Statistics and Probability Letters 112 (2016) 124–130

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Bias corrections for moment estimators in Poisson INAR(1) and INARCH(1) processes

Christian H. Weiß [a,*], Sebastian Schweer [b]

[a] Department of Mathematics and Statistics, Helmut Schmidt University, 22008 Hamburg, Germany
[b] University of Heidelberg, Institute of Applied Mathematics, 69120 Heidelberg, Germany

ARTICLE INFO

Article history:
Received 16 November 2015
Received in revised form 22 January 2016
Accepted 25 January 2016
Available online 11 February 2016

Keywords:
Count-data time series
Poisson INAR(1) model
Poisson INARCH(1) model
Asymptotic properties
Moment estimators
Bias corrections

ABSTRACT

For Poisson INAR(1) or INARCH(1) count data time series, explicit asymptotic approximations for bias and standard deviation of common moment estimators are derived. Their finite-sample performance is shown by simulations. Bias corrected estimators are applied within examples.

1. Introduction

Moment estimators applied to time series data $X_1, \ldots, X_T$ often lead to biased parameter estimates, which is particularly problematic if the given time series is short. For continuous-valued time series, a lot of work has been done in this area, especially for the most elementary model, the first-order (Gaussian) *a*uto*r*egressive (AR(1)) model, which is defined by the recursion $X_t = \alpha \cdot X_{t-1} + \varepsilon_t$. One of the earliest accounts in this regard is the work by Kendall (1954), who derived the asymptotic mean of a particular variation of the moment estimator of α as $\alpha - \frac{1}{T} \cdot (1 + 4\alpha)$. So the estimator exhibits a negative bias, which further increases with α. Many further results, also for higher-order models and different estimators, are provided by Shaman and Stine (1988) and the references therein. A related work for the first-order *a*uto*r*egressive *c*onditional *h*eteroskedasticity (ARCH(1)) model is the one by Engle et al. (1985), who considered the maximum likelihood estimator of the ARCH(1)'s autoregressive parameter α. They derived the formula $\alpha - \frac{1}{T} \cdot (1 + 12\alpha)$ for its asymptotic mean, i.e., with increasing α, a much stronger negative bias can be observed than in the AR(1) case.

For count data time series, there exist two equally popular *in*teger-valued counterparts to the AR(1) model, the so-called Poisson INAR(1) model on the one hand, and the Poisson INARCH(1) model on the other hand (details are provided below). Numerous research papers concerning these models have been published during the last years. However, as of yet, the topic of bias correction has received only little attention in the area of count data time series. Notable exceptions include Jung et al. (2005), where the Gaussian AR(1) model's bias correction is applied to simulated Poisson INAR(1) processes,

* Corresponding author.
E-mail addresses: weissc@hsu-hh.de (C.H. Weiß), Sebastian.Schweer@p3-group.com (S. Schweer).

http://dx.doi.org/10.1016/j.spl.2016.01.018
0167-7152/

Weiß and Kim (2013), where a 2-block jackknife is applied to diverse estimators for a binomial AR(1) process, Bourguignon and Vasconcellos (2015), where a bias correction for the squared difference estimator of the innovations' mean of the Poisson INAR(1) model is derived, and Weiß and Schweer (2015), where the bias of the dispersion ratio both for the Poisson INAR(1) and INARCH(1) models is analyzed.

In this article, we provide a comprehensive analysis of asymptotic distributions and bias corrections of common moment estimators within the Poisson INAR(1) and INARCH(1) model. Both are two-parameter models with an AR(1)-like autocorrelation function, but while the Poisson INAR(1) model has an equidispersed marginal distribution, that of the Poisson INARCH(1) model is overdispersed.

The *INAR(1) model* was first proposed by McKenzie (1985). Using the probabilistic operation '$\circ$' of binomial thinning (Steutel and van Harn, 1979), $(X_t)_{\mathbb{Z}}$ is defined by the recursion

$$X_t = \alpha \circ X_{t-1} + \epsilon_t \quad \text{with}\ \alpha \in (0;1), \tag{1}$$

where $(\epsilon_t)_{\mathbb{Z}}$ is an i.i.d. count data process. Here, all thinning operations are performed independently of each other and of $(\epsilon_t)_{\mathbb{Z}}$, and the thinning operations at each time t as well as ϵ_t are independent of $(X_s)_{s<t}$. In the *Poisson* INAR(1) model, the $(\epsilon_t)_{\mathbb{Z}}$ are assumed to be Poisson-distributed, $\epsilon_t \sim \text{Poi}(\beta)$ with $\beta > 0$. As a result, $(X_t)_{\mathbb{Z}}$ is a stationary and ergodic Markov chain with a Poisson marginal distribution, $X_t \sim \text{Poi}\big(\beta/(1-\alpha)\big)$, such that variance and mean are equal to each other, $\sigma^2 = \mu = \beta/(1-\alpha)$ (*equidispersion*). It is also α-mixing with geometrically decreasing weights (Schweer and Weiß , 2014), also see the more general results in Doukhan et al. (2012, 2013). Its autocorrelation function is given by $\rho(k) := \text{Corr}[X_t, X_{t-k}] = \alpha^k$ like in the standard AR(1) case (McKenzie, 1985).

The (Poisson) *INARCH(1) model*, in contrast, belongs to the family of *in*teger-valued ARCH models as introduced by Heinen (2003); Ferland et al. (2006). Although commonly referred to as INARCH model, it exhibits much more analogies to the conventional AR(1) model, and it can also be understood as a generalized linear model with identity link (Fahrmeir and Tutz, 1994; Kedem and Fokianos, 2002). Denoting its two model parameters again by $\beta > 0$ and $0 < \alpha < 1$, $(X_t)_{\mathbb{Z}}$ is defined by the recursion

$$X_t \big| X_{t-1}, X_{t-2} \ldots \sim \text{Poi}(\beta + \alpha \cdot X_{t-1}). \tag{2}$$

The INARCH(1) process defined by (2) is a stationary, ergodic Markov chain (Ferland et al., 2006; Zhu and Wang, 2011) with simple Poisson probabilities as the transition probabilities. According to Neumann (2011), it is β-mixing (and hence also α-mixing) with geometrically decreasing weights. As in the case of the INAR(1) model, all moments exist (Ferland et al., 2006) and its autocorrelation function equals $\rho(k) = \alpha^k$ again (Weiß , 2009). Here, however, the variance $\sigma^2 = \beta/\big((1-\alpha)(1-\alpha^2)\big)$ is larger than the mean $\mu = \beta/(1-\alpha)$ (*overdispersion*).

Let us give a brief outlook on this article: In Section 2, two central limit theorems are derived for vector-valued processes related to the underlying Poisson INAR(1) and Poisson INARCH(1) process, respectively. The asymptotic variances are explicitly calculated, allowing us to express the asymptotic distributions of two common sets of moment estimators in Sections 3 and 4. Furthermore, they enable us to derive explicitly the asymptotic bias of the estimators under consideration. This leads us directly to the proposal of alternative, bias-corrected moment estimators. After investigating the performance of the approximations in a simulation study, we conclude in Section 5.

2. Two central limit theorems

To analyze the asymptotic distribution of important moment estimators for both the Poisson INAR(1) and INARCH(1) model, we consider the vector-valued process defined by

$$\boldsymbol{Y}_t := \big(X_t - \mu,\ X_t^2 - \mu(0),\ X_t X_{t-1} - \mu(1)\big)^\top \quad \text{with}\ \mu(k) := \mathbb{E}[X_t X_{t-k}], \tag{3}$$

which obviously satisfies $\mathbb{E}[\boldsymbol{Y}_t] = \boldsymbol{0}$.

Let us start with the case of a Poisson INAR(1) process $(X_t)_{\mathbb{Z}}$. This process is α-mixing with exponentially decreasing weights (Schweer and Weiß , 2014); since $(\boldsymbol{Y}_t)_{\mathbb{Z}}$ emerges from a measurable function of $(X_t)_{\mathbb{Z}}$ in (3), it is also α-mixing with exponentially decreasing weights. Thus, Theorem 1.7 of Ibragimov (1962) is applicable to the vector-valued process $(\boldsymbol{Y}_t)_{\mathbb{Z}}$ and leads to the following central limit theorem.

Theorem 2.1. *Let $(X_t)_{\mathbb{Z}}$ be a Poisson INAR(1) process with $\mu = \frac{\beta}{1-\alpha}$, define $\boldsymbol{Y}_t$ as in formula (3). Then*

$$\begin{aligned} &\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{Y}_t \xrightarrow{\mathcal{D}} \text{N}(\boldsymbol{0}, \boldsymbol{\Sigma}) \quad \text{with}\ \boldsymbol{\Sigma} = \big(\sigma_{ij}\big)\ \text{given by} \\ &\sigma_{ij} = \mathbb{E}\left[Y_{0,i} Y_{0,j}\right] + \sum_{k=1}^{\infty} \big(\mathbb{E}[Y_{0,i} Y_{k,j}] + \mathbb{E}[Y_{k,i} Y_{0,j}]\big), \end{aligned} \tag{4}$$