Brazilian Journal of Probability and Statistics
2018, Vol. 32, No. 1, 147–171
https://doi.org/10.1214/16-BJPS335

Noise-indicator nonnegative integer-valued autoregressive time series of the first order

Vladica Stojanović[a], Dragan Randjelović[b] and Kristijan Kuk[b]

[a]*Faculty of Sciences and Mathematics*
[b]*Academy of Criminalistics and Police Studies*

Abstract. This paper presents a modification and, at the same time, a generalization of the linear first order nonnegative integer-valued autoregressive processes, well-known as INAR(1) processes. By using the so-called Noise-Indicator, a nonlinear model with the threshold regime and with more complex structure than the appropriate linear models was obtained. The new model, named NIINAR(1) process, has been investigated in terms of the most general, the power series distribution of its innovations. Basic stochastic properties of the NIINAR(1) model (e.g., correlation structure, over-dispersion conditions and distributional properties) are given. Also, besides of some standard parameters estimators, a novel estimation techniques, together with the asymptotic properties of the obtained estimates is described. At last, a Monte Carlo study of this process is also given, as well as its application in the analysis of dynamics of two empirical dataset.

1 Introduction & motivation

The discrete-valued time series attracted a lot of attention over the last few decades. Starting from the pioneer work of Al-Osh and Alzaid (1987), where the so-called first order INteger-AutoRegressive (INAR(1)) process was introduced, many results related to this (and some similar) models are obtained nowadays. We point out only the most recently works of Bakouch and Ristić (2010), Jazi, Jones and Lai (2012) and Schweer and Weiß (2014) where INAR(1) models with various innovations distributions were investigated, as well as Ristić, Nastić and Bakouch (2012), Ristić, Nastić and Djordjević (2016) where INAR-type models with a some specified marginal distributions were introduced. On the other hand, Weiß (2008, 2009), Weiß and Pollet (2014) and Ristić, Nastić and Miletić (2013) studied the general properties of the INAR(1) models, with emphasis on the binomial thinning operator, while Ristić, Bakouch and Nastić (2009) and Nastić, Ristić and Bakouch (2012) considered INAR models with negative binomial thinning. At last, some extensions, in the terms of increasing the models' order, as well as estimation of their parameters, can be found in Drost, Akken and Werker (2008), Silva and Silva (2009), Nastić (2014) or Martin, Tremayne and Jung (2014). In this paper, starting

Key words and phrases. Noise-indicator, power series distribution, NIINAR(1) process, parameters estimation.
Received January 2016; accepted September 2016.

point is the work of Bourguignon and Vasconcellos (2015), where authors define the PSINAR(1) model, with a very general form of the so-called power series (PS) innovations. Usage of these kind of innovations has a many advantages. First of all, they have the following two basic properties:

1. The PS-innovations contain, as their special cases, the most of the other well-known non-negative integer valued distributions. For example, the independent identically distributed (i.i.d.) time series with binomial, geometric, Poisson, negative binomial or logarithmic distribution can be interpreted as the PS-innovations (see Table 1 in Bourguignon and Vasconcellos (2015)).
2. They allow modeling an empirical nonnegative discrete time series where has been observed an over-(equi- or under-)dispersion, truncated or zero-inflated distributions, etc.

The main aim of this paper is to propose a new INAR(1)-type model, based on the PS-innovations, but which will be more flexible and more applicable to explain the dynamics of some real-life data series (see Section 5, bellow). For instance, it can be situations when the innovations of some time series can be modelled with an equal-dispersed distribution (such as Poisson's), but observations have a pronounced over-dispersion, or highly frequented values in zero, the property known as the zero-inflated distribution. The main goal of our model is to explain these behaviours, that is, to generate marginal distribution which, in the sense of the over-dispersion or the zero-inflated conditions, can be (potentially) different from the appropriate innovations. For this purpose, we set so-called *the Noise-Indicator*, similarly as it was done in modification of the Stochastic Permanent Breaking (STOPBREAK) processes, described in Stojanović, Popović and Popović (2011, 2014, 2015), Stojanović, Milovanović and Jelić (2016), as well as the Stochastic Volatility (SV) models, described in Stojanović, Popović and Milovanović (2016). The basic idea is to allow that innovations sequence has a property of optionality and, according to this, the observations will be have a thresholds, non-linear dynamics. For these reasons, a stochastic structure of our model is somewhat more complex than the standard INAR-processes. At the same time, it differs from the other related models introduced, for instance, in Pavlopulos and Dimitris (2008) or Li, Wang and Zhang (2015).

This paper is organized as follows. The definition of the Noise-Indicator INAR(1) model, in short NIINAR(1) model, as its basic stochastic properties, are described in the following, Section 2. Some parameters estimation procedures of this process, especially a novel, the probability generating function (PGF) technique, is considered in Section 3. Monte Carlo simulations of the PGF estimates are described in Section 4. Finally, Section 5 is devoted to an application of the previous estimation procedure in fitting some empirical data series.

2 Definition & properties of the model

In this section, we first define the general form of the Power Series (PS) distributions, as well as the principle of introducing the Noise-Indicators. In the second part, we describe the NIINAR(1) model itself, and its basic stochastic properties.

2.1 Power series & noise-indicator's distributions

Let (ε_t), $t \in \mathbf{Z}$ be a sequence of independent identically distributed (i.i.d.) random variables. We say that this series has a *Power Series* (*PS*) *distribution* if its probability mass function (p.m.f.) is given by the following expression:

$$p_\varepsilon(x;\theta) := P\{\varepsilon_t = x\} = \frac{a(x)\theta^x}{f(\theta)}, \qquad x \in \mathcal{S}, \tag{2.1}$$

where $\mathcal{S} \subseteq \mathbf{Z}^+ = \{0, 1, 2, \dots, \}$ is the discrete support of the r.v.s ε_t. In the same manner as in Bourguignon and Vasconcellos (2015), here we denote:

(i) $a(x) \geq 0$ is the function which depends (only) of x;
(ii) $\theta > 0$ is the (unknown) parameter;
(iii) $f(\theta) := \sum_{x=0}^{\infty} a(x)\theta^x$ is the function of θ, such that $0 < f(\theta) < +\infty$, where $\theta \in (0, R)$, $R > 0$.

Note that the expression (2.1) represents a very general form of the non-negative, integer-valued p.m.f. It enables that, for some particular choices of functions $a(x)$, $f(\theta)$ and the parameter θ, can be obtained a most of the well-known distributions (see Table 1, below). In addition, according to assumption (iii), it is obvious that power series $f(\theta)$ converges, in fact, on the interval $(-R, R)$. Nevertheless, we introduced an usual assumption that the parameter θ is positive, that is, that the power series $f(\theta)$ converges on interval $(0, R)$. According to these, the function $f(\theta)$ is positive and increasing at this interval. Moreover, for all $n \in \mathbf{N}$ it has a positive derivatives

$$f^{(n)}(\theta) := \frac{d^n f(\theta)}{d\theta^n} = \sum_{x=n}^{\infty} x(x-1)\cdots(x-n+1)a(x)\theta^{x-n}. \tag{2.2}$$

Equality (2.2) can be useful for determining the recurrence relations between the moments of series (ε_t), as it described, for instance, in Noack (1950). In order to obtain an explicit expression for the moments $\mu_n^{(\varepsilon)} := E[\varepsilon_t^n]$, we have used a different approach, based on the moment-generating function (MGF) $M(u) = \mathrm{E}[\exp(u\varepsilon_t)] = f(\theta \mathrm{e}^u)/f(\theta)$. Using the well-known properties of MGFs, after some computations, it follows

$$\mu_1^{(\varepsilon)} = \mathrm{E}[\varepsilon_t] = \frac{d}{du}M(0) = \theta\frac{f'(\theta)}{f(\theta)},$$

$$\mu_2^{(\varepsilon)} = \mathrm{E}\big[\varepsilon_t^2\big] = \frac{d^2}{du^2}M(0) = \theta\frac{f'(\theta)}{f(\theta)} + \theta^2\frac{f''(\theta)}{f(\theta)},$$

$$\vdots$$

$$\mu_n^{(\varepsilon)} = \mathrm{E}\big[\varepsilon_t^n\big] = \frac{d^n}{du^n}M(0) = \frac{1}{f(\theta)}\sum_{k=1}^{n} b_{n;k}\theta^k f^{(k)}(\theta),$$

where the coefficients $b_{n;k}$, for each $n \in \mathbf{N}$, can be calculated recursively:

$$b_{n;1} = 1, \qquad b_{n;k+1} = \frac{(k+1)^{n-1}}{k!} - \sum_{j=1}^{k}\frac{b_{n;j}}{(k-j+1)!}, \qquad k = 1, 2, \dots, n-1.$$

According the first two moments, mean and variance of the r.v.s ε_t are

$$\mu_\varepsilon := \mathrm{E}[\varepsilon_t] = \mu_1^{(\varepsilon)} = \theta g'(\theta), \tag{2.3}$$

$$\sigma_\varepsilon^2 := \mathrm{Var}[\varepsilon_t] = \mu_2^{(\varepsilon)} - \big(\mu_1^{(\varepsilon)}\big)^2 = \mu_\varepsilon + \theta^2 g''(\theta), \tag{2.4}$$

where we set $g(\theta) = \log f(\theta)$. Notice that over-dispersion of ε_t depends (only) on convexity of the function $g(\theta)$. More precisely, the series (ε_t) is over-dispersed, i.e., the condition $D_\varepsilon(\theta) := \sigma_\varepsilon^2 - \mu_\varepsilon > 0$ holds if and only if $g''(\theta) > 0$, $\forall\theta \in (0, R)$. Naturally, many of the well-known, nonnegative integer-valued distributions do not satisfy this condition (see Table 1).

In order to improve the over-dispersion conditions, as well as the properties of zero-inflated distribution, we modify the PS-distributions in the following way. We introduce the series $\eta_t = q_{t-1}\varepsilon_t$, $t \in \mathbf{Z}$, where $q_t = q_t(c)$ is so-called *the Noise-Indicator*, defined as

$$q_t(c) := I(\varepsilon_t \geq c) = \begin{cases} 1, & \varepsilon_t \geq c, \\ 0, & \varepsilon_t < c. \end{cases}$$

The indicators (q_t) impose an involvement of the value of (ε_t) when it is enough statistically significant. If the value of ε_{t-1} is relatively small, then $q_{t-1} = 0$, and, thereby, the realization of η_t at time t equals zero. On the other hand, in the case of pronounced fluctuation caused by the previous realization ε_{t-1}, the indicator value is $q_{t-1} = 1$, and the value η_t matches to the value ε_t. Level of significance in realizations of the sequence (ε_t) determines *the critical value of reaction* $c > 0$, for which we can write

$$m_c := \mathrm{E}\big[I(\varepsilon_t \geq c)\big] = P\{\varepsilon_t \geq c\} = 1 - F_\varepsilon(c),$$

where $F_\varepsilon(\cdot)$ is the cumulative distribution function (c.d.f.) of ε_t. Therefore, for a given value $c > 0$, it can be determined the constant m_c, and vice versa. On the other hand, indicators $q_t(c)$ generalize the PS-distributed series (ε_t), which could be obtained from the series (η_t), when $c \to 0$. Nevertheless, we will assume that the nontrivial condition $0 < F_\varepsilon(c) < 1$ is always fulfilled. Then, the basic properties of the series (η_t) can be formulated by the following proposition.

Table 1 *Comparison of some known distributions, through the PS and NIPS over-dispersion conditions*

Distributions	$\mathcal{S}$	$a(x)$	θ	$f(\theta)$	$D_\varepsilon(\theta)$	$D_\eta(\theta)$
1. Bernoulli	$\{0, 1\}$	1	$(0, \infty)$	$1+\theta$	$-\frac{\theta^2}{(1+\theta)^2}$	$-\frac{m_c^2\theta^2}{(1+\theta)^2}$
2. Binomial	$\{0, \ldots, n\}$	$\binom{n}{x}$	$(0, \infty)$	$(1+\theta)^n$	$-\frac{n\theta^2}{(1+\theta)^2}$	$\frac{m_c n\theta^2(nF_\varepsilon(c)-1)}{(1+\theta)^2}$
3. Poisson	$\{0, \ldots, \infty\}$	$\frac{1}{x!}$	$(0, \infty)$	$\exp(\theta)$	0	$m_c\theta^2 F_\varepsilon(c)$
4. Geometric	$\{0, \ldots, \infty\}$	1	$(0, 1)$	$\frac{1}{1-\theta}$	$\frac{\theta^2}{(1-\theta)^2}$	$\frac{\theta^2[1-(F_\varepsilon(c))^2]}{(1-\theta)^2}$
5. Neg. Binomial	$\{0, \ldots, \infty\}$	$\frac{\Gamma(x+r)}{x!\Gamma(r)}$	$(0, 1)$	$\frac{1}{(1-\theta)^r}$	$\frac{r\theta^2}{(1-\theta)^2}$	$\frac{m_c r\theta^2(rF_\varepsilon(c)+1)}{(1-\theta)^2}$
6. Pascal	$\{r, \ldots, \infty\}$	$\binom{x-1}{r-1}$	$(0, 1)$	$\frac{\theta^r}{(1-\theta)^r}$	$\frac{r(2\theta-1)}{(1-\theta)^2}$	$\frac{m_c r(F_\varepsilon(c)+2\theta-1)}{(1-\theta)^2}$
7. Logarithmic	$\{1, \ldots, \infty\}$	x^{-1}	$(0, 1)$	$\log(1-\theta)^{-1}$	$\frac{-\theta^2[1+\log(1-\theta)]}{(1-\theta)^2\log^2(1-\theta)}$	$\frac{m_c\theta^2[F_\varepsilon(c)-1-\log(1-\theta)]}{(1-\theta)^2\log^2(1-\theta)}$

Theorem 2.1. *The series* (η_t), $t \in \mathbf{Z}$ *is the sequence of the uncorrelated r.v.s with the p.m.f.*

$$p_\eta(x;\theta) = \begin{cases} 1 + m_c\big(p_\varepsilon(x;\theta) - 1\big), & x = 0, \\ m_c p_\varepsilon(x;\theta), & x = 1, 2, \ldots, \end{cases} \tag{2.5}$$

where $p_\eta(x;\theta) := P\{\eta_t = x\}$. *In addition, the moments of the r.v.s* η_t *are* $\mu_n^{(\eta)} := \mathrm{E}[\eta_t^n] = m_c \mu_n^{(\varepsilon)}$, *and these r.v.s are over-dispersed if and only if the condition*

$$g''(\theta) \geq 0 \vee F_\varepsilon(c) > -\frac{g''(\theta)}{(g'(\theta))^2} \tag{2.6}$$

is fulfilled.

Proof. See the Appendix. □

Remark 2.1. According to the previous theorem, the main goals of introducing the Noise-Indicators Power-Series (NIPS) distribution can be seen. First of all, the NIPS-distributed r.v.s (η_t) are uncorrelated, so then they can be used as an innovations of some theoretical INAR-type model. Moreover, the equation (2.5) and the condition $0 < m_c < 1$ gives, for $\varepsilon_t \neq I_0$,

$$p_\eta(0;\theta) - p_\varepsilon(0;\theta) = (1 - m_c)\big[1 - p_\varepsilon(0;\theta)\big] > 0,$$

that is, r.v.s (η_t) have a more pronounced zero-inflated distribution than (ε_t). For this reason, they are most adequate in the fitting some empirical datasets which have a pronounced zero-values (see Section 5). At last, over-dispersion condition (2.6) is more “weak” than the appropriate condition of the PS-distributed series (ε_t).

Table 1 presents the functions $a(x)$ and $f(\theta)$, as well as the differences $D_\varepsilon(\theta)$ and $D_\eta(\theta)$, for the most of well-known nonnegative integer-valued distributions. Let us remark that the under-dispersion conditions for both the series (ε_t) and (η_t) holds only for the most simplest, Bernoulli’s distribution. In the case of binomial distribution, where the PS-distributed r.v.s (ε_t) are under-dispersed, the NIPS-distributed r.v.s (η_t) can be over-dispersed for sufficiently large $n \in \mathbf{N}$, that is, if and only if $nF_\varepsilon(c) > 1$. On the other hand, the series (ε_t) with equal-dispersed Poisson distribution always generate the over-dispersed NIPS-distributed sequence (η_t). An interesting fact is that the series with negative binomial, over-dispersed distribution (as well as geometric distribution as its special case), gives the NIPS-series with somewhat smaller over-dispersion. Finally, over-dispersion properties of the Pascal and logarithmic distribution can vary for the both series, although they are more pronounced in the case of NIPS-series (η_t).

2.2 The NIINAR(1) model

According to previously assumptions, let (X_t), $t \in \mathbf{Z}$ be a series defined by the recurrence relation

$$X_t = \alpha \circ X_{t-1} + \eta_t = \begin{cases} \alpha \circ X_{t-1} + \varepsilon_t, & \varepsilon_{t-1} \geq c, \\ \alpha \circ X_{t-1}, & \varepsilon_{t-1} < c. \end{cases} \tag{2.7}$$

Here, $\alpha \in (0, 1)$ and $\alpha \circ X := \sum_{j=1}^{X} B_j(\alpha)$ is the binomial thinning operator, where X is a nonnegative integer-valued r.v., and for any $t \in \mathbf{Z}$ the r.v.s $B_j = B_j(\alpha)$ are mutually independent, also independent of X, with Bernoulli's distribution $P\{B_j = 1\} = 1 - P\{B_j = 0\} = \alpha$. Then, we say that the sequence (X_t) represents the *Noise-Indicator INAR(1) process*, or in short, the *NIINAR(1) process*.

As we can see from the equality (2.7), the series (X_t) can be interpreted as the INAR(1) model with an "optional" PS-innovations (ε_t). Namely, if the value of ε_{t-1} is less than the critical value c, it follows that $\eta_t = 0$. Then, the realization of X_t at time t depends only on X_{t-1}. On the other hand, in the case when ε_{t-1} is enough statistically significant, that is, greater than or equal c, the value X_t is realized as a standard INAR(1) model. This feature allows that our model has a specific threshold, non-linear structure, which depends, in fact, on its NIPS-innovations (η_t). Based on the above-mentioned properties of this series, it can be shown, as we shall see in the following, some particular properties of our model which are different compared with standard INAR-type models.

First, note that k-step conditional measures of X_{t+k} on X_t depend not only of X_t, but also of the Noise-Indicator q_t (i.e., of the level of significance the PS-innovation ε_t). According to some well-known properties of the binomial thinning operator (see, for instance, Silva and Oliveira (2004)), it can be easily obtained the first-step conditional mean

$$\mathrm{E}[X_{t+1}|X_t] = \alpha X_t + \mu_\varepsilon q_t = \begin{cases} \alpha X_t + \theta g'(\theta), & \varepsilon_t \geq c, \\ \alpha X_t, & \varepsilon_t < c, \end{cases}$$

as well as conditional variance

$$\begin{aligned} \mathrm{Var}[X_{t+1}|X_t] &= \alpha(1-\alpha)X_t + \sigma_\varepsilon^2 q_t \\ &= \begin{cases} \alpha(1-\alpha)X_t + \theta g'(\theta) + \theta^2 g''(\theta), & \varepsilon_t \geq c, \\ \alpha(1-\alpha)X_t, & \varepsilon_t < c. \end{cases} \end{aligned}$$

In the general case, using the induction method and after some computations, k-step conditional measures can be calculated for each $k \in \mathbf{N}$. In this way, it follows

Theorem 2.2. *Let (X_t) be a NIINAR(1) model, that is, the sequence of r.v.s defined by equation (2.7). Then, the k-step conditional mean and variance of X_{t+k} on X_t*

are, respectively,

$$\mathrm{E}[X_{t+k}|X_t] = \alpha^k X_t + \alpha^{k-1}\mu_\varepsilon q_t + \frac{1-\alpha^{k-1}}{1-\alpha} m_c \mu_\varepsilon, \tag{2.8}$$

$$\begin{aligned} \mathrm{Var}[X_{t+k}|X_t] &= \alpha^k(1-\alpha^k)X_t + \alpha^{k-1}(1-\alpha^{k-1})\mu_\varepsilon q_t + \alpha^{2(k-1)}\sigma_\varepsilon^2 q_t \\ &\quad + \frac{(\alpha-\alpha^{k-1})(1-\alpha^{k-1})\mu_\eta + (1-\alpha^{2(k-1)})\sigma_\eta^2}{1-\alpha^2}, \end{aligned} \tag{2.9}$$

and the autocorrelation function (a.c.f.) at lag k is $\rho(k) = \alpha^k$.

If we put that $k \to \infty$, equalities (2.8) and (2.9) give the unconditional mean and variance of X_t, respectively,

$$\mu_X := \mathrm{E}[X_t] = \lim_{k\to\infty} \mathrm{E}[X_{t+k}|X_t] = \frac{m_c\mu_\varepsilon}{1-\alpha} = \frac{m_c\theta g'(\theta)}{1-\alpha}, \tag{2.10}$$

$$\begin{aligned} \sigma_X^2 &:= \mathrm{Var}[X_t] = \lim_{k\to\infty} \mathrm{Var}[X_{t+k}|X_t] = \frac{\alpha\mu_\eta + \sigma_\eta^2}{1-\alpha^2} \\ &= \mu_X + \frac{m_c\theta^2}{1-\alpha^2}\big[g''(\theta) + F_\varepsilon(c)\big(g'(\theta)\big)^2\big]. \end{aligned} \tag{2.11}$$

Notice that differences $D_X(\theta) := \sigma_X^2 - \mu_X$ and $D_\eta(\theta) = \sigma_\eta^2 - \mu_\eta$ satisfy relation $D_X(\theta) = D_\eta(\theta)/(1-\alpha^2)$. According to this, the following statement is valid.

Corollary 2.1. *The over-dispersion conditions of the series (X_t) and (η_t) are equivalent, that is, (X_t) is over-(equal- or under-)dispersed if and only if the same properties has the series (η_t).*

Thus, the series (X_t) will be over-dispersed if and only if the series (η_t) is over-dispersed, that is, if and only if the condition (2.6) is fulfilled. Moreover, above-mentioned over-dispersion properties of the NIPS-series (η_t) represent, at the same time, the over-dispersion properties of the series (X_t). In this way, the over-dispersion of the NIINAR(1) model can be investigated (only) based on the over-dispersion of its NIPS-innovations.

In the following, we prove some of the distributional properties of the NIINAR(1) model. First of all, we have investigated, so-called, infinite-order INteger Moving Average (INMA) representation of the series (X_t). After that, we give an explicit expression of its probability generating function (PGF), for an arbitrary order $r \in \mathbf{N}$. At last, we consider its Markovian properties, as well as a marginal distribution of our model.

Theorem 2.3. *Let suppose that PS-series (ε_t) has a finite moments of the first two order, which are uniformly bounded on $\theta \in (0, R)$. Then, for an arbitrary $\alpha \in$*

$(0, 1)$, *series* (X_t) *has an INMA*(∞) *representation*

$$X_t \overset{d}{=} \sum_{k=0}^{\infty} \alpha^k \circ \eta_{t-k}, \tag{2.12}$$

where the sum above converges in mean-square and almost surely.

Proof. See the Appendix. □

Remark 2.2. Using the p.m.f of η_t, given by the equation (2.5), their PGF is as one

$$\Psi_\eta(u) = \big[(1 - m_c)\Psi_{I_0} + m_c \Psi_\varepsilon\big](u) = 1 + m_c\left(\frac{f(u\theta)}{f(\theta)} - 1\right),$$

where $\Psi_{I_0}(u) \equiv 1$ and $\Psi_\varepsilon(u) = f(u\theta)(f(\theta))^{-1}$ are the PGFs of $I_0 \overset{\text{a.s.}}{=} 0$ and ε_t, respectively. Substituting the last expression in the equation (A.4), the PGF of X_t can be rewritten as

$$\Psi_X(u) = \prod_{k=0}^{\infty}\left[1 + m_c\left(\frac{f((1 + \alpha^k(u - 1))\theta)}{f(\theta)} - 1\right)\right].$$

Moreover, let $\mathbf{u} = (u_1, \ldots, u_r)' \in \mathbf{R}^r$, $r \geq 2$ and $\mathbf{X}_t^{(r)} := (X_t, \ldots, X_{t+r-1})'$, $t \in \mathbf{Z}$ be the so-called *overlapping blocks* of the process (X_t). Then, we can introduce *the* r*-dimensional PGF* of the random vector $\mathbf{X}_t^{(r)}$ as

$$\Psi_X^{(r)}(\mathbf{u}; \Theta) := \mathrm{E}\big[u_1^{X_t} \cdots u_r^{X_{t+r-1}}\big], \tag{2.13}$$

where $\Theta = (\theta, \alpha, m_c)'$ is a vector of unknown parameters of NIINAR(1) process. Substituting $k = 1, \ldots, r-1$ in equation (A.5), and after some computation similar as above, it can be obtain an explicit expression of this PGF:

$$\Psi_X^{(r)}(\mathbf{u}; \Theta) = \Psi_X\left(\prod_{k=0}^{r-1}(1 + \alpha^k(u_{k+1} - 1))\right)\prod_{\ell=2}^{r}\Psi_\eta\left(\prod_{k=0}^{r-\ell}(1 + \alpha^k(u_{k+\ell} - 1))\right).$$

The function $\Psi_X^{(r)}(\mathbf{u}; \Theta)$ will be used later in the estimation procedure of the parameters of our model (see Section 3, below).

Theorem 2.4. *The NIINAR*(1) *series* (X_t) *is a homogeneous Markovian process. The first-step transition probabilities* $p_{jk} := P\{X_t = k | X_{t-1} = j\}$ *can be written as*

$$\begin{aligned} p_{jk} &= (1 - m_c)\binom{j}{k}\alpha^k(1 - \alpha)^{j-k} I(N_2 \geq k) \\ &\quad + m_c \sum_{i=N_1}^{N_2}\binom{j}{i}\alpha^i(1 - \alpha)^{j-i} p_\varepsilon(k - i), \end{aligned} \tag{2.14}$$

where $p_\varepsilon(x;\theta)$ *is the p.m.f. of the series* (ε_t), *and*

$$(N_1, N_2) = \begin{cases} (0, \min\{j, k-s\}), & \text{if } \mathcal{S} = \{s, s+1, \dots\}, \\ (\max\{0, k-s\}, \min\{j, k\}), & \text{if } \mathcal{S} = \{0, 1, \dots, s\} \end{cases}$$

for some fixed $s \in \mathbf{Z}^+$.

Proof. See the Appendix. □

Remark 2.3. In the following, we suppose that the condition $0 \in \mathcal{S}$ (which is the usual in zero-inflated distributions) is always fulfilled. Notice that in this case $I(k \le N_2) = I(k \le j)$. It means that the first, singular part in equality (2.14) is greater than zero if and only if the process (X_t) exceeds from state $X_{t-1} = j$ to "non-increasing" state $X_t = k \le j$, and vice versa.

At the end of this section, notice that above-mentioned theorems, first of all the equalities (2.12) and (2.14), imply the following important properties of the NI-INAR(1) model.

Corollary 2.2. *Series* (X_t) *is strictly stationary and ergodic process, and its marginal p.m.f* $p_X(x;\theta) := P\{X_t = x\}$ *is*

$$\begin{aligned} p_X(x;\theta) &= (1-m_c)\sum_{j=x}^{\infty} p_X(j;\theta)p_j(x;\alpha) \\ &\quad + m_c \sum_{j=0}^{\infty} p_X(j;\theta) \sum_{i=N_1}^{N_2} p_j(i;\alpha)p_\varepsilon(x-i). \end{aligned} \tag{2.15}$$

3 Estimation of the model's parameters

Procedure of the parameters estimation of NIINAR(1) model, because of its specific structure, is much more complex than with the most of the similar linear INAR models. For example, according to the equation (2.8), conditional mean $\mathrm{E}[X_{t+k}|X_t]$ depends on (not observable) realizations of the indicator $q_t(c)$. Consequently, one of the typical estimation methods, such as conditional last squares (CLS) method, cannot be used here. On the other hand, according to Theorem 2.2 and the equations (2.10)–(2.11), the Yule–Walker (YW) estimators $\widetilde{\theta}$, $\widetilde{\alpha}$, $\widetilde{m}_c$ of the unknown parameters θ, α, m_c, respectively, can be obtain as it follows:

$$\begin{aligned} \widetilde{\alpha} &= \widehat{\gamma}(1)/\widehat{\gamma}(0), \\ \widetilde{m}_c\widetilde{\theta}g'(\widetilde{\theta}) &= (1-\widetilde{\alpha})\bar{X}_T, \\ \widehat{\gamma}(0) &= \frac{\widetilde{m}_c\widetilde{\theta}g'(\widetilde{\theta})}{1-\widetilde{\alpha}} + \frac{\widetilde{m}_c\widetilde{\theta}^2 g''(\widetilde{\theta})}{1-\widetilde{\alpha}^2}\big[g''(\widetilde{\theta}) + (1-\widetilde{m}_c)\big(g'(\widetilde{\theta})\big)^2\big], \end{aligned} \tag{3.1}$$

where $\bar{X}_T := \frac{1}{T}\sum_{t=1}^{T} X_t$, $\widehat{\gamma}(k) := \frac{1}{T}\sum_{t=1}^{T-k}(X_t - \bar{X}_T)(X_{t+k} - \bar{X}_T)$, $k = 0, 1$ and $\{X_1, \dots, X_T\}$ is some realization of the series (X_t). Using some well-known facts about the asymptotic distribution of the INAR(p) processes (cf. Silva and Silva, 2006), as well as continuity of the stochastic convergences (cf. Serfling, 1980, pp. 24, 118), the strong consistency and asymptotic normality of the YW estimates can be proved.

In order to get the more efficient estimators of the parameters of the NIINAR(1) model, we use a novel estimation technique, which we called *the probability generating functions (PGF) method*. The main aim of this method is to minimize "the distance" between the theoretical PGF, defined with equation (2.13), and the appropriate *Empirical PGF* (of order $r \in \mathbf{N}$):

$$\widetilde{\Psi}_T(\mathbf{u}) := \frac{1}{T - r + 1}\sum_{t=1}^{T-r+1} u_1^{X_t} \cdots u_r^{X_{t+r-1}},$$

where $\mathbf{u} = (u_1, \dots, u_r)' \in \mathbf{R}^r$. It this well known that theoretical PGF $\Psi_X^{(r)}(\mathbf{u}; \theta)$ converges at least for all $\mathbf{u} \in [-1, 1]^r$. Thus, the objective function can be defined as

$$S_T^{(r)}(\Theta) := \int_{-1}^{1} \cdots \int_{-1}^{1} g(\mathbf{u}) \big|\Psi_X^{(r)}(\mathbf{u}; \theta) - \widetilde{\Psi}_T^{(r)}(\mathbf{u})\big|^2 \,\mathrm{d}\mathbf{u}, \tag{3.2}$$

where $\Theta = (\theta, \alpha, m_c)'$ is a vector of unknown parameters, $\mathrm{d}\mathbf{u} := \mathrm{d}u_1 \cdots \mathrm{d}u_r$ and $g : \mathbf{R}^r \to \mathbf{R}^+$ is a some weight function, integrable on $[-1, 1]^r$. Estimates based on the PGF method will be obtained by the minimization the objective function (3.2) with respect to Θ. More accurate, they represent the solutions of the minimization equation

$$\widehat{\Theta}_T^{(r)} = \arg\min_{\Theta \in K} S_T^{(r)}(\Theta), \tag{3.3}$$

where $K := (0, R) \times (0, 1)^2$ is the parameter space of the nontrivial and stationary NIINAR(1) process. We point out that the basic idea of the PGF estimation method is a similar as in the so-called *Empirical characteristic function (ECF) method* (see, for instance, Yu, 2004). It is based on the fact that PGF of the series (X_t), as its CF, has the same information about its distribution. Similarly like that, we shall investigate, under some conditions, strong consistency and asymptotic normality (AN) of the PGF estimators of NIINAR(1) model's parameters.

Theorem 3.1. *Let Θ_0 be the true value of the parameters set K, and for an arbitrary $T = 1, 2, \dots$, let $\widehat{\Theta}_T^{(r)}$ be solutions of the equation* (3.3). *Additionally, let suppose that the following regularity conditions are fulfilled*:

(i) $\Theta_0 \in K$ *and* $\widehat{\Theta}_T^{(r)} \in K$, *for T large enough*;

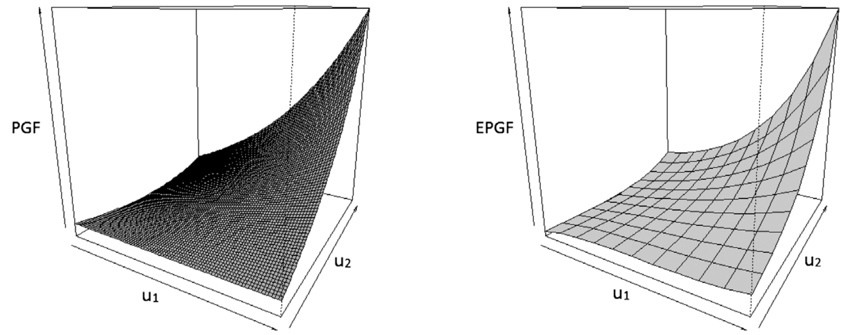

Figure 1 *Graphs of the two-dimensional PGF (panel left) and the appropriate EPGF (panel right) of the series* $\mathbf{X}_t^{(2)} = (X_t, X_{t+1})$. *Innovations are Poisson distributed r.v.s with* $\theta = 1$.

(ii) *Function*

$$S_0^{(r)}(\Theta) := \int_{-1}^{1} \cdots \int_{-1}^{1} g(\mathbf{u}) \left| \Psi_X^{(r)}(\mathbf{u}; \Theta) - \Psi_X^{(r)}(\mathbf{u}; \Theta_0) \right|^2 \mathrm{d}\mathbf{u}$$

has an unique minima at $\Theta = \Theta_0$;

(iii) $\dfrac{\partial^2 S_T^{(r)}(\Theta_0)}{\partial \Theta \, \partial \Theta'}$ *is a regular matrix.*

Then, $\widehat{\Theta}_T^{(r)}$ *is strictly consistent and asymptotically normal estimator for* Θ.

Proof. See the Appendix. □

Remark 3.1. In order to obtain the PGF estimators of true values of the parameter $\Theta = \Theta_0$, and using the same deliberation as in the ECF estimates of the AR(1)-processes (see, for instance, Knight and Yu, 2002), in the further we will use the PGF of order $r = 2$ (Figure 1). Therefore, the PGF procedure will be based on realizations of the two-dimensional random vector $\mathbf{X}_t^{(2)} := (X_t, X_{t+1})'$, and the objective function $S_T^{(2)}$ represents a double integral with respect to the weight function $g : \mathbf{R}^2 \to \mathbf{R}^+$. Consequently, as we shall see in the following, it can be numerically approximated by using some cubature formulas.

4 Numerical simulations

In this section, we consider the practical application of the above-mentioned procedures of estimation of the unknown parameters $\Theta = (\theta, \alpha, m_c)'$ of the NIINAR(1) process:

$$X_t = \alpha \circ X_{t-1} + q_{t-1}\varepsilon_t, \qquad t = 1, \ldots, T. \tag{4.1}$$

For this purpose, we have taken two different types of distribution of the innovations (ε_t). We firstly assumed that it was the i.i.d. sequence with Poisson distribution, and then that (ε_t) has a geometric distribution. For the both of them, as for each of the sample size $T \in \{100, 500, 2500\}$, we generated 500 independent Monte Carlo simulations of the series (ε_t) and $(q_{t-1}\varepsilon_t)$. Thereafter, according to equation (4.1), we have obtained the appropriate realizations $\{X_0, X_1, \dots, X_T\}$ of the series (X_t), where we set $X_0 \overset{\text{a.s.}}{=} 0$.

For the parameters' estimators of our model, we have first used the YW estimates, as solutions of the equations (3.1). In the case of Poisson innovations (ε_t), the appropriate estimators are:

$$\begin{aligned}
\widetilde{\alpha} &= \widehat{\gamma}(1)/\widehat{\gamma}(0), \\
\widetilde{\theta} &= (1+\widetilde{\alpha})\big[\widehat{\gamma}(0)/\bar{X}_T - 1\big] + (1-\widetilde{\alpha})\bar{X}_T, \\
\widetilde{m}_c &= (1-\widetilde{\alpha})\bar{X}_T/\widetilde{\theta},
\end{aligned}$$

and for geometric distributed innovations (ε_t), the YW estimates are:

$$\begin{aligned}
\widetilde{\alpha} &= \widehat{\gamma}(1)/\widehat{\gamma}(0), \\
\widetilde{\theta} &= 1 - 2\big[(1-\widetilde{\alpha})(1+\bar{X}_T) + (1+\widetilde{\alpha})\widehat{\gamma}(0)/\bar{X}_T\big]^{-1}, \\
\widetilde{m}_c &= \widetilde{\theta}^{-1}(1-\widetilde{\theta})(1-\widetilde{\alpha})\bar{X}_T.
\end{aligned}$$

The numerical results obtained by YW estimates, that is, their means (Mean), minima (Min.), maxima (Max.) and the standard estimating errors (SEE), are presented in the left parts of Table 2 and Table 3, respectively. In both cases, the YW estimates of NIINAR(1) model are convergent, and the standard estimating errors are decreasing as sample size is increasing.

In following, we apply the PGF method, with the initial values that were obtained in the previous estimation procedure. In this way, the PGF estimates $\widehat{\theta}$, $\widehat{\alpha}$, $\widehat{m}_c$ are computed according to the minimization of the (double) integral

$$S_T^{(2)}(\theta) = \iint_{[-1,1]^2} g(u_1,u_2)\big|\Psi_X^{(2)}(u_1,u_2;\theta) - \widetilde{\Psi}_T(u_1,u_2)\big|^2 \,\mathrm{d}u_1\,\mathrm{d}u_2, \qquad (4.2)$$

with respect to the weight function $g : [-1,1]^2 \to \mathbf{R}^+$. Integral in (4.2) can be numerically approximated by using some N-point cubature formula

$$I(f;g) := \iint_{[-1,1]^2} g(u_1,u_2)f(u_1,u_2)\,\mathrm{d}u_1\,\mathrm{d}u_2 \approx C_N(f) := \sum_{j=1}^{N} \omega_j f(u_{1j},u_{2j}),$$

where (u_{1j}, u_{2j}) are the cubature nodes, and ω_j are the corresponding weight coefficients. In our simulations study, we used Gauss–Legendre cubature formulas with $g(u_1;u_2) \equiv 1$ and $N = 36$ nodes. The numerical construction of these formulas was done by the MATHEMATICA package "`Orthogonal Polynomials`" (cf. Cvetković and Milovanović, 2004). After that, the objective function (4.2) is

Table 2 *Estimated values of the parameters of* NIINAR(1) *model with Poisson innovations*. (*True values of the parameters are*: $\theta = 1$, $\alpha = 0.5$, $m_c = 0.2642\ldots$)

Sample		YW estimates				PGF estimates			
		$\widetilde{\theta}$	$\widetilde{\alpha}$	$\widetilde{m}_c$	$S_T^{(2)}$	$\widehat{\theta}$	$\widehat{\alpha}$	$\widehat{m}_c$	$S_T^{(2)}$
$T = 100$	Min.	0.5551	0.1470	0.1284	0.0020	0.6187	0.2339	0.1387	0.0010
	Mean	1.0646	0.4285	0.2552	0.1550	1.0461	0.4674	0.2579	0.0104
	Max.	1.8780	0.6865	0.4170	0.7552	1.4354	0.6754	0.3999	0.0394
	SEE	0.2504	0.0997	0.2400	0.1615	0.2054	0.0856	0.1612	0.0075
$T = 500$	Min.	0.7371	0.2695	0.1781	0.0010	0.7723	0.3397	0.1582	0.0006
	Mean	1.0585	0.4441	0.2698	0.1536	1.0381	0.4689	0.2613	0.0074
	Max.	1.3840	0.5637	0.3025	0.7050	1.2980	0.6581	0.3667	0.0353
	SEE	0.1195	0.0476	0.0929	0.0773	0.1045	0.0636	0.0925	0.0072
$T = 2500$	Min.	0.8988	0.3936	0.2289	0.0004	0.9122	0.4492	0.2463	7.30E-5
	Mean	1.0317	0.4945	0.2653	0.0523	0.9894	0.5019	0.2646	1.16E-4
	Max.	1.1930	0.5354	0.3082	0.3502	1.1061	0.5425	0.2828	1.69E-3
	SEE	0.0378	0.0153	0.0144	0.0074	0.0345	0.0061	0.0447	1.51E-5

minimized by a Nelder–Mead method, and the estimation procedure is realized by the original authors' codes written in statistical programming language "R".

Summary statistics of the PGF estimates, computed via this estimation procedure are presented in the right parts of Tables 2 and 3. For both types of estimates, the values of the objective function $S_T^{(2)}$, as the reference estimation error, were also shown. In comparison to initial estimates, it is obvious that means of the PGF estimates are somewhat closer to the true values of parameters, as well as they have a smaller estimation errors.

5 Application of the model

We describe here a practical application of the NIINAR(1) process in modeling the dynamics of some actual time series. For this purpose, we analyze two real-life datasets. The first one, supplied by the Statistical Office of the Republic of Serbia, contains the number of daily mortality in Niš, the second largest city in Republic of Serbia, in the period from 1992 to 2009. The empirical distribution of this time series is shown in the Figure 2 (panel left), and at first glance it has a typical (equal-dispersed) Poisson distribution. The second time series contains the daily number of casualties in road traffic crashes on the territory of Belgrade, the Serbian capital, from 2004 to 2013, according to a statistics collected by Ministry of Interior of the Republic of Serbia. The right panel in Figure 2 represents the empirical distribution of this series, and it is obvious pronounced adapting with the (over-dispersed) negative binomial distribution (i.e., geometric, in a special case).

Table 3 *Estimated values of the parameters of* NIINAR(1) *model with geometric innovations*. (*True values of the parameters are*: $\theta = \alpha = 0.5$, $m_c = 0.25$)

Sample		YW estimates $\widetilde{\theta}$	$\widetilde{\alpha}$	$\widetilde{m}_c$	$S_T^{(2)}$	PGF estimates $\widehat{\theta}$	$\widehat{\alpha}$	$\widehat{m}_c$	$S_T^{(2)}$
$T = 100$	Min.	0.2161	0.3151	0.0595	0.1321	0.2887	0.3249	0.1430	0.0014
	Mean	0.4314	0.4058	0.2741	0.3686	0.5254	0.5481	0.2511	0.0048
	Max.	0.6796	0.6671	0.6363	0.9747	0.6609	0.6379	0.4773	0.0091
	SEE	0.0986	0.1334	0.0890	0.2081	0.0853	0.1160	0.0447	0.0015
$T = 500$	Min.	0.3233	0.3642	0.0649	0.0912	0.4131	0.4317	0.2075	0.0032
	Mean	0.4452	0.4414	0.2573	0.3437	0.5024	0.5031	0.2494	0.0048
	Max.	0.6033	0.6056	0.5336	0.5649	0.5703	0.5763	0.3186	0.0073
	SEE	0.0378	0.0596	0.0786	0.1107	0.0304	0.0554	0.0186	0.0006
$T = 2500$	Min.	0.4401	0.4128	0.1386	0.0355	0.4597	0.4232	0.2268	0.0028
	Mean	0.5050	0.4953	0.2485	0.1218	0.5010	0.5016	0.2497	0.0039
	Max.	0.5633	0.5787	0.3897	0.2020	0.5408	0.5789	0.2765	0.0059
	SEE	0.0191	0.0253	0.0370	0.0968	0.0141	0.0262	0.0080	0.0003

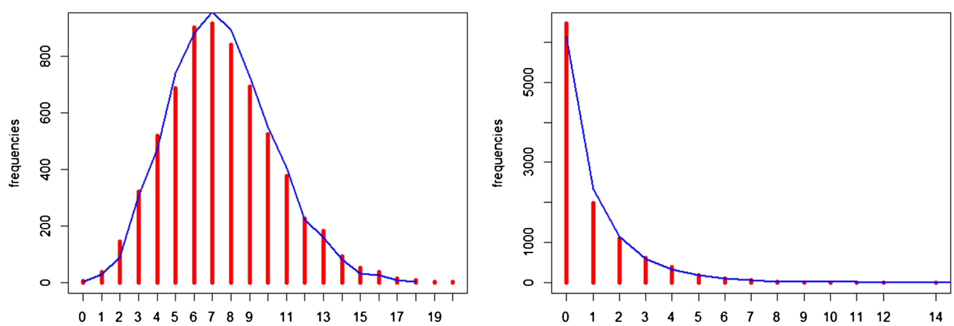

Figure 2 *Empirical distributions of the two real data series, fitted by the Poisson distribution* (*panel left*), *and the negative binomial distribution* (*panel right*).

A simple descriptive statistical analysis of both series, denoted as Series A and Series B, respectively, is shown in Table 4. It can be recognized that in both of them there exists an over-dispersion. On the other hand, Series B has a highly frequented values in zero (more than half of its realizations), i.e., the zero-inflated distribution. Finally, the both series have a positive autocorrelation, which is a typical characteristic of INAR-models.

For the both of time series, we analyze the modeling with the the NIINAR(1) model, in comparison to the standard INAR(1) process. For the Series A, according to the aforementioned facts, the Poisson innovations are considered, as for the Series B, we assumed the geometric distributed innovations.

Estimated values of the parameters for both series (and for both INAR-type models) are shown in the top half of Tables 5 and 6. As in previous simulations

Table 4 *Summary statistics of the real data series*

Statistics	Series A	Series B
Sample size	6,575	10,959
Min.	0.000	0.000
1st Qu.	5.000	0.000
Median	7.000	0.000
3rd Qu.	9.000	1.000
Max.	20.000	14.000
Mean	7.476	0.908
Variance	8.801	2.189
ACF(1)	0.165	0.477
ACF(2)	0.136	0.435
ACF(3)	0.150	0.441
...	...	...
ACF(10)	0.119	0.429

study, for the initial values of parameters $\Theta = (\theta, \alpha, m_c)'$ we have used the YW estimates and, therefore, we have computed the PGF estimates. At last, solving the equations $P\{\varepsilon_t \geq c\} = \widetilde{m}_c$ and $P\{\varepsilon_t \geq c\} = \widehat{m}_c$ with respect to c, we obtain the estimate of the critical value, that is,

$$\widetilde{c} := \inf_{x \in \mathbf{N}_0} P\{\varepsilon_t < x\} \geq 1 - \widetilde{m}_c, \qquad \widehat{c} := \inf_{x \in \mathbf{N}_0} P\{\varepsilon_t < x\} \geq 1 - \widehat{m}_c.$$

It can be easily seen that, for both empirical data series, YW and PGF estimates of parameter m_c give the nontrivial estimated values $c > 0$. At first glance, it is some confirmation and justification for the NIINAR model introducing. It is also noticeable that estimated values of the parameter α of the Series B are relatively "small" ($\alpha < 0.05$), which is expected as a consequence of the zero-inflated distribution of this series.

Furthermore, we analyze the efficiency of fitting for both empirical data series (and for both of INAR-type models), when the PGF method, as well as the Yule–Walker method of moments were applied. For this purpose, we generated 500 independent simulations of INAR(1) and NIINAR(1) models, for all estimated values of their parameters. In order to check the efficiency of proposed models, we computed two typical goodness-of-fit statistics: the root mean squares of differences of observations and predicted values (RMS), as well as the Akaike Information Criterion (AIC). The average values of the all of these statistics are shown in the lower part of Tables 5 and 6. Notice that, in the case of Series A, the standard INAR(1) model and the NIINAR(1) model have similar efficiency, because of the RMS and AIC statistics of both models are relatively close, for the both of estimation methods which were used here. However, there is somewhat better fitting with the NIINAR(1) model. On the other hand, it is evident that in the case of the

Table 5 *Estimated values of parameters and error statistics of the empirical data* (*Series A*)

	YW estimates		PGF estimates	
Parameters	INAR(1)	NIINAR(1)	INAR(1)	NIINAR(1)
θ	6.2430	6.4494	5.4170	5.7204
α	0.1649	0.1649	0.2701	0.2370
m_c	1.0000	0.9680	1.0000	0.9930
c	0.0000	2.0000	0.0000	1.0000
$S_T^{(2)}$	1.96E-5	4.84E-7	7.79E-6	1.03E-7
RMS	0.4012	0.3970	0.3978	0.3944
AIC	329.57	329.50	329.57	329.47
DM	0.5066		0.8146	
(p-value)	(0.3027)		(0.2077)	

Table 6 *Estimated values of parameters and error statistics of the empirical data* (*Series B*)

	YW estimates		PGF estimates	
Parameters	INAR(1)	NIINAR(1)	INAR(1)	NIINAR(1)
θ	0.3219	0.3561	0.5888	0.5882
α	0.0477	0.0477	0.0493	0.0434
m_c	1.0000	0.3712	1.0000	0.2582
c	0.0000	1.0000	0.0000	3.0000
$S_T^{(2)}$	0.2544	0.0137	0.1990	2.62E-4
RMS	0.2190	0.1762	0.2184	0.1716
AIC	396.86	288.82	396.78	288.35
DM	4.1602		8.9784	
(p-value)	(1.78E-5)		($<$2.20E-16)	

Series B, estimated values of the fitting statistics are generally less when the NIINAR(1) model has been applied, as an appropriate fitting model, than when the standard INAR(1) model has been applied.

Similar conclusions can be reached on the basis of analysis of the forecast accuracy between the INAR(1) and NIINAR(1) model. In this sense, we have compared the accuracy of these two models by using the well-known, one-sided Diebold–Mariano test for predictive accuracy: the null hypothesis is that these two models have the same forecast accuracy, as the alternative is that NIINAR(1) is more accurate model. Test statistics (labeled as DM), as well as the corresponding p-values, were computed by using the R-package “forecast” (Hyndman, 2016). In the case of Series A, dataset from 2010 to 2013 was used as the forecast horizon. Realized values of the DM statistics, shown in Table 5, indicate that the both of models have the similar accuracy (i.e., the null hypothesis is valid). On the other hand, in the

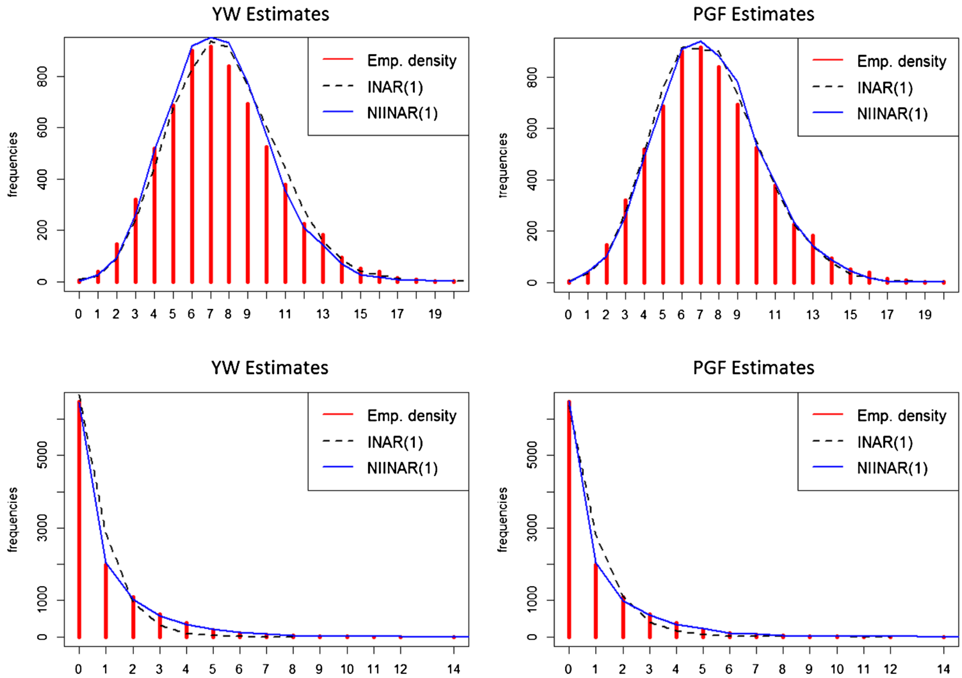

Figure 3 *Empirical and fitted p.m.f.s of the INAR*(1) *and NIINAR*(1) *processes*: *Series A* (*graphs above*) *and Series B* (*graphs below*).

case of Series B, where the forecast horizon from 2014 to 2015 was used, there is a significantly greater accuracy in prediction when the NIINAR(1) model was used (Table 6).

Some of these, aforementioned facts can be also seen in Figure 3, where the empirical p.m.f.s of the both of actual time series, along with the appropriate p.m.f.s of INAR(1) and NIINAR(1) processes, are shown. It can be easily seen that in the case of Series A the similar efficiency has been achieved by using the both of models. In the case of Series B, it is obvious that the NIINAR(1) model provide better match to the empirical p.m.f. in comparison with the standard INAR(1) model.

Appendix

Proof of Theorem 2.1. According the definition of the series (η_t), and using the conditional probability, we have

$$\begin{aligned} p_\eta(x;\theta) &= P\{I_0 = x\}P\{q_{t-1} = 0\} + P\{\varepsilon_t = x\}P\{q_{t-1} = 1\} \\ &= (1 - m_c)p_0(x) + m_c p_\varepsilon(x;\theta), \end{aligned}$$

where $x = 0, 1, 2, \ldots,$ and $p_0(x)$ is the p.m.f. of the r.v. $I_0 \overset{\text{a.s.}}{=} 0$. From here, it is easily to show that equality (2.5) holds. Also, it is obviously that $q_t^n \overset{d}{=} q_t$, for each $n \in \mathbf{N}$, which implies the equalities $\mu_n^{(\eta)} = m_c \mu_n^{(\varepsilon)}$. From this, as well as the expressions (2.3) and (2.4) for the mean and the variance of r.v.s ε_t, we obtain the mean and the variance of r.v.s η_t, respectively,

$$\mu_\eta := \mathrm{E}[\eta_t] = m_c \mu_\varepsilon = m_c \theta g'(\theta) = m_c \theta \frac{f'(\theta)}{f(\theta)}, \tag{A.1}$$

$$\begin{aligned} \sigma_\eta^2 := \mathrm{Var}[\eta_t] &= m_c \frac{\theta f'(\theta) + \theta^2 f''(\theta)}{f(\theta)} - m_c^2 \left(\theta \frac{f'(\theta)}{f(\theta)} \right)^2 \\ &= m_c \theta \frac{f'(\theta)}{f(\theta)} + m_c \theta^2 \frac{f''(\theta) f(\theta) - m_c (f'(\theta))^2}{[f(\theta)]^2} \\ &= \mu_\eta + m_c \theta^2 \big[g''(\theta) + F_\varepsilon(c) \big(g'(\theta) \big)^2 \big]. \end{aligned} \tag{A.2}$$

As we assumed that $F_\varepsilon(c) \in (0, 1)$, the inequality $D_\eta(\theta) := \sigma_\eta^2 - \mu_\eta > 0$ holds if and only if the condition (2.6) is valid. Finally, for an arbitrary $k \in \mathbf{N}$, it follows

$$\begin{aligned} \mathrm{Cov}[\eta_t, \eta_{t+k}] &= \mathrm{E}[\eta_t \eta_{t+k}] - \big(\mathrm{E}[\eta_t] \big)^2 \\ &= \begin{cases} \mathrm{E}[q_{t-1}] \mathrm{E}[q_t \varepsilon_t] \mathrm{E}[\varepsilon_{t+1}] - \mu_\eta^2, & k = 1, \\ \mathrm{E}[q_{t-1}] \mathrm{E}[q_{t+k-1}] \mathrm{E}[\varepsilon_t] \mathrm{E}[\varepsilon_{t+k}] - \mu_\eta^2, & k > 1, \end{cases} \\ &= \begin{cases} m_c \mu_\varepsilon \mathrm{E}\big[(1 - m_c) I_0 + m_c \varepsilon_t \big] - m_c^2 \mu_\varepsilon^2, & k = 1, \\ m_c^2 \mu_\varepsilon^2 - m_c^2 \mu_\varepsilon^2, & k > 1, \end{cases} \\ &= 0. \end{aligned}$$

In this way, the theorem is completely proven. $\square$

Proof of Theorem 2.3. According to the assumptions given in the theorem, there exists a constant $M > 0$ such that $0 < \mu_\varepsilon(\theta) \le M < +\infty$ for any $\theta \in (0, R)$. Then,

$$\sum_{k=1}^{\infty} P\{\varepsilon_t \ge k\} = \sum_{k=1}^{\infty} k P\{\varepsilon_t = k\} = \mu_\varepsilon(\theta) \le M < +\infty,$$

that is, the sum above converges uniformly on $\theta \in (0, R)$. Furthermore, after some simple calculations, it follows $P\{\eta_t \ge k\} = m_c P\{\varepsilon_t \ge k\}$, that is,

$$\sum_{k=1}^{\infty} P\{\eta_t \ge k\} = m_c \mu_\varepsilon(\theta) < \mu_\varepsilon(\theta) < +\infty.$$

On the other hand, as the sequence $b_k = \frac{1}{k}$, $k = 1, 2, \ldots$ is a monotone and bounded, the Abel's convergence criteria implies

$$\sum_{k=1}^{\infty} \frac{1}{k} P\{\eta_t \ge k\} \le \sum_{k=1}^{\infty} \frac{1}{k} P\{\varepsilon_t \ge k\} < +\infty, \tag{A.3}$$

uniformly on θ. As it is known from Alzaid and Al-Osh (1990), the inequality (A.3) is necessary and sufficient condition for the equality

$$\Psi_X(u) = \prod_{k=0}^{\infty} \Psi_\eta\big(1 + \alpha^k(u-1)\big), \tag{A.4}$$

where $\Psi_X := \mathrm{E}[u^{X_t}]$ and $\Psi_\eta := \mathrm{E}[u^{\eta_t}]$ are the PGFs of the r.v.s X_t and η_t, respectively, and the product above converges absolutely at least for all $u \in [-1, 1]$. Hence, the equality (A.4) is equivalent to the INMA(∞) representation (2.12).

In order to prove the second part of theorem, notice that, according to the definition (2.7) of the series (X_t), for an arbitrary $k \in \mathbf{N}$ holds

$$X_t \overset{d}{=} \alpha^k \circ X_{t-k} + \sum_{j=0}^{k-1} \alpha^j \circ \eta_{t-j}. \tag{A.5}$$

This implies

$$\mathrm{E}\left[X_t - \sum_{j=0}^{k-1} \alpha^j \circ \eta_{t-j}\right]^2 = \alpha^{2k}\mathrm{E}\big[X_{t-k}^2\big] + \alpha^k(1-\alpha^k)\mathrm{E}[X_{t-k}] \longrightarrow 0, \qquad k \to \infty,$$

and the mean-square convergence of the sum in (2.12) is confirmed. Now, if we define the event $A := \{\lim_{k\to\infty} \sum_{j=0}^{k-1} \alpha^j \circ \eta_{t-j} = X_t\}$, i.e., $A = \bigcup_{n=1}^{\infty} A_n$, where $A_n := \bigcap_{k=n}^{\infty}\{\alpha^k \circ X_{t-k} = 0\}$, then we have

$$\begin{aligned}
P(A_n) &= \lim_{m\to\infty}\left(P\{\alpha^{m+n} \circ X_{t-m-n} = 0\} \times \prod_{k=n}^{m+n-1} P\{\alpha^k \circ \eta_{t-k} = 0\}\right)\\
&= \lim_{m\to\infty}\left(\sum_{j=0}^{\infty}(1-\alpha^{m+n})^j P\{X_{t-m-n} = j\}\right)\\
&\quad \times \lim_{m\to\infty} \prod_{k=n}^{m+n-1}\left(\sum_{j=0}^{\infty}(1-\alpha^k)^j P\{\eta_{t-k} = j\}\right)\\
&= \lim_{m\to\infty} \Psi_X\big(1-\alpha^{m+n}\big) \times \lim_{m\to\infty} \prod_{k=n}^{m+n-1} \Psi_\eta\big(1-\alpha^k\big)\\
&= \Psi_X(1) \times \prod_{k=n}^{\infty} \Psi_\eta\big(1-\alpha^k\big)\\
&= \prod_{k=n}^{\infty} \Psi_\eta\big(1-\alpha^k\big).
\end{aligned}$$

According to the continuity of probability, as well as the convergence of the product proven in (A.4), it follows $P(A) = \lim_{n\to\infty} P(A_n) = 1$, that is, the almost surely convergence of the sum in (2.12) holds. □

Proof of Theorem 2.4. For an arbitrary $n \in \mathbf{N}$, let us denote $p_n(x;\alpha) := \binom{n}{x}\alpha^x(1-\alpha)^{n-x}$, $x = 0, 1, \ldots, n$ the p.m.f. of the binomial $\mathcal{B}(n;\alpha)$ distributed r.v. Then, the conditional distribution of X_t for a given X_{t-1} can be written as the convolution of the binomial distribution and the distribution of NIPS-innovations (see, for more details, Bourguignon and Vasconcellos (2015)):

$$p_{jk} = \sum_{x=N_1}^{N_2} p_j(x;\alpha)p_\eta(k-x;\theta) = \sum_{x=N_1}^{N_2} \binom{j}{\alpha}\alpha^x(1-\alpha)^{j-x}p_\eta(k-x;\theta).$$

According to equality (2.5) for the p.m.f. of η_t, it follows

$$\begin{aligned} p_{jk} &= \sum_{x=N_1}^{N_2} p_j(x;\alpha)\big[(1-m_c)p_0(k-x) + m_c p_\varepsilon(k-x;\theta)\big] \\ &= (1-m_c)\sum_{x=N_1}^{N_2} p_j(x;\alpha)I(k=x) + m_c\sum_{x=N_1}^{N_2} p_j(x;\alpha)p_\varepsilon(k-x;\theta) \\ &= (1-m_c)p_j(k;\alpha)I(N_1 \le k \le N_2) + m_c\sum_{x=N_1}^{N_2} p_j(x;\alpha)p_\varepsilon(k-x;\theta). \end{aligned}$$

By definition of the numbers N_1, N_2 it can be easily seen that inequality $k \ge N_1$ is fulfilled for each $k \in \mathbf{Z}^+$, that is, (2.14) is valid. □

Proof of Theorem 3.1. In order to prove the consistency of $\widehat{\Theta}_T^{(r)}$, we check sufficient consistency conditions of extremum estimators (see, for instance Newey and McFadden, 1994). Under assumption (i), the set $\overline{K} = [0, R] \times [0, 1]^2$ is the compact, and $\Theta_0 \in \mathrm{int}(\overline{K})$. As the series (X_t) is ergodic and $\widetilde{\Psi}_T(\mathbf{u})$ is an unbiased estimator of $\Psi_X^{(r)}(\mathbf{u};\Theta_0)$, that is, $E[\widetilde{\Psi}_T^{(r)}(\mathbf{u})] = \Psi_X^{(r)}(\mathbf{u};\Theta_0)$, the strong law of large numbers gives $\widetilde{\Psi}_T(\mathbf{u}) \xrightarrow{\text{a.s.}} \Psi_X^{(r)}(\mathbf{u};\Theta_0)$. Hence, it follows

$$\sup_{\Theta\in K}\big|\widetilde{\Psi}_T(\mathbf{u}) - \Psi_X^{(r)}(\mathbf{u};\Theta_0)\big| \xrightarrow{\text{a.s.}} 0, \qquad T \to +\infty.$$

Further on, notice that $\Psi_T^{(r)}(\mathbf{u};\Theta)$ is a continuous function on the compact $[-1, 1]^r \times \overline{K}$, and $\widetilde{\Psi}_T^{(r)}(\mathbf{u})$ is a continuous on the compact $[-1, 1]^r$. Hence, for some $M_1, M_2 > 0$, the inequalities

$$\max_{(\mathbf{u},\Theta)\in[-1,1]^r\times\overline{K}}\big|\Psi_T^{(r)}(\mathbf{u};\Theta)\big| \le M_1 < +\infty, \qquad \max_{\mathbf{u}\in[-1,1]^r}\big|\widetilde{\Psi}_T^{(r)}(\mathbf{u})\big| \le M_2 < +\infty,$$

hold, and according to these, it can be easily obtain

$$\big|S_T^{(r)}(\Theta) - S_0^{(r)}(\Theta)\big| \le (3M_1 + M_2)\int_{-1}^{1}\cdots\int_{-1}^{1} g(\mathbf{u})\big|\widetilde{\Psi}_T^{(r)}(\mathbf{u}) - \Psi_X^{(r)}(\mathbf{u};\Theta_0)\big|\,d\mathbf{u}.$$

Thus,

$$\sup_{\Theta\in\overline{K}}\left|S_T^{(r)}(\Theta)-S_0^{(r)}(\Theta_0)\right|\xrightarrow{\text{a.s.}}0,\qquad T\to+\infty,$$

that is, $S_T^{(r)}(\Theta)$ uniformly converges almost surely to $S_0^{(r)}(\Theta)$. According to the aforementioned facts and the assumption (ii), Theorem 2.1 in Newey and McFadden (1994) imply

$$\widehat{\Theta}_T^{(p)}-\Theta_0\xrightarrow{\text{a.s.}}0,\qquad T\to+\infty.$$

In order to show the AN, notice that function $S_T^{(r)}(\Theta)$ has the continuous partial derivatives up to the second order, for any component of the vector Θ. Thus, the Taylor expansion of $\partial S_T^{(r)}(\Theta)/\partial\Theta$ at $\Theta=\Theta_0$ gives

$$\frac{\partial S_T^{(r)}(\Theta)}{\partial\Theta}=\frac{\partial S_T^{(r)}(\Theta_0)}{\partial\Theta}+\frac{\partial^2 S_T^{(r)}(\Theta_0)}{\partial\Theta\,\partial\Theta'}\cdot(\Theta-\Theta_0)+o(\Theta-\Theta_0).$$

Substituting for T large enough Θ with $\widehat{\Theta}_T^{(p)}$, under assumption (iii) and the fact that $\partial S_T^{(r)}(\widehat{\Theta}_T^{(r)})/\partial\Theta=0$, we have

$$\widehat{\Theta}_T^{(r)}-\Theta_0=-\left[\frac{\partial^2 S_T^{(r)}(\Theta_0)}{\partial\Theta\,\partial\Theta'}\right]^{-1}\frac{\partial S_T^{(r)}(\Theta_0)}{\partial\Theta}+o\big(\widehat{\Theta}_T^{(r)}-\Theta_0\big).$$

According to the above-mentioned properties, the function $S_T^{(r)}(\Theta)$ can be differentiated under the integral sign, i.e.

$$\frac{\partial S_T^{(r)}(\Theta)}{\partial\Theta}=2\int_{-1}^{1}\cdots\int_{-1}^{1}g(\mathbf{u})\big[\Psi_X^{(r)}(\mathbf{u};\Theta)-\widetilde{\Psi}_T(\mathbf{u})\big]\frac{\partial\Psi_X^{(p)}(\mathbf{u};\Theta)}{\partial\Theta}\,d\mathbf{u}\tag{A.6}$$

and

$$\begin{aligned}\frac{\partial^2 S_T^{(r)}(\Theta)}{\partial\Theta\,\partial\Theta'}=2\int_{-1}^{1}\cdots\int_{-1}^{1}g(\mathbf{u})\bigg\{&\frac{\partial\Psi_X^{(r)}(\mathbf{u};\Theta)}{\partial\Theta}\frac{\partial\Psi_X^{(r)}(\mathbf{u};\Theta)}{\partial\Theta'}\\&+\big[\Psi_X^{(r)}(\mathbf{u};\Theta)-\widetilde{\Psi}_T(\mathbf{u})\big]\frac{\partial^2\Psi_X^{(p)}(\mathbf{u};\Theta)}{\partial\Theta\,\partial\Theta'}\bigg\}\,d\mathbf{u}.\end{aligned}\tag{A.7}$$

Now, equations (A.6)–(A.7) give

$$E\left[\frac{\partial S_T^{(r)}(\Theta_0)}{\partial\Theta}\right]=0,\qquad E\left[\frac{\partial^2 S_T^{(r)}(\Theta_0)}{\partial\Theta\,\partial\Theta'}\right]=2\mathbf{V},\tag{A.8}$$

where

$$\mathbf{V}=\int_{-1}^{1}\cdots\int_{-1}^{1}g(\mathbf{u})\frac{\partial\Psi_X^{(r)}(\mathbf{u};\Theta_0)}{\partial\Theta}\frac{\partial\Psi_X^{(r)}(\mathbf{u};\Theta_0)}{\partial\Theta'}\,d\mathbf{u}.$$

As the function $h(\mathbf{u}) := \partial\Psi_X^{(r)}(\mathbf{u};\Theta_0)/\partial\Theta \times \partial\Psi_X^{(r)}(\mathbf{u};\Theta_0)/\partial\Theta'$ is continuous on the compact $[-1,1]^r$, it follows

$$\left|h(\mathbf{u})\right| \le M < +\infty, \qquad \mathbf{u} \in [-1,1]^r$$

for some $M > 0$. Hence, the inequalities

$$0 < \|\mathbf{V}\| \le M \int_{-1}^{1} \cdots \int_{-1}^{1} g(\mathbf{u})\, \mathrm{d}\mathbf{u} < +\infty$$

hold, and consequently

$$\left(\frac{\partial S_T^{(r)}(\Theta_0)}{\partial\Theta}, \frac{\partial^2 S_T^{(r)}(\Theta_0)}{\partial\Theta\,\partial\Theta'} \right) \xrightarrow{\text{a.s.}} (0, 2\mathbf{V}), \qquad T \to +\infty. \tag{A.9}$$

Further on, we shall write the gradient of $S_T^{(r)}(\Theta)$ as

$$\frac{\partial S_T^{(r)}(\Theta)}{\partial\Theta} = \frac{2}{T-r+1} \sum_{t=1}^{T-1} C_t(\Theta),$$

where

$$C_t(\Theta) = \int_{-1}^{1} \cdots \int_{-1}^{1} g(\mathbf{u}) \big[\Psi_X^{(r)}(\mathbf{u};\Theta) - \widetilde{\Psi}_T(\mathbf{u})\big] \frac{\partial\Psi_X^{(r)}(\mathbf{u};\Theta)}{\partial\Theta}\, \mathrm{d}\mathbf{u}.$$

It can be shown (see, for instance, Yu (2004)), that the finite nonzero limit

$$\begin{aligned} \mathbf{W}^2 &:= \lim_{T\to\infty} \frac{1}{(T-r+1)^2} \mathrm{Var}\left[\sum_{t=1}^{T-r+1} C_t(\Theta_0)\right] \\ &= \lim_{T\to\infty} \frac{1}{(T-r+1)^2} \sum_{t=1}^{T-r+1} \sum_{s=1}^{T-r+1} \mathrm{Cov}\big[C_t(\Theta_0) C_s(\Theta_0)\big] \end{aligned}$$

exists if the series $\gamma_X(k) := \mathrm{Cov}(X_t, X_{t+k})$, $k = 0, \pm1, \pm2, \ldots$ has the finite, non-zero sum. In the case of our, NIINAR(1) model, we have

$$C := \sum_{k=-\infty}^{+\infty} \gamma_X(k) = \gamma_X(0) \left(2 \sum_{k=0}^{+\infty} \alpha^k - 1 \right) = \sigma_X^2 \frac{1+\alpha}{1-\alpha},$$

and the inequalities $0 < C < +\infty$ hold for each $\Theta \in K$. Now, if we applying the central limit theorem for stationary processes, we obtain

$$\sqrt{T-r+1} \frac{\partial S_T^{(r)}(\Theta_0)}{\partial\Theta} \xrightarrow{d} \mathcal{N}(0, 4\mathbf{W}^2), \qquad T \to +\infty.$$

This convergence and the equations (A.8)–(A.9) imply

$$\sqrt{T-r+1} (\widehat{\Theta}_T^{(p)} - \Theta_0) \xrightarrow{d} \mathcal{N}(0, \mathbf{V}^{-1}\mathbf{W}^2\mathbf{V}^{-1}), \qquad T \to +\infty,$$

and the theorem is completely proven. □

Acknowledgements

The authors are grateful to Ministry of Interior of the Republic of Serbia for consigning of datasets and for giving the constructive information regarding them. Also, authors would like to thank anonymous referee for a valuable suggestions and comments that have allowed to improve the paper. Finally, we thank to academician Gradimir V. Milovanović for his generous help and advices about the cubature formulas which were used in this paper.

This paper is part of the research project no. III 44007 and subproject 3 in project no. III 47016, supported by the Ministry of Education, Science and Technological development of the Republic of Serbia.

References

Al-Osh, M. A. and Alzaid, A. A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis* **8**, 261–275. MR0903755

Alzaid, A. A. and Al-Osh, M. A. (1990). An integer-valued pth-order autoregressive structure (INAR(p)) process. *Journal of Applied Probability* **27**, 314–324. MR1052303

Bakouch, H. S. and Ristić, M. M. (2010). Zero truncated Poisson integer valued AR(1) model. *Metrika* **72**, 265–280. MR2725101

Bourguignon, M. and Vasconcellos, K. L. P. (2015). First order non-negative integer valued autoregressive processes with power series innovations. *Brazilian Journal of Probability and Statistics* **29**, 71–93. MR3299108

Cvetković, A. S. and Milovanović, G. V. (2004). *The Mathematica Package "Orthogonal Polynomials"*. *Facta Univesritatis—Series Mathematics and Informatics* **19**, 17–36. MR2122752

Drost, F. C., Akken, R. V. D. and Werker, B. J. M. (2008). Local asymptotic normality and efficient estimation for INAR(p) models. *Journal of Time Series Analysis* **29**, 783–801. MR2450896

Hyndman, R. (2016). Forecasting functions for time series and linear models. R package version 7.1. Available at http://CRAN.R-project.org/package=forecast.

Jazi, M. A., Jones, G. and Lai, C.-D. (2012). First-order integer valued AR processes with zero inflated Poisson innovations. *Journal of Time Series Analysis* **33**, 954–963. MR2991911

Knight, J. L. and Yu, J. (2002). Empirical characteristic function in time series estimation. *Econometric Theory* **18**, 691–721. MR1906331

Li, C., Wang, D. and Zhang, H. (2015). First-order mixed integer-valued autoregressive processes with zero-inflated generalized power series innovations. *Journal of the Korean Statistical Society* **44**, 232–246. MR3342635

Martin, V. L., Tremayne, A. R. and Jung, R. C. (2014). Efficient method of moments estimators for integer time series models. *Journal of Time Series Analysis* **35**, 491–516. MR3300684

Nastić, A. S. (2014). On suitability of negative binomial marginals and geometric counting sequence in some applications of combined INAR(p) model. *Facta Univesritatis—Series Mathematics and Informatics* **29**, 25–42. MR3213327

Nastić, A. S., Ristić, M. M. and Bakouch, H. S. (2012). A combined geometric INAR(p) model based on negative binomial thinning. *Mathematical and Computer Modelling* **55**, 1665–1672. MR2899113

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. In *Handbook of Econometrics, Vol. 4*. Amsterdam: North-Holland. MR1315971

Noack, A. (1950). A class of random variables with discrete distributions. *The Annals of Mathematical Statistics* **21**, 127–132. MR0033472

Pavlopulos, H. and Dimitris, K. (2008). INAR(1) modeling of overdispersed count series with an environmental application. *Environmetrics* **19**, 369–393. MR2440038

Ristić, M. M., Bakouch, H. S. and Nastić, A. S. (2009). A new geometric first-order integer valued autoregressive (NGINAR(1)) process. *Journal of Statistical Planning and Inference* **139**, 218–226. MR2507983

Ristić, M. M., Nastić, A. S. and Bakouch, H. S. (2012). Estimation in an integer-valued autoregressive process with negative binomial marginals (NBINAR(1)). *Communications in Statistics—Theory and Methods* **41**, 606–618. MR2874378

Ristić, M. M., Nastić, A. S. and Miletić, A. V. (2013). A geometric time series model with dependent Bernoulli counting series. *Journal of Time Series Analysis* **34**, 466–476. MR3070869

Ristić, M. M., Nastić, A. S. and Djordjević, M. S. (2016). An INAR model with discrete Laplace marginal distributions. *Brazilian Journal of Probability and Statistics* **30**, 107–126. MR3453517

Schweer, S. and Weiß, C. H. (2014). Compound Poisson INAR(1) processes: Stochastic properties and testing for over-dispersion. *Computational Statistics and Data Analysis* **77**, 267–284. MR3210062

Serfling, R. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley. MR0595165

Silva, I. and Silva, M. E. (2006). Asymptotic distribution of the Yule-Walker estimator for INAR(p) processes. *Statistics and Probability Letters* **76**, 1655–1663. MR2248854

Silva, I. and Silva, M. E. (2009). Parameter estimation for INAR processes based on high-order statistics. *REVSTAT—Statistical Journal* **7**, 105–117. MR2501373

Silva, M. E. and Oliveira, V. L. (2004). Difference equations for the higher-order moments and cumulants of the INAR(1) model. *Journal of Time Series Analysis* **25**, 317–333. MR2062676

Stojanović, V., Milovanović, G. V. and Jelić, G. (2016). Distributional properties and parameters estimation of GSB process: An approach based on characteristic functions. *Latin American Journal of Probability and Mathematical Statistics* **13**, 835–861. MR3546383

Stojanović, V. S., Popović, B. Č. and Popović, P. M. (2011). The Split-BREAK model. *Brazilian Journal of Probability and Statistics* **25**, 44–63. MR2746492

Stojanović, V. S., Popović, B. Č. and Popović, P. M. (2014). *Stochastic Analysis of GSB Process*. *Publications de l'Institut Mathématique* **95**, 109. MR3221223

Stojanović, V. S., Popović, B. Č. and Popović, P. M. (2015). Model of general Split-BREAK process. *REVSTAT—Statistical Journal* **13**, 145–168. MR3366965

Stojanović, V. S., Popović, B. Č. and Milovanović, G. V. (2016). The Split-SV model. *Computational Statistics and Data Analysis* **100**, 560–581. MR3505818

Weiß, C. H. (2008). Thinning operations for modeling time series of counts—A survey. *Advances in Statistical Analysis* **92**, 319–341. MR2426093

Weiß, C. H. (2009). A new class of autoregressive models for time series of binomial counts. *Communications in Statistics—Theory and Methods* **38**, 447–460. MR2510797

Weiß, C. H. and Pollet, P. K. (2014). Binomial autoregressive processes with density-dependent thinning. *Journal of Time Series Analysis* **35**, 115–132. MR3166349

Yu, J. (2004). Empirical characteristic function estimation and its applications. *Econometric Review* **23**, 93–123. MR2075094

V. Stojanović
Faculty of Sciences and Mathematics
Kosovska Mitrovica
Serbia
E-mail: vladica.stojanovic@pr.ac.rs

D. Randjelović
K. Kuk
Academy of Criminalistics and Police Studies
Belgrade
Serbia
E-mail: dragan.randjelovic@kpa.edu.rs
kristijan.kuk@kpa.edu.rs