International Journal of Statistics and Probability; Vol. 12, No. 6; November 2023
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Integer-Valued First Order Autoregressive (INAR(1)) Model With Negative Binomial (NB) Innovation For The Forecasting Of Time Series Count Data

Nasiru Mukaila Olakorede[1] and Samuel Olayemi Olanrewaju[1]

[1]Department of Statistics, Faculty of Science, University of Abuja, Abuja, Nigeria

Correspondence: Nasiru Mukaila Olakorede, Department of Statistics, Faculty of Science, University of Abuja, Abuja, Nigeria. E-mail: nasiru.olakorede@uniabuja.edu.ng and olanrewaju.samuel@uniabuja.edu.ng

Received: April 18, 2023 Accepted: November 30, 2023 Online Published: December 27, 2023

doi:10.5539/ijsp.v12n6p23 URL: https://doi.org/10.5539/ijsp.v12n6p23

Abstract

This paper is about the theoretical investigation of integer-valued first order autoregressive (INAR(1)) model with negative binomial (NB) innovation for the forecasting of time series count data. The study makes use of the Conditional Least squares (CLS) estimator to estimate the parameter of INAR(1) model, and Maximum Likelihood Estimator (MLE) to estimate the mean (μ) and the dispersion parameter (K) of the NB distribution. A simulation experiment based on theoretical generated data were addressed under different parameter values α=0.2, 0.6, 0.8, different sample sizes n=30, 90, 120, 600 for the class of INAR(1) model, and μ =0.85, 1.5, 2, K=1,2, 4 for the NB distribution. The Monte Carlo simulations were conducted with codes written in R, all results were based on 1000 runs. The estimation of parameter for the class of INAR(1) model gives a better result when the number of observations is small and the parameter value is high. The NB estimation gives a better result when the number of observations is small and with large K values. The forecasting accuracy of the model at different lead time period l=1, 3, 5, 7, 9, 15 were investigated with codes written in R. The results showed that the minimum mean square error (MMSE) produced when the number of lead times forecasts is between one and five were less than that produced when the numbers of lead times forecast were greater than five. The MMSE increased when the number of lead time periods increases. This result indicates that forecasting with this class of model is better with short time frame of predictions. The study was applied to the number of deaths arising from COVID-19 in Nigeria which consist of count time series data of 48 observations (weekly data), from January 2021 to December 2021.

Keywords: INAR(1) model, NB distribution, count data, CLS estimation, MLE estimation, Covid-19, Forecasting

1. Introduction

Integer Autoregressive Moving Average (INARMA) models has recently received wider attention in the literature. The necessity for such investigations arises from the fact that, INARMA models are capable of modelling and forecasting Time Series Count data that appears in several diverse scientific especially for low frequency count with overdispersed data.

A time series is a set of observations y_t, observed sequentially with time t. For continuous time series, the observations are measured continuously over some time interval, for example, T=[0,1], for a discrete-time series, the observations are measured at a sequential integer values over a fixed time intervals.

Discrete variate time series for counts occur in many contexts either as counts of events, for example, the number of road accidents in a given period of time, the number of births at a hospital in a given period of time, the number of deaths arising from a particular disease, or, of individuals for example, the number of people in a queue waiting to receive a service at a particular time. The INARMA model was originally introduced in the 1980s Mc Kenzie (1985), Al-Osh and Alzaid (1987). The INARMA models have been proposed for forecasting time series of counts, and have received wider attentions in the last three decades. This model has been shown to be analogous to well-known conventional time series model namely Autoregressive Moving Average (ARMA) models by Box et al. (1994) for modelling continuous data.

The study and analysis of count time series poses several problems and questions. For instance, a common distribution that is used in practice to model the response time series, is the Poisson distribution. Such an assumption is sensible because the Poisson distribution is the simplest discrete distribution, yet its properties are satisfactory to cover a large class of problems (as cited in Christou, 2013).

Researchers have investigated the classes of INARMA models with the assumption that the innovation distribution are Poisson distribution. But the Poisson distribution has a feature of equal mean-variance relationship which makes it inadequate for modeling time series count data because most of the count data have properties of overdispersion. In this research, we investigate INAR(1) model with the assumption that the innovation distributions are Negative Binomial distribution. The Negative Binomial distribution is capable of taking into account the overdispersion found in time series count data. This research will fill a major gap in the literature.

Modeling discrete -valued time series is the most challenging and, yet, least well-developed of all areas of research in time series. The fact that variate values are integer, renders most traditional representations of dependence either impossible or impractical. In the past there have been a number of imaginative attempts to develop a suitable class of models (as cited in McKenzie, 2000). In recent times, Fokianos (2012), Davis and Liu (2015) have made an effort to the development of models appropriate for discrete valued time series. Such data usually occur in the form of counts rendering the traditional ARMA-type models impractical. Steutel and Harn (1979) proposed the most popular count time series models that are based on the notion of binomial thinning. These models namely the integer-valued autoregressive (INAR) processes, were introduced by McKenzie (1985), Al-Osh and Alzaid (1987) as a convenient way to transform the usual autoregressive structure to discrete valued-time series. Several attempts have been made to extent and generalize the simplest INAR(1) process. One of the most interesting but less developed generalization that have appeared in the literature is the extension of INAR-type models to the multi-dimensional space. Most attempts to this direction considered the bivariate case (n=2) since the complexity of the model increases rapidly for n>2. Xanthi and Dimitris (2014) considered a simplified version of the multivariate INAR(1) process proposed by Pedeli and Karlis (2013) where the innovation distribution are assumed to be independent random variables. Therefore, cross-correlation between the series of the multivariate process is only due to the non-diagonal autocorrelation matrix A. It is shown that such a specification is extremely advantageous in terms of practical implementation without significant precision losses. They used some multivariate time series earthquakes to illustrate the model. Its appropriateness for syndromic surveillance and outbreak detection purposes is also discussed.

For the Innovation distribution ε_t, of INARMA models, many models have been proposed in the literature for the integer-valued time series count data. The Poisson distribution is often assumed as the distribution of ε_t in the INARMA models. The Poisson distribution has a characteristic of equidispersion. In practice, however, count data are overdispersed in nature relative to the Poisson distribution. For this reason, the INARMA models with Poisson innovations is not always suitable for modeling integer-valued time series, therefore, several models which describe the over-dispersion phenomena have been discussed in the statistical literature.

One common approach is to change the thinning operation in the INAR(1) model. Weiß (2018) summarized several alternative thinning operators, such as random coefficient thinning, iterated thinning and quasi-binomial thinning operator to the extended binomial case.

Changing the distribution of innovations is also used to modify the INAR(1) model. Jung et al. (2005) indicated that the INAR(1) model with negative binomial innovation (NB- INAR(1)) is appropriate for generating overdispersion. Jazi et al.(2012) defined a zero-inflated Poisson ZIP(p, λ)for innovation (ZIP- INAR(1)), because a frequent occurrence in overdispersion is that the incidence of zero counts is generated than expected from the Poisson distribution. Jazi et al. (2012) proposed a modification of INAR(1) model with Geometric innovation (G-INAR(1)) for modeling overdispersed count data. Schwer and Weiß (2014) investigated the compound Poisson INAR(1) (CP- INAR(1)) model, which is suitable for fitting data sets with overdispersion. According to Schwer and Weiß (2014) the negative binomial distribution and the geometric distribution both belonging to the compound Poisson distribution. Livio et al. (2018) presented the INAR(1) model with the Poisson-Lindely innovations, that is, PL-INAR(1) model. Bourgnignon et al. (2019) introduced the INAR(1) model with double Poisson (DP- INAR(1)) and generalized Poisson innovations (GP-INAR(1)) model. Qi et al. (2019) considered zero-order one-inflated INAR(1)-type models, and Cunha et al. (2021) introduced an INAR(1) model with Borel innovation to model zero truncated count time series. Huang and Zim (2021) introduced a new INAR(1) model with Bell innovations (BL- INAR(1)). Huang and Zim (2021) used a relative simple distribution introduced by Castellares et al. (2018) for innovation. Mahmoudi and Rostami (2020) introduced a first-order nonnegative integer-valued moving average process with power series innovations based on a Poison thinning operator (PINMAPS(1)) for modeling overdispersed and underdispersed count time series. Bouguinon and Vasconcellos (2015) introduced INAR(1) processes with power series innovations. Yu and Wang (2021) introduced a new overdispersed integer-valued moving average model with dependent counting series. In this paper we investigate the theoretical properties of INAR(1) model with NB innovation, and assess the practical validity and applicability of the main results of the study on real life data.

2. Methodology

2.1 The Binomial Thinning Operator

Before introducing the INAR(1) model, we first introduced the meaning of Binomial thinning operation and it properties.

The binomial thinning operation was defined by Steutel and Harn (1979). Suppose Y is a non-negative integer-valued random variable. Then, for any $\alpha \in [0,1]$, the thinning operation "∘" is defined by:

$$\alpha \circ Y = \sum_{i=1}^{y} x_i \tag{2.1}$$

Where $\{X_i\}$ is a sequence of i.i.d. Bernoulli random variables, independent of Y, and with a constant probability that the variable will take the value of unity:

$$P\ X_i = 1 - P\ X_i = 0 = \alpha \tag{2.2}$$

Some of the properties of the thinning operation can be obtained as follows:

(1) $0 \circ Y = 0$

(2) $1 \circ Y = Y$

(3) $\alpha \circ (\beta \circ Y) \overset{d}{=} (\alpha\beta) \circ Y$

(4) $(\circ Y) = \alpha\, E(Y)$

(5) $E\ (\alpha \circ Y)^2 = \alpha^2 (Y^2 \quad + \alpha\ 1 - \alpha\ (Y)$

(6) $\mathrm{var}\ \alpha \circ Y = \alpha^2 \mathrm{var}\ Y + \alpha\ 1 - \alpha\ E(Y)$

2.2 Integer-Valued First Order Autoregressive (INAR(1)) Model

The Integer-valued first order Autoregressive INAR(1) model is defined by

$$y_t = \alpha \circ y_{t-1} + z_t \tag{2.3}$$

Where $\alpha \in (0,1)$, and z_t is a sequence of i.i.d non-negative integer-valued random variables, independent of $y_t \sim (\mu_{z,} \sigma^2{}_{z,})$, z_t and y_{t-1} are assumed to be stochastically independent for all points in time, and the thinning operator "∘" is defined via:

$$\alpha \circ y = \sum_{i=1}^{y} x_i \tag{2.4}$$

Where x_i is a sequence of independently and identically distributed (i.i.d.), Bernoulli random variables, independent of y, and with a constant probability that the variable will take value of unity.

$$\mathrm{P}(x_t=1) = 1-\mathrm{p}(x_t=0) = \alpha \tag{2.5}$$

The process obtained by equation (2.3) is stationary and it resembles the Gaussian AR(1) process except that it is nonlinear due to the thinning operation "∘" replacing the scalar multiplication in continuous models.

Equation (2.3) shows that, based on the definition of the thinning operation, the memory of an INAR(1) model decays exponentially as has been shown (Al-Osh and Alzaid, 1987).

2.3 Method of Estimation

The Conditional Least Square (CLS) estimation method was employed in this research. Lawrence and Paul (1978) developed the Conditional Least Square (CLS) estimation procedure for stochastic processes based on the minimization of a sum of squared deviations about conditional expectation.

It can be easily seen that in the INAR(1) model, Y_t given Y_{t-1} is still a random variable due to the definition of the thinning operation. The conditional mean of Y_t given Y_{t-1}, which is the best one-step-ahead predictor as has been shown (Brännäs and Hall, 2001) is:

$$E\ (Y_t / Y_{t-1} = \alpha Y_{t-1} + \lambda = (\boldsymbol{\theta}, Y_{t-1}) \tag{2.6}$$

where $\boldsymbol{\theta} = (\alpha, \lambda)'$ is the vector of parameters to be estimated. Al-Osh and Alzaid (1987) employed a procedure developed by Klimko and Nelson (1978) and derived the estimators for α given by:

$$\hat{\alpha} = \frac{\sum_{t=1}^{n} Y_t Y_{t-1} - (\sum_{t=1}^{n} Y_t \sum_{t=1}^{n} Y_{t-1})/n}{\sum_{t=1}^{n} Y_{t-1}^2 - (\sum_{t=1}^{n} Y_{t-1})^2/n} \tag{2.7}$$

2.4 Forecasting Method

One of the objectives of a time series models is to forecast the future values of a time series observations.

2.4.1 Minimum Mean Square Error (MMSE) Forecasts

The conditional expectation has been the most commonly used forecasting procedure discussed in the time series literature (Freeland and McCabe, 2004b). The main advantage of this method, apart from being simple, is that it produces forecasts with minimum mean square error (MMSE).

Minimum mean square error (MMSE) forecasts are used to find $\hat{Y}_{T+h,} h = 1,2,\dots,H$ of the processes Y_t based on the observed series of $\{Y_1,\dots,Y_T\}$. The MMSE forecast of the process is given by:

$$\hat{Y}_{T+h,} = E(Y_{T+h}|Y_T,\dots,Y_1) \tag{2.8}$$

This method yields forecasts with minimum MSE. For an INAR(p) model, we have:

$$\hat{Y}_{T+h} = \alpha_1 Y_{T+h-1} + \alpha_2 Y_{T+h-2} + \cdots + \alpha_p Y_{T+h-p} + \mu \tag{2.9}$$

Where the Y values on the RHS of equation (2.9) may be either actual or forecast values as has been shown (Du and Li, 1991; Jung and Tremayne, 2006b).

2.4.2 Lead Time Forecasting for an INAR(1) Model

For the INAR(1) process of $Y_t = \alpha \circ Y_{t-1} + Z_t$, the cumulative Y over lead time l is given by:

$$\sum_{j=1}^{l+1} Y_{t+j} = Y_{t+1} + Y_{t+2} + \cdots + Y_{t+l+1}$$

$$= (\alpha \circ Y_{t-1} + Z_{t+1}) + (\alpha^2 \circ Y_t + \alpha \circ Z_{t+1} + Z_{t+2})$$

$$+ \cdots + (\alpha^{l+1} \circ Y_t + \alpha^l \circ Z_{t+1} + \alpha^{l-1} \circ Z_{t+2} + \cdots + Z_{t+l+1}) \tag{2.10}$$

Because $\alpha \circ X + \beta \circ X \neq (\alpha + \beta) \circ X$, the above equation can be written as:

$$\sum_{j=1}^{l+1} Y_{t+j} = \sum_{j=1}^{l+1} \sum_{i=1}^{n_j^1} \psi_{ij}^1 \circ Y_t + \sum_{j=1}^{l+1} \sum_{i=1}^{n_j^2} \psi_{ij}^2 \circ Z_{t+k_{ij}} \tag{2.11}$$

Where $n_j^1$is the number of Y_t terms in each of $\{Y_{t+j}\}_{j=1}^{l+1}$ in equation (2.11), ψ_{ij}^1 is the corresponding coefficient for each Y_t, n_j^2 is the number of $Z_{t+k_{ij}}$ terms in each of $\{Y_{t+j}\}_{j=1}^{l+1}$ in equation (2.10), ψ_{ij}^2 is the corresponding coefficient for each $Z_{t+k_{ij}}$. All of these terms are explained below.

It can be seen that because the process is an integer autoregressive of order one, each of $\{Y_{t+j}\}_{j=1}^{l+1}$ yields only one Y_t, in equation (2.11); therefore, n_j^1=1. The corresponding coefficient for Y_t in each of $\{Y_{t+j}\}_{j=1}^{l+1}$ (sayY_{t+2}) is obtained from α thinned the coefficient of Y_t in the previous term (in this case Y_{t+1}). As a result, ψ_{ij}^1= α^j.

It can be seen from equation (2.10) that due to the repeated substitution of Y_{t+j}, the number of $Z_{t+k_{ij}}$ increases in each of $\{Y_{t+j}\}_{j=1}^{l+1}$. This number, shown by n_j^2, can be obtained from n_{j-1}^2+1. This means that each of $\{Y_{t+j}\}_{j=1}^{l+1}$ (sayY_{t+2}) has one of more Z compared to the previous one (which is Y_{t+1}in this case). The corresponding coefficient for each $Z_{t+k_{ij}}$ shown by Ψ_{ij}^2, is α thinned the corresponding coefficient in the previous term $\alpha \circ \Psi_{1(j+1)}^2$.) $t + k_{ij}$ is the subscript of innovation terms in each of $\{Y_{t+j}\}_{j=1}^{l+j}$ and from equation (2.11) it can be easily seen that k_{ij} is given by

$$k_{ij} = \{_{j}^{k_{i(j-1)}} \;\; {}_{for\; n_{j-1}^{2}<i \leq n_j^2}^{for\; 1 \leq i \leq n_{j-1}^{2}} \; for\; j = 1, \ldots, l+1 \tag{2.12}$$

Based on equation (2.11), the conditional expected value of the aggregated process:

$$E\left(\sum_{j=1}^{l+1} Y_{t+j} \mid Y_t\right) = \left(\sum_{j=1}^{l+1}\sum_{i=1}^{n_j^1} \Psi_{ij}^{1}\right) Y_t + \left(\sum_{j=1}^{l+1}\sum_{i=1}^{n_j^2} \Psi_{ij}^{2}\right)\mu = \frac{\alpha(1-\alpha^{l+1})}{1-\alpha}\, Y_t$$

$$+ \left(\sum_{j=1}^{l+1}\sum_{i=1}^{j} \alpha^{i-1}\right)\mu = \frac{\alpha(1-\alpha^{l+1})}{1-\alpha}\, Y_t + \frac{\mu}{1-a}\left[(l+1) - \sum_{j=1}^{l+1}\alpha^{j}\right] \tag{2.13}$$

Therefore, at time T, when Y_T is observed, the lead time forecast can be obtained from:

$$E\left(\sum_{j=1}^{l+1} Y_{T+j} \mid Y_T\right) = \frac{\alpha(1-\alpha^{l+1})}{1-\alpha}\, Y_T + \frac{\mu}{1-a}\left[(l+1) - \sum_{j=1}^{l+1}\alpha^{j}\right] \tag{2.14}$$

2.5 The Negative Binomial (NB) Distribution

The innovation distribution assumed in this research is the negative binomial distribution. The negative binomial distribution has two parameters: the mean μ and the shape parameter or the dispersion parameter k, which is commonly considered to be fixed to measure overdispersion. For a sample of counts X that fits a negative binomial distribution (X ~ NB(μ, k)), the variance of the distribution is

$\mu + \mu^2 / k$. The probability that the variable *X* takes the value x is:

$$\text{Prb}[X=x]= \frac{\Gamma(x+k)}{x!\Gamma(k)}\left(\frac{\mu}{\mu+k}\right)^{x}\left(1+\frac{\mu}{k}\right)^{-k} = \frac{(x+k-1)(x+k-2)\ldots(k+1)k}{x!}\left(\frac{\mu}{\mu+k}\right)^{x}\left(1+\frac{\mu}{k}\right)^{-k}, \; \mu, k>0,\; x=0,1,2,\ldots \tag{2.15}$$

Where $\Gamma(.)$ denotes the gamma function defined by:

$$\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt. \tag{2.16}$$

From the probability density function of the negative binomial distribution, it can be seen that k is an essential part of the model. Estimation of k is thus important given a sample of counts.

In this research, the method of maximum likelihood estimator (MLE) is adopted to estimate the mean and the dispersion parameter of the NB. According to Fisher, the log-likelihood function from a sample of independent identically distributed (i.i.d.) variate (x_i', s) is proportional to:

$$l(\text{k},\; \mu) = \frac{1}{n}\sum_{i=1}^{n} log\left(\frac{\Gamma(x_i+k)}{\Gamma(k)}\right) + \bar{x}\log(\text{u}) - (\bar{x}+k)\log(1+\frac{\mu}{\text{k}}) \tag{2.17}$$

Where μ is again the mean of the negative binomial distribution. The sample variate are integers in practice, which yields:

$$\frac{\Gamma(x+k)}{\Gamma(k)} = (x+k-1)(x+k-2)\ldots(k+1)k. \text{ the term } log\left(\frac{\Gamma(x_i+k)}{\Gamma(k)}\right) \tag{2.18}$$

then can be written as:

$$log\left(\frac{\Gamma(x_i+k)}{\Gamma(k)}\right) = \sum_{v=0}^{x_i-1} klog\left(1+\frac{v}{k}\right) \tag{2.19}$$

Without call to the gamma function.

Thus, the log-likelihood function can finally be expressed by:

$$l(\text{k},\; \mu) = \frac{1}{n}\sum_{i=1}^{n}\sum_{v=0}^{x_i-1} klog\left(1+\frac{\mu}{k}\right) + \bar{x}\log(\mu) - (\bar{x}+k)\log(1+\frac{\mu}{\text{k}}) \tag{2.20}$$

With gradient elements

$$\nabla_{\mu}\text{l} = \frac{\bar{x}}{\mu} - \frac{1+\bar{x}/k}{1+\mu/k} \text{ and}$$

$$\nabla_{k}\text{l} = \frac{1}{n}\sum_{i=1}^{n}\sum_{v=0}^{x_i-1}\left(\frac{v}{1+v/k}\right) + k^2\log(1+\frac{\text{u}}{\text{k}}) - \frac{\mu(\bar{x}+k)}{1+\mu/k}. \tag{2.21}$$

From the gradient element, setting $\nabla_{\mu} l=0$ yields $\hat{\mu}=\bar{x}$. Then the MLE of k can be obtained via a nonlinear root-finder by setting $\nabla_{k} l=0$ and given $\mu = \hat{\mu}$.

3. Results and Interpretations

This section focuses on the result which is based on the simulation study of theoretical investigation of the class of INAR(1) model with NB innovation. The study makes use of the Conditional Least squares (CLS) estimate to estimate the parameter of INAR(1) model, and Maximum Likelihood Estimate (MLE) to estimate the mean and the dispersion parameter of the NB distribution.

3.1 Estimation of Parameter For INAR(1) Model and NB Distribution

A simulation experiment based on theoretical generated data were addressed under different parameter values and different sample sizes. The Monte Carlo simulations were conducted with a code written in R, all results were based on 1000 runs.

In estimating the parameter of INAR(1) model, we make use of equation (2.7), with the following parameter setting: α=0.2, 0.6, and 0.8. n=30, 90, 120, and 600, with number of replication r= 1000 times.

In estimating the mean and dispersion parameter of the NB distribution, equation (2.20) were used with the following parameter setting: μ =0.85, 1.5 and 2 respectively. K:1,2, and 4 for each μ, with number of replication r= 1000 times. The results are shown below:

Table 3.1. Parameter Estimate of CLS Estimator for INAR(1) Series Replication=1000

Estimator	**Parameter setting (α)**	**Parameter Estimate and S.E**	**Sample Size (n)**			
			30	**90**	**120**	**600**
Conditional Least Square	**0.2**	$\hat{\alpha}$	0.0924	0.0943	0.0945	0.0937
		S.E	0.1771	0.1805	0.1808	0.1797
	0.6	$\hat{\alpha}$	0.5488	0.5469	0.5465	0.5464
		S.E	0.1482	0.1538	0.1555	0.1566
	0.8	$\hat{\alpha}$	0.8133	0.8170	0.7975	0.8094
		S.E	0.1161	0.1134	0.1203	0.1181

Table 3.1 presents the results of the parameter estimates of the Integer Auto regressive of order 1 (INAR(1)) model. The first row reports the parameter estimates of the model, while the second row reports the standard errors (S.E) of the estimates obtained by simulation. Results are based on 1000 replication.

The results confirmed that, the standard error (SE) produced by the Conditional Least Squares (CLS) increases as the number of samples increases. The SE reduces as the parameter values increases. This means that estimating the parameter of INAR(1) model is better when the number of observations is small and the parameter value is high.

Table 3.2. Estimation of K and μ of Negative Binomial Distribution for n=30,90,120, and 600

Sample Size (n)	μ	K=1	K=2	K=4
30	0.85	$\hat{k}$=0.4965 $\hat{\mu}$=1.0666 AIC=87.8352	$\hat{k}$=0.5387 $\hat{\mu}$=0.600 AIC=66.8343	$\hat{k}$=0.4967 $\hat{\mu}$=0.5334 AIC=62.7172
	1.5	$\hat{k}$=0.6576 $\hat{\mu}$=1.7331 AIC=87.8352	$\hat{k}$=0.5936 $\hat{\mu}$=1.000 AIC=84.8273	$\hat{k}$=0.9652 $\hat{\mu}$=0.9667 AIC=85.7712
	2.0	$\hat{k}$=0.6942 $\hat{\mu}$=2.2664 AIC=123.8003	$\hat{k}$=0.4612 $\hat{\mu}$=1.5340 AIC=102.8366	$\hat{k}$=0.9775 $\hat{\mu}$=1.3001 AIC=98.4756
90	0.85	$\hat{k}$=0.7585 $\hat{\mu}$=0.9332 AIC=244.4199	$\hat{k}$=3.1373 $\hat{\mu}$=0.7999 AIC=222.3790	$\hat{k}$=2.2996 $\hat{\mu}$=0.8667 AIC=232.7188
	1.5	$\hat{k}$=0.8171 $\hat{\mu}$=1.5669 AIC=312.3411	$\hat{k}$=2.2563 $\hat{\mu}$=1.4444 AIC=296.6447	$\hat{k}$=2.7858 $\hat{\mu}$=1.5556 AIC=303.4139
	2.0	$\hat{k}$=1.2249	$\hat{k}$=1.5143	$\hat{k}$=4.5232

		$\hat{\mu}$=1.9217 AIC=341.3869	$\hat{\mu}$=2.0444 AIC=348.6402	$\hat{\mu}$=2.1221 AIC=335.6395
120	0.85	$\hat{k}$=0.8268 $\hat{\mu}$=0.9501 AIC=327.8408	$\hat{k}$=2.9590 $\hat{\mu}$=1.008 AIC=331.9965	$\hat{k}$=2.5928 $\hat{\mu}$=0.9418 AIC=321.632
	1.5	$\hat{k}$=0.8957 $\hat{\mu}$=1.6167 AIC=421.4639	$\hat{k}$=2.1894 $\hat{\mu}$=1.6752 AIC=421.6663	$\hat{k}$=2.6640 $\hat{\mu}$=1.5831 AIC=406.9943
	2.0	$\hat{k}$=1.1353 $\hat{\mu}$=1.9331 AIC=455.4091	$\hat{k}$=1.9922 $\hat{\mu}$=2.2333 AIC=476.1481	$\hat{k}$=2.8983 $\hat{\mu}$=2.1333 AIC=458.4345
600	0.85	$\hat{k}$=1.0327 $\hat{\mu}$=0.7931 AIC=1481.292	$\hat{k}$=1.8040 $\hat{\mu}$=0.8583 AIC=1532.379	$\hat{k}$=1.7264 $\hat{\mu}$=0.8099 AIC=1443.602
	1.5	$\hat{k}$=1.0882 $\hat{\mu}$=1.450 AIC=1991.472	$\hat{k}$=2.2187 $\hat{\mu}$=1.5402 AIC=2010.441	$\hat{k}$=6.5003 $\hat{\mu}$=1.4667 AIC=1901.274
	2.0	$\hat{k}$=1.0480 $\hat{\mu}$=1.9848 AIC=2287.914	$\hat{k}$=2.3424 $\hat{\mu}$=2.0767 AIC=2279.278	$\hat{k}$=5.2949 $\hat{\mu}$=1.9482 AIC=2149.624

Table 3.2 present the result of the estimation results for K and μ of NB at different simple sizes. Comparing the AIC of the result at different K values and at different sample sizes, the estimation produced less AIC with low sample sizes especially when n=30. However, as the number of k increases the result showed a decreased in the value of AIC. This means that estimation of K of NB distribution is better when the number of observations is small and the more the dispersion the better for the estimation.

3.2 Forecasting in INAR(1) Model With NB Innovation

This section concentrate on the investigation of the forecasting accuracy of INAR(1) model, with NB innovation. The forecast accuracy at different lead time period l=1, 3, 5, 7, 9, and 15 were investigated with codes written in R statistical package. All results were based on 1000 runs.

At time T, when Y_T is observed, the lead time forecast is obtained using equation (2.14), with the following Parameter values: l=1, 3, 5, 7, 9, and 15. α= 0.83, μ=0.85, $j = 1, \ldots, l + 1$, and Y_T=30, 90, 120, 600, number of replication r=1000. The result is summarized in the table 3.3.

Table 3.3. MMSE of lead Time Forecasts For INAR(1) Series With mean of NB Distribution

Y_T	l=1	l=3	l=5	l=7	l=9	l=15
30	0.140044	0.140044	0.140051	0.140058	0.140102	0.165879
90	0.03265801	0.03265801	0.0326592	0.03266839	0.03267158	0.03269515
120	0.03928588	0.03928588	0.039355	0.03942412	0.03949324	0.0397006
600	0.009020943	0.009020943	9.032981e-03	9.045019e-03	9.057057e-03	9.093171e-03

Table 3.3 presents the results of the MMSE forecasts for INAR(1) series with NB innovation. The results showed that, the MMSE produced when the number of lead times forecasts between one and five were less than that produced when the numbers of lead times forecast were greater than five. The MMSE increased when the number of lead time periods increases. This result indicates that, forecasts with this class of model is better with short time frame of predictions.

3.3 Application to COVID-19 Data Set

The theoretical investigations result obtained in this study was applied to the number of deaths arising from COVID-19 in Nigeria. The count time series data consists of 48 observations (weekly data), from January 2021 to December 2021. The data was obtained from the Nigeria Centre for Disease Control (NCDC), and analyzed with the aid of R statistical package. The results and the interpretation of the analysis are presented below:

Table 3.4. Descriptive Statistic of COVID-19 Death Cases

	Covid19 Death Cases
Mean	28
Median	14.5 i.e. 15
Maximum	90
Minimum	1
Variance	876.16
Observations	48

Table 3.4 depicts the summary statistic of the number of deaths arising from COVID-19 in Nigeria in 2021. From the table, the mean and the variance respectively are 28 and 876.16 which is evident of overdispersion.

Table 3.5. Preliminary Test

Test type	Test value	p-value	Decision
Overdispersion	Z=6.3476	1.094e-10	Reject H_0
Autocorrelation	Chi-square=73.062	5.107e-15	Reject H_0

Table 3.5 presents the Preliminary test of the number of deaths from COVID-19 in Nigeria in the year 2021. The results suggest that the null hypotheses (H0) (i.e. no true dispersion and autocorrelation in the series and its residuals respectively) cannot be accepted, thus there is true dispersion in the Covid19 death series and presence of autocorrelation in the residuals of the Covid19 death series, which corroborates the descriptive analysis. Hence, a negative binomial distribution is assumed for the innovation.

Fig3.1 and Fig3.2 respectively depicts the plots of ACF and PACF respectively. Based on the information supplied by the plots, the candidates' models in table3.6 were suggested. Comparing the AIC of the models in table3.6. the INAR(1) model gives the minimum AIC and hence an INAR(1) model best fit the data set.

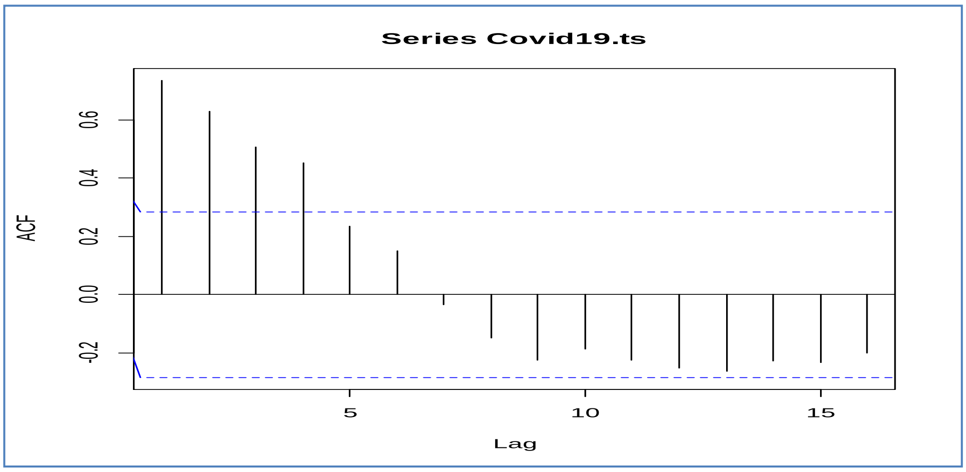

Figure 3.1. ACF Plot of Covid-19 Death cases

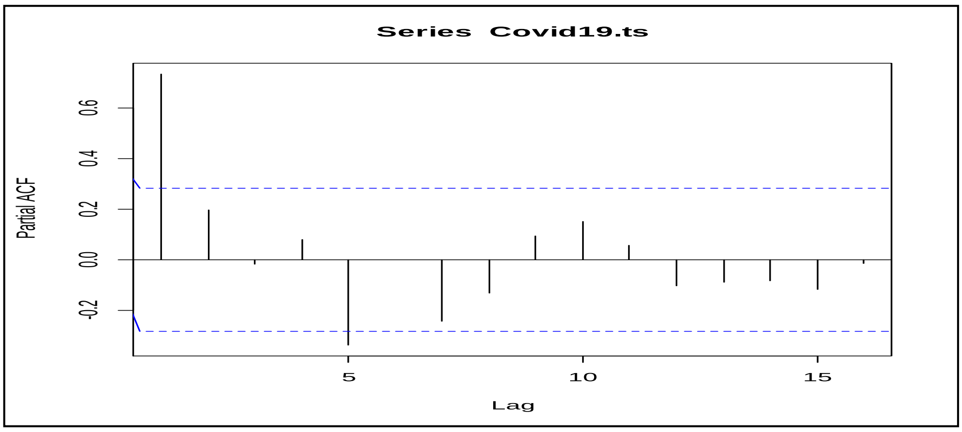

Figure 3.2. PACF plot of Covid19 Death cases

Table 3.6. Candidates of INARMA Models

Candidate Models	AIC
INARMA(1,0)	414
INARMA(0,1)	416
INARMA(2,0)	417
INARMA(1,1)	416
INARMA(2,1)	419

Table 3.7 Parameter Estimate of INMA(1) Model

Model	Parameter estimate	Std. Error	p-value
INAR(1)	-0.32582	0.13628	0.01681

Table3.7 presents the estimate of parameter of INAR(1) model. In line with the simulation result, Conditional Least Square (CLS) estimation method was employed. The parameter value α = -0.32582 with standard error of 0.13628, the model is found to be statistically significant at 5% level of significance ($p< 0.05$).

Table 3.8. Lead Time Forecast of the Covid-19 Data Using INAR(1) Model

Lead time	1	3	5	7	9
Forecast	38.5552	38.5161	38.5119	38.5115	38.5114
MMSE	9.3760	9.3760	9.3761	9.3772	9.3871

Table 3.8 depicts the lead time forecast of the number of death arising from Covid-19 in Nigeria, using the fitted INAR(1) model. The accuracy of the forecast is measured by the MMSE. The error of forecast increases as the number of lead time increased. This result is in line with the theoretical investigations obtained in this study. This shows that forecasting with this class of model is better with short time prediction. The results show that, the number of death shows a decreasing trends as the number of leads time increases,

4. Conclusion

From our findings, the following conclusions were drawn:

The results of the estimation of parameter of the INAR(1) model confirmed that the standard error (SE) produced by the Conditional Least Squares (CLS) increases as the number of samples increases, the error reduced as the parameter values increases. This means that estimating the parameter of INAR(1) model is better when the number of observations is small and the parameter value is high.

The result of the estimation of the parameters (K and μ) of NB distribution at different simple sizes. Comparing the Akaike Information Criterial (AIC) of the result at different K values and at different sample sizes, the estimation produced less AIC with low sample sizes especially when n=30. However, as the number of k increases the result showed a decreased in the value of AIC. This means that estimation of K of NB distribution is better when the number of observations is small and the more the dispersion the better for the estimation.

The forecasting accuracy were measured by the MMSE. The results showed that, the MMSE produced when the number of lead time forecasts between one and five were less than that produced when the numbers of lead times forecast were greater than five. The MMSE increased when the number of lead time periods increases. This result indicates that, forecasts with this class of model is better with short time frame of predictions.

Lastly, the theoretical investigations were validated with a real life data using the number of death arising from Covid-19 in Nigeria. The results obtained corroborates the results from the theoretical investigations.

Refrences

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

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis: Forecasting and control.* (3rd ed.). Prentice Hall: New Jersey.

Brännäs, K., & Hall, A. (2001). Estimation in integer-valued moving average models. *Applied Stochastic Models in Business and Industry, 17*(3), 277-291. https://doi.org/10.1002/asmb.445

Castellares, F., & Ferrari, S. L. P., & Lemonte, A. J. (2018). On the Bell distribution and its associated regression model for count data. *Appl. Math. Model. 56,* 172–185. https://doi.org/10.1016/j.apm.2017.12.014

Cunha, E. T. D., Bourguignon, M., & Vasconcellos, K. L. P. (2021). On shifted integer-valued autoregressive model for count time series showing equidispersion, underdispersion or overdispersion. Commun. Stat. https://doi.org/10.1080/03610926.2020.1725822

Davis, R. A., & Liu, H. (2015). Theory and inference for a class of observation-driven models with application to time series of counts. Statistical sinica. https://doi.org/10.5705/ss.2014.145t

Du, J. G., & Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model. *Journal of Time Series Analysis, 12*(2), 129-142. https://doi.org/10.1111/j.1467-9892.1991.tb00073.x

Eddie Mc Kenzie. (2000). Discrete Variate Time Series.

Eisa, M., & Ameneh, R. (2020). First-order integer-valued moving average process with power series innovation. *Journal of Statistical Theory and applications, 19*(3), 415-431. https://doi.org/10.2991/jsta.d.200917.001

Fokianos, K. (2012) Count time series models. In T Subba Rao S Subba Rao and CR Rao(eds). Handbook of statistics 30: Time series – methods and applications, pages 315–47. Amsterdam: Elsevier B.V. https://doi.org/10.1016/B978-0-444-53858-1.00012-0

Freeland, R. K., & McCabe, B. P. M. (2004b). Forecasting discrete valued low count time series. *International Journal of Forecasting, 20*(3), 427-434. https://doi.org/10.1016/S0169-2070(03)00014-1

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,* 954–63. https://doi.org/10.1111/j.1467-9892.2012.00809.x

Jie, H., & Fukang, Z. (2021). A new first-order integer-valued autoregressive model with Bell innovations. *Entropy (Basel), 23*(6), 713. https://doi.org/10.3390/e23060713

Jung, R. C., & Tremayne, A. R. (2006b). Coherent forecasting in integer time series models. *International Journal of Forecasting, 22*(2), 223-238. https://doi.org/10.1016/j.ijforecast.2005.07.001

Jung, R. C., Ronning, G., Tremayne, A. R. (2005) Estimation in conditional first order autoregression with discrete support. *Statistical Papers, 46*(2), 195–224. https://doi.org/10.1007/BF02762968

Kaizhi, Y., & Huiqiao, W. (2021). A new overdispersed integer-valued moving average model with dependent counting series. *Entropy, 23*, 706. https://doi.org/10.3390/e23060706

Lawrence, A., Klimko, & Paul, I. N. (1978). On Conditional Least Square estimation for stochastic processes. *The Anals of statistics, 6*(3), 629-642. https://doi.org/10.1214/aos/1176344207

Livio, T.; Mamode, K. N., Bourgignon, M., & Bakouch, H. S. (2018). An INAR(1) model with Poisson–Lindley innovations. *Econ. Bull, 38*, 1505–1513.

Marcelo, B., & Klaus, L. P. V. (2015). First order non-negative integer valued autoregressive processes with power series innovations. *Braz. J. Probab. Stat., 29,* 71-93. https://doi.org/10.1214/13-BJPS229

Marcelo. Bourguignon, J. Rodrigues and M. Santos-Neto (2019). Extended Poisson INAR(1) processes with equidispersion, underdispersion and overdispersion. Journal of Applied Statistics vol.46-issue 1. https://doi.org/10.1080/02664763.2018.1458216

McKenzie, E. (1985). Some simple models for discrete variate series. *Water Resources Bulletin, 21*(4), 645-650.

https://doi.org/10.1111/j.1752-1688.1985.tb05379.x

Pedeli, X., & Karlis, D. (2013a) On composite likelihood estimation of a multivariate INAR(1) model. *Journal of Time Series Analysis, 34,* 206–20. https://doi.org/10.1111/jtsa.12003

Qi, X., Li, Q., & Zhu, F. (2019). 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. 346*, 572-590. https://doi.org/10.1016/j.cam.2018.07.043

Schweer, S., & Weiß, C. H. (2014). Compound Poisson INAR(1) processes: stochastic properties and testing for overdispersion. *Computational Statistics and Data Analysis, 77*, 267–84. https://doi.org/10.1016/j.csda.2014.03.005

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

Vasiliki, C. (2013). Statistical Theory for mixed Poisson time series models. A PhD thesis submitted to the University of Cyprus Nicosia, Cyprus.

Weiß, C. H. (2018). *An introduction to discrete-valued time series* (1st ed.). Chichester, England: John Wiley & Sons, Inc. https://doi.org/10.1002/9781119097013

Xanthi, P., & Dimitris, K. (2014). A bivariate INAR(1) process with application- SAGE journals vol.11. issue 4.