HERCMA 2001
E. A. Lipitakis (Editor)
2001, L.E.A.
Printed in Hellas

On Some Discrete- Valued Time Series Models based on Mixtures and Thinning

Nikkie Yiokari, Evdokia Xekalaki and Dimitris Karlis

Abstract— **Time series have received considerable attention as a tool for the treatment of practical situations in several fields of statistics. However, most of the time series based statistical methodologies have been designed for data of a continuous nature (continuous-valued time series), while little attention has been given to time series models for data of a discrete nature (discrete-valued time series). This paper is concerned with discrete-valued time series. The two main classes of models that have been considered in the literature, are discussed. These are models based on mixtures and models based on thinning operators. Moreover, a Discrete Autoregressive model of first order with a generalized Waring marginal distribution is proposed, based on negative hypergeometric thinning.**

Keywords— **INAR model; Generalized Waring Distribution; Binomial thinning; negative hypergeometric thinning;**

I. Introduction

Time series models have been extensively used in a wide range of applications. The majority of such models have been developed to cope with continuous data, usually under a normality assumption. For non-normal data, the literature is rather sparse. However, in certain situations the nature of the data is discrete in the sense that they represent counts. In such cases, standard time series models based on the normality assumption are inappropriate.

Time series models for describing discrete data have been proposed in the literature. Such models are usually referred to as discrete-valued time series models. MacDonald and Zucchni [16] provide a survey on a variety of discrete-valued time series models. The models discussed include Markov chains, higher-order Markov chains, models based on mixtures and models based on the idea of thinning. Further, Markov regression models parameter-driven models, state-space models and Hidden Markov models along with their extensions were presented and described as models capable of dealing with discrete-valued time series.

This paper focuses on a particular class of discrete-valued time series models. More specifically models based on thinning operators are discussed. Such models are the discrete counterparts of the common autoregressive models for normal data suitably defined to deal with the discrete nature of the data. A brief review of discrete valued time series models is given in section II. The integer valued autoregressive model (INAR(1)) is discussed in section III. In section IV a new model belonging to this class is developed. This model has a Univariate Generalized Waring distribution (UGWD) as a marginal distribution and thus it serves as a potential alternative to existing models. Some concluding remarks are given in section V.

Department of Statistics, Athens University of Economics and Business. 76 Patision St., 10434. Athens. GREECE

II. Discrete-valued Time Series Models

A. Models based on Mixtures

An attempt to provide a class of models for discrete-valued time series was made by Jacobs and Lewis [11],[12],[13],[14]. Jacobs and Lewis presented a simple scheme for constructing stationary sequences of dependent random variables. The construction of such models is based on probabilistic mixtures of independent and identically distributed (i.i.d) discrete random variables. Moreover, these models are characterized by a specified marginal distribution and correlation structure. The structure of these models is analogous to the continuous-valued Gaussian autoregressive moving average (ARMA) process. Therefore, these models were introduced as discrete mixed autoregressive moving average (DARMA) models. Their definition mimics that of ARMA models for continuous time series, suitably defined to fit the discrete nature of the data. For more details the interested reader is referred to Jacobs and Lewis [11],[12],[13],[14], McKenzie [20], Al-Osh and Alzaid [2] among others.

B. Models based on Binomial Thinning

Another class of models, which has been developed, in order to deal with discrete-valued time series are based on the idea of the binomial thinning operator. They mirror the structure and the correlation of the well-known autoregressive moving-average (ARMA) processes used to model time-series with Gaussian marginals, which are quite attractive since they involve a few parameters and have a simple linear structure.

Such models can be generalized easily by defining other thinning operators, as described in the sequel. In this paper, we focus on the INAR process as it combines an easy interpretation with useful properties.

III. The Integer-valued Autoregressive process

McKenzie [17] and Al-Osh and Alzaid [1] defined the Integer-valued autoregressive ($INAR$) process as follows:

Definition 1: A sequence of random variables $\{X_t\}$ is an $INAR(1)$ process if it satisfies a difference equation of the form

$$X_t = a * X_{t-1} + R_t, \qquad \text{for } t = 0, \pm 1, \pm 2, \ldots \quad (1)$$

where R_t is a sequence of uncorrelated non-negative integer-valued random variables having mean μ and finite variance σ^2. The operator '$*$' is defined by

$$a * X = \sum_{t=1}^{X} Y_t,$$

where Y_t are Bernoulli random variables with $P(Y_t = 1) = a = 1 - P(Y_t = 0)$, $a \in [0, 1]$.

Thus, conditional on X, $a * X$ is a binomial random variable. It represents the number of successes in X independent trials when the probability of success in each trial is a.

The term R_t is referred to as the *innovation term* and must be independent of $a * X_{t-1}$.

From the above definition, one can see that the model mimics the normal autoregressive model and belongs to a more general family of autoregressive models discussed in Grunwald *et al.* [8]

The basic ingredient of the INAR model is that it assumes that the realization of the process at time t is composed of the survivals of the elements of the process, at time $t-1$, each with probability of survival α and the elements R_t which entered the system in the interval $[t-1, t]$ The innovation R_t follows some discrete distribution.

The mean and variance of the $INAR(1)$ process $\{X_t\}$ are:

$$E(X_t) = aE(X_{t-1}) + \mu$$

$$\begin{aligned} Var(X_t) &= a^{2t}Var(X_0) + (1-a) \\ & \sum_{j=1}^{t} a^{2j-1}E(X_{t-j}) + \sigma^2 \sum_{j=1}^{t} a^{2(j-1)}. \end{aligned}$$

In order for second-order stationarity to hold, the initial value of the process, X_0, must have:

$$E(X_0) = \tfrac{\mu}{1-a} \quad \text{and} \quad Var(X_0) = \tfrac{a\mu+\sigma^2}{1-a^2}.$$

For any non-negative integer k, the covariance $\gamma(k)$, at lag k for the process is

$$\gamma(k) = Cov(X_{t-k}, X_t) = a^k \gamma(0).$$

From the covariance function, it is easy to obtain the autocorrelation function $\rho(k)$ as follows:

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = a^k.$$

Thus, the autocorrelation function $\rho(k)$ decays exponentially with lag k.

Alzaid and Al-Osh [3] studied the regression behavior of the $INAR(1)$ process, assuming that $\mathrm{Var}(X_t) < \infty$. Thus, the regression of X_t on $X_{t-1} = x$ is given by,

$$E(X_t | X_{t-1} = x) = ax + \mu$$

and the conditional variance of X_t is given by,

$$Var(X_t | X_{t-1} = x) = a(1-a)x + \sigma^2.$$

Steutel and Van Harn [21] proposed analogues for the concept of self-decomposability for distributions on non-negative integers. Based on this definition of "discrete self-decomposable" distributions, Alzaid and Al-Osh [3] showed that every discrete self-decomposable distribution can arise as the limiting distribution for an $INAR(1)$ process. Such distributions are the Poisson, the negative binomial and the geometric distributions. Thus, McKenzie[17], [20] and Alzaid and Al-Osh [3] dealt with the *Poisson* $INAR(1)$ process, McKenzie [18] dealt with the *negative binomial* $INAR$ process whereas McKenzie [17], [18] and Alzaid and Al-Osh [3] constructed the *geometric* $INAR$ process.

The $INAR(1)$ model can be extended to the $INAR(p)$ model, which is similar in form to the Gaussian $AR(p)$ process and was defined by Alzaid and Al-Osh [4].

In practice, any discrete distribution can serve as the distribution of the innovation term. The problem is that the resulting marginal distribution may not be of a known form for all values of t. Thus, one may define INAR models either by specifying a discrete distribution for the innovation term or by determining this distribution so as to obtain a known form of distribution as a marginal distribution. Both methods are equivalent to defining the process via the transition probabilities between X_t and X_{t-1}. Consider, for example, the case of a Poisson INAR model. This model implies that, the marginal mean and the marginal variance coincide. This, however, is not realistic in many applications. So, in order to achieve overdispersion (variance> mean), one may assume an overdispersed distribution for the innovation term, reltive to the Poisson distribution. For example, Franke and Seligmann, [7] used a 2-finite Poisson mixture innovation distribution, defining a Switching INAR model.

A. Related Models

A.1 Models based on Quasi-Binomial Thinning

The $INAR$ model generally assumes that the probability of retaining an element is constant. Nevertheless, in many real data on counting processes it seams reasonable to assume that the probability of retaining an element is not constant but may depend on time and/or the number of elements already retained or it may be a random variable itself.

Alzaid and Al-Osh [5] considered the development of $ARMA$ models with Generalized Poisson marginals. The key assumption in the development of these models is that the probability of retaining an element is a linear function of the number of elements being retained. In particular, given $X_{t-1} = n$, the number of retained elements at time t has a quasi-binomial distribution with parameters (p, θ, n), denoted by $QB(p, \theta, n)$. More details about the Quasi-Binomial distribution can be found in Consul and Mittal [6].

A.2 Models based on Hypergeometric Thinning

The binomial distribution, in contrast to the Poisson, negative binomial and geometric distributions does not enjoy the property of self-decomposability and thus may not serve as a marginal distribution for the $INAR$ process. Hence, in order to construct models with an ARMA structure and a binomial marginal distribution, another approach must be considered. Al-Osh and Alzaid [2] introduced a hypergeometric thinning operator on which they based the construction of $ARMA$ processes with a binomial marginal. This hypergeometric thinning operator assumes that the probability of retaining an element present at time X_{t-1}depends on the elements that have already been retained and hence it is not constant as in the case of the binomial thinning operator.

A.3 Models based on Negative Hypergeometric Thinning

The main idea of the two alternative thinning operators discussed above, was the fact that the probability of retaining an existing element depends on X_{t-1}. Alternatively, one may consider that α is not constant but it is itself a random variable, having a density, say $g(.)$. For this reason, the resulting operator can be considered as a mixed binomial thinning operator. In this case, the conditional expectation is given by

$$E(X_t \mid X_{t-1}=x)=E(\alpha)x+\mu=$$

$$=[E(\alpha)]^t E(X_0)+\mu\sum_{j=0}^{t-1}[E(\alpha)]^j .$$

In a similar fashion, the conditional variance is given as

$$Var(X_t \mid X_{t-1}=x)=E\left(\alpha(1-\alpha)\right)x+\sigma^2$$

and thus the unconditional variance is given as

$$\begin{aligned} Var(X_t) &= Var\left[E(X_t \mid X_{t-1}=x)\right]+ \\ &+ E\left[Var(X_t \mid X_{t-1}=x)\right]= \\ &= [E(\alpha)]^2\,Var(X_{t-1})+ \\ &+ E\left(\alpha(1-\alpha)\right)E(X_{t-1})+\sigma^2. \end{aligned}$$

Using these formulae one can deduce the stationarity conditions as

$$E(X_0)=\frac{\mu}{1-E(a)} \qquad \text{and}$$

$$Var(X_0)=\frac{1}{1-[E(\alpha)]^2}\left[\frac{E(\alpha)-E(\alpha^2)}{1-E(\alpha)}\mu+\sigma^2\right].$$

For constant α, the above stationarity conditions reduce to those for the simple INAR model with binomial thinning. It can be also seen that the autocorrelation function is given as

$$\rho(k)=[E(\alpha)]^k .$$

An interesting result regarding the mixed binomial thinning operator is the fact that the index of dispersion of X_t, denoted by $ID(X_t)$, is given as

$$\begin{aligned} ID(X_t) &= \frac{Var(X_t)}{E(X_t)}= \\ &= \frac{E(\alpha)-E(\alpha^2)}{1-[E(\alpha)]^2}+\frac{ID(R_t)}{1+E(\alpha)}= \\ &= \frac{E(\alpha)+ID(R_t)}{1+E(\alpha)}+\frac{Var(\alpha)}{1-[E(\alpha)]^2} \end{aligned}$$

For constant α, we obtain that

$$ID(X_t)=\frac{\alpha+ID(R_t)}{1+\alpha}.$$

Usually, when treating count data, the index of dispersion is interpreted as an index of overdispersion relative to the Poisson distribution whose index of dispersion is equal to 1. Hence, by considering an overdispersed innovation distribution one can deduce an overdispersed marginal distribution for the series. Since $0\leq Var(\alpha)/(1-[E(\alpha)]^2)\leq 1$, one can see that the mixed binomial thinning operator adds overdispersion to the observed series, that, however, cannot be higher than 1.

The most common mixed binomial thinning operator is the negative hypergeometric operator which results if a beta distribution is assumed for α. In order to overcome the complexity of the innovation term of the negative binomial $INAR$ process, McKenzie [18] developed a more flexible negative binomial $AR(1)$ model, based on the negative hypergeometric thinning operator. The construction of this negative binomial $AR(1)$ model is based on the reproducibility property of the negative binomial distribution, asserting that the sum of two independent negative binomial variables with parameters k,α and m,α, denoted as $NB(,k,\alpha)$ and $NB(m,\alpha)$ respectively has the $NB(k+m,\alpha)$ distribution.

As mentioned earlier, letting $Y=a*X$, is equivalent to assuming that $Y|X\sim Binomial(X,a)$ where $a\in[0,1]$ is the probability of success. McKenzie [18] assumes that a is not a constant but follows a Beta(p,q) distribution. Then, the conditional probability distribution of $Y=a*X$ is defined by

$$P_{Y|X}(y|x)=\binom{x}{y}\frac{B\left(p+y,q+x-y\right)}{B(p,q)}, \qquad (2)$$

where $y=0,1,...x$. This distribution is known as the negative hypergeometric or beta-binomial distribution (see, e.g. [15]).

This approach generalizes the idea of the binomial thinning operator, with α as a constant. The resulting operator is refrred to as the *negative hypergeometric thinning operator* and is denoted by $S(X)$. Based on this operator, McKenzie defined the Negative Binomial $AR(1)$ process as follows:

Definition 2: If $X_{t-1}\sim NB(\beta,\lambda)$, $A_t\sim Beta(a,\beta-a)$, $R_t\sim NB(\beta-\alpha,\lambda)$ and X_{t-1},A_t,R_t are mutually independent, then,

$$X_t=S(X_{t-1})+R_t \qquad (3)$$

defines a stationary process $\{X_t\}$ with the $NB(\beta\lambda)$ distribution as marginal distribution.

Clearly, the binomial thinning operator can be generalized in several other ways by allowing the distribution of α to have various other forms.

IV. Discrete-valued Time Series with Univariate Generalized Waring Marginal Distribution

The $UGWD$ was introduced by Irwin [9],[10] in the context of accident theory. This can arise as a mixture on λ of the $NB(\beta,\lambda)$ distribution and as such it provides a more flexible model for the interpretation of data yielding chance mechanisms in a number of diverse fields ranging from accident theory, psychometry and linguistics through to econometrics and operational research. It would therefore be interesting to examine the implications of the assumption of a variable λ on the model defined by (3). As shown in section V, this leads to an $INAR$ model that allows for a correlation between the number of retained elements and the number of innovations. In the remaining of the precent section certain properties of the $UGWD$ are discussed that are of underlying importance to the developement of the new model. For a more detailed account the interested reader is referred to Xekalaki [22], [23],[24], [25].

A. The Univariate Generalized Waring Distribution and its Properties

The probability function of the UGWD is given by:

$$P(X=x)=\frac{\rho_{[k]}}{(a+\rho)_{[k]}}\frac{\alpha_{[k]}k_{[x]}}{(a+k+\rho)_{[x]}}\frac{1}{x!}\ ,\ x=0,1,... \quad (4)$$

with parameters $\alpha,k,\rho>0$ where $x_{[k]}=x(x+1)...(x+k-1)=\frac{\Gamma(x+k)}{\Gamma(x)}$. We denote this distribution as $UGWD(a,k,\rho)$.

The r-th factorial moment of the $UGWD$ is given by:

$$\mu_{[r]}=\frac{\alpha_{[r]}k_{[r]}}{(\rho-1)(\rho-2)...(\rho-r)}.$$

The moments about any constant, including central moments, can be determined by the above formula by the usual transformation formulae. In particular, the mean and variance of the generalized Waring distribution are:

$$E(X)=\frac{ak}{(\rho-1)},\rho>1$$

and

$$\sigma^2=\frac{ka(\rho+a-1)(\rho+k-1)}{(\rho-1)^2(\rho-2)},\ \rho>2.$$

The UGWD can arise via a variety of mechanisms (see, [22]). Its derivation as a Poisson mixture is quite important. It is well known that the negative binomial distribution can be obtained as a mixture of a Poisson distribution by allowing the parameter of the Poisson distribution to follow a gamma distribution. The UGWD results by considering a further mixing in the sense of allowing the parameter of the negative binomial distribution to be a random variable. The full scheme is the following:

$$X\mid\lambda\sim Poisson(\lambda)$$
$$\lambda\mid\nu\sim Gamma\left(k,\frac{1}{\nu}\right)$$
$$\nu\sim BetaTypeII(a,\rho)$$

This can be shown to lead to the $UGWD$ with probability function as given by (4).

Apart from its mixture nature, the UGWD can be given an interpretation in terms of conditionality models which provide a framework appropriate for our purpose [22]. Conditionality models are mixture models with discrete mixture distributions.

Let X and Y be non-negative integer-valued random variables such that the conditional distribution of $Y|(X=x)$ is the negative hypergeometric distribution with parameters x,m and N and probability function given by:

$$P(Y=y|X=x)=\frac{\binom{-m}{y}\binom{-N+m}{x-y}}{\binom{-N}{x}}, \quad (5)$$

where $m,N>0,\quad y=0,1,...x$.

Moreover, let the distribution of X be the $UGWD(\alpha,N;\rho)$. Then, Xekalaki [22] showed that the distribution of Y is the $UGWD(\alpha,m,\rho)$.

Hence, the $UGWD$ is reproducible with respect to the negative hypergeometric family of distributions. It is interesting to note that for certain limiting values of the parameters, the $UGWD$ tends to the negative binomial distribution. In this case, the reproducibility with respect to the negative hypergeometric family of distributions is preserved. The above model led to the following characterization theorem of the $UGWD$, derived by Xekalaki [22].

Theorem 1: Let X and Y be non-negative , integer-valued random variables such that the conditional distribution of $Y\ |(X=x)$ is the negative hypergeometric with parameters x,m and N defined by (5). Then the distribution of X is the $UGWD(a,N,\rho)$ if and only if the distribution of Y is the $UGWD(a,m,\rho)$.

Xekalaki [22] obtained various other results pertaining to the genesis of the $UGWD$, such as its derivation based on an urn model, a mixed confluent hypergeometric distribution, etc.

Discrete self-decomposability of the $UGWD$, in the sense of Steutel and Van Harn, was proven by Xekalaki [24]. This property is of central importance to the problem investigated in the present paper as it implies that the $UGWD$ can lend itself as a marginal distribution in an $INAR$ model. In addition, the UGWD is infinitely divisible on $\{0,1,2,...\}$.

B. The Bivariate Generalized Waring distribution

Xekalaki [25] introduced the bivariate Generalized Waring distribution whose probability function is given by:

$$P(X=x,Y=y)=$$
$$\frac{\rho_{(k+m)}}{(a+\rho)_{(k+m)}}\frac{\alpha_{(x+y)}k_{(x)}m_{(y)}}{(a+k+m+\rho)_{(x+y)}}\frac{1}{x!}\frac{1}{y!}$$

where,

$$a,k,m,\rho>0 \text{ and } x,y=0,1,2,...$$

This distribution is denoted as $BGWD(a;k,m,\rho)$.

Consider a random vector (X, Y) of non-negative and integer-valued components such that $(X, Y) \sim BGWD(a, k, m, \rho)$. Then, the following properties hold:
1. $X \sim UGWD(a, k, \rho)$ and $Y \sim UGWD(a, m, \rho)$
2. $X + Y \sim UGWD(a, k + m, \rho)$
3. $X|(Y = y) \sim UGWD(a + y, k, \rho + m)$
4. $Y|(X = x) \sim UGWD(a + x, m, \rho + k)$

The proofs of the above properties can be found in Xekalaki [25]. These properties will be useful for the construction of the new process.

V. A New INAR(1) Model with a UGWD as Marginal Distribution

As already indicated, being discrete self-decomposable, the $UGWD$ may serve as a marginal distribution of an $INAR$ process, defined by

$$X_t = \alpha * X_{t-1} + R_t, \qquad \text{for } t = 0, \pm 1, \pm 2, \ldots$$

where R_t is independent of X_{t-1}, $0 \leq a \leq 1$ and $X_t \sim UGWD(a, k; \rho)$.

Since X_t and X_{t-1} have the same distribution, the innovation term R_t must be a random variable, with alternating probability generating function given as :

$$G_{R_t}(1 - z) = G_{X_{t-1}}(1 - z)/G_{\alpha * X_{t-1}}(1 - z).$$

where $G_U(\cdot)$ denotes the probability generating function of a random variable U (Mc Kenzie, [18]).

In the sequel, an $INAR(1)$ model is constructed with an UGW marginal distribution based on the negative hypergeometric thinning operator. The approach followed here is analogous to that followed by McKenzie [18] for the construction of an INAR model with negative binomial marginals.

Consider a discrete autoregressive process of the form,

$$X_t = A_t * X_{t-1} + R_t, \tag{6}$$

where $X_{t-1} \sim UGWD(k, N, \rho)$, $A_t \sim Beta(m, N)$ and $R_t \sim UGWD(k, N - m, \rho)$.

Then, according to McKenzie, $A_t * X_{t-1}$ conditional on X_{t-1} has a Negative Hypergeometric(x, m, N) distribution.

Moreover, due to the properties of the $UGWD$ we have that:

$$A_t * X_{t-1} \sim UGWD(k, m, \rho).$$

The random variable $A_t * X_{t-1}$ corresponds to the number of survivors which are present at time $t - 1$. In this case, the survival probability is not constant but can be regarded as being a random variable that follows the beta distribution. The random variable R_t corresponds to the number of new entrants to the system between the times $t - 1$ and t.

In order for (6) to lead to a sequence of random variables, following the $UGWD(k, N, \rho)$, it is natural to regard the bivariate random vector $(A_t * X_{t-1}, R_t)$ as having a $BGWD(k, m, N - m, \rho)$.

According to the properties of the $BGWD$:

$$A_t * X_{t-1} + R_t \sim UGWD(k, N, \rho) \Leftrightarrow$$
$$X_t \sim UGWD(k, N, \rho).$$

Thus, the full model is:
$X_t = A_t * X_{t-1} + R_t$,
where,

$$\begin{aligned} A_t &\sim Beta(m, N) \\ R_t &\sim UGWD(k, N - m, \rho) \\ X_{t-1} &\sim UGWD(k, N, \rho) \\ (A_t * X_{t-1}, R_t) &\sim BGWD(k, m, N - m, \rho). \end{aligned}$$

It is important to note that the approach considered above is not conformal to the general framework of discrete-valued AR models, which assume independence of the random variables X_{t-1} and R_t. In particular, the number of individuals that enter the system at time t is not independent of the number of individuals that are in the system up to time t. This assumption can reflect situations that occur in real-life

The above model can be written using the negative hypergeometric thinning operator as

$$X_t = S(X_{t-1}) + R_t.$$

The stationary mean can be derived using the results of section III, since the correlation structure between $A_t * X_{t-1}$ and R_t does not affect the mean. This is not true for the variance and the autocorrelation function of the series.

The UGWD is known to have heavy tails (see, [10]) and thus it could be helpful for modeling data series with long tails.

VI. Concluding Remarks

In this paper, a new autoregressive model for integer valued time series is developed. The marginal distribution of the model is the UGWD. Its derivation allows for a correlation between the retained elements of the series and the innovations. The value of the new model lies in that it illustrates the general direction in which $INAR$ models can be extended so as to be applicable in situations where the requirement of independence between X_{t-1} and R_t imposed by he usual $INAR$ model conditions can be prohibitive of its implementation in practice. Thus, in a plethora of practical situations, the number R_t of new entrants at time t may not be reasonably considered to be independent of the number X_{t-1} of elements already existing in the system. Would a long queue, i.e. a large value of X_{t-1} in technical terms, not have a deterring effect on the new arrivals, thus influencing the value of R_t negatively? The opposite may well be true. The value of X_{t-1} may have a contagion effect on the magnitude of the value of R_t as, for example, in the case where the length of the queue is indicative of the quality of services offered. The results obtained in this paper have mainly focused on the modelling aspect of the integer-valued time series with a UGWD marginal proposed above. Results referring to the general type of such time series allowing for correlated $S(X_{t-1})$ and R_t as well as asymptotic properties and estimation procedures are currently under research.

REFERENCES

[1] Al-Osh, M.A. and Alzaid, A.A. (1987). First-Order Integer-valued Autoregressive (INAR(1)) process. *Journal of Time Series Analysis*, 8, 261-275.

[2] Al-Osh, M.A. and Alzaid, A.A. (1991). Binomial autoregressive moving average models. *Stochastic Models*, 7, 261-282.

[3] 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.

[4] Alzaid, A.A. and Al-Osh, M.A. (1990). An integer-valued pth-order autoregressive structure (INAR(p)) process. *Journal of Applied Probability*, 27, 314-324.

[5] Alzaid, A.A. and Al-Osh, M.A. (1993). Some autoregressive moving average processes with generalized Poisson marginal distributions. *Annals of the Institute of Statistical Mathematics*, 45, 223-232.

[6] Consul, P.C. and Mittal, S.P. (1975). A new urn model with predetermined strategy. *Biometrical Journal*, 17, 67-75.

[7] Franke, J. and Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modeling epileptic seizure counts. *In "Developments in Time Series Analysis", ed. T. Subba Rao, Chapman and Hall*, 310-330.

[8] Grunwald, G.K., Hyndman, R.J., Tedesco, L. and Tweedie, R.L.(2000). Non-Gaussian conditional linear AR(1) models. *Australian and New Zealand Journal of Statistics*, 42, 479-495.

[9] Irwin, J.O. (1968). The Generalized Waring distribution applied to accident theory. *Journal of the Royal Statistical Society*, A 131, 205-225.

[10] Irwin, J.O. (1975). The Generalized Waring distribution. Parts 1-3. *Journal of the Royal Statistical Society*, A 138, 18-31, 204-227, 374-384.

[11] Jacobs, P.A. and Lewis, P.A.W. (1978a). Discrete time series generated by mixtures 1: Correlational and runs properties. *Journal of the Royal Statistical Society*, B 40, 94-105.

[12] Jacobs, P.A. and Lewis, P.A.W. (1978b). Discrete time series generated by mixtures 2: Asymptotic properties. *Journal of the Royal Statistical Society*, B 40, 222-228.

[13] Jacobs, P.A. and Lewis, P.A.W. (1978c) Discrete time series generated by mixtures 3: Autoregressive processes (DAR(p)). *Technical report NPS55-78-022, Naval Postgraduate School, Monterey, California.*

[14] Jacobs, P.A. and Lewis, P.A.W. (1983). Stationary discrete autoregressive-moving average time series generated by mixtures. *Journal of Time Series Analysis*, 4, 19-36.

[15] Johnson, N., Kotz, S. and Kemp, A.W. (1992). *Univariate Discrete Distributions*. Wiley- New York.

[16] MacDonald, I.L., and Zucchini, Z. (1997). *Hidden Markov and other models for discrete-valued time series*. Chapman & Hall.

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

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

[19] McKenzie, E. (1987). Innovation distributions for gamma and negative binomial autoregressions. *Scandinavian Journal of Statistics*, 14, 79-85.

[20] McKenzie, E. (1988). Some ARMA models for dependent sequences of Poisson counts. *Advances in Applied Probability*, 20, 822-835.

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

[22] Xekalaki, E. (1981). Chance mechanisms for the univariate generalized Waring distribution and related characterizations. *Statistical Distributions in Scientific Work*, 4, 157-171.

[23] Xekalaki, E. (1983a). The univariate generalized Waring distribution in relation to accident theory: proneness, spells or contagion? *Biometrics*, 39, 887-895.

[24] Xekalaki, E. (1983). Infinite divisibility, completeness and regression properties of the univariate generalized Waring distribution. *Annals of the Institute of Statistical Mathematics*, 35, 279-289.

[25] Xekalaki, E. (1984). The bivariate generalized Waring distribution and its application to accident Theory. *Journal of Royal Statistical Society*, A 147, 488-498.