axioms

Article

Parameters Estimation in Non-Negative Integer-Valued Time Series: Approach Based on Probability Generating Functions

Vladica Stojanović [1,*,†], Eugen Ljajko [2,†] and Marina Tošić [2,†]

[1] Department of Informatics & Computer Sciences, University of Criminal Investigation and Police Studies, 11000 Belgrade, Serbia
[2] Department of Mathematics, Faculty of Sciences & Mathematics, University of Priština in Kosovska Mitrovica, 38220 Kosovska Mitrovica, Serbia
* Correspondence: vladica.stojanovic@kpu.edu.rs
† These authors contributed equally to this work.

Abstract: This manuscript deals with a parameter estimation of a non-negative integer-valued (NNIV) time series based on the so-called probability generating function (PGF) method. The theoretical background of the PGF estimation technique for a very general, stationary class of NNIV time series is described, as well as the asymptotic properties of the obtained estimates. After that, a particular emphasis is given to PGF estimators of independent identical distributed (IID) and integer-valued non-negative autoregressive (INAR) series. A Monte Carlo study of the thus obtained PGF estimates, based on a numerical integration of the appropriate objective function, is also presented. For this purpose, numerical quadrature formulas were computed using Gegenbauer orthogonal polynomials. Finally, the application of the PGF estimators in the dynamic analysis of some actual data is given.

Keywords: integer-valued time series; parameter estimation; probability generating functions; asymptotic properties; simulation; numerical integration; application

MSC: 62M10; 65D30

Citation: Stojanović, V.; Ljajko, E.; Tošić, M. Parameters Estimation in Non-Negative Integer-Valued Time Series: Approach Based on Probability Generating Functions. *Axioms* **2023**, *12*, 112. https://doi.org/10.3390/axioms12020112

Academic Editor: Paolo Pagnottoni

Received: 19 December 2022
Revised: 14 January 2023
Accepted: 18 January 2023
Published: 21 January 2023

1. Introduction

Non-negative integer-valued (NNIV) time series represent one of the most important stochastic models that have attracted the attention of numerous studies (for some more recent ones, see, e.g., [1–7]). Due to the specific structure of an NNIV series, the procedure for estimating their parameters may differ from that used for similar non-integer stochastic models. Thus, in the case of the most famous NNIV series, the so-called integer non-negative autoregressive (INAR) processes, various estimation techniques have been developed. For instance, in [8], conditional least square estimators were described, while in [9], estimators based on the conditional maximum likelihood method were discussed. On the other hand, the authors in [10,11] considered, among others, moment-based estimation techniques, i.e., the well-known Yule–Walker (YW) equations. Finally, some generalizations of these methods based on the high-order statistics can be found in [12,13].

However, in order to apply any of the mentioned methods, the estimation functions must usually be given in a finite, closed form and (most often) bounded within a certain parameter space. Unfortunately, sometimes, such conditions cannot be met, both in the case of independently identically distributed (IID) and stochastically dependent time series. Therefore, an estimation technique based on *probability generating functions (PGFs)* can be a useful alternative approach, since PGFs exist and are bounded for any NNIV time series, at least on the interval $[-1, 1]$. Another advantage of PGF estimators is that the estimation procedure is close to the *empirical characteristic function (ECF)* method introduced in [14,15] and developed later in many studies, cf. [16–18]. Generally speaking, the idea behind the PGF method is based on the fact that PGF contains the same amount of information about the distribution of some random variable (RV) as the appropriate characteristic function (CF). Thus, it is expected that the minimization of 'the distance' between the PGF of some time series (X_t) and its empirical PGF gives an efficient parameter estimator. In this way, the PGF method, similar to the ECF method for non-integer time series, could be applied to various kinds of stationary NNIV series. For these reasons, our main goal here is to obtain efficient PGF parameter estimates for the case of a rather general, stationary class of NNIV time series, with applications that will also be described in the following.

Let us point out that, to the best of our knowledge, there are only few studies where the PGF estimation procedure has been discussed so far. First, Esquìvel [19] has shown the uniform laws of large numbers and the functional limit theorem for the empirical PGFs. Thereafter, Stojanović et al. [20] used a more general PGF estimation procedure, based on numerical integration techniques, in order to obtain parameter estimators of the so-called noise-indicator integer non-negative autoregressive (NIINAR) time series model. Our main motive here is to generalize these results, i.e., to give theoretical foundations of the PGF estimation procedure in the most general form. According to that, as it will be seen, the PGF estimates can be practically implemented in the estimation of some of the most important and often used NNIV times series, as well as in modelling the dynamics of an actual data set of the number of COVID-19 death cases.

For this purpose, in Section 2, we give some basic facts that are necessary to formulate the PGF estimation method. Additionally, under some regulatory conditions, we investigate the asymptotic properties and efficiency of the PGF estimators for an arbitrary stationary NNIV time series. After that, in Section 3, we analyze some specific PGF estimates for cases of independent identically distributed (IID) and integer-valued autoregressive (INAR) time series. For both types of time series, Monte Carlo simulations of the PGF estimates are computed and their asymptotic properties, such as efficiency and asymptotic normality, are examined too. Section 4 presents an application of the PGF method in parameters and distribution estimation of death cases in the period before and after the appearance of the SARS-CoV2 virus infection on the territory of the Republic of Serbia. Finally, Section 5 contains some concluding remarks.

2. PGF Estimates: Definition and Asymptotic Properties

Let (X_t), $t \in \mathbb{Z}$ be a non-negative, integer-valued, and stationary time series with finite moments and probability mass distribution (PMF)

$$p_X(x;\theta) := P\{X_t = x\}, \quad x \in \mathbb{Z}^+ = \{0,1,2,\dots\},$$

which depends on a vector of unknown parameters $\theta = (\theta_1, \dots, \theta_r)' \in \mathbb{R}^r$, $r \geq 1$. Additionally, let $\{X_1, X_2, \dots, X_T\}$ be some finite realization of the series (X_t), of the length T. In order to obtain the PGF estimates of θ, we introduce the following terms.

Let $\mathbf{u} = (u_1, \dots, u_r)' \in \mathbb{R}^r$ and $\mathbf{X}_t^{(r)} := (X_t, \dots, X_{t+r-1})'$, $t \in \mathbb{Z}$ be *the overlapping blocks* of the process (X_t), and

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

be *the r-dimensional PGF* of the random vector $\mathbf{X}_t^{(r)}$. Due to the stationarity of series (X_t), it is obvious that the function in Equation (1) does not depend on the choice of overlapping blocks, i.e., moment of time t. The main aim of the PGF method is to minimize 'the distance' between the theoretical PGF, given by Equation (1), and the corresponding *Empirical PGF* (of order $r \geq 1$):

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

As the series (X_t) is stationary, it follows that $\widetilde{\Psi}_T^{(r)}(\mathbf{u})$ is an unbiased estimator of $\Psi_X^{(r)}(\mathbf{u};\theta_0)$, i.e.,

$$\mathrm{E}\left[\widetilde{\Psi}_T^{(r)}(\mathbf{u})\right] = \Psi_X^{(r)}(\mathbf{u};\theta_0), \tag{2}$$

where θ_0 is the true (but unknown) value of parameter θ. It is also well-known that theoretical PGF $\Psi_X^{(r)}(\mathbf{u};\theta)$ is well defined at least for all values $\mathbf{u} \in [-1,1]^r$ (see, e.g., [20]). Thus, the objective function can be defined as

$$S_T^{(r)}(\theta) := \int_{-1}^{1} \cdots \int_{-1}^{1} \omega(\mathbf{u}) \Big(\Psi_X^{(r)}(\mathbf{u};\theta) - \widetilde{\Psi}_T^{(r)}(\mathbf{u}) \Big)^2 \mathrm{d}\mathbf{u}, \tag{3}$$

where $\mathrm{d}\mathbf{u} := \mathrm{d}u_1 \cdots \mathrm{d}u_r$ and $\omega : \mathbb{R}^r \to \mathbb{R}^+$ is a weight function, integrable on $[-1,1]^r$. The PGF estimates will be obtained by minimizing the objective function in Equation (3) with respect to θ, i.e., they are solutions of the minimization equation

$$\widehat{\theta}_T^{(r)} = \arg\min_{\theta\in\Theta} S_T^{(r)}(\theta), \tag{4}$$

where $\Theta \subseteq \mathbf{R}^r$ is a parameter space. In order to solve this equation, we use numerical methods, i.e., procedures based on numerical integration, which is discussed in Section 3. In the sequel, we will examine strong consistency and asymptotic normality (AN) of the PGF estimators under some regularity conditions.

Theorem 1. *Let θ_0 be the exact value of the parameter θ, and let $\widehat{\theta}_T^{(r)}$, for an arbitrary $T = 1,2,\ldots$, be the solutions of Equation (4). Additionally, let us suppose that the following regularity conditions are fulfilled:*

(A_1) *Θ is a compact set, where $\theta_0 \in \mathrm{int}(\Theta)$, and $\widehat{\theta}_T^{(r)} \in \mathrm{int}(\Theta)$ for T large enough.*

(A_2) *At the point $\theta = \theta_0$ function*

$$S_0^{(r)}(\theta) := \int_{-1}^{1} \cdots \int_{-1}^{1} \omega(\mathbf{u}) \Big(\Psi_X^{(r)}(\mathbf{u};\theta) - \Psi_X^{(r)}(\mathbf{u};\theta_0) \Big)^2 \mathrm{d}\mathbf{u}$$

has a unique minimum $S_0^{(r)}(\theta_0) = 0$.

(A_3) $\dfrac{\partial^2 S_T^{(r)}(\theta_0)}{\partial\theta\,\partial\theta'}$ *is a regular matrix.*

(A_4) $\dfrac{\partial\Psi_X^{(r)}(\mathbf{u};\theta_0)}{\partial\theta}\,\dfrac{\partial\Psi_X^{(r)}(\mathbf{u};\theta_0)}{\partial\theta'}$ *is a non-zero matrix, uniformly bounded by the strictly positive ω-integrable function $\mathcal{L} : \mathbb{R}^r \to \mathbb{R}^+$.*

(A_5) *The covariance function $\gamma_K(h;\theta) := \mathrm{Cov}\Big[K_t^{(r)}(\theta),\, K_{t+h}^{(r)}(\theta)\Big]$ of the series*

$$K_t^{(r)}(\theta) := \int_{-1}^{1} \cdots \int_{-1}^{1} \omega(\mathbf{u}) \left[\Psi_X^{(r)}(\mathbf{u};\theta) - u_1^{X_t} \cdots u_r^{X_{t+r-1}} \right] \frac{\partial\Psi_X^{(r)}(\mathbf{u};\theta)}{\partial\theta}\, \mathrm{d}\mathbf{u}. \tag{5}$$

is absolutely summable when $\theta = \theta_0$, i.e.,

$$\sum_{h=-\infty}^{\infty} |\gamma_K(h;\theta_0)| < +\infty.$$

At the same time, the integral in Equation (5) is non-singular and exists for each $\theta \in \Theta$.

Then, $\widehat{\theta}_T^{(r)}$ is a strictly consistent and AN estimator for θ.

Proof. Firstly, we check the sufficient consistency conditions of the so-called extremum estimators (for more details, see Newey and McFadden [21]). According to given assumptions, the series (X_t) is ergodic, so that Equation (2) and the strong law of large numbers (SLLN) give

$$\widetilde{\Psi}_T^{(r)}(\mathbf{u}) \xrightarrow{\text{as}} \Psi_X^{(r)}(\mathbf{u};\theta_0), \quad T \to \infty, \tag{6}$$

where "as" denotes the almost sure convergence. Furthermore, according to assumption (A_1), the function $\Psi_X^{(r)}(\mathbf{u};\theta)$ is continuous on the compact $[-1,1]^r \times \Theta$, and $\widetilde{\Psi}_T^{(r)}(\mathbf{u})$ is continuous on the compact Θ. Hence, there exist constants $C_1, C_2 > 0$ such that the inequalities

$$\max_{(\mathbf{u};\theta)\in[-1,1]^r\times\Theta}\left|\Psi_X^{(r)}(\mathbf{u};\theta)\right| \le C_1 < +\infty, \qquad \max_{\mathbf{u}\in[-1,1]^r}\left|\widetilde{\Psi}_T^{(r)}(\mathbf{u})\right| \le C_2 < +\infty$$

hold. According to these, one obtains

$$\begin{aligned}
\left|S_T^{(r)}(\theta) - S_0^{(r)}(\theta)\right| &= \bigg|\int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\Big(\Psi_X^{(r)}(\mathbf{u};\theta) - \widetilde{\Psi}_T^{(r)}(\mathbf{u})\Big)^2 \mathrm{d}\mathbf{u} \\
&\qquad - \int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\Big(\Psi_X^{(r)}(\mathbf{u};\theta) - \Psi_X^{(r)}(\mathbf{u};\theta_0)\Big)^2 \mathrm{d}\mathbf{u}\bigg| \\
&= \bigg|\int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\Big[\Big(\widetilde{\Psi}_T^{(r)}(\mathbf{u})\Big)^2 - \Big(\Psi_X^{(r)}(\mathbf{u};\theta_0)\Big)^2\Big]\mathrm{d}\mathbf{u} \\
&\qquad + 2\int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\Psi_X^{(r)}(\mathbf{u};\theta)\Big(\Psi_X^{(r)}(\mathbf{u};\theta_0) - \widetilde{\Psi}_T^{(r)}(\mathbf{u})\Big)\mathrm{d}\mathbf{u}\bigg| \\
&\le (C_1 + C_2)\int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\left|\widetilde{\Psi}_T^{(r)}(\mathbf{u}) - \Psi_X^{(r)}(\mathbf{u};\theta_0)\right|\mathrm{d}\mathbf{u} \\
&\qquad + 2\,C_1\int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\left|\Psi_X^{(r)}(\mathbf{u};\theta_0) - \widetilde{\Psi}_T^{(r)}(\mathbf{u})\right|\mathrm{d}\mathbf{u} \\
&\le (3C_1 + C_2)\int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\left|\widetilde{\Psi}_T^{(r)}(\mathbf{u}) - \Psi_X^{(r)}(\mathbf{u};\theta_0)\right|\mathrm{d}\mathbf{u},
\end{aligned}$$

so the last inequality and Equation (6) imply

$$\sup_{\theta\in\Theta}\left|S_T^{(r)}(\theta) - S_0^{(r)}(\theta_0)\right| \xrightarrow{\text{as}} 0, \quad T \to +\infty.$$

Thus, $S_T^{(r)}(\theta)$ uniformly converges almost surely toward $S_0^{(r)}(\theta)$. According to this, as well as the assumption (A_2) and Theorem 1 in Newey and McFadden [21], it follows

$$\widehat{\theta}_T^{(r)} - \theta_0 \xrightarrow{\text{as}} 0, \quad T \to +\infty,$$

i.e., $\widehat{\theta}_T^{(r)}$ is strictly consistent estimator for θ.

Furthermore, let us notice that function $S_T^{(r)}(\theta)$ has continuous partial derivatives of the first and second orders, with respect to any component of the vector θ. Hence, 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 θ with $\widehat{\theta}_T^{(r)}$, for a T large enough, under assumption (A_3) and the fact that $\partial S_T^{(r)}(\widehat{\theta}_T^{(r)})/\partial\theta = 0$, we have

$$\widehat{\theta}_T^{(r)} - \theta_0 = -\left[\frac{\partial^2 S_T^{(r)}(\theta_0)}{\partial\theta\,\partial\theta'}\right]^{-1}\frac{\partial S_T^{(r)}(\theta_0)}{\partial\theta} + o\big(\widehat{\theta}_T^{(r)} - \theta_0\big). \tag{7}$$

Furthermore, the function $S_T^{(r)}(\theta)$ can be differentiated under the integral sign, i.e.,

$$\frac{\partial S_T^{(r)}(\theta)}{\partial\theta} = 2\int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\left[\Psi_X^{(r)}(\mathbf{u};\theta)-\widetilde{\Psi}_T(\mathbf{u})\right]\frac{\partial\Psi_X^{(r)}(\mathbf{u};\theta)}{\partial\theta}\mathrm{d}\mathbf{u}, \tag{8}$$

$$\frac{\partial^2 S_T^{(r)}(\theta)}{\partial\theta\,\partial\theta'} = 2\int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\left\{\frac{\partial\Psi_X^{(r)}(\mathbf{u};\theta)}{\partial\theta}\frac{\partial\Psi_X^{(r)}(\mathbf{u};\theta)}{\partial\theta'} + \left[\Psi_X^{(r)}(\mathbf{u};\theta)-\widetilde{\Psi}_T(\mathbf{u})\right]\frac{\partial^2\Psi_X^{(r)}(\mathbf{u};\theta)}{\partial\theta\,\partial\theta'}\right\}\mathrm{d}\mathbf{u}. \tag{9}$$

Now, Equations (8) and (9) give

$$E\left[\frac{\partial S_T^{(r)}(\theta_0)}{\partial\theta}\right] = \mathbf{0}_{r\times 1}, \qquad E\left[\frac{\partial^2 S_T^{(r)}(\theta_0)}{\partial\theta\,\partial\theta'}\right] = 2\mathbf{V}, \tag{10}$$

where

$$\mathbf{V} = \int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\frac{\partial\Psi_X^{(r)}(\mathbf{u};\theta_0)}{\partial\theta}\frac{\partial\Psi_X^{(r)}(\mathbf{u};\theta_0)}{\partial\theta'}\mathrm{d}\mathbf{u}.$$

According to assumption (A_4), there exists an ω-integrable function $\mathcal{L}:\mathbb{R}^r\to\mathbb{R}^+$ such that

$$0 < \left\|\frac{\partial\Psi_X^{(r)}(\mathbf{u};\theta_0)}{\partial\theta}\frac{\partial\Psi_X^{(r)}(\mathbf{u};\theta_0)}{\partial\theta'}\right\| \le \mathcal{L}(\mathbf{u}), \quad \mathbf{u}\in[-1,1]^r,$$

where $\|\cdot\|$ is a matrix norm on $\mathbb{R}^r\times\mathbb{R}^r$. Hence, the inequalities

$$0 < \|\mathbf{V}\| \le \int_{-1}^{1}\cdots\int_{-1}^{1}\omega(\mathbf{u})\mathcal{L}(\mathbf{u})\mathrm{d}\mathbf{u} < +\infty$$

hold, so Equation (10) and the SLLN give

$$\left(\frac{\partial S_T^{(r)}(\theta_0)}{\partial\theta}, \frac{\partial^2 S_T^{(r)}(\theta_0)}{\partial\theta\,\partial\theta'}\right) \xrightarrow{as} (0, 2\mathbf{V}), \quad T\to+\infty. \tag{11}$$

To prove the AN property, note that using Equations (5) and (8), we can write the gradient of $S_T^{(r)}(\theta)$ as

$$\frac{\partial S_T^{(r)}(\theta)}{\partial\theta} = \frac{2}{T-r+1}\sum_{t=1}^{T-r+1}K_t^{(r)}(\theta). \tag{12}$$

Therefore, according to assumption (A_5), equality $\mathrm{E}[K_t(\theta_0)] = 0$ holds and it follows

$$\begin{aligned}
\mathbf{W}^2 &:= \lim_{T\to\infty}\mathrm{Var}\left[\frac{1}{\sqrt{T-r+1}}\sum_{t=1}^{T-r+1}K_t^{(r)}(\theta_0)\right] = \lim_{T\to\infty}\frac{1}{T-r+1}\mathrm{E}\left[\sum_{t=1}^{T-r+1}K_t^{(r)}(\theta_0)\right]^2 \\
&= \lim_{T\to\infty}\frac{1}{T-r+1}\sum_{t=1}^{T-r+1}\sum_{s=1}^{T-r+1}\mathrm{E}\left[K_t^{(r)}(\theta_0)\,K_s^{(r)}(\theta_0)\right] \\
&= \lim_{T\to\infty}\frac{1}{T-r+1}\sum_{h=-T+r-1}^{T-r+1}(T-r+1-h)\gamma_K(h;\theta_0) \\
&= \lim_{T\to\infty}\sum_{h=-T}^{T}\left(1-\frac{h}{T}\right)\gamma_K(h;\theta_0) = \sum_{h=-\infty}^{\infty}\gamma_K(h;\theta_0) < +\infty.
\end{aligned}$$

Note that the last equality follows from Kronecker's lemma (see, c.f. Fuller [22]). Thus, the proven convergence and Equation (12) imply

$$\sqrt{T-r+1}\,\frac{\partial S_T^{(r)}(\theta_0)}{\partial\theta} \xrightarrow{d} \mathcal{N}(\mathbf{0}_{r\times 1}, 4\mathbf{W}^2), \quad T\to+\infty, \tag{13}$$

where "d" denotes the convergence in distribution. Now, applying the central limit theorem for stationary processes, as well as Equations (7), (11) and (13), one obtains

$$\sqrt{T-r+1}\left(\widehat{\theta}_T^{(r)}-\theta_0\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}_{r\times 1}, \mathbf{V}^{-1}\mathbf{W}^2\mathbf{V}^{-1}), \quad T\to+\infty,$$

and the statement of theorem is completely proven. □

At the end of this section, we point out that PGF estimators of the r-dimensional parameter θ are usually obtained using overlapping blocks of the same dimensions. This is also consistent with some of the basic assumptions of the ECF method (for more details, see [14]). For these reasons, below, we examine in detail the PGF estimation procedures for the parameter order (as well as the corresponding PGFs order) $r=1$ and $r=2$.

3. PGF Estimations of IID and INAR Time Series

In this section, we discuss the construction of PGF estimates for two important classes of stationary NNIV time series. First, using a very general form of the so-called *power series (PS) distributions*, the estimation procedure of independent identical distributed (IID) time series was examined. After that, using the obtained results, PGF estimators of integer-valued autoregressive (INAR) time series are considered.

3.1. Estimation of IID Time Series

The simplest case of the stationary NNIV time series are the IID series, which we denote as (ε_t), $t\in\mathbb{Z}$. Similar to that in Stojanović et al. [20], and Bourguignon and Vasconcellos [23], we say that the series (ε_t) has a *PS distribution* if its PMF is given as follows:

$$p_\varepsilon(x;\theta) := P\{\varepsilon_t = x\} = \frac{a(x)\,\theta^x}{f(\theta)}, \quad x\in\mathcal{S}. \tag{14}$$

Here, $\mathcal{S}\subseteq\mathbb{Z}^+=\{0,1,2,\dots\}$ is the so-called discrete support of the RV ε_t and

(i) $a(x)\geq 0$ is a function that depends (only) on $x\in\mathcal{S}$;

(ii) $\theta>0$ is a (one-dimensional) unknown parameter;

(iii) $f(\theta) := \sum_{x\in\mathcal{S}} a(x)\,\theta^x$ is a function on θ, so that $0<f(\theta)<+\infty$ when $\theta\in(0,R)$, $R>0$.

The expression given by Equation (14) represents the so-called *PS distributions* that, for some special choices of $a(x)$, $f(\theta)$ and θ, allow for obtaining the most of the known distributions (see Table 1, below). In addition, according to assumption (iii), it is obvious that, in fact, the power series $f(\theta)$ converges on the interval $(-R,R)$. However, a common assumption is $\theta>0$, so we say that power series $f(\theta)$ converges on $(0,R)$. Accordingly, the function $f(\theta)$ is positive and increasing at this interval. Moreover, for an arbitrary $u\in[-1,1]$, we denote by $\Psi_\varepsilon^{(1)}(u;\theta)=\Psi_\varepsilon(u;\theta)$ the PGF (of the first order) of RVs (ε_t). Thus, according to Equation (14), one obtains

$$\Psi_\varepsilon(u;\theta) := \mathrm{E}[u^{\varepsilon_t}] = \frac{1}{f(\theta)}\sum_{x\in\mathcal{S}} a(x)\,(\theta u)^x = \frac{f(\theta u)}{f(\theta)}, \tag{15}$$

where the sum above converges when $u<R/\theta$. Note that Equation (15) allows for a simple computation of first-order PGFs for some of the well-known NNIV distributions, which are also given in Table 1.

Table 1. NNIV distributions, along with their PGFs, interpreted in the form of PS distributions.

Distributions	$\mathcal{S}$	$a(x)$	$(0, R)$	$f(\theta)$	$\Psi_\varepsilon(u;\theta)$	R/θ
1. Bernoulli	$\{0,1\}$	1	$(0,\infty)$	$1+\theta$	$\dfrac{1+\theta u}{1+\theta}$	∞
2. Binomial	$\{0,\dots,n\}$	$\binom{n}{x}$	$(0,\infty)$	$(1+\theta)^n$	$\left(\dfrac{1+\theta u}{1+\theta}\right)^n$	∞
3. Poisson	$\{0,\dots,\infty\}$	$\dfrac{1}{x!}$	$(0,\infty)$	$\exp(\theta)$	$\exp(\theta(u-1))$	∞
4. Geometric	$\{0,\dots,\infty\}$	1	$(0,1)$	$\dfrac{1}{1-\theta}$	$\dfrac{1-\theta}{1-\theta u}$	$1/\theta$
5. Neg. Binomial	$\{0,\dots,\infty\}$	$\dfrac{\Gamma(x+n)}{x!\,\Gamma(n)}$	$(0,1)$	$\dfrac{1}{(1-\theta)^n}$	$\left(\dfrac{1-\theta}{1-\theta u}\right)^n$	$1/\theta$
6. Pascal	$\{n,\dots,\infty\}$	$\binom{x-1}{n-1}$	$(0,1)$	$\left(\dfrac{\theta}{1-\theta}\right)^n$	$\left(\dfrac{(1-\theta)u}{1-\theta u}\right)^n$	$1/\theta$

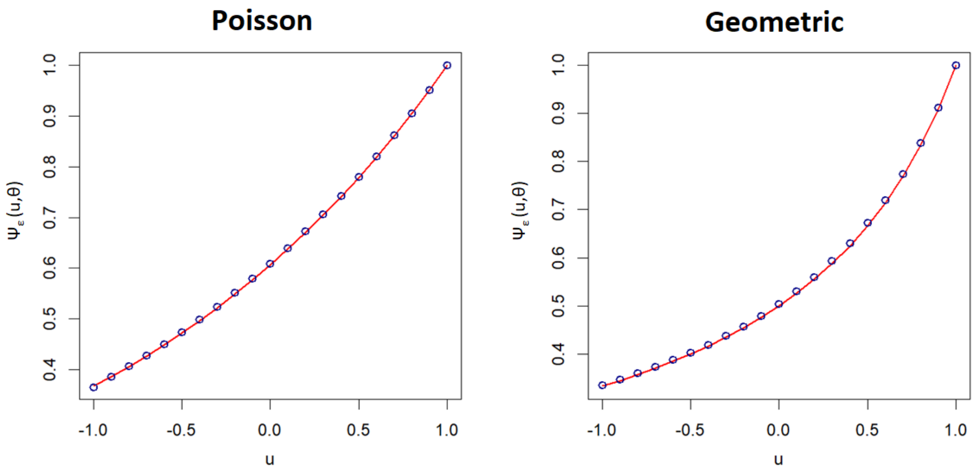
Furthermore, let $\widetilde{\Psi}_T(u) := \frac{1}{T}\sum_{t=1}^{T} u^{\varepsilon_t}$ be the appropriate first-order empirical PGF of series (ε_t), obtained by its realization $\varepsilon_1,\dots,\varepsilon_T$. Thus, the objective function can be written as

$$\begin{aligned} S_T(\theta) &:= \int_{-1}^{1} \omega(u)\Big(\Psi_\varepsilon(u;\theta) - \widetilde{\Psi}_T(u)\Big)^2 \mathrm{d}u = \frac{1}{T^2}\int_{-1}^{1}\omega(u)\left[\sum_{t=1}^{T}\left(\frac{f(\theta u)}{f(\theta)} - u^{\varepsilon_t}\right)\right]^2 \mathrm{d}u \\ &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\int_{-1}^{1}\omega(u)\left(\frac{f(\theta u)}{f(\theta)} - u^{\varepsilon_t}\right)\left(\frac{f(\theta u)}{f(\theta)} - u^{\varepsilon_s}\right)\mathrm{d}u, \end{aligned} \tag{16}$$

where $\omega : [-1,1] \to \mathbb{R}^+$ is the weight function. Finally, in this case, the minimization Equation (5) reads as follows

$$\widehat{\theta}_T = \arg\min_{0<\theta<R} S_T(\theta). \tag{17}$$

As an illustration, Figure 1 shows plots of the PGFs of some typical NNIV IID time series, along with their empirical PGFs.



(**a**) (**b**)

Figure 1. Graphs of PGFs (lines) and empirical PGFs (dots) of the IID series: (**a**) Poisson distribution; (**b**) geometric distribution. (True parameter value is $\theta = 1/2$.)

In the following, we assume that $R < +\infty$, so the set $\overline{(0, R)} = [0, R]$ will be a compact. Thus, assuming that other regularity conditions in Theorem 1 are fulfilled, we can apply the PGF method. For this purpose, we have taken two different types of distribution of the series (ε_t). Firstly, we assume that (ε_t) is the IID series with Poisson distribution and, later, that it has a geometric distribution. In the both cases, for estimating the parameter $\theta > 0$, we generated $N = 500$ independent Monte Carlo replications of the series (ε_t), of the length $T = 1000$. Note that this is close to the length of actual series of COVID-19 death cases, which will be considered in the following. To examine the convergence, i.e., the quality of the PGF estimators, the appropriate estimation errors will also be investigated.

The PGF estimates are computed according to a minimization of the integral in Equation (16). Note that it can be numerically approximated by using an N-point quadrature formula:

$$I(\varphi;\omega) := \int_{-1}^{1} \omega_k(u)\varphi(u)du \approx Q_N(h) := \sum_{j=1}^{N} \omega_k^{(j)} \varphi(u_j),$$

where (u_j) are the quadrature nodes and $\omega_k^{(j)}$, $j = 1, \ldots, N$ are the corresponding weight coefficients. In our simulation study, quadrature formulas based on Gegenbauer orthogonal polynomials, with weights $\omega_k(u) = (1 - u^2)^{(k-1)/2}$ and $N = 20$ nodes, were used. More specifically, we consider Gaussian quadratures based on three well-known special cases of Gegenbauer polynomials, i.e., the appropriate weights:

(i) Chebyshev polynomials (of the first kind): $\omega_0(u) = (1 - u^2)^{-1/2}$;

(ii) Legendre polynomials: $\omega_1(u) \equiv 1$;

(iii) Chebyshev polynomials (of the second kind): $\omega_2(u) = (1 - u^2)^{1/2}$.

The construction of all these quadrature formulas was realized using the Wolfram Mathematica package "OrthogonalPolinomials", authored by Cvetković and Milovanović [24]. After that, the objective function in Equation (17) is minimized by the "R" procedure "optimize", based on the Brent minimization algorithm [25] and realized by the authors' original codes in "R".

Summary statistics of the PGF estimates $\widehat{\theta}_k$, $k = 0, 1, 2$, obtained with the appropriate weights $\omega_k(u)$ and computed via aforementioned estimation procedure, are presented in Table 2. More precisely, it contains the mean values (Mean), minima (Min.), maxima (Max.), and the mean squared estimating errors (MSEE) of obtained PGF estimates of the IID series with Poisson and geometric distribution. (These two distributions are taken as an illustration of the application of the PGF method, that is, to describe the dynamics of actual data in Section 4, below). The values of objective functions $\widehat{S}_{T_k} = S_T(\widehat{\theta}_k)$, $k = 0, 1, 2$, as the reference estimation errors, are also shown. According to presented results, it is obvious that PGF estimates of both IID series are generally the efficient parameter estimators and have very similar properties. Slightly lower estimation errors are observed in the case of the Gauss–Chebyshev quadrature of the second type, i.e., the weight function $\omega_2(u)$. This is expected, since it is known that this weight forces the points around the coordinate origin and ignores those near the ends of the interval $[-1, 1]$. On the other hand, the somewhat larger estimation error in the case of Gauss–Chebyshev quadrature of the first kind is a consequence of 'forcing' the ends of the interval $[-1, 1]$, where the weight is infinite. In addition, the results of AN testing of the obtained PGF estimates are also presented in Table 2. For this purpose, the Anderson–Darling normality test was used and its test statistic (labelled as AD), along with the corresponding p-values, were calculated using the procedure from the R-package "nortest", authored by Gross [26]. According to the values thus obtained, it is evident that the AN property is valid for all PGF estimators, which is in accordance with the previous theoretical results, i.e., the statement of Theorem 1.

Table 2. Estimated parameters' values of IID series (ε_t) series with Poisson and geometric distribution. (True parameter value is $\theta = 0.5$.)

	Statistics	PGF Estimates					
		$\widehat{\theta}_0$	$\widehat{S}_{T_0}$	$\widehat{\theta}_1$	$\widehat{S}_{T_1}$	$\widehat{\theta}_2$	$\widehat{S}_{T_2}$
Poisson	Min.	0.4222	4.63×10^{-8}	0.4293	3.20×10^{-9}	0.4267	3.83×10^{-8}
	Mean	0.5006	9.19×10^{-5}	0.4994	4.47×10^{-5}	0.5005	2.80×10^{-5}
	Max.	0.5967	7.17×10^{-4}	0.6098	4.00×10^{-4}	0.5938	2.38×10^{-4}
	MSEE	0.0220	–	0.0200	–	0.0196	–
	AD	0.3849	–	0.1965	–	0.2911	–
	p-value	0.3919	–	0.8889	–	0.6076	–
Geometric	Min.	0.4491	1.09×10^{-6}	0.4520	3.27×10^{-7}	0.4467	2.40×10^{-7}
	Mean	0.5004	2.30×10^{-4}	0.5001	1.06×10^{-4}	0.4996	6.17×10^{-5}
	Max.	0.5535	2.71×10^{-3}	0.5353	3.62×10^{-3}	0.5580	5.60×10^{-4}
	MSEE	0.0129	–	0.0121	–	0.0117	–
	AD	0.4583	–	0.2214	–	0.2490	–
	p-value	0.2627	–	0.8307	–	0.7467	–

3.2. Estimation of INAR Time Series

In this part of our work, using the previous assumptions, we define *the integer-valued autoregressive process (of the first order)* or, simply, *INAR(1) process* (X_t), $t \in \mathbb{Z}$ by the recurrence relation

$$X_t = \alpha \circ X_{t-1} + \varepsilon_t. \tag{18}$$

Here, (ε_t) is the IID series of RVs commonly called *innovations*, $\alpha \in (0,1)$ is an unknown parameter, and

$$\alpha \circ X := \sum_{j=1}^{X} B_j(\alpha) \tag{19}$$

is the so-called *binomial thinning operator* (for more details, see [27]). More precisely, X is an arbitrary NNIV RV such that, for any $t \in \mathbb{Z}$, the RVs $B_j = B_j(\alpha)$ are mutually independent (also independent of X) with Bernoulli's distribution

$$P\{B_j = 1\} = 1 - P\{B_j = 0\} = \alpha.$$

In that way, it is simply shown that the INAR(1) process (X_t) is stationary for each $\alpha \in (0,1)$, with the autocorrelation function (ACF)

$$\rho_X(k) := \mathrm{Corr}[X_t, X_{t+k}] = \alpha^k, \quad k = 0,1,2,\ldots \tag{20}$$

As an illustration, Figure 2 presents sample frequency distributions of the INAR(1) process (X_t), along with their corresponding innovations (ε_t), for Poisson and geometric distributions. Furthermore, it will be of interest to use the first-order PGF of the RV $\alpha \circ X$. For that purpose, we prove the following simple statement:

Lemma 1. *For an arbitrary $\alpha \in (0,1)$ and NNIV RV X, the first-order PGF of the RV $\alpha \circ X$ reads as follows:*

$$\Psi_{\alpha \circ X}(u;\alpha) = \Psi_X(1 + \alpha(u-1)). \tag{21}$$

Proof. According to the definition of the first-order PGF, Equation (19) and the law of total expectation (see, e.g., Theorem 34.4 in [28]), we have

$$\begin{aligned}\Psi_{\alpha\circ X}(u;\alpha) &:= E\left[u^{\alpha\circ X}\right] = E\left[u^{\sum_{j=1}^{X} B_j(\alpha)}\right] = E\left[E\left(u^{\sum_{j=1}^{X} B_j(\alpha)}\,\middle|\, X\right)\right] \\ &= E\left[\prod_{j=1}^{X}\Psi_j(u;\alpha)\right] = E\left[\left(\Psi_j(u;\alpha)\right)^X\right].\end{aligned} \tag{22}$$

Here, $\Psi_j(u;\alpha) := E\left[u^{B_j(\alpha)}\right] = 1+\alpha(u-1)$ is the first-order PGF of the RVs $B_j(\alpha)$. Hence, substituting this expression in Equation (22) obviously gives Equation (21). □

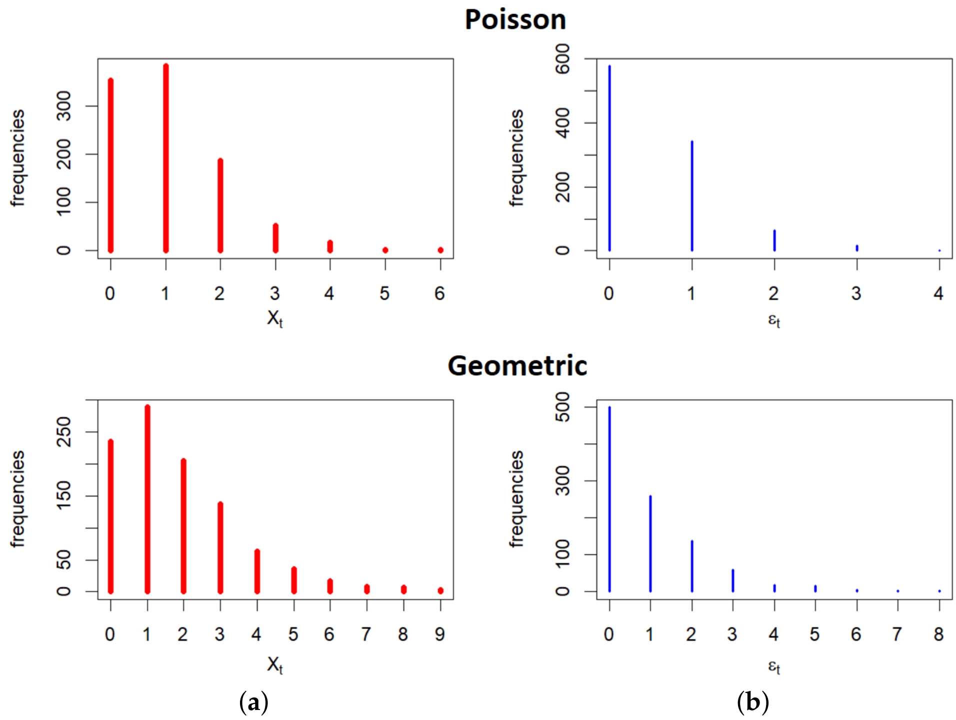


Figure 2. Plots of the sample frequency distributions of INAR(1) processes (**a**) and their innovations (**b**) with Poisson distribution (plots above) and geometric distribution (plots below). True parameters' values are $\theta = \alpha = 1/2$.

Let us also note that some of the important properties of the INAR(1) process can be found in many studies, cf. [29–36]. This kind of stochastic model was first introduced in the pioneer work of Al-Osh and Alzaid [37]. Here, among others, the first order PGF of the INAR(1) process was derived. In the case of PS-distributed innovations (ε_t), we generalized this result in the following way:

Theorem 2. *Let us assume that the innovations (ε_t) of the INAR(1) process have the PS distribution given by Equation (14). Additionally, suppose that the function $g(\theta) = \log f(\theta)$ has a finite derivative, uniformly bounded on a finite interval $\theta \in (0, R)$. Then, an arbitrary $\alpha \in (0,1)$ INAR (1) series (X_t), given by Equation (19), has a PGF of the first order:*

$$\Psi_X(u;\theta,\alpha) = \prod_{k=0}^{\infty}\left(\frac{f\left(\left(1+\alpha^k(u-1)\right)\theta\right)}{f(\theta)}\right), \tag{23}$$

as well as PGFs of order $r \geq 2$:

$$\Psi_X^{(r)}(\mathbf{u};\theta,\alpha) = \Psi_X\left(\prod_{k=0}^{r-1}\left(1+\alpha^k(u_{k+1}-1)\right)\right)\prod_{\ell=2}^{r}\Psi_\varepsilon\left(\prod_{k=0}^{r-\ell}\left(1+\alpha^k(u_{k+\ell}-1)\right)\right). \tag{24}$$

Proof. According to a definition of the PS-distributed RVs (ε_t), given by Equation (14), as well as their first-order PGF, given by Equation (15), the mathematical expectation of the series (ε_t) can be obtained as follows:

$$\mu_\varepsilon(\theta) := \mathrm{E}[\varepsilon_t] = \frac{\partial \Psi_\varepsilon(u;\theta)}{\partial u}\bigg|_{u=1} = \theta\,\frac{f'(\theta)}{f(\theta)} = \theta g'(\theta).$$

Based on this and the assumptions of the theorem, for some constant $M > 0$ and any $\theta \in (0, R)$, we have

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

where the sum above converges uniformly on $\theta \in (0, R)$. On the other hand, the sequence $1/k$, $k = 1, 2, \ldots$ is monotone and bounded, so Abel's convergence criterion implies that, when uniformly on θ, the following is valid:

$$\sum_{k=1}^{\infty} \frac{1}{k} P\{\varepsilon_t \geq k\} < +\infty. \tag{25}$$

As it is known from Alzaid and Al-Osh [38], the inequality (25) represents a necessary and sufficient condition for the infinite-order integer moving average (INMA) representation of the series (X_t):

$$X_t \overset{d}{=} \sum_{k=0}^{\infty} \alpha^k \circ \varepsilon_{t-k}, \tag{26}$$

where the sum above converges almost surely. According to Equation (26) and the previous Lemma, we have the following:

$$\Psi_X(u;\theta,\alpha) = \prod_{k=0}^{\infty} \Psi_\varepsilon\Big(1 + \alpha^k(u-1);\theta\Big), \tag{27}$$

and the product above converges absolutely at least for all $u \in [-1, 1]$. Hence, substituting Equation (15) in Equation (27), we have Equation (23). In order to prove the second part of the theorem, notice that, according to the definition (18) of the series (X_t), for an arbitrary $k \in \mathbb{N}$, the following holds:

$$X_{t+k} \overset{d}{=} \alpha^k \circ X_t + \sum_{j=0}^{k-1} \alpha^j \circ \varepsilon_{t+j}. \tag{28}$$

Substituting $k = 1, \ldots, r-1$ into Equation (28) and after some computation similar to the previous one, an explicit expression of the PGF in Equation (24) can be easily obtained. □

Remark 1. *Note that the distribution of the INAR(1) process depends on two unknown parameters θ, α. Therefore, in order to estimate them, it is of interest to determine the two-dimensional PGF of (X_t), which is simply obtained by replacing $r = 2$ in Equation (24):*

$$\Psi_X^{(2)}(u_1, u_2;\theta,\alpha) = \Psi_X\big(u_1(1 + \alpha(u_2 - 1));\theta,\alpha\big)\Psi_\varepsilon(u_2;\theta). \tag{29}$$

In the following, similar to the case of IID series, the procedure for estimating the parameters of the INAR(1) model with Poisson and geometric innovations (ε_t) will be described. Using Poisson distribution for series (ε_t), as well as Equations (23) and (29), the one- and two-dimensional PGFs of the INAR(1) process (X_t), after some simple computations, can be obtained as follows:

$$\Psi_X(u;\theta,\alpha) = \prod_{k=0}^{\infty} \frac{\exp\left[\theta\left(1+\alpha^k(u-1)\right)\right]}{\exp(\theta)} = \exp\left[\frac{\theta(u-1)}{1-\alpha}\right],$$
$$\Psi_X^{(2)}(u_1,u_2;\theta,\alpha) = \exp\left[\theta\left(u_1+u_2-1+\frac{\alpha u_1 u_2 - 1}{1-\alpha}\right)\right].$$

In a similar way, for the appropriate PGFs of the INAR(1) process with geometric innovations, one obtains

$$\Psi_X(u;\theta,\alpha) = \prod_{k=0}^{\infty} \frac{1-\theta}{1-\theta(1+\alpha^k(u-1))} = \exp\left[-\sum_{k=0}^{\infty} \ln\left(1-\frac{\theta(u-1)}{1-\theta}\alpha^k\right)\right]$$
$$= \exp\left[\sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j(1-\alpha^j)}\left(\frac{\theta(u-1)}{1-\theta}\right)^j\right],$$
$$\Psi_X^{(2)}(u_1,u_2;\theta,\alpha) = \exp\left[\sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j(1-\alpha^j)}\left(\frac{\theta(u_1(1+\alpha(u_2-1)))-1)}{1-\theta}\right)^j\right] \cdot \frac{1-\theta}{1-\theta u_2}.$$

Three-dimensional plots of the theoretical and empirical PGFs of INAR(1) processes with Poisson and geometric innovations, respectively, are shown in Figure 3.

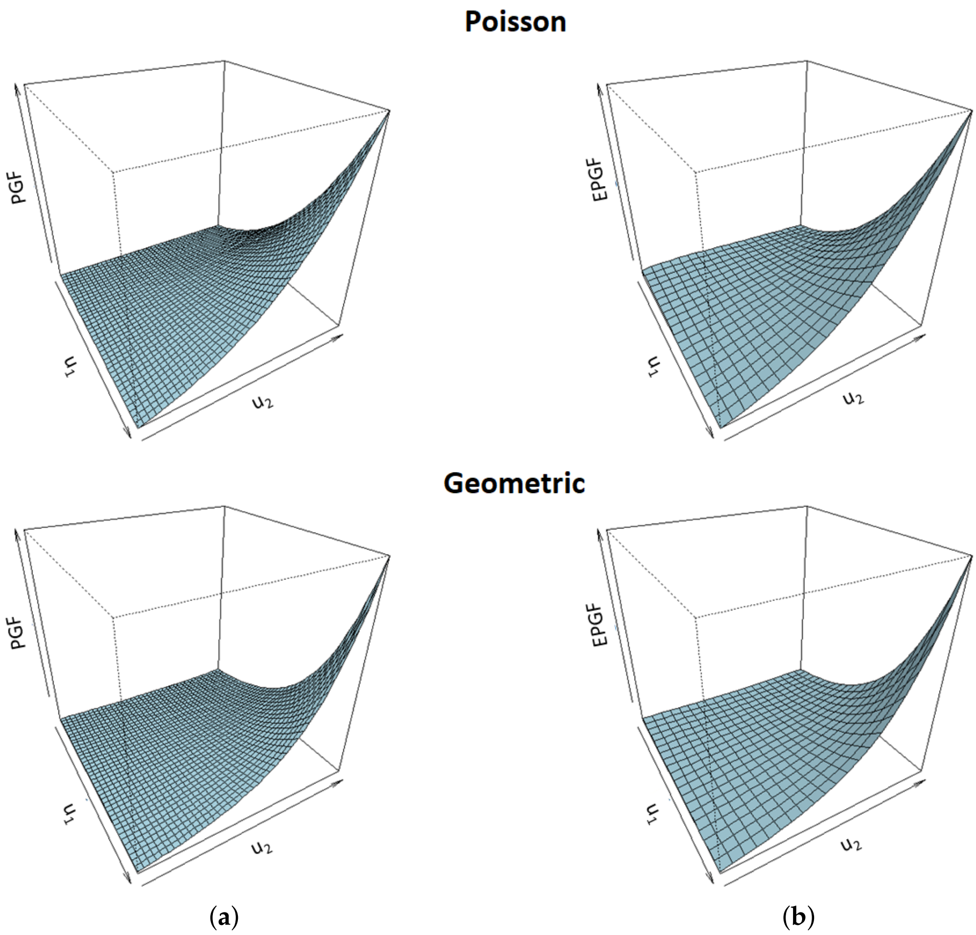


Figure 3. Three-dimensional plots of the two-dimensional PGFs (**a**) and empirical PGFs (**b**) of the INAR(1) series with Poisson innovations (panels above) and geometric innovations (panels below). True parameters' values are $\theta = \alpha = 1/2$.

Therefore, notice that the PGFs of the INAR(1) process with geometric innovations are not in closed form. However, they depend on alternating sums and can be easily approximated by finite k-term sums. Thus, for instance, in the case of a one-dimensional PGF, the approximation error amounts to

$$R_{k+1}(u;\theta,\alpha) \leq \left.\frac{\theta^{k+1}(u-1)^{k+1}}{(k+1)(1-\alpha^{k+1})(1-\theta)^{k+1}}\right|_{u=\frac{1}{\theta}} \leq \frac{1}{(k+1)(1-\alpha^{k+1})}.$$

Similar to that in the previous estimation procedure, the PGF estimates $(\widehat{\theta},\widehat{\alpha})$ of unknown parameters (θ,α) are computed according to a minimization of the double integral

$$S_T^{(2)}(\theta,\alpha) = \iint_{[-1,1]^2} \omega(u_1,u_2)\Big(\Psi_X^{(2)}(u_1,u_2;\theta,\alpha) - \widetilde{\Psi}_T(u_1,u_2)\Big)^2 \mathrm{d}u_1\mathrm{d}u_2, \qquad (30)$$

with respect to some weight function $\omega : [-1,1]^2 \to \mathbb{R}^+$. Similar to earlier, the integral in (30) can be numerically approximated by using some N-point cubature formula:

$$I(\varphi;\omega) := \iint_{[-1,1]^2} \omega(u_1,u_2)\varphi(u_1,u_2)du_1du_2 \approx \sum_{j=1}^{N} \omega_j\, \varphi(u_{1j},u_{2j}),$$

where (u_{1j},u_{2j}) are the cubature nodes, and ω_j are the corresponding weight coefficients. In our simulations study, we used the same kind of (two-dimensional) weights as in the IID case, i.e., the Gaussian qubature formulas based on Gegenbauer orthogonal polynomials, with weights

$$\omega_k(u_1,u_2) = \Big((1-u_1^2)(1-u_2^2)\Big)^{(k-1)/2}$$

and $N = 36$ nodes. After that, the objective function (30) is minimized by "R" procedure for linearly constrained minimization "constrOptim", based on the Nelder–Mead optimization method [39].

As the initial values of this estimation procedure, we used the Yule–Walker (YW) estimates, obtained by solving the equations

$$\mu_X = \overline{X}_T, \qquad \rho_X(1) = \widehat{\rho}(1).$$

Here,

$$\mu_X := E[X_t] = \frac{\mu_\varepsilon}{1-\alpha}, \qquad \rho_X(1) := \mathrm{Corr}[X_t, X_{t+1}] = \alpha$$

are, respectively, the mean and first autocorrelation of the INAR(1) series (X_t), while

$$\overline{X}_T := \frac{1}{T}\sum_{t=1}^{T} X_t, \qquad \widehat{\rho}(1) := \frac{T^{-1}\sum_{t=1}^{T} X_t X_{t-1} - \overline{X}_T^2}{T^{-1}\sum_{t=1}^{T} X_t^2 - \overline{X}_T^2}$$

are their sample estimators. Thus, in the case of Poisson distributed innovations (ε_t), the YW estimators are obtained as follows:

$$\widetilde{\alpha} = \widehat{\rho}(1), \qquad \widetilde{\theta} = (1-\widetilde{\alpha})\overline{X}_T.$$

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

$$\widetilde{\alpha} = \widehat{\rho}(1), \qquad \widetilde{\theta} = \frac{(1-\widetilde{\alpha})\overline{X}_T}{1+(1-\widetilde{\alpha})\overline{X}_T}.$$

Summary statistics of the PGF estimates $(\widehat{\theta}_k, \widehat{\alpha}_k)$, obtained using the aforementioned numerical estimation procedure with the appropriate weights $\omega_k(u_1,u_2)$, $k = 0,1,2$, are presented in the following Table 3. Similar to in the IID case, two different types of innovations (ε_t), with Poisson and geometric distributions, were used. Thus, in both cases, the unknown parameters (θ, α) were estimated through the set of $N = 500$ independent Monte Carlo simulations of the innovation series (ε_t) and INAR(1) process (X_t) of the length $T = 1000$.

In the same way as before, Table 3 contains the mean values (Mean), minima (Min.), maxima (Max.), as well as the mean squared estimating errors (MSEE) and the objective function $S_T^{(2)}$ for all obtained PGF estimates. Notice that the thus obtained PGF estimators have very similar properties as in the one-dimensional case, with somewhat higher estimation errors, which is expected. Namely, they are computed using a two-step procedure, using the YW estimated values as the initial ones.

Table 3. Estimated parameter values of the INAR(1) process (X_t) with Poisson and geometric innovations. (True parameters' values are $\theta = \alpha = 1/2$).

	Statistics	PGF Estimates								
		$\widehat{\theta}_0$	$\widehat{\alpha}_0$	$S_{T_0}^{(2)}$	$\widehat{\theta}_1$	$\widehat{\alpha}_1$	$S_{T_1}^{(2)}$	$\widehat{\theta}_2$	$\widehat{\alpha}_2$	$S_{T_2}^{(2)}$
Poisson	Min.	0.3845	0.3606	2.12×10^{-5}	0.3653	0.3352	5.48×10^{-6}	0.3710	0.3365	1.65×10^{-6}
	Mean	0.5003	0.4986	8.08×10^{-4}	0.5002	0.5006	1.73×10^{-4}	0.5015	0.4951	7.36×10^{-5}
	Max.	0.6711	0.5971	7.12×10^{-3}	0.6948	0.6292	1.32×10^{-3}	0.7049	0.6251	5.13×10^{-4}
	MSEE	0.0329	0.0350	–	0.0364	0.0337	–	0.0381	0.0321	–
	AD	0.7468	0.3089	–	0.7272	0.4153	–	0.8609	0.4724	–
	p-value	0.0515	0.5574	–	0.0576	0.3326	–	0.0269 *	0.2427	–
Geometric	Min.	0.3962	0.3592	3.96×10^{-5}	0.3673	0.3129	7.98×10^{-6}	0.3965	0.3199	2.75×10^{-6}
	Mean	0.5002	0.4974	7.31×10^{-4}	0.4981	0.4977	2.23×10^{-4}	0.4986	0.5005	8.98×10^{-5}
	Max.	0.6071	0.6646	9.02×10^{-3}	0.5960	0.6037	1.57×10^{-3}	0.6033	0.6747	7.39×10^{-4}
	MSEE	0.0291	0.0521	–	0.0265	0.0446	–	0.0271	0.0425	–
	AD	0.6155	0.8198	–	0.3150	0.5174	–	0.4126	0.7898	–
	p-value	0.1088	0.0340 *	–	0.5426	0.1882	–	0.3375	0.0403 *	–

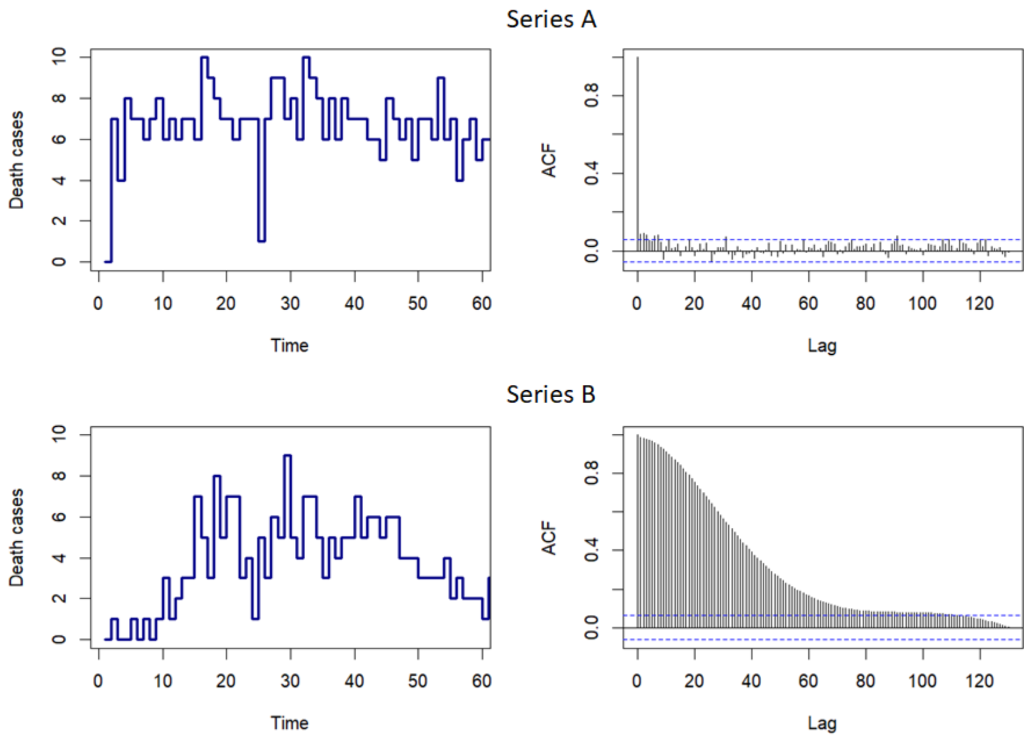
* $p < 0.05$.

Similar to earlier, the summarized values of AD test statistics, along with the appropriate p-values, are also shown in Table 3. It can be seen that the AN property is confirmed in almost all cases, that is, for most of the PGF estimators of the parameters (θ, α), as well as the corresponding weight functions $\omega_k(u_1, u_2)$. We assume, in the same way as it is shown in Stojanović et al. [40], that the variations that occur in some simulations, at the significant level of $0.01 < p < 0.05$, are also reasonable as a consequence of this specific, two-stage estimation procedure.

4. Application of the PGF Estimates

In this section, we indicate some possibilities of the practical application of PGF estimators in stochastic modeling of real data. Additionally, our motivation is also to point out the different dynamics of time series of death cases in the time period before and after the emergence of the SARS-CoV2 virus. We emphasize that much attention has been paid to the COVID-19 disease, from the beginning of the pandemic until today. For these reasons, various theoretical models were proposed for research of this still current issue. Rigorous mathematical ones, usually based on solving coupled partial equations, have been proposed, e.g., in [41–43]. Some other authors [44–46] used both deterministic and stochastic approaches to predict the various effects caused by the COVID-19 pandemic. A particular interesting approach is given in [47,48] where the NNIV time series were used for modelling the dynamic of some the COVID-19 count data. In this regard, here, we explore some more possibilities in modeling the dynamics of such data.

More precisely, we observed two real data sets, one of which refers to the period before and the other after the appearance of the SARS-CoV2 virus infection on the territory of the Republic of Serbia. The first time series, labelled as Series A, contains the number of daily mortality in Niš, the third largest city in the Republic of Serbia, in the period from 1 January 2017 until 31 December 2019. The second one (Series B) contains the daily number of deaths per day in the Republic of Serbia, from 20 March 2020 (the date of the first death caused by the disease COVID-19) to 1 December 2022. Dynamics of the observed time series, lengths $T = 1095$ and $T = 990$, respectively, are shown on the left plots in Figure 4. On the other hand, the right plots in Figure 4 show the autocorrelation functions (ACFs) of these time series. In the case of Series A, it is obvious that there is a very weak correlation, so it can be seen as an IID time series. On the contrary, with Series B, it is evident that there is a positive and significant but exponentially decreasing correlation, depending on the lag $k = 0, 1, 2, \ldots$. Therefore, in accordance with the previous theoretical results, primarily Equation (20), it could be expected that the INAR(1) process is an adequate stochastic model for describing the dynamics of the number of deaths caused by the COVID-19 disease.



(**a**) (**b**)

Figure 4. (**a**) Dynamics of the number of deaths before and after the onset of the disease COVID-19 in Serbia. (**b**) The corresponding ACFs of the observed time series.

Some confirmation of these observations can also be seen in Table 4, where its upper part shows the basic descriptive statistical analysis of the both time series. In the case of Series A, the existence of an equal-dispersion property can be recognized through 'close' values of its sample mean and variance. Thus, this series can be modeled with a Poisson distribution. On the other hand, with Series B, there is a wide range and variability of observed data. Additionally, it can be seen that frequencies of 'smaller' numerical values are emphasized (its mode equals 1). Therefore, it could be assumed that, in the INAR(1) model, innovations with a Poisson distribution (and a 'small' parameter θ) can be taken, as well as with a geometric distribution. In accordance with the mentioned facts, we modeled the dynamics of the Series B by using the INAR(1) process (X_t) with geometric distributed innovations (ε_t). Finally, the lower part of Table 4 shows some of the estimated values for the ACFs of both observed time series. Obviously, the obtained autocorrelation values of Series A are not statistically significant, so it can be considered as an IID time series.

Conversely, the estimated ACF values of Series B indicate the possibility of modeling the dynamics of this time series using the INAR(1) process.

Table 4. Summary statistics of the total number of deaths per day before and after the appearance of the COVID-19 disease in the Republic of Serbia.

Statistics	Series A	Series B
Sample size	1095	990
Minimum	0	0
Mode	6	1
1st Quartile	4	3
Median	6	9
Mean	5.758	17.58
3rd Quartile	7	27
Maximum	24	79
St. deviation	2.428	18.48
Variance	5.895	341.6
Skewness	1.797	1.159
Kurtosis	10.803	3.161
ACF(1)	0.085	0.987
...	...	...
ACF(10)	0.020	0.911
...	...	...
ACF(20)	−0.024	0.754
...	...	...
ACF(50)	0.049	0.254
...	...	...
ACF(100)	−0.020	0.078

In the upper part of Table 5, the estimated values of parameter $\theta > 0$ for Series A, as well as estimates (θ, α) of the INAR(1) process for Series B, are shown. As in the previous simulation study, PGF estimates for IID Series A were obtained using an unconstrained Brent's optimization procedure. On the other hand, for Series B, PGF estimates were computed using the two-step estimation procedure described in the previous section, that is, by linearly constrained optimization and YW estimates as initial parameter values. In both cases, as before, weight functions $\omega_k(u_1, u_2)$, $k = 0, 1, 2$ based on Gegenbauer orthogonal polynomials were used.

It can be easily seen that in the case of Series A, estimated parameter θ is 'between' to sample mean and variance, which is expected. For Series B, the estimated values of the parameter θ are relatively 'small'. This is obviously a consequence of the emphasized frequencies of smaller integer values of the observed time series (which can also be seen in the following Figure 5). Conversely, the estimated values of the parameter α are close to 1. This can also be explained by the previously described properties of this time series, which has a significantly high estimated value of the first correlation.

Furthermore, we analyzed the fitting efficiency for both time series, when all weights are applied. For this purpose, we generated 500 independent simulations for the IID Series A, as well as the INAR(1) model for Series B, for all estimated parameters' values. In order to check the efficiency of proposed models, we computed two typical goodness-of-fit statistics: the mean squared estimation error (MSEE) and the Akaike information criterion (AIC). The mean values of these statistics are shown in the middle part of Table 5. It is noticeable that, in all the mentioned cases, the PGF estimates have similar efficiency because the MSEE and AIC statistics have relatively close and small estimated values.

Table 5. Estimated values of parameters, along with error and predictive statistics of the actual data.

Parameters/Statistics	Series A			Series B		
	ω_0	ω_1	ω_2	ω_0	ω_1	ω_2
θ	5.8069	5.8218	5.8176	0.1812	0.1808	0.1810
α	–	–	–	0.9872	0.9865	0.9871
$S_T^{(2)}$	1.06×10^{-3}	1.06×10^{-3}	7.16×10^{-5}	6.48×10^{-3}	3.46×10^{-3}	1.92×10^{-3}
MSEE	0.0726	0.0789	0.0771	0.0221	0.0288	0.0233
AIC	4.5992	4.6023	4.6016	5.3035	5.3041	5.3042
DM	-0.3641	0.1039	0.0974	0.5855	0.7797	0.6456
p-value	0.3579	0.5414	0.5388	0.7208	0.7821	0.7406

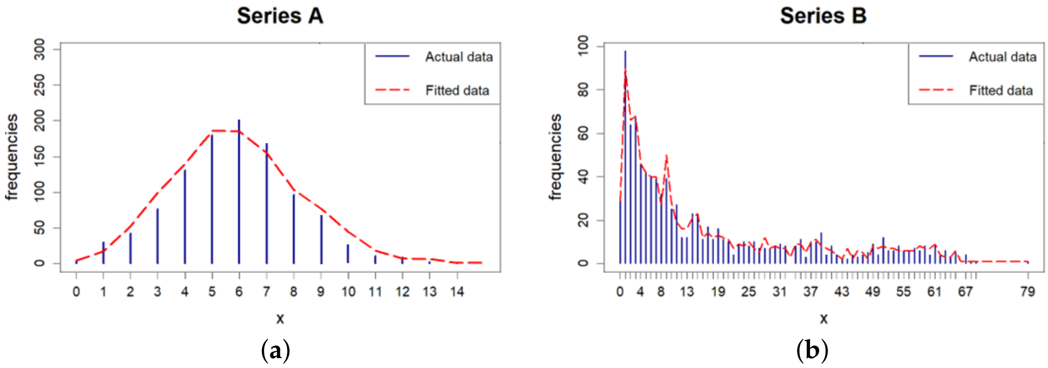


Figure 5. Frequency distribution of the actual and fitted data, obtained by (**a**) IID Poisson process; (**b**) INAR(1) process with geometric innovations.

Finally, an analysis of the forecast accuracy of models obtained with PGF estimates was also checked. For this purpose, we used the one-tailed Diebold–Mariano test [49] for prediction accuracy. The null hypothesis is that the fitted time series have the same accuracy as the actual data, while the alternative is that the series obtained by the PGF estimation procedure are less accurate. The test statistic (denoted as DM), as well as the corresponding p-values, were calculated using the R-package "forecast", authored by Hyndman [50]. In the case of Series A, the data set from 1 January 2020 to 20 March 2020 was used, as a forecast horizon of length $h = 80$. For Series B, the forecast horizon from 1 December 2022 to 9 January 2023 was used, the length of which is equal to $h = 40$. As a result, it is noticeable that the realized values of the DM statistics, shown in the lower part of Table 5, indicate that all models fitted by PGF estimators have the same forecast accuracy as empirical data, i.e., the null hypothesis is valid. Some of these aforementioned facts can be also seen in Figure 5, where the empirical frequencies distribution, along with the appropriate distributions of IID and INAR(1) processes, are shown.

5. Conclusions

In this paper, an estimation method based on probability generating functions (PGFs) is proposed. For this purpose, the stochastic properties of this method were analyzed, as well as its application in estimating the parameters of some typical stationary time series. Notice that the PGF method, with certain modifications, may be suitable for estimating and fitting time series of a similar, or some other kind to those considered here. These can be, for instance, INAR processes with negative binomial thinning [51,52], integer generalized autoregressive (IN-GARCH) processes [53], or the so-called self-exciting processes [54]. Moreover, the application of the PGF estimates is also particularly interesting in the case of some non-linear time series, when the usual estimation methods cannot be used. We point out again that, in Stojanović et al. [20], the PGF method was applied in parameter estimation of the so-called noise-indicator INAR process. Consequently, the application of the PGF estimators presented here can be a starting point for some future research related to these estimators. Therefore, notice that, in the case of Series B, the estimated values of the parameter α indicate that the hypothesis $H_0 : \alpha = 1$ is potentially valid. This could indicate a possible non-stationarity in the dynamics of the observed time series and could be the subject of some further research.

Author Contributions: Conceptualization, V.S. and E.L.; methodology, V.S. and M.T.; software, V.S. and E.L.; validation, E.L.; formal analysis, V.S.; data curation, V.S. and M.T.; writing—original draft preparation, V.S., E.L. and M.T.; writing—review and editing, E.L.; visualization, V.S. and M.T.; supervision, V.S.; project administration, E.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Official data on deaths from the disease COVID-19 can be found on the website of the Office for IT and e-Government of the Republic of Serbia (http://covid19.data.gov.rs/?locale=en, accessed on 6 December 2022).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACF	Auto Correlation Function
AD test	Anderson–Darling test
AIC	Akaike Information Criterion
AN	Asymptotic Normality
CF	Characteristic Function
ECF	Empirical Characteristic Function
IID series	Independent Identical Distributed series
INAR process	Integer-Valued Autoregressive process
INMA	Integer Moving Average
MSEE	Mean Squared Estimating Error
NIINAR process	Noise-Indicator Integer-Valued Autoregressive process
NNIV series	Non-Negative Integer-Valued series
PGF	Probability Generating Function
PMF	Probability Mass Distribution
PS	Power Series
RV	Random Variable
SLLN	Strong Law of Large Numbers
YW	Yule–Walker

References

1. Johnson, N.L.; Kemp, A.W.; Kotz, S. *Univariate Discrete Distributions*, 3rd ed.; Wiley Series in Probability and Statistics; Wiley: Washington, DC, USA, 2005.
2. Xu, H.-Y.; Xie, M.; Goh, T.N.; Fu, X. A Model for Integer–Valued Time Series With Conditional Overdispersion. *Comput. Stat. Data Anal.* **2012**, *56*, 4229–4242. [CrossRef]
3. Jazi, M.A.; Jones, G.; Lai, C.-D. First-Order Integer Valued AR Processes With Zero Inflated Poisson Innovations. *J. Time Ser. Anal.* **2012**, *33*, 954–963. [CrossRef]
4. Weiß, C.H.; Pollet, P.K. Binomial Autoregressive Processes With Density-Dependent Thinning. *J. Time Ser. Anal.* **2014**, *35*, 115–132. [CrossRef]
5. Graziadei, H.; Lijoi, A.; Lopes, H.F.; Marques, P.C.; Prünster, I. Prior Sensitivity Analysis in a Semi-Parametric Integer-Valued Time Series Model. *Entropy* **2020**, *22*, 69. [CrossRef] [PubMed]
6. Khoo, W.C.; Ong, S.H.; Atanu, B. Coherent Forecasting for a Mixed Integer-Valued Time Series Model. *Mathematics* **2022**, *10*, 2961. [CrossRef]
7. El-Morshedy, M. A Discrete Linear-Exponential Model: Synthesis and Analysis with Inference to Model Extreme Count Data. *Axioms* **2022**, *11*, 531. [CrossRef]
8. Du, J.-G.; Li, Y. The Integer-Valued Autoregressive (INAR(p)) Model. *J. Time Ser. Anal.* **1991**, *12*, 129–142. [CrossRef]

9. Franke, J.; Seligmann, T.H. Conditional Maximum Likelihood Estimates for INAR(1) Processes and Their Application to Modelling Epileptic Seizure Counts. In *Developments in Time Series*; Chapman & Hall: London, UK, 1993; pp. 310–330.
10. Latour, A. Existence and Stochastic Structure of a Non-negative Integer-valued Autoregressive Process. *J. Time Ser. Anal.* **1998**, *19*, 439–455.
11. Silva, I.; Silva, M.E. Asymptotic Distribution of the Yule-Walker Estimator for INAR(p) Processes. *Stat. Prob. Lett.* **2006**, *76*, 1655–1663. [CrossRef]
12. Silva, I.; Silva, M.E. Parameter Estimation for INAR Processes Based on High-Order Statistics. *Revstat* **2009**, *7*, 105–117. [CrossRef]
13. Martin, V.L.; Tremayne, A.R.; Jung, R.C. Efficient Method of Moments Estimators for Integer Time Series Models. *J. Time Ser. Anal.* **2014**, *35*, 491–516.
14. Knight, J.L.; Yu, J. Empirical Characteristic Function in Time Series Estimation. *Econom. Theory* **2002**, *18*, 691–721. [CrossRef]
15. Yu, J. Empirical Characteristic Function Estimation and Its Applications. *Econom. Rev.* **2004**, *23*, 93–123. [CrossRef]
16. Kotchoni, R. Applications of the Characteristic Function-Based Continuum GMM in Finance. *Comput. Stat. Data Anal.* **2012**, *56*, 3599–3622. [CrossRef]
17. Meintanis, S.G.; Swanepoel, J.; Allison, J. The probability weighted characteristic function and goodness-of-fit testing. *J. Stat. Plan. Infer.* **2014**, *146*, 122–132. [CrossRef]
18. Stojanović, V.; Milovanović, G.V.; Jelić, G. Distributional Properties and Parameters Estimation of GSB Process: An Approach Based on Characteristic Functions. *ALEA Lat. Am. J. Probab. Math. Stat.* **2016**, *13*, 835–861. [CrossRef]
19. Esquìvel, M.L. Some Applications of Probability Generating Function Based Methods to Statistical Estimation. *Discuss. Math.* **2009**, *29*, 131–153. [CrossRef]
20. Stojanović, V.; Randjelović, D.; Kuk, K. Noise-Indicator Nonnegative Integer-Valued Autoregressive Time Series of the First Order. *Braz. J. Probab. Stat.* **2018**, *32*, 147–171. [CrossRef]
21. Newey, W.K.; McFadden, D. Large Sample Estimation and Hypothesis Testing. In *Handbook of Econometrics*; Elsevier: Amsterdam, The Netherlands, 1994; Volume 4. [CrossRef]
22. Fuller, W.A. *Introduction to Statistical Time Series*; John Wiley & Sons: New York, NY, USA, 1996.
23. Bourguignon, M.; Vasconcellos, K.L.P. First Order Non-Negative Integer Valued Autoregressive Processes With Power Series Innovations. *Braz. J. Probab. Stat.* **2015**, *29*, 71–93.
24. Cvetković, A.S.; Milovanović, G.V. The Mathematica Package "Orthogonal Polynomials". *Facta Univ. Ser. Math. Inform.* **2004**, *19*, 17–36. [CrossRef]
25. Brent, R. *Algorithms for Minimization Without Derivatives*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1973.
26. Gross, L. Tests for Normality. R Package Version 1.0-2. 2013. Available online: http://CRAN.R-project.org/package=nortest (accessed on 6 December 2022).
27. Weiß, C.H. Thinning Operations for Modelling Time Series of Counts—A Survey. *Adv. Statist. Anal.* **2008**, *92*, 319–341.
28. Billingsley, P. *Probability and Measure*; John Wiley & Sons: New York, NY, USA, 1995. [CrossRef]
29. Silva, M.E.; Oliveira, V.L. Difference Equations for the Higher-Order Moments and Cumulants of the INAR(1) Model. *J. Time Ser. Anal.* **2004**, *25*, 317–333.
30. Jung, R.C.; Ronning, G.; Tremayne, A.R. Estimation in Conditional First Order Autoregression with Discrete Support. *Stat. Pap.* **2005**, *46*, 195–224. [CrossRef]
31. Ristić, M.M.; Nastić, A.S.; Miletić, A.V. A Geometric Time Series Model with Dependent Bernoulli Counting Series. *J. Time Ser. Anal.* **2013**, *34*, 466–476. [CrossRef]
32. Schweer, S.; Weiß, C.H. Compound Poisson INAR(1) Processes: Stochastic Properties and Testing for Over-Dispersion. *Comput. Stat. Data Anal.* **2014**, *77*, 267–284. [CrossRef]
33. Bermúdez, L.; Karlis, D. Multivariate INAR(1) Regression Models Based on the Sarmanov Distribution. *Mathematics* **2021**, *9*, 505. [CrossRef]
34. Li, Q.; Chen, H.; Liu, X. A New Bivariate Random Coefficient INAR(1) Model with Applications. *Symmetry* **2022**, *14*, 39. [CrossRef]
35. Maya, R.; Chesneau, C.; Krishna, A.; Irshad, M.R. Poisson Extended Exponential Distribution with Associated INAR(1) Process and Applications. *Stats* **2022**, *5*, 755–772. [CrossRef]
36. Maya, R.; Irshad, M.R.; Chesneau, C.; Nitin, S.L.; Shibu, D.S. On Discrete Poisson–Mirra Distribution: Regression, INAR(1) Process and Applications. *Axioms* **2022**, *11*, 193. [CrossRef]
37. Al-Osh, M.A.; Alzaid, A.A. First-Order Integer-Valued Autoregressive (INAR(1)) Process. *J. Time Ser. Anal.* **1987**, *8*, 261–275. [CrossRef]
38. Alzaid, A.A.; Al-Osh, M.A. An Integer-Valued p^{th}-order Autoregressive Structure (INAR(p)) Process. *J. App. Prob.* **1990**, *27*, 314–324. [CrossRef]
39. Lange K. Numerical Analysis for Statisticians. In *Book Series: Statistics and Computing*; Springer: New York, NY, USA, 2001. [CrossRef]
40. Stojanović, V.; Popović B.Č.; Milovanović G.V. The Split-SV model. *Comput. Statist. Data Anal.* **2016**, *100*, 560–581.
41. Vaz, S.; Torres, D.F.M. A Discrete-Time Compartmental Epidemiological Model for COVID-19 with a Case Study for Portugal. *Axioms* **2021**, *10*, 314. [CrossRef]
42. Ghosh, S.; Volpert, V.; Banerjee, M. An Epidemic Model with Time Delay Determined by the Disease Duration. *Mathematics* **2022**, *10*, 2561. [CrossRef]

43. Sivakumar, B.; Deepthi, B. Complexity of COVID-19 Dynamics. *Entropy* **2022**, *24*, 50. [CrossRef]
44. Hassan, S.M.; Riveros Gavilanes, J.M. First to React Is the Last to Forgive: Evidence from the Stock Market Impact of COVID 19. *J. Risk Financ. Manag.* **2021**, *14*, 26. [CrossRef] [PubMed]
45. Zakharov, V.; Balykina, Y.; Ilin, I.; Tick, A. Forecasting a New Type of Virus Spread: A Case Study of COVID-19 with Stochastic Parameters. *Mathematics* **2022**, *10*, 3725. [CrossRef]
46. Jovanović, M.; Stojanović, V.; Kuk, K.; Popović, B.; Čisar, P. Asymptotic Properties and Application of GSB Process: A Case Study of the COVID-19 Dynamics in Serbia. *Mathematics* **2022**, *10*, 3849. [CrossRef]
47. Stapper, M. Count Data Time Series Modelling in Julia—The CountTimeSeries.jl Package and Applications. *Entropy* **2021**, *23*, 666. [CrossRef]
48. Eliwa, M.S.; Tyagi, A.; Almohaimeed, B.; El-Morshedy, M. Modelling Coronavirus and Larvae Pyrausta Data: A Discrete Binomial Exponential II Distribution with Properties, Classical and Bayesian Estimation. *Axioms* **2022**, *11*, 646. [CrossRef]
49. Diebold, F.X.; Mariano, R.S. Comparing Predictive Accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263. [CrossRef]
50. Hyndman, R. Forecasting Functions for Time Series and Linear Models. R Package Version 7.1. 2016. Available online: http://CRAN.R-project.org/package=forecast (accessed on 9 January 2023).
51. Nastić, A.S.; Ristić, M.M.; Bakouch, H.S. A Combined Geometric INAR(p) Model Based on Negative Binomial Thinning. *Math. Comput. Model.* **2012**, *55*, 1665–1672.
52. Ristić, M.M.; Nastić, A.S.; Bakouch, H.S. Estimation in an Integer-Valued Autoregressive Process With Negative Binomial Marginals (NBINAR(1)). *Comm. Stat. Theory Methods* **2012**, *41*, 606–618. [CrossRef]
53. Weiß, C. INGARCH Models for Count Time Series. In *An Introduction to Discrete-Valued Time Series*; Weiß, C., Ed.; Elsevier: Amsterdam, The Netherlands, 2018. [CrossRef]
54. Reinhart, A. A Review of Self-Exciting Spatio-Temporal Point Processes and Their Applications. *Stat. Sci.* **2018**, *33*, 299–318. [CrossRef]