Annals of the Institute of Statistical Mathematics (2019) 71:1033–1057
https://doi.org/10.1007/s10463-018-0676-7

CrossMark

# CUSUM test for general nonlinear integer-valued GARCH models: comparison study

**Youngmi Lee**[1] · **Sangyeol Lee**[1]

Received: 9 July 2017 / Revised: 31 May 2018 / Published online: 23 June 2018


**Abstract**
This study considers the problem of testing 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 · Exponential family · Autoregressive models · Parameter change test · CUSUM test · Comparison of tests

## 1 Introduction

Time series models of counts have been intensively studied in recent years, given their widespread applications in many research fields (e.g., economics, finance, environmental science, epidemiology). Integer-valued autoregressive-type (INAR-type) time series models based on a binomial thinning operation appear in the literature as a counterpart of AR models; they have been extensively studied by many researchers. See McKenzie (1985, 2003), Alzaid and Al-Osh (1990), Al-Osh and Aly (1992) and Weiß (2008). Other than the INAR models, different approaches have been taken, such as Poisson AR or nonlinear integer-valued generalized autoregressive conditional heteroscedastic (INGARCH) models. See Heinen (2003), Ferland et al. (2006), Fokianos et al. (2009), Neumann (2011) and Doukhan et al. (2012, 2013).

For conditional distributions, some researchers also considered the use of distributions other than the Poisson distribution. For example, Davis and Wu (2009), Zhu

✉ Sangyeol Lee
sylee@stats.snu.ac.kr

1 Department of Statistics, Seoul National University, Seoul 08826, South Korea

(2011), and Christou and Fokianos (2014) considered negative binomial INGARCH (NB-INGARCH) models, and Zhu (2012a, b) and Lee et al. (2016) considered zero-inflated generalized Poisson INGARCH models. Davis and Liu (2016) have recently extended the Poisson AR model to one-parameter exponential family AR models, called general nonlinear INGARCH models, and established its stationarity and ergodicity, as well as the asymptotic properties of the conditional maximum likelihood estimator (CMLE) under some regularity conditions. Since many time series often experience structural changes in their underlying models, the change point detection problem has been a core issue in the time series context. See Csörgö and Horváth (1997) for a general review. The change point test for integer-valued time series has attracted many researchers. We refer the reader to Kang and Lee (2009), Fokianos and Fried (2010, 2012), Franke et al. (2012), Hudecová (2013), and Fokianos et al. (2014). Kang and Lee (2014) have recently studied the change point test for Poisson AR models, especially two types of CUSUM tests: an estimate-based test that uses the CMLE and a residual-based test. Lee et al. (2016) extended the test to the zero-inflated generalized Poisson AR models. Diop and Kengne (2017) considered an estimate-based CUSUM test for general nonlinear INGARCH models. We also refer the reader to Hudecová et al. (2016) for recent relevant studies. While the aforementioned works are devoted to the retrospective parameter change problem, others considered online detection problem using control charts aimed at an early detection of parameter changes. See Weiß and Testik (2009, 2011), Huh et al. (2017), Kim and Lee (2017), and the papers cited in these articles.

Although the estimate-based CUSUM test generally performs well, the estimate-based test occasionally suffers from severe size distortions and cannot be completely reliable (Kang and Lee 2014; Lee et al. 2016, 2018). 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 is not always satisfactory, and a great power loss can occur, particularly when dealing with a parameter change in conditional locations (Oh and Lee 2018). As an alternative, one can use the score vector-based CUSUM test (Berkes et al. 2004; Oh and Lee 2017), because it might outperform the residual-based CUSUM test in terms of power. This study additionally considers the residual-based CUSUM test using the "standardized" residuals, as doing so can to a great extent enhance the test performance in terms of power; this is seen in the results of our simulation studies. The current study pays a special attention to comparing the performance of the score vector-, (standardized) residual-, and estimate-based CUSUM tests for general nonlinear INGARCH models. For this task, however, we make the effort to derive the limiting null distributions for obtaining critical values, as used in Monte Carlo simulations. Our findings show that among the CUSUM tests studied, the standardized residual-based test performs best.

The remainder of this paper is organized as follows: Sect. 2 introduces the one-parameter exponential family AR models and establishes the asymptotic results for the CMLE and the CUSUM tests based on the score vectors, (standardized) residuals, and estimates. Section 3 discusses a simulation study for comparison. Section 4 provides concluding remarks. Finally, all proofs are provided in "Appendix."

## 2 Main Results

### 2.1 Models and asymptotic properties of CMLE

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

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

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

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

wherein $\eta$ is the natural parameter, $A(\eta)$ and $h(y)$ are known functions, and $A'(\cdot)$ exists and is strictly increasing, and further, $\eta_t = (A')^{-1}(X_t)$. We express $B(\eta) = A'(\eta)$. Then, $B(\eta_t)$ and $B'(\eta_t)$ are the conditional mean and variance of $Y_t$, respectively, and $X_t = B(\eta_t)$. To emphasize the role of $\theta$, we also use notation $X_t(\theta)$ and $\eta_t(\theta)$ to stand for $X_t$ and $\eta_t$.

In what follows, we assume

(**A0**) For all $x, x' \geq 0$ and $y, y' \in \mathbb{N}_0$,

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

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

Davis and Liu (2016) showed that this assumption ensures the strict stationarity and ergodicity of $\{(X_t, Y_t)\}$ and the existence of a measurable function $f_\infty^\theta : \mathbb{N}_0^\infty = \{(n_1, n_2, \ldots), n_i \in \mathbb{N}_0, i = 1, 2, \ldots\} \to [0, \infty)$ such that $X_t(\theta) = f_\infty^\theta(Y_{t-1}, Y_{t-2}, \ldots)$ a.s. The conditional likelihood function of model (1), based on the observations $Y_1, \ldots, Y_n$, 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

$$\widetilde{X}_t(\theta) = f_\theta(\widetilde{X}_{t-1}(\theta), Y_{t-1}), \quad t = 2, 3, \ldots, \quad \widetilde{X}_1(\theta) = \widetilde{X}_1,$$

with an arbitrarily chosen initial random variable $\widetilde{X}_1$. In what follows, $\theta_0$ denotes the true value of $\theta$. We obtain the CMLE of $\theta_0$ 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 $\theta_0$. Further, we use notation $\tilde{\eta}_t = \tilde{\eta}_t(\theta)$ for simplicity.

- (**A1**) $\theta_0$ is an interior point in the compact parameter space $\Theta \in \mathbb{R}^d$.
- (**A2**) For any $\theta \in \Theta$ and $\mathbf{y} \in \mathbb{N}_0^{\infty}$, $f_{\infty}^{\theta}(\mathbf{y}) \geq x_{\theta}^* \in \mathcal{R}(B)$, where $\mathcal{R}(B)$ is the range of $B(\eta)$. Moreover, $x_{\theta}^* \geq x^* \in \mathcal{R}(B)$ for all $\theta$.
- (**A3**) For any $\mathbf{y} \in \mathbb{N}_0^{\infty}$, the mapping $\theta \mapsto f_{\infty}^{\theta}(\mathbf{y})$ is twice continuously differentiable with respect to $\theta$.
- (**A4**) $E\left\{Y_1 \sup_{\theta\in\Theta} |\eta_1(\theta)|\right\} < \infty$.
- (**A5**) If there exists a $t > 1$ such that $X_t(\theta) = X_t(\theta_0)$ a.s., then $\theta = \theta_0$.
- (**A6**) $E\left(\sup_{\theta\in\Theta} X_1^2(\theta)\right) < \infty$.
- (**A7**)

$$E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial X_1(\theta)}{\partial\theta}\right\|^4\right) < \infty \quad \text{and} \quad E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial^2 X_1(\theta)}{\partial\theta\partial\theta^T}\right\|^2\right) < \infty.$$

- (**A8**)

$$E\left[\sup_{\theta\in\Theta}\left\|B'(\eta_1)\left(\frac{\partial\eta_1}{\partial\theta}\cdot\frac{\partial\eta_1}{\partial\theta^T}\right)\right\|\right] < \infty, \ E\left[\sup_{\theta\in\Theta}\left\|(Y_1 - B(\eta_1))\frac{\partial^2\eta_1}{\partial\theta\partial\theta^T}\right\|\right] < \infty.$$

- (**A9**) For all $t$,

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

- (**A10**) For all $t$, $\sup_{\theta\in\Theta}\sup_{0\leq\delta\leq 1} B'((1-\delta)\eta_t + \delta\tilde{\eta}_t) \geq \underline{c}$ for some constant $\underline{c} > 0$.
- (**A11**) For all $t$, $\sup_{\theta\in\Theta} |B'(\tilde{\eta}_t) - B'(\eta_t)| \leq V\rho^t$ a.s.
- (**A12**) For all $t$, $\sup_{\theta\in\Theta} B'(\eta_t)^{-3/2}B''(\eta_t) \leq K$ for some $K > 0$.
- (**A13**) $\nu^T\frac{\partial X_1(\theta)}{\partial\theta} = 0$ a.s. (or equivalently, $\nu^T\frac{\partial\eta_1(\theta)}{\partial\theta} = 0$ a.s.) if and only if $\nu = 0$.

Conditions (**A1**)–(**A5**) and (**A8**) can be found in Davis and Liu (2016). They also derived the asymptotic properties of the CMLE. The proposition below can be proven using Lemma 2 in "Appendix," 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); see the proof in "Appendix."

**Proposition 1** *Suppose that conditions (**A0**)–(**A13**) hold. Then, as $n \to \infty$,*

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

*and*

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

*where*

$$I(\theta_0)=E\left(\frac{\partial \ell_t(\theta_0)}{\partial \theta}\frac{\partial \ell_t(\theta_0)}{\partial \theta^T}\right)=-E\left(\frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta^T}\right)=E\left(B^{'}(\eta_t(\theta_0))\frac{\partial \eta_t(\theta_0)}{\partial \theta}\frac{\partial \eta_t(\theta_0)}{\partial \theta^T}\right)$$

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

### 2.2 Change point test

In this subsection, we study the score vector-, residual-, standardized residual-, and estimate-based CUSUM tests used to assess the hypotheses

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

#### 2.2.1 Score vector-based CUSUM test

The score vector-based CUSUM test is given by

$$T_n^{\text{score}}=\max_{1\le k\le n}\frac{1}{n}\left(\sum_{t=1}^{k}\frac{\partial \tilde{\ell}_t(\hat{\theta}_n)}{\partial \theta}\right)^T \hat{I}_n^{-1}\left(\sum_{t=1}^{k}\frac{\partial \tilde{\ell}_t(\hat{\theta}_n)}{\partial \theta}\right), \tag{2}$$

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 theorem.

**Theorem 1** *Suppose that conditions* (**A0** )–(**A13**) *hold. Then, under* $H_0$*, as* $n\to\infty$*,*

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

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

#### 2.2.2 Residual-based CUSUM test

We consider the two types of residuals: $\epsilon_{t,1}=Y_t-X_t(\theta_0)$ and $\epsilon_{t,2}=(Y_t-X_t(\theta_0))/\sqrt{B'(\eta_t(\theta_0))}$. The former is considered by Franke et al. (2012), Kang and Lee (2014), and Lee et al. (2016, 2018) 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

$$\sup_{0<s<1}\frac{1}{\sqrt{n}\tau_i}\left|\sum_{t=1}^{[ns]}\epsilon_{t,i}-\frac{k}{n}\sum_{t=1}^{n}\epsilon_{t,i}\right|\xrightarrow{w}\sup_{0\le s\le 1}|\mathbf{B}_1^{\circ}(s)|, \tag{3}$$

where $\tau_1^2 = Var(\epsilon_{1,1})$ and $\tau_2^2 = 1$. Note that $\epsilon_{t,i}$ is not observable, but it is possible to compute $\hat{\epsilon}_{t,1} = Y_t - \hat{X}_t$ or $\hat{\epsilon}_{t,2} = (Y_t - \hat{X}_t)/\sqrt{B'(\hat{\eta}_t)}$, where $\hat{X}_t = f_{\hat{\theta}_n}(\hat{X}_{t-1}, Y_{t-1})$, $\hat{\eta}_t = B^{-1}(\hat{X}_t)$ for $t \geq 2$ and $\hat{X}_1$ is an arbitrarily chosen value. We thus consider the tests

$$T_n^{\mathrm{res},i} = \max_{1\leq k\leq n} \frac{1}{\sqrt{n}\hat{\tau}_{n,i}} \left| \sum_{t=1}^{k} \hat{\epsilon}_{t,i} - \frac{k}{n} \sum_{t=1}^{n} \hat{\epsilon}_{t,i} \right|, \tag{4}$$

where $\hat{\tau}_{n,1}^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{\epsilon}_{t,1}^2$ and $\hat{\tau}_{n,2}^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{\epsilon}_{t,2}^2$. Then, we can obtain the following theorem, the proof of which is similar to that of Kang and Lee (2014) and is omitted for brevity.

**Theorem 2** *Suppose that conditions* (**A0**)–(**A13**) *hold. Then, under $H_0$, as $n \to \infty$,*

$$T_n^{\mathrm{res},i} \xrightarrow{w} \sup_{0\leq s\leq 1} |\mathbf{B}_1^{\circ}(s)|, \quad i = 1, 2.$$

In our simulation study (Sect. 3) the following estimate-based CUSUM tests are compared to score vector- and (standardized) residual-based CUSUM tests

$$T_n^{\mathrm{est},1} = \max_{1\leq k\leq n} \frac{k^2}{n} (\hat{\theta}_k - \hat{\theta}_n)^T \hat{I}_n (\hat{\theta}_k - \hat{\theta}_n), \tag{5}$$

where $\hat{\theta}_k$ is the CMLE of $\theta_0$ based on $Y_1, \ldots, Y_k$ and

$$T_n^{\mathrm{est},2} = \max_{v_n\leq k\leq n-v_n} \frac{1}{q^2(k/n)} \frac{k^2(n-k)^2}{n^3} (\hat{\theta}_k - \tilde{\theta}_k)^T \hat{I}_n' (\hat{\theta}_k - \tilde{\theta}_k), \tag{6}$$

where $\tilde{\theta}_k$ are the CMLE of $\theta_0$ based on the observations $Y_{k+1}, \ldots, Y_n$,

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

$q : (0, 1) \to (0, \infty)$ is a non-decreasing/non-increasing weight function on a neighborhood of zero/one such that $\inf_{\eta<\tau<1-\eta} q(\tau) > 0$ for all $0 < \eta < 1/2$, and $\{u_n : n \geq 1\}$ and $\{v_n : n \geq 1\}$ are sequences of integers diverging to $\infty$ with $u_n/n, v_n/n \to 0$ as $n \to \infty$. Note that, under conditions (**A0**)–(**A13**) and $H_0$, $T_n^{\mathrm{est},i}$ converges weakly to $\sup_{0\leq s\leq 1} \|\mathbf{B}_d^{\circ}(s)\|$; see Diop and Kengne (2017).

### 2.3 INGARCH(1,1) models

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

$$Y_t|\mathcal{F}_{t-1} \sim p(y|\eta_t), \quad X_t = \omega + \alpha X_{t-1} + \beta Y_{t-1}, \tag{7}$$

where $X_t = B(\eta_t) = E(Y_t|\mathcal{F}_{t-1})$ and $\theta = (\omega, \alpha, \beta)$ satisfies $\omega > 0, \alpha \geq 0, \beta \geq 0$, and $\alpha + \beta < 1$. The process $\{(X_t, Y_t); t \geq 1\}$ has then a strictly stationary and ergodic solution. To ensure Proposition 1 in this case, (**A1**) can be replaced with the following:

(**A1'**) The true parameter $\theta_0$ lies in a compact neighborhood $\Theta \in \mathbb{R}_+^3$ of $\theta_0$, where

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

for some $\epsilon > 0$.

Note that, by iterating (7),

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

where the initial value is taken as $\tilde{X}_1 = \omega/(1-\alpha)$. Hence, (**A2**) is satisfied, because $X_t(\theta) \geq \omega/(1-\alpha) \geq x^* = \omega_L/(1-\epsilon)$. Concerning (**A3**)–(**A9**) and (**A13**), we refer the reader to Kang and Lee (2014), Davis and Liu (2016), and Diop and Kengne (2017). Below, we summarize some of the most typical examples wherein Conditions (**A10**)–(**A12**) are found to hold—namely Poisson, negative binomial, and binomial distributions.

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

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

In this model, $\eta_t = \log(X_t(\theta))$ and $A(\eta) = e^\eta$. Since $X_t(\theta) \geq \omega_L$, $\tilde{X}_t(\theta) \geq \omega_L$, and $B'(\eta) = e^\eta$ is increasing, (**A10**) holds. Moreover, since $B'(\eta_t) = X_t(\theta)$, (**A11**) is satisfied. Finally, (**A12**) holds, due to (**A10**) and the fact that $B'(\eta) = B''(\eta)$.

***Example 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

$$P(Y_t = k) = \binom{k+r-1}{r-1} (1-p)^k p^r, \qquad k = 0, 1, 2, \ldots.$$

Here, $r$ is assumed to be known. In this model, $\eta_t = \log(X_t(\theta)/(X_t(\theta) + r))$ and $A(\eta) = -r\log(r/(1-e^\eta))$. 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, (**A10**) holds. Next, since $B'(\eta_t) = X_t(\theta)(X_t(\theta) + r)/r$,

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

owing to (**A6**) and Lemma 1 in "Appendix," which in turn implies (**A11**). Finally, (**A12**) 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 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,

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

Note that $\eta_t = \log\left(X_t(\theta)/(m - X_t(\theta))\right)$ and $A(\eta) = m\log(1+e^{\eta})$. Since $p_t \in (0, 1)$, (**A10**) and (**A12**) hold; furthermore, given the fact that $B'(\eta_t) = X_t(\theta)\,(1 - X_t(\theta)/m)$, it can be shown that (**A11**) holds, similar to the case with the NB-INGARCH(1,1) model.

***Remark 1*** Besides the above linear models, one may consider nonlinear models. For example, the threshold AR model was considered in Kang and Lee (2014) (see also Chen and Lee 2016; Davis and Liu 2016). In general, it is challenging to introduce new nonlinear INGARCH models with the flexibility, practicality, and popularity in diverse applications.

## 3 Simulation study

In this section, we report our simulation results and evaluate the performance of the tests proposed in Sect. 2.2. We consider the INGARCH(1,1) model in Sect. 2.3. In this simulation study, we use $n = 300, 500, 1000$ and repetition number 1000. The critical values are obtained through Monte Carlo simulations (cf., Lee et al. 2003) at the nominal level of 0.05 using 10000 repetitions. For $T_n^{\mathrm{est},1}$, $T_n^{\mathrm{est},2}$, and $T_n^{\mathrm{score}}$ in (5), (6), and (2), they are 3.004, and for $T_n^{\mathrm{res},1}$ and $T_n^{\mathrm{res},2}$ in (4), they are 1.353. The $T_n^{\mathrm{est},2}$ 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

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

with $k_L = 20$, instead of (5).

**Table 1** Empirical sizes for Poisson INGARCH(1,1) models at the nominal level 0.05

| $(\alpha,\beta)$ | $n$ | $\omega=1$ | | | | | $\omega=0.3$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $T_n^{\text{est},1}$ | $T_n^{\text{est},2}$ | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ | $T_n^{\text{est},1}$ | $T_n^{\text{est},2}$ | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ |
| (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 |
| (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 |
| (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 |
| (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 |
| (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 |
| (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 |

### 3.1 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 1. As pointed out in Kang and Lee (2014), $T_n^{\text{est},1}$ exhibits severe size distortions when $\alpha+\beta \approx 1$ and $T_n^{\text{est},2}$ behave similarly. On the contrary, $T_n^{\text{score}}$, $T_n^{\text{res},1}$, and $T_n^{\text{res},2}$ 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 Table 2), because the estimate-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

**Table 2** Empirical powers for Poisson INGARCH(1,1) models at the nominal level 0.05 when a parameter change occurs at $t = [n\tau]$

| | | $\tau = 1/3$ | | | $\tau = 1/2$ | | | $\tau = 2/3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ |
| $(\alpha, \beta)$ | $n$ | $\omega = 1$ changes to $\omega' = 0.3$ and $(\alpha, \beta)$ does not change. | | | | | | | | |
| (0.1,03) | 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 |
| (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 |
| (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 |
| (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 |
| (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 |
| (0.4,0.5) | 300 | 0.062 | 0.430 | 0.174 | 0.044 | 0.230 | 0.210 | 0.048 | 0.032 | 0.128 |
| | 500 | 0.226 | 0.754 | 0.606 | 0.074 | 0.718 | 0.696 | 0.120 | 0.076 | 0.540 |
| | 1000 | 0.816 | 0.996 | 0.994 | 0.596 | 0.944 | 1.000 | 0.688 | 0.474 | 1.000 |
| $(\alpha', \beta')$ | $n$ | $(\alpha, \beta) = (0.1, 0.5)$ changes to $(\alpha', \beta')$ when $\omega = 1$. | | | | | | | | |
| (0.1,0.5) | 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 |
| (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 |
| (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 |
| (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 |
| (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 |

many cases are close to 1, but the power becomes smaller when $\alpha + \beta \approx 1$—that is, $(\alpha, \beta) = (0.1, 0.8), (0.4, 0, 5)$. In most cases, among the CUSUM tests, $T_n^{\mathrm{res},2}$ appears to produce the largest powers.

### 3.2 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); see 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 estimate-based tests give rise to severe size distortions, while the others produce no size distortions; furthermore, $T_n^{\mathrm{res},2}$ produces the largest powers (see Tables

**Table 3** Empirical sizes for negative binomial INGARCH(1,1) models at the nominal level 0.05

| $(\alpha, \beta)$ | $n$ | $r = 1$ | | | | | $r = 8$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $T_n^{\mathrm{est},1}$ | $T_n^{\mathrm{est},2}$ | $T_n^{\mathrm{score}}$ | $T_n^{\mathrm{res},1}$ | $T_n^{\mathrm{res},2}$ | $T_n^{\mathrm{est},1}$ | $T_n^{\mathrm{est},2}$ | $T_n^{\mathrm{score}}$ | $T_n^{\mathrm{res},1}$ | $T_n^{\mathrm{res},2}$ |
| (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 |
| (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 |
| (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 |
| (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 |
| (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 |
| (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 4** Empirical powers for negative binomial INGARCH(1,1) models at the nominal level 0.05 when a parameter change occurs at $t = [n/2]$

| | | $r = 1$ | | | $r = 8$ | | |
|---|---|---|---|---|---|---|---|
| | | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ |
| $(\alpha, \beta)$ | $n$ | $\omega = 1$ changes to $\omega' = 0.3$. | | | | | |
| (0.1,0.3) | 300 | 0.648 | 0.500 | 0.968 | 0.662 | 0.886 | 0.994 |
| | 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 |
| (0.1,0.5) | 300 | 0.760 | 0.180 | 0.932 | 0.780 | 0.676 | 0.990 |
| | 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 |
| (0.1,0.8) | 300 | 0.666 | 0.008 | 0.700 | 0.182 | 0.028 | 0.438 |
| | 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 |
| (0.3,0.2) | 300 | 0.558 | 0.412 | 0.968 | 0.510 | 0.904 | 0.998 |
| | 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 |
| (0.3,0.4) | 300 | 0.562 | 0.104 | 0.836 | 0.306 | 0.412 | 0.908 |
| | 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 |
| (0.4,0.5) | 300 | 0.172 | 0.018 | 0.378 | 0.052 | 0.076 | 0.218 |
| | 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 |
| $(\alpha', \beta')$ | $n$ | $(\alpha, \beta) = (0.1, 0.5)$ changes to $(\alpha', \beta')$ when $\omega = 1$. | | | | | |
| (0.1,0.3) | 300 | 0.112 | 0.360 | 0.990 | 0.270 | 0.814 | 1.000 |
| | 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 |
| (0.1,0.8) | 300 | 0.124 | 0.018 | 0.488 | 0.492 | 0.040 | 0.258 |
| | 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 |
| (0.3,0.2) | 300 | 0.186 | 0.504 | 0.974 | 0.364 | 0.920 | 0.992 |
| | 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 |
| (0.3,0.4) | 300 | 0.122 | 0.156 | 0.740 | 0.210 | 0.694 | 0.932 |
| | 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 |
| (0.4,0.5) | 300 | 0.450 | 0.660 | 0.898 | 0.654 | 0.970 | 0.996 |
| | 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 |

3, 4). Overall, our simulation results confirm the validity of the 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.

### 3.3 Binomial INGARCH(1,1) models

We consider the binomial INGARCH(1,1) model

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

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 $m = 10$. Table 5 shows the sizes derived from the tests. As with the two aforementioned cases, the results of the estimate-based tests exhibit severe size distortions. $T_n^{\text{score}}$ has a somewhat larger size whenever the $(\alpha, \beta)$ is small or the sample size is small, whereas the 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 $\tau = 1/3, 1/2, 2/3$:

$$\text{Case 1}: \omega \text{ changes to } \omega' \text{ and } (\alpha, \beta) \text{ does not change.}$$
$$\text{Case 2}: (\alpha, \beta) = (0.1, 0.2) \text{ changes to } (\alpha', \beta'), \text{ and } \omega \text{ does not change.}$$

It appears that the powers of $T_n^{\text{score}}$, $T_n^{\text{res},1}$, and $T_n^{\text{res},2}$ are similar (see Tables 6, 7). In Case 1, the powers are close to 1, except when $m = 1$ and the sample size is small, regardless of $\tau$. In Case 2, the powers are small when $m = 1$ and $\omega$ 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.

## 4 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 through a simulation study. We derived their limiting null distributions under certain conditions. The simulations showed that the score vector- and residual-based CUSUM tests can serve as promising alternative methods to the estimate-based CUSUM tests; in particular, the standardized residual-based CUSUM mostly outperforms the other tests.

**Table 5** Empirical sizes for binomial INGARCH(1,1) models at the nominal level 0.05

| $(\alpha, \beta)$ | $n$ | $T_n^{\text{est},1}$ | $T_n^{\text{est},2}$ | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ | $T_n^{\text{est},1}$ | $T_n^{\text{est},2}$ | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $m = 1$ | | $\omega = 0.1$ | | | | | $\omega = 0.3$ | | | | |
| (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 |
| (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 |
| (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 |
| (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 |
| $m = 5$ | | $\omega = 0.5$ | | | | | $\omega = 1$ | | | | |
| (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 |
| (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 |
| (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 |
| (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 |
| $m = 10$ | | $\omega = 1$ | | | | | $\omega = 3$ | | | | |
| (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 |
| (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 |
| (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 |
| (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 6** Empirical powers for binomial INGARCH(1,1) models at the nominal level 0.05 when $\omega$ changes to $\omega'$ at $t = [n\tau]$ and $(\alpha, \beta)$ does not change

| $(\alpha, \beta)$ | $n$ | $\tau = 1/3$ | | | $\tau = 1/2$ | | | $\tau = 2/3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $T_n^{\text{score}}$ | $T_n^{\text{res.1}}$ | $T_n^{\text{res.2}}$ | $T_n^{\text{score}}$ | $T_n^{\text{res.1}}$ | $T_n^{\text{res.2}}$ | $T_n^{\text{score}}$ | $T_n^{\text{res.1}}$ | $T_n^{\text{res.2}}$ |
| $m = 1$ | | $\omega = 0.1$ changes to $\omega' = 0.3$ | | | | | | | | |
| (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 |
| (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 |
| (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 |
| (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 |
| $m = 5$ | | $\omega = 0.5$ changes to $\omega' = 1$ | | | | | | | | |
| (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 |
| (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 |
| (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 |
| (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 |
| $m = 10$ | | $\omega = 1$ changes to $\omega' = 3$ | | | | | | | | |
| (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 |
| (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 |
| (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 |

**Table 6** continued

| $m = 10$ | | $\omega = 1$ changes to $\omega' = 3$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| (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 7** Empirical powers for binomial INGARCH(1,1) models at the nominal level 0.05 when $(\alpha, \beta) = (0.1, 0.2)$ changes to $(\alpha', \beta')$ at $t = [n/2]$ and $\omega$ does not change

| $(\alpha', \beta')$ | $n$ | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ | $T_n^{\text{score}}$ | $T_n^{\text{res},1}$ | $T_n^{\text{res},2}$ |
|---|---|---|---|---|---|---|---|
| $m = 1$ | | $\omega = 0.1$ | | | $\omega = 0.3$ | | |
| (0.1,0.5) | 300 | 0.264 | 0.142 | 0.070 | 0.894 | 0.816 | 0.840 |
| | 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 |
| (0.4,0.2) | 300 | 0.140 | 0.186 | 0.208 | 0.628 | 0.880 | 0.828 |
| | 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 |
| (0.3,0.4) | 300 | 0.232 | 0.416 | 0.318 | 1.000 | 0.926 | 0.988 |
| | 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 |
| $m = 5$ | | $\omega = 0.5$ | | | $\omega = 1$ | | |
| (0.1,0.5) | 300 | 0.572 | 0.664 | 0.516 | 0.900 | 0.968 | 0.960 |
| | 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 |
| (0.4,0.2) | 300 | 0.528 | 0.764 | 0.790 | 0.800 | 0.964 | 0.958 |
| | 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 |
| (0.3,0.4) | 300 | 0.654 | 0.950 | 0.938 | 0.712 | 0.988 | 0.974 |
| | 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 |
| $m = 10$ | | $\omega = 1$ | | | $\omega = 3$ | | |
| (0.1,0.5) | 300 | 0.846 | 0.940 | 0.868 | 0.946 | 0.994 | 0.992 |
| | 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 |
| (0.4,0.2) | 300 | 0.794 | 0.940 | 0.960 | 0.848 | 0.998 | 0.978 |
| | 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 |
| (0.3,0.4) | 300 | 0.776 | 0.970 | 0.974 | 1.000 | 1.000 | 1.000 |
| | 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 |

**Acknowledgements** We thank the Editor, an AE, and two referees for their careful reading and valuable comments. This research is supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (No. 2018R1A2A2A05019433).

## Appendix : Proofs of Theorems 1 and 2

Before proving Theorems 1 and 2, we prepare some lemmas. In what follows, we use notation $\eta_t^0 = \eta_t(\theta_0)$ and $\eta_t^n = \eta_t(\hat{\theta}_n)$.

**Lemma 1** *Suppose that conditions (**A0**), (**A6**), and (**A10**) hold. Then, we have*

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

***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 (**A10**), we have

$$\begin{aligned}|\tilde{\eta}_t - \eta_t| &= |B^{-1}(\tilde{X}_t(\theta)) - B^{-1}(X_t(\theta))| \le \frac{\omega_1^{t-1}}{B'(\eta_t^*)} |\tilde{X}_1 - X_1(\theta)| \\ &\le \frac{\omega_1^{t-1}}{\underline{c}} |\tilde{X}_1 - X_1(\theta)|,\end{aligned}$$

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 (**A6**), we establish the lemma. □

**Lemma 2** *Suppose that (**A0**)–(**A13**) hold. Then, under $H_0$, as $n \to \infty$,*

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

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

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

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

*where $\theta_n^*$ is the intermediate point $\hat{\theta}_n$ and $\theta_0$.*

***Proof*** (i) It suffices to show that

$$\sup_{\theta\in\Theta}|\tilde{\ell}_t(\theta)-\ell_t(\theta)|\to 0 \ \ a.s. \ \text{ as } \ t\to\infty. \tag{8}$$

Note that, by the mean value theorem, (**A10**), and Lemma 1,

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

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

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

according to (**A6**), we can show that (8) holds.

(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

$$\begin{aligned}&\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^n\frac{\partial\tilde{\ell}_t(\theta_0)}{\partial\theta}-\frac{1}{\sqrt{n}}\sum_{t=1}^n\frac{\partial\ell_t(\theta_0)}{\partial\theta}\right\|\\ &\le\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^n\tilde{U}_t(\theta_0)\left(\frac{\partial\tilde{\eta}_t^0}{\partial\theta}-\frac{\partial\eta_t^0}{\partial\theta}\right)\right\|+\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^n\left(\tilde{U}_t(\theta_0)-U_t(\theta_0)\right)\frac{\partial\eta_t^0}{\partial\theta}\right\|.\end{aligned}\tag{9}$$

Using (**A6**), (**A7**), (**A9**)–(**A11**), Lemma 1, and the following facts: $|\tilde{U}_t(\theta)|\le|U_t(\theta)|+V\rho^t$ and

$$\begin{aligned}\left\|\frac{\partial\tilde{\eta}_t^0}{\partial\theta}-\frac{\partial\eta_t^0}{\partial\theta}\right\| &=\left\|\frac{1}{B'(\tilde{\eta}_t^0)}\frac{\partial\tilde{X}_t(\theta_0)}{\partial\theta}-\frac{1}{B'(\eta_t^0)}\frac{\partial X_t(\theta_0)}{\partial\theta}\right\|\\ &\le\frac{1}{B'(\tilde{\eta}_t^0)}\left\|\frac{\partial\tilde{X}_t(\theta_0)}{\partial\theta}-\frac{\partial X_t(\theta_0)}{\partial\theta}\right\|+\left|\frac{1}{B'(\tilde{\eta}_t^0)}-\frac{1}{B'(\eta_t^0)}\right|\left\|\frac{\partial X_t(\theta_0)}{\partial\theta}\right\|,\end{aligned}$$

we can see that the first term of (9) is $o_P(1)$.

On the other hand, by Lemma 1, the second term of (9) is bounded by

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

which is $o_P(1)$ due to (**A7**) and (**A10**).

(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

$$\begin{aligned}\left\|\frac{\partial^2 \tilde{\ell}_t(\theta)}{\partial\theta\partial\theta^T} - \frac{\partial^2 \ell_t(\theta)}{\partial\theta\partial\theta^T}\right\| &\le \left\|B'(\tilde{\eta}_t)\left(\frac{\partial\tilde{\eta}_t}{\partial\theta} - \frac{\partial\eta_t}{\partial\theta}\right)\frac{\partial\tilde{\eta}_t}{\partial\theta^T}\right\| + \left\|B'(\tilde{\eta}_t)\frac{\partial\eta_t}{\partial\theta}\left(\frac{\partial\tilde{\eta}_t}{\partial\theta^T} - \frac{\partial\eta_t}{\partial\theta^T}\right)\right\| \\ &\quad + \left\|\left\{B'(\tilde{\eta}_t) - B'(\eta_t)\right\}\frac{\partial\eta_t}{\partial\theta}\frac{\partial\eta_t}{\partial\theta^T}\right\| + \left\|\tilde{U}_t(\theta)\left(\frac{\partial^2\tilde{\eta}_t}{\partial\theta\partial\theta^T} - \frac{\partial^2\eta_t}{\partial\theta\partial\theta^T}\right)\right\| \\ &\quad + \left\|\left\{\tilde{X}_t(\theta) - X_t(\theta)\right\}\frac{\partial^2\eta_t}{\partial\theta\partial\theta^T}\right\|. \qquad (10)\end{aligned}$$

Since $B'(\eta_t)\partial\eta_t/\partial\theta = \partial X_t(\theta)/\partial\theta$, the first and second terms of the RHS of (10) converge to 0 $a.s.$ as $t \to \infty$ because of (**A7**), (**A9**), and (**A10**). On the other hand, the fourth and fifth terms converge to 0 $a.s.$ as $t \to \infty$ owing to (**A6**), (**A7**), (**A9**), and Lemma 1. Due to (**A11**), we have

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

Henceforth, the third term converges to 0 $a.s.$ as $t \to \infty$ owing to (**A7**) and (**A10**).

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

**Lemma 3** *Suppose that conditions* (**A0**)–(**A13**) *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 $\theta_0$, it suffices to show that

$$\frac{1}{\sqrt{n}}\max_{1\le k\le n}\left\|\tilde{S}_k(\hat{\theta}_n) - \tilde{S}_k(\theta_0) - \frac{k}{n}\left\{\tilde{S}_n(\hat{\theta}_n) - \tilde{S}_n(\theta_0)\right\}\right\| = o_P(1). \qquad (11)$$

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 $\theta_0$ and $\hat{\theta}_n$. Thus, we have

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

Note that, by Proposition 1 and (iv) in Lemma 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 (**A7**), for some $0 < \gamma < 1/2$,

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

Furthermore, since

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

owing to (iv) in Lemma 2, we can show that $I_n = o_P(1)$. This asserts (11), and the lemma is validated. □

***Proof of Theorem 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 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 (ii) in Lemma 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 3, we obtain

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

This establishes the theorem. □

***Proof of Theorem 2*** We consider $T_n^{\text{res},2}$ only. (The proof of $T_n^{\text{res},1} \xrightarrow{w} \sup_{0\le s\le 1} |\mathbf{B}_1^\circ(s)|$ is similar to that of Kang and Lee (2014).)

We write

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

Noting (3), it suffices to show that

$$\max_{1\le k\le n} \frac{1}{\sqrt{n}} \left| \sum_{t=1}^{k} (\hat{\epsilon}_{t,2} - \epsilon_{t,2}) - \frac{k}{n} \sum_{t=1}^{n} (\hat{\epsilon}_{t,2} - \epsilon_{t,2}) \right| = o_P(1), \tag{12}$$

i.e., for $i = 1, 2, 3$,

$$\max_{1\le k\le n} \frac{1}{\sqrt{n}} \left| \sum_{t=1}^{k} R_{t,i} - \frac{k}{n} \sum_{t=1}^{n} R_{t,i} \right| = o_P(1). \tag{13}$$

Firstly, 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

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

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

Using the mean value theorem with intermediate points $\theta^*_{n,1}$ and $\theta^*_{n,2}$ between $\hat{\theta}_n$ and $\theta_0$, we have

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

where we have used Proposition 1 and (**A7**), (**A10**), and (**A12**). 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

$$\begin{aligned} I_{n,2} &\le \frac{1}{\sqrt{n}} \frac{1}{\underline{c}^{3/2}} \sum_{t=1}^{n} \left| \left( X_t(\theta_0) - X_t(\hat{\theta}_n) \right) \left( B'(\eta_t^n) - B'(\hat{\eta}_t) \right) \right| \\ &\le \sqrt{n} \|\hat{\theta}_n - \theta_0\| \frac{V}{n\underline{c}^{3/2}} \sum_{t=1}^{n} \rho^t \left\| \frac{\partial X_t(\theta^*_{n,1})}{\partial \theta} \right\| = O_p(1) \cdot o_P(1) = o_P(1) \end{aligned}$$

with intermediate point $\theta^*_{n,1}$ between $\hat{\theta}_n$ and $\theta_0$, due to (**A7**), (**A10**), and (**A11**). Furthermore, note that $|\hat{X}_t - X_t(\hat{\theta}_n)| \le V\rho^t$ a.s. since owing to (**A0**),

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

Then, by using this and (**A10**),

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

Thus, (13) holds for $i = 1$.

Secondly, we express

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

where

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

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

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

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

Similarly to the case of $I_{n,2}$, we have

$$\begin{aligned}II_{n,1}&\le\frac{2}{\sqrt{n}}\sum_{t=1}^{n}\left|\epsilon_{t,1}\left(\frac{1}{\sqrt{B'(\hat{\eta}_t)}}-\frac{1}{\sqrt{B'(\eta_t^n)}}\right)\right|\\&\le\frac{1}{\sqrt{n}}\frac{1}{\underline{c}^{3/2}}\sum_{t=1}^{n}\left|\epsilon_{t,1}(B'(\hat{\eta}_t)-B'(\eta_t^n))\right|\le\frac{V}{\sqrt{n}}\frac{1}{\underline{c}^{3/2}}\sum_{t=1}^{n}|\epsilon_{t,1}|\rho^t=o_P(1).\end{aligned}$$

Using Taylor's theorem, we have

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

where $\theta_n^*$ is an intermediate point between $\hat{\theta}_n$ and $\theta_0$, and $\zeta_t(\theta)=B''(\eta_t)B'(\eta_t)^{-3/2}\frac{\partial\eta_t}{\partial\theta}$, so that

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

where

$$II'_{n,2}=\sqrt{n}\|\hat{\theta}_n-\theta_0\|\max_{1\le k\le n}\frac{k}{n}\left|\frac{1}{k}\sum_{t=1}^{k}\epsilon_{t,1}\zeta_t(\theta_0)-\frac{1}{n}\sum_{t=1}^{n}\epsilon_{t,1}\zeta_t(\theta_0)\right|,$$

$$II''_{n,2}=\sqrt{n}\|\hat{\theta}_n-\theta_0\|\frac{2}{n}\sum_{t=1}^{n}|\epsilon_{t,1}|\left\|\zeta_t(\theta_n^*)-\zeta_t(\theta_0)\right\|.$$

Since $\{\epsilon_{t,1}\zeta_t(\theta_0)\}$ is ergodic, $\sqrt{n}\|\hat{\theta}_n-\theta_0\|=O_P(1)$, and $E|\epsilon_{t,1}|\|\zeta_t(\theta_0)\|<\infty$ owing to (**A6**), (**A7**), (**A10**), and (**A12**), we have $II'_{n,2}=o_P(1)$. Moreover, because

$$II''_{n,2}\le\sqrt{n}\|\hat{\theta}_n-\theta_0\|\frac{2}{n}\sum_{t=1}^{n}|\epsilon_{t,1}|\sup_{\|\theta-\theta_0\|\le\|\hat{\theta}_n-\theta_0\|}\|\zeta_t(\theta)-\zeta_t(\theta_0)\|$$

and $E\sup_{\theta\in\Theta}\|\zeta_t(\theta)\|^2 < \infty$ owing to (**A7**), (**A10**), and (**A12**), using (**A3**), (**A6**), and the dominated convergence theorem, we can have $II''_{n,2} = o_P(1)$ (cf., Proposition 5 in Lee et al. (2016)). Thus, (13) holds for $i = 2$.

Finally, using Taylor's theorem, we have

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

for some $\theta_n^*$ between $\hat{\theta}_n$ and $\theta_0$. Then, similarly to the case of $II_{n,2}$, we can show that (13) holds for $i = 3$. Hence, (12) is verified. □

## References

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

Alzaid, A., Al-Osh, M. (1990). An integer-valued pth-order autoregressive structure (INAR(p)) process. *Journal of Applied Probability*, *27*, 314–324.

Berkes, I., Horvth, L., Kokoszka, P. (2004). Testing for parameter constancy in GARCH(p, q) models. *Statistics & Probability Letters*, *70*(4), 263–273.

Chen, C. W., Lee, S. (2016). Generalized Poisson autoregressive models for time series of counts. *Computational Statistics & Data Analysis*, *99*, 51–67.

Christou, V., Fokianos, K. (2014). Quasi-likelihood inference for negative binomial time series models. *Journal of Time Series Analysis*, *35*(1), 55–78.

Csörgö, M., Horváth, L. (1997). *Limit theorems in change-point analysis* (Vol. 18). New York: Wiley.

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

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

Diop, M. L., Kengne, W. (2017). Testing parameter change in general integer-valued time series. *Journal of Time Series Analysis*, *38*(6), 880–894.

Doukhan, P., Fokianos, K., Tjøstheim, D. (2012). On weak dependence conditions for Poisson autoregressions. *Statistics & Probability Letters*, *82*(5), 942–948.

Doukhan, P., Fokianos, K., Tjøstheim, D. (2013). Correction to "On weak dependence conditions for Poisson autoregressions". *Statistics & Probability Letters*, *83*(8), 1926–1927.

Ferland, R., Latour, A., Oraichi, D. (2006). Integer-valued GARCH process. *Journal of Time Series Analysis*, *27*(6), 923–942.

Fokianos, K., Fried, R. (2010). Interventions in INGARCH processes. *Journal of Time Series Analysis*, *31*(3), 210–225.

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

Fokianos, K., Rahbek, A., Tjøstheim, D. (2009). Poisson autoregression. *Journal of American Statistical Association*, *104*(488), 1430–1439.

Fokianos, K., Gombay, E., Hussein, A. (2014). Retrospective change detection for binary time series models. *Journal of Statistical Planning & Inferences*, *145*, 102–112.

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

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

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

Hudecová, Š. (2013). Structural changes in autoregressive models for binary time series. *Journal of Statistical Planning & Inferences*, *143*(10), 1744–1752.

Hudecová, Š., Hudecová, M., Meintanis, S. (2016). Change detection in INARCH time series of counts. In R. Cao, W. Gonzalez Manteiga, J. Romo (Ed.), *Nonparametric Statistics* (Vol. 175, pp. 47–58).

Huh, J., Kim, H., Lee, S. (2017). Monitoring parameter shift with Poisson integer-valued GARCH models. *Journal of Statistical Computation & Simulation*, *87*(9), 1754–1766.

Kang, J., Lee, S. (2009). Parameter change test for random coefficient integer-valued autoregressive processes with application to polio data analysis. *Journal of Time Series Analysis*, *30*(2), 239–258.

Kang, J., Lee, S. (2014). Parameter change test for Poisson autoregressive models. *Scandinavian Journal of Statistics*, *41*(4), 1136–1152.

Kim, H., Lee, S. (2017). On first-order integer-valued autoregressive process with Katz family innovations. *Journal of Statistical Computation & Simulation*, *87*(3), 546–562.

Lee, S., Lee, J. (2015). Parameter change test for nonlinear time series models with GARCH type errors. *Journal of Korean Mathematical Society*, *52*(23), 503–553.

Lee, S., Ha, J., Na, O., Na, S. (2003). The cusum test for parameter change in time series models. *Scandinavian Journal of Statistics*, *30*(4), 781–796.

Lee, S., Tokutsu, Y., Maekawa, K. (2004). The cusum test for parameter change in regression models with ARCH errors. *Journal of Japan Statistical Society*, *34*(2), 173–188.

Lee, S., Lee, Y., Chen, C. W. (2016). Parameter change test for zero-inflated generalized Poisson autoregressive models. *Statistics*, *50*(3), 1–18.

Lee, Y., Lee, S., Tjøstheim, D. (2018). Asymptotic normality and parameter change test for bivariate Poisson INGARCH models. *TEST*, *27*(1), 52–69.

McKenzie, E. (1985). Some simple models for discrete variate time series1. *Journal of American Water Resource Association*, *21*(4), 645–650.

McKenzie, E. (2003). Ch. 16. Discrete variate time series. *Handbook of statistics*, *21*, 573–606.

Neumann, M. H. (2011). Absolute regularity and ergodicity of Poisson count processes. *Bernoulli*, *17*(4), 1268–1284.

Oh, H., Lee, S. (2017). *On score vector- and residual-based CUSUM tests in ARMA-GARCH models*. Statistical Methods & Applications: Online published.

Oh, H., Lee, S. (2018). Modified residual CUSUM test for location-scale time series models with heteroscedasticity. *Accepted*.

Weiß, C. H. (2008). Thinning operations for modeling time series of counts-a survey. *AStA-Advances Statistical Analysis*, *92*(3), 319–341.

Weiß, C. H., Testik, M. C. (2009). Cusum monitoring of first-order integer-valued autoregressive processes of Poisson counts. *Journal of Quality Technology*, *41*(4), 389–400.

Weiß, C. H., Testik, M. C. (2011). The Poisson INAR(1) CUSUM chart under overdispersion and estimation error. *IIE Transactions*, *43*(11), 805–818.

Zhu, F. (2011). A negative binomial integer-valued GARCH model. *Journal of Time Series Analysis*, *32*(1), 54–67.

Zhu, F. (2012a). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *Journal of Mathematical Analysis & Applications*, *389*(1), 58–71.

Zhu, F. (2012b). Zero-inflated Poisson and negative binomial integer-valued GARCH models. *Journal of Statistical Planning & Inferences*, *142*(4), 826–839.