

AD-A039 467

NAVAL POSTGRADUATE SCHOOL MONTEREY CALIF
DISCRETE TIME SERIES GENERATED BY MIXTURES I: CORRELATIONAL AND--ETC(U)
JAN 77 P A LEWIS, P A JACOBS
NPS55-77-1

F/G 12/1

NSF-75-02026

NL

UNCLASSIFIED

| OF |
AD
A039 467



END

DATE
FILMED
6-77

AD A 039467

2

NPS55-77-1

NAVAL POSTGRADUATE SCHOOL

Monterey, California



DISCRETE TIME SERIES GENERATED BY MIXTURES I:
CORRELATIONAL AND RUNS PROPERTIES

by

P. A. W. Lewis

and

P. A. Jacobs

January 1977

AD NO. _____
DDC FILE COPY

Approved for public release: distribution unlimited.

Prepared for:

Chief of Naval Research
Arlington, VA 22217

National Science Foundation
Washington, DC 20360

DDC
FORM 1
MAY 15 1977
B
R

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral Isham Linder
Superintendent

Jack T. Borsting
Provost

The work reported herein was supported by funds provided directly from the Chief of Naval Research under Grant NR-42-284, and the National Science Foundation under Grant NSF-75-02026.

Reproduction of all or part of this report is authorized.

Prepared by:

Peter A. W. Lewis

Peter A. W. Lewis, Professor
Department of Operations Research

P. A. Jacobs

P. A. Jacobs

Reviewed by:

Michael G. Sovereign

Michael G. Sovereign, Chairman
Department of Operations Research

Released by:

Robert Fossum

Robert Fossum
Dean of Research

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

→ as some properties for the lengths of runs. The special case of a process in which the variables take on only two values is useful as a model for the counting process in a discrete-time point process. An application to the modelling of errors in the transmission of binary data is briefly discussed. ↖

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

DISCRETE TIME SERIES GENERATED BY MIXTURES I:

CORRELATIONAL AND RUNS PROPERTIES

by

P. A. Jacobs*
Stanford University, Stanford, Ca.

and

P. A. W. Lewis**
Naval Postgraduate School, Monterey, Ca.

NTIS	White Section	<input checked="" type="checkbox"/>
DDC	Ref Section	<input type="checkbox"/>
UNANNOUNCED		<input type="checkbox"/>
JUSTIFICATION.....		
BY.....		
DISTRIBUTION/AVAILABILITY CODES		
Dist.	AVAIL.	and/or SPECIAL
A		

ABSTRACT

A broad but parametrically simple model for a stationary sequence of dependent discrete random variables is given and several submodels are discussed. The structure of the model is specified by the marginal distribution of the random variables and several other parameters. The sequence of random variables is formed by a probabilistic linear combination of independent, identically distributed discrete random variables and is in general not Markovian. Second-order joint moments and spectra are obtained for the model, as well as some properties for the lengths of runs. The special case of a process in which the variables take on only two values is useful as a model for the counting process in a discrete-time point process. An application to the modelling of errors in the transmission of binary data is briefly discussed.

* Support under contract number NSF-75-02026 is gratefully acknowledged.

** Support under the Office of Naval Research contract NRO42-284 is gratefully acknowledged.

1. INTRODUCTION

In this paper we will introduce a simple method for obtaining a stationary sequence of dependent random variables having a specified marginal distribution and correlation structure (second-order joint moments). One advantage of the model is that the specification of these two aspects of the model is independent. Another advantage is that the sequence is obtained as a very simple transformation of a sequence of independent random variables. The model is analogous to models for dependent sequences of exponential random variables introduced in Jacobs and Lewis (1977) and Lawrance and Lewis (1977).

We will now define some quantities which will be used throughout the paper. Let $\{Y_n\}$ be a sequence of independent random variables taking values in a discrete space E each having the distribution π . Let $\{U_n\}$ and $\{V_n\}$ be independent sequences of independent random variables taking the values 0 and 1 with

$$(1.1) \quad P\{U_n = 1\} = \beta \quad \text{and} \quad P\{V_n = 1\} = \rho$$

for fixed $0 \leq \beta \leq 1$ and $0 \leq \rho < 1$. Let $\{S_n\}$ be a sequence of independent identically distributed random variables taking the values $0, 1, 2, \dots, N$ with distribution F , where N is a fixed integer.

The most general case which we will consider here is a sequence of random variables $\{X_n\}$ which is formed according to the probabilistic linear model

$$(1.2) \quad X_n = U_n Y_{n-S_n} + (1 - U_n) A_{n-(N+1)}$$

for $n = 1, 2, \dots$, where

$$(1.3) \quad A_n = V_n A_{n-1} + (1 - V_n) Y_n.$$

The model of (1.2) and (1.3) will be termed DARMA(1, N+1), (discrete mixed autoregressive-moving average process with autoregression of order 1 and moving average of order N+1).

If we start the process with $A_{-(N+1)}$ having the distribution π independent of $\{Y_n; n \geq -N\}$, $\{U_n\}$, $\{V_n\}$, and $\{S_n\}$, then the X_n 's, $n = 1, 2, \dots$, will be shown to form a stationary sequence of dependent discrete random variables having marginal distribution π . This stationary sequence is in general not Markovian, although it will be so if $\beta = 0$. Its correlation structure is determined by the parameters ρ and β and the distribution F . Note that π can be any distribution. Some cases of discrete distributions of particular interest are obtained by choosing π to be geometric or Poisson.

Certain special cases of the DARMA(1, N+1) process are of particular interest and their consideration will make the nomenclature clear.

(i) The DAR(1) process.

If $\beta = 0$, then

$$(1.4) \quad X_n = A_{n-(N+1)} = \begin{cases} A_{n-(N+1)-1} & \text{with probability } \rho, \\ Y_{n-(N+1)} & \text{with probability } (1-\rho). \end{cases}$$

2

$\{A_n\}$ is called the DAR(1) process (discrete autoregressive process of order 1).

(ii) The DMA(N) process

If $\beta = 1$, then $X_n = Y_{n-S_n} = Y_{n-k}$ with probability $F(k)$ for $k = 0, 1, \dots, N$ where F is the distribution of S_n . In this case, $\{X_n\}$ is called a DMA(N) process (discrete moving average process of order N). Note that if $\{X_n\}$ is a DMA(1) process, then

$$X_n = \begin{cases} Y_n & \text{with probability } F(0), \\ Y_{n-1} & \text{with probability } 1 - F(0). \end{cases}$$

(iii) The DARMA(1,1) process

Finally, if $N = 0$, then $\{X_n\}$ will be termed a DARMA(1,1) process (discrete mixed autoregressive-moving average process both of order 1) with parameters β and ρ ; that is,

$$X_n = \begin{cases} Y_n & \text{with probability } \beta, \\ A_{n-1} & \text{with probability } (1-\beta). \end{cases}$$

Note that the DMA(1) process is a special case of the DARMA(1,1) process when $\rho = 0$.

(iv) Independent process

If $\beta = 1, N = 0$ or $\beta = 0, \rho = 0$, then $\{X_n\}$ is a sequence of independent random variables with common distribution π .

The model of (1.2) and (1.3) is really the backward DARMA model. The forward model is defined in a similar fashion. However, the two, while similar, are not necessarily equivalent. This is because $\{X_n\}$

is not in general time reversible in the sense that $\{X_1, \dots, X_k\}$ will not in general have the same distribution as $\{X_{-k}, X_{-k+1}, \dots, X_{-1}\}$. The properties of one model can be derived by the same techniques as those of the other, so we will only consider the backward model.

Note that the DARMA process may be defined using any sequence of independent identically distributed random variables $\{Y_n\}$, not necessarily discrete. However, (1.2) and (1.3) show that $\{X_n\}$ is obtained as a mixture of the $\{Y_n\}$ sequence. As a result, even if the distribution of Y_n is continuous, a realization of the sequence $\{X_n\}$ will in general contain many runs of a single value. This seems to be the major drawback to using this scheme to obtain a sequence of dependent random variables with a specified continuous marginal distribution and correlation structure. Other schemes for obtaining sequences of dependent exponential and gamma random variables have been proposed which look more promising; cf. Lawrance and Lewis (1977), Gaver and Lewis (1977), and Jacobs and Lewis (1977).

One motivation behind the DARMA models was to provide a simple scheme for obtaining models with which to analyse stationary sequences of dependent discrete random variables with specified marginal distribution and correlation structure. In general, there is not much beyond a Markov chain model which is overparametrized for statistical purposes for modelling dependent sequences of random variables. In addition it is very often simple to show from data that the correlation structure of the sequence is not Markovian. The DARMA model can be used to model nonMarkovian sequences of discrete random variables; an observed sequence of this kind is discussed in the last section.

Another motivation for the development of this process was to provide models for point processes in which the data is given in terms of counts in fixed time intervals rather than the exact times of arrivals. Most models for point processes beyond the Poisson process are most easily described in terms of times of arrivals or times between arrivals and it is often hard to obtain results concerning the joint distribution of counts in different fixed time intervals. We feel that the DARMA models will be of use in such situations. There is also the possibility of modelling directly the binary counting process in a discrete time point process. This is discussed in Section 6.

2. SOME PRELIMINARY PROPERTIES OF THE DARMA(1,N+1) PROCESS

In this section we will give some properties for the DARMA(1,N+1) process. Unless otherwise indicated we will assume throughout the paper that $A_{-(N+1)}$ has a distribution π and is independent of $\{Y_n\}$, $\{U_n\}$, $\{V_n\}$, and $\{S_n\}$.

2.1. The marginal distribution of X_n .

We will first show that X_n as defined by (1.2) has distribution π for all n . To this end we note from the expression (1.3) that the random variable A_n can be expanded backwards to the initial value $A_{-(N+1)}$ to give $A_n = Y_{n-j}$ with probability $\rho^j(1-\rho)$ for $0 \leq j \leq N+n$ and $A_n = A_{-(N+1)}$ with probability ρ^{N+n+1} ; that is, A_n is a mixture of $Y_n, Y_{n-1}, \dots, Y_{-N}$, and $A_{-(N+1)}$. Hence, for i in the state space E

$$P\{A_n=i\} = \sum_{j=0}^{N+n} P\{Y_{n-j}=i\} \rho^j(1-\rho) + P\{A_{-(N+1)}=i\} \rho^{N+n+1} = \pi(i)$$

for $n = -N, -N+1, \dots$. Similarly,

$$P\{Y_{n-S_n}=i\} = \sum_{j=0}^N P\{Y_{n-j}=i, S_n=j\} = \pi(i) \sum_{j=0}^N P\{S_n=j\} = \pi(i)$$

for $i \in E$ and $n = 1, 2, \dots$. From (1.2) and (1.3) it now follows that

$$P\{X_n=i\} = \beta P\{Y_{n-S_n}=i\} + (1-\beta) P\{A_{n-(N+1)}=i\} = \pi(i)$$

for $i \in E$ and $n = 1, 2, \dots$. Hence, the marginal distribution of the X_n 's like those of the Y_n 's is π .

2.2. Correlational properties of $\{X_n\}$

Although the X_n 's have a stationary distribution π , the X_n 's are not independent, as are the Y_n 's. This can be seen by the following calculation of the covariance between X_n and X_{n+j} . After some simplification

$$\begin{aligned} (2.1) \quad E[X_{n+j}X_n] - E[X_{n+j}]E[X_n] &= \beta^2 \{E[Y_{n+j-S_{n+j}} Y_{n-S_n}] - E[Y_{n+j-S_{n+j}}] E[Y_{n-S_n}]\} \\ &\quad + \beta(1-\beta) \{E[Y_{n+j-S_{n+j}} A_{n-(N+1)}] - E[Y_{n+j-S_{n+j}}] E[A_{n-(N+1)}]\} \\ &\quad + \beta(1-\beta) \{E[A_{n+j-(N+1)} Y_{n-S_n}] - E[A_{n+j-(N+1)}] E[Y_{n-S_n}]\} \\ &\quad + (1-\beta)^2 \{E[A_{n+j-(N+1)} A_{n-(N+1)}] - E[A_{n+j-(N+1)}] E[A_{n-(N+1)}]\}. \end{aligned}$$

The covariance of Y_{n-S_n} and $Y_{n+1-S_{n+1}}$ is

$$\begin{aligned} E[Y_{n+1-S_{n+1}} Y_{n-S_n}] - E[Y_{n+1-S_{n+1}}] E[Y_{n-S_n}] \\ = \sum_{k=0}^N P\{S_{n+1}=k\} P\{S_n=k+1\} \text{Var } Y_{n-S_n}. \end{aligned}$$

In the case $\beta = 1$, (the DMA(N) model), only the first term in (2.1) is nonzero and we can get the correlation from the above result.

Putting $F(k) = P\{S_n=k\}$ we have for the correlation

$$\begin{aligned}
 (2.2) \quad \rho_{MA}(1) &= \text{corr}(Y_{n-S_n}, Y_{n+1-S_{n+1}}) \\
 &= \{E[Y_{n+1-S_{n+1}} Y_{n-S_n}] - E[Y_{n+1-S_{n+1}}] E[Y_{n-S_n}]\} / \text{Var } Y_{n-S_n} \\
 &= \sum_{k=0}^{N-1} F(k) F(k+1) .
 \end{aligned}$$

By similar reasoning, for $j \leq N$

$$(2.3) \quad \rho_{MA}(j) = \text{corr}(Y_{n-S_n}, Y_{n+j-S_{n+j}}) = \sum_{k=0}^{N-j} F(k) F(k+j)$$

and for $j > N$, $\rho_{MA}(j) = 0$. Note that these expressions do not depend on n and thus the DMA(N) process is second order covariance stationary.

We will now compute the covariance of A_n and A_{n-1} which appears in (2.1); this will incidentally give us the correlation structure of the DAR(1) process.

$$\begin{aligned}
 &E[A_n A_{n-1}] - E[A_n] E[A_{n-1}] \\
 &= \rho \{E[A_{n-1} A_{n-1}] - E[A_{n-1}] E[A_{n-1}]\} \\
 &\quad + (1-\rho) \{E[Y_n A_{n-1}] - E[Y_n] E[A_{n-1}]\} \\
 &= \rho \text{Var } A_{n-1} .
 \end{aligned}$$

since the second term is zero because Y_n and A_{n-1} are independent. This is because A_{n-1} is a function only of Y_{n-1}, Y_{n-2}, \dots . By an induction argument we obtain for the correlation of A_n and A_{n+j}

$$(2.4) \quad \rho_A(j) = \text{corr}(A_n, A_{n+j}) = \rho^j$$

for $j \geq 1$. Because of the assumption that $A_{-(N+1)}$ has distribution π , (2.4) does not depend on n and thus the autoregressive process is second-order covariance stationary.

To complete the result for the general DARMA(1, N+1) process, we compute the cross covariances between the sequences $\{Y_{n-S_n}\}$ and $\{A_n\}$. We obtain

$$(2.5) \quad E[Y_{n+j-S_{n+j}} A_{n-(N+1)}] - E[Y_{n+j-S_{n+j}}] E[A_{n-(N+1)}] = 0$$

for $j \geq 1$ since S_{n+j} only takes on the values $\{0, 1, \dots, N\}$. For $0 \leq j \leq N$

$$(2.6) \quad E[Y_{n-S_n} A_{n-N+j}] - E[Y_{n-S_n}] E[A_{n-N+j}] \\ = \sum_{k=N-j}^N E[Y_{n-k} A_{n-N+j}; S_n=k] - E[Y_{n-k}; S_n=k] E[A_{n-N+j}] \\ = [(1-\rho) F(N-j) + \rho(1-\rho) F(N-j+1) + \dots + \rho^j(1-\rho) F(N)] \text{Var}(Y_1)$$

For $j > N$

$$\begin{aligned}
 (2.7) \quad & E[Y_{n-S_n} A_{n-N+j}] - E[Y_{n-S_n}] E[A_{n-N+j}] \\
 &= \sum_{k=0}^N E[Y_{n-k} A_{n-N+j}; S_n=k] - E[Y_{n-k}; S_n=k] E[A_{n-N+j}] \\
 &= [\rho^{j-N}(1-\rho) F(0) + \rho^{j-N+1}(1-\rho) F(1) + \dots + \rho^j(1-\rho) F(N)] \text{Var}(Y_1).
 \end{aligned}$$

Putting everything together in (2.1) we obtain the correlation of X_n and X_{n+j}

$$\rho_M(j) = \{E[X_n X_{n+j}] - E[X_n] E[X_{n+j}]\} / \text{Var} X_1.$$

We have

$$(2.8) \quad \rho_M(1) = \beta^2 \sum_{k=0}^{N-1} F(k) F(k+1) + \beta(1-\beta) F(N) (1-\rho) + (1-\beta)^2 \rho.$$

For $1 \leq j \leq N$

$$\begin{aligned}
 (2.9) \quad & \rho_M(j) \\
 &= \beta^2 \sum_{k=0}^{N-j} F(k) F(k+j) + \beta(1-\beta)(1-\rho) \{F(N-j+1) + \rho F(N-j+2) + \dots + \rho^{j-1} F(N)\} \\
 & \quad + (1-\beta)^2 \rho^j.
 \end{aligned}$$

For $j \geq N+1$

$$(2.10) \quad \rho_M(j) = \rho^{j-N-1} \{ \beta(1-\beta)(1-\rho)[F(0) + \rho F(1) + \dots + \rho^N F(N)] + (1-\beta)^2 \rho^{N+1} \} .$$

Note that $0 \leq \rho_M(j) \leq 1$ and for $j > N$, $\rho_M(j)$ decreases geometrically if $\rho > 0$ and $\rho < 1$. Since $\rho_M(j)$ is independent of n , the DARMA(1,N+1) process is second order covariance stationary.

2.3. Invariance under transformations

From its definition we note that X_n is a mixture of the random variables $Y_n, Y_{n-1}, \dots, Y_{-N}$ and $A_{-(N+1)}$; i.e., it is a random selection of one and only one of these random variables. Thus, if we transform each of the random variables $Y_n, Y_{n-1}, \dots, Y_{-N}, A_{-(N+1)}$ by the same function, each X_n will be transformed in the same way and its distribution will be that of the transformed Y_n 's.

Similar remarks apply if we transform the X_n 's. Note that in applying a common transformation individually to the X_n 's we do not affect the selection procedure and therefore the correlation structure of the transformed process is the same as that of the untransformed process. This (marginal) transformation invariance is important for statistical analysis of the process.

3. THE AUTOREGRESSIVE PROCESS DAR(1)

In this section we will give some properties of the DAR(1) process $\{A_n\}$. As usual we will assume that $A_{-(N+1)}$ has distribution π . By the results of Section 2, $\{A_n\}$ is a stationary sequence of random variables with marginal distribution π and correlations

$$(3.1) \quad \rho_A(j) = \text{corr}(A_n, A_{n+j}) = \rho^j, \quad j \geq 1.$$

The spectrum of the process is thus

$$(3.2) \quad f(\omega) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{j=1}^{\infty} \rho_A(j) \cos(\omega j) \right\} = \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos \omega}$$

It follows from (1.3) that $\{A_n\}$ is a Markov chain; that is, $P\{A_{n+1} = i | A_1, \dots, A_n\} = P\{A_{n+1} = i | A_n\}$ for any i in the state space E . Further, it is not hard to show that the transition matrix P is given by

$$(3.3) \quad P\{A_{n+1} = i | A_n = k\} = P(k, i) = \begin{cases} (1-\rho)\pi(i), & \text{for } k \neq i, \\ \rho + (1-\rho)\pi(i) & \text{for } k = i. \end{cases}$$

Note that we have started from the opposite direction from that usually taken in Markov chain theory; we have specified the stationary distribution associated with the chain first and specified the (Markovian) dependency structure by a single parameter ρ . Moreover changing ρ

does not affect π . When $\rho = 0$ we have a stationary sequence of independent identically distributed random variables with distribution π .

The fact that $\{A_n\}$ is a Markov chain with a particularly simple transition function P makes many calculations quite easy. For example, in discrete time series, runs of given values of the random variables X_n are useful in statistical analyses. Properties of these runs are easy to obtain for the DAR(1) process. Thus fix a state $i \in E$ and let $T_i = \inf\{n \geq 1: A_n \neq i\} - 1$; T_i is the length of a run of i 's starting at time 1, where length can be $0, 1, \dots$. Then

$$P\{T_i \geq n\} = P\{A_1 = A_2 = \dots = A_n = i\} = \pi(i) P(i, i)^{n-1}$$

for $n \geq 1$ and $P\{T_i = 0\} = 1 - \pi(i)$. Thus

$$(3.4) \quad E[T_i] = \frac{\pi(i)}{1 - P(i, i)} = \frac{\pi(i)}{(1-\rho)[1 - \pi(i)]}.$$

If $\rho = 0$, then $A_n = Y_n$ for $n \geq 1$ and $E[T_i] = \pi(i)/[1 - \pi(i)]$ as expected since $\{A_n\}$ is a sequence of independent random variables in this case. Note that for $0 \leq \rho < 1$

$$E[T_i] \geq \frac{\pi(i)}{1 - \pi(i)};$$

that is, the expected length of a run of i 's for a DAR(1) process is always greater than or equal to the expected run length for a sequence of independent random variables. Moreover the inflation in the expected

length of runs is uniform for all states. This is a consequence of the fact that we are dealing with a one parameter Markov chain.

It is also not hard to calculate the generating function for T_i . We have for $0 \leq z \leq 1$

$$\begin{aligned} \Phi(z) &= \sum_{n=0}^{\infty} z^n P\{T_i=n\} = [1-\pi(i)] + \sum_{n=1}^{\infty} z^n [P(i,i)]^{n-1} [1-P(i,i)] \\ &= [1-\pi(i)] + \frac{z[1-P(i,i)] \pi(i)}{1 - zP(i,i)} \\ &= \frac{[1-\pi(i)][1-z\rho]}{1 - z\pi(i) - z\rho[1-\pi(i)]} \end{aligned}$$

Again, if $\rho = 0$, $\Phi(z)$ reduces to $\frac{[1-\pi(i)]}{1-z\pi(i)}$ the expression for the generating function of a length of run of i for a sequence of independent random variables with marginal distribution π .

4. THE DMA PROCESS

In this section we will consider the DMA(N) process $X_n = Y_{n-S_n}$. Note that, unlike the DAR(1) process, the DMA(N) process is not Markovian in general.

4.1. Correlation properties.

By results in section 2, $\{X_n\}$ is a stationary sequence of random variables with marginal distribution π and correlations

$$(4.1) \quad \rho_{MA}(j) = \text{corr}(X_n, X_{n+j}) = \frac{\sum_{k=0}^{N-j} F(k) F(k+j)}{\sum_{v=j}^N F(v) F(v-j)}$$

for $1 \leq j \leq N$. Also $\rho_{MA}(j) = 0$ for $j > N$ and $\rho_{MA}(0) = 1$. Note that when $N = 1$, the maximum value of the first order serial correlation, $\max_{F(0)} \rho_{MA}(1) = \max_{F(0)} \{F(0) [1-F(0)]\} = 1/4$. In fact one can show that for any $N \geq 1$ the maximum first order serial correlation $\rho_{MA}(1)$ that can be achieved is $1/4$. One can also maximize the correlation at any point, say j , by making $F(j) = F(0) = 1/2$. However, all the other correlations are zero.

For the spectrum of the DMA(N) process we have

$$(4.2) \quad f(\omega) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{j=1}^{\infty} \rho_{MA}(j) \cos(\omega_j) \right\};$$

then if we define $\rho_{MA}(-j) = \rho_{MA}(j)$, we have

$$(4.3) \quad \begin{aligned} f(\omega) &= \frac{1}{2\pi} \sum_{j=-\infty}^{+\infty} e^{i\omega j} \sum_{v=|j|}^N F(v) F(v-j) \\ &= \frac{1}{2\pi} \left[\left(\sum_{j=0}^N e^{-i\omega j} F(j) \right) \left(\sum_{j=0}^N e^{i\omega j} F(j) \right) + 1 - \sum_{j=0}^N F(j)^2 \right] \\ &= \frac{1}{2\pi} \left[\varphi_S(\omega) \varphi_S(-\omega) + 1 - \sum_{j=0}^N F(j)^2 \right] \\ &= \frac{1}{2\pi} \left[|\varphi_S(\omega)|^2 + 1 - \sum_{j=0}^N F(j)^2 \right] \end{aligned}$$

where φ_S is the characteristic function of the distribution F of the random variable S . Thus we can model a broad class of spectra $f(\omega)$. If $F(0) = 1$ we have an independent identically distributed sequence and a flat (constant) spectrum.

By way of example, it is worth noting that we have restricted S to have finite support. Then (4.3) is a polynomial in $\cos \omega$ just like any moving average process. The finite support was necessary to allow inclusion of the autoregressive tail (1.2). If one does not want to add this tail, then there is no reason to restrict the range of S . One then gets a much broader class of models for which (4.3) in particular holds, although the model is still a random index model. This extended model is not as broad as the DARMA(1,N+1) model in the sense that one cannot, as in linear models (see for example,

Anderson, 1970) represent the tail (1.3) as a random index model in which the random indices for each n are independent random variables.

To continue with the example, in the extended moving average model let S have a geometric distribution

$$P\{S=j\} = F(j) = p(1-p)^j \quad j = 0, 1, \dots$$

Then

$$\varphi_S(\omega) = \frac{p}{1 - (1-p)e^{i\omega}}$$

and

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi} \left[\frac{p}{[1 - (1-p)e^{i\omega}]} \frac{p}{[1 - (1-p)e^{-i\omega}]} + 1 - \sum_{j=0}^{\infty} p^2(1-p)^{2j} \right] \\ &= \frac{1}{2\pi} \left[\frac{p^2}{[1 + (1-p)^2 - 2(1-p) \cos \omega]} + \frac{2p(1-p)}{[1 - (1-p)^2]} \right] \end{aligned}$$

The initial point in the spectrum is related to the amount of long term dependence there is in the process. One could measure this by an index of dispersion (Cox and Lewis, 1966, p. 71)

$$J_k = \frac{\text{Var}(X_1 + \dots + X_k)}{k\{E[X]\}^2}$$

and

$$\lim_{k \rightarrow \infty} J_k = 2\pi \frac{\text{Var } X}{E[X]^2} f(0+).$$

For the moving average process, $f(0+) = \frac{1}{2\pi} [2 - \sum_{k=0}^{\infty} F(k)^2]$ which takes values between $1/2\pi$ and π . To compare the moving average process to the DAR(1) we note that from (3.2) for the DAR(1) process

$f(0+) = \frac{1}{2\pi} [(1-\rho^2)/(1-\rho)^2] = \frac{1}{2\pi} \frac{1+\rho}{(1-\rho)}$ which is always greater than $1/2\pi$ if $\rho > 0$ and increases with ρ to infinity. Note that $f(0+)$ for a sequence of independent random variables is $1/2\pi$. Thus both the moving average process and the DAR process give more long term dependence than a sequence of independent identically distributed random variables. The DAR(1) process allows more long term dependence than the moving average process.

4.2. Joint distributions and time reversibility

Unless otherwise indicated we will restrict our attention to the DMA(1) process in the remainder of this section; that is, if $\alpha = P\{S_n=0\}$, then

$$(4.8) \quad X_n = \begin{cases} Y_n & \text{with probability } \alpha \\ Y_{n-1} & \text{with probability } (1-\alpha). \end{cases}$$

In this case $\rho_{MA}(1) = \alpha(1-\alpha)$ and $\rho_{MA}(j) = 0$ for $j \geq 2$.

It is not hard to calculate the joint Laplace-Stieltjes transforms of the joint distributions of random variables in the DMA(1) sequence but it is tedious. For example, if $\gamma(s) = E[e^{-sX_1}]$, then from (4.8)

$$\begin{aligned}
 \psi_2(s_1, s_2) &= E[\exp\{-s_1 X_1 - s_2 X_2\}] \\
 &= (1-\alpha) \gamma(s_1) \gamma(s_2) + \alpha(1-\alpha) \gamma(s_1+s_2) + \alpha^2 \gamma(s_1) \gamma(s_2) \\
 &= \gamma(s_1) \gamma(s_2) [1 - \alpha(1-\alpha)] + \alpha(1-\alpha) \gamma(s_1 + s_2).
 \end{aligned}$$

Similarly, by conditioning arguments we obtain

$$\begin{aligned}
 \psi_3(s_1, s_2, s_3) &= E[\exp \{s_1 X_1 - s_2 X_2 - s_3 X_3\}] \\
 &= (1-\alpha) r(s_1) \psi_2(s_2, s_3) + \alpha(1-\alpha) r(s_1 + s_2) r(s_3) \\
 &\quad + \alpha^2(1-\alpha) r(s_1) r(s_2 + s_3) + \alpha^3 r(s_1) r(s_2) r(s_3) \\
 &= r(s_1) r(s_2) r(s_3) [1 - 2\alpha(1-\alpha)] \\
 &\quad + [r(s_1) r(s_2 + s_3) + r(s_1 + s_2) r(s_3)] \alpha / (1-\alpha)
 \end{aligned}$$

and

$$\begin{aligned}
 \psi_4(s_1, s_2, s_3, s_4) &= E[\exp\{-s_1 X_1 - s_2 X_2 - s_3 X_3 - s_4 X_4\}] \\
 &= (1-\alpha) r(s_1) \psi_3(s_1, s_2, s_3) + \alpha(1-\alpha) r(s_1 + s_2) \psi_2(s_3, s_4) \\
 &\quad + \alpha^2(1-\alpha^2) r(s_1) r(s_2 + s_3) r(s_4) + \alpha^3(1-\alpha) r(s_1) r(s_2) r(s_3 + s_4) \\
 &\quad + \alpha^4 r(s_1) r(s_2) r(s_3) r(s_4) \\
 &= r(s_1) r(s_2) r(s_3) r(s_4) [1 - 3\alpha(1-\alpha) + \alpha^2(1-\alpha) - \alpha^3(1-\alpha)] \\
 &\quad + r(s_1) r(s_2) r(s_3 + s_4) [\alpha(1-\alpha) - \alpha^2(1-\alpha) + \alpha^3(1-\alpha)] \\
 &\quad + r(s_1) r(s_2 + s_3) r(s_4) \alpha(1-\alpha) \\
 &\quad + r(s_1 + s_2) r(s_3) r(s_4) [\alpha(1-\alpha) - \alpha^2(1-\alpha) + \alpha^3(1-\alpha)] \\
 &\quad + r(s_1 + s_2) r(s_3 + s_4) \alpha^2(1-\alpha)^2
 \end{aligned}$$

One interest in the joint distributions of the random variables is to look at the time reversibility of the process. One reason for concern with time-reversibility is the following. The EMA(1) process (exponential moving average of order 1) of Lawrance and Lewis (1977)

is not time reversible even though this fact cannot be determined from second-order properties of the process. Consequently one cannot distinguish between α and $(1-\alpha)$ in the spectrum of the EMA(1) process. However, by using higher order moments it is possible to distinguish between α and $(1-\alpha)$. For the DMA(1) process the fact that $\rho_{MA}^{(1)} = \alpha(1-\alpha)$ means we cannot use it to distinguish between α and $(1-\alpha)$. The time reversibility for the DMA(1) process would mean that we might not be able to distinguish between α and $(1-\alpha)$ even by using higher order moments.

Since $\psi_n(s_1, s_2, \dots, s_n) = \psi_n(s_n, s_{n-1}, \dots, s_1)$ for $n = 2, 3, 4$, it seems likely that the DMA(1) process is time reversible. In order to show time reversibility we need to show that $\psi_n(s_1, s_2, \dots, s_n) = \psi_n(s_n, s_{n-1}, \dots, s_1)$ for any nonnegative s_1, \dots, s_n . For simplicity consider the terms $a = \gamma(s_1) \gamma(s_2 + s_3) \gamma(s_4) \dots \gamma(s_n)$ and $b = \gamma(s_1) \gamma(s_2) \dots \gamma(s_{n-3}) \gamma(s_{n-2} + s_{n-1}) \gamma(s_n)$ of $\psi_n(s_1, s_2, \dots, s_n)$. In the expression for $\psi_n(s_1, s_2, \dots, s_n)$ the term a has a coefficient of

$$\begin{aligned} & \sum_{j=3}^n P\{S_1=0 \text{ or } 1, S_2=0, S_3=+1, S_4=1, \dots, S_j=1, S_{j+1}=0, \dots, S_n=0\} \\ &= \sum_{j=3}^n P\{S_1=0, S_2=0, S_3=+1, S_4=+1, \dots, S_j=+1, S_{j+1}=0, \dots, S_n=0\} \\ &+ \sum_{j=3}^n P\{S_1=+1, S_2=0, S_3=+1, S_4=+1, \dots, S_j=+1, S_{j+1}=0, \dots, S_n=0\} \end{aligned}$$

since the event associated with the term a is that X_2 and X_3 pick the same Y_j and that all the other X_i 's pick distinct Y_j 's from each other. Similarly, the term b has a coefficient of

$$\sum_{j=1}^{n-2} P\{S_n=0 \text{ or } 1, S_{n-1}=1, S_{n-2}=0, S_{n-3}=0, \dots, S_j=0, S_{j-1}=1, \dots, S_1=1\}$$

Since the S_n 's are independent and only take the values 0 and 1 and the Y_n 's are independent, the coefficients of the terms a and b of $\psi_n(s_1, s_2, \dots, s_n)$ are equal.

Similar arguments can be used to show that the coefficients of the terms

$$\prod_{i=1}^{j_1-1} \gamma(s_i) \gamma(s_{j_1+s_{j_1+1}}) \prod_{i=2}^{j_2-j_1-1} \gamma(s_{j_1+i}) \gamma(s_{j_2+s_{j_2+1}}) \\ \dots \prod_{i=2}^{j_k-j_{k-1}-1} \gamma(s_{j_{k-1}+i}) \gamma(s_{j_k+s_{j_k+1}}) \prod_{i=j_k+2}^n \gamma(s_i)$$

and

$$\prod_{i=0}^{j_1-2} \gamma(s_{n-i}) \gamma(s_{n-j_1+1+s_{n-j_1}}) \prod_{i=1}^{j_2-j_1-2} (s_{n-j_1-i}) \gamma(s_{n-j_2+1+s_{n-j_2}}) \\ \dots \prod_{i=1}^{j_k-j_{k-1}-2} \gamma(s_{n-j_{k-1}-i}) \gamma(s_{n-j_k+1+s_{n-j_k}}) \prod_{i=1}^{n-j_k-1} \gamma(s_{n-j_k-i})$$

of $\psi_n(s_1, \dots, s_n)$ are equal for any sequences of integers

$$1 \leq j_1 < j_1+1 < j_2 < j_2+1 < j_3 < \dots < j_k < j_k+1 \leq n$$

and $k \geq 1$. Thus $\psi_n(s_1, \dots, s_n) = \psi_n(s_n, \dots, s_1)$ for all n.

Hence the DMA(1) process is time reversible.

4.3. Run lengths for the DMA(1) process.

We will now consider length of runs for a DMA(1) process.

Fix a state i in E and let $T_i = \inf\{n \geq 1: X_n \neq i\} - 1$ the length of a run of i initiated at time 1 where length can be $0, 1, \dots$.

We will first compute $E[T_i]$. Let $a_0 = 1$ and $a_n = P\{X_1 = X_2 = \dots = X_n = i\}$ for $n \geq 1$. Then

$$a_1 = P\{X_1 = i\}$$

$$a_2 = P\{X_1 = X_2 = i\} = (1-\alpha) \pi(i) a_1 + \alpha(1-\alpha) \pi(i) a_0 + \alpha^2 \pi(i)^2$$

and by induction

$$\begin{aligned} a_{n+1} &= P\{X_1 = i, \dots, X_{n+1} = i\} \\ &= (1-\alpha) \pi(i) a_n + \alpha(1-\alpha) \pi(i) a_{n-1} + \alpha^2(1-\alpha) \pi(i)^2 a_{n-2} + \dots \\ &\quad + \alpha^n(1-\alpha) \pi(i)^n a_0 + \alpha^{n+1} \pi(i)^{n+1}. \end{aligned}$$

Thus

$$\begin{aligned} E[T_i] &= \sum_{n=1}^{\infty} a_n \\ &= \pi(i) + (1-\alpha) \pi(i) \sum_{n=1}^{\infty} a_n + \alpha(1-\alpha) \pi(i) \frac{1}{1-\alpha\pi(i)} \sum_{n=1}^{\infty} a_n \\ &\quad + \alpha^2 \pi(i)^2 \frac{1}{1-\alpha\pi(i)} + \alpha(1-\alpha) \pi(i) \frac{1}{1-\alpha\pi(i)} a_0 \\ &= \pi(i) + \frac{\alpha^2 \pi(i)^2 + \alpha\pi(i) - \alpha^2\pi(i)}{1 - \alpha\pi(i)} + \left[(1-\alpha) \pi(i) + \frac{\alpha(1-\alpha)\pi(i)}{1-\alpha\pi(i)} \right] E[T_i] \end{aligned}$$

Solving the last equation for $E[T_i]$ we obtain

$$(4.9) \quad E[T_i] = \frac{p(i)}{1-p(i)} = \frac{\pi(i)[1 + \alpha(1-\alpha)(1 - \pi(i))]}{[1 - \pi(i)][1 - \alpha(1-\alpha)\pi(i)]}$$

where

$$(4.10) \quad \begin{aligned} p(i) &= \pi(i)[1 + \alpha(1-\alpha)] - \pi(i)^2 \alpha(1-\alpha) \\ &= \pi(i) + \alpha(1-\alpha) \pi(i)[1 - \pi(i)] . \end{aligned}$$

If α is either 0 or 1, then $\{X_n\}$ is a sequence of independent random variables and $E[T_i] = \pi(i)/(1-\pi(i))$ as expected. Note that $E[T_i] \geq \pi(i)/[1-\pi(i)]$ for $0 \leq \alpha \leq 1$; that is, the expected length of a run of i for a DMA(1) process is greater than the expected length of a run for a sequence of independent random variables. For a given distribution π , the maximum value for $E[T_i]$ occurs when $\alpha = 1/2$. In this case

$$E[T_i] = \frac{\pi(i)}{1 - \pi(i)} \left[1 + \frac{1}{4 - \pi(i)} \right] .$$

We now turn our attention to the generating function of T_i .

Fix $j \neq i$ in the state space E and let

$$b_n = P\{X_1 = \dots = X_n = i, X_{n+1} = j\}, \quad n \geq 1$$

and

$$b_0 = P\{X_1 = j\} .$$

Using an induction argument we obtain for $n \geq 1$

$$b_n = (1-\alpha) \pi(i) b_{n-1} + \alpha \pi(i) (1-\alpha) b_{n-2} + \alpha^2 \pi(i)^2 (1-\alpha) b_{n-3} \\ + \dots + \alpha^{n-1} \pi(i)^{n-1} (1-\alpha) b_0 + \alpha^n \pi(i)^n \alpha \pi(j) .$$

Thus

$$\sum_{n=0}^{\infty} z^n b_n = b_0 + \sum_{n=0}^{\infty} z^n b_n \{ z(1-\alpha) \pi(i) + (1-\alpha) z \sum_{n=1}^{\infty} \alpha^n \pi(i)^n z^n \} \\ + \alpha \pi(j) \sum_{n=1}^{\infty} \alpha^n \pi(i)^n z^n \\ = b_0 + \sum_{n=0}^{\infty} z^n b_n \left\{ z(1-\alpha) \pi(i) + \frac{z^2 \alpha (1-\alpha) \pi(i)}{1 - \alpha \pi(i) z} \right\} + \frac{\alpha^2 z \pi(i) \pi(j)}{1 - \alpha z \pi(i)} .$$

After some simplification we obtain

$$(4.11) \quad \Phi(z) = \sum_{n=0}^{\infty} z^n P\{T_1=n\} = \frac{[1-\pi(i)][1 - z\pi(i) \alpha(1-\alpha)]}{1 - z\pi(i) - z^2 \pi(i)[1-\pi(i)] \alpha(1-\alpha)} .$$

Note that for $\alpha = 0$ or 1 , $\Phi(z) = [1-\pi(i)]/[1-z\pi(i)]$ as expected.

Higher order moments of the run lengths can be obtained from (4.11).

5. THE BINARY DARMA(1,1) PROCESS

In this section we will consider a DARMA(1,1) process in which X_n takes only the values 0 and 1; that is,

$$(5.1) \quad X_n = \begin{cases} Y_n & \text{with probability } \beta, \\ A_{n-1} & \text{with probability } (1-\beta) \end{cases}$$

and

$$A_n = \begin{cases} A_{n-1} & \text{with probability } \rho, \\ Y_n & \text{with probability } (1-\rho), \end{cases}$$

where $\{Y_n\}$ is a sequence of independent random variables taking the values 0 and 1 with common distribution π . Note that the DARMA(1,1) process is not Markovian in general.

Time series of binary random variables are of particular importance for modelling the differential counting process in discrete time point processes. Klotz (1973) and Kanter (1975) have given a model which is different from the binary DARMA(1,1) process.

Setting $N = 0$ in (2.8)-(2.10) we obtain the correlation of X_n and X_{n+j} , $j \geq 1$,

$$\rho_M(j) = \text{corr}(X_n, X_{n+j}) = \rho^{j-1} [\beta(1-\beta)(1-\rho) + (1-\beta)^2 \rho].$$

The spectrum of the process is thus

$$(5.2) \quad f(\omega) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{j=1}^{\infty} \rho_M(j) \cos(\omega j) \right\}$$

$$= \frac{1}{2\pi} \left[\frac{2c(\beta, \rho) [\cos \omega - \rho] - 2\rho \cos \omega + \rho^2 + 1}{1 - 2\rho \cos \omega + \rho^2} \right]$$

where

$$c(\beta, \rho) = \beta(1-\beta)(1-\rho) + (1-\beta)^2 \rho .$$

We will now consider some properties of lengths of runs.

For fixed $i \in \{0, 1\}$, let $T_i = \inf\{n \geq 1; X_n \neq i\} - 1$ as before.

We will calculate $E[T_i]$ and the generating function for T_i . To

begin, note that although $\{X_n\}$ is not a Markov chain, $\{(A_n, X_n); n=1, 2, \dots\}$

is a Markov chain. For i, ℓ, j, k in $\{0, 1\}$

$$P\{A_{n+1}=j, X_{n+1} = k | A_n = i, X_n = \ell\} = Q_k(i, j)$$

independent of ℓ . Letting Q_k denote the matrix whose (i, j) entry

is $Q_k(i, j)$ we have

$$Q_0 = \begin{bmatrix} \rho(1-\beta) + [1-\rho(1-\beta)] \pi(0) & (1-\beta)(1-\rho)\pi(1) \\ \beta(1-\rho) \pi(0) & \beta\rho \pi(0) \end{bmatrix}$$

and

$$Q_1 = \begin{bmatrix} \beta\rho \pi(1) & \beta(1-\rho) \pi(1) \\ (1-\rho)(1-\beta) \pi(0) & (1-\beta)\rho + [1-(1-\beta)\rho] \pi(1) \end{bmatrix} .$$

Note that $P\{T_0 \geq n | A_0 = i\} = Q_0^n(i,0) + Q_0^n(i,1) \equiv Q_0^n(i,E)$
for $i = 0,1$. Hence

$$E\{T_0 | A_0 = i\} = \sum_{n=1}^{\infty} Q_0^n(i,E) = R_0(i,E) - 1$$

where $R_0(i,j) = \sum_{n=0}^{\infty} Q_0^n(i,j)$ with $I = Q_0^0$ being the identity matrix
and $R_0(i,E) = R_0(i,0) + R_0(i,1)$. It is not hard to show that

$$R_0 = (I - Q_0)^{-1}$$

where

$$(I - Q_0)^{-1} = \frac{1}{\Delta} \begin{bmatrix} 1 - \beta\rho\pi(0) & (1-\beta)(1-\rho)\pi(1) \\ \beta(1-\rho)\pi(0) & 1 - \rho(1-\beta) - [1 - \rho(1-\beta)]\pi(0) \end{bmatrix}$$

with

$$\Delta = \det(I - Q_0) = [1 - \pi(0)][1 - \rho(1-\beta)][1 - \beta\pi(0)] - \beta\pi(0)[1 - \beta(1-\rho)]$$

Thus

(5.3)

$$\begin{aligned} E\{T_0\} &= \pi(0) R_0(0,E) + \pi(1) R_0(1,E) - 1 \\ &= \frac{\pi(0)\{1 - \beta\rho + \beta(1-\rho-\beta+2\beta\rho)[1 - \pi(0)]\}}{[1 - \pi(0)][1 - \rho(1-\beta)][1 - \beta\pi(0)] - \beta\pi(0)[1 - \beta(1-\rho)]} \\ &= \frac{\pi(0)\{1 + \text{corr}(X_n, X_{n+1}) + \beta\rho - \rho\} + \pi(0)^2 \{-\text{corr}(X_n, X_{n+1}) - 2\beta\rho + \rho\}}{1 - \rho(1-\beta) + \pi(0)\{-1 + 2\rho - 3\beta\rho - \text{corr}(X_n, X_{n+1})\} + \pi(0)^2 \{\text{corr}(X_n, X_{n+1}) + 2\beta\rho - \rho\}} \\ &= \frac{\pi(0)\{\text{corr}(X_n, X_{n+1}) + 1 - \rho + \beta\rho - \pi(0)(\text{corr}(X_n, X_{n+1}) + 2\beta\rho - \rho)\}}{[1 - \pi(0)][1 - \rho + \beta\rho - \pi(0)(\text{corr}(X_n, X_{n+1}) + 2\beta\rho - \rho)]} \end{aligned}$$

after some simplification.

Similarly one can show that

(5.4)

$$\begin{aligned}
 E[T_1] &= \frac{\pi(1)\{1-\beta\rho + \beta(1-\rho-\beta+2\beta\rho)[1-\pi(1)]\}}{[1-\pi(1)]\{1-\rho(1-\beta)[1-\beta\pi(1)] - \beta\pi(1)[1-\beta(1-\rho)]\}} \\
 &= \frac{\pi(1)\{1 + \text{corr}(X_n, X_{n+1}) + \beta\rho - \rho\} + \pi(1)^2\{-\text{corr}(X_n, X_{n+1}) - 2\beta\rho + \rho\}}{1-\rho(1-\beta) + \pi(1)\{-1 + 2\rho - 3\beta\rho - \text{corr}(X_n, X_{n+1})\} + \pi(1)^2\{\text{corr}(X_n, X_{n+1}) + 2\beta\rho - \rho\}} \\
 &= \frac{\pi(1)\{\text{corr}(X_n, X_{n+1}) + 1 - \rho + \beta\rho - \pi(1)(\text{corr}(X_n, X_{n+1}) + 2\beta\rho - \rho)\}}{[1-\pi(1)]\{1-\rho + \beta\rho - \pi(1)(\text{corr}(X_n, X_{n+1}) + 2\beta\rho - \rho)\}}
 \end{aligned}$$

Note that the expected length of a run of i for the binary DARMA(1,1) process is always greater than or equal to the expected length of a run of i for the independent case.

We now turn our attention to the computation of

$$\Phi_i(z) = \sum_{n=0}^{\infty} z^n P\{T_i=n\} \quad \text{for } i = 0, 1 \quad \text{and } 0 \leq z \leq 1. \quad \text{To begin, note that}$$

$$P\{T_0=n|A_0=i\} = \sum_j Q_0^n(i, j) [1 - Q_0(j, E)].$$

Thus

$$\sum_{n=0}^{\infty} z^n P\{T_0=n|A_0=i\} = \sum_j \sum_{n=0}^{\infty} z^n Q_0^n(i, j) [1 - Q_0(j, E)].$$

It is not hard to show that $\sum_{n=0}^{\infty} z^n Q_0^n = (I - zQ_0)^{-1}$, where

$$(I - zQ_0)^{-1} = \frac{1}{\Delta(z)} \begin{bmatrix} 1 - z\beta\rho\pi(0) & z(1-\beta)(1-\rho)\pi(1) \\ z\beta(1-\rho)\pi(0) & 1 - z[\rho(1-\beta) + [1-\rho(1-\beta)]\pi(0)] \end{bmatrix}$$

with

$$\Delta(z) = 1 - z\rho(1-\beta) + z\pi(0) \{-\beta-1+\rho(1-\beta)\} \\ + z^2\pi(0)[1-\pi(0)]\{-\beta(1-\beta) + 2\rho\beta(1-\beta)\} + z^2\pi(0)^2\beta\rho .$$

After some manipulation we obtain

$$\Phi_0(z) \\ = \frac{1-z\rho(1-\beta)+\pi(0)[1-\pi(0)]\{z\beta^2-z\beta^2\rho-z\beta-z\beta^2\rho\} + \pi(0)[z\rho-1] - z\beta\rho\pi(0)^2}{1-z\rho(1-\beta)+z\pi(0)\{-\beta\rho-1+\rho(1-\beta)\}+z^2\pi(0)[1-\pi(0)]\{-\beta(1-\beta)+2\rho\beta(1-\beta)\}+z^2\pi(0)^2\beta\rho}$$

In a similar manner one can show that

$$\Phi_1(z) \\ = \frac{1-z\rho(1-\beta) + \pi(1)[1-\pi(1)]\{z\beta^2-z\beta^2\rho-z\beta-z\beta^2\rho\} + \pi(1)[z\rho-1] + z\beta\rho\pi(1)^2}{1-z\rho(1-\beta)+z\pi(1)\{-\beta\rho-1+\rho(1-\beta)\}+z^2\pi(1)[1-\pi(1)]\{-\beta(1-\beta)+2\rho\beta(1-\beta)\}+z^2\pi(1)^2\beta\rho}$$

Higher moments for the runs can be obtained from the generating functions. These are important in determining to what degree the binary DARMA(1,1) process differs from Markov models, e.g. the DAR(1) model, the differentiation considered here being to what extent the distribution of run lengths departs from a geometric distribution.

6. TELEPHONE ERROR DATA

We discuss here very briefly a case in which the binary DARMA(1,1) model may be of use; in particular, we do this to illustrate some of the formulas given in the previous section. Another recent example of the need for discrete time-series models is Gaver, Lavenberg and Price (1976).

Cox and Lewis (1966, p. 173) discussed data consisting of errors which occurred during the transmission of binary data over a telephone line. Let $X_n = 1$ indicate that the n th transmitted bit is in error, while $X_n = 0$ indicates that no error occurred. Models which postulate that bit errors occur independently or according to a Markov chain such as the binary DAR(1) process predict that the runs of ones and runs of zero will both have geometric distributions. But the runs of zeros are just the intervals (or number of bits) between errors without the times consisting of 1 bit between errors, and these were shown by Berger and Mandelbrot (1963) and Lewis and Cox (1966) to be nongeometric. In fact they are highly skewed, long-tailed distributions which led Berger and Mandelbrot (1963) to postulate a model in which intervals between errors were assumed to be independent with Pareto-type distributions.

The problem with the Berger-Mandelbrot model was that the intervals between errors were found to be dependent (Lewis and Cox, 1966). Moreover there are a disproportionate number of 1-bit between error intervals; 128 out of 673 intervals, while the longest interval between errors is 85,993 bits. This suggests that modelling the binary X_n process may be a better approach than modelling the intervals, although the modelling process must be nonMarkovian.

The binary DARMA(1,1) process is a candidate process for modelling this process; in particular one would like to know whether the runs of zeros for β between zero (Markovian) and one (independent) produce highly-skewed run-length distributions. The question is too broad to be considered here, involving also estimation of suitable values of β and ρ , and will be considered elsewhere. Here we will examine only the effect of β on $E(T_0)$ and the plausibility of the model.

For the error data, 672 bits out of 1,106,148 transmitted were in error, so that we can estimate $\pi(1)$ as

$$\hat{\pi}(1) = 1 - \hat{\pi}(0) = \frac{672}{1,106,148} = 0.0006075 .$$

Thus from (3.4) with $\rho = 0$ we compute that the expected lengths of runs of zeros and ones, given that they occur, are respectively $1/\hat{\pi}(1) = 1,645.09$ and $1/\hat{\pi}(0) = 1.000608$; the observed values from the data are 1,911.27 and 1.235, both much longer than predicted under independence assumption ($\rho = 0$).

In Table 1 we give values of $E(T_0)$ computed from the formula (5.3).

Note that for small ρ , e.g. $\rho = 0.1$, the value of $E(T_0)$ first increases with increasing β , and then decreases. This is characteristic behavior for the process when it is almost a moving average. For large ρ , we find $E(T_0)$ decreasing with β . In particular β has a large effect on $E(T_0)$; it remains to be seen how β effects the whole distribution of runs.

TABLE 1. Expected length of runs of zeros in a binary DARMA(1,1) process for various values of the parameters β and ρ with $\pi(1)$ fixed at the value 0.0006075 estimated for the telephone error data. The column $\beta = 1.0$ gives the independence case.

β	$E(T_0)$									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.1	1966.9	2088.2	2179.8	2230.5	2232.4	2184.0	2089.9	1960.4	1808.0	1645.1
0.2	2161.0	2243.2	2299.2	2307.0	2277.6	2206.6	2098.9	1962.8	1808.2	1645.1
0.3	2403.1	2431.3	2430.1	2396.6	2330.3	2233.2	2109.6	1965.7	1808.5	1645.1
0.4	2713.4	2664.4	2594.2	2503.1	2392.6	2264.7	2122.5	1969.3	1808.9	1645.1
0.5	3125.5	2960.9	2796.4	2631.8	2467.3	2302.8	2138.4	1973.9	1809.5	1645.1
0.6	3699.3	3350.8	3051.6	2790.4	2558.7	2349.8	2158.4	1979.9	1810.2	1645.1
0.7	4553.1	3886.4	3384.0	2990.6	2672.8	2409.0	2184.5	1988.2	1811.3	1645.1
0.8	5959.2	4668.0	3834.6	3251.4	2819.6	2486.2	2219.9	2000.3	1813.2	1645.1
0.9	8703.0	5915.8	4480.3	3604.9	3015.2	2590.8	2270.5	2019.7	1816.8	1645.1
1.00	16449.5	8224.3	5482.5	9111.6	3289.1	2740.8	2349.1	2055.3	1826.8	1645.1

The estimated first and second serial correlation coefficients for the data, $\hat{\rho}(1)$ and $\hat{\rho}(2)$ are 0.190 and 0.121 respectively. This is consistent with the model's restriction to positive serial correlation coefficients. From the expressions for $\rho_M(1)$ and $\rho_M(2)$ given in Section 5 we see that $\hat{\rho}(2)/\hat{\rho}(1) = 0.64$ is a rough estimate for ρ ; thus ρ is relatively large for this data. With the proviso that ρ is relatively large it might be possible to find unique values of ρ and β which, with the estimated $\pi(1)$, would make (5.3) and (5.4) equal to the estimated $E(T_0)$ and $E(T_1)$. (This is not possible in general since the expressions (5.3) and (5.4) are not single-valued functions of β for small fixed ρ .) An alternative is to use the estimate of ρ and $E(T_0)$, with $\hat{\pi}(0)$ and $\hat{\rho}(1)$, in (5.3) and solve for β . The rough estimate obtained this way is $\beta = 0.84$.

It appears to be possible to estimate β and ρ in a more systematic way using higher order joint moments of the X_n 's. This will be discussed elsewhere.

7. REFERENCES

- [1] ANDERSON, T.W. (1970). The Statistical Analysis of Time Series, Wiley, New York.
- [2] BERGER, J.M. and MANDELBROT, B. (1963). A new model for error clustering in telephone circuits. IBM Journal of Research and Development, 7, 224-36.
- [3] COX, D.R. and LEWIS, P.A.W. (1966). The Statistical Analysis of Series of Events. Methuen, London and Wiley, New York.
- [4] GAVER, D.P., LAVENBERG, S.S., and PRICE, Jr., T.G. (1976). Exploratory analysis of access path length data for a data base management system. IBM J. Research and Development.
- [5] GAVER, D.P. and LEWIS, P.A.W. (1977). First order autoregressive Gamma sequences and point processes. To appear.
- [6] JACOBS, P.A. and LEWIS, P.A.W. (1977). A mixed autoregressive-moving average exponential sequence and point process (EARMA 1,1). To appear in Adv. Appl. Prob.
- [7] KANTER, M. (1975). Autoregression for discrete processes mod 2. J. Appl. Prob., 12, 37-375.
- [8] KLOTZ, J. (1973). Statistical inference in Bernoulli trials with dependence. Ann. Statist., 1, 373-379.
- [9] LAWRENCE, A.J. and LEWIS, P.A.W. (1977). A moving average exponential point process (EMAL). To appear in J. Appl. Prob.
- [10] LEWIS, P.A.W. and COX, D.R. (1966). A statistical analysis of telephone circuit error data. IEEE Transactions on Communications Systems, Com-14, 382-9.

INITIAL DISTRIBUTION LIST

	Copies		Copies
Dean of Research Code 012 Naval Postgraduate School Monterey, CA 93940	2	Director Office of Naval Research Branch Office 1030 East Green Street Attn: Dr. A. R. Laufer Pasadena, CA 91101	1
Library, Code 0212 Naval Postgraduate School Monterey, CA 93940	2	Office of Naval Research Branch Office 1030 East Green Street Attn: Dr. Richard Lau Pasadena, CA 91101	1
Library, Code 55 Naval Postgraduate School Monterey, CA 93940	2		
P. A. W. Lewis, Code 55Lw Naval Postgraduate School Monterey, CA 93940	10	Office of Naval Research San Francisco Area Office 760 Market Street San Francisco, CA 94102	1
Professors G. G. Brown, Code 55Bw	1		
R. W. Butterworth, 55Bd	1	Technical Library	1
J. D. Esary, 55Ey	1	Naval Ordnance Station	
D. P. Gaver, 55Gv	1	Indian Head, MD 20640	
K. T. Marshall, 55Mt	1		
P. R. Milch, 55Mh	1	Office of Naval Research Branch Office 1030 East Green Street Attn: Dr. D. Osteyee Pasadena, CA 91101	1
F. R. Richards, 55Rh	1		
Naval Postgraduate School Monterey, CA 93940			
Statistics and Probability Program Office of Naval Research Attn: Dr. B. J. McDonald Arlington, VA 22217	3	Naval Ship Engineering Center Philadelphia Division Technical Library Philadelphia, PA 19112	1
Defense Documentation Center Cameron Station Alexandria, VA 22314	2	Bureau of Naval Personnel Department of the Navy Technical Library Washington, D. C. 20370	1
Technical Information Division Naval Research Laboratory Washington, D. C. 20390	6	Director, Naval Research Laboratory Attn: Library, Code 2029 (ONRL) Washington, D. C. 20390	6
Office of Naval Research New York Area Office 715 Broadway Attn: Dr. Robert Grafton New York, New York 10003	1		
Director Office of Naval Research Branch Office 495 Summer Street Attn: Dr. A. L. Powell Boston, MA 02210	1	Director Office of Naval Research Branch Office 536 South Clark Street Attn: Dr. A. R. Dawe Chicago, IL 60605	1

Copies

Library	1
Naval Electronics Laboratory Center San Diego, CA 92152	
Naval Undersea Center Technical Library San Diego, CA 92132	1
Applied Mathematics Laboratory Naval Ship Research and Development Center Attn: Mr. Gene H. Gleissner Washington, D. C. 20007	1