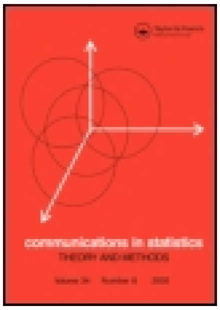

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

First order autoregressive time series with negative binomial and geometric marginals

Mohamed A. Al-Osh [a] & Emad-Eldin A.A. Aly [b]

[a] Department of Statistics, King Saud University, P.O Box 2455, Riyadh, Saudi Arabia

[b] Department of Statistics & Applied Probability, The University of Alberta Edmonton, Edmonton, Alberta, T6G 2G1, Canada

Version of record first published: 27 Jun 2007.

To cite this article: Mohamed A. Al-Osh & Emad-Eldin A.A. Aly (1992): First order autoregressive time series with negative binomial and geometric marginals, Communications in Statistics - Theory and Methods, 21:9, 2483-2492

To link to this article: http://dx.doi.org/10.1080/03610929208830925

COMMUN. STATIST.—THEORY METH., 21(9), 2483-2492 (1992)

FIRST ORDER AUTOREGRESSIVE TIME SERIES WITH NEGATIVE BINOMIAL AND GEOMETRIC MARGINALS

MOHAMED A. AL-OSH and EMAD-ELDIN A.A. ALY

Department of Statistics
King Saud University
P.O. Box 2455
Riyadh, Saudi Arabia

Department of Statistics
& Applied Probability
The University of Alberta
Edmonton, Alberta
Canada T6G 2G1

Key Words and Phrases: Negative binomial processes, generalized geometric AR(1) processes, binomial compounding.

ABSTRACT

In this paper we present first order autoregressive (AR(1)) time series with negative binomial and geometric marginals. These processes are the discrete analogues of the gamma and exponential processes introduced by Sim (1990). Many properties of the processes discussed here, such as autocorrelation, regression and joint distributions, are studied.

1. INTRODUCTION

In recent years there has been increasing interest in developing stationary time series models with discrete and non-Gaussian marginal distributions. These models may be used in modelling time series with such marginal distributions and in the simulation of sequences of dependent variates. Among the models developed for the continuous case are the two-parameter exponential first order autoregressive (NEAR(1)) model of Lawrance and Lewis (1981), the gamma models of Gaver and Lewis (1980) and the beta-gamma models of Lewis, McKenzie and Hugus (1989). Some of the models developed for the discrete case are the negative binomial and the geometric models of McKenzie (1986) and the Poisson and geometric models of Alzaid and Al-Osh (1988).

Recently, Sim (1990) introduced an autoregressive model for stationary gamma processes as a Poisson compounding of exponentials. This model has been used in the study of stochastic reservoir systems with Markovian inputs. Sim (1990) has also generalized the NEAR(1) model to an exponential AR(1) with three parameters.

The purpose of this paper is to introduce processes with negative binomial and geometric marginals which are analogues to the models introduced in Sim (1990). The models introduced here can be utilized for generating correlated negative binomial sequences and for modeling discrete-valued time series with such marginal distribution. The negative binomial distribution is one of the most frequently applied discrete distribution (see, for example, Johnson and Kotz (1969)). The negative binomial process can be viewed as the number of counts in a Poisson process during time intervals whose length is a Gamma process (see Lampard (1968) and McKenzie (1986)).

In section 2 we consider a first order representation of a stationary negative binomial process which is constructed as a random binomial compounding of geometric variates. In section 3 we present a three-parameter first order geometric process. This process is a generalization of the geometric AR(1) process mentioned above. Several properties of each of our two introduced models, such as joint distributions, regression and time reversibility are investigated. Our approach in developing the introduced models follows closely that of Sim (1990).

Throughout the rest of this article we will use $\overline{d}$ to denote $(1-d)$, where d is a generic quantity.

2. A NEGATIVE BINOMIAL AR(1) MODEL

The discrete analogue of the gamma AR(1) model of Sim (1990) is the process defined by:

$$X_n = \alpha * X_{n-1} + \varepsilon_n, \tag{2.1}$$

where the operator $*$ is defined as

$$\alpha * X = \sum_{i=0}^{N(X)} W_i$$

with

(i) W_i are i.i.d geometric random variables with parameter $\frac{\alpha}{1+\alpha}$, $W_0 = 0$

(ii) For each fixed non-negative integer-value x of $X, N(x)$ is a binomial random variable with parameters x and $\lambda, \lambda = \alpha p$ and $0 \leq p \leq 1$,

and the ε_n are i.i.d negative binomial $\left(\frac{\alpha}{1+\alpha}\ , v\right)$ random variables with $v > 0$, and ε_n is independent of $\alpha * X_{n-1}$.

For derivation of the distributional properties of the process (2.1) it is more convenient to use the function $\phi_x(\overline{s}) = E[(1-s)^X]$. Bondesson (1979) has drawn attention to the use of such function and McKenzie (1986) used it in a context similar to ours and called it the alternate probability generating function (p.g.f). Conditional on $X = x$, the alternate p.g.f. of $\alpha * X$ is given by

$$\begin{aligned}
\phi_{\alpha * X|X=x}(\overline{s}) &= E\left[(1-s)^{\sum_{i=0}^{N(X)} W_i} | X = x\right] \\
&= \sum_{n=0}^{x} \left(\frac{\alpha}{\alpha+s}\right)^n \binom{x}{n} \lambda^n (1-\lambda)^{x-n} \\
&= \left(1 - \frac{\lambda s}{\alpha+s}\right)^x,
\end{aligned}$$

where $\phi_{W_i}(\overline{s}) = \alpha/(\alpha+s)$.

By the independence of ϵ_n and $\alpha * X_{n-1}$, the alternate p.g.f. of X_n given $X_{n-1} = x$ is

$$\phi_{X_n|X_{n-1}=x}(\overline{s}) = \left(\frac{\alpha}{\alpha+s}\right)^v \left(1 - \frac{\lambda s}{\alpha+s}\right)^x. \tag{2.2}$$

By (2.2) and the assumption that $\lambda = \alpha p$ we get

$$\phi_{X_n}(\overline{s}) = \left(\frac{\alpha}{\alpha+s}\right)^v \phi_{X_{n-1}}\left(1 - \frac{\alpha p s}{\alpha+s}\right) \tag{2.3}$$

Solving (2.3) recursively gives

$$\phi_{X_n}(\bar{s}) = \left[1 + \frac{1}{\alpha\bar{p}}(1-p^n)s\right]^{-v} \phi_{X_0}\left(1 - \frac{p^n s}{1 + \frac{(1-p^n)s}{\alpha\bar{p}}}\right).$$

As $n \to \infty$ we have

$$\phi_{X_n}(\bar{s}) = \left(\frac{\alpha(1-p)}{\alpha(1-p)+s}\right)^v. \tag{2.4}$$

By (2.4), the limiting distribution of X_n is the negative binomial $\left(\frac{\alpha\bar{p}}{1+\alpha\bar{p}}, v\right)$ distribution.

In the remainder of this section, we will assume that X_0 has the negative binomial $\left(\frac{\alpha\bar{p}}{1+\alpha\bar{p}}, v\right)$ distribution. In this case (2.1) defines a stationary AR(1) process with negative binomial $\left(\frac{\alpha\bar{p}}{1+\alpha\bar{p}}, v\right)$ marginal distribution.

The autocorrelation function ρ_j of the process (2.1) decays exponentially as that of the Gaussian AR(1) and it is given by

$$\rho_j = p^j, \quad j = 0, 1, 2, \ldots$$

The process $\{X_n\}$ of (2.1) is a Markov chain with transition probabilities of the form

$$P\Big(X_n = y | X_{n-1} = x\Big) = \frac{\alpha^v(1-\alpha p)^x}{(1+\alpha)^{v+y}} \left\{ \binom{v+y-1}{y} + \sum_{k=1}^{x} \left(\frac{\alpha}{1+\alpha}\right)^k \right.$$
$$\left. \times \left[\sum_{j=0}^{y} \binom{k+j-1}{j} \binom{v+y-j-1}{y-j}\right] \times \binom{x}{k} \left(\frac{\alpha p}{1-\alpha p}\right)^k \right\}. \tag{2.5}$$

The joint distribution of X_n and X_{n-1} can be determined directly from the conditional probabilities of (2.5). However, to investigate some additional properties of the process $\{X_n\}$ of (2.1) we consider the joint alternate p.g.f. of X_n and X_{n-1} which is given below.

$$\phi_{X_n, X_{n-1}}(\bar{s}_2, \bar{s}_1) = \left[\frac{\alpha^2(1-p)}{\alpha^2(1-p) + \alpha(s_1+s_2) + (1-\alpha p)s_1 s_2}\right]^v. \tag{2.6}$$

The joint alternate p.g.f. of (2.6) is identical to the joint Laplace transform of X_n and X_{n-1} of (2.3) of Sim (1990) except for the coefficient of $s_1 s_2$. The same remark applies to (5.11) of Lampard (1969). Since $\phi_{X_n, X_{n-1}}(\bar{s}_2, \bar{s}_1)$ is symmetric in s_1 and s_2 we conclude that the process $\{X_n\}$ of (2.1) is time reversible. Consequently, the forward and backward regression are identical. The same remark applies to the conditional variances. It can be shown that

$$E\{X_n|X_{n-1}=x\}=px+v/\alpha$$

and

$$Var\{X_n|X_{n-1}=x\}=xp(2+\alpha\overline{p})/\alpha+(1+\alpha)v/\alpha^2.$$

We note here that both $E\{X_n|X_{n-1}=x\}$ and $Var\{X_n|X_{n-1}=x\}$ are linear in x.

Higher order distribution properties of the process can be derived from the joint alternate p.g.f. $\phi_n(\overline{s}_n,\overline{s}_{n-1},\ldots,\overline{s}_1)$, which satisfies the relation:

$$\begin{aligned}\phi_n(\overline{s}_n,\ldots,\overline{s}_1)&=E\left[\prod_{i=0}^{n-1}(1-s_{n-i})^{X_{n-i}}\right]\\&=\left(\frac{\alpha}{\alpha+s_n}\right)^v\phi_{n-1}(\overline{s}_{n-1}\times\overline{G}(s_n),\overline{s}_{n-2},\ldots,\overline{s}_1),\end{aligned}\tag{2.7}$$

where $\overline{G}(s_j)=1-G(s_j)=1-\frac{\alpha p s_j}{\alpha+s_j}$.

The joint alternate p.g.f. can be used to identify the distribution of the total number of counts occuring during the interval $(0,n]$ by setting $s_i=s$ for $i=1,2,\ldots n$ in the final form of $\phi_n(\overline{s}_n,\overline{s}_{n-1},\ldots,\overline{s}_1)$. The bivariate distribution of X_j and X_{k+j} can be obtained from $\phi_n(\overline{s}_n,\overline{s}_{n-1},\ldots,\overline{s}_1)$ by setting $s_r=0$ for $r\neq j,k+j$.

3. A GENERALIZED GEOMETRIC AR(1) MODEL

A generalization of the geometric AR(1) model mentioned in Section 1, is given by:

$$X_n=\alpha I_n\circ X_{n-1}+K_n\circ Z_n,\tag{3.1}$$

where the operator $\circ$ is defined by:

$$\alpha\circ X=\sum_{i=1}^{X}Y_i,$$

where $\{Y_i\}$ is a sequence of i.i.d. binary random variables such that $P(Y_i=1)=\alpha=1-P(Y_i=0)$ and

(i) the $\{Z_n\}$ are i.i.d. geometric random variables with parameter $1-\pi$

(ii) $\{(I_n, K_n) : n = 1, 2, \dots\}$ is a bivariate sequence of i.i.d. random coefficients in which the variables I_n and K_n are correlated with the same probability structure of (3.2) of Sim (1990),

$I_n \setminus K_n$	b	1	total
0	$\frac{\beta b}{1-b}$	$\frac{1-\beta-b}{1-b}$	$1-\beta$
1	$\frac{\beta(\alpha-b)}{1-b}$	$\frac{\beta(1-\alpha)}{1-b}$	β
Total	$\frac{\alpha\beta}{1-b}$	$\frac{1-\alpha\beta-b}{1-b}$	1

and the parameters α, β, b are non-negative with

$$0 \leq b \leq \min(1-\beta, \alpha) < 1; \; 0 \leq \alpha, \beta \leq 1;$$

(iii) Z_n and (I_n, K_n) are mutually independent.

Some properties of the operator $\circ$ which will be used in this section are: (i) $0 \circ X = 0$, (ii) $1 \circ X = X$ and $\alpha \circ (\beta \circ X) \stackrel{d}{=} \alpha\beta \circ X$ (see Al-Osh and Alzaid, 1987). Having these properties, it can be seen that for $\beta = 1$, and consequently $b = 0$, model (3.1) reduces to the geometric AR(1) process of McKenzie (1986) and Alzaid and Al-Osh (1988). When $b = \alpha(1-\beta)$, I_n and K_n are independent variables and the process $\{X_n\}$ reduces to a geometric AR(1) analogues to the NEAR(1) of Lawrance and Lewis (1981).

To investigate the marginal distribution of the process $\{X_n\}$ we consider the alternate p.g.f. of X_n which is given by:

$$\phi_{X_n}(\bar{s}) = (1-\beta)\left(\frac{\lambda}{\lambda + bs}\right)\left(\frac{\lambda + b_1 s}{\lambda + s}\right) + \beta\left(\frac{\lambda}{\lambda + bs}\right)\left(\frac{\lambda + \alpha s}{\lambda + s}\right)\phi_{X_{n-1}}(\overline{\alpha s}), \qquad (3.2)$$

where $\lambda = (1-\pi)/\pi$ and $0 \leq b_1 = b/(1-\beta) \leq 1$. By solving (3.2) recursively and taking X_0 with geometric $(1-\pi)$ distribution we get:

$$\phi_{X_n}(\bar{s}) = \frac{\lambda}{\lambda + s},$$

implying that $\{X_n\}$ is a stationary process with geometric $(1-\pi)$ marginal distribution. The atuocorrelation function ρ_j has the form:

$$\rho_j = (\alpha\beta)^j, \qquad j = 0, 1, 2, \ldots$$

The process $\{X_n\}$ is Markovian with transition probabilities given by:

$$\begin{aligned} P(X_n = y | X_{n-1} = x) &= \frac{\overline{\pi}\pi^{y-x}}{1-b}\{\beta\left[\frac{b}{1-\pi(1-b)}\right]^{y+1}\pi^x + (1-\beta-b)\pi^x \\ &+ \frac{\beta(\alpha-b)}{1-\pi(1-b)}\sum_{k=0}^{y}\binom{x}{k}\alpha^k(\pi\overline{\alpha})^{x-k}\left[\frac{b}{1-\pi(1-b)}\right]^{y-k} \\ &+ \beta\overline{\alpha}\sum_{k=0}^{y}\binom{x}{k}\alpha^k(\pi\overline{\alpha})^{x-k}\} \qquad \text{for } y \le x \\ &= \frac{\overline{\pi}\pi^{y-x}}{1-b}\{\beta\left[\frac{b}{1-\pi(1-b)}\right]^{y+1}\pi^x + (1-\beta-b)\pi^x \\ &+ \beta(\alpha-b)b^{y-x}\left[\frac{1}{1-\pi(1-b)}\right]^{y+1}[\alpha+\pi(b-\alpha)]^x \\ &+ \beta\overline{\alpha}(\alpha+\pi\overline{\alpha})^x\} \qquad \text{for } y > x \end{aligned} \tag{3.3}$$

The joint probability function of X_n and X_{n-1} can be derived from (3.3). However, we consider the alternate p.g.f. of X_{n+j} and X_n which is of the form:

$$\begin{aligned} \phi_j(\overline{s}_1, \overline{s}_2) = \phi_{X_{n+j},X_n}(\overline{s}_1, \overline{s}_2) &= E[(1-s_1)^{X_{n+j}}(1-s_2)^{X_n}] \\ &= \beta\left(\frac{\lambda}{\lambda+bs_1}\right)\left(\frac{\lambda+\alpha s_1}{\lambda+s_1}\right)\phi_{j-1}(\overline{\alpha s_1}, \overline{s}_2) \\ &+ (1-\beta)\left(\frac{\lambda}{\lambda+bs_1}\right)\left(\frac{\lambda+b_1s_1}{\lambda+s_1}\right)\left(\frac{\lambda}{\lambda+s_2}\right), \end{aligned} \tag{3.4}$$

where $\phi_0(s_1, s_2) = E\{[(1-s_1)(1-s_2)]^{X_n}\} = \phi_{X_n}(\overline{s_1+s_2-s_1s_2})$. For $j = 1$, equation (3.4) reduces to

$$\begin{aligned} \phi_{X_{n+1},X_n}(\overline{s}_1, \overline{s}_2) &= \beta\left(\frac{\lambda}{\lambda+bs_1}\right)\left(\frac{\lambda+\alpha s_1}{\lambda+s_1}\right)\left(\frac{\lambda}{\lambda+(\alpha s_1+s_2)-\alpha s_1 s_2}\right) \\ &+ (1-\beta)\left(\frac{\lambda}{\lambda+bs_1}\right)\left(\frac{\lambda+b_1s_1}{\lambda+s_1}\right)\left(\frac{\lambda}{\lambda+s_2}\right). \end{aligned} \tag{3.5}$$

From (3.5) it is clear that $\phi_{X_{n+1},X_n}(\overline{s}_1, \overline{s}_2)$ is not symmetric in s_1 and s_2 and consequently the process $\{X_n\}$ is not time reversible. The forward regression is given by

$$E(X_n | X_{n-1} = x) = \alpha\beta x + \frac{\pi}{1-\pi}(1-\alpha\beta),$$

which is linear in x. The backward regression,obtained by inverting (3.5), is of the form:

$$E(X_{n-1}|X_n = x) = \frac{\pi}{\overline{\pi}} + \frac{\alpha\beta}{\overline{\alpha}\overline{b}\overline{\pi}} + \frac{\alpha\beta}{\overline{\alpha}\,\overline{\pi}(b-\alpha)}\left[\frac{\alpha}{1-\overline{\alpha}\pi}\right]^{x+1} + \frac{\alpha\beta}{\overline{b}\overline{\pi}(\alpha-b)}\left[\frac{b}{1-\overline{b}\pi}\right]^{x+1}, \quad x \geq 0$$

which is not linear in x. This is in agreement with the result of Alzaid and Al-Osh (1988) who proved that in the class of integer-valued AR(1), which is of the form of the standard AR(1), only for the Poisson process the backward regression is linear in X_n.

4. COMMENTS

The autoregression models presented in this paper are the discrete analogues of those of Sim (1990). Whereas the gamma AR(1), as Sim (1990) pointed out, has the same probability structure as the time intervals between events of a certain stochastic reversible counter system, the negative binomial process has the probability structure of the events (counts) on these intervals. The analogy between the two types of processes makes several features common between them. An important feature, common between the processes of Sim (1990) and the processes discussed here, is the fact that their correlation structure can be adjusted independent of the parameters in their respective marginal distributions. This flexibility is particularly useful for simulation.

The negative binomial process of (2.1) can be used to develop an Algorithm parallel to Algorithm II of Sim and Lee (1989) for generating correlated negative binomial variates.

It is worth mentioning that it is possible to derive a negative binomial process analogous to the gamma AR(1) process of Sim (1986). The model in this case has the form

$$X_n = U_n \circ X_{n-1} + \epsilon_n ,$$

where $\{U_n\}$ is a sequence of independent random coefficients with standard Power-function distribution, $F_{U_n}(u) = u^\alpha, \alpha > 0$, defined on the interval $[0,1)$, and U_n is independent of ϵ_n. Following the approach of Sim (1986), it can be shown that if ϵ_n is geometric with parameter p, then X_n is a stationary process with negative binomial $(p, \alpha+1)$ marginal distribution. This process is related to the negative binomial AR(1) process which is the analogue of the gamma beta process discussed by McKenzie (1986).

Among other aspects of the models presented here which need to be investigated are specification, estimation and checking the adequency of the model. For estimating the parameters of the models, two approaches seem to be feasible. The first is the conditional least squares approach as developed by Klimko and Nelson (1978). This approach was considered by Al-Osh and Alzaid (1987) for a similar class of discrete time series. The second approach is the optimal estimating equations as was applied by Thavaneswaran and Abraham (1988) to the problem of estimation for non-linear time series models. Detailed study of these aspects will be reported elsewhere.

ACKNOWLEDGEMENTS

This research was done during a visit of the first author to the University of Alberta supported by an NSERC Canada operating grant of the second author.

REFERENCES

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

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

3. Bondesson, L. (1979). On generalized gamma and generalized negative binomial convolutions. Part I. Scand. Actuarial J., 125–146.

4 Gaver, D.P. and Lewis, P.A.W. (1980). First order autoregressive gamma sequences and point processes. Adv. Appl. Prob. **12**, 727–745.

5. Johnson, N.L. and Kotz, S. (1969) *Discrete Distributions.* Houghton-Mifflin, Boston.

6. Klimko, L.A. and Nelson, P.I. (1978). On conditional least squares estimation for stochastic processes. Ann. Statist. **6**, 629–642.

7. Lampard, D.G. (1968). A stochastic process whose successive intervals between events form a first order Markov Chain I . J. Appl. Probab. **5**, 648–668.

8. Lawrance, A.J. and Lewis, P.A.W. (1981). A new autoregressive time series model in exponential variables (NEAR(1)). Adv. Appl. Prob. **13**, 826–845.

9. Lewis, P.A.W., McKenzie, E. and Hugus, D.K. (1989). Gamma Processes. Stochastic Models, **5**, 1-30.

10. McKenzie, E. (1986). Autoregressive moving-average processes with negative binomial and geometric marginal distributions. Adv. Appl. Prob. **18**, 679–705.

11. Sim, C.H. (1990). First-order autoregressive models for gamma and exponential processes. J. Appl. Prob. **27**, 325–332.

12. Sim, C.H. (1986). Simulation of Weibull and gamma autoregressive stationary process. Commun. Statist-Simulation **15**, 1141–1146.

13. Sim, C.H. and Lee, P.A. (1989). Simulation of negative binomial processes. J. Statist. Comput. Simul. **34**, 29–42.

14. Thavaneswaran, A. and Abraham, B. (1988). Estimation for non-linear time series models using estimating equations. J. Time Series Analysis **9**, 99–108.

Received December 1991; Revised March 1992