Article

A Noncentral Lindley Construction Illustrated in an INAR(1) Environment

Johannes Ferreira [1,2,*] and **Ané van der Merwe** [1]

[1] Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria 0028, South Africa; ane.neethling@up.ac.za
[2] Centre of Excellence in Mathematical and Statistical Sciences, Johannesburg 2050, South Africa
* Correspondence: johan.ferreira@up.ac.za

Abstract: This paper proposes a previously unconsidered generalization of the Lindley distribution by allowing for a measure of noncentrality. Essential structural characteristics are investigated and derived in explicit and tractable forms, and the estimability of the model is illustrated via the fit of this developed model to real data. Subsequently, this model is used as a candidate for the parameter of a Poisson model, which allows for departure from the usual equidispersion restriction that the Poisson offers when modelling count data. This Poisson-noncentral Lindley is also systematically investigated and characteristics are derived. The value of this count model is illustrated and implemented as the count error distribution in an integer autoregressive environment, and juxtaposed against other popular models. The effect of the systematically-induced noncentrality parameter is illustrated and paves the way for future flexible modelling not only as a standalone contender in continuous Lindley-type scenarios but also in discrete and discrete time series scenarios when the often-encountered equidispersed assumption is not adhered to in practical data environments.

Keywords: compounding; maximum likelihood; moments; Poisson; transition probability

1. Introduction

The analysis of time series has enjoyed sustained interest, in particular the case when the data generating process emanates from an underlying counting model, resulting in the considered innovation process to follow some stochastic discrete distribution. Practical examples of such scenarios include the number of daily transactions on a stock exchange, number of days patients spend in a health care facility, or the daily counts of positive COVID-19 cases in a specific geographical area.

Under this discrete innovation assumption, the often considered assumption of normality leaves the practitioner in an unideal position due to this continuous assumption for a process which is inherently discrete in nature. In particular, Ref. [1] provides a seminal guideline on the use and implementation of a discrete-based time series analysis with discrete innovations, and various authors (such as [2–5] and others) have considered the first-order integer-valued (stationary) autoregressive process (INAR(1)) as a more viable contender compared to the usual (continuous) first-order autoregressive process (AR(1)), where the error term $\epsilon_t > 0$ is oftentimes characterised by a Poisson random variable. Suppose that X_t represents the (non-negative) discrete time series count at time t ($t = 0, 1, 2, \ldots$), then the INAR(1) process is defined by

$$X_t = p \circ X_{t-1} + \epsilon_t, t \in \mathbb{Z} \tag{1}$$

where $p \in (0, 1)$ and $\circ$ denote the binomial thinning operator, defined by $p \circ X_{t-1} := \sum_{i=1}^{X_{t-1}} W_i$ where W_i is a Bernoulli random variable such that $P(W_i = 1) = 1 - P(W_i = 0) = p$. For the innovations ϵ_t in (1), the mean is indicated by μ_ϵ and finite variance by σ_ϵ^2; and for the INAR(1) model itself in (1), the mean is given by $\mu_{X_t} = \frac{\mu_\epsilon}{1-p}$ and the variance

by $\sigma_{X_t}^2 = \frac{\sigma_\epsilon^2 + p\mu_\epsilon}{1 - p^2}$ (see also [1,6,7]). The choice of the Poisson model for ϵ_t was introduced by [8], but has been shown to be inefficient for accurate inference due to the equidispersion property, where $\mu_\epsilon = \sigma_\epsilon^2$ of the Poisson distribution.

To address this, Refs. [6,9–11] introduced alternative considerations for the choice of ϵ_t, some of which emanates from the Lindley distribution or its generalisations incorporated into a discrete model for the INAR(1) innovation process. The Lindley distribution in itself has shown significant versatility as a continuous lifetime model, and a variety of generalisations of this model has been studied and reported on in literature (see [12–14]). In addition to this, the incorporation of the Lindley distribution and its generalisations within a Poisson framework to obtain generalised discrete models has also experienced a sustained interest, mainly due to the attractive feature of being able to model departures that data might exhibit from equidispersion. In this study, a meaningful noncentrality parameter is systematically induced into the Lindley distribution from where a discrete counting model is also derived. This counting model is then illustrated as a candidate for ϵ_t within a discrete time series environment. In this way, this work enriches the existing literature with a previously unconsidered noncentral candidate with a parameter that has leverage and meaning within understanding and modelling dispersion in discrete models, apart from a standalone contribution to the ever-growing literature on statistical distribution theory.

Therefore the contribution of this paper is threefold:

1. A noncentral Lindley distribution is systematically constructed and some statistical characteristics are derived;
2. A discrete counting model based on this noncentral Lindley distribution is derived via compounding with the Poisson distribution, together with essential statistical characteristics;
3. This discrete counting model is implemented and illustrated as innovation structure (i.e., ϵ_t) within an INAR(1) time series environment, and juxtaposed against often-considered contenders for the innovation structures.

The particular value of interpretable and systematically induced noncentrality in these contributions is highlighted. These contributions pave the way for flexible statistical modelling in practice not only for meaningful inference, but also continuing the understanding, expansion, and development of the statistical theoretic universe. These models, as competitors, exhibits meaningful and tractable choices for the practitioner to employ within not only usual positive real-data modelling, but also discrete modelling, and thirdly, within an INAR counting time series environment.

The paper is outlined as follows. In Section 2, the systematic construction of the noncentral Lindley (ncL) distribution is discussed and characteristics of this model are derived, followed by the construction and characteristics of the Poisson-noncentral Lindley (PncL) model. Section 3 describes and contains the modelling results of applying the newly developed models to real data, including a fitting and discussion around the implementation of the PncL within the INAR(1) (1) model. The final thoughts are contained in Section 4.

2. Building Blocks of New Model

The Lindley distribution (initially introduced by [15]) has seen a resurgence of attention over the span of the last 15 years, which could be attributed to its elegant construction as a weighted two-mixture distribution. This mixture consists of an exponential distribution with parameter $\beta > 0$ as the first component (with density function denoted by $f_{exp}(y)$), and a gamma distribution with scale and shape parameters $(\beta, 2)$ (with density function denoted by $f_{gam}(y)$). The weight function for this two-mixture distribution is given by $\omega = \frac{\beta}{1+\beta}$. In this paper, we consider a generalised weight function with additional parameters α and θ for additional flexibility; that is, $\omega = \frac{\alpha\beta}{\alpha\beta + \theta}$ where $\theta > 0$ and $\alpha\beta + \theta > 0$ (see [3,16]). This leaves a random variable Y with density function:

$$f(y) \quad = \omega f_{exp}(y) + (1 - \omega) f_{gam}(y)$$

$$= \frac{\beta^2}{\alpha\beta + \theta} \exp(-\beta y)(\alpha + y\theta), \quad y > 0. \tag{2}$$

Our point of departure for the contribution of this paper is to consider a noncentral gamma distribution with noncentral parameter $\lambda > 0$ instead of the usual gamma choice in (2). This leads to our proposed noncentral Lindley construction (ncL) construction, defined below.

2.1. A Noncentral Lindley Construction

We construct an ncL distribution by considering a random variable Y which follows the noncentral gamma distribution with scale, shape, and noncentral parameters $(\beta, 2, \lambda)$ with density function (see [17]):

$$f_{ncgam}(y) = \sum_{i=1}^{\infty} \frac{\exp(-\frac{\lambda}{2})(\frac{\lambda}{2})^i}{i!} \frac{\beta^{2+i}}{\Gamma(2+i)} y^{2+i-1} \exp(-\beta y), \quad y > 0 \tag{3}$$

where $\Gamma(\cdot)$ denotes the gamma function. Various authors (see for example, [17–20]) have considered and employed the noncentral gamma distribution as an essential generalisation of the gamma distribution, and also as a direct generalisation which emanates from the popular noncentral chi-square distribution. In particular, the hierarchical representation of (3) is theoretically attractive as a discrete mixture of the (usual) gamma distribution.

The consideration of this noncentral gamma distribution (3) as an enriched mixture component in (2), instead of the usual gamma distribution, is thus a systematic and organic parameter enrichment to result in a noncentral Lindley alternative. In this case, the density function of the resulting ncL distribution is given by:

$$f(y) \quad = \omega f_{exp}(y) + (1 - \omega) f_{ncgam}(y)$$

$$= \frac{\beta^2}{\alpha\beta + \theta} \exp(-\beta y) \left[\alpha + \theta y \exp\left(-\frac{\lambda}{2}\right) \sum_{i=0}^{\infty} \frac{(\frac{\lambda}{2}\beta y)^i}{i!} \frac{1}{(2)_i} \right] \tag{4}$$

$$= \frac{\beta^2}{\alpha\beta + \theta} \exp(-\beta y) \left[\alpha + y\theta \exp\left(-\frac{\lambda}{2}\right) {}_0F_1\left(2; \frac{\lambda}{2}\beta y\right) \right]$$

where ${}_0F_1(\cdot)$ denotes the confluent hypergeometric function, and $(a)_j$ denotes the Pochammer coefficient (see [21]). In this case, we denote $Y \sim \text{ncL}(\beta, \alpha, \theta, \lambda)$. When $\lambda = 0$, (5) reduces to the considered generalised Lindley model of [3] with density function (2). Furthermore, when $\alpha = \beta = 1$, (5) reflects the usual Lindley density function (see [15]).

The distribution function, moment generating function, and moments of $Y \sim \text{ncL}(\beta, \alpha, \theta, \lambda)$ with density (5) is presented in the following theorem. The proof is contained in Appendix A.

Theorem 1. *Suppose that the random variable Y is distributed as ncL with density function (5). Then, the distribution function, moment generating function, and moments of Y is given by:*

1.

$$F_Y(y) = \omega(1 - \exp(-\beta y)) + (1 - \omega) \exp\left(-\frac{\lambda}{2}\right) \sum_{i=0}^{\infty} \frac{(\frac{\lambda}{2})^i}{(2)_i} \gamma(i + 2, \beta y), \tag{5}$$

2.

$$M_Y(t) = \frac{\beta^2}{(\alpha\beta + \theta)(\beta - t)^2} \left(\alpha(\beta - t) + \theta \exp\left(-\frac{\lambda}{2}\left(\frac{t}{\beta - t}\right)\right) \right), \tag{6}$$

3.

$$E(Y^r) = \frac{\beta^2}{\alpha\beta + \theta} \frac{\alpha\Gamma(r+1)}{\beta^{r+1}} + \frac{\theta\exp(-\frac{\lambda}{2})\Gamma(r+2)}{(\alpha\beta+\theta)\beta^r} {}_1F_1\left(r+2; 2; \frac{\lambda}{2}\right) \tag{7}$$

where $r > 0$, $t < \beta$, $\gamma(\cdot, \cdot)$ denotes the incomplete lower gamma function (see [21]) and ${}_1F_1(\cdot)$ represents the hypergeometric function with one lower and one upper parameter (see [21]). Note that $\gamma(i+2, \beta y)$ in (5) may be simplified further into a closed form as $(i+1)!\left(1 - \exp(-\beta y)\sum_{m=0}^{i+1} \frac{(\beta y)^m}{m!}\right)$ using [21], equation 8.352.6.

These expressions exhibit the elegant incorporation of the noncentral parameter λ, and using all three characteristics (5), (6), and (7) by setting $\lambda = 0$, the corresponding characteristics of the usual generalised Lindley of [22] (see also [3]) may be obtained.

In Figure 1, some particular shapes of the density function (5) are illustrated for various combinations of parameters. The following effects are observed for changes in the different parameters:

- λ affects the location and spread of the distribution. As λ increases (comparing Figure 1a–c relative to Figure 1d–f), the distribution flattens out, shifts to the right and becomes less skewed;
- β affects the scale of the distribution. As β increases, the scale of the distribution increases (comparing the densities in Figure 1a and Figure 1d, respectively);
- α and θ affect the shape of the distribution;
  - As α increases positively (refer to Figure 1b and Figure 1e, respectively), the distribution tends to the shape of an exponential distribution;
  - In contrast to the latter, the shape of the distribution tends to that similar to an exponential distribution as θ tends to zero (refer to Figure 1c and Figure 1f, respectively);
  - The origin of the distribution is shifted for $y > 0$ when $\alpha > 0$.

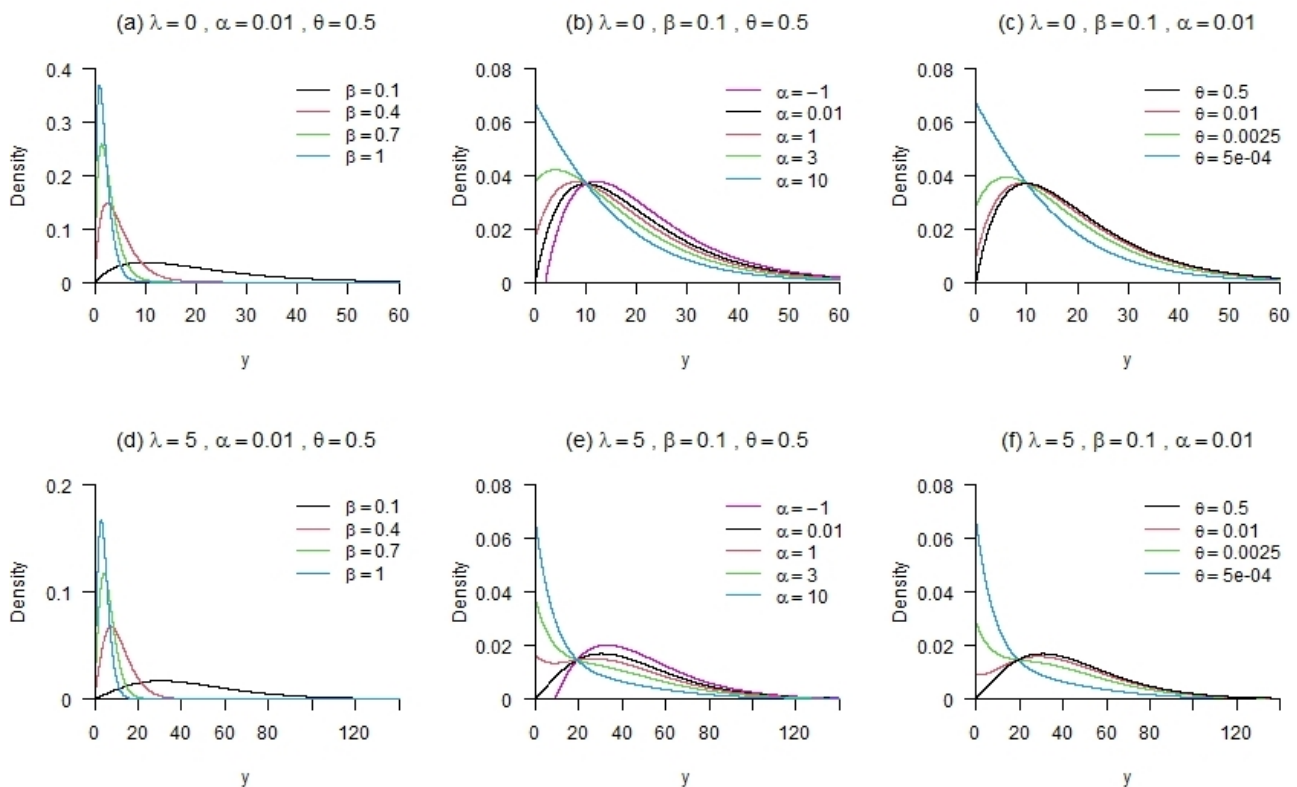**Figure 1.** Shapes of the density function (5) for arbitrary parameter choices. (**a**–**f**) some particular shapes of the density function for various combinations of parameters.

2.2. A Counting Model: PncL Distribution

For discrete (count data) modelling, the Poisson distribution is a well-known and popular choice due to its tractability and ease of implementation. The author of [23] illustrated the possibility of allowing the Poisson parameter (say μ) to be distributed according to the Lindley distribution of [15], and investigated characteristics related to this. When considering the dispersion index $DI = \frac{E(X)}{\text{Var}(X)}$, it is noted that the equidispersion property (when $DI = 1$) of the Poisson distribution is often restrictive—in this sense, the contribution of the ncL distribution as a viable candidate for the distribution of the Poisson parameter μ is discussed in this section. This proves to be particularly valuable and directly insightful as the practitioner may allow for leverage on the mean (or rather, noncentrality) of the model via the use of λ and thus inherently be able to model departures from equidispersion which the data may exhibit.

Suppose a variable X follows a Poisson distribution with parameter $\mu > 0$, and let μ be distributed as ncL with a density function as in (5). Then, using the compounding method with [21], the mass function for X is given by:

$$
\begin{aligned}
p(x) &= \int_0^\infty \frac{\exp(-\mu)\mu^x}{x!} \frac{\beta^2}{\alpha\beta+\theta} \exp(-\beta\mu)\left(\alpha + \theta\exp(-\frac{\lambda}{2})\mu \;\; {}_0F_1(2;\frac{\lambda}{2}\beta\mu)\right)d\mu \\
&= \frac{\beta^2}{\alpha\beta+\theta} \frac{1}{x!}\left[\alpha \int_0^\infty \mu^x \exp(-\mu(\beta+1))d\mu \right. \\
&\quad \left. +\theta\exp(-\frac{\lambda}{2})\int_0^\infty \exp(-\mu(\beta+1))\mu^{x+1} \;\; {}_0F_1(2;\frac{\lambda}{2}\beta\mu)\right] \\
&= \frac{\alpha\beta^2}{(\alpha\beta+\theta)(\beta+1)^{x+1}} + \frac{\theta\beta^2\exp(-\frac{\lambda}{2})\Gamma(x+2)}{(\alpha\beta+\theta)(\beta+1)^{x+2}x!} \;\; {}_1F_1\left(x+2,2;\frac{\frac{\lambda}{2}\beta}{\beta+1}\right)
\end{aligned}
\tag{8}
$$

for $x = 0, 1, 2, \ldots$ representing counts. In this case, we define $X \sim \text{PncL}(\beta, \alpha, \theta, \lambda)$.

The probability generating function and factorial moments of the distribution in (9) are presented in the following theorem. The proof is contained in Appendix A.

Theorem 2. *Suppose that the random variable X is distributed as a PncL variable with mass function (9). Then, the probability generating function and factorial moment of X are given by:*

1.

$$
G_X(s) = \frac{\beta^2}{(\alpha\beta+\theta)(\beta+1-s)^2}\left[\alpha(\beta+1-s) + \theta\exp\left(-\frac{\lambda}{2}\left(\frac{s-1}{\beta+1-s}\right)\right)\right],
\tag{9}
$$

2.

$$
E((X)_r) = \frac{\beta^2}{\alpha\beta+\theta}\frac{\Gamma(r+1)}{(\beta+1)^{r+1}}\left(\alpha + \theta\exp\left(-\frac{\lambda}{2}\right) {}_1F_1\left(2,r+1;\frac{\frac{\lambda}{2}\beta}{\beta+1}\right)\right)
\tag{10}
$$

where $s < \beta + 1$ and $r > 0$. Note that by replacing s with $\exp(s)$ in (9), the moment generating function of the distribution in (9) is obtained.

Our systematically constructed discrete model includes the model of [23] as a special case, particularly when $\alpha = \theta = 1$ and $\lambda = 0$ and also reflects a similar model as that of [3]. It is, however, important to note that [3] relied on the survival discretisation method instead of compounding to obtain a discrete candidate based on their considered underlying generalised Lindley model.

Table 1 illustrates the behaviour that can be captured by the PncL model (9) by calculating the theoretical values of the mean, variance, DI, skewness, and kurtosis for different values of λ. It is valuable to note the correspondence between the mean and λ in

particular, as this provides a unique insight for the practitioner of the leverage related to the *DI* via the incorporation of λ. In this case, the skewness and kurtosis of the variable X is given by $\frac{E(X^3)-3E(X^2)E(X)+2(E(X))^3}{(\text{Var}(X))^{3/2}}$ and $\frac{E(X^4)-4E(X^2)E(X)+6E(X^2)(E(X))^2-3(E(X))^4}{(\text{Var}(X))^2}$, respectively, and calculated using the moments (A1)–(A5) as given in Appendix B.

Table 1. Summary statistics for the PncL distribution for $\alpha = 0.01$, $\beta = 0.9$, $\theta = 0.5$, and various values of λ.

λ	0	5	10	15	20	25	30
Mean	2.203	4.931	7.660	10.389	13.117	15.846	18.575
Variance	4.671	13.705	23.007	32.576	42.414	52.520	62.894
DI	2.121	2.779	3.004	3.136	3.233	3.314	3.386
Skewness	1.495	1.132	0.900	0.746	0.627	0.528	0.440
Kurtosis	6.240	4.772	4.108	3.785	3.608	3.513	3.471

In addition to Table 1, Figure 2 illustrates some particular shapes of the mass function (9) for different values of λ for the following combinations of β, α, and θ:

1. $\beta = 0.1$, $\alpha = 0.02$ and $\theta = 0.3$, illustrated in Figure 2a–c;
2. $\beta = 1$, $\alpha = 0.02$ and $\theta = 0.3$, illustrated in Figure 2d–f;
3. $\beta = 1$, $\alpha = 0.8$ and $\theta = 1.3$, illustrated in Figure 2g–i.

The following is observed from Figure 2:

- λ affects the location (evident from the shifts in the means illustrated in red) and spread of the distribution. As λ increases, the distribution flattens out, shifts to the right and becomes less skewed. This is also evident from Table 1;
- β affects the scale of the distribution. As β increases, the scale of the distribution increases (comparing the mass functions in Figure 2a–c to Figure 2d–f, respectively);
- α and θ jointly affect the shape of the distribution (comparing the mass functions in Figure 2d–f to Figure 2g–i, respectively).
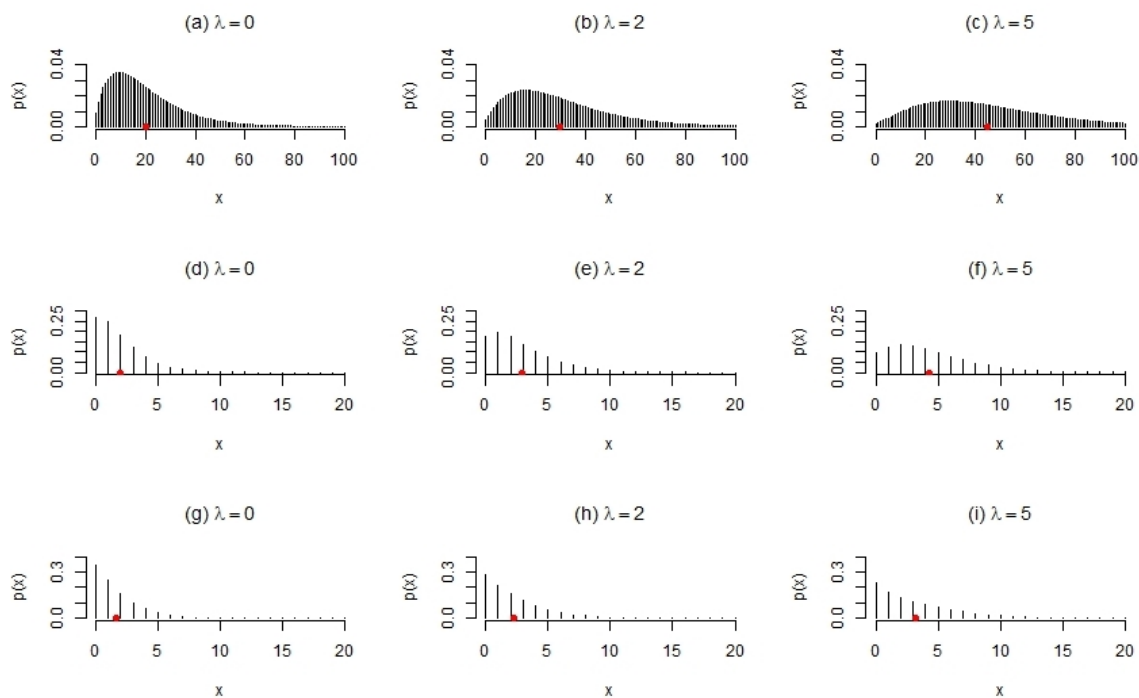

Figure 2. Shapes of the mass function (9) for arbitrary parameter choices, with means indicated in red. (**a–i**) some particular shapes of the mass function (9) for different values of λ for the following combinations of β, α, and θ.

3. Numerical Illustrations

In this section, data fitting applications are presented in the case of the (continuous) ncL, (discrete) PncL, as well as the implementation of the PncL in an INAR(1) environment, to illustrate the position that the newly developed models in Sections 2.1 and 2.2 takes up in the literature. The models are fitted using maximum likelihood estimation for ncL and PncL, and conditional maximum likelihood estimation for the illustration and fit of the INAR(1) process. These fits are comparatively investigated with other popular contenders using criteria which includes the negative maximum log-likelihood (-ℓ) and Akaike's information criterion (AIC). The datasets considered below can be found within the original published works, but for the convenience of the reader, the data has been uploaded to a Google Drive link available at the end of this paper. All computations are carried out using R 4.1.0 in a Win 64 environment with a 1.30 GHz/Intel(R) Core(TM) i7-1065G7 CPU Processor and 8.0 GB RAM.

3.1. Fitting of the ncL

The value of the ncL distribution (5) is illustrated here by fitting this distribution to two datasets:

1. Waiting times (in minutes) before service of 100 banking customers;
2. Survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli.

The fits are compared to the usual Lindley model of [22] (i.e., (5) when $\alpha = \theta = 1$ and $\lambda = 0$) and the three parameter Lindley (3PL) model of [16] (i.e., (5) when $\lambda = 0$). The maximum likelihood estimates for the parameters of these models are reported in Tables 2 and 3, together with the goodness of fit statistics. In addition to the -ℓ and AIC statistics, the Kolmogorov–Smirnov (*KS*) nonparametric test is implemented to compare the data to the estimated models. The *KS* test statistic is defined as:

$$KS = \max_y |F_n(y) - F(y)|$$

where $F_n(y)$ and $F(y)$ represent the empirical and estimated distribution functions, respectively [24]. Since the *KS* test statistic measures the largest absolute distance between the empirical and estimated distribution functions, smaller *KS* test statistics (with larger *p*-values) suggest a better fit.

Table 2. Estimated parameters and goodness of fit statistics of considered models for the banking waiting times dataset.

Model	β	α	θ	λ	$-\ell$	AIC	*KS*	*p*-Value
Lindley	0.187	n/a	n/a	n/a	319.04	640.07	0.068	0.749
3PL	0.211	−0.554	1.477	n/a	316.93	639.85	0.057	0.904
ncL	0.211	−0.531	1.413	0.001	316.93	641.85	0.057	0.903

Table 3. Estimated parameters and goodness of fit statistics of considered models for the guinea pigs dataset.

Model	β	α	θ	λ	$-\ell$	AIC	*KS*	*p*-Value
Lindley	0.011	n/a	n/a	n/a	429.28	860.56	0.17	0.031
3PL	0.012	−287.59	40.167	n/a	428.11	862.21	0.152	0.072
ncL	0.027	−6.071	55.985	5.48	426.05	860.1	0.101	0.455

With similar -ℓ and AIC statistics, Table 2 shows that the proposed ncL model compares well to its Lindley and 3PL contenders for the banking waiting times dataset. With a small (close to zero) estimated λ, it can be concluded that the noncentrality is insignificant for this data, hence the ncL model performs similar to the Lindley and 3PL models. These results are verified in Figure 3.
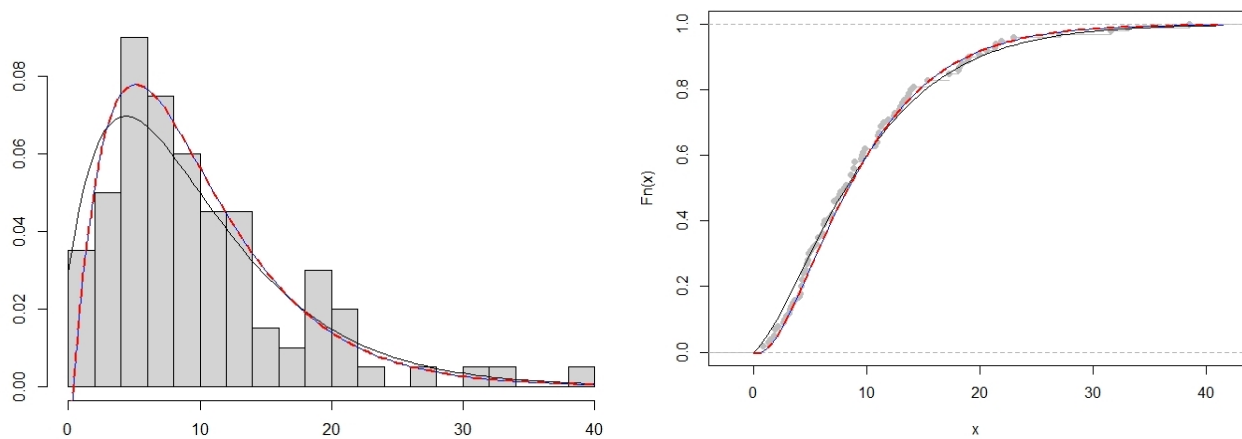
Figure 3. Fitted Lindley (black), 3PL (blue), and ncL (red) models with corresponding estimated distribution functions for the banking waiting times dataset.

According to all goodness of fit statistics, Table 3 shows that the proposed ncL model outperforms both its Lindley and 3PL contenders for the guinea pigs dataset. This suggests that the noncentrality within this dataset is captured effectively, which is also evident from Figure 4.
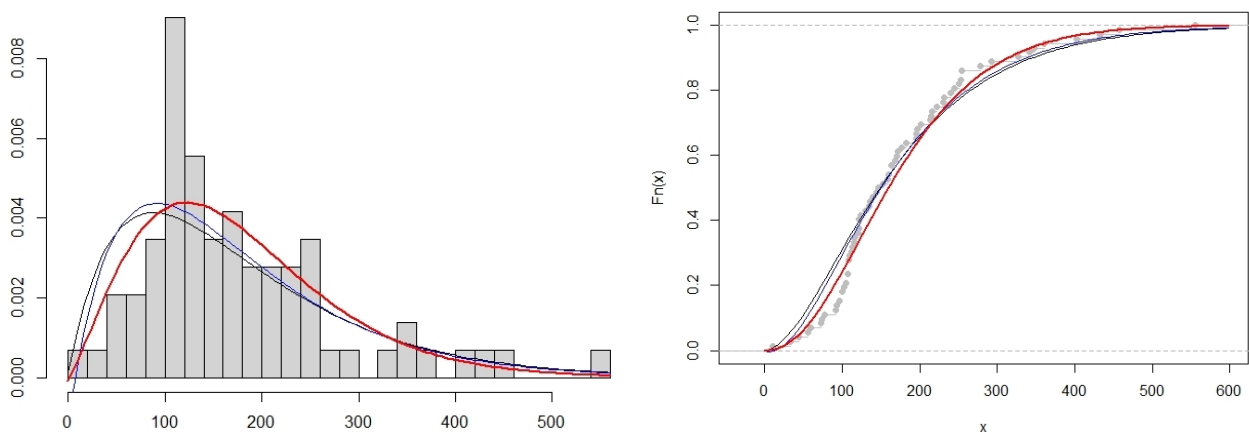

Figure 4. Fitted Lindley (black), 3PL (blue), and ncL (red) models with corresponding estimated distribution functions for the guinea pigs dataset.

3.2. Fitting of the PncL

Here, the discrete PncL (9), in comparison to the Poisson–Lindley (PL) and Poisson three parameter Lindley (P-3PL) of [3,23], respectively, is fitted to the following datasets:

1. The length of stay of 67 patients on a psychiatric ward (see [3]);
2. The survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli.

It is of interest to note that the ncL model was also fitted to the second dataset mentioned above. The values of this dataset are spread among a higher range and could be considered continuous (see previous section), but in essence, the integer nature of the data could also be considered as counts. In this spirit, we analyse the same dataset using the PncL as well. Tables 4 and 5 report the results of the psychiatric data fitting and the estimated statistics of interest, such as the mean, variance, DI, skewness, and kurtosis, respectively. Comparing the goodness of fit statistics in Table 4, it is clear that both the P-3PL and PncL perform similarly, outperforming the PL. Considering the estimated statistics in Table 5, it is evident that the proposed PncL describes the data best.

Table 4. Estimated parameters and goodness of fit statistics of considered discrete models for the psychiatric dataset.

Model	β	α	θ	λ	$-\ell$	AIC
PL	0.616	n/a	n/a	n/a	138.37	278.74
P-3PL	0.726	−1.145	4.031	n/a	129.39	264.78
PncL	1.665	−0.802	4.58	2.711	129.36	266.73

Table 5. Mean, variance, *DI*, skewness, and kurtosis values of considered discrete models for the psychiatric dataset.

Model	Mean	Variance	*DI*	Skewness	Kurtosis
PL	2.629	7.52	2.861	1.647	6.879
P-3PL	2.597	3.829	1.474	1.38	5.94
PncL	2.597	3.684	1.419	1.285	5.481
Data	2.587	3.668	1.413	0.98	0.231

Considering the discrete fit for the guinea pigs dataset, Tables 6 and 7 report the estimation results and estimated statistics of interest, such as the mean, variance, *DI*, skewness, and kurtosis, respectively. Comparing the goodness of fit statistics in Table 6, it is clear that the proposed PncL outperforms both PL and P-3PL models, capturing the characteristics of the data well (evident from Table 7).

Table 6. Estimated parameters and goodness of fit statistics of considered discrete models for the guinea pigs dataset.

Model	β	α	θ	λ	$-\ell$	AIC
PL	0.011	n/a	n/a	n/a	429.47	860.94
P-3PL	0.012	−119.963	15.541	n/a	428.07	862.13
PncL	0.031	8.216	73.062	7.089	425.94	859.89

Table 7. Mean, variance, *DI*, skewness, and kurtosis values of considered discrete models for the guinea pigs dataset.

Model	Mean	Variance	*DI*	Skewness	Kurtosis
PL	176.803	15,980.8	90.388	1.414	6.001
P-3PL	176.821	14,178.9	80.188	1.428	6.03
PncL	176.809	9517.5	53.829	0.92	4.174
Data	176.819	10,702.9	60.53	1.371	2.225

3.3. Fitting of the PncL within INAR(1)

Here, we model the INAR process as given in (1), where the one-step transition probability of the process is given by:

$$P(X_t = k | X_{t-1} = l) = \sum_{i=1}^{\min(k,l)} P(B_l^p = i) \times P_{PncL}(\epsilon_t = k - i) \tag{11}$$

where $B_l^p \sim Binomial(p, l)$ and $p \in (0, 1)$, and in the case of this study, $P_{PncL}(\epsilon_t = k - i)$, is given by (9). This proposed process aims to allow the researcher a more interpretable approach to leverage information of the mean via the noncentrality parameter λ, and subsequently incorporate this as part of the estimation of the model when applied to time series data that illustrates departures from equidispersion.

Inspired by [7,25], the following remark is presented without proof, and captures meaningful properties of an INAR(1) process when subjected with PncL innovations with mass function in (9):

Remark 1. *If X_t is the INAR(1) process as defined in (1) with transition probability given by (11), then:*

1. $E(X_t|X_{t-1}) = pX_{t-1} + E(X)$;
2. $Var(X_t|X_{t-1}) = p(1-p)X_{t-1} + E(X^2) - (E(X))^2$;
3. $\mu_{X_t} = \frac{E(X)}{1-p}$;
4. $\sigma^2_{X_t} = \frac{E(X^2)-(E(X))^2+pE(X)}{1-p^2}$;
5. $\gamma_k = p^k Var(X_{t-k}), \quad \rho_k = p^k$;

for lag $k > 0$. Here, $E(X)$ is given by (A1), and $E(X^2)$ is given by (A2).

We consider the following time series datasets for the conditional maximum likelihood estimation of the PncL distribution within an INAR(1) context, in comparison to its well-known Poisson, negative binomial (NB), Bell (see [7,26]), and P-3PL (see [3]) contenders:

1. Daily number of downloads of certain software for the period June 2006 up to February 2007 (sample size $T = 267$) (see [27]);
2. Monthly number of strikes leading to at least 1000 workers being idle (published by the U.S. Bureau of Labor Statistics, http://www.bls.gov/wsp/ (accessed on 1 December 2021)). The time period January 1994 to December 2002 (sample size $T = 108$) is considered.

The conditional log-likelihood function of our INAR(1) process with innovations given by (9) is defined as:

$$
\begin{aligned}
\ell(\Theta) &= \sum_{t=2}^{T} \ln(P(X_t = k|X_{t-1} = l)) \\
&= \sum_{t=2}^{T} \ln\left(\sum_{i=1}^{\min(X_t,X_{t-1})} \binom{X_{t-1}}{i} p^i (1-p)^{X_{t-1}-i} p_{PncL}(X_t - i) \right)
\end{aligned}
\tag{12}
$$

where $\Theta = (p, \alpha, \beta, \theta, \lambda)$ is the vector of maximum likelihood estimates of the parameters and $p_{PncL}(\cdot)$ denotes the mass function in (9).

Figures 5 and 6 illustrate the time series plots of both the downloads' and strikes' data, respectively, together with the observed marginal distributions, autocorrelation functions (ACFs), and partial autocorrelation functions (PACFs). The sample mean and variance for the downloads data is given by 2.401 and 7.534, respectively, whereas the sample mean and variance for the strikes' data is given by 4.944 and 7.922, respectively. After evaluating the sample ACFs and PACFs, it is concluded that an AR(1)-like process is an appropriate choice for fitting a time series model to both these datasets (suggested by the significant sample partial autocorrelations at lags 1 only). Furthermore, since both these time series are discrete in nature, an INAR(1) process is suggested with an error term ϵ_t characterised by a discrete distribution allowing for overdispersion (due to a sample $DI > 1$ for both time series).
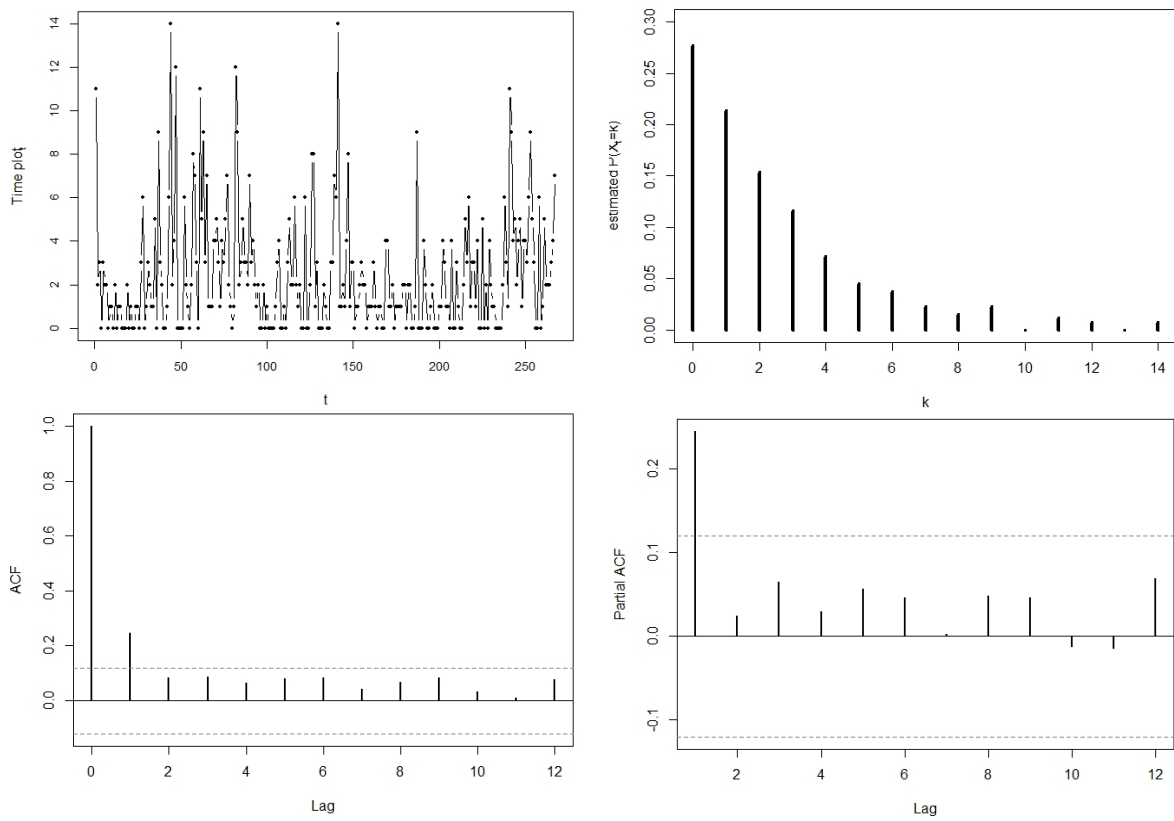
Figure 5. The time plot, observed marginal distribution, ACF, and PACF of the downloads dataset.
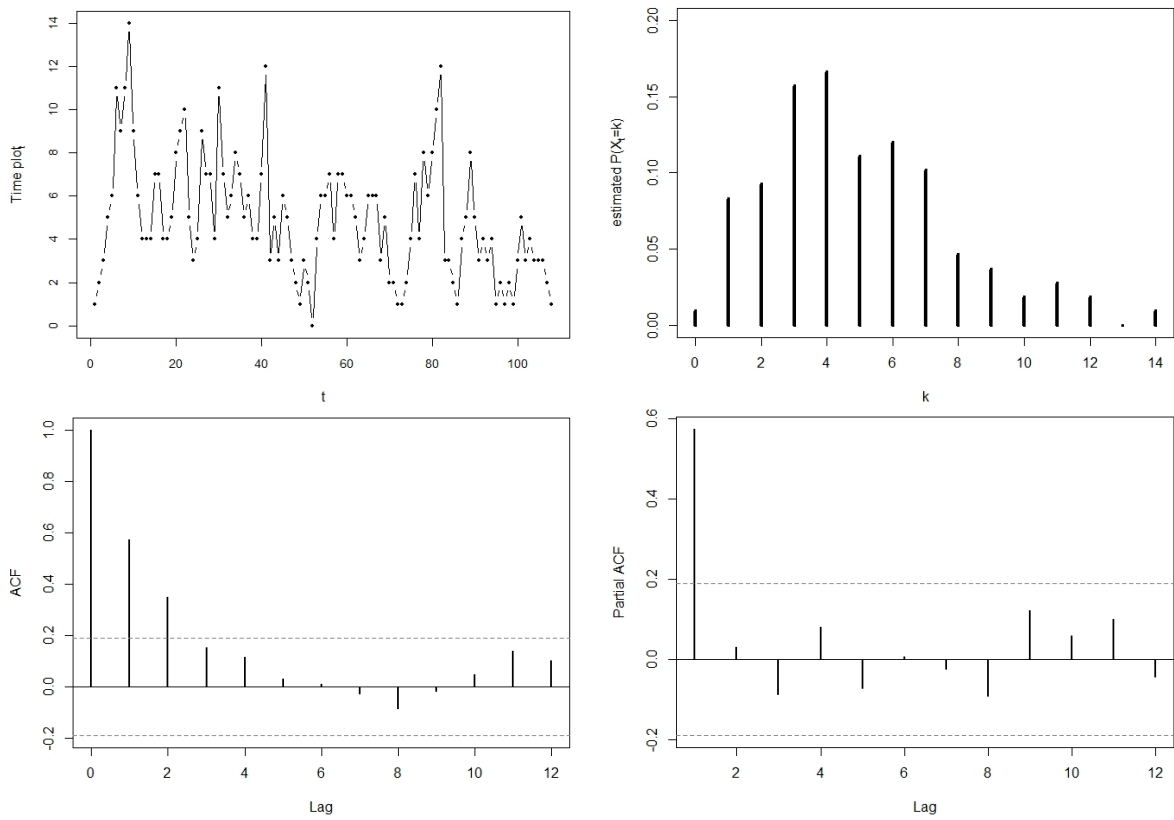

Figure 6. The time plot, observed marginal distribution, ACF, and PACF of the strikes dataset.

Table 8 shows the estimation results for the downloads data. With the PncL INAR(1) model having the lowest -ℓ value and an estimated mean and variance of 2.366 and 6.995 (compared to the sample mean and variance of 2.4 and 7.5), respectively, we conclude that the proposed PncL INAR(1) model competes well in relation to its NB INAR(1) contender, having the smallest AIC value. This estimated PncL INAR(1) model is given by:

$$X_t = 0.157 \circ X_{t-1} + \epsilon_t$$

where the error term structure is estimated as $\epsilon_t \sim \text{PncL}(1.023, 2.32, 1.29, 3.919)$.

Table 8. Estimated parameters, goodness of fit statistics, fitted mean, and fitted variance of considered models for the download counts time series dataset.

Model	Parameter	Estimates	$-\ell$	AIC	Mean (μ_{X_t})	Variance ($\sigma^2_{X_t}$)
Poisson INAR(1)	p	0.172	634.1	1272	2.365	2.365
	μ	1.959				
NB INAR(1)	p	0.154	537.9	1082	2.366	7.187
	n	0.850				
	π	0.298				
Bell INAR(1)	p	0.108	554.2	1112	2.367	4.240
	θ	0.877				
P-3PL INAR(1)	p	0.138	538.3	1085	2.366	6.601
	α	16.180				
	β	0.399				
	θ	0.001				
PncL INAR(1)	p	0.157	537.6	1085	2.366	6.995
	α	2.320				
	β	1.023				
	θ	1.290				
	λ	3.919				

Considering the estimation results for the strikes' data, Table 9 shows that the proposed PncL INAR(1) model competes well to its previously proposed competitors. With the lowest -ℓ value and an estimated mean and variance of 4.981 and 6.922 (compared to the sample mean and variance of 4.9 and 7.9), respectively, we can conclude that the PncL INAR(1) model is an adequate fit for the strikes' data with an estimated function given by:

$$X_t = 0.558 \circ X_{t-1} + \epsilon_t$$

where the error term structure is estimated as $\epsilon_t \sim \text{PncL}(3.843, 0.29, 13.083, 14.191)$. Note that even though the Bell INAR(1) yields the smallest AIC value, it overestimates the variance within the time series. In this table, the mean and variance are rounded to three decimal places and in this case the means present almost equal for all of the fitted models.

In order to examine the accuracy of the proposed candidate, it remains necessary to analyse residuals from the fitted model, considering the standardised Pearson residuals for $t = 2, \ldots, T$ defined by:

$$e_t = \frac{x_t - E[X_t|x_{t-1}]}{\sqrt{\text{Var}[X_t|x_{t-1}]}} \tag{13}$$

where $E[X_t|x_{t-1}] = pX_{t-1} + \mu_\epsilon$ and $\text{Var}[X_t|x_{t-1}] = p(1-p)X_{t-1} + \sigma^2_\epsilon$ (see [1]), with μ_ϵ and σ^2_ϵ denoting the mean and variance of the innovation ϵ_t distribution calculated using the moments (A1) and (A2) given in Appendix B. For the model to be an adequate fit, it is required that these residuals are uncorrelated with a mean close to zero and a variance close to one. Note that a variance deviating from one indicates that the model does not sufficiently capture all dispersion present in the data. The mean and variance of the Pearson residuals obtained from the PncL INAR(1) model for the downloads data are -0.001 and

0.631, respectively. Considering the Pearson residuals for the strikes' data, we observe a mean and variance of −0.001 and 0.522, respectively. With these values being close to the desired values, and together with the respective ACFs of the Pearson residuals showing no significant autocorrelation in Figures 7 and 8, it is concluded that the proposed PncL INAR(1) model fits both time series adequately.

Table 9. Estimated parameters, goodness of fit statistics, fitted mean, and fitted variance of considered models for the strikes counts time series dataset.

Model	Parameter	Estimates	$-\ell$	AIC	Mean (μ_{X_t})	Variance ($\sigma^2_{X_t}$)
Poisson INAR(1)	p	0.506	234.5	473.1	4.981	4.981
	μ	2.460				
NB INAR(1)	p	0.548	231.8	469.7	4.981	6.858
	n	3.858				
	π	0.632				
Bell INAR(1)	p	0.579	232.1	468.2	4.981	7.741
	θ	0.875				
P-3PL INAR(1)	p	0.548	232.0	472.1	4.981	7.139
	α	−0.322				
	β	0.737				
	θ	8.276				
PncL INAR(1)	p	0.558	231.8	473.6	4.981	6.922
	α	0.290				
	β	3.843				
	θ	13.083				
	λ	14.191				
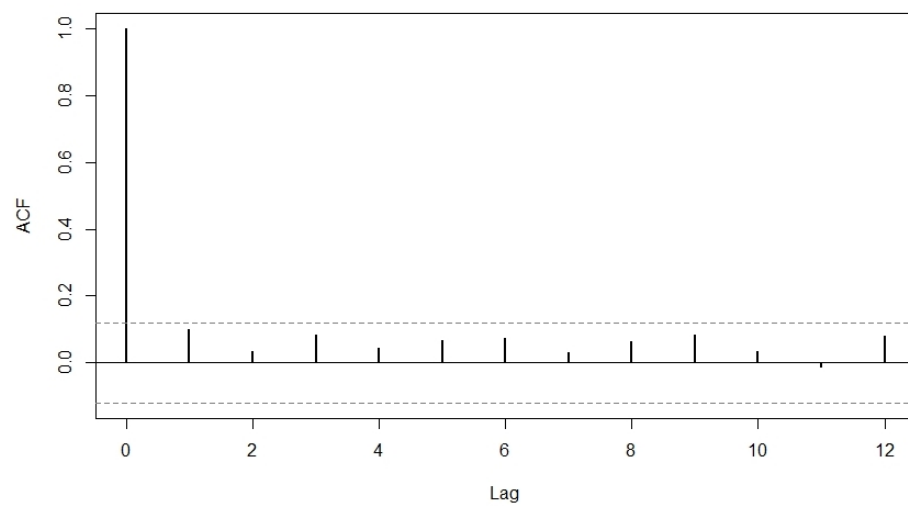
Figure 7. The ACF of the Pearson residuals (13) for the downloads dataset.
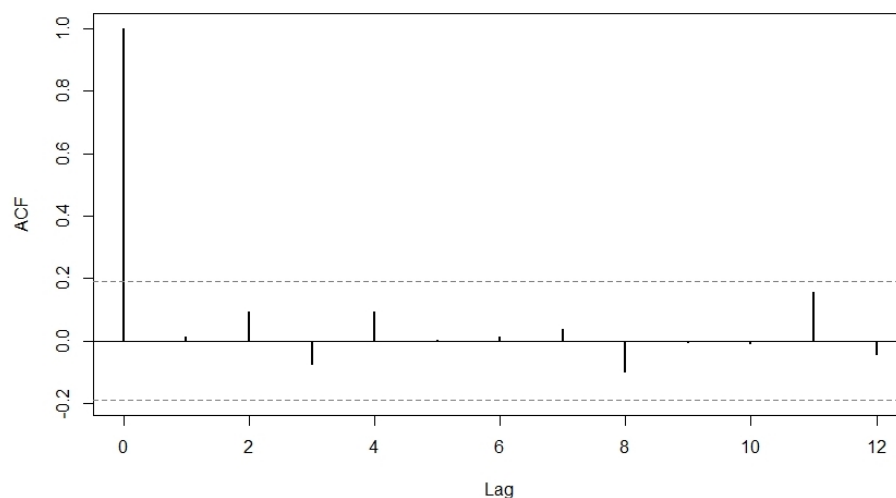
Figure 8. The ACF of the Pearson residuals ([13](#)) for the strikes dataset.

4. Conclusions

This study presents an enriched noncentral Lindley distribution (ncL) containing a (noncentral) parameter which has been systematically introduced via the noncentral gamma distribution. The characteristics of this newly developed model received attention, which includes the cumulative distribution function, moment generating function, and the moments of the distribution which are derived in closed form analytical expressions. This model retains mathematical elegance with closed forms for the characterisations, and is shown to contain existing models in the literature for specific choices of the parameters. In addition, a discrete counterpart was derived by using this ncL model via compounding with a Poisson variable. In this way, the interpretability of the noncentral parameter is inherited within a discrete environment as well, and paves the way for the practitioner to leverage information regarding the mean (or rather, noncentrality) of any suitable given data to provide insight on the dispersion index of the data in question. This model is incorporated within a time series context as innovation terms in an INAR(1) environment and juxtaposed against other popular models.

It is valuable to note that no one distribution will always outperform another, due to the data-driven nature of statistical model fitting. As such, the contributions in this paper does not exclusively outperform competing models in terms of the considered datasets—but results indicate that it is as good a model, if not occasionally slightly better, than the usual choices. The true value of this paper lies within this systematic construction of a previously unconsidered continuous (ncL) and discrete (PncL) distribution, and the inclusion of a noncentral parameter which is estimable and interpretable in a location context for data. Additionally, these models contain usual considered models as special cases, and thus acts as a unifying consideration for practitioners not only with continuous or discrete interests, but also with discrete time series interests. In future, further considerations and comparisons of other unconsidered discrete models (such as the refreshing recent contribution of [28] as well as [29]) within a discrete time series environment may be pursued, in conjunction with alternate thinning considerations (such as [30]).

Appendix A. Proofs

Proof of Theorem 1.

1. See that

$$
\begin{aligned}
F_Y(y) &= \frac{\beta^2}{\alpha\beta + \theta} \int_0^y \exp(-\beta t)\left[\alpha + t\theta \exp\left(-\frac{\lambda}{2}\right) {}_0F_1\left(2; \frac{\lambda}{2}\beta t\right)\right] dt \\
&= \frac{\alpha\beta^2}{\alpha\beta + \theta} \int_0^y \exp(-\beta t)dt + \frac{\theta\beta^2}{\alpha\beta + \theta} \exp\left(-\frac{\lambda}{2}\right) \int_0^y t \exp(-\beta t) {}_0F_1\left(2; \frac{\lambda}{2}\beta t\right) dt \\
&= \frac{\alpha\beta}{\alpha\beta + \theta}(1 - \exp(-\beta x)) + \frac{\theta}{\alpha\beta + \theta} \exp\left(-\frac{\lambda}{2}\right) \sum_{i=0}^\infty \frac{\left(\frac{\lambda}{2}\right)^i}{(2)_i i!} \gamma(i+2, \beta x) \\
&= \omega(1 - \exp(-\beta x)) + (1 - \omega) \exp\left(-\frac{\lambda}{2}\right) \sum_{i=0}^\infty \frac{\left(\frac{\lambda}{2}\right)^i}{(2)_i i!} \gamma(i+2, \beta x).
\end{aligned}
$$

2. Using [21] p. 815, see that

$$
\begin{aligned}
M_Y(y) &= \int_0^\infty \exp(ty) \frac{\beta^2}{\alpha\beta + \theta} \exp(-\beta y)\left[\alpha + y\theta \exp\left(-\frac{\lambda}{2}\right) {}_0F_1\left(2; \frac{\lambda}{2}\beta y\right)\right] dy \\
&= \frac{\alpha\beta^2}{\alpha\beta + \theta} \int_0^\infty \exp(-(\beta - t)y)dy \\
&\quad + \frac{\theta\beta^2}{\alpha\beta + \theta} \exp\left(-\frac{\lambda}{2}\right) \int_0^\infty y \exp(-(\beta - t)y) {}_0F_1\left(2; \frac{\lambda}{2}\beta y\right) dy \\
&= \frac{\beta^2}{\alpha\beta + \theta}\left(\frac{\alpha}{\beta - t} + \frac{\theta \exp\left(-\frac{\lambda}{2}\right)}{(\beta - t)^2} \exp\left(\frac{\frac{\lambda}{2}\beta}{\beta - t}\right)\right),
\end{aligned}
$$

which leaves the final result.

3. Using [21] p. 815, see that

$$E(Y^r)$$

$$= \int_0^\infty y^r \frac{\beta^2}{\alpha\beta+\theta} \exp(-\beta y)\left[\alpha + y\theta\exp\left(-\frac{\lambda}{2}\right){}_0F_1\left(2;\frac{\lambda}{2}\beta y\right)\right]dy$$

$$= \frac{\alpha\beta^2}{\alpha\beta+\theta}\int_0^\infty y^r\exp(-\beta y)dy$$

$$+ \frac{\theta\beta^2}{\alpha\beta+\theta}\exp\left(-\frac{\lambda}{2}\right)\int_0^\infty y^{r+1}\exp(-\beta y)\,{}_0F_1\left(2;\frac{\lambda}{2}\beta y\right)dy$$

$$= \frac{\beta^2}{\alpha\beta+\theta}\frac{\alpha}{\beta^{r+1}}\Gamma(r+1) + \frac{\beta^2}{\alpha\beta+\theta}\frac{\theta}{\beta^{r+2}}\exp\left(-\frac{\lambda}{2}\right)\Gamma(r+2)\,{}_1F_1\left(r+2,2;\frac{\lambda}{2}\right),$$

which leaves the final result.

□

Proof of Theorem 2.

1. See that

$$G_X(s)$$

$$= \sum_{x=0}^\infty s^x \int_0^\infty \frac{\exp(-\mu)\mu^x}{x!}\frac{\beta^2}{\alpha\beta+\theta}\exp(-\beta\mu)\left[\alpha + \mu\theta\exp\left(-\frac{\lambda}{2}\right){}_0F_1\left(2;\frac{\lambda}{2}\beta\mu\right)\right]d\mu$$

$$= \frac{\beta^2}{\alpha\beta+\theta}\int_0^\infty \exp(-s\mu-\mu-\beta\mu)\left[\alpha + \mu\theta\exp\left(-\frac{\lambda}{2}\right){}_0F_1\left(2;\frac{\lambda}{2}\beta\mu\right)\right]d\mu$$

$$= \frac{\beta^2}{\alpha\beta+\theta}\left(\alpha\frac{1}{\beta+1-s} + \theta\exp\left(-\frac{\lambda}{2}\right)\frac{1}{(\beta+1-s)^2}\,{}_1F_1\left(2,2;\frac{\frac{\lambda}{2}\beta}{\beta+1-s}\right)\right)$$

$$= \frac{\beta^2}{\alpha\beta+\theta}\frac{1}{(\beta+1-s)^2}\left(\alpha(\beta+1-s) + \theta\exp\left(-\frac{\lambda}{2}\left(\frac{s-1}{\beta+1-s}\right)\right)\right),$$

which leaves the final result.

2. See that

$$E((X)_r)$$

$$= \sum_{x=0}^{\infty} x(x-1)\dots(x-r+1) \int_0^{\infty} \frac{\exp(-\mu)\mu^x}{x!} \frac{\beta^2}{\alpha\beta+\theta} \exp(-\beta\mu)$$

$$\times \left[\alpha + \mu\theta \exp\left(-\frac{\lambda}{2}\right) {}_0F_1\left(2; \frac{\lambda}{2}\beta\mu\right)\right] d\mu$$

$$= \int_0^{\infty} \left(\sum_{x=r}^{\infty} \frac{\exp(-\mu)\mu^{x-r}}{(x-r)!}\right) \mu^r \exp(-\mu) \frac{\beta^2}{\alpha\beta+\theta} \exp(-\beta\mu)$$

$$\times \left[\alpha + \mu\theta \exp\left(-\frac{\lambda}{2}\right) {}_0F_1\left(2; \frac{\lambda}{2}\beta\mu\right)\right] d\mu$$

$$= \frac{\beta^2}{\alpha\beta+\theta} \int_0^{\infty} \mu^r \exp(-(\beta+1)\mu)\left[\alpha + \mu\theta \exp\left(-\frac{\lambda}{2}\right) {}_0F_1\left(2; \frac{\lambda}{2}\beta\mu\right)\right] d\mu$$

$$= \frac{\alpha\beta^2}{\alpha\beta+\theta} \int_0^{\infty} \mu^r \exp(-(\beta+1)\mu) d\mu$$

$$+ \frac{\beta^2}{\alpha\beta+\theta}\theta \exp\left(-\frac{\lambda}{2}\right) \int_0^{\infty} \mu^r \exp(-(\beta+1)\mu) {}_0F_1\left(2; \frac{\lambda}{2}\beta\mu\right) d\mu$$

$$= \frac{\alpha\beta^2}{\alpha\beta+\theta} \frac{\Gamma(r+1)}{(\beta+1)^{r+1}} + \frac{\beta^2}{\alpha\beta+\theta}\theta \exp\left(-\frac{\lambda}{2}\right) \frac{\Gamma(r+1)}{(\beta+1)^{r+1}} {}_1F_1\left(2, r+1; \frac{\frac{\lambda}{2}\beta}{\beta+1}\right),$$

which again, using [21] p. 815, leaves the final result.

□

Appendix B. Moments

The first four moments of interest of the PncL distribution (9) is given below.

$$E(X) = \frac{\left(\alpha + \frac{\theta}{\beta}\left(\frac{\lambda}{2}+2\right)\right)}{\alpha\beta+\theta}. \tag{A1}$$

$$E(X^2) = \frac{1}{\alpha\beta+\theta}\left[\theta\left(\frac{\lambda}{2\beta} + \frac{2}{\beta} + \frac{3\lambda}{\beta^2} + \frac{\lambda^2}{4\beta^2} + \frac{6}{\beta^2}\right) + \alpha + \frac{2\alpha}{\beta}\right]. \tag{A2}$$

$$E(X^3)$$

$$= \frac{1}{\alpha\beta+\theta}\left[\alpha(1 + \frac{6}{\beta} + \frac{6}{\beta^2})\right. \tag{A3}$$

$$\left. + \theta\left(\frac{\lambda^3}{8\beta^3} + \frac{3\lambda^2}{\beta^3} + \frac{3\lambda^2}{4\beta^2} + \frac{18\lambda}{\beta^3} + \frac{9\lambda}{\beta^2} + \frac{\lambda}{2\beta} + \frac{2}{\beta} + \frac{18}{\beta^2} + \frac{24}{\beta^3}\right)\right].$$

$$E(X^4)$$

$$= \frac{1}{\alpha\beta + \theta}(\alpha(1 + \frac{14}{\beta} + \frac{36}{\beta^2} + \frac{24}{\beta^3}) + \theta \times (\frac{2 + \frac{\lambda}{2}}{\beta} + \frac{42 + 21\lambda + \frac{7}{4}\lambda^2}{\beta^2} \tag{A4}$$

$$+ \frac{144 + 108\lambda + 18\lambda^2 + \frac{3}{4}\lambda^3}{\beta^3} + \frac{120 + 120\lambda + 30\lambda^2 + \frac{5}{2}\lambda^3 + \frac{1}{16}\lambda^4}{\beta^4})).$$

References

1. Weiß, C.H. *An Introduction to Discrete-Valued Time Series*; John Wiley & Sons: Hoboken, NJ, USA, 2018.
2. Lívio, T.; Khan, N.M.; Bourguignon, M.; Bakouch, H.S. An INAR (1) model with Poisson–Lindley innovations. *Econ. Bull.* **2018**, *38*, 1505–1513.
3. Eliwa, M.S.; Altun, E.; El-Dawoody, M.; El-Morshedy, M. A new three-parameter discrete distribution with associated INAR (1) process and applications. *IEEE Access* **2020**, *8*, 91150–91162. [CrossRef]
4. Eliwa, M.; El-Morshedy, M. A one-parameter discrete distribution for over-dispersed data: Statistical and reliability properties with applications. *J. Appl. Stat.* **2021**, 1–21. [CrossRef]
5. Irshad, M.R.; Chesneau, C.; D'cruz, V.; Maya, R. Discrete Pseudo Lindley Distribution: Properties, Estimation and Application on INAR (1) Process. *Math. Comput. Appl.* **2021**, *26*, 76. [CrossRef]
6. Alzaid, A.; Al-Osh, M. First-order integer-valued autoregressive (INAR (1)) process: Distributional and regression properties. *Stat. Neerl.* **1988**, *42*, 53–61. [CrossRef]
7. Huang, J.; Zhu, F. A New First-Order Integer-Valued Autoregressive Model with Bell Innovations. *Entropy* **2021**, *23*, 713. [CrossRef] [PubMed]
8. McKenzie, E. Some simple models for discrete variate time series. *Water Resour. Bull.* **1985**, *21*, 645–650. [CrossRef]
9. Kim, H.; Lee, S. On first-order integer-valued autoregressive process with Katz family innovations. *J. Stat. Comput. Simul.* **2017**, *87*, 546–562. [CrossRef]
10. Al-Osh, M.A.; Alzaid, A.A. First-order integer-valued autoregressive (INAR (1)) process. *J. Time Ser. Anal.* **1987**, *8*, 261–275. [CrossRef]
11. Ristić, M.M.; Bakouch, H.S.; Nastić, A.S. A new geometric first-order integer-valued autoregressive (NGINAR (1)) process. *J. Stat. Plan. Inference* **2009**, *139*, 2218–2226. [CrossRef]
12. Ekhosuehi, N.; Opone, F. A three parameter generalized Lindley distribution: Properties and application. *Statistica* **2018**, *78*, 233–249.
13. Abd El-Monsef, M. A new Lindley distribution with location parameter. *Commun. Stat. Theory Methods* **2016**, *45*, 5204–5219. [CrossRef]
14. Nedjar, S.; Zeghdoudi, H. On gamma Lindley distribution: Properties and simulations. *J. Comput. Appl. Math.* **2016**, *298*, 167–174. [CrossRef]
15. Lindley, D.V. Fiducial distributions and Bayes' theorem. *J. R. Stat. Soc. Ser. Methodol.* **1958**, *20*, 102–107. [CrossRef]
16. Shanker, R.; Shukla, K.K.; Shanker, R.; Tekie, A. A three-parameter Lindley distribution. *Am. J. Math. Stat.* **2017**, *7*, 15–26.
17. Knüsel, L.; Bablok, B. Computation of the noncentral gamma distribution. *SIAM J. Sci. Comput.* **1996**, *17*, 1224–1231. [CrossRef]
18. Bekker, A.; Ferreira, J. Bivariate gamma type distributions for modelling wireless performance metrics. *Stat. Optim. Inf. Comput.* **2018**, *6*, 335–353. [CrossRef]
19. Chen, Z.Y. The S-system computation of non-central gamma distribution. *J. Stat. Comput. Simul.* **2005**, *75*, 813–829. [CrossRef]
20. de Oliveira, I.R.C.; Ferreira, D.F. Computing the noncentral gamma distribution, its inverse and the noncentrality parameter. *Comput. Stat.* **2013**, *28*, 1663–1680. [CrossRef]
21. Gradshteyn, I.S.; Ryzhik, I.M. *Table of Integrals, Series, and Products*; Academic Press: Cambridge, MA, USA, 2014.
22. Ghitany, M.E.; Atieh, B.; Nadarajah, S. Lindley distribution and its application. *Math. Comput. Simul.* **2008**, *78*, 493–506. [CrossRef]
23. Sankaran, M. The discrete Poisson-Lindley distribution. *Biometrics* **1970**, *26*, 145–149. [CrossRef]
24. Massey Jr, F.J. The Kolmogorov-Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **1951**, *46*, 68–78. [CrossRef]
25. Qi, X.; Li, Q.; Zhu, F. Modeling time series of count with excess zeros and ones based on INAR (1) model with zero-and-one inflated Poisson innovations. *J. Comput. Appl. Math.* **2019**, *346*, 572–590. [CrossRef]
26. Castellares, F.; Ferrari, S.L.; Lemonte, A.J. On the Bell distribution and its associated regression model for count data. *Appl. Math. Model.* **2018**, *56*, 172–185. [CrossRef]
27. Weiß, C.H. Thinning operations for modelling time series of counts—A survey. *AStA Adv. Stat. Anal.* **2008**, *92*, 319–341. [CrossRef]
28. Hanandeh, A.; Al-Nasser, A.D. New Distribution for Fitting Discrete Data: The Poisson-Gold Distribution and Its Statistical Properties. *Austrian J. Stat.* **2021**, *50*, 19–35. [CrossRef]

29. Bhati, D.; Sastry, D.; Qadri, P.M. A new generalized Poisson-Lindley distribution: Applications and properties. *Austrian J. Stat.* **2015**, *44*, 35–51. [CrossRef]

30. Liu, Z.; Zhu, F. A new extension of thinning-based integer-valued autoregressive models for count data. *Entropy* **2021**, *23*, 62. [CrossRef] [PubMed]