) - \boldsymbol{F}_t(\theta)\|_p \to 0$ a.s.. Meanwhile, the second term of the RHS of (4.22) is bounded by

$$\sup_{\theta\in\Theta} \|\tilde{\boldsymbol{D}}_t(\theta) - \boldsymbol{D}_t(\theta)\|_p \left\| \frac{\partial \boldsymbol{\lambda}_t(\theta)}{\partial \theta} \right\|_p + \sup_{\theta\in\Theta} \|\tilde{\boldsymbol{D}}_t(\theta)\|_p \left\| \frac{\partial \tilde{\boldsymbol{\Lambda}}_t(\theta)}{\partial \theta} - \frac{\partial \boldsymbol{\Lambda}_t(\theta)}{\partial \theta} \right\|_p,$$

which goes to 0 a.s. as $t \to \infty$ by Lemmas 4.1 and 4.2 and (4.21), and therefore, the lemma is established. □

Proof of Lemma 4.6. See the proof of Proposition 5 of Lee et al. (2016a). □

Proof of Lemma 4.7 Since

$$E\{\partial \ell_t(\theta_0)/\partial \theta | \mathcal{F}_{t-1}\} = E\{\boldsymbol{D}_t(\theta)|\mathcal{F}_{t-1}\}\Lambda_t(\theta) \quad \text{a.s.},$$

to verify $E\{\partial \ell_t(\theta_0)/\partial \theta | \mathcal{F}_{t-1}\} = 0$ a.s., it suffices to show that $E\{\boldsymbol{D}_t(\theta)|\mathcal{F}_{t-1}\} = 0$ a.s.. We claim that

$$E\{D_{t,1}(\theta)|\mathcal{F}_{t-1}\} = 0 \quad \text{a.s..} \tag{4.24}$$

Since

$$E\{D_{t,1}(\theta)|\mathcal{F}_{t-1}\} = \frac{\lambda_{t,1}(\theta_1)}{\lambda_{t,1}(\theta_1) - \varphi} - 1 - E\left[\frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)} \Big| \mathcal{F}_{t-1} \right] \cdot \frac{1}{\lambda_{t,1}(\theta_1) - \varphi},$$

we only have to show that

$$E\left[\frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)}\Big|\mathcal{F}_{t-1}\right] = \varphi \quad \text{a.s..}$$

Here, for notational simplicity, we set $\lambda_{t,i} = \lambda_{t,i}(\theta_i)$ for $i = 1, 2$. Then, we have

$$\begin{aligned}
E\left[\frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)}\Big|\mathcal{F}_{t-1}\right] &= \sum_{n_1,n_2} P(Y_{t,1} = n_1, Y_{t,2} = n_2|\mathcal{F}_{t-1})\frac{g((n_1, n_2)^T, \boldsymbol{\lambda}_t, \varphi, 1)}{g((n_1, n_2)^T, \boldsymbol{\lambda}_t, \varphi, 0)} \\
= \sum_{n_1,n_2\geq 0} e^{-(\lambda_{t,1}+\lambda_{t,2}-\varphi)}\frac{(\lambda_{t,1} - \varphi)^{n_1}}{n_1!}\frac{(\lambda_{t,2} - \varphi)^{n_2}}{n_2!}\sum_{s=0}^{n_1\wedge n_2}\binom{n_1}{s}\binom{n_2}{s}s!sf(\boldsymbol{\lambda}_t, \varphi)^s \\
= \sum_{n_1,n_2\geq 1} e^{-(\lambda_{t,1}+\lambda_{t,2}-\varphi)}\frac{(\lambda_{t,1} - \varphi)^{n_1}}{n_1!}\frac{(\lambda_{t,2} - \varphi)^{n_2}}{n_2!}\sum_{s=1}^{n_1\wedge n_2}\binom{n_1}{s}\binom{n_2}{s}s!sf(\boldsymbol{\lambda}_t, \varphi)^s \\
= \sum_{n_1,n_2\geq 1} \varphi e^{-(\lambda_{t,1}+\lambda_{t,2}-\varphi)}\frac{(\lambda_{t,1} - \varphi)^{n_1-1}}{(n_1-1)!}\frac{(\lambda_{t,2} - \varphi)^{n_2-1}}{(n_2-1)!} \\
\times \sum_{s=1}^{n_1\wedge n_2}\binom{n_1-1}{s-1}\binom{n_2-1}{s-1}(s-1)!f(\boldsymbol{\lambda}_t, \varphi)^{s-1},
\end{aligned}$$

where we have used the fact that $\binom{n}{s} = \binom{n-1}{s-1}\frac{n}{s}$. Therefore, putting $n_1' = n_1 - 1, n_2' = n_2 - 1$ and $s' = s - 1$, we get

$$\begin{aligned}
&E\left[\frac{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 1)}{g(\mathbf{Y}_t, \boldsymbol{\lambda}_t, \varphi, 0)}\Big|\mathcal{F}_{t-1}\right] \\
&= \varphi\sum_{n_1',n_2'\geq 0} e^{-(\lambda_{t,1}+\lambda_{t,2}-\varphi)}\frac{(\lambda_{t,1} - \varphi)^{n_1'}}{n_1'!}\frac{(\lambda_{t,2} - \varphi)^{n_2'}}{n_2'!}\sum_{s'=0}^{n_1'\wedge n_2'}\binom{n_1'}{s'}\binom{n_2'}{s'}s'!f(\boldsymbol{\lambda}_t, \varphi)^{s'} \\
&= \varphi,
\end{aligned}$$

which implies (4.24). Since the same can be proven similarly for $D_{t,2}$ and $D_{t,3}$, the lemma is established. □

Table 4.1: Empirical sizes of $T_n^{est,1}$, T_n^{res} and $T_n^{est,2}$ at the nominal level 0.05.

	$n = 300$			$n = 500$			$n = 1000$		
φ	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$
$\boldsymbol{\theta} = (3, 0.2, 0.1, 0, 1, 0.1, 0, 0.2, \varphi)$									
0	0.084	0.036	0.044	0.092	0.046	0.044	0.050	0.030	0.062
0.3	0.092	0.030	0.108	0.084	0.036	0.066	0.084	0.048	0.072
0.7	0.122	0.044	0.292	0.094	0.034	0.190	0.036	0.060	0.172
$\boldsymbol{\theta} = (3, 0.4, 0.1, 0, 1, 0.3, 0, 0.2, \varphi)$									
0	0.128	0.036	0.066	0.146	0.050	0.066	0.100	0.052	0.064
0.3	0.148	0.040	0.120	0.144	0.038	0.080	0.138	0.054	0.082
0.7	0.158	0.032	0.132	0.132	0.054	0.100	0.122	0.044	0.080
$\boldsymbol{\theta} = (3, 0.2, 0.1, 0.2, 1, 0.1, 0.1, 0.2, \varphi)$									
0	0.098	0.036	0.076	0.086	0.042	0.074	0.068	0.050	0.076
0.3	0.116	0.024	0.088	0.094	0.042	0.078	0.084	0.062	0.098
0.7	0.138	0.052	0.214	0.116	0.044	0.160	0.104	0.040	0.178
$\boldsymbol{\theta} = (3, 0.4, 0.1, 0.2, 1, 0.3, 0.1, 0.2, \varphi)$									
0	0.188	0.048	0.092	0.176	0.040	0.092	0.168	0.054	0.098
0.3	0.170	0.044	0.088	0.178	0.042	0.100	0.164	0.044	0.114
0.7	0.192	0.022	0.082	0.196	0.060	0.140	0.166	0.048	0.112

Table 4.2: Empirical powers of $T_n^{est,1}$, T_n^{res} and $T_n^{est,2}$ at the nominal level 0.05 when **B** is a diagonal matrix and $\boldsymbol{\theta}_0 = (3, 0.2, 0.1, 0, 1, 0.1, 0, 0.2, \varphi)$ changes to $\boldsymbol{\theta}'$ at $t = [n\tau]$.

		$n=300$			$n=500$			$n=1000$		
τ	φ	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$
$(\delta_1, \delta_2) = (3, 1) \to (\delta_1', \delta_2') = (2.7, 1.5)$										
	0	0.664	0.762	0.690	0.850	0.970	0.882	0.994	1	1
1/3	0.3	0.704	0.782	0.730	0.874	0.976	0.926	0.994	1	1
	0.7	0.782	0.878	0.852	0.950	0.992	0.976	1	1	1
	0	0.794	0.892	0.798	0.988	0.998	0.992	1	1	1
1/2	0.3	0.814	0.904	0.848	0.994	1	1	1	1	1
	0.7	0.924	0.948	0.912	0.994	1	0.998	1	1	1
	0	0.706	0.812	0.716	0.876	0.966	0.884	0.998	1	1
2/3	0.3	0.760	0.824	0.748	0.932	0.986	0.948	0.996	1	0.998
	0.7	0.830	0.890	0.840	0.968	0.996	0.956	1	1	1
$(\alpha_1, \alpha_2) = (0.2, 0.1) \to (\alpha_1', \alpha_2') = (0.3, 0.2)$										
	0	0.434	0.348	0.304	0.614	0.648	0.526	0.830	0.950	0.860
1/3	0.3	0.450	0.304	0.356	0.544	0.602	0.564	0.854	0.920	0.862
	0.7	0.418	0.266	0.470	0.576	0.562	0.592	0.824	0.930	0.842
	0	0.466	0.456	0.390	0.702	0.778	0.618	0.934	0984	0.936
1/2	0.3	0.502	0.434	0.444	0.652	0.712	0.604	0.914	0.974	0.920
	0.7	0.542	0.424	0.524	0.656	0.674	0.654	0.912	0.908	0.920
	0	0.440	0.378	0.316	0.558	0.664	0.526	0.880	0.980	0.848
2/3	0.3	0.404	0.330	0.340	0.554	0.644	0.542	0.878	0.958	0.830
	0.7	0.430	0.366	0.470	0.564	0.570	0.612	0.868	0.936	0.832

Table 4.3: Empirical powers of $T_n^{est,1}$, T_n^{res} and $T_n^{est,2}$ at the nominal level 0.05 when $\mathbf{B}$ is a diagonal matrix and $\boldsymbol{\theta}_0 = (3, 0.2, 0.1, 0, 1, 0.1, 0, 0.2, \varphi)$ changes to $\boldsymbol{\theta}'$ at $t = [n\tau]$.

		$n = 300$			$n = 500$			$n = 1000$		
τ	φ	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$
$\varphi = 0 \to \varphi' = 0.3$										
1/3		0.100	0.034	0.074	0.080	0.046	0.046	0.086	0.056	0.056
1/2		0.088	0.034	0.050	0.084	0.036	0.056	0.082	0.030	0.078
2/3		0.112	0.036	0.056	0.088	0.050	0.048	0.072	0.044	0.054
$\varphi = 0 \to \varphi' = 0.7$										
1/3		0.166	0.052	0.148	0.288	0.044	0.210	0.664	0.046	0.626
1/2		0.120	0.030	0.100	0.234	0.030	0.240	0.734	0.046	0.774
2/3		0.150	0.038	0.118	0.154	0.058	0.188	0.388	0.028	0.480

Table 4.4: Empirical powers of $T_n^{est,1}$, T_n^{res} and $T_n^{est,2}$ at the nominal level 0.05 when $\mathbf{B}$ is a non-diagonal matrix and $\boldsymbol{\theta}_0 = (3, 0.2, 0.1, 0.2, 1, 0.1, 0.2, 0.1, \varphi)$ changes to $\boldsymbol{\theta}'$ at $t = [n\tau]$.

		$n = 300$			$n = 500$			$n = 1000$		
τ	φ	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$
$(\delta_1, \delta_2) = (3, 1) \to (\delta_1', \delta_2') = (2.7, 1.5)$										
	0	0.488	0.596	0.530	0.710	0.874	0.754	0.968	1	0.984
1/3	0.3	0.552	0.682	0.582	0.760	0.906	0.804	0.976	1	0.992
	0.7	0.566	0.692	0.658	0.790	0.934	0.858	0.986	1	1
	0	0.610	0.740	0.632	0.818	0.950	0.884	0.992	1	0.994
1/2	0.3	0.634	0.784	0.672	0.840	0.950	0.878	0.994	1	1
	0.7	0.662	0.822	0.770	0.890	0.996	0.928	0.998	1	1
	0	0.532	0.620	0.546	0.738	0.898	0.804	0.976	1	0.976
2/3	0.3	0.574	0.648	0.602	0.786	0.914	0.806	0.990	1	0.996
	0.7	0.646	0.718	0.616	0.820	0.946	0.858	1	1	0.998
$(\alpha_1, \alpha_2) = (0.2, 0.1) \to (\alpha_1', \alpha_2') = (0.3, 0.2)$										
	0	0.520	0.346	0.460	0.750	0.722	0.692	0.956	0.992	0.952
1/3	0.3	0.598	0.358	0.526	0.756	0.686	0.714	0.948	0.974	0.944
	0.7	0.516	0.348	0.516	0.698	0.662	0.712	0.940	0.960	0.930
	0	0.652	0.508	0.570	0.832	0.838	0.784	0.990	0.998	0.992
1/2	0.3	0.638	0.518	0.580	0.848	0.848	0.808	0.984	0.996	0.982
	0.7	0.606	0.464	0.584	0.830	0.816	0.830	0.978	0.992	0.978
	0	0.554	0.428	0.450	0.778	0.772	0.676	0.970	0.996	0.954
1/3	0.3	0.570	0.460	0.536	0.744	0.758	0.716	0.968	0.990	0.944
	0.7	0.520	0.402	0.566	0.730	0.702	0.676	0.948	0.978	0.930

Table 4.5: Empirical powers of $T_n^{est,1}$, T_n^{res} and $T_n^{est,2}$ at the nominal level 0.05 when $\mathbf{B}$ is a non-diagonal matrix and $\boldsymbol{\theta}_0 = (3, 0.2, 0.1, 0.2, 1, 0.1, 0.2, 0.1, \varphi)$ changes to $\boldsymbol{\theta}'$ at $t = [n\tau]$.

		$n = 300$			$n = 500$			$n = 1000$		
τ	φ	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$	$T_n^{est,1}$	T_n^{res}	$T_n^{est,2}$
$\varphi = 0 \to \varphi' = 0.3$										
1/3		0.096	0.036	0.068	0.116	0.032	0.074	0.106	0.036	0.102
1/2		0.096	0.040	0.052	0.090	0.030	0.052	0.106	0.062	0.106
2/3		0.096	0.026	0.060	0.102	0.040	0.066	0.074	0.052	0.106
$\varphi = 0 \to \varphi' = 0.7$										
1/3		0.148	0.038	0.116	0.200	0.052	0.154	0.404	0.056	0.408
1/2		0.170	0.058	0.116	0.196	0.036	0.178	0.420	0.046	0.483
2/3		0.116	0.052	0.074	0.170	0.048	0.128	0.242	0.028	0.306

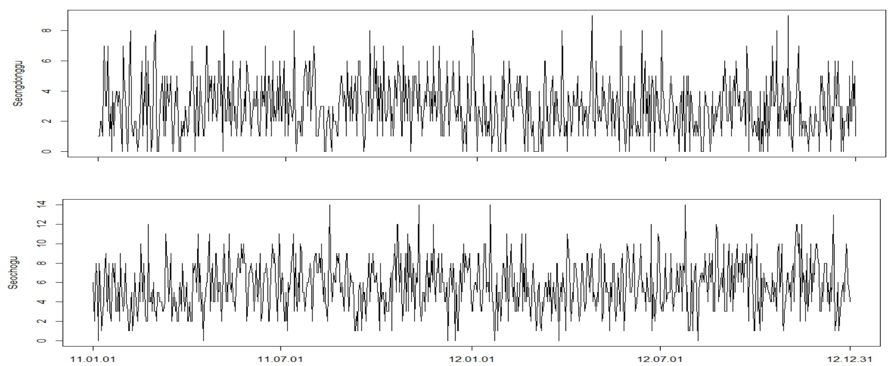

Figure 4.1: Plot of counts series for the traffic accidents of Seongdong and Seocho counties.

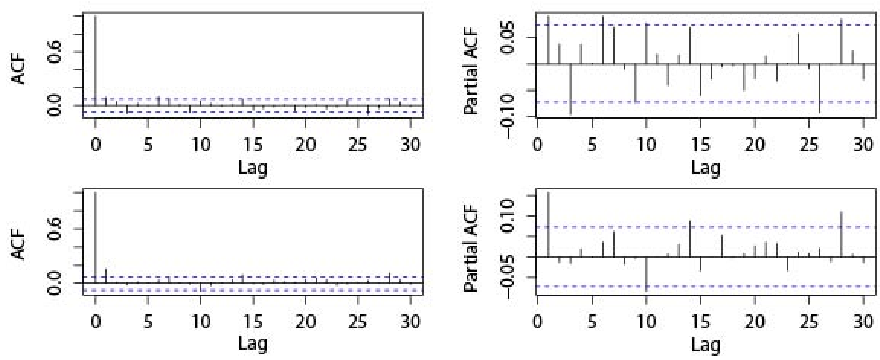

Figure 4.2: Plot of the sample autocorrealation and the sample partial autocorrealation traffic accidents of Seongdong (upper) and Seocho (lower) counties.

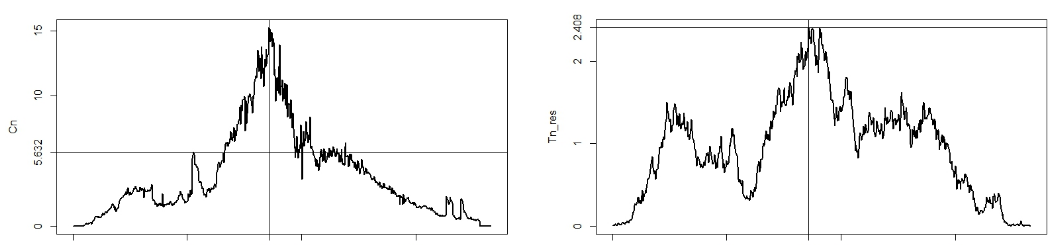

Figure 4.3: Plot of Plot of $T_n^{est,2}$ and T_n^{res} for the traffic accidents of Seongdong and Seocho counties.

Table 4.6: Estimated parameters for the counts of traffic accidents of Seongdong (Y_1) and Seocho (Y_2) counties in Seoul, Korea, based on a Bivariate Poisson INGARH(1,1) model. Standard errors are shown in parentheses.

	Mean	Variance	$\hat{\delta}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\varphi}$
Full Data							
Y_1	2.927	3.574	2.117	0.181	0.095	0.001	0.529
			(0.629)	(0.212)	(0.034)	(0.023)	(0.126)
Y_2	5.661	7.164	4.704	0.001	0.001	0.167	
			(0.993)	(0.176)	(0.048)	(0.035)	
First period : Jan.01.2011-Dec.08.2011							
Y_1	3.164	3.645	2.843	0.009	0.092	0.001	0.489
			(0.928)	(0.281)	(0.052)	(0.037)	(0.192)
Y_2	5.594	7.164	4.501	0.008	0.001	0.186	
			(2.494)	(0.454)	(0.070)	(0.052)	
Second period : Dec.09.2011 - Dec.31.2012							
Y_1	2.720	3.429	1.556	0.329	0.062	0.017	0.638
			(0.935)	(0.364)	(0.046)	(0.028)	(0.168)
Y_2	5.720	7.399	4.882	0.001	0.035	0.143	
			(1.169)	(0.203)	(0.068)	(0.048)	

Chapter 5

Comparison Study on CUSUM Tests for General Nonlinear Integer-valued GARCH Models

5.1 Introduction

In this study, special attention is paid to comparing the performance of the score vector- and (standardized) residual-, and estimates-based CUSUM tests empirically for GN-INGARCH models. For this task, however, we make an effort to derive their limiting null distributions to obtain the critical values, used for Monte Carlo simulations. Our findings show that the standardized residual-based test performs the best among the CUSUM tests under consideration.

This paper is organized as follows. Section 5.2 introduces the one-parameter exponential family AR models and establishes the asymptotic results for the CMLE. Section 5.3 introduces the CUSUM tests based on score vectors, (standardized) residuals, and estimates, and then derives their limiting null distributions. Sections 5.4 implements a simulation study for comparison and we analyze real data in Section 5.5. Section 5.6 provides concluding remarks. Finally, all the proofs are provided in Section 5.7.

5.2 Models and likelihood inferences

5.2.1 Basic set-up and asymptotics

Let $\{Y_t, t \geq 1\}$ be the GN-INGARCH time series of counts with the conditional distribution of the one-parameter exponential family:

$$Y_t|\mathcal{F}_{t-1} \sim p(y|\eta_t), \;\; X_t := E(Y_t|\mathcal{F}_{t-1}) = f_\theta(X_{t-1}, Y_{t-1}), \tag{5.1}$$

where $\mathcal{F}_t$ is the σ-field generated by $\eta_1, Y_1, \ldots, Y_t$, and $f_\theta(x, y)$ is a nonnegative bivariate function defined $[0, \infty) \times \mathbb{N}_0$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, depending on the parameter $\theta \in \Theta \subset \mathbb{R}^d$. Here, $p(\cdot|\cdot)$ is a probability mass function given by

$$p(y|\eta) = \exp\{\eta y - A(\eta)\}h(y), \quad y \geq 0,$$

where η is the natural parameter and $A(\eta)$ and $h(y)$ are known functions. If $B(\eta) = A'(\eta)$, $B(\eta_t(\theta_0))$ and $B'(\eta_t(\theta_0))$ are, then, the conditional mean and variance of Y_t, respectively.

In what follows, we assume

(**C0**) For all $x, x' \geq 0$ and $y, y' \in \mathbb{N}_0$ where $\omega_1, \omega_2 \geq 0$ satisfying $\omega_1 + \omega_2 < 1$,

$$\sup_{\theta\in\Theta} |f_\theta(x, y) - f_\theta(x', y')| \leq \omega_1|x - x'| + \omega_2|y - y'|.$$

Davis and Liu (2016) shows that this assumption ensures the strict stationarity and ergodicity of $\{(X_t, Y_t)\}$. The conditional likelihood function of model (5.1), based on the observation $Y_1, \ldots, Y_n$, is given by

$$\widetilde{\mathcal{L}}(\theta|Y_1, \ldots, Y_n, \tilde{\eta}_1) = \prod_{t=1}^{n} \exp\{\tilde{\eta}_t(\theta)Y_t - A(\tilde{\eta}_t(\theta))\}h(Y_t),$$

where $\tilde{\eta}_t(\theta) = B^{-1}(\widetilde{X}_t(\theta))$ is recursively updated through the equations:

$$\widetilde{X}_t(\theta) = f_\theta(\widetilde{X}_{t-1}(\theta), Y_{t-1}),$$

with an arbitrarily chosen initial random variable $\widetilde{X}_1$. In what follows, θ_0 denotes the true value of θ. We obtain the CMLE of θ_0 by

$$\hat{\theta}_n = \arg\max_{\theta\in\Theta} \widetilde{\mathcal{L}}(\theta) = \arg\max_{\theta\in\Theta} \widetilde{L}_n(\theta) = \arg\max_{\theta\in\Theta} \sum_{t=1}^{n} \tilde{\ell}_t(\theta),$$

where $\tilde{\ell}_t(\theta) = \log p(Y_t|\tilde{\eta}_t(\theta)) = \tilde{\eta}_t(\theta)Y_t - A(\tilde{\eta}_t(\theta))$.

To ensure the strong consistency and asymptotic normality of the CMLE, we impose some regularity conditions, wherein V and $\rho \in (0, 1)$ stand for a generic integrable random variable and constant, respectively; symbol $\|\cdot\|$ denotes the L^1 norm for matrices and vectors; and $E(\cdot)$ is taken under θ_0. Further, we use the notation $\eta_t = \eta_t(\theta)$ and $\tilde{\eta}_t = \tilde{\eta}_t(\theta)$ for simplicity.

(**C1**) θ_0 is an interior point in the compact parameter space $\Theta \in \mathbb{R}^d$

(**C2**) For any $\theta \in \Theta$, $f_\infty^\theta \geq x_\theta^* \in \mathcal{R}(B)$, where $\mathcal{R}(B)$ is the range of $B(\eta)$. Moreover, $x_\theta^* \geq x^* \in \mathcal{R}(B)$ for all θ.

(**C3**) For any $\mathbf{y} \in [0, \infty)^\infty$ or $\mathbb{N}_0^\infty$, the mapping $\theta \mapsto f_\infty^\theta(\mathbf{y})$ is continuous.

(**C4**) $f(x, y)$ is increasing in (x, y) if Y_t given $\mathcal{F}_{t-1}$ has a continuous distribution.

(**C5**) $E\left\{Y_1 \sup_{\theta\in\Theta} B^{-1}(f_\infty^\theta(Y_0, Y_{-1}, \ldots))\right\} < \infty$

(**C6**) If there exists a $t > 1$ such that $X_t(\theta) = X_t(\theta_0)$ a.s., then $\theta = \theta_0$.

(**C7**) $E\left(\sup_{\theta\in\Theta} X_1(\theta)\right) < \infty$ and $E\left(\sup_{\theta\in\Theta} \tilde{X}_1(\theta)\right) < \infty$.

(**C8**) The mapping $\theta \mapsto f_\infty^\theta$ is twice continuously differentiable with respect to θ and satisfies

$$E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial f_\infty^\theta(Y_{t-1}, Y_{t-2}, \ldots)}{\partial\theta}\right\|\right)^2 < \infty,\ E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial^2 f_\infty^\theta(Y_{t-1}, Y_{t-2}, \ldots)}{\partial\theta\partial\theta^T}\right\|\right)^2 < \infty.$$

(**C9**) For all t,

$$E\left[\sup_{\theta\in\Theta}\left\|B'(\eta_t)\left(\frac{\partial\eta_t}{\partial\theta}\cdot\frac{\partial\eta_t}{\partial\theta^T}\right)\right\|\right] < \infty,\ E\left[\sup_{\theta\in\Theta}\left\|(Y_t - B(\eta_t))\frac{\partial^2\eta_t}{\partial\theta\partial\theta^T}\right\|\right] < \infty.$$

(**C10**) For all t, a.s.,

$$\sup_{\theta\in\Theta}\left\|\frac{\partial\tilde{X}_t(\theta)}{\partial\theta} - \frac{\partial X_t(\theta)}{\partial\theta}\right\| \le V\rho^t \ \text{ and } \ \sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\eta}_t(\theta)}{\partial\theta\partial\theta^T} - \frac{\partial^2\eta_t}{\partial\theta\partial\theta^T}\right\| \le V\rho^t,$$

(**C11**) For some constant $\underline{c} > 0$, $\sup_{\theta\in\Theta}\sup_{0\ge\delta\ge1} B'((1-\delta)\eta_t + \delta\tilde{\eta}_t) \ge \underline{c}$, for all t.

(**C12**) For all t, a.s., $\sup_{\theta\in\Theta}|B'(\tilde{\eta}_t) - B'(\eta_t)| \le V\rho^t$.

(**C13**) For some constant $K > 0$, $\sup_{\theta\in\Theta} B'(\eta_t)^{-3/2}B''(\eta_t) \le K$, for all t.

Conditions (**C1**)–(**C9**) can be found in Davis and Liu (2016). They also derive the asymptotic properties of the CMLE. The proposition below can be proven using Lemma 5.2 in Section 5.7, in a manner similar to that seen with their Theorems 1 and 2. Although the definition of our CMLE is similar to theirs, a subtle difference exists in the condition and proof, because we are taking the approach of Francq and Zakoïan (2004).

Proposition 5.1. *Suppose that conditions* (**C0**)–(**C13**) *hold. Then, as* $n \to \infty$,

$$\hat{\theta}_n \longrightarrow \theta_0 \ \ a.s.,$$

and

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{w} N(0, I(\theta_0)^{-1}),$$

where

$$I(\theta_0) = E\left(\frac{\partial \ell_t(\theta_0)}{\partial \theta}\frac{\partial \ell_t(\theta_0)}{\partial \theta^T}\right) = -E\left(\frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta^T}\right)$$

and $\ell_t(\theta_0) = \eta_t(\theta_0)Y_t - A(\eta_t(\theta))$.

5.2.2 INGARCH(1,1) models

In this subsection, we focus on the INGARCH(1,1) model:

$$Y_t|\mathcal{F}_{t-1} \sim p(y|\eta_t), \ \ X_t = \omega + \alpha X_{t-1} + \beta Y_{t-1}, \tag{5.2}$$

where $X_t = B(\eta_t) = E(Y_t|\mathcal{F}_{t-1})$ and $\theta = (\omega, \alpha, \beta)$ satisfy $\omega > 0, \alpha \geq 0, \beta \geq 0$ and $\alpha + \beta < 1$. The process $\{(X_t, Y_t); t \geq 1\}$ has then a strictly stationary and ergodic solution. To ensure Proposition 5.1 in this case, (**C1**) can be replaced with the following:

(**C1'**) The true parameter θ_0 lies in a compact neighborhood $\Theta \in \mathbb{R}_+^3$ of θ_0, where

$$\Theta \in \{\theta = (\omega, \alpha, \beta)^T \in \mathbb{R}_+^3 : 0 < \omega_L \leq \omega \leq \omega_U, \epsilon \leq \alpha + \beta \leq 1 - \epsilon\} \text{ for some } \epsilon > 0.$$

Note that, by iterating (5.2),

$$X_t(\theta) = \frac{\omega}{1-\alpha} + \beta \sum_{k=0}^{\infty} \alpha^k Y_{t-k-1}, \ \ \tilde{X}_t(\theta) = \frac{\omega}{1-\alpha} + \beta \sum_{k=0}^{t-2} \alpha^k Y_{t-k-1},$$

where the initial value is taken as $\tilde{X}_1(\theta) = \omega/(1-\alpha)$. Hence, (**C2**) is satisfied, because $X_t(\theta) \geq \omega/(1-\alpha) \geq x^* = \omega_L/(1-\epsilon)$. Below, we summarize some of the most typical examples wherein Conditions (**C11**)–(**C13**) are found to hold—namely, Poisson, negative binomial, and binomial distributions. For (**C3**)–(**C10**), see Kang and Lee (2014), Davis and Liu (2016), and Diop and Kengne (2017).

Example 5.1 (Poisson INGARCH(1,1) model)**.** The Poisson INGARCH(1,1) model is given by

$$Y_t|\mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t), \ \ \lambda_t = \omega + \alpha\lambda_{t-1} + \beta Y_{t-1}.$$

In this model, $\eta_t = \log(X_t(\theta))$ and $A(\eta_t) = e^{\eta_t(\theta)}$. Since $X_t(\theta) \geq \omega_L$, $\tilde{X}_t(\theta) \geq \omega_L$, and $B'(\eta) = e^\eta$ is increasing, (**C11**) holds. Moreover, since $B'(\eta_t) = X_t(\theta)$, (**C12**) is satisfied. Finally, (**C13**) holds, due to (**C11**) and the fact that $B'(\eta) = B''(\eta)$.

Example 5.2 (NB-INGARCH(1,1) model)**.** The NB-INGARCH(1,1) model is defined as

$$Y_t|\mathcal{F}_{t-1} \sim \text{NB}(r, p_t), \quad X_t = \frac{r(1-p_t)}{p_t} = \omega + \alpha X_{t-1} + \beta Y_{t-1},$$

where $r \in \mathbb{N}$ and $Y \sim \text{NB}(r,p)$ denotes the negative binomial distribution, with the probability mass function given by

$$P(Y_t = k) = \left(\begin{array}{c} k+r-1 \\ r-1 \end{array} \right) (1-p)^k p^r, \qquad k = 0, 1, 2, \ldots.$$

Here, r is assumed to be known. In this model, $\eta_t = \log(X_t(\theta)/(X_t(\theta)+r)$ and $A(\eta_t) = -r\log(r/(1-e^{\eta_t}))$. Since $X_t(\theta) \geq \omega_L$, $\tilde{X}_t(\theta) \geq \omega_L$, and $B'(\eta) = re^\eta/(1-e^\eta)^2$ is increasing, (**C11**) holds. Next, since $B'(\eta_t) = X_t(\theta)(X_t(\theta)+r)/r$,

$$|B'(\tilde{\eta}_t) - B'(\eta_t)| \leq \left(\tilde{X}_t(\theta) + X_t(\theta) + 1\right) |\tilde{X}_t(\theta) - X_t(\theta)|/r \leq V\rho^t,$$

owing to (**C7**) and Lemma 5.1 in Section 5.7, which in turn implies (**C12**). Finally, (**C13**) is established, owing to the fact that $\log \omega_L/(\omega_L + r) \leq \eta_t < 1$ and $B''(\eta) = re^{\eta}(1+e^{\eta})/(1-e^{\eta})^3$.

Example 5.3 (Binomial INGARCH(1,1) model)**.** The binomial INGARCH(1,1) model is given by

$$Y_t|\mathcal{F}_{t-1} \sim \mathrm{B}(m, p_t), \quad X_t = mp_t = \omega + \alpha X_{t-1} + \beta Y_{t-1},$$

where $\omega > 0, \alpha \geq 0, \beta \geq 0$ and $\omega + \alpha m + \beta m \leq m$ are assumed to ensure $p_t \in (0,1)$. When $m = 1$, the model is considered a Bernoulli INGARCH(1,1) model. In this case, since $p_t \in (0,1)$, the parameter space becomes

$$\Theta = \left\{(\omega, \alpha, \beta)^T : 0 < \omega_L \leq \omega \leq \omega_U,\ \epsilon \leq \alpha + \beta \leq 1 - \epsilon\right\} \text{ for some } \epsilon > \omega_U/m.$$

In particular, for the Bernoulli INGARCH(1,1) model,

$$\Theta = \left\{\theta = (\omega, \alpha, \beta)^T \in \mathbb{R}_+^3 : \epsilon \leq \omega + \alpha + \beta \leq 1 - \epsilon\right\} \quad \text{for some} \quad 0 < \epsilon < 1.$$

Note that $\eta_t = \log\left(X_t(\theta)/(m - X_t(\theta))\right)$ and $A(\eta) = m \log(1 + e^{\eta})$. Since $p_t \in (0,1)$, (**C11**) and (**C13**) hold; furthermore, given the fact that $B'(\eta_t) = X_t(\theta)\left(1 - X_t(\theta)/m\right)$, it can be shown that (**C12**) holds, similar to the case with the NB-INGARCH(1,1) model.

5.3 Change point test

In this section, we introduce the score vector-, residual-, standardized residual-, and estimates-based CUSUM tests used to assess the hypotheses:

$$H_0 : \theta \text{ does not change over } Y_1, \cdots, Y_n \ \ vs. \ \ H_1 : \text{not } H_0.$$

The asymptotic results of these CUSUM test are proved in Section 5.7.

5.3.1 Score vector-based CUSUM test

The score vector-based CUSUM test is given by:

$$T_n^{score} = \max_{1\le k\le n} \frac{1}{n}\left(\sum_{t=1}^{k} \frac{\partial \tilde{\ell}_t(\hat{\theta}_n)}{\partial\theta}\right)^T \hat{I}_n^{-1}\left(\sum_{t=1}^{k} \frac{\partial \tilde{\ell}_t(\hat{\theta}_n)}{\partial\theta}\right),$$

where

$$\hat{I}_n = -\frac{1}{n}\sum_{t=1}^{n} \frac{\partial^2 \tilde{\ell}_t(\hat{\theta}_n)}{\partial\theta\partial\theta^T},$$

is a consistent estimator of $I(\theta_0)$. Then, we obtain the following: see the proof in Section 5.7.

Theorem 5.1. *Suppose that conditions* **(C0)**–**(C13)** *hold. Then, under* H_0*, as* $n\to\infty$*,*

$$T_n^{score} \xrightarrow{w} \sup_{0\le s\le 1} \|\mathbf{B}_d^\circ(s)\|^2, \tag{5.3}$$

where $\{\mathbf{B}_d^\circ(s), 0<s<1\}$ *is a* $d-$*dimensional Brownian bridge.*

5.3.2 Residual-based CUSUM test

We consider the two types of residuals:

$$\epsilon_{t,1} = Y_t - X_t(\theta_0) \ \text{ and } \ \epsilon_{t,2} = (\mathrm{Y_t} - \mathrm{X_t}(\theta_0))/\sqrt{\mathrm{B}'(\eta_t(\theta_0))}.$$

The former is considered by Franke et al. (2012), Kang and Lee (2014), and Lee et al. (2016a,b) in some Poisson AR models, whereas the latter is newly considered here. Since $\{\epsilon_{t,i}, \mathcal{F}_t\}$, $i=1,2$, are stationary ergodic martingale difference sequences, using a functional central limit theorem, we can derive

$$\sup_{0<s<1} \frac{1}{\sqrt{n}\tau_i}\left|\sum_{t=1}^{[ns]} \epsilon_{t,i} - \frac{k}{n}\sum_{t=1}^{n}\epsilon_{t,i}\right| \xrightarrow{w} \sup_{0\le s\le 1} |\mathbf{B}_1^\circ(s)|, \tag{5.4}$$

where $\tau_i^2 = Var(\epsilon_{1,i})$ (actually $\tau_2^2 = 1$). However, since $\epsilon_{t,i}$ is not observable, we consider the tests:

$$T_n^{res,i} = \max_{1\le k\le n} \frac{1}{\sqrt{n}\hat{\tau}_{n,i}} \left| \sum_{t=1}^{k} \hat{\epsilon}_{t,i} - \frac{k}{n}\sum_{t=1}^{n} \hat{\epsilon}_{t,i} \right|,$$

where, for $t \ge 2$, $\hat{\epsilon}_{t,1} = Y_t - \hat{X}_t$; $\hat{\epsilon}_{t,2} = (Y_t - \hat{X}_t)/\sqrt{B'(\hat{\eta}_t)}$; $\hat{X}_t = f_{\hat{\theta}_n}(\hat{X}_{t-1}, Y_{t-1})$, $\hat{\eta}_t = B^{-1}(\hat{X}_t)$; $\hat{X}_1$ is an arbitrarily chosen initial variable; and $\hat{\tau}_{n,1}^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{\epsilon}_{t,1}^2$ and $\hat{\tau}_{n,2}^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{\epsilon}_{t,2}^2$. Then, we can obtain the following, the proof of which is similar to that of Kang and Lee (2014) and is omitted for brevity.

Theorem 5.2. *Suppose that conditions* (**C0**)–(**C13**) *hold. Then, under* H_0*, as* $n \to \infty$,

$$T_n^{res,1} \xrightarrow{w} \sup_{0\le s\le 1} \|\mathbf{B}_1^{\circ}(s)\|.$$

Moreover, using (5.4), we can obtain the following: see the proof in Section 5.7.

Theorem 5.3. *Suppose that conditions* (**C0**)–(**C13**) *hold. Then, under* H_0*, as* $n \to \infty$,

$$T_n^{res,2} \xrightarrow{w} \sup_{0\le s\le 1} \|\mathbf{B}_1^{\circ}(s)\|.$$

In our simulations study, the two following estimates-based CUSUM tests are compared to score vector- and (standardized) residual-based CUSUM tests:

$$T_n^{est,1} = \max_{1\le k\le n} \frac{k^2}{n}(\hat{\theta}_k - \hat{\theta}_n)^T \hat{I}_n (\hat{\theta}_k - \hat{\theta}_n), \tag{5.5}$$

where $\hat{\theta}_k$ is the CMLE of θ_0 based on $Y_1, \ldots, Y_k$, and

$$T_n^{est,2} = \max_{v_n\le k\le n-v_n} \frac{k^2(n-k)^2}{n^3}(\hat{\theta}_k - \tilde{\theta}_k)^T \hat{I}'_n (\hat{\theta}_k - \tilde{\theta}_k),$$

where $\tilde{\theta}_k$ are the CMLE of θ_0 based on the observations $Y_{k+1}, \ldots, Y_n$,

$$\hat{I}'_n = -\frac{1}{2}\left[\frac{1}{u_n}\sum_{t=1}^{u_n} \frac{\partial^2 \tilde{\ell}_t(\hat{\theta}_{u_n})}{\partial\theta\partial\theta^T} + \frac{1}{n-u_n}\sum_{t=u_n+1}^{n} \frac{\partial^2 \tilde{\ell}_t(\tilde{\theta}_{u_n})}{\partial\theta\partial\theta^T} \right],$$

and $\{u_n : n \geq 1\}$ and $\{v_n : n \geq 1\}$ are sequences of integers diverging to ∞, such that $u_n/n, v_n/n \to 0$ as $n \to \infty$. Then, under **(C0)**–**(C13)** and H_0, $T_n^{est,i}$ converges weakly to $\sup_{0\leq s\leq 1} \|\mathbf{B}_d^{\circ}(s)\|$; for the latter, see Diop and Kengne (2017).

5.4 Simulation study

In this section, we report our simulation results and evaluate the performance of the tests proposed in Section 5.3. We consider the INGARCH(1,1) model in Section 5.2.2. In this simulation study, we employ the nominal level of 0.05, $n = 300, 500, 1000$, and 1,000 as the number of repetitions. The critical value for this nominal level is obtained through Monte Carlo simulations (cf. Lee et al. (2003)): for $T_n^{est,1}, T_n^{est,2}, T_n^{score}$, it is 3.004; for $T_n^{res,1}, T_n^{res,2}$, it is 1.353. The $T_n^{est,2}$ is calculated with $q \equiv 1$ and $u_n = v_n = [(\log n)^2]$. Since $\hat{\theta}_k$ is inaccurate for small k values, we use the test statistic:

$$T_n^{est,1} = \max_{k_L \leq k \leq n} \frac{k^2}{n} (\hat{\theta}_k - \hat{\theta}_n)^T \hat{I}_n (\hat{\theta}_k - \hat{\theta}_n),$$

with $k_L = 20$, instead of (5.5).

5.4.1 Test for Poisson INGARCH(1,1) models

We consider the Poisson INGARCH(1,1) model:

$$Y_t|\mathcal{F}_{t-1} \sim \text{Poisson}(X_t), \quad X_t = \omega + \alpha X_{t-1} + \beta Y_{t-1},$$

where X_1 is set to be 0. To calculate empirical size, we consider the parameters $\omega = 1, 0.3$ and $(\alpha, \beta) = (0.1, 0.3), (0.1, 0.5), (0.1, 0.8), (0.3, 0.2), (0.3, 0.4), (0.4, 0.5)$. The empirical sizes are listed in Table 5.1. As pointed out in Kang and Lee (2014),

$T_n^{est,1}$ exhibits severe size distortions when $\alpha+\beta\approx 1$ and $T_n^{est,2}$ behave similarly. On the contrary, $T_n^{score}, T_n^{res,1}$, and $T_n^{res,2}$ have no severe size distortions.

To examine power, we consider the case that $\theta=(\omega,\alpha,\beta)$ changes to $\theta'=(\omega',\alpha',\beta')$ at $[n\tau]$ with $\tau=1/3, 1/2, 2/3$:

$$\text{Case 1}: \omega=1 \text{ changes to } \omega'=0.3 \text{ and } (\alpha,\beta) \text{ does not change.}$$

$$\text{Case 2}: (\alpha,\beta)=(0.1,0.5) \text{ changes to } (\alpha',\beta'), \text{ and } \omega=1 \text{ does not change.}$$

We compare only the results of the score vector- and residual-based tests (see Tables 5.2 and 5.3), because the estimates-based tests have severe size distortions. Therein, we can see that the sizes are smaller in Case 1 than in Case 2, and that the powers in many cases are close to 1, but the power becomes smaller when $\alpha+\beta\approx 1$—that is, (0.1,0.8) and (0.4,0,5). In most cases, among the CUSUM tests, $T_n^{res,2}$ appears to produce the largest powers.

5.4.2 Test for NB-INGARCH(1,1) models

We consider the NB-INGARCH(1,1) model:

$$Y_t|\mathcal{F}_{t-1}\sim \mathrm{NB}(r,p_t),\quad X_t=\frac{r(1-p_t)}{p_t}=\omega+\alpha X_{t-1}+\beta Y_{t-1},$$

where X_1 is set to be 0. We assume that r is known. However, in practice, r is unknown and should be estimated—using, for example, an information criterion such as the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) Davis and Wu (2009).

To examine empirical size and power, we use the same settings as in the previous case, except that we deal only with $\tau=1/2$. In particular, we consider the cases of $r=1$ and $r=8$. As seen in the Poisson INGARCH(1,1) model case, our findings show that the estimates-based tests give rise to severe size distortions, while the others produce no size distortions; furthermore, $T_n^{res,2}$ produces the largest powers (see Tables 5.4-5.6). Overall, our simulation results confirm the validity of score vector- and residual-based CUSUM tests in terms of stability and power. In particular, these results advocate the superiority of the standardized residual-based CUSUM test over the other tests.

5.4.3 Test for binomial INGARCH(1,1) models

We consider the binomial INGARCH(1,1) model:

$$\mathrm{Y}_t|\mathcal{F}_{t-1} \sim \mathrm{B}(m, p_t), \quad X_t = mp_t = \omega + \alpha X_{t-1} + \beta Y_{t-1},$$

where X_1 is set to be 0 and m is known. We consider the cases of $m = 1, 5, 10$ and the parameters $(\alpha, \beta) = (0.1, 0.2), (0.1, 0.4), (0.2, 0.1), (0.3, 0.2)$, with $\omega = 0.1, 0.3$ for $m = 1$, $\omega = 0.5, 1$ for $m = 5$, and $\omega = 1, 3$ for $m = 10$. Tables 5.7-5.9 shows the sizes derived from the tests. As with the two aforementioned cases, the results of the estimates-based tests exhibit severe size distortions. T_n^{score} has a somewhat larger size whenever the (α, β) is small or the sample size is small, whereas residual-based tests produce no severe size distortion.

To examine empirical power, we consider the case that $\theta = (\omega, \alpha, \beta)$ changes to $\theta' = (\omega', \alpha', \beta')$ at $[n\tau]$ with $\tau = 1/3, 1/2, 2/3$:

Case 1 : ω changes to ω' and (α, β) does not change.

Case 2 : $(\alpha, \beta) = (0.1, 0.2)$ changes to (α', β'), and ω does not change.

It appears that the powers of $T_n^{score}, T_n^{res,1}$, and $T_n^{res,2}$ are similar (see Tables 5.10-5.14). In Case 1, the powers are close to 1, except when $m = 1$ and the sample size is small, regardless of τ. In Case 2, the powers are small when $m = 1$ and ω is small, but close to 1 in the other, remaining cases. Overall, among the tests studied, the standardized residual-based CUSUM test appears to perform best.

5.5 Real data analysis

In this section, we provide a real data example. We analyze daily data set of car accidents that occurred in Seondong county in Seoul, Korea during the period from January 1, 2011 to December 31, 2012 (the sample size = 731). The time series plot, the autocorrelation and partial autocorrelation functions are given in Figure 5.1. The mean and variance are 2.927 and 3.574, indicating over-dispersion.

We perform the CUSUM test and obtain that $T_n^{est,1} = 5.008$, $T_n^{est,2} = 35.870$, $T_n^{score} = 2.542$, $T_n^{res,1} = 1.502$ and $T_n^{res,2} = 1.497$, which reject the null hypothesis, H_0, except T_n^{score}. The results are presented in Figure 5.2. $T_n^{est,1}, T_n^{est,2}$ and $T_n^{res,2}$ are maximized at $t = 349$, and T_n^{score} and $T_n^{res,1}$ are maximized at $t = 363$. Since $T_n^{res,1}$ has the second highest value, 1.499, at $t = 349$, we can assume that the change occurs at $t = 349$, December 15, 2011. The parameter estimates are summarized in Table 5.15 for the full data under the null hypothesis and two sub-data before/after the change under the alternative hypothesis.

5.6 Concluding remarks

In this study, we considered CUSUM tests based on score vectors and residuals, and compared their performance for general integer-valued time series models. We derived their limiting null distributions under certain conditions and demonstrated their validity through a simulation study. Our findings show that score vector- and residual-based CUSUM tests can serve as promising alternative methods to estimates-based CUSUM tests; in particular, the standardized residual-based CUSUM mostly outperforms the other tests. Practitioners are therefore urged to use this test, unless they are planning a special mission.

5.7 Proofs

In this section, we provide the proofs of the theorems stated in the previous sections. In what follows, we use notation $\eta_t^0 = \eta_t(\theta_0)$ and $\eta_t^n = \eta_t(\hat{\theta}_n)$.

Lemma 5.1. *Suppose that conditions* **(C0)**, **(C7)** *and* **(C11)** *hold. Then, we have*

$$|\tilde{X}_t(\theta) - X_t(\theta)| \le V\rho^t, \quad |\tilde{\eta}_t - \eta_t| \le V\rho^t.$$

Proof. Note that

$$\begin{aligned} |\tilde{X}_t(\theta) - X_t(\theta)| &= \Big| f_\theta(\tilde{X}_{t-1}(\theta), Y_{t-1}) - f_\theta(X_{t-1}(\theta), Y_{t-1}) \Big| \\ &\le \omega_1 |\tilde{X}_{t-1}(\theta) - X_{t-1}(\theta)| \le \omega_1^{t-1} |\tilde{X}_1 - X_1(\theta)|. \end{aligned}$$

Then, using the mean value theorem and **(C11)**, we have

$$|\tilde{\eta}_t - \eta_t| = |B^{-1}(\tilde{X}_t(\theta)) - B^{-1}(X_t(\theta))| = \frac{\omega_1^{t-1}}{B'(\eta_t^*)} |\tilde{X}_1 - X_1(\theta)|$$

where $\eta_t^* = B^{-1}(X_t^*)$ and X_t^* is an intermediate point between $\tilde{X}_t(\theta)$ and $X_t(\theta)$. Hence, using by **(C7)**, the proof is completed. □

Lemma 5.2. *Suppose that* **(C0)**-**(C13)** *hold. Then, under H_0, we have that as $n \to \infty$,*

(i) $\displaystyle \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n \tilde{\ell}_t(\theta) - \frac{1}{n} \sum_{t=1}^n \ell_t(\theta) \right| \longrightarrow 0 \quad a.s.;$

(ii) $\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \tilde{\ell}_t(\theta_0)}{\partial \theta} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_t(\theta_0)}{\partial \theta} \right\| = o_P(1);$

(iii) $\sup_{\theta \in \Theta} \left\| \frac{\partial^2 \tilde{\ell}_t(\theta)}{\partial \theta \partial \theta^T} - \frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta^T} \right\| \longrightarrow 0 \quad a.s.;$

(iv) $-\frac{1}{n} \frac{\partial^2 \tilde{L}_n(\theta)}{\partial \theta \partial \theta^T} \longrightarrow I(\theta_0) \quad a.s..$

Proof. (i) It suffices to show that

$$\sup_{\theta \in \Theta} |\tilde{\ell}_t(\theta) - \ell_t(\theta)| \to 0 \ \ a.s. \ \text{ as } \ t \to \infty. \tag{5.6}$$

Note that, by the mean value theorem, **(C11)** and Lemma 5.1,

$$\begin{aligned}
\{\tilde{\ell}_t(\theta) - \ell_t(\theta)\}^+ &= \{\tilde{\ell}_t(\theta) - \ell_t(\theta)\} \vee 0 \\
&\leq |\tilde{\eta}_t - \eta_t| Y_t + |A(\tilde{\eta}_t) - A(\eta_t)| \\
&= |B^{-1}(\tilde{X}_t(\theta)) - B^{-1}(X_t(\theta))| Y_t + |A(B^{-1}(\tilde{X}_t(\theta))) - A(B^{-1}(X_t(\theta)))| \\
&\leq \frac{Y_t + X_t^*}{B'(\eta_t^*)} |\tilde{X}_t(\theta) - X_t(\theta)| \leq \frac{Y_t + X_t^*}{\underline{c}} V \rho^t
\end{aligned}$$

for some intermediate points X_t^* between $X_t(\theta)$ and $\tilde{X}_t(\theta)$ and $\eta_t^* = B^{-1}(X_t^*)$. Since

$$Y_t + X_t^* \leq Y_t + X_t(\theta) + |\tilde{X}_t(\theta) - X_t(\theta)| \leq Y_t + X_t(\theta) + V\rho^t,$$

according to **(C7)**, $\sup_{\theta \in \Theta} \{\tilde{\ell}_t(\theta) - \ell_t(\theta)\}^+ \to 0, \ \ a.s.$ as $t \to \infty$. Similarly, it can be seen that $\sup_{\theta \in \Theta} \{\tilde{\ell}_t(\theta) - \ell_t(\theta)\}^- \to 0 \ \ a.s.$ as $t \to \infty$, which yields (5.6).

(ii) Note that

$$\frac{\partial \ell_t(\theta)}{\partial \theta} = (Y_t - B(\eta_t)) \frac{\partial \eta_t}{\partial \theta} := U_t(\theta) \frac{\partial \eta_t}{\partial \theta}.$$

Hence, we have

$$\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \tilde{\ell}_t(\theta_0)}{\partial \theta} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_t(\theta_0)}{\partial \theta} \right\| \tag{5.7}$$

$$\leq \left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \tilde{U}_t(\theta_0) \left(\frac{\partial \tilde{\eta}_t^0}{\partial \theta} - \frac{\partial \eta_t^0}{\partial \theta} \right) \right\| + \left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(\tilde{U}_t(\theta_0) - U_t(\theta_0) \right) \frac{\partial \eta_t^0}{\partial \theta} \right\|.$$

From **(C7)**,**(C10)** and Lemma 5.1 , it can be seen that the first term of (5.7) is $o_P(1)$, since

$$|\tilde{U}_t(\theta)| = |Y_t - \tilde{X}_t(\theta)| \leq |Y_t - X_t(\theta)| + |X_t(\theta) - \tilde{X}_t(\theta)| \leq |Y_t - X_t(\theta)| + V\rho^t.$$

Since the second term of (5.7) is bounded by

$$\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(\tilde{X}_t(\theta_0) - X_t(\theta_0) \right) \frac{\partial \eta_t^0}{\partial \theta} \right\| \leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\| \frac{\partial \eta_t^0}{\partial \theta} \right\| V\rho^t,$$

it becomes $o_P(1)$ owing to Lemma 5.1 and **(C8)**.

(iii) Note that

$$\frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta^T} = -B'(\eta_t) \frac{\partial \eta_t}{\partial \theta} \frac{\partial \eta_t}{\partial \theta^T} + U_t(\theta) \frac{\partial^2 \eta_t}{\partial \theta \partial \theta^T}$$

Therefore, we have

$$\begin{aligned} \left\| \frac{\partial^2 \tilde{\ell}_t(\theta)}{\partial \theta \partial \theta^T} - \frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta^T} \right\| \leq & \left\| B'(\tilde{\eta}_t) \left(\frac{\partial \tilde{\eta}_t}{\partial \theta} - \frac{\partial \eta_t}{\partial \theta} \right) \frac{\partial \tilde{\eta}_t}{\partial \theta} \right\| + \left\| B'(\tilde{\eta}_t) \frac{\partial \eta_t}{\partial \theta} \left(\frac{\partial \tilde{\eta}_t}{\partial \theta} - \frac{\partial \eta_t}{\partial \theta} \right) \right\| \\ + & \left\| \{B'(\tilde{\eta}_t) - B'(\eta_t)\} \frac{\partial \eta_t}{\partial \theta} \frac{\partial \eta_t}{\partial \theta^T} \right\| + \left\| \tilde{U}_t(\theta) \left(\frac{\partial^2 \tilde{\eta}_t}{\partial \theta \partial \theta^T} - \frac{\partial^2 \eta_t}{\partial \theta \partial \theta^T} \right) \right\| \\ + & \left\| \left\{ \tilde{X}_t(\theta) - X_t(\theta) \right\} \frac{\partial^2 \eta_t}{\partial \theta \partial \theta^T} \right\|. \end{aligned} \tag{5.8}$$

Since $B'(\eta_t)\partial\eta_t/\partial\theta = \partial X_t(\theta)/\partial\theta$, the first and second term of the RHS of (5.8) converge to 0 $a.s.$ as $n \to \infty$ because of **(C8)** and **(C10)**. On the other hand, the forth and fifth terms converge to 0 $a.s.$ owing to **(C7)**, **(C10)** and Lemma 5.1, respectively. Due to **(C12)**, we have

$$\left\| \{B'(\tilde{\eta}_t) - B'(\eta_t)\} \frac{\partial \eta_t}{\partial \theta} \frac{\partial \eta_t}{\partial \theta^T} \right\| \leq \left\| \frac{1}{B'(\eta_t)^2} \frac{\partial X_t(\theta)}{\partial \theta} \frac{\partial X_t(\theta)}{\partial \theta^T} \right\| \cdot V\rho^t.$$

Henceforth, the third term converges to 0 $a.s.$ owing to **(C8)** and **(C11)**.

(iv) This can be proven similarly to the proof of Proposition 5 of Lee et al. (2016a). □

Lemma 5.3. *Suppose that conditions* **(C0)**-**(C13)** *hold. Then, under* H_0*, as* $n \to \infty$*,*

$$\frac{1}{\sqrt{n}} \max_{1\le k\le n} \left\| \tilde{S}_k(\hat{\theta}_n) - \left\{ \tilde{S}_k(\theta_0) - \frac{k}{n}\tilde{S}_n(\theta_0) \right\} \right\| = o_P(1),$$

where $\tilde{S}_k(\theta) = \sum_{t=1}^{k} \partial\tilde{\ell}_t(\theta)/\partial\theta$*.*

Proof. As $\hat{\theta}_n$ is the CMLE of θ_0, we show that

$$\frac{1}{\sqrt{n}} \max_{1\le k\le n} \left\| \tilde{S}_k(\hat{\theta}_n) - \tilde{S}_k(\theta_0) - \frac{k}{n} \left\{ \tilde{S}_n(\hat{\theta}_n) - \tilde{S}_n(\theta_0) \right\} \right\| = o_P(1).$$

By Taylor's theorem, we have

$$\tilde{S}_k(\hat{\theta}_n) = \tilde{S}_k(\theta_0) + \sum_{t=1}^{k} \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T}(\hat{\theta}_n - \theta_0),$$

where $\theta_n *$ is an intermediate point between θ_0 and $\hat{\theta}_n$. Thus we have

$$\begin{aligned}
&\frac{1}{\sqrt{n}} \max_{1\le k\le n} \left\| \tilde{S}_k(\hat{\theta}_n) - \tilde{S}_k(\theta_0) - \frac{k}{n} \left\{ \tilde{S}_n(\hat{\theta}_n) - \tilde{S}_n(\theta_0) \right\} \right\| \\
\le\ & \frac{1}{\sqrt{n}} \max_{1\le k\le n} \left\| \sum_{t=1}^{k} \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T}(\hat{\theta}_n - \theta_0) - \frac{k}{n} \sum_{t=1}^{n} \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T}(\hat{\theta}_n - \theta_0) \right\| \\
\le\ & \frac{1}{\sqrt{n}} \max_{1\le k\le n} \left\| \sum_{t=1}^{k} \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T}(\hat{\theta}_n - \theta_0) + kI(\theta_0)(\hat{\theta}_n - \theta_0) \right\| \\
& + \frac{1}{\sqrt{n}} \max_{1\le k\le n} \left\| \frac{k}{n} \sum_{t=1}^{n} \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T}(\hat{\theta}_n - \theta_0) + kI(\theta_0)(\hat{\theta}_n - \theta_0) \right\| \\
\le\ & \max_{1\le k\le n} \frac{k}{n} \left\| \frac{1}{k} \sum_{t=1}^{k} \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T} + I(\theta_0) \right\| \cdot \sqrt{n}\|\hat{\theta}_n - \theta_0\|
\end{aligned}$$

$$+ \max_{1\le k\le n} \frac{k}{n} \left\| \frac{1}{n}\sum_{t=1}^{n} \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T} + I(\theta_0) \right\| \cdot \sqrt{n}\|\hat{\theta}_n - \theta_0\| := I_n + II_n.$$

Note that, by Proposition 5.1 and (iv) in Lemma 5.2,

$$II_n \le \left\| \frac{1}{n}\sum_{t=1}^{n} \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T} + I(\theta_0) \right\| \cdot \sqrt{n}\|\hat{\theta}_n - \theta_0\| = o_p(1)\cdot O_P(1) = o_P(1).$$

Meanwhile, due to **(C8)**, for some $0 < \gamma < 1/2$,

$$\max_{1\le k\le n^\gamma} \frac{k}{n} \left\| \frac{1}{k}\sum_{t=1}^{k} \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T} + I(\theta_0) \right\| \le \frac{n^\gamma}{n}\sum_{t=1}^{n^\gamma} \left\| \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T} \right\| + \frac{n^\gamma}{n}\|I(\theta_0)\| = o_P(1).$$

Furthermore, since

$$\max_{n^\gamma < k\le n} \frac{k}{n} \left\| \frac{1}{k}\sum_{t=1}^{k} \frac{\partial^2 \tilde{\ell}_t(\theta_n^*)}{\partial\theta\partial\theta^T} + I(\theta_0) \right\| \to 0 \quad a.s.$$

owing to (iv) in Lemma 5.2, we can show that $I_n = o_P(1)$. This asserts the lemma. □

Proof of Theorem 5.1. Since $\{\partial\ell_t(\theta_0)/\partial\theta, \mathcal{F}_t\}$ forms a sequence of stationary ergodic martingale differences, using a functional central limit theorem, we can have

$$I(\theta_0)^{-1/2}\frac{1}{\sqrt{n}}S_{[ns]}(\theta_0) \xrightarrow{w} \mathbf{B}_d(s),$$

where $S_k(\theta) = \sum_{t=1}^{k} \partial\ell_t(\theta)/\partial\theta$ and $\{\mathbf{B}_d(s), 0 < s < 1\}$ is a $d-$dimensional standard Brownian motion. Further, from (i) in Lemma 5.2, we have

$$I(\theta_0)^{-1/2}\frac{1}{\sqrt{n}}\tilde{S}_{[ns]}(\theta_0) \xrightarrow{w} \mathbf{B}_d(s).$$

Then, using Lemma 5.3, we obtain

$$\hat{I}_n^{-1/2}\frac{1}{\sqrt{n}}\tilde{S}_{[ns]}(\hat{\theta}_n) \xrightarrow{w} \mathbf{B}_d(s).$$

This establishes the theorem. □

Proof of Theorem 5.3. We verify that

$$\max_{1\le k\le n}\frac{1}{\sqrt{n}}\left|\sum_{t=1}^{k}(\hat{\epsilon}_{t,2}-\epsilon_{t,2})-\frac{k}{n}\sum_{t=1}^{n}(\hat{\epsilon}_{t,2}-\epsilon_{t,2})\right|=o_P(1). \tag{5.9}$$

Note that

$$\begin{aligned}
\hat{\epsilon}_{t,2}-\epsilon_{t,2} &= \frac{Y_t-\hat{X}_t}{\sqrt{B'(\hat{\eta}_t)}}-\frac{Y_t-X_t(\theta_0)}{\sqrt{B'(\eta_t^0)}}\\
&= (X_t(\theta_0)-\hat{X}_t)\left(\frac{1}{\sqrt{B'(\hat{\eta}_t)}}-\frac{1}{\sqrt{B'(\eta_t^0)}}\right)\\
&\quad +\epsilon_{t,1}\left(\frac{1}{\sqrt{B'(\hat{\eta}_t)}}-\frac{1}{\sqrt{B'(\eta_t^0)}}\right)+\frac{1}{\sqrt{B'(\eta_t^0)}}(X_t(\theta_0)-\hat{X}_t))\\
&:= R_{t,1}+R_{t,2}+R_{t,3}.
\end{aligned}$$

It suffices to show that for $i=1,2,3$,

$$\max_{1\le k\le n}\frac{1}{\sqrt{n}}\left|\sum_{t=1}^{k}R_{t,i}-\frac{k}{n}\sum_{t=1}^{n}R_{t,i}\right|=o_P(1). \tag{5.10}$$

We express

$$\max_{1\le k\le n}\frac{1}{\sqrt{n}}\left|\sum_{t=1}^{k}R_{t,1}-\frac{k}{n}\sum_{t=1}^{n}R_{t,1}\right|\le\frac{2}{\sqrt{n}}\sum_{t=1}^{n}|R_{t,1}|\le I_{n,1}+I_{n,2}+I_{n,3},$$

where

$$\begin{aligned}
I_{n,1} &= \frac{2}{\sqrt{n}}\sum_{t=1}^{n}\left|\left(X_t(\theta_0)-X_t(\hat{\theta}_n)\right)\left(\frac{1}{\sqrt{B'(\eta_t^0)}}-\frac{1}{\sqrt{B'(\eta_t^n)}}\right)\right|,\\
I_{n,2} &= \frac{2}{\sqrt{n}}\sum_{t=1}^{n}\left|\left(X_t(\theta_0)-X_t(\hat{\theta}_n)\right)\left(\frac{1}{\sqrt{B'(\hat{\eta}_t)}}-\frac{1}{\sqrt{B'(\eta_t^n)}}\right)\right|,\\
I_{n,3} &= \frac{2}{\sqrt{n}}\sum_{t=1}^{n}\left|\left(X_t(\hat{\theta}_n)-\hat{X}_t\right)\left(\frac{1}{\sqrt{B'(\hat{\eta}_t)}}-\frac{1}{\sqrt{B'(\eta_t^n)}}\right)\right|.
\end{aligned}$$

Using the mean value theorem with intermediate points $\theta^*_{n,1}$ and $\theta^*_{n,2}$ between $\hat{\theta}_n$, we have

$$\begin{aligned}
I_{n,1} &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left|(\hat{\theta}_n-\theta_0)^T\frac{\partial X_t(\theta^*_{n,1})}{\partial\theta}\cdot\frac{B''(\eta_t(\theta^*_{n,2}))}{B'(\eta_t(\theta^*_{n,2}))^{3/2}}\frac{1}{B'(\eta_t(\theta^*_{n,2}))}(\hat{\theta}_n-\theta_0)^T\frac{\partial X_t(\theta^*_{n,2})}{\partial\theta}\right| \\
&\leq n\|\hat{\theta}_n-\theta_0\|^2\frac{1}{\underline{c}}\frac{1}{n\sqrt{n}}\sum_{t=1}^{n}\left|\sup_{\theta\in\Theta}\frac{B''(\eta_t)}{B'(\eta_t)^{3/2}}\right|\cdot\left\|\sup_{\theta\in\Theta}\frac{\partial X_t(\theta)}{\partial\theta}\right\|^2=o_P(1),
\end{aligned}$$

where we have used Theorem 1.1 and **(C8)**, **(C11)** and **(C13)**. Since $\hat{\eta}_t$ can be represented as $\tilde{\eta}_t(\hat{\theta}_n)=B^{-1}(\tilde{X}_t(\hat{\theta}_n))$ with $\tilde{X}_1(\hat{\theta}_n)=\hat{X}_1$, we have,

$$\begin{aligned}
I_{n,2} &\leq \frac{1}{\sqrt{n}}\frac{1}{2\underline{c}\sqrt{\underline{c}}}\sum_{t=1}^{n}\left|\left(X_t(\theta_0)-X_t(\hat{\theta}_n)\right)(B'(\eta^n_t)-B'(\hat{\eta}_t))\right| \\
&\leq \sqrt{n}\|\hat{\theta}_n-\theta_0\|\frac{V}{2\underline{c}\sqrt{\underline{c}}}\sum_{t=1}^{n}\rho^t\left\|\frac{\partial X_t(\theta^*_{n,1})}{\partial\theta}\right\|=O_p(1)\cdot o_P(1)=o_P(1),
\end{aligned}$$

with intermediate point $\theta^*_{n,1}$ between $\hat{\theta}_n$ and θ_0, due to **(C8)**, **(C11)** and **(C12)**. Furthermore, note that $|\hat{X}_t-X_t(\hat{\theta}_n)|\leq V\rho^t$ a.s. since owing to **(C0)**,

$$\begin{aligned}
|\hat{X}_t-X_t(\hat{\theta}_n)| &= \left|f_{\hat{\theta}_n}(\hat{X}_{t-1},Y_{t-1})-f_{\hat{\theta}_n}(X_{t-1}(\hat{\theta}_n),Y_{t-1})\right| \\
&\leq \omega_1|\hat{X}_{t-1}-X_{t-1}(\hat{\theta}_n)|\leq\omega_1^{t-1}|\hat{X}_1-X_1(\hat{\theta}_n)|.
\end{aligned}$$

Then, by using this and **(C11)**,

$$I_{n,3} \leq \frac{2}{\sqrt{n}}\sum_{t=1}^{n}V\rho^t\left\|\sup_{\theta\in\Theta}\frac{2}{\sqrt{B'(\eta_t)}}\right\|\leq\frac{4V}{\sqrt{\underline{c}}\sqrt{n}}\sum_{t=1}^{n}\rho^t=o_P(1).$$

Now, we show that (5.10) holds for $i=2$. We express

$$\max_{1\leq k\leq n}\frac{1}{\sqrt{n}}\left|\sum_{t=1}^{k}R_{t,2}-\frac{k}{n}\sum_{t=1}^{n}R_{t,2}\right|\leq\frac{2}{\sqrt{n}}\sum_{t=1}^{n}|R_{t,2}|\leq II_{n,1}+II_{n,2},$$

where

$$II_{n,1}=\max_{1\leq k\leq n}\frac{1}{\sqrt{n}}\left|\sum_{t=1}^{k}\epsilon_{t,1}\left(\frac{1}{\sqrt{B'(\hat{\eta}_t)}}-\frac{1}{\sqrt{B'(\eta^n_t)}}\right)-\frac{k}{n}\sum_{t=1}^{n}\epsilon_{t,1}\left(\frac{1}{\sqrt{B'(\hat{\eta}_t)}}-\frac{1}{\sqrt{B'(\eta^n_t)}}\right)\right|,$$

$$II_{n,2} = \max_{1\le k\le n} \frac{1}{\sqrt{n}} \left| \sum_{t=1}^{k} \epsilon_{t,1} \left(\frac{1}{\sqrt{B'(\eta_t^n)}} - \frac{1}{\sqrt{B'(\eta_t^0)}} \right) - \frac{k}{n} \sum_{t=1}^{n} \epsilon_{t,1} \left(\frac{1}{\sqrt{B'(\eta_t^n)}} - \frac{1}{\sqrt{B'(\eta_t^0)}} \right) \right|.$$

First, we can see that, using **(C7)**, **(C11)** and **(C12)**,

$$\begin{aligned} II_{n,1} &\le \frac{2}{\sqrt{n}} \sum_{t=1}^{n} \left| \epsilon_{t,1} \left(\frac{1}{\sqrt{B'(\hat\eta_t)}} - \frac{1}{\sqrt{B'(\hat\eta_t^n)}} \right) \right| \\ &\le \frac{1}{\sqrt{n}} \frac{1}{\underline{c}\sqrt{\underline{c}}} \sum_{t=1}^{n} |\epsilon_{t,1}(B'(\hat\eta_t) - B'(\eta_t^n))| \le \frac{V}{\sqrt{n}} \frac{1}{\underline{c}\sqrt{\underline{c}}} \sum_{t=1}^{n} |\epsilon_{t,1}|\rho^t = o_P(1). \end{aligned}$$

Next, using Taylor's theorem, we can have

$$B'(\eta_t^n)^{-1/2} = B'(\eta_t^0)^{-1/2} - \frac{1}{2} Z_t(\theta_0)(\hat\theta_n - \theta_0)^T \frac{\partial \eta_t^0}{\partial\theta} - \frac{1}{2}(\hat\theta_n - \theta_0)^T \left(Z_t(\theta_n^*) \frac{\partial \eta_{t,n}^*}{\partial\theta} - Z_t(\theta_0) \frac{\partial \eta_t^0}{\partial\theta} \right),$$

where $\eta_{t,n}^* = \eta_t(\theta_n^*)$ and θ_n^* is an intermediate point between $\hat\theta_n$ and θ_0, and $Z_t(\theta) = B''(\eta_t)B'(\eta_t)^{-3/2}$, so that

$$II_{n,2} \le II'_{n,2} + II''_{n,2}$$

where

$$\begin{aligned} II'_{n,2} &= \sqrt{n}\|\hat\theta_n - \theta_0\| \frac{k}{n} \left| \frac{1}{k} \sum_{t=1}^{k} \epsilon_{t,1} Z_t(\theta_0) \frac{\partial \eta_t^0}{\partial\theta} - \frac{1}{n} \sum_{t=1}^{n} \epsilon_{t,1} Z_t(\theta_0) \frac{\partial \eta_t^0}{\partial\theta} \right|, \\ II''_{n,2} &= \sqrt{n}\|\hat\theta_n - \theta_0\| \frac{2}{n} \sum_{t=1}^{n} |\epsilon_{t,1}| \left\| Z_t(\theta_n^*) \frac{\partial \eta_{t,n}^*}{\partial\theta} - Z_t(\theta_0) \frac{\partial \eta_t^0}{\partial\theta} \right\|. \end{aligned}$$

Since $\{\epsilon_{t,1} Z_t(\theta_0) \partial\eta_t^0/\partial\theta\}$ is ergodic and $\sqrt{n}\|\hat\theta_n - \theta_0\| = O_p(1)$, we have $II'_{n,2} = o_P(1)$ by using ergodic theorem. Further, because

$$II''_{n,2} \le \sqrt{n}\|\hat\theta_n - \theta_0\| \frac{2}{n} \sum_{t=1}^{n} |\epsilon_{t,1}| \sup_{\|\theta-\theta_0\| \le \|\hat\theta_n - \theta_0\|} \left\| Z_t(\theta) \frac{\partial \eta_t}{\partial\theta} - Z_t(\theta_0) \frac{\partial \eta_t^0}{\partial\theta} \right\|$$

and $E\sup_{\theta\in\Theta} |Z_t(\theta)| \cdot \|\partial\eta_t/\partial\theta\| < \infty$, we have $II''_{n,2} = o_P(1)$, which implies $II_{n,2} = o_P(1)$, and thus, (5.10) for $i = 2$.

Finally, using Taylor's theorem, we write

$$X_t(\hat{\theta}_n) = X_t(\theta_0) + (\hat{\theta}_n - \theta_0)^T \frac{\partial X_t(\theta_0)}{\partial \theta} + (\hat{\theta}_n - \theta_0)^T \left(\frac{\partial X_t(\theta_n^*)}{\partial \theta} - \frac{\partial X_t(\theta_0)}{\partial \theta} \right)$$

for some θ_n^* between $\hat{\theta}_n$ and θ_0. Then, similarly to the case of $II_{n,2}$, we can show that (5.10) holds for $i = 3$. Hence, (5.9) is verified. The theorem is then a direct result of (5.4). □

Table 5.1: Empirical sizes for Poisson INGARCH (1,1) models.

	$\omega = 1$					$\omega = 0.3$				
n	$T_n^{est,1}$	$T_n^{est,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	$T_n^{est,1}$	$T_n^{est,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
$(\alpha, \beta) = (0.1, 0.3)$										
300	0.110	0.122	0.074	0.036	0.026	0.090	0.110	0.076	0.032	0.048
500	0.086	0.090	0.064	0.032	0.036	0.112	0.100	0.066	0.036	0.038
1000	0.055	0.080	0.036	0.045	0.040	0.080	0.095	0.070	0.050	0.055
$(\alpha, \beta) = (0.1, 0.5)$										
300	0.100	0.114	0.040	0.038	0.048	0.107	0.124	0.056	0.032	0.040
500	0.063	0.068	0.028	0.042	0.030	0.075	0.110	0.042	0.054	0.048
1000	0.040	0.055	0.050	0.045	0.038	0.050	0.060	0.045	0.050	0.060
$(\alpha, \beta) = (0.1, 0.8)$										
300	0.322	0.454	0.028	0.048	0.030	0.250	0.430	0.032	0.036	0.038
500	0.244	0.362	0.024	0.038	0.040	0.234	0.348	0.050	0.040	0.038
1000	0.210	0.170	0.038	0.025	0.044	0.190	0.265	0.040	0.050	0.025
$(\alpha, \beta) = (0.3, 0.2)$										
300	0.210	0.240	0.032	0.036	0.030	0.172	0.236	0.042	0.024	0.026
500	0.204	0.222	0.054	0.038	0.038	0.210	0.246	0.052	0.036	0.038
1000	0.120	0.205	0.046	0.020	0.040	0.175	0.175	0.055	0.040	0.045
$(\alpha, \beta) = (0.3, 0.4)$										
300	0.230	0.240	0.026	0.022	0.046	0.226	0.242	0.014	0.038	0.038
500	0.182	0.180	0.028	0.034	0.036	0.184	0.196	0.024	0.020	0.020
1000	0.220	0.185	0.034	0.035	0.038	0.165	0.170	0.055	0.045	0.040
$(\alpha, \beta) = (0.4, 0.5)$										
300	0.388	0.528	0.022	0.014	0.014	0.378	0.542	0.016	0.024	0.026
500	0.330	0.498	0.044	0.032	0.024	0.268	0.488	0.036	0.048	0.038
1000	0.195	0.320	0.040	0.040	0.048	0.270	0.375	0.035	0.045	0.035

Table 5.2: Empirical powers for Poisson INGARCH (1,1) models when $\omega = 1$ changes to $\omega' = 0.3$ at $t = [n\tau]$ and (α, β) does not change.

	$\tau = 1/3$			$\tau = 1/2$			$\tau = 2/3$		
n	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
$(\alpha, \beta) = (0.1, 0.3)$									
300	0.750	0.972	0.984	0.660	0.960	1.000	0.794	0.698	0.982
500	0.996	1.000	1.000	0.998	1.000	1.000	0.992	0.978	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(\alpha, \beta) = (0.1, 0.5)$									
300	0.628	0.920	0.962	0.772	0.844	0.992	0.914	0.580	0.990
500	0.988	1.000	1.000	0.994	1.000	1.000	1.000	0.964	0.998
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(\alpha, \beta) = (0.1, 0.8)$									
300	0.060	0.192	0.180	0.068	0.094	0.222	0.140	0.016	0.168
500	0.114	0.402	0.530	0.184	0.214	0.704	0.420	0.042	0.580
1000	0.060	0.070	0.065	0.812	0.688	1.000	0.984	0.236	0.996
$(\alpha, \beta) = (0.3, 0.2)$									
300	0.748	0.986	0.990	0.516	0.960	0.998	0.508	0.596	0.968
500	0.998	1.000	1.000	0.986	1.000	1.000	0.970	0.988	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(\alpha, \beta) = (0.3, 0.4)$									
300	0.406	0.858	0.862	0.286	0.704	0.930	0.552	0.314	0.856
500	0.890	0.996	1.000	0.802	0.962	1.000	0.926	0.742	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(\alpha, \beta) = (0.4, 0.5)$									
300	0.074	0.426	0.176	0.034	0.178	0.224	0.038	0.040	0.116
500	0.030	0.026	0.022	0.058	0.434	0.696	0.116	0.076	0.538
1000	0.030	0.055	0.050	0.600	0.976	1.000	0.748	0.484	1.000

Table 5.3: Empirical powers for Poisson INGARCH (1,1) models when $(\alpha, \beta) = (0.1, 0.5)$ change to (α', β') at $t = [n\tau]$ and ω does not change.

	$\tau = 1/3$			$\tau = 1/2$			$\tau = 2/3$		
n	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
$(\alpha', \beta') = (0.1, 0.3)$									
300	0.790	0.998	1.000	0.610	0.942	1.000	0.836	0.648	0.994
500	0.998	1.000	1.000	0.968	1.000	1.000	0.994	0.954	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(\alpha', \beta') = (0.1, 0.8)$									
300	0.370	0.044	0.092	0.724	0.112	0.144	0.780	0.156	0.232
500	0.808	0.076	0.144	0.976	0.114	0.222	0.972	0.210	0.242
1000	1.000	0.088	0.244	1.000	0.155	0.330	1.000	0.328	0.312
$(\alpha', \beta') = (0.3, 0.2)$									
300	0.804	0.998	0.994	0.706	0.988	1.000	0.866	0.822	0.994
500	1.000	1.000	1.000	0.976	1.000	1.000	0.992	0.992	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(\alpha', \beta') = (0.3, 0.4)$									
300	0.444	0.810	0.854	0.670	0.830	0.962	0.768	0.634	0.900
500	0.892	0.988	0.994	0.988	0.988	1.000	0.988	0.976	0.998
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(\alpha', \beta') = (0.4, 0.5)$									
300	0.886	1.000	0.992	0.822	0.996	0.998	0.856	0.908	0.976
500	1.000	1.000	1.000	0.980	1.000	1.000	0.998	0.998	1.000
1000	1.000	1.000	1.000	0.990	1.000	1.000	1.000	1.000	1.000

Table 5.4: Empirical sizes for negative binomial INGARCH (1,1) models.

	$\omega = 1$					$\omega = 0.3$				
n	$T_n^{est,1}$	$T_n^{est,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	$T_n^{est,1}$	$T_n^{est,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
$(\alpha, \beta) = (0.1, 0.3)$										
300	0.136	0.122	0.072	0.032	0.032	0.120	0.120	0.088	0.032	0.034
500	0.112	0.138	0.080	0.026	0.032	0.078	0.098	0.048	0.028	0.046
1000	0.084	0.118	0.074	0.056	0.050	0.084	0.096	0.050	0.038	0.052
$(\alpha, \beta) = (0.1, 0.5)$										
300	0.126	0.124	0.030	0.024	0.028	0.118	0.104	0.062	0.024	0.036
500	0.114	0.086	0.038	0.020	0.044	0.112	0.088	0.048	0.024	0.028
1000	0.100	0.126	0.060	0.048	0.046	0.076	0.070	0.034	0.040	0.036
$(\alpha, \beta) = (0.1, 0.8)$										
300	0.168	0.176	0.042	0.024	0.034	0.198	0.320	0.022	0.026	0.030
500	0.154	0.172	0.038	0.028	0.038	0.208	0.256	0.042	0.032	0.036
1000	0.132	0.156	0.046	0.040	0.032	0.180	0.200	0.058	0.038	0.042
$(\alpha, \beta) = (0.3, 0.2)$										
300	0.188	0.258	0.024	0.018	0.022	0.170	0.264	0.038	0.042	0.030
500	0.228	0.296	0.046	0.036	0.032	0.194	0.228	0.050	0.024	0.028
1000	0.186	0.268	0.050	0.040	0.038	0.162	0.234	0.038	0.044	0.038
$(\alpha, \beta) = (0.3, 0.4)$										
300	0.234	0.290	0.026	0.032	0.034	0.216	0.258	0.024	0.030	0.032
500	0.232	0.228	0.038	0.028	0.034	0.178	0.226	0.034	0.028	0.026
1000	0.142	0.174	0.034	0.046	0.044	0.180	0.168	0.048	0.046	0.044
$(\alpha, \beta) = (0.4, 0.5)$										
300	0.300	0.484	0.036	0.030	0.028	0.336	0.496	0.028	0.040	0.030
500	0.306	0.388	0.034	0.044	0.038	0.334	0.466	0.024	0.028	0.018
1000	0.230	0.310	0.056	0.040	0.042	0.252	0.336	0.038	0.046	0.036

Table 5.5: Empirical powers for negative binomial INGARCH (1,1) models when the parameter change occurs at $t = [n/2]$.

		$r=1$			$r=8$		
		T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
(α, β)	n	$\delta = 1 \longrightarrow \delta' = 0.3$					
	300	0.648	0.500	0.968	0.662	0.886	0.994
(0.1,0.3)	500	0.988	0.814	1.000	0.996	0.998	1.000
	1000	1.000	0.996	1.000	1.000	1.000	1.000
	300	0.760	0.180	0.932	0.780	0.676	0.990
(0.1,0.5)	500	0.992	0.350	0.996	0.992	0.954	1.000
	1000	1.000	0.804	1.000	1.000	1.000	1.000
	300	0.666	0.008	0.700	0.182	0.028	0.438
(0.1,0.8)	500	0.958	0.006	0.944	0.574	0.046	0.786
	1000	1.000	0.004	0.990	0.980	0.078	1.000
	300	0.558	0.412	0.968	0.510	0.904	0.998
(0.3,0.2)	500	0.930	0.814	0.996	0.986	1.000	1.000
	1000	1.000	0.992	1.000	1.000	1.000	1.000
	300	0.562	0.104	0.836	0.306	0.412	0.908
(0.3,0.4)	500	0.926	0.208	0.996	0.820	0.844	1.000
	1000	1.000	0.584	1.000	0.998	1.000	1.000
	300	0.172	0.018	0.378	0.052	0.076	0.218
(0.4,0.5)	500	0.602	0.028	0.770	0.168	0.090	0.656
	1000	0.992	0.014	0.990	0.826	0.248	0.996

Table 5.6: Empirical powers for negative binomial INGARCH (1,1) models when the parameter change occurs at $t = [n/2]$.

		$r=1$			$r=8$		
		T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
(α', β')	n	$(\alpha, \beta) = (0.1, 0.5) \longrightarrow (\alpha', \beta')$					
	300	0.112	0.360	0.990	0.270	0.814	1.000
(0.1,0.3)	500	0.222	0.544	0.998	0.638	0.994	1.000
	1000	0.472	0.892	1.000	0.952	1.000	1.000
	300	0.124	0.018	0.488	0.492	0.040	0.258
(0.1,0.8)	500	0.306	0.018	0.758	0.962	0.084	0.356
	1000	0.718	0.044	0.970	1.000	0.152	0.532
	300	0.186	0.504	0.974	0.364	0.920	0.992
(0.3,0.2)	500	0.306	0.814	0.998	0.622	0.998	1.000
	1000	0.706	0.988	1.000	0.956	1.000	1.000
	300	0.122	0.156	0.740	0.210	0.694	0.932
(0.3,0.4)	500	0.234	0.424	0.956	0.460	0.968	1.000
	1000	0.558	0.798	0.998	1.000	1.000	1.000
	300	0.450	0.660	0.898	0.654	0.970	0.996
(0.4,0.5)	500	0.798	0.942	0.996	0.932	1.000	1.000
	1000	0.988	0.998	1.000	1.000	1.000	1.000

Table 5.7: Empirical sizes for binomial INGARCH (1,1) models when $m = 1$.

	$\omega = 0.1$					$\omega = 0.3$				
n	$T_n^{est,1}$	$T_n^{est,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	$T_n^{est,1}$	$T_n^{est,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
$(\alpha, \beta) = (0.1, 0.2)$										
300	0.136	0.244	0.116	0.020	0.022	0.100	0.174	0.082	0.030	0.028
500	0.112	0.224	0.098	0.042	0.040	0.098	0.208	0.076	0.032	0.030
1000	0.120	0.204	0.084	0.052	0.040	0.080	0.192	0.060	0.036	0.040
$(\alpha, \beta) = (0.1, 0.4)$										
300	0.106	0.178	0.072	0.026	0.032	0.092	0.148	0.062	0.034	0.032
500	0.108	0.144	0.070	0.042	0.052	0.090	0.110	0.050	0.036	0.042
1000	0.072	0.092	0.052	0.032	0.044	0.076	0.108	0.044	0.068	0.072
$(\alpha, \beta) = (0.2, 0.1)$										
300	0.140	0.388	0.146	0.016	0.020	0.156	0.250	0.072	0.044	0.044
500	0.150	0.350	0.106	0.032	0.038	0.124	0.204	0.076	0.036	0.038
1000	0.168	0.300	0.072	0.036	0.040	0.104	0.184	0.036	0.036	0.036
$(\alpha, \beta) = (0.3, 0.2)$										
300	0.202	0.320	0.054	0.026	0.028	0.210	0.292	0.070	0.040	0.042
500	0.198	0.280	0.034	0.046	0.050	0.168	0.238	0.052	0.046	0.046
1000	0.164	0.200	0.036	0.028	0.028	0.192	0.276	0.072	0.060	0.064

Table 5.8: Empirical sizes for binomial INGARCH (1,1) models when $m = 5$.

	$\omega = 0.5$					$\omega = 1$				
n	$T_n^{est,1}$	$T_n^{est,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	$T_n^{est,1}$	$T_n^{est,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
$(\alpha, \beta) = (0.1, 0.2)$										
300	0.116	0.198	0.088	0.032	0.036	0.092	0.168	0.088	0.022	0.024
500	0.114	0.170	0.082	0.036	0.034	0.082	0.136	0.084	0.034	0.032
1000	0.084	0.148	0.076	0.048	0.048	0.084	0.164	0.044	0.048	0.048
$(\alpha, \beta) = (0.1, 0.4)$										
300	0.092	0.146	0.056	0.030	0.034	0.130	0.120	0.066	0.020	0.024
500	0.098	0.114	0.074	0.042	0.038	0.096	0.090	0.060	0.032	0.032
1000	0.072	0.088	0.044	0.024	0.028	0.108	0.096	0.048	0.044	0.048
$(\alpha, \beta) = (0.2, 0.1)$										
300	0.132	0.344	0.150	0.044	0.044	0.132	0.172	0.060	0.024	0.024
500	0.156	0.288	0.110	0.038	0.040	0.156	0.172	0.068	0.046	0.044
1000	0.132	0.256	0.092	0.044	0.048	0.132	0.132	0.048	0.032	0.028
$(\alpha, \beta) = (0.3, 0.2)$										
300	0.202	0.252	0.052	0.036	0.032	0.154	0.210	0.092	0.038	0.036
500	0.184	0.246	0.062	0.036	0.032	0.156	0.196	0.070	0.042	0.042
1000	0.164	0.256	0.068	0.036	0.044	0.156	0.192	0.072	0.076	0.076

Table 5.9: Empirical sizes for binomial INGARCH (1,1) models when $m = 10$.

	$\omega = 0.5$					$\omega = 1$				
n	$T_n^{est,1}$	$T_n^{est,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	$T_n^{est,1}$	$T_n^{est,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
$(\alpha, \beta) = (0.1, 0.2)$										
300	0.104	0.190	0.098	0.028	0.030	0.090	0.080	0.154	0.032	0.032
500	0.102	0.146	0.088	0.032	0.032	0.074	0.086	0.110	0.040	0.040
1000	0.088	0.140	0.072	0.040	0.036	0.064	0.096	0.080	0.048	0.048
$(\alpha, \beta) = (0.1, 0.4)$										
300	0.126	0.124	0.070	0.032	0.036	0.104	0.084	0.090	0.042	0.034
500	0.084	0.100	0.050	0.046	0.044	0.066	0.068	0.088	0.024	0.024
1000	0.120	0.108	0.068	0.048	0.048	0.052	0.044	0.040	0.032	0.036
$(\alpha, \beta) = (0.2, 0.1)$										
300	0.140	0.184	0.074	0.020	0.022	0.062	0.084	0.172	0.050	0.050
500	0.134	0.188	0.062	0.040	0.042	0.064	0.090	0.180	0.066	0.066
1000	0.100	0.164	0.036	0.036	0.032	0.064	0.080	0.144	0.052	0.052
$(\alpha, \beta) = (0.3, 0.2)$										
300	0.158	0.212	0.092	0.044	0.044	0.136	0.120	0.190	0.068	0.066
500	0.174	0.226	0.054	0.024	0.032	0.142	0.134	0.186	0.034	0.032
1000	0.140	0.212	0.064	0.048	0.068	0.124	0.168	0.088	0.056	0.056

Table 5.10: Empirical powers for binomial INGARCH (1,1) models when $m = 1$ and $\omega = 0.1$ changes to $\omega' = 0.3$ at $t = [n\tau]$ and (α, β) does not change.

	$\tau = 1/3$			$\tau = 1/2$			$\tau = 2/3$		
n	T_n^{score}	$T_n^{res.1}$	$T_n^{res.2}$	T_n^{score}	$T_n^{res.1}$	$T_n^{res.2}$	T_n^{score}	$T_n^{res.1}$	$T_n^{res.2}$
$(\alpha, \beta) = (0.1, 0.2)$									
300	0.670	0.616	0.716	0.600	0.662	0.726	0.478	0.724	0.714
500	0.928	0.768	0.916	0.780	0.866	0.926	0.740	0.956	0.942
1000	1.000	0.988	1.000	0.896	1.000	1.000	0.960	1.000	1.000
$(\alpha, \beta) = (0.1, 0.4)$									
300	0.712	0.880	0.900	0.760	0.952	0.962	0.608	0.888	0.882
500	0.970	0.994	0.996	0.970	0.998	0.998	0.944	0.994	0.994
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(\alpha, \beta) = (0.2, 0.1)$									
300	0.720	0.388	0.560	0.528	0.518	0.648	0.404	0.602	0.586
500	0.932	0.626	0.858	0.678	0.816	0.912	0.660	0.926	0.906
1000	1.000	0.992	0.996	0.956	1.000	1.000	0.948	0.996	0.996
$(\alpha, \beta) = (0.3, 0.2)$									
300	0.770	0.754	0.744	0.542	0.828	0.802	0.408	0.812	0.718
500	0.936	0.914	0.924	0.718	0.970	0.950	0.648	0.974	0.938
1000	0.992	0.996	0.996	0.844	1.000	1.000	0.884	1.000	1.000

Table 5.11: Empirical powers for binomial INGARCH (1,1) models when $m = 5$ and $\omega = 0.5$ changes to $\omega' = 1$ at $t = [n\tau]$ and (α, β) does not change.

	$\tau = 1/3$			$\tau = 1/2$			$\tau = 2/3$		
n	T_n^{score}	$T_n^{res.1}$	$T_n^{res.2}$	T_n^{score}	$T_n^{res.1}$	$T_n^{res.2}$	T_n^{score}	$T_n^{res.1}$	$T_n^{res.2}$
$(\alpha, \beta) = (0.1, 0.2)$									
300	0.838	0.750	0.878	0.752	0.858	0.910	0.644	0.870	0.870
500	0.984	0.914	0.960	0.928	0.960	0.984	0.912	0.984	0.980
1000	1.000	0.992	1.000	0.988	1.000	1.000	0.980	1.000	1.000
$(\alpha, \beta) = (0.1, 0.4)$									
300	0.844	0.932	0.968	0.832	0.968	0.986	0.724	0.940	0.938
500	0.992	1.000	1.000	0.996	1.000	1.000	0.984	1.000	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(\alpha, \beta) = (0.2, 0.1)$									
300	0.938	0.688	0.868	0.806	0.792	0.866	0.682	0.858	0.862
500	0.992	0.836	0.952	0.910	0.908	0.964	0.850	0.954	0.936
1000	1.000	0.996	1.000	1.000	1.000	1.000	0.988	1.000	1.000
$(\alpha, \beta) = (0.3, 0.2)$									
300	0.868	0.808	0.898	0.730	0.902	0.928	0.610	0.914	0.906
500	0.990	0.948	0.990	0.918	0.974	0.982	0.856	0.996	0.990
1000	1.000	1.000	1.000	0.992	1.000	1.000	0.992	1.000	1.000

Table 5.12: Empirical powers for binomial INGARCH (1,1) models when $m = 10$ and $\omega = 1$ changes to $\omega' = 3$ at $t = [n\tau]$ and (α, β) does not change.

	$\tau = 1/3$			$\tau = 1/2$			$\tau = 2/3$		
n	T_n^{score}	$T_n^{res.1}$	$T_n^{res.2}$	T_n^{score}	$T_n^{res.1}$	$T_n^{res.2}$	T_n^{score}	$T_n^{res.1}$	$T_n^{res.2}$
$(\alpha, \beta) = (0.1, 0.2)$									
300	0.970	0.814	0.938	0.812	0.882	0.944	0.772	0.940	0.934
500	0.998	0.922	0.984	0.934	0.954	0.978	0.906	0.970	0.954
1000	1.000	0.984	1.000	1.000	1.000	1.000	0.992	1.000	1.000
$(\alpha, \beta) = (0.1, 0.4)$									
300	0.952	0.986	0.996	0.910	0.996	1.000	0.890	0.998	0.998
500	1.000	1.000	1.000	0.994	1.000	1.000	0.996	1.000	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(\alpha, \beta) = (0.2, 0.1)$									
300	0.986	0.746	0.932	0.916	0.840	0.912	0.792	0.836	0.822
500	1.000	0.870	0.980	0.962	0.924	0.960	0.944	0.950	0.932
1000	1.000	0.972	1.000	0.992	1.000	1.000	1.000	1.000	0.996
$(\alpha, \beta) = (0.3, 0.2)$									
300	0.974	0.886	0.964	0.770	0.932	0.948	0.724	0.954	0.924
500	1.000	0.954	0.992	0.932	0.990	0.992	0.914	0.994	0.988
1000	1.000	1.000	1.000	0.992	1.000	1.000	0.996	1.000	1.000

Table 5.13: Empirical powers for binomial INGARCH (1,1) models when $(\alpha, \beta) = (0.1, 0.2)$ changes to (α', β') at $t = [n/2]$ and ω does not change.

(α', β')	n	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
$m = 1$		$\omega = 0.1$			$\omega = 0.3$		
	300	0.264	0.142	0.070	0.894	0.816	0.840
(0.1,0.5)	500	0.516	0.284	0.168	0.990	0.974	0.976
	1000	0.904	0.588	0.368	1.000	1.000	1.000
	300	0.140	0.186	0.208	0.628	0.880	0.828
(0.4,0.2)	500	0.270	0.432	0.434	0.852	0.994	0.970
	1000	0.660	0.836	0.824	0.984	1.000	1.000
	300	0.232	0.416	0.318	1.000	0.926	0.988
(0.3,0.4)	500	0.568	0.712	0.592	1.000	0.970	0.996
	1000	0.980	0.984	0.928	1.000	0.996	1.000
$m = 5$		$\omega = 0.5$			$\omega = 1$		
	300	0.572	0.664	0.516	0.900	0.968	0.960
(0.1,0.5)	500	0.916	0.914	0.822	1.000	1.000	1.000
	1000	1.000	1.000	0.988	1.000	1.000	1.000
	300	0.528	0.764	0.790	0.800	0.964	0.958
(0.4,0.2)	500	0.892	0.968	0.974	0.962	0.998	0.998
	1000	0.996	1.000	1.000	1.000	1.000	1.000
	300	0.654	0.950	0.938	0.712	0.988	0.974
(0.3,0.4)	500	0.988	0.998	1.000	0.942	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 5.14: Empirical powers for binomial INGARCH (1,1) models when $m = 10$, $(\alpha, \beta) = (0.1, 0.2)$ changes to (α', β') at $t = [n/2]$ and ω does not change.

(α', β')	n	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$	T_n^{score}	$T_n^{res,1}$	$T_n^{res,2}$
$m = 10$			$\omega = 1$			$\omega = 3$	
	300	0.846	0.940	0.868	0.946	0.994	0.992
(0.1,0.5)	500	0.998	0.996	0.996	1.000	0.984	0.984
	1000	1.000	1.000	1.000	1.000	1.000	1.000
	300	0.794	0.940	0.960	0.848	0.998	0.978
(0.4,0.2)	500	0.970	0.998	1.000	0.970	0.996	0.996
	1000	1.000	1.000	1.000	1.000	1.000	1.000
	300	0.776	0.970	0.974	1.000	1.000	1.000
(0.3,0.4)	500	0.988	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

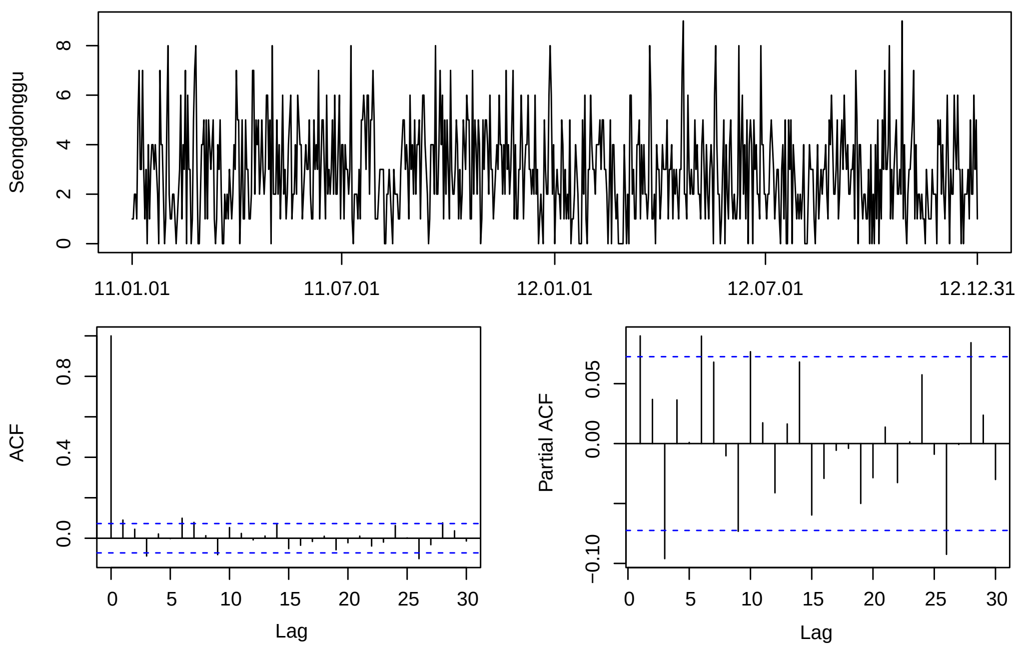

Figure 5.1: Plot of counts series, the sample autocorrealation and the sample partial autocorrealation for the traffic accidents of Seongdong county.

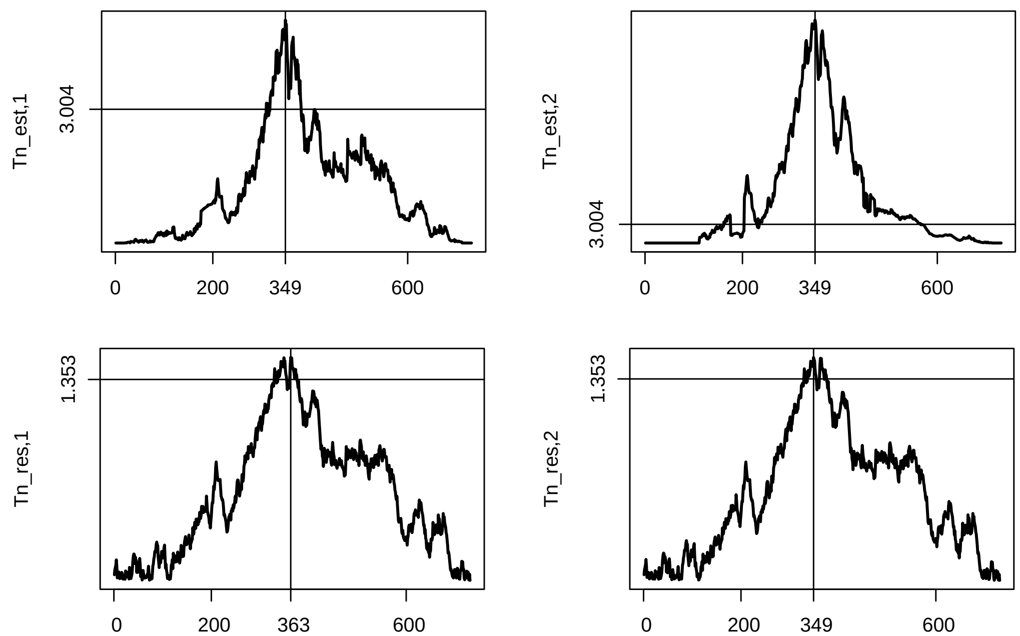

Figure 5.2: Plot of Plot of $T_n^{est,i}$ and $T_n^{res,i}$, $i = 1, 2$, for the traffic accidents of Seongdong county.

Table 5.15: Estimated parameters for the counts of traffic accidents of Seongdong county in Seoul, Korea, based on a Poisson INGARH(1,1) model. Standard errors are shown in parentheses.

	Mean	Variance	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
Full period	2.927	3.574	2.203	0.155	0.094
			(0.942)	(0.326)	(0.030)
First period	3.163	3.613	2.843	0.013	0.090
Jan.01.2011-Dec.14.2011			(1.769)	(0.556)	(0.047)
Second period	2.713	3.451	1.947	0.329	0.062
Dec.15.2011 - Dec.31.201			(1.539)	(0.582)	(0.038)

Bibliography

Al-Osh, M. A. and Aly, E.-E. A. (1992). First order autoregressive time series with negative binomial and geometric marginals. *Commun. Stat.-Theory Methods* **21**(9), 2483–2492.

Alzaid, A. and Al-Osh, M. (1990). An integer-valued pth-order autoregressive structure (INAR(p)) process. *J. Appl. Probab.* **27**, 314–324.

Andreassen, C. M. (2013). Models and inference for correlated count data. PhD thesis. Aarhus University.

Berkes, I., Horvth, L. and Kokoszka, P. (2004). Testing for parameter constancy in GARCH(p,q) models. *Stat. Probab. Lett.* **70**(4), 263–273.

Christou, V. and Fokianos, K. (2014). Quasi-likelihood inference for negative binomial time series models. *J. Time Ser. Anal.* **35**(1), 55–78. doi: 10.1111/jtsa.12050.

Consul, P. C. and Famoye, F. (1992). Generalized poisson regression model. *Commun. Stat.-Theory Methods* **21**(1), 89–109.

Consul, P. C. and Jain, G. C. (1973). A generalization of the poisson distribution. *Technometrics* **15**(4), 791–799.

Csörgö, M. and Horváth, L. (1997). *Limit theorems in change-point analysis*. Vol. 18. John Wiley & Sons Inc.

Davis, A., R. and Liu, H. (2016). Theory and inference for a class of observation-driven models with application to time series of counts. *Stat. Sin.* **26**(4), 1673–1707.

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

Davis, R., Dunsmuir, W. and Wang, Y. (2000). On autocorrelation in a Poisson regression model. *Biometrika* **87**(3), 491–505. doi: 10.1093/biomet/87.3.491.

De Pooter, M. and van Dijk, D. (2004). Testing for changes in volatility in heteroskedastic time series - a further examination. Econometric Institute Research Papers EI 2004-38. Erasmus University Rotterdam, Erasmus School of Economics (ESE), Econometric Institute.

Denuit, M. and Lambert, P. (2005). Constraints on concordance measures in bivariate discrete data. *J. Multi. Anal.* **93**(1), 40–57.

Diop, M. L. and Kengne, W. (2017). Testing parameter change in general integer-valued time series. *J. Time Ser. Anal.* pp. n/a–n/a. doi: 10.1111/jtsa.12240.

Doukhan, P., Fokianos, K. and Tjøstheim, D. (2012). On weak dependence conditions for Poisson autoregressions. *Stat. Probab. Lett.* **82**(5), 942–948.

Doukhan, P., Fokianos, K. and Tjøstheim, D. (2013). Correction to " On weak dependence conditions for Poisson autoregressions ". *Stat. Probab. Lett.* **83**(8), 1926–1927.

Doukhan, P. and Kengne, W. (2015). Inference and testing for structural change in general poisson autoregressive models. *Electron. J. Stat.* **9**(1), 1267–1314.

Efron, B. (1986). Double exponential families and their use in generalized linear regression. *J. Am. Stat. Assoc.* **81**(395), 709–721.

Fahrmeir, L. and Tutz, G. (2001). *Models for Multicategorical Responses: Multivariate Extensions of Generalized Linear Models*. Springer New York. pp. 69–137.

Famoye, F. (1993). Restricted generalized poisson regression model. *Commun. Stat.-Theory Methods* **22**(5), 1335–1354. doi: 10.1080/03610929308831089.

Famoye, F., Wulu, J. T. and Singh, K. P. (2004). On the generalized Poisson regression model with an application to accident data. *J. Data Sci.* **2**(2004), 287–295.

Ferland, R., Latour, A. and Oraichi, D. (2006). Integer-valued GARCH process. *J. Time Ser. Anal.* **27**(6), 923–942.

Fokianos, K. and Fried, R. (2010). Interventions in INGARCH processes. *J. Time Ser. Anal.* **31**(3), 210–225.

Fokianos, K. and Fried, R. (2012). Interventions in log-linear Poisson autoregression. *Stat. Model.* **12**(4), 299–322.

Fokianos, K., Gombay, E. and Hussein, A. (2014). Retrospective change detection for binary time series models. *J. Stat. Plan. Infer.* **145**, 102–112.

Fokianos, K., Rahbek, A. and Tjøstheim, D. (2009). Poisson autoregression. *J. Am. Stat. Assoc.* **104**(488), 1430–1439.

Fokianos, K. and Tjøstheim, D. (2012). Nonlinear poisson autoregression. *Ann. Inst. Stat. Math.* **64**(6), 1205–1225.

Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* **10**(4), 605–637.

Franke, J., Kirch, C. and Kamgaing, J. T. (2012). Changepoints in times series of counts. *J. Time Ser. Anal.* **33**(5), 757–770.

Gupta, P. L., Gupta, R. C. and Tripathi, R. C. (1995). Inflated modified power series distributions with applications. *Commun. Stat.-Theory Methods* **24**(9), 2355–2374. doi: 10.1080/03610929508831621.

Gupta, P. L., Gupta, R. C. and Tripathi, R. C. (1996). Analysis of zero-adjusted count data. *Comput. Stat. Data Anal.* **23**(2), 207 – 218.

Heinen, A. (2003). Modelling time series count data: an autoregressive conditional Poisson model. *Available at SSRN 1117187* .

Heinen, A. and Rengifo, E. (2003). Multivariate modelling of time series count data: an autoregressive conditional Poisson model. Technical report. Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Heinen, A. and Rengifo, E. (2007). Multivariate autoregressive modeling of time series count data using copulas. *J. Empir. Financ.* **14**(4), 564–583.

Hudecová, Š. (2013). Structural changes in autoregressive models for binary time series. *J. Stat. Plan. Infer.* **143**(10), 1744–1752.

Jin-Guan, D. and Yuan, L. (1991). The integer-valued autoregressive (INAR(p)) model. *J. Time Ser. Anal.* **12**(2), 129–142.

Jung, R. C., Kukuk, M. and Liesenfeld, R. (2006). Time series of count data: modeling, estimation and diagnostics. *Comput. Stat. Data Anal.* **51**(4), 2350–2364.

Jung, R. C. and Tremayne, A. (2011). Useful models for time series of counts or simply wrong ones?. *AStA-Adv. Stat. Anal.* **95**(1), 59–91.

Kang, J. and Lee, S. (2009). Parameter change test for random coefficient integer-valued autoregressive processes with application to polio data analysis. *J. Time Ser. Anal.* **30**(2), 239–258.

Kang, J. and Lee, S. (2014). Parameter change test for Poisson autoregressive models. *Scand. J. Stat.* **41**(4), 1136–1152.

Kirch, C. and Tadjuidje Kamgaing, J. (2014). Detection of change points in discrete valued time series. *Handbook of discrete valued time series. In: Davis RA, Holan SA, Lund RB, Ravishanker N* .

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* **34**(1), 1–14. doi: 10.1080/00401706.1992.10485228.

Lee, S., Ha, J., Na, O. and Na, S. (2003). The cusum test for parameter change in time series models. *Scand. J. Stat.* **30**(4), 781–796.

Lee, S. and Lee, J. (2015). Parameter change test for nonlinear time series models with GARCH type errors. *J. Korean. Math. Soc.* **52**(23), 503–553.

Lee, S., Lee, Y. and Chen, C. W. (2016a). Parameter change test for zero-inflated generalized Poisson autoregressive models. *Statistics* **50**(3), 1–18.

Lee, S., Tokutsu, Y. and Maekawa, K. (2004). The cusum test for parameter change in regression models with arch errors. *J. Japan Stat. Soc.* **34**(2), 173–188. doi: 10.14490/jjss.34.173.

Lee, Y., Lee, S. and Tjøstheim, D. (2016b). Asymptotic normality and parameter change test for bivariate Poisson INGARCH models. *TEST* pp. 1–18. doi: 10.1007/s11749-016-0510-6.

Li, W. K. (1994). Time series models based on generalized linear models: Some further results. *Biometrics* **50**(2), 506–511.

Liu, H. (2012). Some models for time series of counts. PhD thesis. Columbia University.

McKenzie, E. (2003). Ch. 16. Discrete variate time series. *Handbook of statistics* **21**, 573–606.

Na, O., Lee, J. and Lee, S. (2012). Change point detection in copula armagarch models. *J. Time Ser. Anal.* **33**(4), 554–569. doi: 10.1111/j.1467-9892.2011.00763.x.

Neumann, M. H. (2011). Absolute regularity and ergodicity of Poisson count processes. *Bernoulli* **17**(4), 1268–1284.

Oh, H. and Lee, S. (2017a). On score vector- and residual-based CUSUM tests in ARMA-GARCH models. *Submitted for publication* .

Oh, H. and Lee, S. (2017b). Modified residual CUSUM test for location-scale time series models with heteroscedasticity. *Submitted for publication* .

Pedeli, X. and Karlis, D. (2011). A bivariate INAR(1) process with application. *Stat. Model.* **11**(4), 325–349.

Pedeli, X. and Karlis, D. (2013a). On composite likelihood estimation of a multivariate INAR(1) model. *J. Time Ser. Anal.* **34**(2), 206–220.

Pedeli, X. and Karlis, D. (2013b). On estimation of the bivariate Poisson INAR process. *Commun. Stat.-Simul. Comput.* **42**(3), 514–533.

Puig, P. and Valero, J. (2006). Count data distributions: Some characterizations with applications. *J. Am. Stat. Assoc.* **101**(473), 332–340.

Quoreshi, A. S. (2006). Bivariate time series modeling of financial count data. *Commun. Stat.-Theory Methods* **35**(7), 1343–1358.

Song, J. (2008). Statistical inference for discrete and continuous time stochastic processes: robust estimation and change point analysis. PhD thesis. Seoul National University.

Weiß, C. H. (2008). Thinning operations for modeling time series of counts-a survey. *AStA-Adv. Stat. Anal.* **92**(3), 319–341.

Weiß, C. H. (2009). Modelling time series of counts with overdispersion. *Stat. Method. Appl.* **18**(4), 507–519.

Winkelmann, R. (2008). *Econometric analysis of count data.* Springer Science & Business Media.

Zeger, S. L. and Qaqish, B. (1988). Markov regression models for time series: A quasi-likelihood approach. *Biometrics* **44**(4), 1019–1031.

Zhu, F. (2011). A negative binomial integer-valued GARCH model. *J. Time Ser. Anal.* **32**(1), 54–67. doi: 10.1111/j.1467-9892.2010.00684.x.

Zhu, F. (2012a). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *J. Math. Anal. Appl.* **389**(1), 58 – 71.

Zhu, F. (2012b). Zero-inflated Poisson and negative binomial integer-valued GARCH models. *J. Stat. Plan. Infer.* **142**(4), 826–839. doi: 10.1016/j.jspi.2011.10.002.

Zhu, F. (2012c). Modeling time series of counts with COM-Poisson INGARCH models. *Math. Comput. Model.* **56**(9-10), 191–203. doi: 10.1016/j.mcm.2011.11.069.

Zhu, F. and Wang, D. (2011). Estimation and testing for a Poisson autoregressive model. *Metrika* **73**(2), 211–230. doi: 10.1007/s00184-009-0274-z.

Zhu, R. and Joe, H. (2006). Modelling count data time series with Markov processes based on binomial thinning. *J. Time Ser. Anal.* **27**(5), 725–738. doi: 10.1111/j.1467-9892.2006.00485.x.

국문초록

본 논문에서는 이산형 시계열 모형에서의 모수 변화 검정을 고려하였다. 먼저, 영과잉 일반화된 포아송 (ZIGP) 자기회귀 모형에서의 모수 변화 검정에 대한 문제에 대해 살펴보았다. ZIGP 자기회귀 모형이 정상적이고 에르고딕하며, 조건부 최대우도 추정량의 강일치성과 점근적 정규성을 보였다. 이 결과를 바탕으로 추정량과 잔차를 기반으로 하는 누적합 검정통계량을 만들고, 귀무가설 하에서 각각의 극한 분포가 서로 독립인 브라우니안 브리지의 함수 형태라는 것을 보였다. 모의실험을 통해 검정통계량이 잘 작동함을 확인하고, 호주의 범죄 자료를 이용한 분석을 시행하였다. 두 번째로 이변량 포아송 INGARCH(1,1) 모형에서의 모수 변화 검정에 대한 문제를 고려하였다. 조건부 최대우도추정량이 근사적으로 정규분포를 따른다는 것을 보이고, 이에 기반 하여 추정량과 잔차에 기반을 둔 누적합 통계량을 만들고, 이 통계량들이 각각 귀무가설 하에서 서로 독립인 브라우니안 브리지의 함수로 수렴한다는 것을 보였다. 모의실험을 시행하고, 서울 성수구와 서초구의 자동차 사고 자료를 이용한 분석을 시행하였다. 마지막으로 조건부 분포가 일변량 지수족을 따르는 시계열 모형에서의 모수 변화 검정의 문제를 연구하였다. 점수함수와 잔차 그리고 추정량에 기반을 둔 누적합 통계량을 만들고, 각각이 귀무가설 하에서 서로 독립인 브라우니안 브리지의 함수로 수렴하는 것을 보였다. 이를 바탕으로 몬테 칼로 모의실험을 통해 위의 통계량의 성과를 비교하였다. 그 결과 표준화 된 잔차를 기반으로 하는 검정이 다른 검정보다 뛰어나다는 것을 확인하였다.

주요어 : 이산형 시계열, 영과잉 일반화된 포아송 자기회귀 모형, 정수값을 갖는 GARCH 모형, 모수 변화 검정, 누적합 검정, 브라우니안 브리지로의 약수렴, 이변량 포아송 INGARCH 모형, 지수족, 통계량 비교.

학번 : 2003-20356