



Imputation-based semiparametric estimation for INAR(1) processes with missing data

Wei Xiong¹ , Dehui Wang^{*1} , Xinyang Wang² 

¹*School of Mathematics, Jilin University, Changchun 130012, P.R. China.*

²*School of Mathematics and Systematic Sciences, Shenyang Normal University, Shenyang 110034, P.R. China.*

Abstract

In applied problems parameter estimation with missing data has risen as a hot topic. Imputation for ignorable incomplete data is one of the most popular methods in integer-valued time series. For data missing not at random (MNAR), estimators directly derived by imputation will lead results that is sensitive to the failure of the effectiveness. In view of the first-order integer-valued autoregressive (INAR(1)) processes with MNAR response mechanism, we consider an imputation based semiparametric method, which recommends the complete auxiliary variable of Yule-Walker equation. Asymptotic properties of relevant estimators are also derived. Some simulation studies are conducted to verify the effectiveness of our estimators, and a real example is also presented as an illustration.

Mathematics Subject Classification (2010). 62M10, 62F12

Keywords. Integer-valued autoregressive, semiparametric likelihood, first-step imputation, missing not at random

1. Introduction

Integer-valued time series of counts plays an important role in various practical fields, such as the number of patients who are contracted the same infection every day in the same area, or the number of people in a queue waiting for service at a certain moment. Al-Osh and Alzaid [1] proposed first-order integer-valued autoregressive (INAR(1)) processes, Du and Li [6] generalized the order to p . Since then, researches for integer-valued time series have been developing rapidly. For example, Jung et al. [11] discussed the estimation for conditional linear first-order autoregressive (CLAR(1)) processes with a Poisson marginal model; Zheng et al. [27] proposed p th-order random coefficient integer-valued autoregressive (RCINAR(p)) processes and made some statistical inferences; Zhang et al. [26] considered the statistical properties of integer-valued autoregressive processes with signed generalized power series thinning operator, and Yang et al. [25] proposed negative binomial thinning based integer-valued threshold autoregressive processes driven by independent negative-binomial distributed random variables.

*Corresponding Author.

Email addresses: xiongwei16@mails.jlu.edu.cn (W. Xiong), wangdh@jlu.edu.cn (D. Wang), wxy_math@163.com (X. Wang)

Received: 05.11.2019; Accepted: 15.09.2020

In practice, it is common to encounter missing data in many areas of statistics. Statistical analysis of missing data has attracted wide attentions in recent years. Rubin [21] introduced the statistical inference of response mechanism named missing at random (MAR); Little and Rubin [15] studied the propensity and statistical inference for both ignorable and nonignorable missing data in general regressive model. Ignorable missing data has been widely concerned in time series, Pourahmadi [19] proposed an interpolation algorithm handling missing data in stationary time series; Andersson and Karlis [2] studied the statistical inference of missing data based on imputation in INAR(1) model; Jia et al. [10] made some researches of ignorable missing data in PINAR(1) $_T$ processes.

Nonignorable missing data, which is also named as data missing not at random (MNAR), has been considered widely in recent years. Kim and Yu [13] discussed the modeling of nonignorable nonresponse and estimated the mean of response variable in the i.i.d. case, Tang et al. [23] extended this to estimate more general parameters based on empirical likelihood. Cui et al. [5] provided conditionals of identification with MNAR mechanism in generalized linear models. Shao and Wang [22] proposed a semiparametric inverse propensity weighting method using the nonresponse instrumental variable assumption of Wang et al. [24], which discussed the identification of parameter for MNAR data as well. Riddles et al. [20] proposed a propensity score adjustment method for MNAR data with a specified parametric model for the conditional distribution of respondents, and Morikawa et al. [17] loosen this restriction to a semiparametric estimation.

However, there are few literatures for nonignorable missing data in integer-valued autoregressive processes. Postulated a set of time sequence $\{X_n, n \in \mathbb{N}\}$, a missingness indicator δ_n where $\delta_n = 1$ means the value of X_n is observed and $\delta_n = 0$ otherwise, nonignorable missing data, whose response probability π_n depends on all the data not later than time n , regardless of whether it is missing, can be regard as

$$\delta_n | \mathcal{F}_n \sim \text{Bernoulli}(\pi_n), \quad n = 1, 2, \dots,$$

where \mathcal{F}_n is the σ -field generated by $\{X_n, X_{n-1}, \dots, X_0\}$. Handling nonignorable missing data is quite challenging because it is difficult to determine when it is missing. Although we know the dependent of data, estimation may be seriously biased without proper treatment. Therefore, finding the effective estimation for nonignorable nonresponse problem in time series has been developed as a main objective.

In this paper, we consider the parameter estimation of Poisson INAR(1) processes with MNAR data. We propose a semiparametric likelihood estimation for nonignorable nonresponse by employing the idea of Morikawa et al. [17], which is based on the first-step imputation. A parametric model for π_n and a nonparametric model for the distribution of observed data are allowed. Furthermore, we establish conditional observed likelihood by the joint conditional density of $\{X_n, X_{n-1}, \dots, X_1\}$ given X_0 to produce the effective estimator in INAR(1) processes.

This paper is organised as follows. In Section 2, we propose the thought of first-step imputation for substituting in INAR(1) processes. In Section 3, a semiparametric estimation for Poisson INAR(1) processes is discussed and we represent the theoretical properties of parameter estimation in Section 4. Simulation studies and a real data example to prove the efficiency of our algorithm are mentioned in Section 5 and Section 6, respectively. In addition, we consider this method for other INAR models as presenting in Section 7.

2. First-step imputation of missing data

We consider INAR(1) processes $\{X_n, n \in \mathbb{N}\}$:

$$X_n = \alpha \circ X_{n-1} + \varepsilon_n, \quad (2.1)$$

where $\{\varepsilon_n, n \in \mathbb{N}\}$ constitutes a periodic sequence of independent Poisson-distributed random variables with mean λ ($\lambda \geq 0$), which is assumed to be independent of X_{n-1} for

each n . $\alpha \in (0, 1)$. The thinning operator “ \circ ” is defined as

$$\alpha \circ X = \sum_{i=1}^X W_i, \quad W_i \stackrel{i.i.d.}{\sim} \text{Bernoulli}(\alpha),$$

and W_i is independent of X . $\theta = (\alpha, \lambda)^\top$ is the unknown parameter to be estimated.

Let (x_1, x_2, \dots, x_n) be the observations of $\{X_n, n \in \mathbb{N}\}$ with $X_0 = x_0$ and δ_i be the response indicator of x_i ($i = 1, 2, \dots, n$). Since (2.1) is strict stationary ergodicity (Du and Li [6]), the response probability for INAR(1) can be represented as

$$\pi(\phi; x_i, x_{i-1}) := p(\delta_i = 1 | X_i = x_i, X_{i-1} = x_{i-1}; \phi), \quad i = 1, 2, \dots, n, \quad (2.2)$$

where $\phi = (\phi_1, \phi_2, \dots, \phi_q)^\top \in \Phi$ is the unknown parameter and $\Phi \subseteq \mathbb{R}^r$, $q, r \in \mathbb{N}^+$.

First of all, it is necessary to estimate ϕ of the nonignorable missing mechanism. Let $f(\cdot | x_0)$ be the conditional probability given $X_0 = x_0$, the estimator of ϕ could be derived by the following conditional observed likelihood function by Kim and Shao [14]:

$$L_{obs}(\phi, \theta) = \int f(x_1, x_2, \dots, x_n, \delta | x_0; \phi, \theta) d\mu(\mathbf{x}_{mis}), \quad (2.3)$$

where \mathbf{x}_{mis} represents the missing part of X_1, \dots, X_n and μ be the Lebesgue measure.

However, expressing the explicit function of the integral in (2.3) is quite difficult because of the uncertainty of the missing time n . Notice that (2.1) admits the representation

$$\widetilde{\mathbf{X}}_n = \alpha \circ \widetilde{\mathbf{X}}_{n-1} + \varepsilon_n \quad (2.4)$$

by Yule–Walker equation with $\widetilde{\mathbf{X}}_n = (X_1, X_2, \dots, X_n)^\top$, $\widetilde{\mathbf{X}}_{n-1} = (X_0, X_1, \dots, X_{n-1})^\top$ and $\varepsilon_n = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^\top$. To avoid confusion, we denote

$$\alpha \circ \widetilde{\mathbf{X}}_{n-1} := (\alpha \circ X_0, \alpha \circ X_1, \dots, \alpha \circ X_{n-1})^\top,$$

where “ \top ” denotes the matrix transpose. Consequently, the integrand in (2.3) can be rewritten as the following from:

$$\begin{aligned} f(x_1, x_2, \dots, x_n, \delta | x_0; \phi, \theta) &= f(\widetilde{\mathbf{x}}_n, \delta | \widetilde{\mathbf{x}}_{n-1}; \phi, \theta) \\ &= f(\widetilde{\mathbf{x}}_n | \widetilde{\mathbf{x}}_{n-1}; \theta) p(\delta | \widetilde{\mathbf{X}}_n = \widetilde{\mathbf{x}}_n, \widetilde{\mathbf{X}}_{n-1} = \widetilde{\mathbf{x}}_{n-1}; \phi). \end{aligned}$$

Notice that the missing data exists in the both sides of (2.4). Traditional imputation estimation, which supplying the missingness in both $\widetilde{\mathbf{X}}_n$ and $\widetilde{\mathbf{X}}_{n-1}$, will cause the failure of the result under MNAR hypothesis, and another difficulty for incomplete INAR data is the inexistence of a full-observed sequence. Based on these, our major idea is only filling the missing part in the right side of (2.4), but allowing $\widetilde{\mathbf{X}}_n$ remains as the original structure, i.e. for any $0 < k < n$, X_{n-k} is on the both sides of (2.4), if the value of X_{n-k} is unobserved, we can fill this gap in $\widetilde{\mathbf{X}}_{n-1}$ by some ways and the site of X_{n-k} in $\widetilde{\mathbf{X}}_n$ remains missing. Reasonable imputation provides an interpolated sequence with approximate dynamic lag relationship, hence it could be regarded as the condition to make inference. Moreover, estimation with missing $\widetilde{\mathbf{X}}_n$ remedies effectiveness of directly imputation conversely. Motivated by this thought, several interpolation methods can be considered. For instance, Andersson and Karlis [2] proposed imputation based on the bridge imputation (BI), Jia et al. [10] considered subgroup mean (SM) imputation. For ease of description, we assume the interpolated value of $\widetilde{\mathbf{x}}_{n-1}$ is $\mathbf{x}_{n-1}^* = (x_0^*, x_1^*, \dots, x_{n-1}^*)$, where x_{i-1}^* be the interpolation value if $\delta_{i-1} = 0$ and $x_{i-1}^* = x_{i-1}$ ($i = 1, 2, \dots, n$) otherwise. Since we would rather making statistical inference for (2.3) by imputing $\widetilde{\mathbf{X}}_{n-1}$ at first, we name this as “first-step imputation”.

3. Semiparametric estimation

The conditional observed likelihood function (2.3) could be represented as following:

$$L_{obs}(\phi, \theta) = \int f(\tilde{\mathbf{x}}_n | \tilde{\mathbf{x}}_{n-1}; \theta) p(\delta | \tilde{\mathbf{X}}_n = \tilde{\mathbf{x}}_n, \tilde{\mathbf{X}}_{n-1} = \tilde{\mathbf{x}}_{n-1}; \phi) d\mu(\mathbf{x}_{mis}). \tag{3.1}$$

For time series data, we can see that

$$p(\delta | \tilde{\mathbf{X}}_n = \tilde{\mathbf{x}}_n, \tilde{\mathbf{X}}_{n-1} = \tilde{\mathbf{x}}_{n-1}; \phi) = \prod_{i=1}^n p(\delta_i | X_i = x_i, \tilde{X}_{i-1} = \tilde{x}_{i-1}; \phi)$$

and

$$f(\tilde{\mathbf{x}}_n | \tilde{\mathbf{x}}_{n-1}; \theta) = \prod_{i=1}^n f(x_i | \tilde{x}_{i-1}; \theta)$$

by Markovian property, thus we have

$$L_{obs}(\phi, \theta) = \prod_{\delta_i=1} f(x_i | x_{i-1}^*; \theta) \pi(\phi; x_i, x_{i-1}^*) \int \prod_{\delta_i=0} f(x_i | x_{i-1}^*; \theta) \pi_0(\phi; x_i, x_{i-1}^*) d\mu(x_i) \tag{3.2}$$

by replacing the data of $\tilde{\mathbf{X}}_{n-1}$ with \mathbf{x}_{n-1}^* , where $\pi_0(\phi; x_i, x_{i-1}^*) = 1 - \pi(\phi; x_i, x_{i-1}^*)$.

Morikawa et al. [17] proposed a semiparametric method for (3.2) in the case of independent and identical distribution data. Based on the first-step imputation, a similar way (refer to Louis [16]) is considered to establish the estimation equation for ϕ by solving

$$\frac{1}{n} \sum_{i=1}^n [\delta_i \mathbf{s}_i(\phi; x_{i-1}^*, \delta_i) + (1 - \delta_i) E\{\mathbf{s}_i(\phi; x_{i-1}^*, \delta_i) | x_{i-1}^*, \delta_i = 0\}] = \mathbf{0}, \tag{3.3}$$

where $\mathbf{s}_i(\phi; x_{i-1}^*, \delta_i) = \frac{\partial}{\partial \phi} \log[\pi^{\delta_i}(\phi; x_i, x_{i-1}^*) \pi_0^{1-\delta_i}(\phi; x_i, x_{i-1}^*)]$.

Under some regularity conditions, the conditional expectation in (3.3) can be employed the kernel smoothing to perform the nonparametric estimator by $\hat{\mathbf{C}}_{si}(\phi) / \hat{D}_i(\phi)$ that

$$\begin{aligned} \hat{\mathbf{C}}_{si}(\phi) &= (nh)^{-1} \sum_{j=1}^n \delta_j K((x_{j-1}^* - x_{i-1}^*)/h) O_j(\phi; x_{i-1}^*) \mathbf{s}_j(\phi; x_{i-1}^*, \delta_i), \\ \hat{D}_i(\phi) &= (nh)^{-1} \sum_{j=1}^n \delta_j K((x_{j-1}^* - x_{i-1}^*)/h) O_j(\phi; x_{i-1}^*), \end{aligned} \tag{3.4}$$

where $O_j(\phi; x_{i-1}^*) = \pi_0(\phi; x_j, x_{i-1}^*) / \pi(\phi; x_j, x_{i-1}^*)$, $K(\cdot)$ is a kernel function defined on \mathbb{R} with the bandwidth h . Therefore, (3.3) could be written as

$$\frac{1}{n} \sum_{i=1}^n [\delta_i \mathbf{s}_i(\phi; x_{i-1}^*, \delta_i) + (1 - \delta_i) \hat{\mathbf{C}}_{si}(\phi) / \hat{D}_i(\phi)] = \mathbf{0}, \tag{3.5}$$

and the estimator will be given as $\hat{\phi}_n$.

To estimate θ , we consider a strategy inspired by conditional maximum likelihood (CML) estimation in complete case for INAR(1) processes. Let $\hat{\theta}_{CML}$ be the CML estimator,

$$\hat{\theta}_{CML} = \arg \max_{\theta} \prod_{i=1}^n f(x_i | x_{i-1}; \theta).$$

The conditional probability density function $f(x_i | x_{i-1}; \theta)$ is given by Al-Osh and Alzaid [1], thus $\hat{\theta}$ in this paper by maximizing $L_{obs}(\hat{\phi}, \theta)$, which is equivalent to solving the following score equation:

$$\frac{1}{n} \sum_{i=1}^n [\delta_i \mathbf{P}_i(\theta) + (1 - \delta_i) E\{\mathbf{l}_i(\theta) | x_{i-1}^*, \delta_i = 0\}] = \mathbf{0}, \tag{3.6}$$

where $\mathbf{l}_i(\theta) = \frac{\partial}{\partial \theta} \log f(x_i | x_{i-1}; \theta)$.

Equation (3.6) can be solved in a similar way because $E\{\mathbf{l}_i(\boldsymbol{\theta})|x_{i-1}^*, \delta_i = 0\}$ could be approximated by the kernel smoothing method again. The left side of (3.6) can be estimated by

$$\frac{1}{n} \sum_{i=1}^n [\delta_i \mathbf{l}_i(\boldsymbol{\theta}) + (1 - \delta_i) \hat{\mathbf{C}}_{li}(\boldsymbol{\theta}, \hat{\boldsymbol{\phi}}) / \hat{D}_i(\hat{\boldsymbol{\phi}})] = \mathbf{0}, \tag{3.7}$$

where

$$\hat{\mathbf{C}}_{li}(\boldsymbol{\theta}, \hat{\boldsymbol{\phi}}) = (nh)^{-1} \sum_{j=1}^n \delta_j K((x_{j-1}^* - x_{i-1}^*)/h) O_j(\hat{\boldsymbol{\phi}}; x_{i-1}^*) \mathbf{l}_j(\boldsymbol{\theta}).$$

and an effective estimator $\hat{\boldsymbol{\theta}}_n$ will be obtained.

4. Asymptotic properties

In this section, we discuss the asymptotic properties of proposed estimation. We focus on the consistency and asymptotic normality of $\hat{\boldsymbol{\phi}}_n$ and $\hat{\boldsymbol{\theta}}_n$ with sample size n . In what follows, $E_0[\cdot] := E[\cdot|\delta = 0]$, $E_1[\cdot] := E[\cdot|\delta = 1]$, $f_1(\cdot) := f(\cdot|\delta = 1)$, and $\|\cdot\|$ denotes the Euclidean norm (for example, for a matrix $A = (a_{ij})$, $\|A\| = (\sum_{i,j} a_{ij}^2)^{1/2}$). Firstly, the following necessary regularity conditions are given to study properties of $\hat{\boldsymbol{\phi}}_n$:

- (A1) Let \mathfrak{X} be a compact set that is contained in the support of \mathbf{X} , for all $n = 1, 2, \dots$ and all $\boldsymbol{\phi}$, $X_i \in \mathfrak{X}$ and $\pi(\boldsymbol{\phi}; x_n, x_{n-1}) \in (0, 1)$;
- (A2) $E[\frac{\partial}{\partial \boldsymbol{\phi}} \pi(\boldsymbol{\phi}_0; x_n, x_{n-1}) \frac{\partial}{\partial \boldsymbol{\phi}^\top} \pi(\boldsymbol{\phi}_0; x_n, x_{n-1})]$ is positive definite, where $\boldsymbol{\phi}_0 \in \boldsymbol{\Phi}$ is the true value of $\boldsymbol{\phi}$;
- (A3) $\pi(\boldsymbol{\phi}; x_n, x_{n-1})$ is identifiable, continuously differentiable until order 3 with probability one, $\|\frac{\partial}{\partial \boldsymbol{\phi}} \pi(\boldsymbol{\phi}; x_n, x_{n-1})\|$, $\|\frac{\partial^2}{\partial \boldsymbol{\phi}^\top \partial \boldsymbol{\phi}} \pi(\boldsymbol{\phi}; x_n, x_{n-1})\|$ are bounded on an open set containing \mathfrak{X} and $E[\|\frac{\partial}{\partial \boldsymbol{\phi}} \log \pi(\boldsymbol{\phi}; x_n, x_{n-1})\|]$, $E[\|\frac{\partial^2}{\partial \boldsymbol{\phi}^\top \partial \boldsymbol{\phi}} \log \pi(\boldsymbol{\phi}; x_n, x_{n-1})\|]$, $E[\|\frac{\partial^3}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^\top \partial \boldsymbol{\phi}} \log \pi(\boldsymbol{\phi}; x_n, x_{n-1})\|]$ are finite;
- (A4) The kernel $K(u)$ has bounded derivative until order 3 satisfies that $\int K(u)du = 1$, and has zero moments of $m \geq 4$ and a nonzero m th order moment;
- (A5) For $p \geq 4$, $E_1[\|\frac{\partial}{\partial \boldsymbol{\phi}} \log \pi(\boldsymbol{\phi}; x_n, x_{n-1})\|^p]$, $E_1[\|\frac{\partial}{\partial \boldsymbol{\phi}} \log \pi(\boldsymbol{\phi}; x_i, x_{n-1})\|^p |x_{n-1}] f_1(x_{n-1})$, $E_1[\|\pi^{-1}(\boldsymbol{\phi}; x_n, x_{n-1})\|^p]$ and $E_1[\|\pi^{-1}(\boldsymbol{\phi}; x_n, x_{n-1})\|^p |x_{n-1}] f_1(x_{n-1})$ are bounded on an open set containing \mathfrak{X} ;
- (A6) The bandwidth h satisfies that: (i) $h \rightarrow 0$, (ii) for $p \geq 4$, $n^{1-\frac{2}{p}} h / \log n \rightarrow \infty$, (iii) $\sqrt{nh}^6 / \log n \rightarrow \infty$ and (iv) for $m \geq 4$, $\sqrt{nh}^{2m} \rightarrow \infty$.

Conditions (A4)-(A6) provide the kernel conditions for the nonparameter estimator (3.6). Newey and McFadden [18] proved that for $i = 1, 2, \dots, n$ and $\boldsymbol{\phi} \in \boldsymbol{\Phi}$,

$$\begin{aligned} \|\hat{\mathbf{C}}_{si}(\boldsymbol{\phi}) - \mathbf{C}_{si}(\boldsymbol{\phi})\|_k &= o_p(n^{-1/4}), \\ \|\hat{D}_i(\boldsymbol{\phi}; x_{i-1}^*) - D_i(\boldsymbol{\phi}; x_{i-1}^*)\|_k &= o_p(n^{-1/4}), \end{aligned} \tag{4.1}$$

where $\|\cdot\|_k$ is defined as Sobolev norm which $\|\gamma(x)\|_k := \max_{j \leq k} \sup_{x \in \mathfrak{X}} \|\frac{\partial^j}{\partial x^j} \gamma(x)\|$ for all the distinct j th-order partial derivatives of all elements of $\gamma(x)$ and nonnegative integer k .

The asymptotic properties of $\hat{\boldsymbol{\phi}}_n$ can be summarized from the following theorem:

Theorem 4.1. *Under conditions (A1)-(A6), there exists a maximizer $\hat{\boldsymbol{\phi}}_n \in \boldsymbol{\Phi}$ of (3.2) such that*

- (i) (Consistency) $\hat{\boldsymbol{\phi}}_n \xrightarrow{p} \boldsymbol{\phi}_0$, as $n \rightarrow \infty$,
- (ii) (Asymptotic Normality) $\sqrt{n}(\hat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0) \xrightarrow{L} N(\mathbf{0}, \mathbf{I}_\phi^{-1} \mathbf{W}_\phi \mathbf{I}_\phi^{-1})$, as $n \rightarrow \infty$,

where \mathbf{I}_ϕ , \mathbf{W}_ϕ are presented in the proof.

Let $\hat{\mathbf{S}}_n(\phi) = n^{-1} \sum_{i=1}^n [\delta_i \mathbf{s}_i(\phi; x_{i-1}^*, \delta_i) + (1 - \delta_i) \hat{\mathbf{C}}_{s_i}(\phi) / \hat{D}_i(\phi)]$, it can be reformed to a linearization via Lemma 4.2 of Morikawa et al [17] and Khashimov [12], that is

$$\hat{\mathbf{S}}_n(\phi) = n^{-1} \sum_{i=1}^n \left[1 - \frac{\delta_i}{\pi(\phi; x_i, x_{i-1}^*)} \right] E_0 \{ \mathbf{s}_i(\phi; x_{i-1}^*, \delta_i) | x_{i-1}^* \} + o_p(n^{-1/2}). \tag{4.2}$$

Denote the influence functions of $\hat{\mathbf{S}}_n(\phi)$ by $\hat{\mathbf{S}}_n^l(\phi)$, it can be proved that

$$\sup_{\phi} \| \hat{\mathbf{S}}_n^l(\phi) - \mathbf{S}_n(\phi) \| \rightarrow 0, \quad \text{as } n \rightarrow \infty, \tag{4.3}$$

where $\mathbf{S}_n(\phi) = n^{-1} \sum_{i=1}^n [\delta_i \mathbf{s}_i(\phi; x_{i-1}^*, \delta_i) + (1 - \delta_i) E_0 \{ \mathbf{s}_i(\phi; x_{i-1}^*, \delta_i) | x_{i-1}^* \}]$. There remains to verify the asymptotic properties of $\hat{\phi}_n$ in $\hat{\mathbf{S}}_n^l(\phi)$, it requires us give a following lemma:

Lemma 4.2. *Let $\mathcal{F}_n = \sigma(X_n, X_{n-1}, \dots, X_0)$ be a σ -field, under (A1)-(A6), for $n \rightarrow \infty$, $\{ \hat{\mathbf{S}}_n^l(\phi), \mathcal{F}_n, n \geq 1 \}$ is an asymptotic martingale sequence.*

Proof of Lemma 4.2. Let $M_i(\phi)$ label the i -th elements of $n\mathbf{S}_n(\phi)$ and $\mathbf{S}_n^{(1)}(\phi) = \sum_{i=1}^n M_i(\phi)$, we consider a σ -field as $\mathcal{F}_{\delta_n} = \sigma(X_n, \delta_n, X_{n-1}, \delta_{n-1}, \dots, \delta_1, X_0)$ such that $\mathcal{F}_n \subset \mathcal{F}_{\delta_n}$. Following Theorem 2.5 of Kim and Shao [14] we can obtain

$$\begin{aligned} E\{ \mathbf{S}_n^{(1)}(\phi) | \mathcal{F}_{n-1} \} &= E\{ E[\mathbf{S}_n^{(c)}(\phi) | \mathcal{F}_{\delta_n}] | \mathcal{F}_{n-1} \} \\ &= E\{ \mathbf{S}_{n-1}^{(c)}(\phi) | \mathcal{F}_{\delta_{n-1}} \} \\ &= \mathbf{S}_{n-1}^{(1)}(\phi), \end{aligned} \tag{4.4}$$

where $\mathbf{S}_n^{(c)}(\phi)$ is the score function with complete data of INAR(1) satisfies

$$E\{ \mathbf{S}_n^{(c)}(\phi) | \mathcal{F}_{n-1} \} = \mathbf{S}_{n-1}^{(c)}(\phi) + E\left\{ \frac{\partial}{\partial \phi} \log f(x_n, \delta_n = 1 | x_{n-1}; \phi, \theta) | \mathcal{F}_{n-1} \right\},$$

and

$$E\left\{ \frac{\partial}{\partial \phi} \log f(x_n, \delta_n = 1 | x_{n-1}; \phi, \theta) | \mathcal{F}_{n-1} \right\} = \mathbf{0},$$

which implies that $\{ \mathbf{S}_n(\phi), \mathcal{F}_n, n \geq 1 \}$ be a martingale, therefore, as n tends to infinity, $\{ \hat{\mathbf{S}}_n^l(\phi), \mathcal{F}_n, n \geq 1 \}$ is an asymptotic martingale by (4.3). \square

Lemma 4.2 proved the martingale sequence from the construction of (4.2). The laws of large number and center limits theorem of martingale could be used to prove theorem 4.1.

Proof of Theorem 4.1. Let

$$\hat{M}_i^l(\phi) = \left[1 - \frac{\delta_i}{\pi(\phi; x_i, x_{i-1}^*)} \right] E_0 \{ \mathbf{s}_i(\phi; x_{i-1}^*, \delta_i) | x_{i-1}^* \},$$

from lemma 4.2, it is easy to see that

$$\hat{\mathbf{S}}_n^l(\phi) \xrightarrow{p} E[\mathbf{M}_n(\phi)], \quad \text{as } n \rightarrow \infty,$$

by the laws of large number. With the strict stationary ergodicity of $\{X_n, n \in \mathbb{N}\}$, we can show that

$$\begin{aligned} E[\mathbf{M}_n(\phi_0)] &\rightarrow \mathbf{0} \quad \text{as } n \rightarrow \infty, \\ \text{Var}(\mathbf{M}_n(\phi_0)) &= E[\mathbf{M}_n(\phi_0) \mathbf{M}_n^\top(\phi_0)] := \mathbf{I}_\phi, \end{aligned}$$

and

$$E\left[\frac{\partial}{\partial \phi^\top} \mathbf{M}_n(\phi_0) \right] = -\mathbf{I}_\phi, \quad \text{as } n \rightarrow \infty.$$

Let

$$\hat{\mathbf{S}}_n^*(\phi) = \sum_{i=1}^n \left\{ \frac{\partial}{\partial \phi^\top} \hat{M}_i^l(\phi) - E\left[\frac{\partial}{\partial \phi^\top} \hat{M}_i^l(\phi) \right] \right\},$$

where $\frac{\partial}{\partial \phi^\top} \hat{M}_i^l(\phi_0)$ denotes $\frac{\partial}{\partial \phi^\top} \hat{M}_i^l(\phi)|_{\phi=\phi_0}$, we can also prove that $\{\hat{S}_n^*(\phi), \mathcal{F}_n, n \geq 1\}$ is a martingale and

$$\hat{S}_n^*(\phi_0) = \sum_{i=1}^n [\frac{\partial}{\partial \phi^\top} \hat{M}_i(\phi_0) + \mathbf{I}_\phi],$$

so that

$$\|\frac{1}{n} \hat{S}_n^*(\phi_0)\| \xrightarrow{p} 0, \quad n \rightarrow \infty.$$

Conditions (B1)-(B3) of Basawa et al. [3] hold and, furthermore, it can be proved that (3.1) has a root $\hat{\phi}_n$ which is consistent to ϕ_0 as $n \rightarrow \infty$. (See also in section 6.2, Hall and Heyde [8])

To derive the asymptotic distribution of $\hat{\phi}_n$, we consider Taylor series expansion at $\phi = \phi_0$ that

$$\mathbf{0} = \hat{S}_n^l(\hat{\phi}_n) = \sqrt{n} \hat{S}_n^l(\phi_0) + \frac{\partial}{\partial \phi^\top} \hat{S}_n^l(\phi_0) \sqrt{n}(\hat{\phi}_n - \phi_0) + \frac{1}{2} \frac{\partial}{\partial \phi \partial \phi^\top} \hat{S}_n^l(\phi_1) \sqrt{n}(\hat{\phi}_n - \phi_0)^2,$$

where ϕ_1 values between $\hat{\phi}_n$ and ϕ_0 . Through the arguments of martingale central limit theorem (Hall and Heyde [8]),

$$\frac{1}{\sqrt{n}} \hat{S}_n^l(\phi_0) \xrightarrow{L} N(\mathbf{0}, \mathbf{W}_\phi),$$

where $\mathbf{W}_\phi = E\{\hat{S}_n^l(\phi_0) \hat{S}_n^{l\top}(\phi_0)\}$. On the other hand,

$$\|\frac{\partial}{\partial \phi \partial \phi^\top} \hat{S}_n^l(\phi_1) - E[\frac{\partial}{\partial \phi \partial \phi^\top} \mathbf{M}_n(\phi_1)]\| \xrightarrow{p} 0, \quad n \rightarrow \infty,$$

therefore

$$\|\frac{\partial^2}{\partial \phi \partial \phi^\top} \hat{S}_n^l(\phi_1)(\hat{\phi}_n - \phi_0)\| = o_p(1)$$

from the consistency of $\hat{\phi}_n$. Therefore, the (ii) of theorem 4.1 follows by the Slutsky theorem. The proof is completed. \square

The following regularity conditions for INAR(1) processes are assumed as below, which are necessary for studying the asymptotic properties of $\hat{\theta}_n$.

- (B1) $\{X_n, n \in \mathbb{N}\}$ has finite moments until order 3, that is, for $k \leq 3$, $EX_n^k < \infty$;
- (B2) There exists $p \geq 4$ such that $E_1[\|\mathbf{l}(\theta)\|^p]$ and $E_1[\|\mathbf{l}(\theta)\|^p | x_{i-1}] f_1(x_{i-1})$ are bounded, where $\mathbf{l}(\theta) = \sum_{i=1}^n \mathbf{l}_i(\theta)$.

Theorem 4.3. Under conditions (A4)-(A6) and (B1)-(B2), the estimator $\hat{\theta}_n$ in (3.9) satisfies

$$(i)(Consistency) \quad \hat{\theta}_n \xrightarrow{p} \theta_0, \quad \text{as } n \rightarrow \infty,$$

$$(ii)(Asymptotic Normality) \quad \sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{L} N(\mathbf{0}, \mathbf{I}_\theta^{-1} \mathbf{W}_\theta \mathbf{I}_\theta^{-1}), \quad \text{as } n \rightarrow \infty,$$

where \mathbf{I}_θ and \mathbf{W}_θ are presented in the proof.

Note that the left side of (3.6) is the score function of the conditional observed likelihood (2.3) for θ , let $\mathbf{U}_n(\theta) = \sum_{i=1}^n [\delta_i \mathbf{l}_i(\theta) + (1 - \delta_i) E_0\{\mathbf{l}_i(\theta) | x_{i-1}^*\}]$, we have $\mathbf{U}_n(\theta) = \frac{\partial}{\partial \phi} L_{obs}(\theta)$. Let $\hat{\mathbf{U}}_n(\theta)$ be the estimation of (3.6), it can also be represented as a linearization that

$$\hat{\mathbf{U}}_n(\theta) = n^{-1} \sum_{i=1}^n [\frac{\delta_i \mathbf{l}_i(\theta)}{\pi(\hat{\phi}_n; x_i, x_{i-1}^*)} + (1 - \frac{\delta_i}{\pi(\hat{\phi}_n; x_i, x_{i-1}^*)}) E_0\{\mathbf{l}_i(\theta) | x_{i-1}^*\}] + o_p(n^{-1/2}),$$

and for all $\theta \in \Theta$, the influence function $\hat{\mathbf{U}}_n^l(\theta)$ converges to $\mathbf{U}_n(\theta)$ as n tends to infinity. Similar to lemma 4.2, we can prove that $\{\hat{\mathbf{U}}_n^l(\theta), \mathcal{F}_n, n \geq 1\}$ is an asymptotic martingale sequence. It inspires us considering as the same proof as theorem 4.1.

Proof of Theorem 4.3. Let

$$\hat{N}_i^l(\boldsymbol{\theta}) = \frac{\delta_i \mathbf{l}_i(\boldsymbol{\theta})}{\pi(\hat{\boldsymbol{\phi}}_n; x_i, x_{i-1}^*)} + (1 - \frac{\delta_i}{\pi(\hat{\boldsymbol{\phi}}_n; x_i, x_{i-1}^*)}) E_0\{\mathbf{l}_i(\boldsymbol{\theta}) | x_{i-1}^*\},$$

$$\mathbf{N}_i(\boldsymbol{\theta}) = \delta_i \mathbf{l}_i(\boldsymbol{\theta}) + (1 - \delta_i) E_0\{\mathbf{l}_i(\boldsymbol{\theta}) | x_{i-1}^*\},$$

one can see that for $n \rightarrow \infty$,

$$\hat{U}_n^l(\boldsymbol{\theta}) \xrightarrow{p} E[\mathbf{N}_n(\boldsymbol{\theta})].$$

Denote

$$\hat{U}_n^*(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ \frac{\partial}{\partial \boldsymbol{\theta}^\top} \hat{N}_i^l(\boldsymbol{\theta}) - E\left[\frac{\partial}{\partial \boldsymbol{\theta}^\top} \hat{N}_i^l(\boldsymbol{\theta}) \right] \right\},$$

it can be found the martingale sequence of $\{\hat{U}_n^*(\boldsymbol{\theta}), \mathcal{F}_n, n \geq 0\}$ and similar with theorem 4.1, $\hat{\boldsymbol{\theta}}_n$ is converge to the true value $\boldsymbol{\theta}_0$ in probability as n increasing to infinity.

Freeland and McCabe [7] proved the asymptotic normality for $\hat{\boldsymbol{\theta}}_{CML}$ in the case of complete data, the proposition 3 can directly derive the martingale sequences for score function $\mathbf{l}_i(\boldsymbol{\theta})$. Since $\{\hat{U}_n^l(\boldsymbol{\theta}), \mathcal{F}_n, n \geq 1\}$ is an asymptotic martingale sequence, thus

$$\frac{1}{\sqrt{n}} \hat{U}_n^l(\boldsymbol{\theta}_0) \xrightarrow{L} N(\mathbf{0}, \mathbf{W}_\theta),$$

by martingale central limit theorem with $\mathbf{W}_\phi = E[\hat{N}_n^l(\boldsymbol{\theta}_0) \hat{N}_n^{l\top}(\boldsymbol{\theta}_0)]$. Similarly, we can prove that

$$\frac{\partial}{\partial \boldsymbol{\theta}^\top} \hat{U}_n^l(\boldsymbol{\theta}_0) \xrightarrow{p} E\left[\frac{\partial}{\partial \boldsymbol{\theta}^\top} \mathbf{l}_n(\boldsymbol{\theta}_0) \right] := \mathbf{I}_\theta$$

and

$$\left\| \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \hat{U}_n^l(\boldsymbol{\theta}_1) (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_1) \right\| = o_p(1),$$

as $n \rightarrow \infty$ with $\boldsymbol{\theta}_1$ values between $\hat{\boldsymbol{\theta}}_n$ and $\boldsymbol{\theta}_0$. Therefore, by Taylor series expansion, the asymptotic distribution for $\hat{\boldsymbol{\theta}}_n$ can be given as

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{L} N(\mathbf{0}, \mathbf{I}_\theta^{-1} \mathbf{W}_\theta \mathbf{I}_\theta^{-1}), \quad n \rightarrow \infty,$$

□

Remark 1 Note that asymptotic properties of estimators are provided by the strictly stationary ergodicity of $\{X_n, n \in \mathbb{N}\}$, therefore, unbiased results by our method will be established under other stationary INAR models in the same way. This will be confirmed in section 7 with numerical studies.

Remark 2 In theorem 4.1 and theorem 4.3, asymptotic variances are constructed by the imputed value x_{i-1}^* ($i = 1, \dots, n$), which indicates that the first-step imputation somehow influences the efficiency of the estimator. In practice, a proper choice for the imputation method depends on the modeling of π_t . The sample size, response rates and $\boldsymbol{\theta}_0$ will also dominate the option in general. Refer to section 5 for related simulation results.

5. Simulation studies

To test the performance of our theory described in previous sections, we perform several simulation studies to examine our algorithm for Poisson INAR(1) processes for 4 kinds of parameter combinations. The response mechanism for each scenario is specified as a log-log model $\pi_t = 1 - \exp(-\exp(\phi_1 + \phi_2 X_t))$. The true value of $\boldsymbol{\theta} = (\alpha, \lambda)$ and $\boldsymbol{\phi}$ is designed as following:

- Model 1. $\boldsymbol{\theta} = (0.3, 1)$ with the mechanism parameter $\boldsymbol{\phi}_1$;
- Model 2. $\boldsymbol{\theta} = (0.3, 2)$ with the mechanism parameter $\boldsymbol{\phi}_2$;
- Model 3. $\boldsymbol{\theta} = (0.6, 1)$ with the mechanism parameter $\boldsymbol{\phi}_3$;
- Model 4. $\boldsymbol{\theta} = (0.6, 2)$ with the mechanism parameter $\boldsymbol{\phi}_4$;

For each model, we use 500 independent replications for different series with the initial value $X_0 \sim \text{Poisson}(\lambda)$. The response rates are controlled at about 85% and 70% respectively by adjusting ϕ . Simulations are carried out to evaluate efficiency of $\hat{\theta}_n$ under consideration of four strategies: (i) Subgroup Mean Imputation estimation (SM); (ii) Bridge-Imputation estimation (BI); (iii) First-Step Imputation Semiparametric estimation by Subgroup Mean Imputation (SM-FISE) and (iv) First-Step Imputation Semiparametric estimation by Bridge-Imputation (BI-FISE).

In the subgroup mean imputation, we use the rounding of $(\sum_i \delta_i)^{-1} \sum_{\delta_i=1} X_i$ to fill up each missing part in \tilde{X}_{n-1} . In the bridge imputation, we consider the algorithm as following:

step 1. Specify an observed series X_{obs} by deleting the missing part of X_n , estimate a start value θ_1 from X_{obs} ;

step 2. For some $k > 1$, denote X_i and X_{i+k} are observed points and all the points that are between are missing. The BI algorithm proceeds by simulating value X_{i+r+1} conditional on the previous point X_{i+r} ($r = 0, \dots, k - 1$) by (2.1) with θ_1 ;

step 3. Simulate the value of X_{i+k} as X_{i+k}^* , if X_{i+k}^* coincides with X_{i+k} to use this path filling all the missingness, otherwise repeat step 2 to generate a new path;

step 4. Use the entire series to simulate an estimator θ_2 . Repeat step 2-4 with the parameter values obtained in the last iteration until $\|\theta_1 - \theta_2\| < \epsilon$ and select each X_{i+r+1}^* in step 3 to replace the missingness of \tilde{X}_{n-1} .

For the semiparametric simulation study, the nonparametric kernel regression estimator of (3.5), (3.8) is computed using a Gaussian kernel function $K(u) = (2\pi)^{-1/2} \exp(-u^2/2)$. The bandwidth is selected as $h = 1.5\hat{\sigma}_x n^{-1/3}$ which is considered in Chen et al. [4], Shao and Wang [22] and Morikawa et al. [17], where $\hat{\sigma}_x$ is the estimated standard deviation of X_n in the sample. Moreover, the convergence criteria of bridge imputation is set as $\epsilon = 10^{-2}$.

Table 1–4 summarized the results for each $\hat{\theta}_n$ from the scenario Monte Carlo samples of size $n = 100, 300$ with 85% and 70% response rates, respectively. In each table the empirical bias (Bias), mean squared errors (MSE), standard deviations (SD) and standard errors (SE) are presented. Otherwise, the approximately 95% confidence intervals (CI) of the parameter with coverage rate (CR) are reported, which is constructed based on $\hat{\theta}_n \pm 1.96SE$ and SE is the square root of diagonal elements of $n^{-1} \sum_{i=1}^n \mathbf{I}_{\hat{\theta}_n}^{-1} \mathbf{W}_{\hat{\theta}_n} \mathbf{I}_{\hat{\theta}_n}^{-1}$. In particular, we compute CI and CR by substituting the true value to the missing part of \tilde{X}_{n-1} to compare the asymptotic normality influenced by different first-step imputations.

It is worth to show that estimations directly from imputation are much more seriously biased obviously biased in terms of nonignorable missing data, especially for $\hat{\lambda}_n$, our method improves the accuracy of imputation estimation to a certain extent. As the sample size increasing, desired results of SE-FISE and BI-FISE are maintained throughout different response rates as well. On the other hand, we found that regardless of the missing rates SE-FISE generally performs less bias and MSE than BI-FISE under the circumstances of 100 sample size in model 1 and model 2, while BI-FISE works better in model 3, 4. The convergence speed of BI-FISE performs faster as sample size increasing as well. This is mainly because the subgroup imputation for INAR(1) reduces the sampling variability in the data. The imputed values with deterministic imputation cause the inaccurate characterization of the missingness as data varying far away from EX_n (as shown in Figure 1–4), the increasing n deeps this problem in the same way. However, bridge imputation can be problematic contrarily. Therefore, when α and λ are specified small, SM-FISE could be considered as an efficient estimation because of saving time for the iterative process. Moreover, the coverage probabilities of 95% confidence intervals is decreasing with the increase of missing rates. Due to different properties of imputations CR of BI-FISE performs a little bit stable and as in direct proportion to α and λ , it computed around 90%

in model 3, 4 but opposite for SM-FISE. In summary, choosing an appropriate imputation method for the semiparametric estimation in terms of preliminary inference of models will lead a better conclusion.

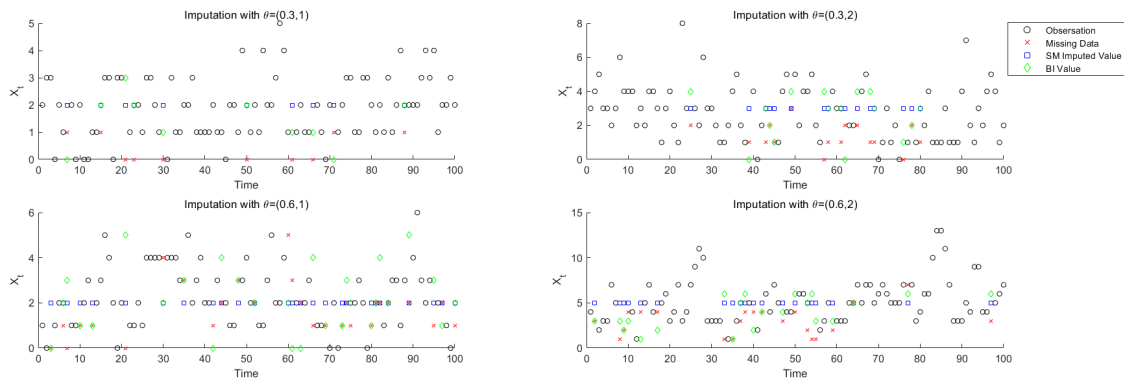


Figure 1. Scattergram of observations and imputations for INAR(1) with 100 sample size, 85% response rates.

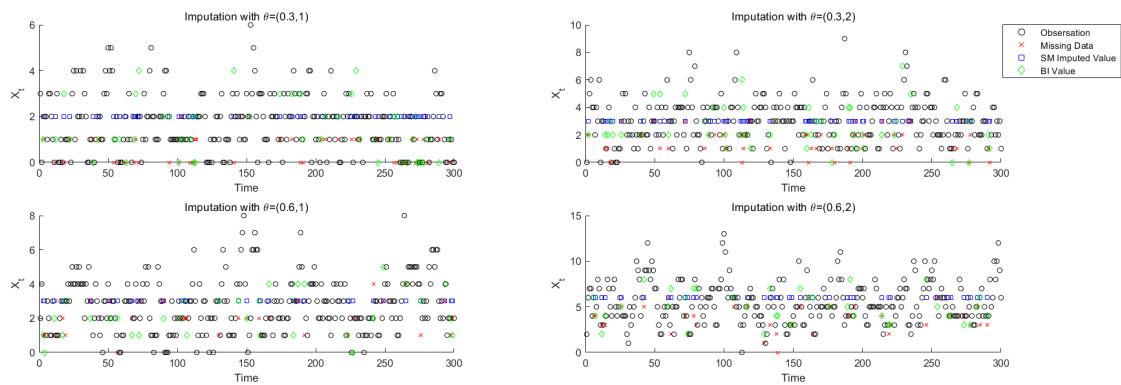


Figure 2. Scattergram of observations and imputations for INAR(1) with 300 sample size, 85% response rates.

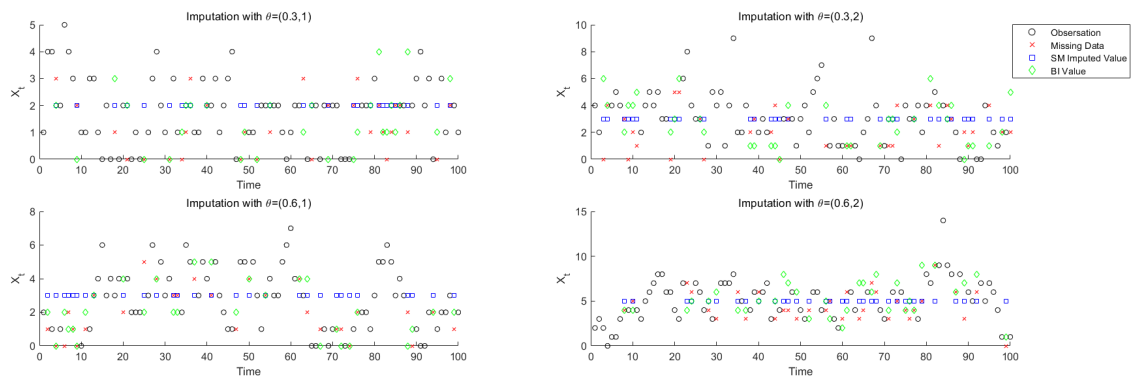


Figure 3. Scattergram of observations and imputations for INAR(1) with 100 sample size, 70% response rates.

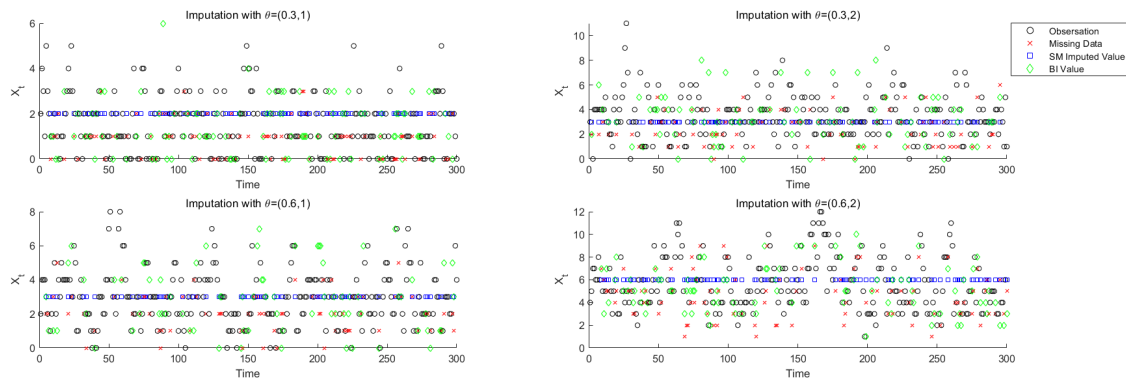


Figure 4. Scattergram of observations and imputations for INAR(1) with 300 sample size, 70% response rates.

6. Empirical analysis

In this section, we apply the proposed algorithm to the monthly counts of Criminal mischiefs in the 21 police car beat in Pittsburgh from January 1990 to December 2001, which could be extracted from the website of Forecasting Principles (<http://www.forecastingprinciples.com>), consisting of 144 monthly observations that the sample path and ACF, PACF plot are presented in Figure 5, Figure 6 and Figure 7, respectively. From ACF and PACF of the criminal data, it is reasonable to show that the analysed data set can be fitted by Poisson INAR(1) processes. The CML estimator is (0.2335, 4.5205) in a complete case, which is specified as a measure of estimations.

We create artificial missing conditions by defining response mechanisms for the sample described above. Specifically, 3 groups of models are used:

- (i) (MCAR) $\pi_t = \phi_1$;
- (ii) (Log-log) $\pi_t = 1 - \exp(-\exp(\phi_1 + \phi_2 X_t))$;
- (iii) (Logistic) $\pi_t = 1/(1 + \exp(\phi_1 + \phi_2 X_t))$;

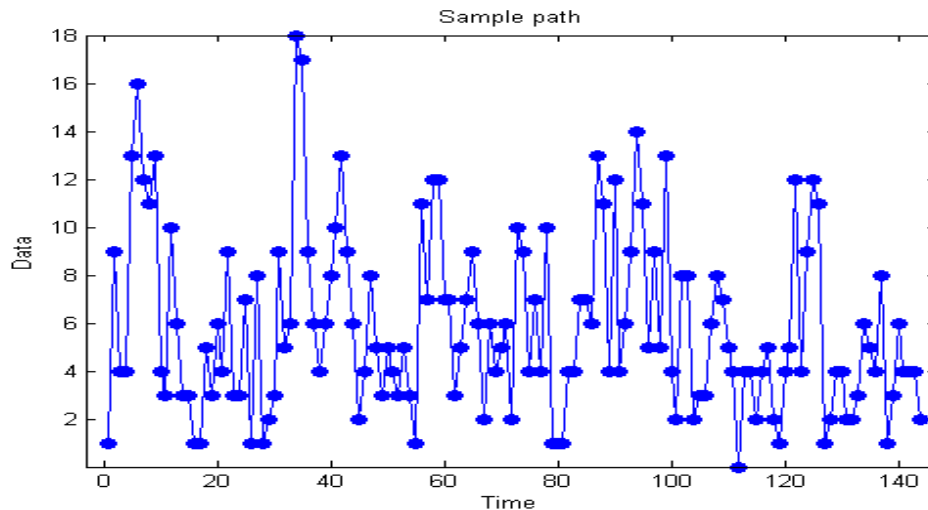


Figure 5. Sample path plot of criminal mischiefs in the 21 police car beat in Pittsburgh.

The missing rates are controlled at about 15% and 30% with different $\phi = (\phi_1, \phi_2)$. For each mechanism, we simulate 50 independent replications and the convergence criteria

$\epsilon = 0.08$ to improve the efficiency of BI. the Bias, MSE, SD and SE of parameters estimated by algorithms mentioned in section 5 are compared.

Table 5–6 report the differences of indicators between each estimator and θ_0 with different response rates. Under response mechanisms (i), (ii) and (iii), the Bias based on two kinds of imputation estimators are consistently larger than our FISE estimators whenever response rate valued 85% or 70% and SD, SE display the same performance as well. Therefore, this empirical study demonstrates a better effectiveness for our method.

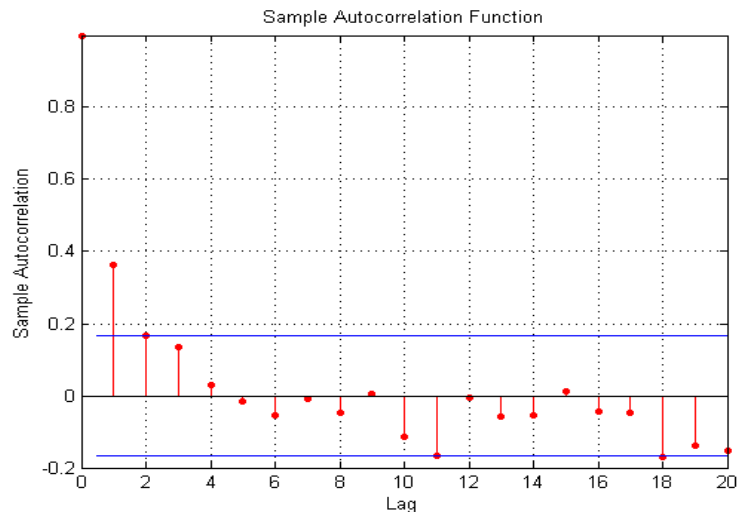


Figure 6. ACF plot of criminal mischiefs in the 21 police car beat in Pittsburgh.

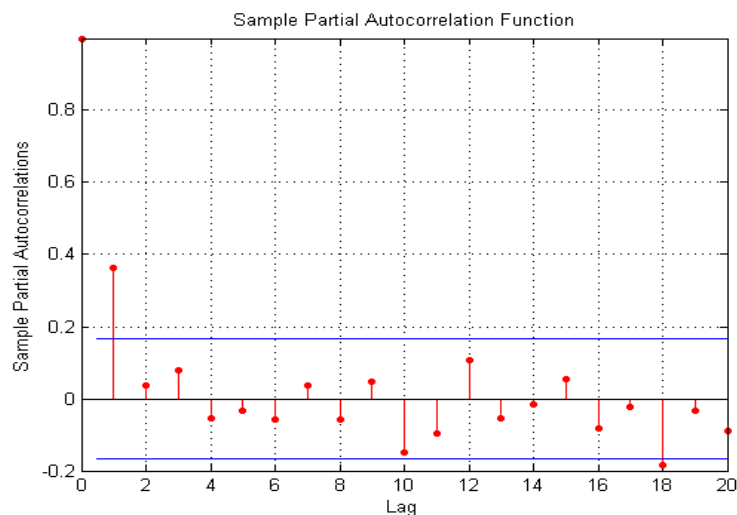


Figure 7. PACF plot of criminal mischiefs in the 21 police car beat in Pittsburgh.

7. Numerical experiments for diverse INAR models

In this section, we perform some numerical experiments for some other integer-valued autoregressive processes to verify that first-step imputation semiparametric estimation is universal. Typically, we consider the estimation for stationary integer-valued autoregressive with geometric innovations (INARG(1)) in terms of nonignorable missing data, which

is well-discussed in Jazi [9]. In this paper we pay our attention to the following processes:

$$X_t = \alpha \circ X_{t-1} + \varepsilon_t, \quad \varepsilon_t \stackrel{i.i.d.}{\sim} \text{Geometric}(\lambda).$$

We make the numerical simulation with parameters $(\alpha, \lambda) = (0.3, 0.5)$ and $(0.5, 0.5)$ respectively, the initial value $X_0 \sim \text{Geometric}(\lambda)$. The nonignorable response mechanism for each model is specified as $\pi_t = 1 - \exp(-\exp(\phi X_t))$, the same indicators are adopted as in section 5 for 500 independent replications with various sample size and response rates as well. As shown in Table 7–8, expected results for first-step imputation semiparametric estimation are performed in the scenarios, where BI-FISE performs best in all scenarios. It concludes that our method will not lose the efficient for other INAR processes.

8. Conclusion

In this article, we introduced an estimation method for which observations are incomplete in Poisson INAR(1) processes. When the response mechanism is nonignorable, the effective semiparametric estimator (3.9) based on a first-step imputation could be obtained. Asymptotic properties, simulation and empirical results have been presented to prove our estimators provide satisfactory performances in general. For different parameters we discuss the division between SM-FISE and BI-IFSE, our method can be seen as an improvement for the directly imputation estimation. It will be suitable for other parameter estimation of stationary integer-valued time series.

There are some more points that we did not pursue in this article. In the nonparametric part, choice of the bandwidth h in kernel function has not been discussed at length in this paper, some algorithms like cross-validation could be considered to improve estimation effects but this still be a computational challenging problem. On the other hand, if the response mechanism π_i becomes unknown or under a class of candidate specifications $\{\pi_{ij}, i = 1, 2, \dots, I; j = 1, 2, \dots, J\}$, propensities of estimation may be challenging. In addition, the identification of parameters in complex time series under nonignorable response mechanisms is also an impeding problem. These remains topics of future research.

Acknowledgment. This work is supported by National Natural Science Foundation of China (No. 11871028, 11731015, 11901053), Natural Science Foundation of Jilin Province (No. 20180101216JC), Program for Changbaishan Scholars of Jilin Province (2015010).

Table 1. Simulation results for the INAR(1)'s parameters with 100 samples, 500 replications for about 85% response rates.

Model	Para.	Indicator	SM	BI	SM-FISE	BI-FISE
Model 1. $\phi_1 = (0, 0.6)$	$\hat{\alpha}_n$	Bias(MSE)	-0.0760(0.0156)	-0.0399(0.0166)	-0.0345(0.0120)	-0.0247(0.0136)
		SD(SE)	0.0996(0.1061)	0.1211(0.0992)	0.1048(0.0990)	0.1142(0.0983)
		CI(CR)	[0.0161,0.4320](0.8900)	[0.0656,0.4545](0.8620)	[0.0715,0.4595](0.9240)	[0.0825,0.4680](0.8940)
	$\hat{\lambda}_n$	Bias(MSE)	0.2493(0.1120)	0.1735(0.0817)	0.0153(0.0420)	0.0236(0.0472)
		SD(SE)	0.2269(0.1926)	0.2315(0.1809)	0.2034(0.1685)	0.2422(0.1683)
		CI(CR)	[0.8719,1.6267](0.7520)	[0.8190,1.5280](0.8000)	[0.6850,1.3455](0.9080)	[0.6937,1.3536](0.8980)
Model 2. $\phi_2 = (-0.5, 0.5)$	$\hat{\alpha}_n$	Bias(MSE)	-0.0722(0.0148)	-0.0466(0.0168)	0.0148(0.0132)	-0.0146(0.0142)
		SD(SE)	0.0964(0.1074)	0.1193(0.1009)	0.1095(0.0957)	0.1148(0.0979)
		CI(CR)	[0.0173,0.4382](0.9120)	[0.0577,0.4531](0.8620)	[0.0977,0.4727](0.8820)	[0.0936,0.4772](0.8700)
	$\hat{\lambda}_n$	Bias(MSE)	0.4206(0.2845)	0.3161(0.2745)	0.0144(0.1453)	0.0300(0.1634)
		SD(SE)	0.3319(0.3594)	0.4213(0.3343)	0.3693(0.3041)	0.4178(0.3084)
		CI(CR)	[1.7162,3.1250](0.8280)	[1.6608,2.9713](0.8020)	[1.4183,2.6105](0.8660)	[1.4256,2.6344](0.8580)
Model 3. $\phi_3 = (0.2, 0.2)$	$\hat{\alpha}_n$	Bias(MSE)	-0.1202(0.0230)	-0.0323(0.0108)	-0.0585(0.0099)	-0.0173(0.0060)
		SD(SE)	0.0944(0.0900)	0.0971(0.0842)	0.0777(0.0771)	0.0733(0.0705)
		CI(CR)	[0.3035,0.6561](0.7440)	[0.4027,0.7328](0.8800)	[0.3905,0.6929](0.9080)	[0.4446,0.7209](0.9280)
	$\hat{\lambda}_n$	Bias(MSE)	0.3620(0.2091)	0.1055(0.0888)	0.0585(0.0428)	0.0052(0.0386)
		SD(SE)	0.2642(0.2477)	0.2672(0.2251)	0.1915(0.1999)	0.1965(0.1843)
		CI(CR)	[0.8766,1.8474](0.7080)	[0.6643,1.5468](0.8820)	[0.6666,1.4504](0.9540)	[0.6440,1.3664](0.9220)
Model 4. $\phi_4 = (-0.27, 0.2)$	$\hat{\alpha}_n$	Bias(MSE)	-0.1300(0.0250)	-0.0467(0.0133)	-0.0642(0.0099)	-0.0236(0.0076)
		SD(SE)	0.0919(0.0910)	0.1043(0.0843)	0.0747(0.0903)	0.0835(0.0788)
		CI(CR)	[0.2917,0.6484](0.7340)	[0.3880,0.7186](0.8560)	[0.3588,0.7128](0.9380)	[0.4221,0.7308](0.8940)
	$\hat{\lambda}_n$	Bias(MSE)	0.7684(0.8402)	0.2865(0.3738)	0.1601(0.1696)	0.0397(0.1525)
		SD(SE)	0.5182(0.4863)	0.5290(0.4383)	0.3711(0.4599)	0.3783(0.4072)
		CI(CR)	[1.8151,3.7216](0.6740)	[1.4272,3.1457](0.8780)	[1.2587,3.0615](0.9680)	[1.2416,2.8378](0.9080)

Table 2. Simulation results for the INAR(1)'s parameters with 300 samples, 500 replications for about 85% response rates.

Model	Para.	Indicator	SM	BI	SM-FISE	BI-FISE
Model 1. $\phi_1 = (0, 0.6)$	$\hat{\alpha}_n$	Bias(MSE)	-0.0679(0.0081)	-0.0211(0.0053)	-0.0257(0.0046)	-0.0092(0.0043)
		SD(SE)	0.0584(0.0638)	0.0701(0.0589)	0.0628(0.0558)	0.0659(0.0553)
		CI(CR)	[0.1071,0.3570](0.8140)	[0.1650,0.3943](0.8820)	[0.1650,0.3836](0.8900)	[0.1825,0.3991](0.8960)
	$\hat{\lambda}_n$	Bias(MSE)	0.2521(0.0826)	0.1319(0.0336)	-0.0041(0.0124)	-0.0127(0.0116)
		SD(SE)	0.1407(0.1147)	0.1261(0.1039)	0.1161(0.0933)	0.1073(0.0929)
		CI(CR)	[1.0287,1.4756](0.3680)	[0.9282,1.3356](0.7180)	[0.8180,1.1788](0.8920)	[0.8051,1.1695](0.8960)
Model 2. $\phi_2 = (-0.5, 0.5)$	$\hat{\alpha}_n$	Bias(MSE)	-0.0582(0.0067)	-0.0323(0.0065)	-0.0065(0.0043)	-0.0092(0.0053)
		SD(SE)	0.0559(0.0639)	0.0714(0.0591)	0.0628(0.0535)	0.0667(0.0539)
		CI(CR)	[0.1165,0.3670](0.8820)	[0.1519,0.3834](0.8380)	[0.1887,0.3984](0.8860)	[0.1851,0.3964](0.8500)
	$\hat{\lambda}_n$	Bias(MSE)	0.3530(0.1573)	0.2758(0.1424)	-0.0218(0.0377)	0.0122(0.0547)
		SD(SE)	0.1765(0.2103)	0.2356(0.1950)	0.1928(0.1675)	0.2225(0.1704)
		CI(CR)	[1.9409,2.7651](0.6020)	[1.8935,2.6581](0.6600)	[1.6499,2.3065](0.8860)	[1.6783,2.3461](0.8540)
Model 3. $\phi_3 = (0.2, 0.2)$	$\hat{\alpha}_n$	Bias(MSE)	-0.1016(0.0133)	-0.0160(0.0037)	-0.0523(0.0045)	-0.0101(0.0020)
		SD(SE)	0.0570(0.0530)	0.0584(0.0495)	0.0454(0.0433)	0.0408(0.0386)
		CI(CR)	[0.3944,0.6024](0.5440)	[0.4870,0.6810](0.8900)	[0.4629,0.6325](0.8380)	[0.5142,0.6657](0.9180)
	$\hat{\lambda}_n$	Bias(MSE)	0.3287(0.1380)	0.0651(0.0302)	0.0452(0.0138)	-0.0086(0.0117)
		SD(SE)	0.1830(0.1461)	0.1592(0.1312)	0.1132(0.1101)	0.1047(0.1011)
		CI(CR)	[1.0432,1.6151](0.3820)	[0.8074,1.3228](0.8720)	[0.8293,1.2611](0.9580)	[0.7931,1.1896](0.9040)
Model 4. $\phi_4 = (-0.27, 0.2)$	$\hat{\alpha}_n$	Bias(MSE)	-0.0998(0.0130)	-0.0233(0.0040)	-0.0513(0.0046)	-0.0104(0.0021)
		SD(SE)	0.0493(0.0528)	0.0598(0.0489)	0.0418(0.0435)	0.0456(0.0378)
		CI(CR)	[0.3967,0.6036](0.5420)	[0.4818,0.6736](0.8960)	[0.4634,0.6340](0.8260)	[0.5154,0.6638](0.9140)
	$\hat{\lambda}_n$	Bias(MSE)	0.6142(0.4613)	0.1679(0.1255)	0.1033(0.0512)	-0.0065(0.0477)
		SD(SE)	0.2698(0.2818)	0.3067(0.2537)	0.1936(0.2152)	0.2131(0.1929)
		CI(CR)	[2.0619,3.1665](0.4040)	[1.6667,2.6690](0.8740)	[1.6815,2.5251](0.9620)	[1.6155,2.3715](0.8940)

Table 3. Simulation results for the INAR(1)'s parameters with 100 samples, 500 replications for about 70% response rates.

Model	Para.	Indicator	SM	BI	SM-FISE	BI-FISE
Model 1.	$\hat{\alpha}_n$	Bias (MSE)	-0.1134(0.0237)	-0.0306(0.0237)	-0.0422(0.0167)	-0.0240(0.0220)
		SD(SE)	0.0958(0.1140)	0.1530(0.0988)	0.1225(0.1058)	0.1486(0.1113)
		CI(CR)	[-0.0369,0.4101](0.8380)	[0.0757,0.4631](0.7700)	[0.0540,0.4651](0.8680)	[0.0579,0.4941](0.8140)
$\phi_1 = (-0.1, 0.2)$	$\hat{\lambda}_n$	Bias (MSE)	0.2719(0.1379)	0.1386(0.0887)	0.0360(0.0708)	0.0648(0.0945)
		SD(SE)	0.2440(0.2008)	0.2687(0.1770)	0.2365(0.1911)	0.3008(0.2028)
		CI(CR)	[0.8783,1.6655](0.6880)	[0.7917,1.4854](0.7800)	[0.6615,1.4105](0.8420)	[0.6673,1.4623](0.8060)
Model 2.	$\hat{\alpha}_n$	Bias (MSE)	-0.0978(0.0199)	-0.0308(0.0255)	-0.0057(0.0213)	-0.0284(0.0232)
		SD(SE)	0.0975(0.1126)	0.1593(0.0968)	0.1399(0.1047)	0.1490(0.1065)
		CI(CR)	[-0.0185,0.4229](0.7860)	[0.0794,0.4590](0.7460)	[0.0891,0.4996](0.8140)	[0.0630,0.4803](0.7760)
$\phi_2 = (-0.4, 0.2)$	$\hat{\lambda}_n$	Bias (MSE)	0.4761(0.3564)	0.2798(0.3681)	0.0669(0.3023)	0.1337(0.3122)
		SD(SE)	0.3535(0.3731)	0.5557(0.3233)	0.5056(0.3438)	0.5339(0.3596)
		CI(CR)	[1.7448,3.2074](0.8060)	[1.6461,2.9134](0.6880)	[1.3930,2.7408](0.7480)	[1.4288,2.8386](0.8140)
Model 3.	$\hat{\alpha}_n$	Bias (MSE)	-0.1930(0.0488)	-0.0506(0.0188)	-0.0919(0.0185)	-0.0328(0.0123)
		SD(SE)	0.1113(0.1004)	0.1234(0.0863)	0.0954(0.0967)	0.1055(0.0845)
		CI(CR)	[0.2103,0.6038](0.5440)	[0.3817,0.7201](0.8300)	[0.3186,0.6975](0.8800)	[0.4015,0.7329](0.8980)
$\phi_3 = (-0.3, 0.2)$	$\hat{\lambda}_n$	Bias (MSE)	0.6229(0.4991)	0.1764(0.1552)	0.1086(0.0740)	0.0328(0.0762)
		SD(SE)	0.3264(0.2873)	0.3433(0.2334)	0.2453(0.2526)	0.2655(0.2268)
		CI(CR)	[1.0598,2.1860](0.4220)	[0.7188,1.6339](0.8220)	[0.6134,1.6037](0.9440)	[0.5937,1.4827](0.8760)
Model 4.	$\hat{\alpha}_n$	Bias (MSE)	-0.1989(0.0508)	-0.0604(0.0196)	-0.0770(0.0153)	-0.0436(0.0116)
		SD(SE)	0.1078(0.0992)	0.1233(0.0827)	0.0926(0.1130)	0.0963(0.1034)
		CI(CR)	[0.2067,0.5955](0.5000)	[0.3775,0.7071](0.8000)	[0.3015,0.7440](0.9280)	[0.3538,0.7590](0.9260)
$\phi_4 = (-0.3, 0.1)$	$\hat{\lambda}_n$	Bias (MSE)	1.1266(1.1614)	0.3608(0.5750)	0.2607(0.2664)	0.1637(0.2670)
		SD(SE)	0.5907(0.5311)	0.6689(0.4267)	0.4153(0.5848)	0.4748(0.5466)
		CI(CR)	[2.0857,4.1675](0.4480)	[1.5245,3.1971](0.7920)	[1.1145,3.4070](0.9500)	[1.0923,3.2351](0.9260)

Table 4. Simulation results for the INAR(1)'s parameters with 300 samples, 500 replications for about 70% response rates.

Model	Para.	Indicator	SM	BI	SM-FISE	BI-FISE
Model 1.	$\hat{\alpha}_n$	Bias(MSE)	-0.0964(0.0132)	-0.0154(0.0086)	-0.0289(0.0064)	-0.0159(0.0074)
		SD(SE)	0.0653(0.0670)	0.0894(0.0575)	0.0740(0.0622)	0.0837(0.0625)
		CI(CR)	[0.0723,0.3349](0.7000)	[0.1718,0.3973](0.7840)	[0.1491,0.3933](0.8660)	[0.1616,0.4065](0.8440)
$\hat{\lambda}_n$	Bias(MSE)	0.2936(0.1238)	0.1102(0.0360)	0.0277(0.0244)	0.0290(0.0227)	
	SD(SE)	0.1860(0.1201)	0.1494(0.1018)	0.1603(0.1098)	0.1557(0.1097)	
	CI(CR)	[1.0582,1.5289](0.3220)	[0.9106,1.3098](0.7040)	[0.8124,1.2429](0.8680)	[0.8142,1.2439](0.8140)	
Model 2.	$\hat{\alpha}_n$	Bias(MSE)	-0.0847(0.0112)	-0.0079(0.0076)	0.0070(0.0076)	-0.0039(0.0074)
		SD(SE)	0.0619(0.0673)	0.0820(0.0572)	0.0861(0.0617)	0.0841(0.0607)
		CI(CR)	[0.0835,0.3471](0.7660)	[0.1801,0.4042](0.7880)	[0.1861,0.4279](0.8420)	[0.1772,0.4150](0.8140)
$\hat{\lambda}_n$	Bias(MSE)	0.4308(0.2274)	0.1930(0.1188)	0.0535(0.1121)	0.0417(0.0839)	
	SD(SE)	0.2017(0.2205)	0.2606(0.1885)	0.3278(0.2000)	0.2685(0.1971)	
	CI(CR)	[1.9987,2.8630](0.5200)	[1.8236,2.5624](0.7000)	[1.6615,2.4455](0.7960)	[1.6553,2.4281](0.8140)	
Model 3.	$\hat{\alpha}_n$	Bias(MSE)	-0.1804(0.0367)	-0.0260(0.0056)	-0.0861(0.0105)	-0.0169(0.0032)
		SD(SE)	0.0648(0.0593)	0.0667(0.0510)	0.0536(0.0529)	0.0517(0.0440)
		CI(CR)	[0.3033,0.5359](0.1560)	[0.4740,0.6740](0.8400)	[0.4102,0.6176](0.6780)	[0.4968,0.6694](0.8980)
$\hat{\lambda}_n$	Bias(MSE)	0.6041(0.4121)	0.1209(0.0534)	0.0853(0.0257)	0.0094(0.0200)	
	SD(SE)	0.2189(0.1587)	0.1876(0.1375)	0.1362(0.1328)	0.1366(0.1155)	
	CI(CR)	[1.2734,1.9348](0.0940)	[0.8513,1.3905](0.7680)	[0.8249,1.3457](0.9340)	[0.7831,1.2357](0.8820)	
Model 4.	$\hat{\alpha}_n$	Bias(MSE)	-0.1710(0.0329)	-0.0228(0.0058)	-0.0699(0.0078)	-0.0161(0.0028)
		SD(SE)	0.0595(0.0579)	0.0646(0.0496)	0.0525(0.0521)	0.0476(0.0433)
		CI(CR)	[0.3156,0.5424](0.1660)	[0.4799,0.6744](0.8360)	[0.4280,0.6322](0.7860)	[0.4989,0.6880](0.9140)
$\hat{\lambda}_n$	Bias(MSE)	0.9697(1.0593)	0.1718(0.1613)	0.2193(0.1139)	0.0465(0.0673)	
	SD(SE)	0.3447(0.3069)	0.3326(0.2585)	0.2598(0.2615)	0.2325(0.2250)	
	CI(CR)	[2.3631,3.5713](0.1220)	[1.6651,2.6786](0.8080)	[1.7066,2.7319](0.9160)	[1.6054,2.4876](0.9240)	

Table 5. Estimated indicators differences between each estimator and θ for real data with response rate=85%.

Method	Estimator	$\phi = 0.85$					$\phi = (-0.3, 0.2)$					$\phi = (0.4, -0.5)$				
		Bias	MSE	SD	SE	SE	Bias	MSE	SD	SE	SE	Bias	MSE	SD	SE	SE
SM	$\hat{\alpha}_n$	-0.0314	0.0022	0.03339	0.0993	0.0993	-0.0163	0.0008	0.0224	0.0586	0.0586	-0.0137	0.0010	0.0285	0.0585	0.0585
	$\hat{\lambda}_n$	0.1766	0.0700	0.1834	0.5880	0.5880	-0.4729	0.2649	0.2035	0.4318	0.4318	0.4696	0.2664	0.2285	0.4318	0.4318
BI	$\hat{\alpha}_n$	0.0203	0.0022	0.0394	0.0910	0.0910	0.0222	0.0018	0.0355	0.0569	0.0569	0.0174	0.0012	0.0233	0.0571	0.0571
	$\hat{\lambda}_n$	-0.1217	0.0747	0.2344	0.5382	0.5382	0.2039	0.0938	0.2032	0.4137	0.4137	0.2350	0.0996	0.1574	0.4159	0.4159
SM-FISE	$\hat{\alpha}_n$	-0.0097	0.0012	0.0315	0.0600	0.0600	-0.0231	0.0012	0.0259	0.0581	0.0581	-0.0166	0.0009	0.0237	0.0581	0.0581
	$\hat{\lambda}_n$	0.0665	0.0420	0.1928	0.4287	0.4287	0.0771	0.0349	0.1533	0.4097	0.4097	0.0918	0.0487	0.2282	0.4155	0.4155
BI-FISE	$\hat{\alpha}_n$	0.0096	0.0015	0.0394	0.0595	0.0595	-0.0096	0.0007	0.0237	0.0583	0.0583	-0.0035	0.0008	0.0236	0.0582	0.0582
	$\hat{\lambda}_n$	-0.0650	0.0537	0.2234	0.4128	0.4128	-0.0120	0.0380	0.2047	0.4058	0.4058	0.1009	0.0526	0.1937	0.4170	0.4170

Table 6. Estimated indicators differences between each estimator and θ for real data with response rate=70%.

Method	Estimator	$\phi = 0.7$					$\phi = (-0.88, 0.2)$					$\phi = (-0.3, -0.1)$				
		Bias	MSE	SD	SE	SE	Bias	MSE	SD	SE	SE	Bias	MSE	SD	SE	SE
SM	$\hat{\alpha}_n$	-0.0587	0.0061	0.0463	0.0573	0.0573	-0.0278	0.0017	0.0311	0.0603	0.0603	-0.0317	0.2033	0.0425	0.0581	0.0581
	$\hat{\lambda}_n$	0.3434	0.1776	0.2304	0.4117	0.4117	0.9392	0.9414	0.2398	0.4686	0.4686	0.4258	0.2803	0.2699	0.4243	0.4243
BI	$\hat{\alpha}_n$	0.0540	0.0058	0.0552	0.0539	0.0539	0.0476	0.0046	0.0540	0.0564	0.0564	0.0727	0.0086	0.0470	0.0540	0.0540
	$\hat{\lambda}_n$	-0.3448	0.2466	0.3595	0.3705	0.3705	0.2999	0.2074	0.3976	0.4234	0.4234	-0.1836	0.1707	0.2939	0.3851	0.3851
SM-FISE	$\hat{\alpha}_n$	-0.0455	0.0049	0.0462	0.0637	0.0637	-0.0458	0.0032	0.0341	0.0604	0.0604	-0.0315	0.0039	0.0433	0.0630	0.0630
	$\hat{\lambda}_n$	0.2583	0.1713	0.2640	0.4712	0.4712	0.1150	0.1292	0.3198	0.4299	0.4299	0.1265	0.2744	0.4259	0.4612	0.4612
BI-FISE	$\hat{\alpha}_n$	0.0246	0.0030	0.0471	0.0634	0.0634	-0.0033	0.0017	0.0456	0.0605	0.0605	0.0279	0.0047	0.0489	0.0616	0.0616
	$\hat{\lambda}_n$	-0.1675	0.1410	0.3337	0.4553	0.4553	-0.0464	0.1522	0.3842	0.4262	0.4262	-0.0748	0.2295	0.4338	0.4664	0.4664

Table 7. Simulation results for parameters of INARG(1) with 85% response rates, 500 replications.

Size	Model	Para.	Indicator	SM	BI	SM-FISE	BI-FISE
n=100	(0.3,0.5)	$\hat{\alpha}_n$	Bias (MSE)	-0.0766(0.0164)	-0.0462(0.0168)	-0.0410(0.0080)	-0.0250(0.0082)
			SD (SE)	0.1025(0.0736)	0.1224(0.0729)	0.0789(0.0716)	0.0839(0.0709)
			CI (CR)	[0.0791,0.3676](0.6940)	[0.1110,0.3966](0.7280)	[0.1187,0.3993](0.8660)	[0.1361,0.4139](0.8620)
	$\hat{\lambda}_n$	Bias (MSE)	-0.0604(0.0060)	-0.0424(0.0046)	0.0041(0.0021)	0.0065(0.0021)	
		SD (SE)	0.0485(0.0133)	0.0529(0.0158)	0.0453(0.0421)	0.0441(0.0419)	
		CI (CR)	[0.4134,0.4657](0.2020)	[0.4266,0.4885](0.3580)	[0.4215,0.5866](0.9280)	[0.4245,0.5886](0.9240)	
	(0.5,0.5)	$\hat{\alpha}_n$	Bias (MSE)	-0.0995(0.0201)	-0.0458(0.0153)	-0.0598(0.0086)	-0.0225(0.0061)
			SD (SE)	0.1012(0.0624)	0.1143(0.0624)	0.0717(0.0652)	0.0770(0.0615)
			CI (CR)	[0.2783,0.5228](0.5840)	[0.3318,0.5766](0.7260)	[0.3125,0.5679](0.8560)	[0.3569,0.5981](0.8840)
	$\hat{\lambda}_n$	Bias (MSE)	-0.0654(0.0067)	-0.0313(0.0046)	0.0004(0.0020)	0.0090(0.0023)	
		SD (SE)	0.0499(0.0117)	0.0593(0.0149)	0.0448(0.0442)	0.0463(0.0435)	
		CI (CR)	[0.4116,0.4576](0.2060)	[0.4395,0.4979](0.3420)	[0.4137,0.5871](0.9360)	[0.4237,0.5943](0.9220)	
n=300	(0.3,0.5)	$\hat{\alpha}_n$	Bias (MSE)	-0.0675(0.0086)	-0.0262(0.0063)	-0.0369(0.0033)	-0.0143(0.0026)
			SD (SE)	0.0643(0.0433)	0.0732(0.0433)	0.0439(0.0420)	0.0474(0.0417)
			CI (CR)	[0.1476,0.3174](0.5760)	[0.1890,0.3586](0.7080)	[0.1807,0.3454](0.8580)	[0.2039,0.3674](0.8840)
	$\hat{\lambda}_n$	Bias (MSE)	0.0624(0.0048)	-0.0415(0.0027)	0.0036(0.0007)	0.0044(0.0007)	
		SD (SE)	0.0308(0.0076)	0.0307(0.0092)	0.0246(0.0247)	0.0252(0.0245)	
		CI (CR)	[0.4227,0.4526](0.0960)	[0.4405,0.4766](0.1980)	[0.4552,0.5519](0.9320)	[0.4563,0.5524](0.9360)	
	(0.5,0.5)	$\hat{\alpha}_n$	Bias (MSE)	-0.0696(0.0083)	-0.0265(0.0051)	-0.0479(0.0037)	-0.0152(0.0015)
			SD (SE)	0.0600(0.0368)	0.0689(0.0370)	0.0348(0.0376)	0.0353(0.0358)
			CI (CR)	[0.3582,0.5026](0.5000)	[0.4010,0.5460](0.6800)	[0.3785,0.5257](0.7820)	[0.4147,0.5548](0.9220)
	$\hat{\lambda}_n$	Bias (MSE)	-0.0563(0.0042)	-0.0299(0.0024)	0.0007(0.0007)	0.0085(0.0008)	
		SD (SE)	0.0318(0.0071)	0.0392(0.0086)	0.0259(0.0258)	0.0264(0.0254)	
		CI (CR)	[0.4297,0.4576](0.0960)	[0.4531,0.4870](0.2780)	[0.4502,0.5512](0.9480)	[0.4586,0.5583](0.9380)	

Table 8. Simulation results for parameters of INARG(1) with 70% response rates, 500 replications.

Size	Model	Para.	Indicator	SM	BI	SM-FISE	BI-FISE
n=100	(0.3,0.5)	$\hat{\alpha}_n$	Bias (MSE)	-0.0985(0.0216)	-0.0392(0.0287)	-0.0336(0.0089)	-0.0211(0.0117)
			SD (SE)	0.1102(0.0770)	0.1644(0.0736)	0.0867(0.0834)	0.1040(0.0790)
			CI (CR)	[0.0492,0.3529](0.6660)	[0.1165,0.4051](0.5780)	[0.1029,0.4299](0.8880)	[0.1240,0.4338](0.8180)
	$\hat{\lambda}_n$	$\phi_1 = 0.14$	Bias (MSE)	-0.0517(0.0063)	-0.0275(0.0061)	0.0005(0.0025)	0.0030(0.0030)
			SD (SE)	0.0594(0.0135)	0.0719(0.0179)	0.0502(0.0465)	0.0548(0.0456)
			CI (CR)	[0.4220,0.4747](0.2720)	[0.4375,0.5075](0.3400)	[0.4093,0.5918](0.9100)	[0.4135,0.5924](0.8700)
	$\hat{\alpha}_n$	(0.5,0.5)	Bias (MSE)	-0.1528(0.0357)	-0.0548(0.0253)	-0.0590(0.0085)	-0.0262(0.0073)
			SD (SE)	0.1120(0.0667)	0.1542(0.0637)	0.00722(0.0719)	0.0817(0.0681)
			CI (CR)	[0.2165,0.4779](0.4100)	[0.3204,0.5701](0.6320)	[0.3000,0.5820](0.8920)	[0.3404,0.6072](0.8720)
	$\hat{\lambda}_n$	$\phi_3 = 0.1$	Bias (MSE)	-0.0750(0.0084)	-0.250(0.0065)	-0.0085(0.0025)	0.0028(0.0029)
			SD (SE)	0.0531(0.0104)	0.0787(0.0159)	0.0489(0.0471)	0.0547(0.0469)
			CI (CR)	[0.4046,0.4453](0.1300)	[0.4438,0.5062](0.2820)	[0.3992,0.5838](0.9200)	[0.4110,0.5947](0.9200)
n=300	(0.3,0.5)	$\hat{\alpha}_n$	Bias (MSE)	-0.0851(0.0115)	-0.0105(0.0081)	-0.0340(0.0035)	-0.0084(0.0033)
			SD (SE)	0.0638(0.0450)	0.0861(0.0437)	0.0470(0.0472)	0.0545(0.0465)
			CI (CR)	[0.1267,0.3032](0.4860)	[0.2039,0.3751](0.6460)	[0.1735,0.3586](0.8960)	[0.2004,0.3828](0.8840)
	$\hat{\lambda}_n$	$\phi_1 = 0.14$	Bias (MSE)	-0.0601(0.0052)	-0.0260(0.0023)	-0.0031(0.0008)	0.0019(0.0010)
			SD (SE)	0.0402(0.0075)	0.0389(0.0104)	0.0277(0.0268)	0.0297(0.0267)
			CI (CR)	[0.4251,0.4547](0.1520)	[0.4537,0.4944](0.3000)	[0.4445,0.5494](0.9220)	[0.4495,0.5542](0.9040)
	$\hat{\alpha}_n$	(0.5,0.5)	Bias (MSE)	-0.1334(0.0219)	-0.0234(0.0074)	-0.0554(0.0049)	-0.0200(0.0029)
			SD (SE)	0.650(0.0392)	0.0839(0.0371)	0.0422(0.0422)	0.0252(0.0274)
			CI (CR)	[0.2898,0.4434](0.1860)	[0.4039,0.5494](0.6120)	[0.3620,0.5272](0.7520)	[0.4016,0.5586](0.8480)
	$\hat{\lambda}_n$	$\phi_3 = 0.1$	Bias (MSE)	-0.0719(0.0060)	-0.0213(0.0025)	-0.0135(0.0009)	0.0018(0.0009)
			SD (SE)	0.0279(0.0062)	0.0432(0.0093)	0.0252(0.0274)	0.0288(0.0277)
			CI (CR)	[0.4160,0.4403](0.0302)	[0.4603,0.4970](0.3120)	[0.4328,0.5402](0.9240)	[0.4474,0.5561](0.9200)

References

- [1] M.A. Al-Osh and A.A. Alzaid, *First-order integer-valued autoregressive (INAR(1)) process*, J. Time Series Anal. **8** (3), 261–275, 1987.
- [2] J. Andersson and D. Karlis, *Treating missing values in INAR(1) models: an application to syndromic surveillance data*, J. Time Series Anal. **31** (1), 12–19, 2010.
- [3] I.V. Basawa, P.D. Feigin and C.C. Heyde, *Asymptotic properties of maximum likelihood estimators for stochastic processes*, Sankhya A **38** (3), 259–270, 1976.
- [4] X. Chen, A.T.K. Wan and Y. Zhou, *Efficient quantile regression analysis with missing observations*, J. Amer. Statist. Assoc. **110**, 723–741, 2015.
- [5] X. Cui, J. Guo and G. Yang, *On the identifiability and estimation of generalized linear models with parametric nonignorable missing data mechanism*, Comput. Statist. Data Anal. **107**, 64–80, 2017.
- [6] J. Du and Y. Li, *The integer-valued autoregressive (INAR(p)) model*, J. Time Series Anal. **12** (2), 129–142, 1991.
- [7] R.K. Freeland and B.P.M. McCabe, *Analysis of low count time series data by poisson autoregression*, J. Time Series Anal. **25** (5), 701–722, 2004.
- [8] P. Hall and C.C. Heyde, *Martingale Limit Theory and Its Application*, Academic Press, New York, 1980.
- [9] M.A. Jazi, G. Jones, and C.D. Lai, *Integer valued AR(1) with geometric innovations*, J. Iran. Stat. Soc. (JIRSS) **2** (2), 173–190, 2012.
- [10] B. Jia, D. Wang and H. Zhang, *A study for missing values in PINAR(1)_T processes*, Comm. Statist. Theory Methods **43** (22), 4780–4789, 2014.
- [11] R. Jung, G. Ronning and A. Tremayne, *Estimation in conditional first order autoregression with discrete support*, Statist. Papers **46** (2), 195–224, 2005.
- [12] S.A. Khashimov, *The central limit theorem for generalized U-statistics for weakly dependent vectors*, Theory Probab. Appl. **38** (3), 563–578, 1993.
- [13] J.K. Kim and C.L. Yu, *A semiparametric estimation of mean functionals with nonignorable missing data*, J. Amer. Statist. Assoc. **106** (493), 157–165, 2012.
- [14] J.K. Kim and J. Shao, *Statistical Methods for Handling Incomplete Data*, CRC Press, Boca Raton, 2013.
- [15] R.J.A. Little and D.B. Rubin, *Statistical Analysis with Missing Data, Second Edition*, John Wiley & Sons, 2002.
- [16] T.A. Louis, *Finding the observed information matrix when using the EM algorithm*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **44**, 226–233, 1982.
- [17] K. Morikawa, J.K. Kim and Y. Kano, *Semiparametric maximum likelihood estimation with data missing not at random*, Canad. J. Statist. **45** (4), 393–409, 2017.
- [18] W.K. Newey and D.L. McFadden, *Large sample estimation and hypothesis testing*, in: Handbook of Econometrics, Vol. IV, Engle R.F., McFadden D.L. editors, North Holland, Amsterdam, 1994.
- [19] M. Pourahmadi, *Estimation and interpolation of missing values of a stationary time series*, J. Time Series Anal. **10** (2), 149–169, 1989.
- [20] M.K. Riddles, J.K. Kim and J. Im, *Propensity-score-adjustment method for nonignorable nonresponse*, Journal of Survey Statistics and Methodology **4**, 215–245, 2016.
- [21] D.B. Rubin, *Inference and missing data*, Biometrika **63** (3), 581–592, 1976.
- [22] J. Shao and L. Wang, *Semiparametric inverse propensity weighting for nonignorable missing data*, Biometrika **103**, 175–187, 2016.
- [23] N. Tang, P. Zhao and H. Zhu, *Empirical likelihood for estimating equations with nonignorably missing data*, Statist. Sinica **24**, 723–747, 2014.
- [24] S. Wang, J. Shao and J.K. Kim, *An instrumental variable approach for identification and estimation with nonignorable nonresponse*, Statist. Sinica **24**, 1097–1116, 2014.

- [25] K. Yang, D. Wang, B. Jia and H. Li, *An integer-valued threshold autoregressive process based on negative binomial thinning*, *Statist. Papers* **59** (3), 1131-1160, 2018.
- [26] H. Zhang, D. Wang and F. Zhu, *Inference for INAR(p) processes with signed generalized power series thinning operator*, *J. Statist. Plann. Inference* **140** (3), 667-683, 2010.
- [27] H. Zheng, I.V. Basawa and S. Datta, *Inference for p th-order random coefficient integer-valued autoregressive processes*, *J. Time Series Anal.* **27** (3), 411-440, 2006.