Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: www.tandfonline.com/journals/tstf20

A new generalized binomial thinning-based INAR(1) process with Poisson–Lindley innovations

Emad-Eldin Aly Ahmed Aly, Naushad Mamode Khan, Muhammed Rasheed Irshad, Veena D'cruz & Radhakumari Maya

To cite this article: Emad-Eldin Aly Ahmed Aly, Naushad Mamode Khan, Muhammed Rasheed Irshad, Veena D'cruz & Radhakumari Maya (2025) A new generalized binomial thinning-based INAR(1) process with Poisson–Lindley innovations, Statistical Theory and Related Fields, 9:4, 384-403, DOI: 10.1080/24754269.2025.2557723

To link to this article: https://doi.org/10.1080/24754269.2025.2557723

Published online: 25 Sep 2025.

STATISTICAL THEORY AND RELATED FIELDS
2025, VOL. 9, NO. 4, 384–403
https://doi.org/10.1080/24754269.2025.2557723

Taylor & Francis
Taylor & Francis Group

A new generalized binomial thinning-based INAR(1) process with Poisson–Lindley innovations

Emad-Eldin Aly Ahmed Aly[a], Naushad Mamode Khan[b], Muhammed Rasheed Irshad [c], Veena D'cruz[c] and Radhakumari Maya[c]

[a]Department of Statistics and Operations Research, Kuwait University, Safat,Kuwait; [b]Department of Economics and Statistics, University of Mauritius, Reduit, Mauritius; [c]Department of Statistics, Cochin University of Science and Technology, Cochin, Kerala, India

ABSTRACT

This paper considers a new parameterization of the generalized binomial thinning operator that is to be incorporated in a simple ordered integer-valued autoregressive process (INAR(1)) with the Poisson–Lindley innovations. The statistical properties of the resulting INAR(1) process are explored along with the estimation procedures. Monte Carlo simulation experiments are executed to assess the consistency of the estimates under the new INAR(1) process. Finally, the importance of the proposed INAR(1) model is confirmed through the analysis of a real data set.

ARTICLE HISTORY
Received 13 January 2025
Accepted 22 August 2025

KEYWORDS
Poisson–Lindley distribution; generalized binomial thinning; conditional least squares; simulation; GBINAR(1) process

1. Introduction

Count data that incorporate time series can be found in a variety of scientific disciplines, such as the insurance industry, sports, medicine, agriculture, and finance, among others. By employing count data models, organizations and researchers can deepen their understanding of events, make informed predictions, and drive impactful decisions. As a result, it is crucial to study and analyse count time series models, which also inspires a brand-new area of research with several practical applications. One of the main approaches used to model such datasets is to use integer-valued autoregressive (INAR) processes. The INAR(1) process was primarily introduced by McKenzie (1985) and Al-Osh and Alzaid (1987) based on the binomial thinning operator with Poisson innovations. The INAR(1) process, $\{Y_t\}_{t\in\mathbb{Z}}$ is defined as

$$Y_t = \alpha \circ Y_{t-1} + \varepsilon_t, \quad 0 \leq \alpha < 1, \tag{1}$$

where the innovations, $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ are independent and identically distributed (i.i.d.) Poisson(λ) random variables (rvs). The operator 'o' in (1) denotes the binomial thinning operator,

CONTACT Muhammed Rasheed Irshad irshadmr@cusat.ac.in Department of Statistics, Cochin University of Science and Technology, Cochin, Kerala, India

introduced by Steutel and van Harn (1979), which is described as

$$\alpha \circ Y_{t-1} = \sum_{j=1}^{Y_{t-1}} C_j,$$

where $\{C_j\}_{j\in\mathbb{Z}}$ is a sequence of i.i.d. Bernoulli rvs with parameter α.

Recently, various INAR(1) models with different innovation processes have been introduced in the statistical literature, such as the INAR(1) process with Poisson–Lindley innovations (INAR(1)PL) introduced by Lívio et al. (2018), the INAR(1) process with discrete three-parameter Lindley as innovation introduced by Eliwa et al. (2020), the INAR(1) process with Poisson quasi xgamma innovations (INAR(1)PQX) proposed by Altun et al. (2021), the INAR(1) process with Bell innovations (INAR(1)BL) introduced by Huang and Zhu (2021), the INAR(1) with discrete pseudo Lindley (INAR(1)DPsL) proposed by Irshad et al. (2021), and the INAR(1) with discrete new XLindley innovation introduced by Maya et al. (2024).

Furthermore, besides the binomial thinning, other thinning operators are also introduced along with INAR(1) models. Weiß (2008) mentions some of the initially developed thinning operators. Al-Osh and Aly (1992) introduced and studied INAR(1) models with geometric and negative binomial marginals based on a negative binomial thinning operator. Aly and Bouzar (1994a) introduced and studied the generalized binomial thinning operator and used it to propose several INAR(1) models. Ristić et al. (2009) introduced a geometric INAR(1) process based on a negative binomial thinning operator. The INAR(p) processes with a signed generalized power series thinning operator are proposed by Zhang et al. (2010). Borges et al. (2016) introduced the geometric time series model with inflated-parameter Bernoulli counting series. Khoo et al. (2017) defined INAR(1) models based on the Pegram mixing and thinning operators. Yang et al. (2019) proposed the generalized Poisson thinning operator. A two-parameter expectation thinning operator based on a linear fractional probability generating function is established by Aly and Bouzar (2019). Extended binomial INAR(1) processes with generalized binomial thinning operator are developed by Kang et al. (2020). Recently, an extended binomial thinning operator is introduced by Liu and Zhu (2020).

In this paper, we propose and study the generalized binomial thinning operator, which is a reparametrization of the thinning operator of Aly and Bouzar (1994a). The resulting INAR(1) process will be denoted by GBINAR(1). The main advantages of this reparametrization are that it has two parameters for having flexible properties, it has the binomial thinning operator as a special case and the computations for mathematical properties as well as estimation procedures are simple compared to other thinning operators.

The Poisson–Lindley (PL) distribution, introduced by Sankaran (1970), is a compound distribution characterized by properties such as unimodality, infinite divisibility, simple forms for the probability mass function (pmf), and various mathematical features. Additionally, it exhibits over-dispersion (Ghitany & Al-Mutairi, 2009). We mount the PL distribution as innovation for the GBINAR(1) process and hence introduce an INAR(1) process with the PL distribution as innovation (i.e., the GBINAR(1)PL process). The PL distribution has been used by various authors for INAR(1) with different thinning operators (see, Lívio et al., 2018, Rostami et al., 2018, etc). Later, we see that the GBINAR(1)PL provides a better fitting criterion than these distributions based on a real data set.

This paper is further arranged as follows. In Section 2, we define the new generalized binomial thinning operator and derive its properties. This new thinning operator is used

to introduce the GBINAR(1)PL process and to derive its various properties in Section 3. The procedures for the estimation of the parameters of the GBINAR(1)PL process are given in Section 4. The results of Monte Carlo simulation studies are reported and discussed in Section 5. As an illustration, we use the GBINAR(1)PL process to analyse a real-life data set in Section 6.

2. The new generalized binomial thinning operator

Aly and Bouzar (1994a) introduced a class of Galton–Watson processes with stationary immigration (GWSI) which generalizes the INAR(1) process defined by Al-Osh and Alzaid (1987). They considered Poisson geometric and negative binomial INAR(1) models with the proposed process. Following the GWSI processes, Aly and Bouzar (1994b) extended those models to some first-order integer-valued autoregressive moving average (INARMA(1)) process. As a result, they proposed Poisson geometric, negative binomial, and Poisson logarithmic INARMA models. All these models are developed with the generalized binomial thinning operator which can be considered as a generalization of the models with the binomial thinning operator of Steutel and van Harn (1979).

The generalized binomial thinning operator of Aly and Bouzar (1994b) is defined as follows,

$$B(\alpha,\theta)\circ X=\sum_{i=1}^{X}\xi_i(\alpha,\theta),$$

where X is a positive integer-valued rv and $\xi_i, i\geq 1$ are i.i.d. rvs such that $\xi_i(\alpha,\theta)\overset{D}{=}M_i\times N_i$, where $(M_i, i\geq 1)$ and $(N_i, i\geq 1)$ are two independent sequences of i.i.d. rvs independent of X such that M_i is Bernoulli (α) and N_i is truncated geometric $((1-\alpha)\theta)$.

Now, we define a new generalized binomial thinning operator as

$$A(\alpha,\theta)\circ X=\sum_{i=1}^{X}W_i(\alpha,\theta),$$

where $W_i, i\geq 1$ are i.i.d. rvs such that $W_i(\alpha,\theta)\overset{D}{=}U_i\times V_i$, and $(U_i, i\geq 1)$ and $(V_i, i\geq 1)$ are two independent sequences of i.i.d. rvs independent of X such that U_i is Bernoulli $\left(\frac{\alpha}{1-\alpha\theta}\right)$ and V_i is truncated geometric $\left(\frac{\theta(1-\alpha)}{1-\alpha\theta}\right)$. Note that $A(\alpha,\theta)\circ$ is a reparametrization of $B(\alpha,\theta)\circ$. We can show that the probability generating function (pgf) of $W_i(\alpha,\theta)$ is given as

$$G_{W_i,\alpha,\theta}(s)=1-\frac{\alpha(1-\theta)(1-s)}{1-\alpha\theta-(1-\alpha)\theta s},\quad 0\leq\alpha\leq 1\quad\text{and}\quad 0\leq\theta\leq 1,\quad |s|\leq 1. \tag{2}$$

The pmf of $W_i(\alpha,\theta)$ is thus

$$P(W_i=k)=\begin{cases}\dfrac{1-\alpha}{1-\alpha\theta}, & \text{if } k=0,\\ \dfrac{\alpha(1-\theta)}{1-\alpha\theta}\left(\dfrac{\theta(1-\alpha)}{1-\alpha\theta}\right)^{k-1}\left(1-\dfrac{\theta(1-\alpha)}{1-\alpha\theta}\right), & \text{if } k=1,2,\ldots.\end{cases}$$

We will call $A(\alpha,\theta)\circ X=\sum_{i=1}^{X}W_i(\alpha,\theta)$ the new generalized binomial thinning operator. Note that the binomial thinning operator is the special case of $A(\alpha,\theta)\circ X$ when $\theta=0$.

Remark 2.1: Note that the new generalized binomial thinning operator is a special case of the fractional thinning operator of Aly and Bouzar (2019) when $r = \frac{1-\theta}{\theta(1-\alpha)}$ and $m = \frac{\alpha}{(1-\alpha\theta)}$.

Theorem 2.1: *For $0 < \alpha, \theta < 1$ and X being a non-negative integer-valued rv, the following properties hold for $A(\alpha,\theta) \circ X$.*

(1)

$$\mathrm{E}(A(\alpha,\theta) \circ X \mid X) = \alpha X,$$
$$\mathrm{Var}(A(\alpha,\theta) \circ X \mid X) = \frac{\alpha(2\theta - \alpha(\theta+1))}{1-\theta} X,$$

and

$$\mathrm{Cov}(A(\alpha,\theta) \circ X, X) = \alpha \mathrm{Var}(X).$$

(2) *Assuming that X and Y are independent non-negative integer-valued rvs, then*

$$A(\alpha,\theta) \circ (X+Y) \stackrel{D}{=} A(\alpha,\theta) \circ X + A(\alpha,\theta) \circ Y.$$

(3) *Letting $0 < \alpha_1, \alpha_2 < 1$,*

$$A(\alpha_1,\theta) \circ (A(\alpha_2,\theta) \circ X) = A(\alpha_1\alpha_2,\theta) \circ X.$$

(4) *Letting $A^{(n)}(\alpha,\theta) \circ X = A^{(n-1)}(\alpha,\theta) \circ (A(\alpha,\theta) \circ X)$, where n is a positive integer ≥ 2, $A^{(1)}(\alpha,\theta) \circ X = A(\alpha,\theta) \circ X$, then*

$$A^{(n)}(\alpha,\theta) \circ X \stackrel{D}{=} A(\alpha^n,\theta) \circ X, \quad n \geq 2.$$

(5)

$$\lim_{\alpha \downarrow 0} A(\alpha,\theta) \circ X \stackrel{D}{=} 0 \quad \text{and} \quad \lim_{\alpha \uparrow 1} A(\alpha,\theta) \circ X \stackrel{D}{=} X.$$

***Proof*:** (1) By the usual definition of expectation, variance, covariance, conditional expectation and differentiating pgf for obtaining moments, we can derive mean, variance and covariance as

$$\mathrm{E}(A(\alpha,\theta) \circ X \mid X) = \sum_{i=1}^{X} \mathrm{E}(W_i(\alpha,\theta) \mid X) = \alpha X,$$
$$\mathrm{Var}(A(\alpha,\theta) \circ X \mid X) = \sum_{i=1}^{X} \mathrm{Var}(W_i(\alpha,\theta) \mid X) = \frac{\alpha(2\theta - \alpha(\theta+1))}{1-\theta} X,$$

and

$$\begin{aligned}\operatorname{Cov}(A(\alpha,\theta)\circ X, X) &= \mathrm{E}\left(\sum_{i=1}^{X} W_i(\alpha,\theta)\right) - \mathrm{E}(X)\mathrm{E}\left(\sum_{i=1}^{X} W_i(\alpha,\theta)\right)\\ &= \mathrm{E}\left(\mathrm{E}\left(\sum_{i=1}^{X} W_i(\alpha,\theta) \mid X\right)\right) - \alpha(\mathrm{E}(X))^2\\ &= \alpha\mathrm{E}(X^2) - \alpha(\mathrm{E}(X))^2\\ &= \alpha\,\operatorname{Var}(X).\end{aligned}$$

(2) Considering the pgf of L.H.S,

$$\mathrm{E}\left(S^{\sum_{i=1}^{X+Y} W_i(\alpha,\theta)}\right) = (G_{X+Y,\alpha,\theta}(s))^{X+Y},$$

where $G_{X,\alpha,\theta}(s)$ is given in (2). The pgf of R.H.S is then,

$$\begin{aligned}&\mathrm{E}\left(S^{\sum_{i=1}^{X} W_i(\alpha,\theta)} + S^{\sum_{i=1}^{Y} W_i(\alpha,\theta)}\right)\\ &= (G_{X,\alpha,\theta}(s))^X (G_{Y,\alpha,\theta}(s))^Y = (G_{X+Y,\alpha,\theta}(s))^{X+Y}.\end{aligned}$$

Therefore, L.H.S and R.H.S have the same pgf implying both are equally distributed.

(3) Suppose $U = \sum_{i=1}^{X} W_i(\alpha_2,\theta)$. The pgf of L.H.S in the equation to prove is

$$\begin{aligned}\mathrm{E}\left(S^{\sum_{i=1}^{U} W_i(\alpha_1,\theta)}\right) &= \mathrm{E}_U\left(\mathrm{E}\left(S^{\sum_{i=1}^{U} W_i(\alpha_1,\theta)} \mid U\right)\right)\\ &= \mathrm{E}_U(1 - G_{X,\alpha_1,\theta}(t))^U\\ &= 1 - G_{X,\alpha_1,\theta}(G_{X,\alpha_2,\theta}(s))\\ &= 1 - \frac{\alpha_1(1-\theta)\left(\frac{\alpha_2(1-\theta)(1-s)}{1-\alpha_2\theta-(1-\alpha_2)\theta s}\right)}{1-\alpha_1\theta-(1-\alpha_1)\theta\left(1-\frac{\alpha_2(1-\theta)(1-s)}{1-\alpha_2\theta-(1-\alpha_2)\theta s}\right)}\\ &= 1 - \frac{\alpha_1\alpha_2(1-\theta)(1-s)}{1-\alpha_1\alpha_2-(1-\alpha_1\alpha_2)\theta s}\\ &= G_{X,\alpha_1\alpha_2,\theta}(s).\end{aligned}$$

The pgf of R.H.S in the equation to prove is

$$\mathrm{E}\left(S^{\sum_{i=1}^{X} W_i(\alpha,\theta)}\right) = G_{X,\alpha_1\alpha_2,\theta}(s),$$

implying L.H.S and R.H.S have the same distribution.

(4) Consider $A^{(n)}(\alpha,\theta)\circ X = A^{(n-1)}(\alpha,\theta)\circ(A(\alpha,\theta)\circ X)$, where n, the exponent is an integer ≥ 2, and $A^{(1)}(\alpha,\theta)\circ X = A(\alpha,\theta)\circ X$. By using point number (3), we can write

$$A(\alpha,\theta)\circ(A(\alpha,\theta)\circ X) = A(\alpha^2,\theta)\circ X.$$

When $n = 2$,

$$\begin{aligned} A^{(2)}(\alpha,\theta)\circ X &= A^{(1)}(\alpha,\theta)\circ(A(\alpha,\theta)\circ X) \\ &= A(\alpha,\theta)\circ(A(\alpha,\theta)\circ X) \\ &= A(\alpha^2,\theta)\circ X. \end{aligned}$$

Also when $n = 3$,

$$\begin{aligned} A^{(3)}(\alpha,\theta)\circ X &= A^{(2)}(\alpha,\theta)\circ(A(\alpha,\theta)\circ X) \\ &= A(\alpha^2,\theta)\circ(A(\alpha,\theta)\circ X) \\ &= A(\alpha^3,\theta)\circ X. \end{aligned}$$

Similarly for large n,

$$\begin{aligned} A^{(n)}(\alpha,\theta)\circ X &= A^{(n-1)}(\alpha,\theta)\circ(A(\alpha,\theta)\circ X) \\ &= A(\alpha^{n-1},\theta)\circ(A(\alpha,\theta)\circ X) \\ &= A(\alpha^n,\theta)\circ X. \end{aligned}$$

(5) When $\alpha \to 0$, pgf of $A(\alpha,\theta)\circ X \to 1$. Also, we know that the pgf of an rv is 1 if it is 0. When $\alpha \to 1$, pgf of $A(\alpha,\theta)\circ X \to S^X$. Pgf of X is $\mathrm{E}(S^X) = S^X$ if X is defined. This completes the proof.

■

Remark 2.2: Now using the first point of Theorem 2.1, the Fisher index of dispersion (DI) of $A(\alpha,\theta)\circ X \mid X$ is given by

$$\begin{aligned} \mathrm{DI}(A(\alpha,\theta)\circ X \mid X) &= \frac{\mathrm{Var}(A(\alpha,\theta)\circ X \mid X)}{\mathrm{E}(A(\alpha,\theta)\circ X \mid X)} \\ &= \frac{2\theta - \alpha(\theta+1)}{1-\theta} \\ &= \frac{2(1-\alpha)}{1-\theta} + \alpha - 2, \end{aligned}$$

implying $A(\alpha,\theta)\circ X \mid X$ can be under or over-dispersed based on the values of α and θ. To be precise, if $\alpha < \frac{3\theta-1}{1+\theta}$, $A(\alpha,\theta)\circ X \mid X$ is over-dispersed or if $\alpha > \frac{3\theta-1}{1+\theta}$, $A(\alpha,\theta)\circ X \mid X$ is under-dispersed and if $\alpha = \frac{3\theta-1}{1+\theta}$, $A(\alpha,\theta)\circ X \mid X$ is equally-dispersed.Hence unlike the binomial thinning operator, the proposed generalized binomial thinning operator can have under or over-dispersed properties.

Remark 2.3: Here through the reparameterization $A(\alpha,\theta)\circ$ provides simple forms and more flexible properties than that of $B(\alpha,\theta)\circ$. Moreover the thinning operator $B(\alpha,\theta)\circ$ was used to form integer-valued moving average processes in Aly and Bouzar (1994a). Here we introduce an INAR(1) process using the thinning operator $A(\alpha,\theta)\circ$ and having the PL innovations.

3. The GBINAR(1)PL process: definition and properties

A sequence $\{Y_t, t \geq 0\}$ of positive integer rvs is said to be INAR(1) model generated by the new generalized binomial thinning operator with PL innovations or GBINAR(1)PL if

$$Y_t = A(\alpha, \theta) \circ Y_{t-1} + \varepsilon_t, \quad t \geq 0, \quad 0 \leq \alpha \leq 1, \quad 0 \leq \theta \leq 1, \tag{3}$$

where $\varepsilon_t, t \geq 0$ are i.i.d. rvs from the PL distribution with parameters λ, ε_t independent from W_i and Y_{t-i} for $t \geq 1$. Using the results of Aly and Bouzar (1994a) we can verify $\{Y_t\}$ is an ergodic Markov chain. Hence, there exists a unique strictly stationary process satisfying (3). Some properties regarding the GBINAR(1)PL process are mentioned in the following lemmas.

Lemma 3.1: *The one step ahead conditional mean and conditional variance of the GBINAR(1)PL process* $\{Y_t\}$ *defined in* (3) *are given by*

$$\mathrm{E}(Y_t|Y_{t-1}) = \alpha Y_{t-1} + \frac{\lambda + 2}{\lambda(\lambda + 1)},$$

and

$$\mathrm{Var}(Y_t|Y_{t-1}) = \frac{\alpha(2\theta - \alpha(\theta + 1))}{1 - \theta} Y_{t-1} + \frac{\lambda^3 + 4\lambda^2 + 6\lambda + 2}{\lambda^2(1 + \lambda)^2}.$$

Proof: Using the usual definition of conditional mean, variance and Theorem 2.1, we have

$$\begin{aligned} \mathrm{E}(Y_t|Y_{t-1}) &= \mathrm{E}(A(\alpha, \theta) \circ Y_{t-1} + \varepsilon_t|Y_{t-1}) \\ &= \mathrm{E}(A(\alpha, \theta) \circ Y_{t-1}|Y_{t-1}) + \mathrm{E}(\varepsilon_t|Y_{t-1}) \\ &= \alpha Y_{t-1} + \frac{\lambda + 2}{\lambda(\lambda + 1)}, \end{aligned} \tag{4}$$

and

$$\begin{aligned} \mathrm{Var}(Y_t|Y_{t-1}) &= \mathrm{Var}(A(\alpha, \theta) \circ Y_{t-1} + \varepsilon_t|Y_{t-1}) \\ &= \mathrm{Var}(A(\alpha, \theta) \circ Y_{t-1}|Y_{t-1}) + \mathrm{Var}(\varepsilon_t|Y_{t-1}) \\ &= \frac{\alpha(2\theta - \alpha(\theta + 1))}{1 - \theta} Y_{t-1} + \frac{\lambda^3 + 4\lambda^2 + 6\lambda + 2}{\lambda^2(1 + \lambda)^2}. \end{aligned} \tag{5}$$

This completes the proof. ■

Lemma 3.2: *For the stationary GBINAR(1)PL process* $\{Y_t\}$ *defined in* (3)*, the mean, variance, covariance at lag h (h is an integer* > 0*) and hence the auto correlation function (ACF) are*

given by

$$\mathrm{E}(Y_t) = \frac{\lambda+2}{\lambda(\lambda+1)(1-\alpha)}, \tag{6}$$

$$\mathrm{Var}(Y_t) = \frac{(1-\alpha)(1-\theta)(\lambda^3+4\lambda^2+6\lambda+2)+\alpha\lambda(\lambda+1)(\lambda+2)(2\theta-\alpha(\theta+1))}{(1-\alpha)^2(\alpha+1)(1-\theta)\lambda^2(\lambda+1)^2}, \tag{7}$$

$$\mathrm{Cov}(Y_t, Y_{t+h}) = \alpha^h \mathrm{Var}(Y_t)r \tag{8}$$

and

$$\gamma_h = \alpha^h.$$

Proof: Using the general results related to mean, variance, covariance and correlation,

$$\begin{aligned}
\mathrm{E}(Y_t) &= \mathrm{E}(\mathrm{E}(Y_t|Y_{t-1})) \\
&= \mathrm{E}(\alpha Y_{t-1}) + \frac{\lambda+2}{\lambda(\lambda+1)} \\
&= \frac{\lambda+2}{\lambda(\lambda+1)(1-\alpha)}, \\
\mathrm{Var}(Y_t) &= \mathrm{Var}(\mathrm{E}(Y_t \mid Y_{t-1})) + \mathrm{E}(\mathrm{Var}(Y_t \mid Y_{t-1})) \\
&= \mathrm{Var}\left(\alpha Y_{t-1} + \frac{\lambda+2}{\lambda(\lambda+1)}\right) \\
&\quad + \mathrm{E}\left(\frac{\alpha(2\theta-\alpha(\theta+1))}{1-\theta} Y_{t-1} + \frac{\lambda^3+4\lambda^2+6\lambda+2}{\lambda^2(1+\lambda)^2}\right) \\
&= \frac{\begin{array}{l}(1-\alpha)(1-\theta)(\lambda^3+4\lambda^2+6\lambda+2) \\ +\alpha\lambda(\lambda+1)(\lambda+2)(2\theta-\alpha(\theta+1))\end{array}}{(1-\alpha)^2(\alpha+1)(1-\theta)\lambda^2(\lambda+1)^2}, \\
\mathrm{Cov}(Y_t, Y_{t-h}) &= \mathrm{Cov}(A(\alpha,\theta)\circ Y_{t-1} + \varepsilon_t, Y_{t-h}) \\
&= \mathrm{Cov}(A(\alpha,\theta)\circ Y_{t-1}, Y_{t-h}) + \mathrm{Cov}(\varepsilon_t, Y_{t-h}) \\
&= \alpha^h\gamma_0 = \alpha^h \mathrm{Var}(Y_t),
\end{aligned}$$

which directly prove the ACF, $\gamma_h = \mathrm{Corr}(Y_t, Y_{t-h}) = \alpha^h$. ■

Remark 3.1: Using (6) and (7), the DI is given as

$$\begin{aligned}
\mathrm{DI}(Y_t) &= \frac{\mathrm{Var}(Y_t)}{\mathrm{E}(Y_t)} \\
&= \frac{(1-\alpha)(1-\theta)(\lambda^3+4\lambda^2+6\lambda+2)+\alpha\lambda(\lambda+1)(\lambda+2)(2\theta-\alpha(\theta+1))}{(1-\alpha)(\alpha+1)(1-\theta)\lambda(\lambda+1)(\lambda+2)}.
\end{aligned}$$

We can show that the GBINAR(1)PL process can be under as well as over-dispersed.

But since over-dispersed datasets are more abundant than under-dispersed ones, we here focus our attention on over-dispersed models.

Lemma 3.3: *The transition probabilities for the GBINAR(1)PL process is given as*

$$P\{Y_{t+1}=k \mid Y_t=m\}=\left(\frac{1-\alpha}{1-\alpha\theta}\right)^m\left\{P\{\varepsilon_{t+1}=k\}+\sum_{j=1}^{k}P\{\varepsilon_{t+1}=k-j\}\left(\frac{(1-\alpha)\theta}{1-\alpha\theta}\right)^j\right.$$
$$\left.\times\sum_{i=0}^{m}\binom{m}{i}\binom{j-1}{i-1}\left(\frac{\alpha(1-\theta)^2}{(1-\alpha)^2\theta}\right)^i\right\}. \tag{9}$$

Proof: The transition probabilities are obtained as that of Poisson geometric in Aly and Bouzar (1994a), that is by taking the pgf of n-step convolution and then derive the transition probabilities which are given in Proposition 3.1 of the same. ■

4. Estimation of the parameters of GBINAR(1)PL

Assume that X_t is a strictly stationary ergodic Markov chain for the GBINAR(1)PL process. In this section, we aim to estimate the unknown parameters of the model. To estimate the parameters of the model, we will use the method of conditional maximum likelihood (CML) and the method of conditional least squares (CLS). Some asymptotic properties regarding the resulting estimators are also given in this section.

4.1. Conditional maximum-likelihood method

The CML method of estimation employs the conditional log likelihood function given by

$$\ell=\sum_{t=1}^{T}\log[P\{Y_{t+1} \mid Y_t\}],$$

where $P\{Y_{t+1} \mid Y_t\}$ is given in (8). The ℓ should be optimized to give the CML estimators, that is, $(\hat{\alpha}_{\mathrm{CML}},\hat{\theta}_{\mathrm{CML}},\hat{\lambda}_{\mathrm{CML}})$ of the parameters α,θ and λ. Since the optimization is analytically difficult, we make use of numerical techniques using R. Using PORT routines in R, the nlminb function is used to obtain the CML estimators of the parameters.

Theorem 4.1: *The CML estimators of the parameters* (α,λ,θ)*, denoted by* $(\hat{\alpha}_{\mathrm{CML}},\hat{\theta}_{\mathrm{CML}},\hat{\lambda}_{\mathrm{CML}})$*, are consistent and asymptotically normally distributed as*

$$\sqrt{T-1}\begin{pmatrix}\hat{\alpha}_{\mathrm{CML}}-\alpha\\ \hat{\theta}_{\mathrm{CML}}-\theta\\ \hat{\lambda}_{\mathrm{CML}}-\lambda\end{pmatrix}\xrightarrow{d}\mathrm{N}(\mathbf{0},\mathbf{I}^{-1}(\alpha,\theta,\lambda)),$$

where $\mathbf{I}(\alpha,\theta,\lambda)$ *denotes the Fisher information matrix.*

Proof: The consistency and asymptotic normality of CML estimators are demonstrated in Andersen (1970), Bu and McCabe (2008), Bu et al. (2008) under some standard regularity conditions. ■

4.2. Conditional least squares method

The CLS estimators of the unknown parameters α and λ can be obtained by minimizing the expression

$$H_1 = \sum_{t=2}^{T} (Y_t - E(Y_t \mid Y_{t-1}))^2. \tag{9}$$

Suppose $\mu = \mathrm{E}(Y_t)$. Then (9) can be written as

$$\begin{aligned} H_1 &= \sum_{t=2}^{T} \left(Y_t - \alpha Y_{t-1} - \frac{\lambda + 2}{\lambda(\lambda + 1)} \right)^2 \\ &= \sum_{t=2}^{T} (Y_t - \alpha Y_{t-1} - (1 - \alpha)\mu)^2. \end{aligned}$$

CLS estimators of α and λ, denoted by $\hat{\alpha}_{\mathrm{CLS}}$, $\hat{\lambda}_{\mathrm{CLS}}$ can be derived by solving

$$\frac{\partial H_1}{\partial \alpha} = 0 \quad \text{and} \quad \frac{\partial H_1}{\partial \lambda} = 0$$

for α and λ. The derived estimators for α, μ and hence λ are

$$\begin{aligned} \hat{\alpha}_{\mathrm{CLS}} &= \frac{(T-1)\sum_{t=2}^{T} Y_t Y_{t-1} - \sum_{t=2}^{T} Y_t \sum_{t=2}^{T} Y_{t-1}}{(T-1)\sum_{t=2}^{T} Y_{t-1}^2 - \left(\sum_{t=2}^{T} Y_{t-1}\right)^2}, \\ \hat{\mu}_{\mathrm{CLS}} &= \frac{\sum_{t=2}^{T} Y_t - \hat{\alpha}_{\mathrm{CLS}} \sum_{t=2}^{T} Y_{t-1}}{(1 - \hat{\alpha}_{\mathrm{CLS}})(T-1)}, \end{aligned}$$

and

$$\hat{\lambda}_{\mathrm{CLS}} = \frac{1 - (1 - \hat{\alpha}_{\mathrm{CLS}})\hat{\mu}_{\mathrm{CLS}} + \sqrt{((1 - \hat{\alpha}_{\mathrm{CLS}})\hat{\mu}_{\mathrm{CLS}} - 1)^2 + 8(1 - \hat{\alpha}_{\mathrm{CLS}})\hat{\mu}_{\mathrm{CLS}}}}{2(1 - \hat{\alpha}_{\mathrm{CLS}})\hat{\mu}_{\mathrm{CLS}}}.$$

Using (9), the estimator of θ cannot be acquired. However, by applying the two-step CLS method introduced by Karlsen and Tjøstheim (1988), the CLS estimator of θ, denoted as $\hat{\theta}_{\mathrm{CLS}}$, can be obtained. By two-step CLS method, $\hat{\theta}_{\mathrm{CLS}}$ can be obtained by minimizing the function

$$\begin{aligned} H_2 &= \sum_{t=2}^{T} [(Y_t - \mathrm{E}(Y_t \mid Y_{t-1}))^2 - \mathrm{Var}(Y_t \mid Y_{t-1})]^2 \\ &= \sum_{t=2}^{T} \left[\left(Y_t - \alpha Y_{t-1} - \frac{\lambda + 2}{\lambda(\lambda + 1)} \right)^2 - \frac{\alpha(2\theta - \alpha(\theta + 1))}{1 - \theta} Y_{t-1} \right. \\ &\quad \left. - \frac{\lambda^3 + 4\lambda^2 + 6\lambda + 2}{\lambda^2(1 - \lambda)^2} \right]^2. \end{aligned}$$

The estimates α and λ should be replaced by the estimators $\hat{\alpha}_{\mathrm{CLS}}$ and $\hat{\lambda}_{\mathrm{CLS}}$.

Theorem 4.2: *The CLS estimators* $(\hat{\alpha}_{\mathrm{CLS}}, \hat{\lambda}_{\mathrm{CLS}})$ *will be asymptotically multivariate normally distributed (MVN) as,*

$$\sqrt{T-1}\begin{pmatrix}\hat{\alpha}_{\mathrm{CLS}}-\alpha\\ \hat{\lambda}_{\mathrm{CLS}}-\lambda\end{pmatrix} \xrightarrow{d} \mathrm{MVN}(\mathbf{0}, \mathbf{V}^{-1}\mathbf{W}\mathbf{V}^{-1}),$$

where

$$\mathbf{W}=\begin{pmatrix}\mathrm{W}_1^2 & \mathrm{W}_{12}\\ \mathrm{W}_{12} & \mathrm{W}_2^2\end{pmatrix}, \quad \mathbf{V}=\begin{pmatrix}\mathrm{V}_{11} & \mathrm{V}_{12}\\ \mathrm{V}_{21} & \mathrm{V}_{22}\end{pmatrix}$$

as

$$\mathrm{W}_1^2=\mathrm{E}\left[Y_1^2\left(Y_2-\alpha Y_1-\frac{\lambda+2}{\lambda(\lambda+1)}\right)^2\right],$$

$$\mathrm{W}_{12}=\mathrm{E}\left[Y_1\left(\frac{2}{\lambda^2}-\frac{1}{(1+\lambda)^2}\right)\left(Y_2-\alpha Y_1-\frac{\lambda+2}{\lambda(\lambda+1)}\right)^2\right],$$

$$\mathrm{W}_2^2=\mathrm{E}\left[\left(\frac{2}{\lambda^2}-\frac{1}{(1+\lambda)^2}\right)^2\left(Y_2-\alpha Y_1-\frac{\lambda+2}{\lambda(\lambda+1)}\right)^2\right],$$

$$\mathrm{V}_{11}=\mathrm{E}(Y_1^2),\ \mathrm{V}_{12}=\mathrm{V}_{21}=\left(\frac{2}{\lambda^2}-\frac{1}{(1+\lambda)^2}\right)\mathrm{E}\left(Y_1\right),\quad \mathrm{V}_{22}=\left(\frac{2}{\lambda^2}-\frac{1}{(1+\lambda)^2}\right)^2,$$

and CLS estimator of θ, $\hat{\theta}_{\mathrm{CLS}}$ *will be having the approximate normal distribution as follows.*

$$\sqrt{T-1}(\hat{\theta}_{\mathrm{CLS}}-\theta) \xrightarrow{d} \mathrm{N}\left(0, \frac{\mathrm{U}^2}{\mathrm{C}^2}\right),$$

where

$$\mathrm{U}^2=\mathrm{E}\left\{\left[\left(Y_2-\alpha Y_1-\frac{\lambda+2}{\lambda(\lambda+1)}\right)^2-\frac{\alpha(1+\theta)(1-\alpha)}{1-\theta}Y_1-\frac{\lambda^3+4\lambda^2+6\lambda+2}{\lambda^2(1-\lambda)^2}\right]^2 \times\left[\frac{2Y_1(1-\alpha)\alpha}{(1-\theta)^2}\right]^2\right\},$$

and

$$\mathbf{V}=\mathrm{E}\left[\frac{2Y_1(1-\alpha)\alpha}{(1-\theta)^2}\right]^2.$$

Proof: Suppose $\mathcal{F}=\sigma\{Y_0, Y_1, \ldots\}$,

$$\begin{aligned}L_T&=-\frac{1}{2}\frac{\partial H_1}{\partial\alpha}\\ &=\sum_{t=2}^{T}Y_{t-1}\left(Y_t-\alpha Y_{t-1}-\frac{\lambda+2}{\lambda(\lambda+1)}\right)\end{aligned}$$

and $L_0 = 0$. Then,

$$\begin{aligned}\mathrm{E}(L_T \mid \mathcal{F}_{T-1}) &= \mathrm{E}\left[L_{T-1} + Y_{T-1}\left(Y_T - \alpha Y_{T-1} - \frac{\lambda+2}{\lambda(\lambda+1)}\right) \mid \mathcal{F}_{T-1}\right] \\ &= L_{T-1} + \mathrm{E}\left[Y_{T-1}\left(Y_T - \alpha Y_{T-1} - \frac{\lambda+2}{\lambda(\lambda+1)}\right) \mid \mathcal{F}_{T-1}\right] \\ &= L_{T-1},\end{aligned}$$

implying $\{L_T, \mathcal{F}_T, T \geq 0\}$ is a martingale, that is, it is a stochastic process where, given the sequence of all previous values, the conditional expectation of future values is equal to the current value. With the help of ergodic theorem (for more details refer Billingsley (1965)), since $\mathrm{E}|Y_T^4| < \infty$ and Y_t is strictly stationary, we can write,

$$\mathrm{E}\left[Y_1^2\left(Y_2 - \alpha - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2\right] < \infty,$$

and then

$$\frac{1}{T-1}\sum_{t=2}^{T} Y_{t-1}^2\left(Y_t - \alpha Y_{t-1} - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2 \xrightarrow{\text{a.s.}} \mathrm{E}\left[Y_1^2\left(Y_2 - \alpha - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2\right] = \mathrm{W}_1^2,$$

where $\xrightarrow{\text{a.s.}}$ denotes "converges almost surely". We proved $\{L_T\}$ is a martingale and hence by martingale central limit theorem,

$$\frac{1}{\sqrt{T-1}} L_T \xrightarrow{d} \mathrm{N}(0, \mathrm{W}_1^2).$$

Likewise, going through the same procedure, we can derive,

$$L'_T = -\frac{1}{2}\sum_{t=2}^{T}\frac{\partial H_1}{\partial \lambda} = \sum_{t=2}^{T}\left[\frac{1}{(1+\lambda)^2} - \frac{2}{\lambda^2}\right]\left[Y_t - \alpha Y_{t-1} - \frac{\lambda+2}{\lambda(\lambda+1)}\right]$$

is also a martingale. Then,

$$\begin{aligned}&\frac{1}{T-1}\sum_{t=2}^{T}\left[\frac{1}{(1+\lambda)^2} - \frac{2}{\lambda^2}\right]^2\left[Y_t - \alpha Y_{t-1} - \frac{\lambda+2}{\lambda(\lambda+1)}\right]^2 \\ &\xrightarrow{\text{a.s.}} \mathrm{E}\left[\left(Y_2 - \alpha Y_1 - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2\right] = \mathrm{W}_2^2\end{aligned}$$

leading to

$$\frac{1}{\sqrt{T-1}} L'_T \xrightarrow{d} \mathrm{N}(0, \mathrm{W}_2^2).$$

Similarly, for any

$$b = (b_1, b_2)^\top \in \mathbb{R}^2/(0,0),$$

we obtain,

$$\begin{aligned}
&\frac{1}{\sqrt{T-1}}\begin{pmatrix} b_1 & b_2 \end{pmatrix}\begin{pmatrix} L_T \\ L_T' \end{pmatrix} \\
&= \frac{1}{\sqrt{T-1}}\sum_{t=2}^{T}\left[b_1 Y_{t-1} + b_2\left(\frac{1}{(1+\lambda)^2} - \frac{2}{\lambda^2}\right)\right]\left[Y_t - \alpha Y_{t-1} - \frac{\lambda+2}{\lambda(\lambda+1)}\right] \\
&\xrightarrow{d} \mathrm{N}\left(0, \mathrm{E}\left\{\left[b_1 Y_{t-1} + b_2\left(\frac{1}{(1+\lambda)^2} - \frac{2}{\lambda^2}\right)\right]^2\left[Y_2 - \alpha Y_1 - \frac{\lambda+2}{\lambda(\lambda+1)}\right]^2\right\}\right).
\end{aligned}$$

Then, by Cramer Wold device (Cramér & Wold, 1936),

$$\frac{1}{\sqrt{T-1}}\begin{pmatrix} L_T \\ L_T' \end{pmatrix} \xrightarrow{d} \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \mathrm{W}_1^2 & \mathrm{W}_{12} \\ \mathrm{W}_{12} & \mathrm{W}_2^2 \end{pmatrix}\right).$$

Moreover, due to the strict stationarity of $\{Y_t\}$ and after some simple algebra,

$$\sqrt{T-1}\begin{pmatrix} \hat{\alpha}_{\mathrm{CLS}} - \alpha \\ \hat{\lambda}_{\mathrm{CLS}} - \lambda \end{pmatrix} \xrightarrow{d} \mathrm{MVN}(\mathbf{0}, \mathbf{V}^{-1}\mathbf{W}\mathbf{V}^{-1}),$$

where

$$\mathbf{W} = \begin{pmatrix} \mathrm{W}_1^2 & \mathrm{W}_{12} \\ \mathrm{W}_{12} & \mathrm{W}_2^2 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} \mathrm{V}_{11} & \mathrm{V}_{12} \\ \mathrm{V}_{21} & \mathrm{V}_{22} \end{pmatrix},$$

as

$$\begin{aligned}
\mathrm{W}_1^2 &= \mathrm{E}\left[Y_1^2\left(Y_2 - \alpha Y_1 - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2\right], \\
\mathrm{W}_{12} &= \mathrm{E}\left[Y_1\left(\frac{2}{\lambda^2} - \frac{1}{(1+\lambda)^2}\right)\left(Y_2 - \alpha Y_1 - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2\right], \\
\mathrm{W}_2^2 &= \mathrm{E}\left[\left(\frac{2}{\lambda^2} - \frac{1}{(1+\lambda)^2}\right)^2\left(Y_2 - \alpha Y_1 - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2\right], \\
\mathrm{V}_{11} &= \mathrm{E}\left(Y_1^2\right), \quad \mathrm{V}_{12} = \mathrm{V}_{21} = \left(\frac{2}{\lambda^2} - \frac{1}{(1+\lambda)^2}\right)\mathrm{E}\left(Y_1\right), \quad \mathrm{V}_{22} = \left(\frac{2}{\lambda^2} - \frac{1}{(1+\lambda)^2}\right)^2.
\end{aligned}$$

Considering the determinant of the matrix $\mathbf{V}$,

$$\mathrm{Det}(\mathbf{V}) = \left(\frac{1}{(1+\lambda)^2} - \frac{2}{\lambda^2}\right)^2 \mathrm{Var}(Y_1) \geq 0,$$

implying $\mathbf{V}$ is invertible. Now, to prove the asymptotic normal approximation of CLS estimator of θ, $\hat{\theta}_{\mathrm{CLS}}$, suppose further

$$\mathcal{F}' = \sigma\{Y_0, Y_1, \ldots, Y_T\},$$

$$L'_T = \frac{1}{2}\frac{\partial H_2}{\partial \theta}$$
$$= \sum_{t=2}^{T}\left[\left(Y_t - \alpha Y_{t-1} - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2 - \frac{\alpha(2\theta-\alpha(\theta+1))}{1-\theta}Y_{t-1} - \frac{\lambda^3+4\lambda^2+6\lambda+2}{\lambda^2(\lambda+1)^2}\right] \times \frac{2Y_{t-1}(1-\alpha)\alpha}{(1-\theta)^2},$$

and $L'_T = 0$. Now,

$$\mathrm{E}(L'_T \mid \mathcal{F}'_{T-1}) = \mathrm{E}\left\{L'_{T-1} + \left[\left(Y_T - \alpha Y_{T-1} - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2 - \frac{\alpha(2\theta-\alpha(\theta+1))}{1-\theta}Y_{T-1} - \frac{\lambda^3+4\lambda^2+6\lambda+2}{\lambda^2(\lambda+1)^2}\right]\frac{2Y_{T-1}(1-\alpha)\alpha}{(1-\theta)^2} \mid \mathcal{F}'_{T-1}\right\}$$
$$= L_{T-1}.$$

$\{L'_T, \mathcal{F}', T \geq 0\}$ is a martingale. Also since, $\mathrm{E}|Y_t^6| \leq \infty$, due to the strict stationarity of $\{Y_t\}$, and by the ergodic theorem, we obtain

$$\mathrm{E}\left\{\left[\left(Y_2 - \alpha Y_1 - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2 - \frac{\alpha(2\theta-\alpha(\theta+1))}{1-\theta}Y_1 - \frac{\lambda^3+4\lambda^2+6\lambda+2}{\lambda^2(\lambda+1)^2}\right]^2\left[\frac{2Y_1(1-\alpha)\alpha}{(1-\theta)^2}\right]^2\right\} < \infty$$

and then

$$\frac{1}{T-1}\sum_{t=2}^{T}\left\{\left[\left(Y_t - \alpha Y_{T-1} - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2 - \frac{\alpha(2\theta-\alpha(\theta+1))}{1-\theta}Y_{t-1} - \frac{\lambda^3+4\lambda^2+6\lambda+2}{\lambda^2(\lambda+1)^2}\right]^2\left[\frac{2Y_{t-1}(1-\alpha)\alpha}{(1-\theta)^2}\right]^2\right\}$$
$$\overset{\text{a.s.}}{\to} \mathrm{E}\left\{\left[\left(Y_2 - \alpha Y_1 - \frac{\lambda+2}{\lambda(\lambda+1)}\right)^2 - \frac{\alpha(2\theta-\alpha(\theta+1))}{1-\theta}Y_1 - \frac{\lambda^3+4\lambda^2+6\lambda+2}{\lambda^2(\lambda+1)^2}\right]^2\left[\frac{2Y_1(1-\alpha)\alpha}{(1-\theta)^2}\right]^2\right\} = \mathrm{U}^2.$$

L'_T is proved to be martingale and hence by martingale central limit theorem,

$$\frac{1}{\sqrt{L_T}} \overset{d}{\to} \mathrm{N}(0, \mathrm{U}^2).$$

Moreover, due to the strict stationarity of $\{Y_t\}$ and after some algebra, we obtain

$$\sqrt{T-1}(\hat{\theta}_{\mathrm{CLS}} - \theta) \overset{d}{\to} \mathrm{N}\left(0, \frac{\mathrm{U}^2}{\mathrm{C}^2}\right),$$

where $\mathrm{C}^2 = \mathrm{E}\left(\frac{2Y_1(1-\alpha)\alpha}{(1-\theta)^2}\right) \geq 0$. ■

Table 1. Simulation results for $\alpha = 0.2, \theta = 0.5, \lambda = 0.1$.

		CML		CLS	
Sample size (n)	Parameters	Bias	MSE	Bias	MSE
	α	0.0243	0.0080	−0.0631	0.0437
25	θ	−0.0069	0.0091	0.0315	0.0156
	λ	0.0368	0.0206	0.0122	0.0319
	α	0.0084	0.0025	−0.0330	0.0196
50	θ	−0.0085	0.0050	0.0025	0.0010
	λ	0.0151	0.0080	−0.0009	0.0098
	α	0.0047	0.0004	−0.0165	0.0099
100	θ	−0.0108	0.0030	−0.0561	0.0536
	λ	0.0051	0.0037	−0.0010	0.0049
	α	−0.0016	0.0003	−0.0108	0.0054
200	θ	−0.0014	0.0029	−0.0100	0.0126
	λ	0.0057	0.0018	−0.0003	0.0025
	α	−0.0008	0.0002	−0.0069	0.0026
400	θ	−0.0002	0.0025	−0.0036	0.0011
	λ	0.0016	0.0008	−0.0012	0.0012

Table 2. Simulation results for $\alpha = 0.5, \theta = 0.6, \lambda = 0.2$.

		CML		CLS	
Sample size (n)	Parameters	Bias	MSE	Bias	MSE
	α	0.0242	0.0646	−0.0979	0.0432
25	θ	0.1290	0.0832	0.1291	0.1233
	λ	−0.0432	0.0629	−0.0172	0.0577
	α	0.0235	0.0557	−0.0573	0.0215
50	θ	0.0998	0.0245	0.2039	0.1343
	λ	−0.0418	0.0122	−0.0198	0.0247
	α	0.0232	0.0548	−0.0272	0.0089
100	θ	0.0956	0.0241	0.3068	0.1551
	λ	−0.0368	0.0080	−0.0089	0.0109
	α	0.0225	0.0521	−0.0114	0.0044
200	θ	0.0599	0.0142	0.3747	0.1696
	λ	−0.0270	0.0050	−0.0003	0.0059
	α	0.0218	0.0500	−0.0070	0.0020
400	θ	0.0351	0.0140	0.4094	0.1798
	λ	−0.0174	0.0028	−0.0002	0.0029

5. Simulation study

The estimation methods CML and CLS are further analysed using the simulation study. For that purpose two sets of parameter values ($\alpha = 0.2, \theta = 0.5, \lambda = 0.1$) and ($\alpha = 0.5, \theta = 0.4, \lambda = 0.2$) are used. For each parameter set, $N = 1000$ replications of random samples of size $n = 25$, 50, 100, 200, and 400 were taken and the bias and mean square errors (MSEs) are calculated for the estimators. Tables 1 and 2 present the simulation results.

Tables 1 and 2 demonstrate that both methods perform quite similarly in estimating the parameters. Additionally, both methods exhibit a decrease in bias and MSE in most cases.

6. Empirical data analysis

A possible application of the proposed process GBINAR(1)PL is discussed in this section using a real data set. The data set used is the number of submissions to animal health laboratories, monthly January 2003 to December 2009, from a region in New Zealand mentioned in Jazi et al. (2012). The submissions contain several classifications for considering symptoms.

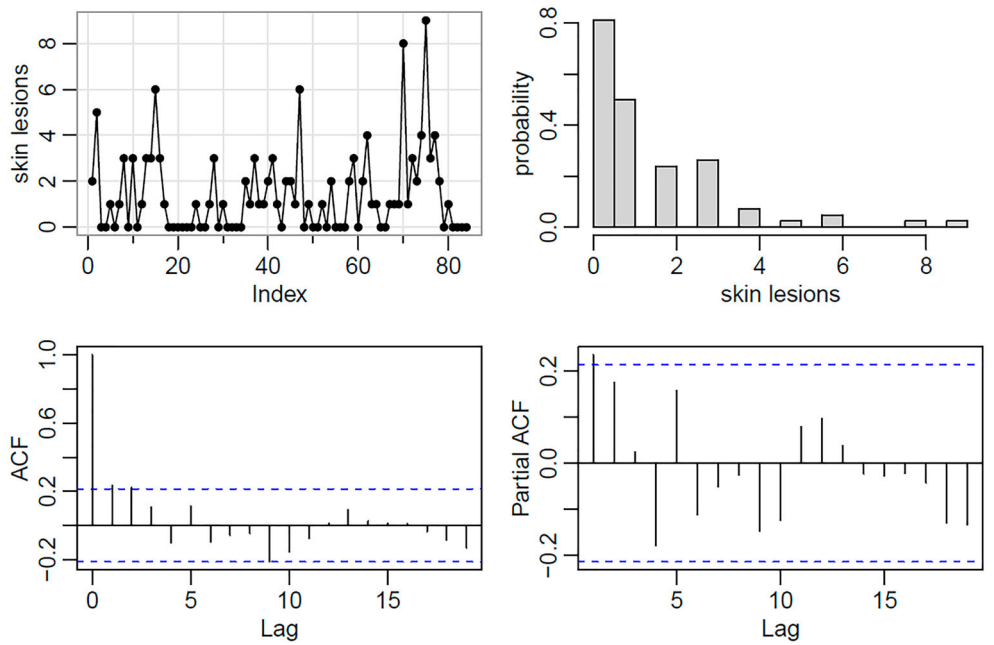

Figure 1. The time series plot, histogram, ACF and PACF of the skin lesions data.

We took the skin lesions data. The mean, variance and DI of the data set used are 1.43, 3.36 and 2.35, respectively. The test of Schweer and Weiß (2014) shows p-value less than 0.001 implying the data set has significant over-dispersion. The time series plot, histogram, ACF and partial ACF (PACF) of the data set are plotted in Figure 1. The PACF plot indicates that only first lag is significant which proves this data can be used for modelling INAR(1) process.

Using this data set we illustrate the better performance of GBINAR(1)PL over INAR(1) with PL innovations based on binomial thinning denoted as, BINAR(1)PL (Lívio et al., 2018) and INAR(1) with PL innovations based on negative binomial thinning denoted as NBINAR(1)PL (Rostami et al., 2018). Also, we used comparison measures such as -log likelihood (-L), Akaike information criterion (AIC), Bayesian information criterion (BIC) and root mean square error (RMSE). The RMSEs represent the square root of sum of squared differences between true values and one step conditional expectations. Furthermore, residual analysis is conducted to assess whether the fitted GBINAR(1)PL process is statistically accurate. For that, Pearson residuals for the GBINAR(1)PL are calculated using

$$e_t = \frac{y_t - \mathrm{E}(Y_t = y_t \mid Y_{t-1} = y_{t-1})}{\mathrm{Var}(Y_t = y_t \mid Y_{t-1} = y_{t-1})^{1/2}},$$

where $\mathrm{E}(Y_t \mid Y_{t-1})$ and $\mathrm{Var}(Y_t \mid Y_{t-1})$ are given in (4) and (5), respectively. The statistical validity of the fitted INAR(1) is proved by acquiring zero mean and unit variance for the uncorrelated Pearson residuals (Harvey & Fernandes, 1989). The fitted GBINAR(1)PL process, BINAR(1)PL and NBINAR(1)PL used for comparison yield the parameter estimates along with SE, -L, AIC, BIC and RMSEs as given in Table 3. The minimum values for -L, AIC, BIC prove GBINAR(1)PL has better performance than other thinnings with PL innovations. Hence, it is convincing that GBINAR(1)PL explains the characteristics of the data set very effectively. The mean and variance of the Pearson residuals of the GBINAR(1)PL process

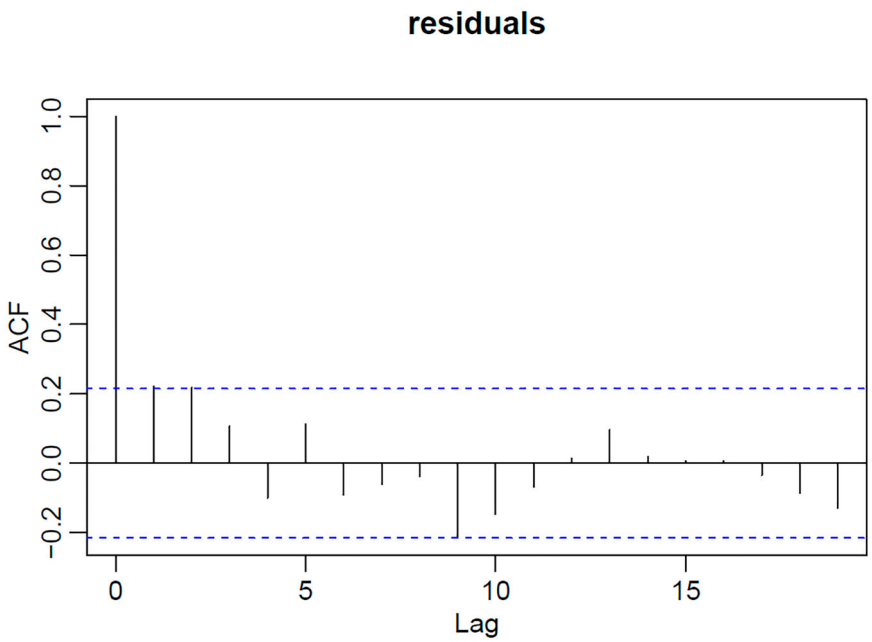

Figure 2. The ACF plot of the Pearson residuals for skin lesions data set.

Table 3. The estimates and modelling adequacy statistics of the fitted distributions for the Skin lesions data.

Model	Parameters	Estimates	SE	-*L*	AIC	BIC	RMSE
	α	0.3835	0.1600	132.7932	271.5865	278.8790	1.82118
GBINAR(1)PL	θ	0.6922	0.1284				
	λ	1.5952	0.3459				
BINAR(1)PL	α	0.1116	0.0769	135.3743	274.7485	279.6102	1.82115
	λ	1.1647	0.1607				
NBINAR(1)PL	α	0.1726	0.1242	135.0272	274.0544	278.9160	1.82116
	λ	1.2391	0.2142				

as 0.0124 and 1.2327, are very close to the desired values, which proves our GBINAR(1)PL process is statistically valid for the data set. Then according to the results of Jazi et al. (2012), the GBINAR(1)PL process for the data is such that,

$$Y_t = A(0.3835, 0.6922) \circ Y_{t-1} + \varepsilon_t,$$

where the innovation process is

$$\varepsilon_t \sim \text{PL}(1.5952).$$

Furthermore, the ACF plot of the Pearson residuals in Figure 2 specifies that there is no presence of autocorrelation between them.

7. Concluding remarks

A new generalized binomial thinning operator having the commonly used binomial thinning operator as a special case is introduced in this paper. Its properties are derived and come out

to be comparatively simple. Then a GBINAR(1) process is defined with its properties. In addition, PL distribution is used as innovation distribution and hence we obtain GBINAR(1)PL process. The estimation of unknown parameters is performed using CML and CLS methods and also asymptotic properties of these estimates are derived. A simulation study proves that both methods are effective. A real-life count data set is considered to prove the applicability of the process GBINAR(1)PL. Hence, the new generalized binomial thinning operator is an effective generalization to the binomial thinning, and also the GBINAR(1)PL process proves to be effective in modelling count datasets.

Conflicts of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Acknowledgments

The authors gratefully acknowledge the editor and referees for their careful reading and insightful comments, which greatly improved the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Muhammed Rasheed Irshad http://orcid.org/0000-0002-5999-1588

References

Al-Osh, M. A., & Aly, E. E. A. (1992). First order autoregressive time series with negative binomial and geometric marginals. *Communications in Statistics-Theory and Methods*, *21*(9), 2483–2492. https://doi.org/10.1080/03610929208830925

Al-Osh, M. A., & Alzaid, A. A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis*, *8*(3), 261–275. https://doi.org/10.1111/j.1467-9892.1987.tb00438.x

Altun, E., Bhati, D., & Khan, N. M. (2021). A new approach to model the counts of earthquakes: INARPQX(1) process. *SN Applied Sciences*, *3*(1), 1–17. https://doi.org/10.1007/s42452-020-04109-8

Aly, E. E. A., & Bouzar, N. (1994a). Explicit stationary distributions for some Galton-Watson processes with immigration. *Stochastic Models*, *10*(2), 499–517. https://doi.org/10.1080/15326349408807305

Aly, E. E. A., & Bouzar, N. (1994b). On some integer-valued autoregressive moving average models. *Journal of Multivariate Analysis*, *50*(1), 132–151. https://doi.org/10.1006/jmva.1994.1038

Aly, E. E. A., & Bouzar, N. (2019). Expectation thinning operators based on linear fractional probability generating functions. *Journal of the Indian Society for Probability and Statistics*, *20*(1), 89–107. https://doi.org/10.1007/s41096-018-0056-x

Andersen, E. B. (1970). Asymptotic properties of conditional maximum-likelihood estimators. *Journal of the Royal Statistical Society: Series B :Statistical Methodology*, *32*(2), 283–301. https://doi.org/10.1111/j.2517-6161.1970.tb00842.x

Billingsley, P. (1965). *Ergodic Theory and Information* (Vol. 1). Wiley.

Borges, P., Molinares, F. F., & Bourguignon, M. (2016). A geometric time series model with inflated-parameter Bernoulli counting series. *Statistics & Probability Letters*, *119*, 264–272. https://doi.org/10.1016/j.spl.2016.08.012

Bu, R., & McCabe, B. (2008). Model selection, estimation and forecasting in INAR(p) models: A likelihood-based Markov chain approach. *International Journal of Forecasting*, *24*(1), 151–162. https://doi.org/10.1016/j.ijforecast.2007.11.002

Bu, R., McCabe, B., & Hadri, K. (2008). Maximum likelihood estimation of higher-order integer-valued autoregressive processes. *Journal of Time Series Analysis*, *29*(6), 973–994. https://doi.org/10.1111/j.1467-9892.2008.00590.x

Cramér, H., & Wold, H. (1936). Some theorems on distribution functions. *Journal of the London Mathematical Society*, *s1-11*(4), 290–294. https://doi.org/10.1112/jlms/s1-11.4.290

Eliwa, M. S., Altun, E., El-Dawoody, M., & El-Morshedy, M. (2020). A new three-parameter discrete distribution with associated INAR(1) process and applications. *IEEE Access*, *8*, 91150–91162. https://doi.org/10.1109/ACCESS.2020.2993593

Ghitany, M. E., & Al-Mutairi, D. K. (2009). Estimation methods for the discrete Poisson-Lindley distribution. *Journal of Statistical Computation and Simulation*, *79*(1), 1–9. https://doi.org/10.1080/00949650701550259

Harvey, A. C., & Fernandes, C. (1989). Time series models for count or qualitative observations. *Journal of Business & Economic Statistics*, *7*(4), 407–417. https://doi.org/10.1080/07350015.1989.10509750

Huang, J., & Zhu, F. (2021). A new first-order integer-valued autoregressive model with Bell innovations. *Entropy*, *23*(6), 1-17. https://doi.org/10.3390/e23060713

Irshad, M. R., Chesneau, C., D'cruz, V., & Maya, R. (2021). Discrete pseudo Lindley distribution: Properties, estimation and application on INAR(1) process. *Mathematical and Computational Applications*, *26*(4), 1-22. https://doi.org/10.3390/mca26040076

Jazi, M. A., Jones, G., & Lai, C. D. (2012). First-order integer valued AR processes with zero inflated Poisson innovations. *Journal of Time Series Analysis*, *33*(6), 954–963. https://doi.org/10.1111/j.1467-9892.2012.00809.x

Kang, Y., Wang, D., & Yang, K. (2020). Extended binomial AR(1) processes with generalized binomial thinning operator. *Communications in Statistics-Theory and Methods*, *49*(14), 3498–3520. https://doi.org/10.1080/03610926.2019.1589519

Karlsen, H., & Tjøstheim, D. (1988). Consistent estimates for the NEAR(2) and NLAR(2) time series models. *Journal of the Royal Statistical Society: Series B (Methodological)*, *50*(2), 313–320. https://doi.org/10.1111/j.2517-6161.1988.tb01730.x

Khoo, W. C., Ong, S. H., & Biswas, A. (2017). Modeling time series of counts with a new class of INAR(1) model. *Statistical Papers*, *58*(2), 393–416. https://doi.org/10.1007/s00362-015-0704-0

Liu, Z., & Zhu, F. (2020). A new extension of thinning-based integer-valued autoregressive models for count data. *Entropy*, *23*(1), 1-17. https://doi.org/10.3390/e23010062

Lívio, T., Khan, N. M., Bourguignon, M., & Bakouch, H. S. (2018). An INAR(1) model with Poisson-Lindley innovations. *Economics Bulletin*, *38*(3), 1505–1513.

Maya, R., Jodrá, P, Aswathy, S, & Irshad, M. R. (2024) The discrete new XLindley distribution and the associated autoregressive process. *International Journal of Data Science and Analytics*, *20*, 1767–1793. https://doi.org/10.1007/s41060-024-00563-4

McKenzie, E. (1985) Some simple models for discrete variate time series. *JAWRA Journal of the American Water Resources Association*, *21*(4), 645–650. https://doi.org/10.1111/j.1752-1688.1985.tb05379.x

Ristić, M. M., Bakouch, H. S., & Nastić, A. S. (2009). A new geometric first-order integer-valued autoregressive (NGINAR(1)) process. *Journal of Statistical Planning and Inference*, *139*(7), 2218–2226. https://doi.org/10.1016/j.jspi.2008.10.007

Rostami, A., Mahmoudi, E, & Roozegar, R (2018) A new integer valued AR(1) process with Poisson-Lindley innovations. arXiv: 1802.00994

Sankaran, M. (1970). The discrete Poisson-Lindley distribution. *Biometrics*, *26*(1), 145–149. https://doi.org/10.2307/2529053

Schweer, S., & Weiß, C. H. (2014). Compound Poisson INAR(1) processes: Stochastic properties and testing for overdispersion. *Computational Statistics & Data Analysis*, *77*, 267–284. https://doi.org/10.1016/j.csda.2014.03.005

Steutel, F. W., & van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability*, *7*(5), 893–899. https://doi.org/10.1214/aop/1176994950

Weiß, C. H. (2008) Thinning operations for modelling time series of counts – a survey. *AStA Advances in Statistical Analysis*, *92*, 319–341. https://doi.org/10.1007/s10182-008-0072-3

Yang, K., Kang, Y, Wang, D, Li, H, & Diao, Y. (2019) Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued autoregressive processes. *Metrika*, *82*, 863–889. https://doi.org/10.1007/s00184-019-00714-9

Zhang, H., Wang, D., & Zhu, F. (2010). Inference for INAR(p) processes with signed generalized power series thinning operator. *Journal of Statistical Planning and Inference*, *140*(3), 667–683. https://doi.org/10.1016/j.jspi.2009.08.012