JIRSS (2012)
Vol. 11, No. 2, pp 173-190

Integer Valued AR(1) with Geometric Innovations

Mansour Aghababaei Jazi[1], Geoff Jones[2], Chin-Diew Lai[2]

[1]Faculty of Mathematics, University of Sistan and Baluchestan, Zahedan, Iran.
[2]Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand.

Abstract. The classical integer valued first-order autoregressive (INAR(1)) model has been defined on the basis of Poisson innovations. This model has Poisson marginal distribution and is suitable for modeling equidispersed count data. In this paper, we introduce an modification of the INAR(1) model with geometric innovations (INARG(1)) for modeling overdispersed count data. We discuss some structural mathematical properties of the process comparing with classical INAR(1). Also, the superiority of the model in contrast with the INAR(1) is shown by some real time series.

Keywords. Binomial thinning operator, conditional maximum likelihood estimation, infinitely divisible distributions, mixture distributions.

MSC: 62M10.

1 Introduction

Recently there has been growing interest in modeling discrete-time dependent integer valued time series. In developing such models the integer valued first-order autoregressive (INAR(1)) process, introduced independently by McKenzie (1985) and Al-Osh and Alzaid (1987), has received

Mansour Aghababaei Jazi(✉)(m.aghababaeijazi@gmail.com), Geoff Jones (G.Jones @massey.ac.nz), Chin Diew Lai (C.Lai@massey.ac.nz)
Received: January 1, 2012; Accepted: May 14, 2012

considerable attention. This model has been defined on the basis of binomial thinning operator $\circ$ of Steutel and van Harn (1979).

Definition 1.1. Let X be a nonnegative integer valued random variable (r.v.). Then for any $\alpha \in [0,1]$ the operator $\circ$ is defined by

$$\alpha \circ X = \sum_{i=1}^{X} B_i(\alpha), \tag{1}$$

where counting series $B_i(\alpha)$ is a sequence of independent identically distributed (iid) binary r.v.'s with $P(B_i(\alpha) = 1) = 1 - P(B_i(\alpha) = 0) = \alpha$.

Definition 1.2. A discrete-time dependent integer valued time series $\{X_t\}$ is said to follow an INAR(1) model, if

$$X_t = \alpha \circ X_{t-1} + W_t, \quad t = 0, 1, 2, ..., \tag{2}$$

where $\{W_t\}$ are iid nonnegative integer valued r.v.'s with some discrete distribution and independent of all the counting series $B_i(\alpha)$'s in (1) and X_{t-1}.

Note that $\alpha \in (0,1)$ implies a stationary dependent integer valued time series, whereas $\alpha = 0$ and $\alpha = 1$ imply independence and non-stationarity for X_t, respectively. The process X_t satisfying (2) is second-order stationary if $0 \leq \alpha < 1$, with autocovariance function $\mathrm{Cov}(X_t, X_{t-k}) = \sigma^2 \alpha^k$ (Al-Osh and Alzaid, 1987; Du and Li, 1991).

The coefficient α can be interpreted as the proportion of observations counted at time $t-1$ that still remains at time t. W_t is known as the innovations (or repositions) produced at time t and $\alpha \circ X_{t-1}$ can be interpreted as the number of survivors at time t from the previous period. The INAR(1) process as defined by (2) can be viewed as both a Markov process and a GaltonWatson branching process with immigration.

Al-Osh and Alzaid (1987) provide the following representation for the marginal distribution of the INAR(1) model expressed in terms of the innovation sequence W_t,

$$X_t = \sum_{i=0}^{\infty} \alpha^i \circ W_{t-i} \quad . \tag{3}$$

The above representation implies that the marginal distribution of a stationary INAR(1) process is discrete self-decomposable, i.e., the probability generating function (pgf) of X_t, denoted by $\phi_X(s)$, satisfies

$$\phi_X(s) = \phi_X(1 - \alpha + \alpha s)\phi_w(s) \tag{4}$$

where ϕ_w is the pgf of innovations W_t. In fact, assuming that the INAR(1) process is stationary, the pgf of X_t satisfies (4). Thus, one can choose any member of the class of discrete self-decomposable distributions as the marginal distribution of a stationary INAR(1) model. It is known that discrete self-decomposability implies infinite divisibility and unimodality (see Steutel and Van Harn, 2004). Subsequently, the stationary marginal distribution of a stationary INAR(1) is infinitely divisible and unimodal. Further, a stationary INAR(1) process has a Poisson marginal distribution (with mean θ) if and only if the innovations also follow a Poisson distribution (with mean $(1-\alpha)\theta$). So, INAR(1) with Poisson innovations is a suitable model for equidispersed count data wherein the mean and variance are the same. For more structural and asymptotic properties in INAR(1) process with Poisson marginal, we refer the reader to Silva and Oliveira (2004), Pavlopoulos and Karlis (2006) and Park and Oh(1997).

Autoregressive moving-average processes with Poisson marginal distribution are appropriate for equidispersed time series of counts (where the mean is the same as the variance). In practice, however, some discrete time dependence count data may be overdispersed, i.e., the variance is greater than the mean. McKenzie (1986) developed autoregressive moving-average processes with negative binomial and geometric marginal distributions as analogues of well-known continuous variates models for gamma and negative exponential variates. Also, Alzaid and Al-Osh (1988) considered INAR(1) processes (2) with geometric marginal distribution for time series of overdispersed counts. They denoted the model by GINAR(1) and showed that the innovations are distributed as a mixture of a degenerated distribution at zero with mass α and a geometric distribution with mass $1-\alpha$.

Time series of overdispersed counts can also be modeled through overdispersed innovations. Let μ_w and σ_w^2 be the mean and variance of the innovations, then the mean and variance of the stationary solution of INAR(1) model (2) are

$$\mu = E(X_t) = \frac{\mu_w}{1-\alpha} \quad and \quad \sigma^2 = Var(X_t) = \frac{\alpha\mu_w + \sigma_w^2}{1-\alpha^2},$$

respectively. Thus, the index of dispersion of the solution, $ID(X) = \sigma^2/\mu$, is related to that of the innovations, $ID(W) = \sigma_w^2/\mu_w$, according to the formula

$$ID(X) = [1 + ID(W)/\alpha]/[1 + 1/\alpha].$$

showing that the marginal distribution of the solution $\{X_t\}$ is overdispersed (i.e., $\sigma^2 > \mu$) if and only if the distribution of the innovations$\{W_t\}$ is overdispersed too (i.e., $\sigma_w^2 > \mu_w$). Also, for any fixed value of α, $ID(X)$ increases when $ID(W)$ does, which in turn implies that in order for INAR(1) to be a suitable model for a series of counts with arbitrary large (sample) $ID(X)$, the distribution of innovations must necessarily be such as to allow for quite large values of $ID(W)$ too, in a rather flexible way. In light of this reasoning, e.g., Pavlopoulos and Karlis (2008) considered INAR(1) process (2) with innovations from finite mixtures of Poisson distributions.

In this paper, we study the INAR(1) process (2) with geometric innovations, denoted by INARG(1). The motivation for such a process arises from its potential in the modeling of some overdispersed integer-valued time series.

We also suppose here that a geometric r.v. (innovation) W, denoted by $W \sim \text{Ge}(\pi)$, has probability mass function (pmf)

$$P(W = k) = \pi(1-\pi)^k I_{\{0,1,\ldots\}}(k), \tag{5}$$

where $\pi \in (0,1)$ and the indicator function $I_A(x)$ equals to 1, if $x \in A$ else equals to zero. Also, the pgf of W is given by

$$\phi_w(s) = \frac{\pi}{1-(1-\pi)s}$$

It is easy to show that the mean and the variance of W with the pmf (5) are $\mu_w = (1-\pi)/\pi$ and $\sigma_w^2 = (1-\pi)/\pi^2$, respectively, with $ID(W) = 1/\pi > 1$.

The contents of this paper are organized as follows. Some mathematical and structural properties of the INARG(1) process are derived in Section 2. In Section 3, we discuss parameter estimation and forecasting for our models. In Section 4, we use simulation to study the marginal distribution of INARG(1) process and comparing two maximum likelihood based estimation approaches. Finally, in Section 5, we fit both Poisson INAR(1) and INARG(1) models on some real time series to show the superiority of the INARG(1) model over the traditional INAR(1) with Poisson innovations.

2 INAR(1) With Geometric Innovations

In this section we introduce an INAR(1) process with geometric innovations and derive some mathematical and structural properties of the

corresponding marginal distribution of the process, e.g., the mean, variance, autocovariance and (conditional) likelihood functions. Also, we compare the average run length of zeros and proportions of zeros in the INARG(1) process and classic INAR(1) process with Poisson innovations to derive a simple check on the distribution of innovations in the INAR(1) processes.

We modify classic INAR(1) model of McKenzie (1985) and Al-Osh and Alzaid (1987) and Alzaid and Al-Osh(1988) as follow.

Definition 2.1. $\{X_t\}$ is said to follow a INAR(1) process with geometric innovation, denoted by INARG(1), if

$$X_t = \alpha \circ X_{t-1} + W_t, \quad t = 0, 1, 2, ..., \tag{6}$$

where $\{W_t\}$ are iid Ge(π) r.v.'s independent of X_{t-1} and operator $\circ$ as defined in (1).

2.1 Some Mathematical Properties

In the following theorems, suppose that $\{X_t\}$ follow an INARG(1) process (6) with Ge(π) innovations.

Theorem 2.1. *The mean and variance of X_t are*

$$\mu = \frac{1-\pi}{\pi(1-\alpha)} \tag{7}$$

and

$$\sigma^2 = \frac{1-\pi}{\pi^2(1-\alpha)}, \tag{8}$$

respectively.

Proof. We have

$$\begin{aligned} E[X_t|X_{t-1}] &= E[\alpha \circ X_{t-1} + W_t|X_{t-1}] \\ &= E[\alpha \circ X_{t-1}|X_{t-1}] + E[W_t|X_{t-1}] \\ &= \alpha X_{t-1} + (1-\pi)/\pi \end{aligned} \tag{9}$$

and

$$\begin{aligned} \mathrm{Var}[X_t|X_{t-1}] &= Var[\alpha \circ X_{t-1} + W_t|X_{t-1}] \\ &= \mathrm{Var}[\alpha \circ X_{t-1}|X_{t-1}] + \mathrm{Var}[W_t|X_{t-1}] \\ &= \alpha(1-\alpha)X_{t-1} + (1-\pi)/\pi^2 \end{aligned} \tag{10}$$

Hence, by (9), we get

$$\mu = EE[X_t|X_{t-1}] = \alpha\mu + (1-\pi)/\pi$$

which implies that the mean of the INARG(1) process is as (7).

Also, by (9) and (10), we have

$$\begin{aligned}\sigma^2 &= \mathrm{Var}(E[X_t|X_{t-1}]) + E(\mathrm{Var}[X_t|X_{t-1}]) \\ &= \mathrm{Var}[\alpha X_{t-1} + (1-\pi)/\pi] + E[\alpha(1-\alpha)X_{t-1} + (1-\pi)/\pi^2] \\ &= \alpha^2\sigma^2 + \alpha(1-\alpha)\mu + (1-\pi)/\pi^2\end{aligned}$$

which implies that the variance of the INARG(1) process is as (8). □

Remark 2.1. By a similar argument used for INAR(1) (see Al-Osh and Alzaid, 1987) we can show that the INARG(1) is second-order stationary for all $0 \le \alpha < 1$, with $\mathrm{Cov}(X_t, X_{t-k}) = \sigma^2\alpha^k$ which depends only on the lag k. Also, the variance-to-mean ratio (index of dispersion) of X_t is $ID(X_t) \equiv \sigma^2/\mu = 1/\pi$ so the INARG(1) is a model for overdispersed integer-valued time series.

The following theorem implies that X_t is a unimodal infinitely divisible r.v.

Theorem 2.2. *X_t is a discrete self-decomposable r.v.*

Proof. Clearly, by a recursive substitution, X_t can be written as (3) wherein the innovations W_t's are iid Ge(π) r.v.'s with the pgf $\phi_w(s) = \pi/[1-(1-\pi)s]$. Thus, the pgf of X_t is

$$\begin{aligned}\phi_X(s) &= \prod_{i=0}^{\infty} \phi_w(1-\alpha^i+\alpha^i s) \\ &= \prod_{i=0}^{\infty} \frac{\pi_i}{1-(1-\pi_i)s} \qquad (11)\end{aligned}$$

where

$$\pi_i = \frac{\pi}{\pi + (1-\pi)\alpha^i}, \quad i = 0, 1, \ldots. \qquad (12)$$

Clearly, the pgf of $\phi_X(s)$ satisfies (4), i.e, X_t is self-decomposable. □

Remark 2.2. By (11), X_t can be written as

$$\begin{aligned} X_t &= \sum_{i=0}^{\infty} Z_i \\ &= Z_0 + \sum_{i=1}^{\infty} Z_i \end{aligned} \tag{13}$$

where Z_i's $\sim$ Ge(π_i) are independent r.v.'s and $Z_0 \sim$ Ge(π_0) has the same distribution as the innovations W_t and, consequently, we have

$$\alpha \circ X_{t-1} \stackrel{d}{=} \sum_{i=1}^{\infty} Z_i \tag{14}$$

with the pgf

$$\phi_X(1-\alpha+\alpha s) = \prod_{i=1}^{\infty} \frac{\pi_i}{1-(1-\pi_i)s}.$$

Therefore, by (13) and (14), the stationary INARG(1) time series X_t and $\alpha \circ X_t$ can be represented as an infinite sum of Ge(π_i) r.v.'s with growing parameter π_i given by (12). It follows that for small α and large π the marginal distribution will itself approximate a geometric.

2.2 Marginal and Joint Distribution

The INARG(1) model with innovations $W_t \sim$ Ge(π) r.v.'s forms a stationary discrete time Markov chain with transition probabilities

$$\begin{aligned} p_{ij} &\equiv P(X_t = j | X_{t-1} = i) \\ &= P(\alpha \circ X_{t-1} + W_t = j | X_{t-1} = i) \\ &= \sum_{k=0}^{\min(i,j)} P(\alpha \circ X_{t-1} = k | X_{t-1} = i) P(W_t = j-k) \\ &= \sum_{k=0}^{\min(i,j)} \binom{i}{k} \alpha^k (1-\alpha)^{i-k} \pi (1-\pi)^{j-k} I_{\{0,1,\ldots\}}(j-k), \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i,j = 0,1,\ldots. \end{aligned}$$

giving the probability of going from state i to state j in a single step.

Then, the marginal probability function of X_t is given by

$$\begin{aligned} p_j &\equiv P(X_t = j) \qquad (15)\\ &= \sum_{i=0}^{\infty} p_{ij} P(X_{t-1} = i) \\ &= \sum_{i=0}^{\infty} \sum_{k=0}^{\min(i,j)} \binom{i}{k} \alpha^k (1-\alpha)^{i-k} [\pi(1-\pi)^{j-k} I_{\{0,1,\ldots\}}(j-k)] p_i, \\ & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j = 0, 1, \ldots. \end{aligned}$$

which is a mixture distribution.

In order to find the joint probability function, we use the first order dependence of the process. It leads to the following simplified expression:

$$\begin{aligned} f(i_1, i_2, \ldots, i_n) &\equiv P(X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n) \qquad (16)\\ &= P(X_1 = i_1) P(X_2 = i_2 | X_1 = i_1) P(X_3 = i_3 | X_2 = i_2) \cdots \\ &\cdots P(X_n = i_n | X_{n-1} = i_{n-1}) \\ &= p_{i_1} \prod_{s=1}^{n-1} \sum_{k=0}^{\min(i_s, i_{s+1})} \binom{i_s}{k} \alpha^k (1-\alpha)^{i_s - k} [\pi(1-\pi)^{i_{s+1}-k} \\ &\times I_{\{0,1,\ldots\}}(i_{s+1} - k)] \end{aligned}$$

2.3 Distribution of Zeros

The introduction of the INARG(1) model is motivated by the presence of overdispersion in a nonnegative integer valued series. Since excess zeros are a common feature of overdispersed count data, we therefore consider the distribution of zero values in such a series. For the INARG(1) process the transition probabilities from a zero to zero and nonzero values are equal to π and $1 - \pi$, respectively. The run length of zeros in the process is defined as the number of zeros between two nonzero values and follows a geometric distribution with termination probability $1 - \pi$. Thus, the average run length of zeros in the INARG(1) time series is independent of α and is given by $\eta = 1/(1 - \pi)$ that is longer than the average run length of zeros for the INAR(1) with Po(λ) innovations, for $\lambda > -log(\pi)$.

Also, by the proof of Theorem 2.2, we have the following theorem.

Theorem 2.3. *The proportion of zeros in the INARG(1) process is*

given by

$$P(X_t = 0) = \prod_{i=0}^{\infty} \frac{\pi}{\pi + (1-\pi)\alpha^i}$$

whereas the proportion of zeros in the INAR(1) process (with Poisson innovations) is $P(X_t = 0) = e^{-\lambda/(1-\alpha)}$.

Remark 2.3. The proportion of zeros in the INAR(1) process (with Po(λ) innovations) is exponentially related to the marginal mean $\lambda/(1-\alpha)$, i.e., $P(X_t = 0) = e^{-E(X_t)}$. But there is no such a relation between the marginal mean $E(X_t) = (1-\pi)/[\pi(1-\alpha)]$ and the proportion of zeros $P(X_t = 0)$ for the marginal of the INARG(1) process given by Theorem 2.3. In fact a simulation study on the INARG(1) process indicates that in this case $P(X_t = 0) > e^{-E(X_t)}$. This suggests a simple check on the distribution of innovations in the INAR(1) processes regarding the proportions of zeros. Consider a time series $x_1, x_2, ..., x_n$ with sample mean $\bar{x}$ and proportion of zeros $f_0 = \#0/n$. Then, $f_0 \cong e^{-\bar{x}}$ favors Poisson innovations while $f_0 \gg e^{-\bar{x}}$ favors geometric innovations.

3 Parameter Estimation and Forecasting

Let $\mathbf{x} = (x_1, x_2, ..., x_n)$ be an observed time series following INAR(1) model (2) with Ge(π) innovations and with mean and variance μ and σ^2 given by (7) and (8), respectively. We shall present two maximum likelihood approaches for estimation of the parameters of the model.

3.1 Maximum Likelihood Estimation

From the joint probability function (16), we can write the likelihood function as

$$\begin{aligned} L(\pi, \alpha|\mathbf{x}) &= f(x_1, x_2, ..., x_n) \\ &= p_{x_1} \prod_{i=1}^{n-1} \sum_{k=0}^{\min(x_i, x_{i+1})} \binom{x_i}{k} \alpha^k (1-\alpha)^{x_i - k} [\pi(1-\pi)^{x_{i+1}-k} \\ &\times I_{\{0,1,...\}}(x_{i+1} - k)], \end{aligned}$$

where p_{x_1} is the pmf for x_1. Since the marginal distribution is intractable in general, a simple approach is to condition on the observed X_1, essentially ignoring its contribution and estimating the parameters by conditional maximum likelihood (CML).

Alternatively, as noted in Remark 2.2, for large π and small α the marginal distribution will itself approximate a Ge(π^*). Equating first moment we find

$$\pi^* = \frac{1}{\mu + 1}$$

so an approximate full maximum likelihood (AFML) estimation is found by maximizing

$$\begin{aligned} L(\lambda, \rho, \alpha|\mathbf{x}) &= \pi^*(1-\pi^*)^{x_1} \prod_{i=1}^{n-1} \sum_{k=0}^{\min(x_i, x_{i+1})} \binom{x_i}{k} \alpha^k (1-\alpha)^{x_i - k} \\ &\times \ [\pi(1-\pi)^{x_{i+1}-k} I_{\{0,1,\ldots\}}(x_{i+1} - k)] \end{aligned}$$

In the next section, we use simulation to study the limiting marginal distribution and comparative performance of AFML and CML estimation.

3.2 Forecasting

One of the most common procedures for forecasting the mean value is to use conditional expectation. Applying the properties of the binomial thinning operator $\circ$ leads to the following predictor of the mean:

$$\hat{X}_{t+h}|X_{t-1} = E[X_{t+h}|X_{t-1}], \qquad h = 0, 1, \ldots.$$

By conditioning arguments as in the proof of Theorem 2.1, and recursive substitution, we have

$$\begin{aligned} E[X_{t+h}|X_{t-1}] &= \alpha^{h+1} X_{t-1} + \mu_w \sum_{k=0}^{h} \alpha^k \\ &= \alpha^{h+1} X_{t-1} + \mu_w \frac{1-\alpha^{h+1}}{1-\alpha} \end{aligned}$$

where μ_w is the mean of W_t's. Clearly, $\lim_{h\to\infty} E[X_{t+h}|X_{t-1}] = \mu_w/(1-\alpha)$, which is the unconditional (marginal) mean of the process.

Also, the conditional variance of X_{t+h} is given by

$$\begin{aligned}
\mathrm{Var}[X_{t+h}|X_{t-1}] &= \mathrm{Var}[\alpha^{h+1} \circ X_{t-1}|X_{t-1}] + \mathrm{Var}[\sum_{k=0}^{h} \alpha^k \circ W_{t+h-k}|X_{t-1}] \\
&= \alpha^{h+1}(1-\alpha^{h+1})X_{t-1} + \sum_{k=0}^{h}[\alpha^k(1-\alpha^k)\mu_w + \alpha^{2k}\sigma_w^2] \\
&= \alpha^{h+1}(1-\alpha^{h+1})X_{t-1} + \mu_w \frac{1-\alpha^{h+1}(1-\alpha^{h+1})}{1-\alpha(1-\alpha)} \\
&+ \sigma_w^2 \frac{1-\alpha^{2h-2}}{1-\alpha^2}
\end{aligned}$$

where σ_w^2 is the variance of W_t's. This variance converges to $\mu_w/[1-\alpha(1-\alpha)] + \sigma_w^2/[1-\alpha^2]$ as $h \to \infty$.

In practice the values of π and α will be replaced by their corresponding maximum likelihood estimates.

4 Simulation

Figure 1 shows the sample paths of simulated INAR(1) processes with Ge(π) innovations for $\pi = 0.1, 0.5$ and $\alpha = 0.2, 0.5$. As we can see from (7) and (8), for larger α and less π we have larger mean and variance so a tendency to yield larger values, but for smaller values of α and larger π sample paths tend to smaller values and frequently returns to zero with less mean and variance.

Figure 2 and the corresponding Table 1 illustrate the marginal distributions of the simulated series as described above. The bar plots show that for large π and small α, the empirical marginal distribution decays geometrically, whereas for smaller π, and particularly for larger α, the marginal distribution begins to resemble a more complex mixture, as we expected by Remark 2.2 and the pmf (15). We propose negative binomial for the limiting marginal distribution of the INARG(1), because negative binomial distribution, nb(r, p), can arise as a mixture of Poisson distributions with mean distributed as a Gamma$(r, (1-p)/p)$ distribution. In this model, the mean and variance are $r(1-p)/p$ and $r(1-p)/p^2$, respectively, wherein p and r can be easily estimated via moment method. The goodness-of-fit tests confirm that in this case the negative binomial gives a reasonable fit to the marginal distribution.

Table 1: Goodness of fit of negative binomial model for marginal distribution of INARG(1)

	$(\pi,\ \alpha)$			
	$(0.1, 0.2)$	$(0.1, 0.5)$	$(0.5, 0.2)$	$(0.5, 0.5)$
$\hat{p}$	0.114	0.139	0.555	0.622
$\hat{r}$	1.431	2.947	1.497	3.305
χ^2-Statistic	58.136	67.279	9.323	6.990
p-value	0.050	0.090	0.156	0.430

Table 2 contains the mean and mean squared error (MSE) of estimates by the conditional maximum likelihood (CML) estimation and approximate full maximum likelihood (AFML) estimation methods. The estimates were computed by simulating 100 series (of length 100) from the INARG(1) processes for some chosen π and α. The AFML, which uses a (geometric) approximation to the marginal pmf (for larger π) for the first observation, performed very slightly better than the CML which ignores (conditions on) the first observation. However for smaller π, when we know that the approximation to the marginal is not as good, the AFML performs slightly worse. Given these results, there seems to be no advantage in using AFML for series of length 80 or more (typical of the real datasets we examine in the next section).

Table 2: Mean (MSE) of estimates from fitting INARG(1) by CML and AFML

$(\ \pi,\ \alpha)$		$\hat{\pi}$	$\hat{\alpha}$
(0.1, 0.2)	AFML	0.09806 (0.00050)	0.20645 (0.00136)
	CML	0.10297 (0.00012)	0.20378 (0.00123)
(0.1, 0.5)	AFML	0.09782 (0.00061)	0.51485 (0.00202)
	CML	0.10436 (0.00016)	0.50751 (0.00072)
(0.5, 0.2)	AFML	0.50398 (0.00183)	0.19586 (0.00689)
	CML	0.50348 (0.00185)	0.19603 (0.00690)
(0.5, 0.5)	AFML	0.50718 (0.00175)	0.50712 (0.00352)
	CML	0.50713 (0.00176)	0.50713 (0.00352)

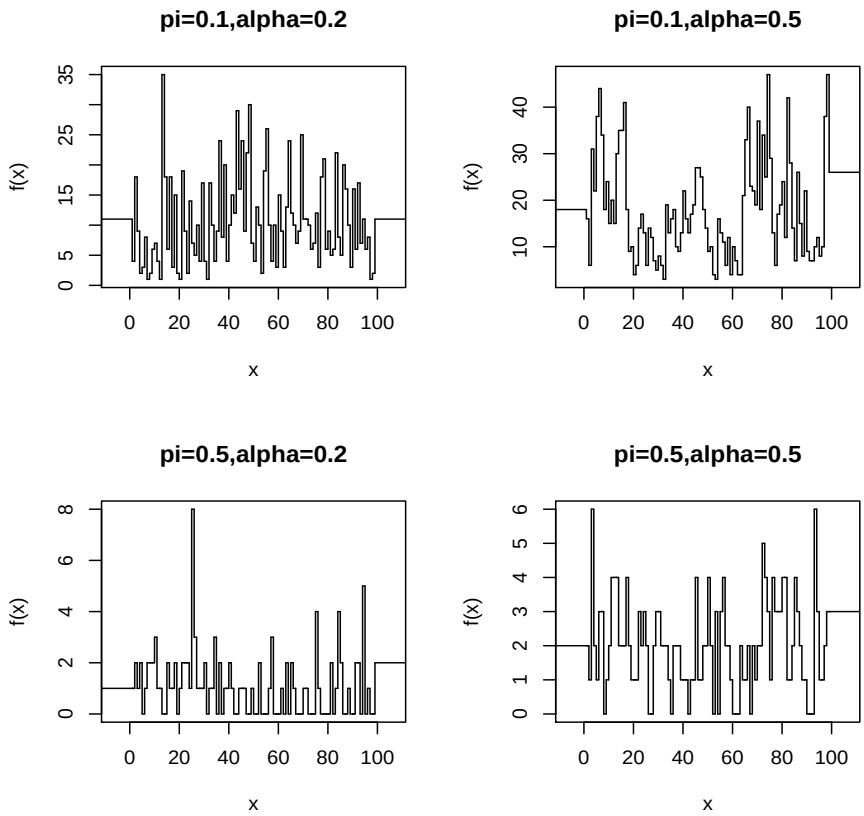

Figure 1: Sample path of INARG(1) process for $\pi = 0.1, 0.5$ and $\alpha = 0.2, 0.5$.

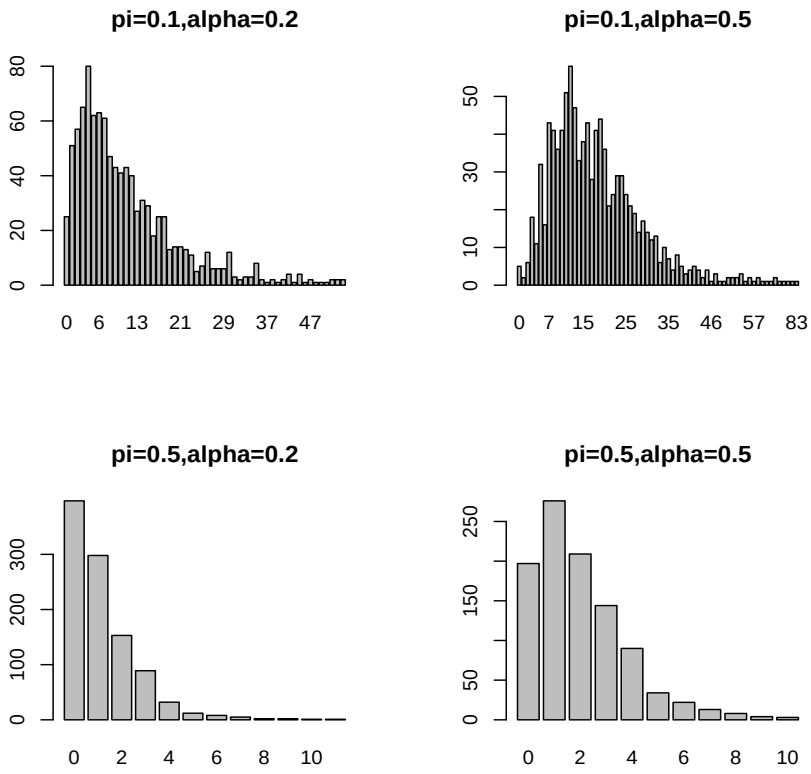

Figure 2: Bar-plots of limiting marginal distribution of INARG(1) for $\pi = 0.1, 0.5$ and $\alpha = 0.2, 0.5$

5 Data analysis

We now illustrate with some real time series data the ability of the INARG(1) to improve on the fit of the traditional INAR(1) with Poisson innovations. Our data give numbers of submissions to animal health laboratories, monthly 2003-2009, from a region in New Zealand. The submissions can be categorized in various ways. Here we consider one series giving the total number of bovine cases, and several others categorized by presenting symptoms. One such is the number of submissions with sudden death, given in Table 3. The sample path and autocorrelation function (ACF) of this series are shown in Figure 3. The series has a large proportion of zero values, and some long runs of zeros. The ACF

suggests first-order dependence. We would expect some positive correlation in such series because the underlying processes causing disease will change smoothly in time.

Table 3: Time series: Sudden death ($\bar{x} = 2.024$, $s^2 = 6.529$, $f_0 = 0.357$, $e^{-\bar{x}} = 0.132$)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2003	2	3	3	0	1	2	3	8	8	8	1	1
2004	2	0	3	5	1	1	2	6	2	2	1	2
2005	0	0	1	2	4	2	0	0	0	3	0	1
2006	0	0	0	0	3	1	1	7	6	4	1	0
2007	0	0	0	0	0	0	0	4	2	3	5	0
2008	0	0	0	0	2	3	9	14	5	3	2	1
2009	0	3	1	1	2	2	2	3	0	0	0	0

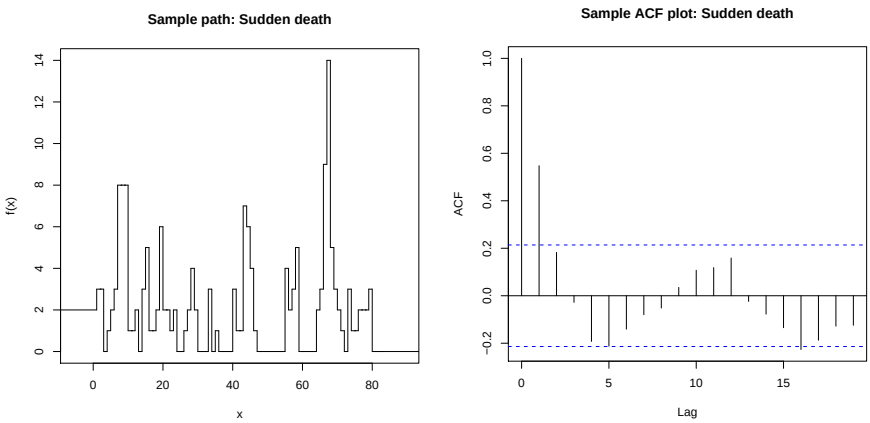

Figure 3: Sample path and ACF plot of Sudden death submissions

We fitted both INAR(1) (with Poisson innovations) and INARG(1) models to Sudden death series by conditional maximum likelihood. By comparing their Akaike's information criterion (AIC) under each model, we conclude that the INARG(1) model with $W_t \sim \text{Ge}(0.421)$ innovations,

$$X_t = 0.317 \circ X_{t-1} + W_t, \tag{17}$$

yields a much better fit (with AIC=305.999) than the traditional INAR(1) model (with AIC=346.4521).

By selected model (17), the predicted values of Sudden death series are

$$\hat{X}_1 = (1-\hat{\pi})/[\hat{\pi}(1-\hat{\alpha})] = 2.014$$

and

$$\begin{aligned} \hat{X}_i &= \hat{\alpha}X_{i-1} + (1-\hat{\pi})/\hat{\pi} \\ &= 0.317X_{i-1} + 1.375, \end{aligned}$$

for $i = 2, 3, ..., 84$. Figure 4 shows the closeness of these predicted values to the sample paths of Sudden death series.

This can be repeated with the other laboratory submission series, giving the results shown in Table 4. We can see that in all cases the improvement from the INARG(1) model is highly significant. We also compare the proportion of zeros f_0 with $e^{-\bar{x}}$ (Table 5), suggesting for each series that the proportion of zeros in the data is greater than expected from INAR(1) with Poisson innovations. It is interesting that even for the series of total bovine cases, which has few zeros, the INARG(1) fits better because the mean is large and so zeros would not be expected under the INAR(1) model.

Table 4: Fitting INAR(1) and INARG(1) models by CML estimation method

	INAR(1)			INARG(1)		
Data	$\hat{\lambda}$	$\hat{\alpha}$	AIC	$\hat{\pi}$	$\hat{\alpha}$	AIC
Bovine	12.023	0.406	1332.840	0.055	0.167	677.1467
Abort	0.982	0.353	309.0855	0.483	0.294	268.6886
Diarrhoea	4.103	0.342	748.2907	0.164	0.187	482.0749
Illthrift	2.422	0.385	512.9148	0.252	0.247	406.6647
Anorexia	0.511	0.385	224.4407	0.637	0.315	195.1021
Skin lesions	1.172	0.173	302.5026	0.444	0.118	273.9057
Sudden death	1.240	0.383	346.4521	0.421	0.317	305.999

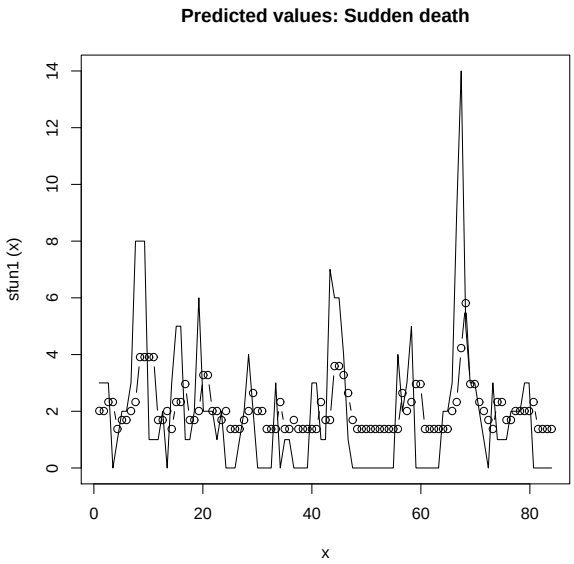

Figure 4: Predicted values of Sudden death series

Table 5: Observed (INAR(1)) proportion of zeros of the laboratory submission series

Data	f_0	$e^{-\bar{x}}$
Bovine	0.071	1.21E-09
Abort	0.440	0.223
Diarrhoea	0.214	0.002
Illthrift	0.238	0.020
Anorexia	0.667	0.440
Skin lesions	0.405	0.240
Sudden death	0.357	0.132

Acknowledgements

We thank Lachlan McIntyre of the Ministry of Agriculture and Forestry, New Zealand for providing the data on animal health laboratory submissions. We also thank an Associate Editor for his/her careful reading of the manuscript and helpful comments.

References

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

Alzaid, A. A. and Al-Osh, M. A. (1988), First-order integer-valued autoregressive (INAR(1)) process: distributional and regression properties. Statistica Neerlandica, **42**(1), 53-61.

Du, J. G. and Li, Y. (1991), The integer-valued autoregressive (INAR(p)) model. Journal of Time Series Analysis, **12**, 129-142.

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

McKenzie, E. (1986), Autoregressive moving-average processes with negative binomial and geometric marginal distributions, Advanced Applied Probability, **18**, 679-705.

Park, Y. and Oh, C. W. (1997), Some asymptotic properties in INAR(1) processes with Poisson marginals. Statistical Papers, **38**, 287-302.

Pavlopoulos, H. and Karlis, D. (2006), On time series of overdispersed count: Modelling, inference, simulation and prediction. Technical Report, Department of Statistics Athens University of Economics and Business, **221**, 1-38.

Pavlopoulos, H. and Karlis, D. (2008), INAR(1) modeling of overdispersed count series with an enviromental application, Environmetrics, **19**, 369-393.

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

Steutel, F. W. and van Harn, K. (1979), Discrete analogues of self-decomposability and stability. The Annals of Probability, **7**(5), 893-899.

Steutel, F. W. and van Harn, K. (2004), Infinite divisibility of probability of distributions on the real line, New York: Marcel Dekker.