Statistica Sinica **32** (2022), 2241-2263
doi:https://doi.org/10.5705/ss.202020.0357

TEST FOR CONDITIONAL VARIANCE OF INTEGER-VALUED TIME SERIES

Yuichi Goto and Kou Fujimori

Kyushu University and Shinshu University

Abstract: We investigate a test for the conditional variance of stationary and ergodic integer-valued time series. This hypothesis testing problem is motivated by the fact that the form of the conditional variance of the process is determined by the conditional distribution and the conditional mean. First, we estimate the unknown parameters of the intensity function using an M-estimator and prove strong consistency and asymptotic normality. Second, we show that the proposed test has asymptotic size α and is consistent. Finally, we discuss the nontrivial power of the proposed test for the local alternative. The proposed test statistic can be applied to various problems, such as specification tests for intensity functions, tests for overdispersion and underdispersion, and goodness-of-fit tests for ergodic and stationary integer-valued time series. A simulation study illustrates the finite-sample performance of the proposed test. Lastly, in a real-data application, we analyze the number of patients with Escherichia coli in Germany.

Key words and phrases: Conditional variance, integer-valued time series, intensity.

1. Introduction

Integer-valued time series are garnering increasing attention in several fields, including analyses of financial data and the number of patients with infectious diseases, among others. One of the most fundamental integer-valued time series is the Poisson process, the conditional distribution of which, given past information, is the Poisson distribution. Based on the Poisson process, we can develop various statistical models, such as the Poisson integer-valued AR model of order p, or, INAR(p), and Poisson integer-valued GARCH model of order p and q, or, INGARCH(p, q). Franke (2010), Neumann (2011), and Doukhan, Fokianos and Tjøstheim (2012) have investigated the stability of these models of Poisson processes. In addition to the Poisson distribution, the negative binomial (NB) distribution is also popular for constructing statistical models of integer-valued time series; see, for examle, Davis and Wu (2009), Zhu and Joe (2010) and Christou and Fokianos (2014).

Corresponding author: Yuichi Goto, Department of Mathematical Sciences, Faculty of Mathematics, Kyushu University, Fukuoka, 819-0395, Japan. E-mail: yuichi.goto@math.kyushu-u.ac.jp.

For spatial point processes or multidimensional count processes, we are interested in the second-order moment functions of the point processes. We can detect whether the observed point pattern is a Poisson, clustering point process, and repulsive point process. That is, the observed points tend to form clusters by observing the second-order moment functions, called the K-function and the pair correlation function, if the process is stationary. Therefore, it is possible to deal with the goodness-of-fit test for a multidimensional stationary Poisson process by using such second-order moment functions of the point processes; see, for example, Heinrich (1991). From this perspective, we can see that the second-order moments of point processes may be essential to constructing the various models of these processes.

Inspired by such research, we focus on the second-order moments, in particular, the variances of the conditional distributions of one-dimensional integer-valued time series. We suppose that the stationary and ergodic integer-valued time series $\{Z_t\}_{t\in\mathbb{Z}}$ on the probability space $(\Omega, \mathcal{F}, P)$ satisfies the following condition:

$$\mathrm{E}(Z_t|\mathcal{F}_{t-1}) = \lambda_t, \quad \mathrm{Var}(Z_t|\mathcal{F}_{t-1}) = \kappa(\lambda_t), \quad t \in \mathbb{Z}, \tag{1.1}$$

where $\{\mathcal{F}_t\}_{t\in\mathbb{Z}}$ is a filtration defined by

$$\mathcal{F}_t = \sigma(Z_s,\ s \leq t), \quad t \in \mathbb{Z},$$

λ_t is an $\mathcal{F}_{t-1}$ measurable random variable, and κ is some function or functional of λ_t. Such models are mentioned by, for example, Aknouche, Bendjeddou and Touche (2018). If the intensity includes unknown parameters, that is, $\lambda_t = \lambda_t(\boldsymbol{\theta})$, with some unknown parameter $\boldsymbol{\theta}$, we can consider parametric models of integer-valued time series. Our model includes various types of integer-valued time series. Actually, for Poisson processes, we can take $\kappa(\lambda_t) = \lambda_t$, and for NB processes, we can take $\kappa(\lambda_t) = \lambda_t(r + \lambda_t)/r$, with a positive parameter r. More generally, if the conditional distribution of $\{Z_t\}$, given past information, is a member of the one-parameter exponential family, that is, $Z_t|\mathcal{F}_{t-1} \sim p(z|\eta_t)$, where $p(z|\eta_t) = \exp(\eta_t z - A(\eta_t))h(z)$, for $z \geq 0$, with an $\mathcal{F}_{t-1}$ measurable random variable η_t, a known function $A(\cdot)$, which is twice differentiable, and a known function $h(\cdot)$, we have that $\mathrm{E}(Z_t|\mathcal{F}_{t-1}) = B(\eta_t)$ and $\mathrm{Var}(Z_t|\mathcal{F}_{t-1}) = B'(\eta_t)$, where $B(\cdot)$ is the first derivative of $A(\cdot)$. Therefore, our model includes such cases, because we can take $\lambda_t = B(\eta_t)$ and the differential operator as κ.

We use an M-estimation method to estimate the unknown parameters for parametric models of integer-valued time series given by (1.1). Davis and Liu (2016) investigated the maximum likelihood estimator for integer-valued time series with conditional distributions that belong to the one-parameter exponential family, as well as and its asymptotic behavior. Poisson and NB quasi maximum likelihood estimations for parametric models are discussed by, for example, Ahmad and Francq (2016) and Aknouche, Bendjeddou and Touche (2018), respectively. Moreover, Aknouche and Francq (2021b) proposed weighted least square estimators for various models. They also proved that the estimator achieves asymptotic efficiency under appropriate conditions.

We also investigate the asymptotic behavior of the general M-estimator, including the (quasi) maximum likelihood estimator, the least square estimator, and the weighted least square estimator.

We propose a hypothesis testing problem for κ, which determines the conditional variance of the integer-valued time series. The testing problem is given as follows:

$$H_0: \ \kappa = \kappa_0, \quad H_1: \ \kappa \neq \kappa_0. \tag{1.2}$$

Considering this test, we can detect, for example, whether the conditional distribution of the observed process is Poisson or NB, among others, for the appropriate function or functional κ_0. We construct the test based on the second-order moment, and derive the asymptotic distribution under the null hypothesis and the consistency of this test. Without making assumptions on the conditional distributions, we can apply the proposed test statistics to various problems investigated in the existing literature. For example, they can be applied to the specification test discussed by Neumann (2011), Fokianos and Neumann (2013), and Leucht and Neumann (2013) for the Poisson INGARCH$(1,1)$, and to the work of Schweer (2016) on the first-order Markov chain models.

The test for overdispersion is also an important example. Weiß, Homburg and Puig (2019) dealt with the overdispersion problem for INAR(1) processes. We consider this problem as a special case of the testing problem provided in (1.2). Finally, we consider the goodness-of-fit test based on our test statistics defined for the testing problem (1.2). Goodness-of-fit tests for integer-valued time series have been investigated intensively . For instance, Meintanis and Karlis (2014) and Hudecová, Hušková and Meintanis (2015) proposed goodness-of-fit tests based on the joint probability generating function for INAR(1) and INARCH(1,1) models, respectively. Unlike their approach, we deal with goodness-of-fit tests as a special case of our testing problem (1.2) under stationarity and the condition that the underlying process belongs to our model.

In summary, this study contributes to the literature in three ways. First, our theory enables us to deal with several problems simultaneously, such as goodness-of-fit tests, specification tests for the intensity functions, and tests for overdispersion and underdispersion. Second, our model encompasses the nonlinear INGARCH(p,q) model. Third, we need not specify the underlying conditional distribution.

The remainder of this paper is organized as follows. In Section 2, we introduce our fundamental setups for parametric models, hypothesis testing problems, test statistics, and some regularity conditions. The main theoretical results are presented in Section 3. We propose the M-estimator and derive the asymptotic behavior of the estimator under appropriate conditions in Subsection 3.1. We also prove that the proposed test statistics are asymptotically normal under the null and the consistency of the test in Subsection 3.2. The nontrivial power of the proposed test is clarified for the local alternative. We provide applications of the proposed test statistics in Section 4. This section shows that our test can be used for the specification test, detection of overdispersion, and goodness-of-fit tests. Section 5 illustrates the finite-sample performance of the test statistics. In Section 6, we analyze the number of patients with Escherichia coli in Germany. All proofs, additional examples of the goodness-of-fit test, and expressions of the higher moments for several distributions are available in the Supplementary Material.

Hereafter, for every $\boldsymbol{v} \in \mathbb{R}^d$ and for $d \in \mathbb{N}$, we write $\|\boldsymbol{v}\|_{\ell_q} := (\sum_{i=1}^d |v_i|^q)^{1/q}$. We use the symbol $\top$ for the transpose of vectors and matrices. For a smooth function $f : \mathbb{R}^d \to \mathbb{R}$, we write the gradient and Hessian of f by $\partial/\partial\boldsymbol{x} f(\boldsymbol{x}) := (\partial/\partial x_1 f(\boldsymbol{x}), \ldots, \partial/\partial x_d f(\boldsymbol{x}))^\top$ and $\partial^2/(\partial\boldsymbol{x}\partial\boldsymbol{x}^\top) f(\boldsymbol{x}) := \left(\partial^2/(\partial x_i \partial x_j) f(\boldsymbol{x})\right)_{1\le i,j\le d}$, respectively. For a random sequence $\{X_n\}$ and a random variable X, $X_n \to^p X$ as $n \to \infty$ denotes the convergence in probability, and $X_n \Rightarrow X$ as $n \to \infty$ denotes the convergence in distribution.

2. Settings

Let $\{Z_t\}$ be an integer-valued or nonnegative time series on the probability space $(\Omega, \mathcal{F}, P)$ with conditional expectation, for any $t \in \mathbb{Z}$,

$$\mathrm{E}\left(Z_t | \mathcal{F}_{t-1}\right) := \lambda(Z_{t-1}, Z_{t-2}, \ldots; \boldsymbol{\theta}_0), \tag{2.1}$$

where $\mathcal{F}_{t-1}$ is the σ-field generated by $\{Z_s, s \le t-1\}$, λ is a known measurable intensity function on $[0,\infty)^\infty \times \mathbb{R}^d$ to $(\delta, +\infty)$, for some $\delta > 0$, and $\boldsymbol{\theta}_0 \in \mathbb{R}^d$ is an unknown parameter. Assuming that the observed stretch $\{Z_1, \ldots, Z_n\}$ is available, we define, for $t \in \mathbb{N} \cup \{0\}$,

$$\lambda_t(\boldsymbol{\theta}) := \lambda(Z_{t-1}, Z_{t-2}, \ldots; \boldsymbol{\theta}), \quad \tilde{\lambda}_t(\boldsymbol{\theta}) := \lambda(Z_{t-1}, Z_{t-2}, \ldots, Z_1, \mathbf{x_0}; \boldsymbol{\theta}),$$

where $\mathbf{x_0} \in [0,\infty)^\infty$ is an initial parameter. Here, $\tilde{\lambda}_t(\boldsymbol{\theta})$ plays a role as a proxy for $\lambda_t(\boldsymbol{\theta})$. Examples of $\mathbf{x_0}$ are given in Remark 2. Let $\hat{\boldsymbol{\theta}}_n$ be an estimator of $\boldsymbol{\theta}_0$ endowed with strong consistency and $\sqrt{n}$-consistency. More precisely, $\hat{\boldsymbol{\theta}}_n$ is satisfied with the following two conditions:

$$\hat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0 \quad \text{a.s. as } n \to \infty \text{ and } \sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) = O_p(1) \text{ as } n \to \infty. \tag{2.2}$$

The construction of such estimators is described in Section 3.

The conditional variance is given by $v_t := \mathrm{Var}\left(Z_t|\mathcal{F}_{t-1}\right) = \mathrm{E}\left(Z_t^2|\mathcal{F}_{t-1}\right) - \lambda_t^2(\boldsymbol{\theta}_0)$. If the conditional distribution of $\{Z_t\}$ follows a Poisson distribution, an NB distribution with parameters r and $r/(r+\lambda_t(\boldsymbol{\theta}_0))$, or an exponential distribution with parameter $1/\lambda_t(\boldsymbol{\theta}_0)$, then the conditional variance is given by $v_t = \lambda_t(\boldsymbol{\theta}_0)$, $\lambda_t(\boldsymbol{\theta}_0)(r+\lambda_t(\boldsymbol{\theta}_0))/r$, or $\lambda_t^2(\boldsymbol{\theta}_0)$, respectively. Thus, the conditional variance can be denoted as $v_t = \kappa(\lambda_t(\boldsymbol{\theta}_0))$, where κ is a measurable function on $[\delta,\infty)$ to $(0,\infty)$. More generally, κ can be some functional. However, for simplicity, we suppose that κ is some function.

In this paper, we discuss the testing problem with the null hypothesis that the conditional variance takes a specific form. More precisely, the null and alternative hypotheses are

$$H_0 : \ \kappa = \kappa_0, \quad H_1 : \ \kappa \neq \kappa_0,$$

where κ_0 is a measurable function. We propose the following test statistic T_n, which can be calculated using the observations $\{Z_t\}_{1\le t\le n}$, for every $n \in \mathbb{N}$:

$$T_n := \frac{1}{\sqrt{M_n}} \sum_{t=1}^{M_n} \left\{ (Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)) \right\}, \tag{2.3}$$

where $\{M_n\}_{n\in\mathbb{N}}$ is an $\mathbb{N}$-valued sequence with $0 < M_n \le n$, and $M_n \to \infty$ as $n \to \infty$. This statistic is motivated by the fact that, under the null H_0, the sequence $\left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0)) : t \in \mathbb{Z}\right\}$ is a martingale difference.

Remark 1. One may think that it is better to use

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\{ (Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)) \right\} \tag{2.4}$$

instead of (2.3). However, we can show that the difference between (2.4) and

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left\{(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))\right\}$$

is not asymptotically negligible as $n\to\infty$, and the asymptotic distribution of the test statistic under the null hypothesis depends on that of the estimator $\hat{\boldsymbol{\theta}}_n$. Hence, we introduce the sequence M_n, and discuss the asymptotic null distributions of the test statistics in the following two cases: (a) $M_n=o(n)$ and (b) $M_n=n$.

We make the following assumptions.

Assumption 1.

(A0) $\{Z_t\}$ *is strictly stationary and ergodic.*

(A1) *There exists a generic positive and integrable random variable* V *and a constant* ρ *such that* $0<\rho<1$,

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\tilde{\lambda}_t(\boldsymbol{\theta})-\lambda_t(\boldsymbol{\theta})\right|\leq V\rho^t \quad a.s.,$$
$$\text{and}\ \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|\kappa_0(\tilde{\lambda}_t(\boldsymbol{\theta}))-\kappa_0(\lambda_t(\boldsymbol{\theta}))|\leq V\rho^t \quad a.s.$$

(A2) $\lambda_t(\boldsymbol{\theta})$ *is differentiable with respect to* $\boldsymbol{\theta}$ *and* κ *is differentiable.*

(M1) *The random variables* Z_t^4, $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\lambda_t^4(\boldsymbol{\theta})$, $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\kappa_0^2(\lambda_t(\boldsymbol{\theta}))$, $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\kappa_0'}^2(\lambda_t(\boldsymbol{\theta}))$, $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|\partial\lambda_t(\boldsymbol{\theta})/\partial\theta_i|^4$, $\left|(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda(\boldsymbol{\theta}_0))\right|^{2+\delta}$, $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|\ell'(Z_t,\lambda_t(\boldsymbol{\theta}))|^4$, *and* $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}|\ell''(Z_t,\lambda_t(\boldsymbol{\theta}))|^2$ *are integrable.*

Remark 2. The broad class of integer-valued and nonnegative time series satisfies (A0) and (A1). We consider the family of distributions $\{Z_\xi:\xi\in\Xi\}$ with mean ξ. The family satisfies the *stochastic equal mean order property* if, for $\xi\leq\xi'$ and any $x\in\mathbb{R}$, $\mathrm{P}(Z_\xi>x)\leq\mathrm{P}(Z_{\xi'}>x)$. By Aknouche and Francq (2021a, Thm. 3.3), given that the nonlinear INGARCH(p,q) model satisfies the contractive condition of the intensity function, the summation of its coefficients is less than one, and the process satisfies the stochastic equal mean order property, there exists a strictly stationary and ergodic solution. More precisely, nonnegative time series $\{Z_t\}$, such that

$$\mathrm{E}\left(Z_t|\mathcal{F}_{t-1}\right):=\lambda(Z_{t-1},\ldots,Z_{t-p},\lambda_{t-1},\ldots,\lambda_{t-q}), \tag{2.5}$$

with the stochastic equal mean order property and the contractive condition, for

$z_1, \ldots, z_p, w_1, \ldots, w_q \in \mathbb{R}$,

$$\begin{aligned} &\left|\lambda(z_1, \ldots, z_p, w_1, \ldots, w_q) - \lambda(z_1', \ldots, z_p', w_1', \ldots, w_q')\right| \\ &\leq \sum_{i=1}^{p} \alpha_i \left|z_i - z_i'\right| + \sum_{j=1}^{q} \beta_j \left|w_j - w_j'\right|, \end{aligned} \tag{2.6}$$

and $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$, there exists a strictly stationary and ergodic solution $\{Z_t\}$. For example, the one-parameter exponential family Davis and Liu (2016), autoregressive conditional duration model, additive duration models, and many zero-inflated distributions satisfy the stochastic equal mean order property (See Aknouche and Francq (2021a) for details). Under the above conditions, we can show that $\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} |\tilde{\lambda}_t(\boldsymbol{\theta}) - \lambda_t(\boldsymbol{\theta})| \leq V\rho^t$ a.s. by Aknouche and Francq (2021a, Lemma A.1). In practice, one needs to choose an initial value $\mathbf{x_0}$. Doukhan and Kengne (2015) put $Z_t = 0$ for all $t \leq 0$ to calculate the $\tilde{\lambda}_t$. Ahmad and Francq (2016) also give examples for INGARCH(1,1) models. In our simulation, we use $Z_0, \tilde{\lambda}_0(\boldsymbol{\theta}) = \sum_{t=1}^n Z_t/n$ and $\partial\lambda_0(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \mathbf{0}$ for the INGARCH(1,1) model. Note that the effect of the initial value is asymptotically negligible by making use of Assumption (A1).

3. Main Theorems

In this section, we present the main results.

3.1. Asymptotic behavior of the estimator for the parameter

In this subsection, we briefly review the asymptotic behavior of the M-estimator, which is essentially developed by Ahmad and Francq (2016), Aknouche, Bendjeddou and Touche (2018), Aknouche and Francq (2021a), and Aknouche and Francq (2021b).

Hereafter, we assume that the estimator $\hat{\boldsymbol{\theta}}_n$ is defined as the following M-estimator:

$$\hat{\boldsymbol{\theta}}_n := \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmax}}\, \tilde{L}_n(\boldsymbol{\theta}), \quad \tilde{L}_n(\boldsymbol{\theta}) := \frac{1}{n}\sum_{t=1}^{n} \ell(Z_t,\tilde{\lambda}_t(\boldsymbol{\theta})), \tag{3.1}$$

where $\ell(\cdot,\cdot)$ is a measurable function. To derive the asymptotic behavior of $\hat{\boldsymbol{\theta}}_n$, we impose the following conditions.

Assumption 2.

(B1) *The function ℓ is almost surely continuous with respect to the second component, and $\lambda_t(\boldsymbol{\theta})$ is almost surely continuous with respect to $\boldsymbol{\theta}$.*

(B2) *It holds that*

$$\left|\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \ell(Z_t,\tilde{\lambda}_t(\boldsymbol{\theta})) - \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \ell(Z_t,\lambda_t(\boldsymbol{\theta}))\right| \to 0 \quad a.s.\ \ as\ t \to \infty.$$

(B3) *It holds that*

$$\mathrm{E}\left(\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right) < \infty.$$

(B4) *The function* $\mathrm{E}\ell(Z_t,\lambda_t(\boldsymbol{\theta}))$ *with respect to* $\boldsymbol{\theta}$ *has a unique maximum at* $\boldsymbol{\theta}_0$*.*

(B5) *The parameter space* $\boldsymbol{\Theta}$ *is a compact set.*

(C6) *The function* ℓ *is twice continuously differentiable with respect to the second component, and* $\lambda_t(\boldsymbol{\theta})$ *is twice continuously differentiable with respect to* $\boldsymbol{\theta}$*.*

(C7) *The following conditions hold true:*

$$\left\|\ell'(Z_t,\tilde{\lambda}_t(\boldsymbol{\theta}))\left(\frac{\partial}{\partial\boldsymbol{\theta}}\tilde{\lambda}_t(\boldsymbol{\theta}) - \frac{\partial}{\partial\boldsymbol{\theta}}\lambda_t(\boldsymbol{\theta})\right)\right\|_{\ell_1} = O(t^{-1/2-\delta}) \quad a.s.\ \ as\ n \to \infty,$$

and, as $n \to \infty$*,*

$$\left\|\frac{\partial}{\partial\boldsymbol{\theta}}\lambda_t(\boldsymbol{\theta})\left(\ell'(Z_t,\tilde{\lambda}_t(\boldsymbol{\theta})) - \ell'(Z_t,\lambda_t(\boldsymbol{\theta}))\right)\right\|_{\ell_1} = O(t^{-1/2-\delta}) \quad a.s.,$$

where $\delta > 0$*.*

(C8) *There exists a neighborhood* $V(\boldsymbol{\theta}_0)$ *of* $\boldsymbol{\theta}_0$ *such that*

$$\mathrm{E}\left(\sup_{\boldsymbol{\theta}\in V(\boldsymbol{\theta}_0)}\left|\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ell(Z_t,\lambda_t(\boldsymbol{\theta}))\right|\right) < \infty, \quad for\ i,j = 1,\dots,d.$$

(C9) *It holds for every* $t \in \mathbb{Z}$ *that*

$$\mathrm{E}\left(\ell'(Z_t,\lambda_t(\boldsymbol{\theta}_0))\middle|\mathcal{F}_{t-1}\right) = 0.$$

(C10) *For every* $i,j = 1,\dots,d$*, it holds that, for some* $\delta > 0$*,*

$$\mathrm{E}\left(\left|\frac{\partial}{\partial\theta_i}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\frac{\partial}{\partial\theta_j}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right|^{1+\delta}\right) < \infty.$$

Hence, the matrix

$$I := \mathrm{E}\left(\frac{\partial}{\partial \boldsymbol{\theta}}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\frac{\partial}{\partial \boldsymbol{\theta}^T}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right)$$

is well defined.

(C11) *The following conditions hold true:*

$$\mathrm{E}\left(\ell''(Z_t, \lambda_t(\boldsymbol{\theta}_0))|\mathcal{F}_{t-1}\right) \neq 0 \quad a.s.,$$

and

$$\mathbf{s}^\top \frac{\partial}{\partial \boldsymbol{\theta}}\lambda_t(\boldsymbol{\theta}_0) = 0 \quad \Rightarrow \ \mathbf{s} = \mathbf{0}.$$

(C12) *The true value $\boldsymbol{\theta}_0$ belongs to the interior of $\boldsymbol{\Theta}$.*

Remark 3. The conditions described in Assumption 2 are fundamental assumptions, such as the identifiability of the intensity function and the compactness of the parameter space, which are required to construct estimators with strong consistency and asymptotic normality (CAN).

Under Assumptions 1 and 2, we have the CAN of $\hat{\boldsymbol{\theta}}_n$.

Theorem 1. *Under Assumptions* 1 (A0) *and* 2 (B1)–(B5)*, it holds that*

$$\hat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0 \quad a.s. \ \ as \ n \to \infty.$$

Theorem 2. *Under Assumption* 1 (A0) *and* 2*, it holds that*

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) \Rightarrow N(0, J^{-1}IJ^{-1}) \quad as \ n \to \infty,$$

where

$$\begin{aligned} I :=&\mathrm{E}\left(\frac{\partial}{\partial \boldsymbol{\theta}}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\frac{\partial}{\partial \boldsymbol{\theta}^\top}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right) \\ =&\mathrm{E}\left(\left(\ell'(Z_t, \lambda_t(\boldsymbol{\theta}_0))\right)^2 \frac{\partial}{\partial \boldsymbol{\theta}}\lambda_t(\boldsymbol{\theta}_0)\frac{\partial}{\partial \boldsymbol{\theta}^\top}\lambda_t(\boldsymbol{\theta}_0)\right) \\ J :=& -\mathrm{E}\left(\frac{\partial^2}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right) \\ =& -\mathrm{E}\left(\ell''(Z_t, \lambda_t(\boldsymbol{\theta}_0))\frac{\partial}{\partial \boldsymbol{\theta}}\lambda_t(\boldsymbol{\theta}_0)\frac{\partial}{\partial \boldsymbol{\theta}^\top}\lambda_t(\boldsymbol{\theta}_0)\right). \end{aligned}$$

Remark 4. Ahmad and Francq (2016) proposed the Poisson quasi maximum likelihood estimator (QMLE), showing that it satisfies the CAN property. This estimator is efficient when the underlying conditional distribution is Poisson. Similarly, Aknouche, Bendjeddou and Touche (2018) investigated the NB QMLE. Moreover, Aknouche and Francq (2021a) suggested the exponential QMLE when the underlying process is a nonnegative time series. Under regularity conditions, the NB QMLE and exponential QMLE satisfy the CAN property and are efficient when the underlying distributions are NB and exponential distributions, respectively. However, the true parameter needs to belong to the interior of the parameter space, and if the underlying conditional distribution is not correct, the QMLEs cannot be efficient. To overcome these drawbacks of QMLEs, Aknouche and Francq (2021b) proposed the weighted least squares estimator and showed that it satisfies the CAN property.

3.2. Asymptotic behaviors of the test statistics

In this subsection, we show that our test has asymptotic size α and is consistent. Furthermore, the proposed test has nontrivial power under the local alternative. We introduce the following notation:

$$\hat{\sigma}_n^2 := \frac{1}{n}\sum_{t=1}^{n}\left\{(Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \kappa_0(\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))\right\}^2$$

and

$$\sigma^2 := \mathrm{E}\left(\left\{(Z_t - \lambda_t(\boldsymbol{\theta}_0))^2 - \kappa_0(\lambda_t(\boldsymbol{\theta}_0))\right\}^2\right).$$

Then, we have the asymptotic null distribution.

Theorem 3. *Suppose that the estimator $\hat{\boldsymbol{\theta}}_n$ of $\boldsymbol{\theta}_0$ is defined by* (3.1). *Under Assumptions* 1 *and* 2 *and the null H_0, the following (a) and (b) hold true:*

(a) Suppose that $M_n = o(n)$. Then, it holds that

$$T_n \Rightarrow N(0, \sigma^2) \quad \text{as } n \to \infty.$$

(b) Suppose that $M_n = n$. Then, it holds that

$$T_n \Rightarrow N(0, \tilde{\sigma}^2) \quad \text{as } n \to \infty,$$

where $\tilde{\sigma}^2$ is defined as follows:

$$\tilde{\sigma}^2 := \sigma^2 + L^\top J^{-1} I J^{-1} L + 2L^\top J^{-1} C_{12},$$

with

$$L := \mathrm{E}\left(\kappa_0'(\lambda_t(\boldsymbol{\theta}_0))\left(\frac{\partial}{\partial\boldsymbol{\theta}}\lambda_t(\boldsymbol{\theta}_0)\right)\right), \tag{3.2}$$

$$C_{12} := \mathrm{E}\left(\left(\frac{\partial}{\partial\boldsymbol{\theta}}\ell(Z_t,\lambda_t(\boldsymbol{\theta}_0))\right)^\top \{(Z_t-\lambda_t(\boldsymbol{\theta}_0))^2-\kappa_0(\lambda_t(\boldsymbol{\theta}_0))\}\right), \tag{3.3}$$

and the matrices I and J defined in Theorem 3.2.

Remark 5. When $M_n := \lfloor cn \rfloor$ for some constant c, such that $0 < c < 1$, T_n is asymptotically normal with mean zero and variance $\sigma^2 + c(L^\top J^{-1} I J^{-1} L + 2L^\top J^{-1} C_{12})$. This result suggests that we choose $M_n := n$ if $\hat{L}^\top \hat{J}^{-1}\hat{I}\hat{J}^{-1}\hat{L} + 2\hat{L}^\top \hat{J}^{-1}\hat{C}_{12} < 0$; otherwise, we take $M_n := o(n)$.

The asymptotic variances σ^2 and $\tilde{\sigma}^2$ can be estimated by $\hat{\sigma}_n^2$ and

$$\hat{\hat{\sigma}}_n^2 = \hat{\sigma}_n^2 + \hat{L}^\top \hat{J}^{-1}\hat{I}\hat{J}^{-1}\hat{L} + 2\hat{L}^\top \hat{J}^{-1}\hat{C}_{12},$$

respectively, where

$$\hat{L} := \frac{1}{n}\sum_{t=1}^{n}\kappa_0'(\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))\frac{\partial}{\partial\boldsymbol{\theta}}\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n),$$
$$\hat{I} := \frac{1}{n}\sum_{t=1}^{n}\frac{\partial}{\partial\boldsymbol{\theta}}\ell(Z_t,\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))\frac{\partial}{\partial\boldsymbol{\theta}^\top}\ell(Z_t,\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)),$$
$$\hat{J} := -\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}\ell(Z_t,\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)),$$

and

$$\hat{C}_{12} := \frac{1}{n}\sum_{t=1}^{n}\left(\frac{\partial}{\partial\boldsymbol{\theta}}\ell(Z_t,\lambda_t(\hat{\boldsymbol{\theta}}_n))\right)^\top \left\{(Z_t-\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2-\kappa_0(\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))\right\},$$

for both cases (a) and (b). When we consider case (a), we obtain the asymptotic size α tests if we reject H_0 when $\hat{\sigma}^{-1}|T_n| \geq z_{\alpha/2}$, where $z_{\alpha/2}$ is the $(1-\alpha/2)$-quantile of the standard normal distribution. The same assertion holds true for case (b).

We can prove the consistency of the test as follows.

Theorem 4. *Suppose that the estimator $\hat{\boldsymbol{\theta}}_n$ of $\boldsymbol{\theta}_0$ is defined by* (3.1). *Under alternative H_1, Assumptions* 1 *and* 2 *and $\kappa(\lambda_t(\boldsymbol{\theta}_0))$ being integrable, and*

$$\mathrm{E}(\kappa(\lambda_t(\boldsymbol{\theta}_0))) \neq \mathrm{E}(\kappa_0(\lambda_t(\boldsymbol{\theta}_0))),$$

where

$$\kappa(\lambda_t(\boldsymbol{\theta}_0)) = \mathrm{E}((Z_t - \lambda_t(\boldsymbol{\theta}_0))^2|\mathcal{F}_{t-1}),$$

it holds for every $C > 0$ *that*

$$P(|T_n| > C|H_1) \to 1, \quad n \to \infty.$$

Moreover, the next theorem shows that the test statistic T_n has nontrivial power.

Theorem 5. *Suppose that the estimator* $\hat{\boldsymbol{\theta}}_n$ *of* $\boldsymbol{\theta}_0$ *is defined by* (3.1)*, and that Assumptions* 1 *and* 2 *hold. Under the local alternative hypothesis*

$$H_{1,n}: \ \kappa(x) = \kappa_n(x), \quad x \in \mathbb{R},$$

where $\kappa_n(x) := \kappa_0(x) + h(x)/\sqrt{M_n}$ *and* $h(x)$ *is a measurable function, such as* $\mathrm{E}|h(\lambda_t(\boldsymbol{\theta}_0))|^{2+\delta} < \infty$ *for some* $\delta > 0$*, it holds that*

$$T_n \Rightarrow \begin{cases} N(\mathrm{E}\left[h(\lambda_t(\boldsymbol{\theta}_0))\right], \sigma^2) & M_n = o(n), \\ N(\mathrm{E}\left[h(\lambda_t(\boldsymbol{\theta}_0))\right], \tilde{\sigma}^2) & M_n = n, \end{cases} \quad \text{as } n \to \infty.$$

From Theorem 5 and the Portmanteau theorem, we can derive the nontrivial power. Under the local alternative $H_{1,n}$ and $M_n = o(n)$, we have

$$\begin{aligned} &\mathrm{P}(\hat{\sigma}_n^{-1}|T_n| > z_{\alpha/2}|H_{1,n}) \\ &\to \mathrm{P}\Big(N(0,1) \in (-\infty, -z_{\alpha/2} - \sigma^{-1}\mathrm{E}\left(h(\lambda_t(\boldsymbol{\theta}_0))\right)) \cup \\ &(z_{\alpha/2} - \sigma^{-1}\mathrm{E}\left(h(\lambda_t(\boldsymbol{\theta}_0))\right), \infty)|H_{1,n}\Big) \quad \text{as } n \to \infty. \end{aligned}$$

This can be rewritten by the simple form $1 - \Phi(z_{\alpha/2} - \sigma^{-1}\mathrm{E}\left(h(\lambda_t(\boldsymbol{\theta}_0))\right)) + \Phi(-z_{\alpha/2} - \sigma^{-1}\mathrm{E}\left(h(\lambda_t(\boldsymbol{\theta}_0))\right))$, where Φ is the cumulative distribution function of the standard normal distribution. Similarly, it holds that, under the local alternative $H_{1,n}$ and $M_n = n$, $\mathrm{P}(\hat{\tilde{\sigma}}_n^{-1}|T_n| > z_{\alpha/2}|H_{1,n}) \to 1 - \Phi(z_{\alpha/2} - \tilde{\sigma}^{-1}\mathrm{E}\left(h(\lambda_t(\boldsymbol{\theta}_0))\right)) + \Phi(-z_{\alpha/2} - \tilde{\sigma}^{-1}\mathrm{E}\left(h(\lambda_t(\boldsymbol{\theta}_0))\right))$ as $n \to \infty$.

Remark 6. Theorem 5 corresponds to (ii) of Proposition 2.3 of Fokianos and Neumann (2013), who proposed a specification test for the intensity function of the Poisson process based on the supremum of the Peason residual. They discuss the nontrivial power for the local alternative.

4. Applications

The proposed test statistics can be applied to various problems. We introduce some of them in this section.

Example 1 (Goodness-of-fit test)**.** The first important application is a goodness-of-fit test. Davis and Liu (2016) proposed the exponential family for integer-valued time series, which is defined as

$$p_{\exp}(z|\eta) := \exp\{\eta z - A(\eta)\}h(z)\mathbb{I}\{z \geq 0\},$$

where η is a natural parameter, and $A(\cdot)$ and $h(\cdot)$ are known functions. If Z_η follows the exponential family with a parameter η, it is known that the mean and variance are given by $\lambda_\eta := \mathrm{E}(Z_\eta) = A'(\eta)$ and $\mathrm{Var}(Z_\eta) = A''(\eta) > 0$, respectively, provided $A(\eta)$ is twice differentiable with respect to η. Thus, $A'(\eta)$ is a strictly increasing function. From Davis and Liu (2016, Prop. A.1), the exponential family satisfies the stochastic equal mean order property; that is, for $\eta \leq \eta'$ (or, equivalently, for $\lambda_\eta \leq \lambda_{\eta'}$), $\mathrm{P}(Z_\eta > x) \leq \mathrm{P}(Z_{\eta'} > x)$, for any $x \in \mathbb{R}$. Thus, there exists a strictly stationary and ergodic solution for the INGARCH(p,q) model under the condition described in Remark 2. The null hypothesis is

$$G_0 : Z_t \text{ follows the target distribution},$$

and the alternative is

$$G_1 : Z_t \text{ does not follow the target distribution}.$$

The derivations of the asymptotic variances of the following test statistics are available in the Supplementary Material. Specific examples are given as follows.

Goodness-of-fit test for Poisson distribution. By setting $\eta = \lambda$, $A(\eta) = \exp(\lambda)$, and $h(z) = 1/z!$, Z_η follows a Poisson distribution with parameter λ, and with mean and variance are λ and λ, respectively. Doukhan, Fokianos, K. and Tjøstheim (2013) showed the existence of a moment of any order under the contractive condition of the intensity function and the summation of its coefficients being less than one. From Theorem 3, it holds that, under the null G_0, the following statistics converge to the standard normal distribution:

$$T_n^{\mathrm{Pois}} := \begin{cases} \hat{\sigma}^{-1}\dfrac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\left\{(Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)\right\} & M_n = o(n) \\ \hat{\hat{\sigma}}^{-1}\dfrac{1}{\sqrt{n}}\sum_{t=1}^{n}\left\{(Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)\right\} & M_n = n, \end{cases}$$

where

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n}\left\{2\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)^2 + \tilde{\lambda}(\hat{\boldsymbol{\theta}}_n)\right\},$$

and $\hat{\hat{\sigma}}^2$ is defined in Theorem 3 for $\kappa_0(\lambda_t(\boldsymbol{\theta})) = \lambda_t(\boldsymbol{\theta})$. Thus, we can construct a goodness-of-fit test for the Poisson distribution.

Goodness-of-fit test for NB distribution If we define $\eta = \log(1-p)$, $A(\eta) = -r\log(1-\exp(\eta))$, and $h(z) = {}_{z+r-1}\mathrm{C}_z$, then Z_η is a NB distribution with known parameter r and unknown parameter p. Then, the mean and variance of Z_η are $\lambda := r(1-p)/p$ and $r(1-p)/p^2 = (\lambda+r)\lambda/r$, respectively. Under an appropriate moment condition and the null G_0, the following test statistic converges to the standard normal distribution as $n \to \infty$:

$$T_n^{\mathrm{NB}} := \begin{cases} \hat{\sigma}^{-1}\dfrac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\left\{(Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - (\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n) + r)\dfrac{\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)}{r}\right\} & M_n = o(n) \\ \hat{\hat{\sigma}}^{-1}\dfrac{1}{\sqrt{M_n}}\sum_{t=1}^{M_n}\left\{(Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - (\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n) + r)\dfrac{\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)}{r}\right\} & M_n = n, \end{cases}$$

where

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n}\frac{(6+2r)\tilde{\lambda}_t^4(\hat{\boldsymbol{\theta}}_n) + (12r+4r^2)\tilde{\lambda}_t^3(\hat{\boldsymbol{\theta}}_n) + (7r^2+2r^3)\tilde{\lambda}_t^2(\hat{\boldsymbol{\theta}}_n) + r^3\tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n)}{r^3},$$

and $\hat{\hat{\sigma}}^2$ is defined in Theorem 3 for $\kappa_0(\lambda_t(\boldsymbol{\theta})) = \lambda_t(\boldsymbol{\theta})(\lambda_t(\boldsymbol{\theta}) + r)/r$. In the case of $r = 1$, we obtain a goodness-of-fit test for the geometric distribution.

In the Supplementary Material, we provide goodness-of-fit tests for binomial and gamma distributions.

Example 2 (Specification test for the intensity function)**.** The next important application is a specification test for the intensity function, as investigated by Neumann (2011). Fokianos and Neumann (2013) proposed a supremum type of specification test, and Leucht and Neumann (2013) advocated the L^2 norm-based test. These papers assume a Poisson INGARCH(1,1) model.

The null hypothesis and the alternative are given by

$$K_0 : \lambda = \lambda^0 \quad \text{and } K_1 : \lambda \neq \lambda^0,$$

respectively. We assume that the form of the conditional variance is known, that is, $v_t = \kappa_0(\lambda_t(\boldsymbol{\theta}_0))$. The test statistic for this test can be defined as

$$T_n^{\text{spec}} := \frac{1}{\sqrt{M_n}} \sum_{t=1}^{M_n} \left\{ (Z_t - \lambda_t^0)^2 - \kappa_0(\lambda_t^0) \right\}.$$

From Theorem 3, it holds that, under the null K_0, $T_n^{\text{spec}} \Rightarrow N(0, \sigma^2)$ as $n \to \infty$. Furthermore, we assume the true intensity function is given by λ^1 and appropriate moment conditions. Then, under the alternative K_1, we observe that

$$\begin{aligned}\frac{1}{\sqrt{M_n}} T_n^{\text{spec}} =& \frac{1}{M_n} \sum_{t=1}^{M_n} \Big((Z_t - \lambda_t^1)^2 - \kappa_0(\lambda_t^1) + (\lambda_t^1 - \lambda_t^0)^2 \\ &+ 2(Z_t - \lambda_t^1)(\lambda_t^1 - \lambda_t^0) + \kappa_0(\lambda_t^1) - \kappa_0(\lambda_t^0) \Big),\end{aligned}$$

which, by the ergodic theorem, converges to $\mathrm{E}\left((\lambda_t^1 - \lambda_t^0)^2 + \kappa_0(\lambda_t^1) - \kappa_0(\lambda_t^0) \right)$ as $n \to \infty$. If this quantity does not equal zero, the consistency of the test holds. Thus, we obtain a size α and consistent test for intensity.

The proposed test statistic does not include Neumann's statistic, given by

$$T_n^{\text{Neumann}} := \left(\frac{2}{n} \sum_{t=1}^{n} \left(\lambda_t^0\right)^2 \right)^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\{ (Z_t - \lambda_t^0)^2 - Z_t \right\}. \tag{4.1}$$

Provided that the underlying conditional distribution is Poisson, T_n^{Neumann} converges to $N(0,1)$ as $n \to \infty$. On the other hand, our statistic for the Poisson hypothesis is defined as

$$T_n^{\text{specPois}} := \hat{\sigma}^{-1} \frac{1}{\sqrt{M_n}} \sum_{t=1}^{M_n} \left\{ (Z_t - \lambda_t^0)^2 - \lambda_t^0 \right\},$$

where $\hat{\sigma}^2 = \sum_{t=1}^{n} \left(2(\lambda_t^0)^2 + \lambda_t^0 \right) / n$, which converges to N(0,1) as $n \to \infty$ under the Poisson assumption.

Note that the intensity function can take the form of the nonlinear INGARCH (p,q) model, and our theory can be applied to distributions other than the Poisson. See also Remark 2.

Our initial attempt is to construct a composite hypothesis for the null hypothesis that λ_t belongs to a given parametric family. Although the consistent test can be constructed in theory, the empirical power of the test is poor unless the discrepancy between the null and the alternative is significant.

Example 3 (Detection of (conditional) overdispersion or underdispersion)**.** Conditional overdispersion (underdispersion) occurs in data when the conditional variance is greater (less) than its conditional expectation. Many models have been proposed for integer-valued time series for data with conditional overdispersion, for example, the NB distribution. It is important to decide whether data are characterized by overdispersion or underdispersion, using a statistical procedure. This can be formulated as follows. The null hypothesis and the alternative are defined as

$$R_0 : \mathrm{E}(Z_t|\mathcal{F}_{t-1}) = \mathrm{Var}(Z_t|\mathcal{F}_{t-1}) \text{ a.s.}$$

and

$$R_1 : \mathrm{P}\left(\mathrm{E}(Z_t|\mathcal{F}_{t-1}) \neq \mathrm{Var}(Z_t|\mathcal{F}_{t-1})\right) > 0,$$

respectively. Putting $\kappa_0(\lambda_t(\boldsymbol{\theta})) = \lambda_t(\boldsymbol{\theta}_0)$, we define the following test statistics:

$$T_n^{\mathrm{disp}} := \begin{cases} \hat{\sigma}^{-1} \dfrac{1}{\sqrt{M_n}} \sum_{t=1}^{M_n} \left\{ (Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n) \right\} & M_n = o(n) \\ \hat{\hat{\sigma}}^{-1} \dfrac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\{ (Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n) \right\} & M_n = n, \end{cases}$$

where

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^{n} \left\{ (Z_t - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n))^2 - \tilde{\lambda}_t(\hat{\boldsymbol{\theta}}_n) \right\}^2,$$

and $\hat{\hat{\sigma}}^2$ is defined in Theorem 3 for $\kappa_0(\lambda_t(\boldsymbol{\theta})) = \lambda_t(\boldsymbol{\theta})$. These two statistics converge to the standard normal distribution under the null hypothesis. Hence, the test can be constructed in the same way as the discussion below Theorem 3. Note that we do not assume that the underlying conditional distribution is Poisson. Several processes with distributions other than Poisson distribution hold the conditional equidispersion property. For example, the conditional expectation and variance can be separately modeled using the double exponential family (see Efron (1986)) in the same way as Heinen (2003).

5. Numerical Study

First, we investigate the finite-sample performance of our methods for the goodness-of-fit test described in Example 1 in Section 4. Here, we assume the intensity follows the INGARCH(1,1) model $\lambda_t = \omega + \alpha Z_{t-1} + \beta\lambda_{t-1}$. The unknown parameters are estimated using a Poisson QMLE. The burn-in period is 1,000. The simulation procedure is as follows. First, we generate n ($n = 50, 100, 200, 300, 600, 900$) samples from Poisson ($\mathrm{Pois}(\lambda_t)$) or NB ($\mathrm{NB}(4, 4/(4 +$

Table 1. The empirical size at the nominal size 0.05 for Poisson INGARCH(1,1) models and κ_{Pois}.

$\omega = 1, \alpha = 0.3, \beta = 0.2$						
Statistic \ n	50	100	200	300	600	900
T_n	0.102	0.092	0.077	0.043	0.039	0.040
T_M	0.090	0.051	0.045	0.045	0.041	0.039
$\omega = 1, \alpha = 0.3, \beta = 0.4$						
n	50	100	200	300	600	900
T_n	0.103	0.088	0.048	0.052	0.049	0.051
T_M	0.075	0.064	0.036	0.037	0.048	0.045
$\omega = 1.5, \alpha = 0.3, \beta = 0.2$						
n	50	100	200	300	600	900
T_n	0.097	0.072	0.064	0.059	0.049	0.048
T_M	0.059	0.046	0.049	0.044	0.046	0.047

Table 2. The empirical size at the nominal size 0.05 for NB INGARCH(1,1) models and κ_{NB}.

$\omega = 1, \alpha = 0.3, \beta = 0.2$						
Statistic \ n	50	100	200	300	600	900
T_n	0.122	0.084	0.083	0.047	0.041	0.047
T_M	0.062	0.052	0.030	0.036	0.038	0.038
$\omega = 1, \alpha = 0.3, \beta = 0.4$						
Statistic \ n	50	100	200	300	600	900
T_n	0.097	0.093	0.073	0.063	0.061	0.040
T_M	0.070	0.060	0.033	0.035	0.035	0.040
$\omega = 1.5, \alpha = 0.3, \beta = 0.2$						
Statistic \ n	50	100	200	300	600	900
T_n	0.107	0.081	0.072	0.074	0.040	0.041
T_M	0.075	0.046	0.045	0.038	0.037	0.046

λ_t))) INGARCH(1,1) models with parameters $(\omega, \alpha, \beta) = (1, 0.3, 0.2), (1, 0.3, 0.4)$, or $(1.5, 0.3, 0.2)$. Then, we calculate the proposed statistics for the null hypothesis $\kappa_{\text{Pois}}(x) := x$ and $\kappa_{\text{NB}}(x) = x(x+4)/4$, with $M_n = n^{4/5}$. We denote the statistics as T_M when $M_n = n^{4/5}$. Finally, we iterate 1,000 times and compute the rejection probability for the significance level 0.05. Note that we use $Z_0, \tilde{\lambda}_0(\boldsymbol{\theta}) = \sum_{t=1}^{n} Z_t/n$ and $\partial\lambda_0(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \mathbf{0}$ as the initial values.

The results are summarized in Tables 1–4. Tables 1 and 2 show that the tests based on the proposed statistics T_n and T_M have good size control overall. For a small sample size, T_M provides better size than that of T_n. The instability of the Poisson QMLE can explain this for small samples. For relatively large sample sizes, the tests' sizes are close to the nominal size of 0.05 and are almost the same.

Table 3. The empirical power at the nominal size 0.05 for Poisson INGARCH(1,1) models and κ_{NB}.

	$\omega = 1, \alpha = 0.3, \beta = 0.2$					
Statistic \ n	50	100	200	300	600	900
T_n	0.734	0.894	0.986	0.999	1.000	1.000
T_M	0.460	0.596	0.782	0.885	0.978	0.997
	$\omega = 1, \alpha = 0.3, \beta = 0.4$					
Statistic \ n	50	100	200	300	600	900
T_n	0.911	0.990	1.000	1.000	1.000	1.000
T_M	0.719	0.893	0.985	0.996	1.000	1.000
	$\omega = 1.5, \alpha = 0.3, \beta = 0.2$					
Statistic \ n	50	100	200	300	600	900
T_n	0.885	0.984	1.000	1.000	1.000	1.000
T_M	0.66	0.8482	0.960	0.989	1.000	1.000

Table 4. The empirical power at the nominal size 0.05 for NB INGARCH(1,1) models and κ_{Pois}.

	$\omega = 1.5, \alpha = 0.3, \beta = 0.2$					
Statistic \ n	50	100	200	300	600	900
T_n	0.003	0.009	0.38	0.8017	0.990	1.000
T_M	0.087	0.190	0.297	0.373	0.564	0.704
	$\omega = 1, \alpha = 0.3, \beta = 0.4$					
Statistic \ n	50	100	200	300	600	900
T_n	0.000	0.03	0.627	0.888	0.991	0.997
T_M	0.164	0.313	0.487	0.594	0.82	0.9192
	$\omega = 1.5, \alpha = 0.3, \beta = 0.2$					
Statistic \ n	50	100	200	300	600	900
T_n	0.001	0.023	0.655	0.921	0.992	0.999
T_M	0.145	0.280	0.445	0.540	0.795	0.902

On the other hand, Tables 3 and 4 show that, as the sample size gets larger, the power of both tests increases. The test based on T_n is more powerful than the test based on T_M.

Next, we illustrate the finite-sample performance of the proposed method for the specification test described in Example 2 in Section 4. The null hypothesis we investigate here is INARCH(1) model $\lambda_t = \omega + \alpha Z_{t-1}$, and the alternative hypothesis INARCH(1) with different coefficients from the null. We set the sample size $n \in \{100, 200, 300, 600, 900\}$, the number of iteration is 1,000, and the significance level $\alpha = 0.05$. We compute our statistics T_n^{spec} and Neumann's statistic T_n^{Neumann} defined in (4.1).

The simulation results are shown in Tables 5–8. Tables 5 and 6 display the empirical sizes of the tests based on T_n^{spec} and T_n^{Neumann} close to 0.05 as the sample

Table 5. The empirical sizes at the nominal size 0.05 for Poisson INARCH(1) model $\lambda_t = \omega + \alpha Z_{t-1}$.

$\omega = 1,\ \alpha = 0.4$					
Statistic \ n	100	200	300	600	900
T_n^{spec}	0.082	0.073	0.066	0.074	0.056
T_n^{Neumann}	0.044	0.047	0.055	0.066	0.044
$\omega = 1,\ \alpha = 0.6$					
T_n^{spec}	0.090	0.059	0.054	0.059	0.052
T_n^{Neumann}	0.046	0.038	0.054	0.053	0.050

Table 6. The empirical sizes at the nominal size 0.05 for NB INARCH(1) model $\lambda_t = \omega + \alpha Z_{t-1}$ with $r = 4$.

$\omega = 1,\ \alpha = 0.4$					
Statistic \ n	100	200	300	600	900
T_n^{spec}	0.128	0.116	0.073	0.073	0.058
T_n^{Neumann}	0.722	0.925	0.985	1.000	1.000
$\omega = 1,\ \alpha = 0.6$					
T_n^{spec}	0.138	0.123	0.109	0.073	0.088
T_n^{Neumann}	0.906	0.994	1.000	1.000	1.000

size becomes larger, except for Neumann's test of the NB case. This is because Neumann's test is constructed using the property of the Poisson distribution.

The empirical power is indicated in Tables 7 and 8. For the Poisson case, both tests have good power when the coefficient of the alternative is larger than that of the null. On the other hand, our proposed test works well for large sample sizes in every case when the conditional distribution follows a NB distribution.

6. Empirical Study

In this section, we analyze the weekly number of patients with Escherichia coli in Germany from January 2001 to May 2013. This data set (called *ecoli*, hereafter) has 646 observations and can be found in *tscount* (Liboschik, Fokianos and Fried (2017)). A plot of the observations and the sample ACF are shown in Figures 1 and 2, respectively. The sample mean and the sample variance are 20.33 and 88.62, respectively, and, thus, *ecoli* exhibits overdispersion.

First, we determine the order $\hat{p}$ and $\hat{q}$ of INGARCH(p,q) by minimizing Takeuchi's information criterion (TIC) (see Takeuchi (1976) and Konishi and Kitagawa (2008, p.60)). That is, $\hat{p}$ and $\hat{q}$ are defined as

$$(\hat{p}, \hat{q}) := \underset{\max(p,q)\leq 5}{\operatorname{argmin}}\ \mathrm{TIC}(p,q), \quad \mathrm{TIC}(p,q) := -2n\tilde{L}_n(\hat{\boldsymbol{\theta}}_n) + 2\mathrm{tr}\left(\hat{J}^{-1}\hat{I}\right).$$

Table 7. The empirical power at the nominal size 0.05 for the null being the Poisson INARCH(1) model $\lambda_t = \omega + \alpha Z_{t-1}$ and the alternative being the Poisson INARCH(1) model $\lambda_t = \omega' + \alpha' Z_{t-1}$.

Statistic \ n	100	200	300	600	900
$\omega = 1$, $\alpha = 0.4$, $\omega' = 1$, $\alpha' = 0.2$,					
T_n^{spec}	0.208	0.252	0.287	0.457	0.533
T_n^{Neumann}	0.060	0.080	0.137	0.209	0.336
$\omega = 1$, $\alpha = 0.4$, $\omega' = 1$, $\alpha' = 0.6$					
T_n^{spec}	0.289	0.627	0.823	0.992	0.999
T_n^{Neumann}	0.290	0.460	0.543	0.758	0.872
$\omega = 1$, $\alpha = 0.4$, $\omega' = 1$, $\alpha' = 0.8$					
T_n^{spec}	0.996	1.000	1.000	1.000	1.000
T_n^{Neumann}	0.990	0.999	1.000	1.000	1.000
$\omega = 1$, $\alpha = 0.6$, $\omega' = 1$, $\alpha' = 0.2$					
T_n^{spec}	0.103	0.105	0.121	0.112	0.138
T_n^{Neumann}	0.398	0.706	0.858	0.994	1.000
$\omega = 1$, $\alpha = 0.6$, $\omega' = 1$, $\alpha' = 0.4$					
T_n^{spec}	0.177	0.194	0.231	0.312	0.388
T_n^{Neumann}	0.077	0.113	0.160	0.317	0.480
$\omega = 1$, $\alpha = 0.6$, $\omega' = 1$, $\alpha' = 0.8$					
T_n^{spec}	0.496	0.855	0.957	1.000	1.000
T_n^{Neumann}	0.512	0.717	0.839	0.981	0.993

Table 8. The empirical power at the nominal size 0.05 for the null being the NB INARCH(1) model $\lambda_t = \omega + \alpha Z_{t-1}$ with $r = 4$ and the alternative being the NB INARCH(1) model $\lambda_t = \omega' + \alpha' Z_{t-1}$ with $r = 4$.

Statistic \ n	100	200	300	600	900
$\omega = 1$, $\alpha = 0.4$, $\omega' = 1$, $\alpha' = 0.2$,					
T_n^{spec}	0.363	0.449	0.539	0.746	0.886
$\omega = 1$, $\alpha = 0.4$, $\omega' = 1$, $\alpha' = 0.6$					
T_n^{spec}	0.130	0.353	0.623	0.940	0.989
$\omega = 1$, $\alpha = 0.4$, $\omega' = 1$, $\alpha' = 0.8$					
T_n^{spec}	0.741	0.933	0.976	0.990	0.992
$\omega = 1$, $\alpha = 0.6$, $\omega' = 1$, $\alpha' = 0.2$					
T_n^{spec}	0.457	0.616	0.760	0.937	0.978
$\omega = 1$, $\alpha = 0.6$, $\omega' = 1$, $\alpha' = 0.4$					
T_n^{spec}	0.383	0.519	0.590	0.797	0.894
$\omega = 1$, $\alpha = 0.6$, $\omega' = 1$, $\alpha' = 0.8$					
T_n^{spec}	0.058	0.166	0.337	0.744	0.883

Table 9 displays TIC values for p and q such that $\max(p, q) \leq 5$, showing that $(p, q) = (5, 4)$ is an appropriate order in terms of the TIC. The estimated parameters are given as follows: $\hat{\omega} = 2.594$ (the perception of INGARCH (5,4)),

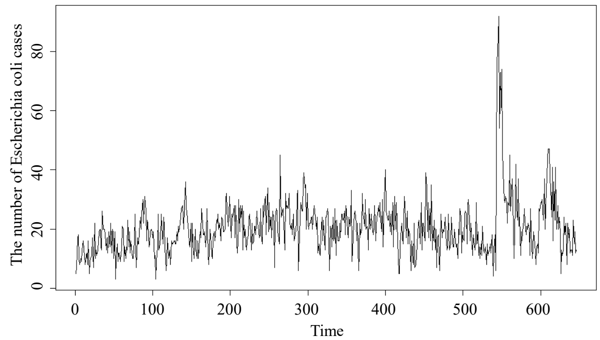

Figure 1. The weekly number of the patients with Escherichia coli in Germany from January 2001 to May 2013.

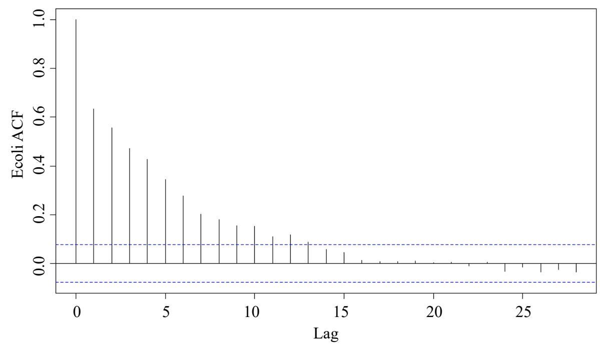

Figure 2. The sample ACF of *ecoli*.

Table 9. TIC values for INGARCH(p,q).

p\q	0	1	2	3	4	5
0	-39320.96	-52888.12	-52895.60	-52944.33	-52919.67	-52962.39
1	-53310.82	-53888.15	-53895.34	-53900.00	-53904.75	-53909.33
2	-53690.14	-53893.83	-53907.72	-53917.73	-53922.09	-53927.31
3	-53798.05	-53900.78	-53915.98	-53920.67	-53915.95	-53927.28
4	-53838.29	-53905.73	-53923.63	-53927.48	-53926.64	-53938.76
5	-53849.16	-53907.63	-53920.39	-53924.59	-53939.10	-53921.30

$(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4, \hat{\alpha}_5) = (3.724\times 10^{-1}, 2.392\times 10^{-4}, 1.853\times 10^{-2}, 3.238\times 10^{-4}, 1.002\times 10^{-5})$ (the autoregression coefficients with respect to $\{Z_t\}$), and $(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4) = (4.763 \times 10^{-1}, 3.505 \times 10^{-3}, 3.660 \times 10^{-4}, 4.260 \times 10^{-5})$ (the regression coefficients with respect to $\{\lambda_t\}$).

Next, we apply our proposed tests based on T_n, T_{M_1}, and T_{M_2} with $M_1 :=$

$\lfloor n^{59/60} \rfloor$ and $M_2 := \lfloor n^{58/60} \rfloor$ for the null hypothesis is that the underlying conditional distribution follows a Poisson distribution or a NB distribution with $r = 1, \ldots, 50$.

As expected, all three tests reject the Poisson hypothesis. When the null hypothesis is NB, any one of three tests reject the hypothesis for $r = 1, \ldots, 14, 24, \ldots, 50$ and all three tests accept the NB distribution with $r = 15, \ldots, 23$. Consequently, a plausible model for *ecoli* is INGARCH(5,4) with an NB conditional distribution for $r = 15, \ldots, 23$.

Supplementary Material

Proofs of Theorems 1–5, additional examples of the goodness-of-fit test, and explicit forms of the higher moments for several distributions are available in the Supplementary Material.

Acknowledgments

The authors are grateful to the editor and two referees for their instructive comments. The author Y.G. is also grateful to Professor Masanobu Taniguchi for his helpful comments and suggestions. This research was supported by Grant-in-Aid for JSPS Research Fellow Grant Number JP201920060 (Yuichi Goto), the Research Institute for Science & Engineering of Waseda University (Masanobu Taniguchi), and JSPS Grant-in-Aid for Scientific Research (S) Grant Number JP18H05290 (Masanobu Taniguchi). This work was carried out when the first author was affiliated with Waseda University.

References

Ahmad, A. and Francq, C. (2016). Poisson QMEL of count time series models. *J. Time Ser. Anal.* **37**, 291–314.

Aknouche, A., Bendjeddou, S. and Touche, N. (2018). Negative binomial quasi-likelihood inference for general integer-valued time series models. *J. Time Ser. Anal.* **39**, 192–211.

Aknouche, A. and Francq, C. (2021a). Count and duration time series with equal conditional stochastic and mean orders. *Econometric Theory* **37**, 248–280.

Aknouche, A. and Francq, C. (2021b). Two-stage weighted least squares estimator of the conditional mean of observation-driven time series models. *J. Econometrics.* In press.

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

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

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

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

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

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

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

Fokianos, K. and Neumann, M. H. (2013). A goodness-of-fit test for poisson count processes. *Electron. J. Stat.* **7**, 793–819.

Franke, J. (2010). Weak dependence of functional INGARCH processes. *Preprint*.

Heinen, A. (2003). Modelling time series count data: An Autoregressive Conditional Poisson model. DOI: `http://dx.doi.org/10.2139/ssrn.1117187`.

Heinrich, L. (1991). Goodness-of-fit tests for the second moment funciton of a stationary multi-dimensional Poisson process. *Statistics* **22**, 245–268.

Hudecová, Š., Hušková, M. and Meintanis, S. G. (2015). Tests for time series of counts based on the probability-generating function. *Statistics* **49**, 316–337.

Konishi, S. and Kitagawa, G. (2008). *Information Criteria and Statistical Modeling*. Springer Science & Business Media, New York.

Leucht, A. and Neumann, M. H. (2013). Degenerate u-and v-statistics under ergodicity: Asymptotics, bootstrap and applications in statistics. *Ann. Inst. Statist. Math.* **65**, 349–386.

Liboschik, T., Fokianos, K. and Fried, R. (2017). tscount: An r package for analysis of count time series following generalized linear models. *J. Statist. Softwr.* **82**, 1–51.

Meintanis, S. G. and Karlis, D. (2014). Validation tests for the innovation distribution in inar time series models. *Comput. Statist.* **29**, 1221–1241.

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

Schweer, S. (2016). A goodness-of-fit test for integer-valued autoregressive processes. *J. Time Ser. Anal.* **37**, 77–98.

Takeuchi, K. (1976). Distribution of information statistics and criteria for adequacy of models. *Math. Sci.* **153**, 12–18 (in Japanese).

Weiß, C. H., Homburg, A. and Puig, P. (2019). Testing for zero inflation and overdispersion in INAR (1) models. *Stat. Pap.* **60**, 823–848.

Zhu, R. and Joe, H. (2010). Negative binomial time series models based on expectation thinning operators. *J. Statist. Plann. Inference* **140**, 1874–1888.

Yuichi Goto

Department of Mathematical Sciences, Faculty of Mathematics, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan.

E-mail: yuichi.goto@math.kyushu-u.ac.jp

Kou Fujimori

Department of Economics, Faculty of Economics and Law, Shinshu University, 3-1-1, Asahi, Matsumoto City, Nagano, 390-8621, Japan.

E-mail: kfujimori@shinshu-u.ac.jp

(Received September 2020; accepted April 2021)