

AD-A012 941

A MOVING AVERAGE EXPONENTIAL POINT PROCESS (EMA1)

P. A. W. Lewis, et al

Naval Postgraduate School

Prepared for:

Office of Naval Research  
National Science Foundation

June 1975

DISTRIBUTED BY:

**NTIS**

**National Technical Information Service  
U. S. DEPARTMENT OF COMMERCE**

AD A 0 1 2 9 4 1

220149

NPS55Lw75061

# NAVAL POSTGRADUATE SCHOOL

## Monterey, California



A MOVING AVERAGE EXPONENTIAL POINT PROCESS (EMA1)

by

A. J. Lawrance

and

P. A. W. Lewis

June 1975

Approved for public release; distribution unlimited.

Reproduced by  
NATIONAL TECHNICAL  
INFORMATION SERVICE  
Springfield, Va. 22151



**UNCLASSIFIED**

**SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)**

times between events. Initial conditions to make the point process count stationary are given, and extensions to higher order moving averages and Gamma point processes are discussed.

**UNCLASSIFIED**

**SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)**

# A MOVING AVERAGE EXPONENTIAL POINT PROCESS (EMA1)

A. J. Lawrance  
University of Birmingham  
England

and

P. A. W. Lewis<sup>\*</sup>  
Naval Postgraduate School  
Monterey, California

## ABSTRACT

A construction is given for a stationary sequence of random variables  $\{X_i\}$  which have exponential marginal distributions and are random linear combinations of order one of an i.i.d. exponential sequence  $\{\epsilon_i\}$ . The joint and trivariate exponential distributions of  $X_{i-1}$ ,  $X_i$  and  $X_{i+1}$  are studied, as well as the intensity function, point spectrum and variance time curve for the point process which has the  $\{X_i\}$  sequence for successive times between events. Initial conditions to make the point process count stationary are given, and extensions to higher order moving averages and Gamma point processes are discussed.

### 1. Introduction

In this paper we discuss the stationary sequence of random variables  $\{X_i\}$  which are formed from an independent and identically distributed exponential sequence  $\{\epsilon_i\}$  according to the linear model

---

\*Support from the Office of Naval Research (Grant NR042-284), the National Science Foundation (Grant AG476) and the United Kingdom Science Research Council is gratefully acknowledged.

$$X_i = \begin{cases} \beta \epsilon_i & \text{with probability } \beta; \\ \beta \epsilon_i + \epsilon_{i+1} & \text{with probability } 1 - \beta. \end{cases} \quad (0 \leq \beta \leq 1, i=0, \pm 1, \pm 2, \dots). \quad (1.1)$$

In fact, the  $\{X_i\}$  form a sequence of exponential random variables, and it will be seen from (1.1) that adjacent members will be correlated. Such a type of first order moving average model arose out of the companion paper, Gaver and Lewis (1975); there the first order autoregressive model

$$\begin{aligned} X_i &= \rho X_{i-1} + \epsilon'_i \quad (i = 0, \pm 1, \pm 2, \dots), \\ &= \sum_{k=0}^{\infty} \rho^k \epsilon'_{i-k} \end{aligned} \quad (1.2)$$

with exponential marginal distributions for the  $\{X_i\}$  is investigated. It is found there that the  $\epsilon'_i$  must be a mixture of a discrete component at zero and an exponential variable. The motivation behind both models (1.1) and (1.2) was three-fold: partly as an alternative to the normality theory of time series, partly as a model for correlated positive random variables with exponential marginal distributions but chiefly as a simple point process model with which to analyze non-Poisson series of events and to study the power of Poisson tests--particularly in situations where there is no obvious physically motivated model.

In the present paper we give a fairly complete picture of the model (1.1), which will be called EMA1 (exponential moving average of order 1), as a stationary point process. Distributions of the sums of the  $X_i$  are obtained and lead to counting properties of the process; the joint distributions of two and three adjacent intervals  $X_i$  are derived and appear to be new bivariate and trivariate exponential distributions. The distributions are investigated through their conditional means and variances, and computations of a conditional correlation

are given. Extensions of the model and estimation problems are briefly discussed.

In developing the properties of the process we will also point out similarities to a backward first order moving average which is defined as

$$X_i = \begin{cases} \beta \epsilon_i & \text{with probability } \beta, \\ \beta \epsilon_i + \epsilon_{i-1} & \text{with probability } 1 - \beta. \end{cases} \quad (0 \leq \beta \leq 1; i=0, \pm 1, \pm 2, \dots). \quad (1.3)$$

Properties of the processes are very similar, but those of the forward model (1.1) have simpler derivations.

It should also be noted that the model (1.1) can be written as a very special type of linear model with random coefficients:

$$X_i = \beta \epsilon_i + I_i \epsilon_{i+1} \quad (0 \leq \beta \leq 1, \quad i = 0, \pm 1, \pm 2, \dots),$$

where the  $I_i$  are i.i.d. Bernoulli random variables which are 1 with probability  $1 - \beta$  and 0 with probability  $\beta$ . This characterization is not very helpful for the first order model; the main point is that since the random coefficient has a probability which is just the parameter  $\beta$ , many of the theorems for linear processes are not applicable.

## 2. Some Basic Aspects of the EMA1 Model

The simplest aspect of the EMA1 model is the exponential marginal distribution of the intervals  $\{X_i\}$ ; in point process terminology (see e.g. Lawrance, 1972) this is the synchronous distribution of intervals and refers to the distribution of the interval from an arbitrarily chosen event to the next event. For the Laplace transform of its probability density function (p.d.f.)  $f_{X_i}(x)$ , we write

$$\begin{aligned} f_{X_i}^*(s) &= E\{e^{-sX_i}\} \\ &= E\{e^{-s\beta\epsilon_i}\}\beta + E\{e^{-s\beta\epsilon_i - s\epsilon_{i+1}}\}(1-\beta) \end{aligned} \quad (2.1)$$

using (1.1). Now the i.i.d. random variables  $\epsilon_i$  have exponential distributions with parameters  $\lambda$ , say, and so their Laplace transform is  $\lambda/(\lambda+s)$ . Thus (2.1) becomes

$$f_{X_i}^*(s) = \frac{\lambda}{\lambda+\beta s} \cdot \beta + \frac{\lambda}{\lambda+\beta s} \cdot \frac{\lambda}{\lambda+s} \cdot (1-\beta) = \frac{\lambda}{\lambda+s}.$$

This demonstrates that the  $X_i$  have identical exponential distributions as asserted. The parameter  $\lambda$  is thus the number of events per unit time, or the rate of the point process.

The correlation between  $X_i$  and  $X_{i+1}$  is easily obtained on considering the product of  $X_i$  from (1.1) with

$$X_{i+1} = \begin{cases} \beta\epsilon_{i+1} & \text{with probability } \beta, \\ \beta\epsilon_{i+1} + \epsilon_{i+2} & \text{with probability } (1-\beta). \end{cases}$$

Thus, again using straightforward conditioning arguments,



$$\begin{aligned}
\text{cov}(X_i, X_{i+1}) &= E(\beta \epsilon_i \epsilon_{i+1}) \beta^2 \\
&+ E(\beta^2 \epsilon_i \epsilon_{i+1} + \beta \epsilon_i \epsilon_{i+2}) \beta (1-\beta) \\
&+ E(\beta^2 \epsilon_i \epsilon_{i+1} + \beta \epsilon_{i+1}^2) \beta (1-\beta) \\
&+ E(\beta^2 \epsilon_i \epsilon_{i+1} + \beta \epsilon_i \epsilon_{i+2} + \epsilon_{i+1} \epsilon_{i+2} + \beta \epsilon_{i+1}^2) (1-\beta)^2,
\end{aligned}$$

and simplification of this result leads to

$$\rho_1 = \text{corr}(X_i, X_{i+1}) = \beta(1-\beta). \quad (2.2)$$

By the construction of EMAl, the higher order serial correlations will be zero, and thus the spectral density of intervals (Cox and Lewis, 1966, p. 70),

$$f_{\pm}(\omega) = \frac{1}{\pi} \left\{ 1 + 2 \sum_{k=1}^{\infty} \rho_k \cos(k\omega) \right\}, \quad (0 \leq \omega \leq \pi),$$

becomes

$$f_{\pm}(\omega) = \frac{1}{\pi} \{ 1 + 2\beta(1-\beta) \cos(\omega) \}. \quad (0 \leq \omega \leq \pi). \quad (2.3)$$

The result (2.2) is the greatest limitation of the EMAl model since it implies that the first order serial correlation is non-negative and bounded by 1/4; this may be compared with an ordinary MA1 model assuming two-sided  $\epsilon_i$  distributions of mean zero for which  $|\rho_1| \leq 1/2$ . In both cases it can be anticipated that the restrictions are a consequence of the linearity of the models.

A further simple aspect of the EMAl model is that the  $\{X_i\}$  sequence reduces to the Poisson process when  $\beta = 0$  or 1, and this gives checks on most of our results. We may also note that the moving average is taken in the forward sense; the backward model (1.3) could equally have been treated,

although producing different but similar results. This serves to emphasize that there is no time-reversibility in the process, in the sense that  $\{X_1, \dots, X_k\}$  does not have the same joint probability distribution as  $\{X_{-1}, \dots, X_{-k}\}$  for all finite  $k$ , where  $k \geq 2$ .

### 3. Distributions of Sums and Counts in $\{X_i\}$ Sequence

In the point process theory of the model, the distribution of the sums  $T_r = X_1 + \dots + X_r$  are very useful; if these can be obtained then the distributions of counts, both in the synchronous and asynchronous mode, can then be derived. As shown in Cox and Lewis (1966, Chapter 4) for instance, these then lead to the second order properties such as the intensity function, the (Bartlett) spectrum of counts and the variance time curve. It is, therefore, a particularly attractive feature of the EMA1 model that the distribution of the  $T_r$  may be obtained, and we shall now give a simple derivation.

Define  $\psi(s)$  as the Laplace transform of the p.d.f. of the  $\epsilon_1$  distribution: except where otherwise remarked this distribution is exponential of parameter  $\lambda$  and so  $\psi(s) = \lambda/(\lambda+s)$ . Define the double Laplace transform (or equivalently the joint moment generating function) of  $T_r$  and  $\epsilon_{r+1}$ , as

$$\phi_r(s_1, s_2) = E\{e^{-s_1 T_r - s_2 \epsilon_{r+1}}\} \text{ for } r = 1, 2, \dots \quad (3.1)$$

For  $r = 1$ , we have

$$\begin{aligned} \phi_1(s_1, s_2) &= E\{e^{-s_1 X_1 - s_2 \epsilon_2}\} = E\{e^{-s_1 \beta \epsilon_1 - s_2 \epsilon_2}\} \beta + E\{e^{-s_1 \beta \epsilon_1 - (s_1 + s_2) \epsilon_2}\} (1-\beta) \\ &= \psi(\beta s_1) [\beta \psi(s_2) + (1-\beta) \psi(s_1 + s_2)] \end{aligned} \quad (3.2)$$

and we shall write

$$\psi(s_1, s_2) = \beta \psi(s_2) + (1-\beta) \psi(s_1 + s_2). \quad (3.3)$$

This is the double Laplace transform of a joint distribution in which the first variable has mass  $\beta$  at zero and with probability  $(1-\beta)$  is exponential distributed. We shall now relate  $\phi_r(s_1, s_2)$  and  $\phi_{r-1}(s_1, s_2)$ . Since

$$T_r = T_{r-1} + X_r$$

$$= \begin{cases} T_{r-1} + \beta \epsilon_r & \text{with probability } \beta \\ T_{r-1} + \beta \epsilon_r + \epsilon_{r+1} & \text{with probability } 1 - \beta, \end{cases}$$

we have

$$\begin{aligned} \phi_r(s_1, s_2) &= E\{e^{-s_1 T_{r-1} - s_1 \beta \epsilon_r - s_2 \epsilon_{r-1}}\} \beta + E\{e^{-s_1 T_{r-1} - s_1 \beta \epsilon_r - (s_1 + s_2) \epsilon_{r+1}}\} (1 - \beta) \\ &= \phi_{r-1}(s_1, \beta s_1) \psi(s_2) \beta + \phi_{r-1}(s_1, \beta s_1) \psi(s_1 + s_2) (1 - \beta) \\ &= [\beta \psi(s_2) + (1 - \beta) \psi(s_1 + s_2)] \phi_{r-1}(s_1, \beta s_1). \end{aligned} \quad (3.4)$$

Solving (3.4) gives,

$$\phi_r(s_1, s_2) = \psi(\beta s_1) [\psi(s_1, \beta s_1)]^{r-1} \psi(s_1, s_2), \quad (3.5)$$

and setting  $s_2 = 0$ , we have for the Laplace transform of the p.d.f. of  $T_r$ ,

$$\phi_r(s) = [\beta \psi(\beta s) + (1 - \beta) \psi(\beta s) \psi(s)] [\beta \psi(\beta s) + (1 - \beta) \psi((1 + \beta)s)]^{r-1} \quad (3.6)$$

$$= \frac{\lambda}{\lambda + s} \left\{ \frac{\lambda(\lambda + 2\beta s)}{(\lambda + 2\beta s)(\lambda + (1 + \beta)s)} \right\}^{r-1}, \quad r \geq 1. \quad (3.7)$$

This is our required result; from (3.7) it will be observed that  $T_r$  is distributed like the sum of  $r$  independently distributed variables, such as in a delayed renewal process, although these are not  $X$  variables. The structure of (3.6) or (3.7) is explained by the fact that the number of intervals which are of the  $\beta \epsilon_1$  form or  $\beta \epsilon_1 + \epsilon_{1+1}$  form are binomially distributed with parameter  $\beta$  or  $1 - \beta$ ; further consideration of the adjacencies of the two types of intervals than leads to the terms in the binomial expansion of (3.6).

We now obtain the distribution of  $N_t^{(f)}$ , the synchronous counting process of number of events occurring in the interval  $(0, t]$  beginning at an arbitrary event; this is related to the distribution of  $T_r$  through the equivalence of the events  $N_t^{(f)} < r$  and  $T_r > t$  for  $r \geq 1$ . Let  $F_r(t)$  denote the distribution of  $T_r$ , and then since

$$\text{Prob}\{N_t^{(f)} = r\} = F_r(t) - F_{r+1}(t), \quad r \geq 0, \quad (3.8)$$

with  $F_0(t) \equiv 1$  for  $t \geq 0$ , we have for the p.d.f. of  $N_t^{(f)}$ ,

$$\begin{aligned} E\{z^{N_t^{(f)}}\} &= \psi_f(z; t) = \sum_{r=0}^{\infty} z^r [F_r(t) - F_{r+1}(t)] \\ &= 1 + (z-1) \sum_{r=1}^{\infty} z^{r-1} F_r(t). \end{aligned} \quad (3.9)$$

Inserting (3.7) in the Laplace transform of (3.9) gives

$$\psi_f^*(z; s) = \frac{\beta(1+\beta)s^2 + [-\beta(1-\beta)z + 2\beta + 1]\lambda s + \lambda^2}{(s+\lambda)[\beta(1+\beta)s^2 + (1+2\beta-2\beta z)\lambda s + (1-z)\lambda^2]}. \quad (3.10)$$

This result is required in Section 4 to follow.

#### 4. The Intensity Function and Spectrum of Counts

The intensity function of a point process is the derivative with respect to  $t$  of  $E\{N_t^{(f)}\}$  and will be denoted by  $m_f(t)$ . The (Bartlett) spectrum of counts, the Fourier transform of the covariance density of the differential counting process, then has the simple expression

$$g_+(\omega) = \frac{\lambda}{\pi} \{1 + m_f^*(i\omega) + m_f^*(-i\omega)\}, \quad (4.1)$$

where  $m_f^*(s)$  is the Laplace transform of  $m_f(t)$ ; this expression for  $g_+(\omega)$  is derived in Cox and Lewis (1966, Section 4.5).

For the EMAL process, the result from (3.10) is that

$$m_f^*(s) = \frac{\lambda(\lambda+\beta s)\{\lambda+(1+\beta)s\}}{\beta(1+\beta)s(\lambda+s)\{s+\lambda/(\beta^2+\beta)\}}. \quad (4.2)$$

In inverting the Laplace transform (4.2) it will be noted that the case  $\beta^2 + \beta = 1$ , i.e.  $\beta = 0.6185$ , must be treated separately since there will then be a factor  $(\lambda+s)^2$  in the denominator. Partial fraction expansions and their inversion then give, for  $t \geq 0$ ,

$$m_f(t) = \lambda \left[ 1 + \frac{\beta(1-\beta)}{\beta^2+\beta-1} \left\{ e^{-\lambda t/(\beta^2+\beta)} - e^{-\lambda t} \right\} \right] \quad (\beta^2+\beta \neq 1), \quad (4.3)$$

$$= \lambda [1 + \beta^3 \lambda t e^{-\lambda t}] \quad (\beta^2+\beta=1). \quad (4.4)$$

We see in both cases that the initial value of  $m_f(t)$  is  $\lambda$  and that they both increase until maximum values are obtained at  $t = \lambda^{-1}(\beta^2+\beta) \times \log[(\beta^2+\beta)/(\beta^2+\beta-1)]$  and at  $t = \lambda^{-1}$  respectively for (4.3) and (4.4); both functions then decrease exponentially to  $\lambda$ . There is no apparent reason for the  $\beta^2 + \beta = 1$  case. When  $\beta = 0$  or  $1$  both functions are constant at  $\lambda$ , as is appropriate to the Poisson process.

The function  $m_f(t)$  is plotted in Figure 1 for several values of  $\beta$ . The spectrum of counts follows easily by inserting (4.2) into (4.1), and has the expressions

$$g_{\pm}(\omega) = \frac{\lambda}{\pi} \left\{ 1 + 2\lambda^2 \frac{\beta(1-\beta)}{\beta^2+\beta-1} \left[ \frac{\beta^2+\beta}{(\beta^2+\beta)^2\omega^2+\lambda^2} - \frac{1}{\omega^2+\lambda^2} \right] \right\} \quad (\beta^2+\beta \neq 1) \quad (4.5)$$

$$= \frac{\lambda}{\pi} \left\{ 1 + 2\lambda^2\beta^3 \left[ \frac{-\omega^2+\lambda^2}{(\omega^2+\lambda^2)^2} \right] \right\} \quad (\beta^2+\beta=1). \quad (4.6)$$

We observe that both these are ratios of 4th order polynomials in  $\omega$ . Estimation of both  $m_f(t)$  and  $g_{\pm}(\omega)$  given an actual sequence of interevent times is considered in Cox and Lewis (1966, Chapter 5); in practice these would then be compared with our given theoretical functions which are graphed in Figure 2.

Note that unlike the 2nd order joint moment functions  $\rho_k$  and  $f_{\pm}(\omega)$  for intervals, the second order moment functions for counts  $m_f(t)$  and  $g_{\pm}(\omega)$  do discriminate between the cases where the parameter is  $\beta$  or  $(1-\beta)$ . However the graphs in Figure 2 indicate that the count spectra of models with  $\beta$  in the range (0.25, 0.75) are fairly close to each other; therefore, the spectrum will not be entirely suitable for discriminating between different  $\beta$  values for small sample sizes.

The variance time curve is considered in Section 7, along with the stationary initial conditions for the process.

## 5. The Joint Distribution of $X_i$ and $X_{i+1}$

We now discuss the joint distribution of  $X_i$  and  $X_{i+1}$  which will be a bivariate exponential distribution. Several authors have discussed bivariate exponential distributions, including Downton (1970), who makes some comparisons with those of Gumbel, Moran and Marshall-Olkin. The distribution to be discussed here does not appear to be one of the earlier ones, although it is fair to say that in common with earlier ones, it is not the 'perfect' bivariate exponential.

The double Laplace transform of the joint pdf of  $X_i$  and  $X_{i+1}$  is easily calculated using (1.1); the required expectation is

$$\begin{aligned}
 E\{e^{-s_1 X_i - s_2 X_{i+1}}\} &= f_{X_i, X_{i+1}}^{**}(s_1, s_2) \\
 &= E\{e^{-\beta s_1 \epsilon_i - \beta s_2 \epsilon_{i+1}}\} \beta^2 + E\{e^{-\beta s_1 \epsilon_i - s_2 (\beta \epsilon_{i+1} + \epsilon_{i+2})}\} \beta(1-\beta) \\
 &\quad + E\{e^{-s_1 (\beta \epsilon_i + \epsilon_{i+1}) - \beta s_2 \epsilon_{i+1}}\} \beta(1-\beta) \\
 &\quad + E\{e^{-s_1 (\beta \epsilon_i + \epsilon_{i+1}) - s_2 (\beta \epsilon_{i+1} + \epsilon_{i+2})}\} (1-\beta)^2, \tag{5.1}
 \end{aligned}$$

which can be written

$$f_{X_i, X_{i+1}}^{**}(s_1, s_2) = \psi(\beta s_1) [\beta \psi(\beta s_2) + (1-\beta) \psi(s_1 + \beta s_2)] [\beta + (1-\beta) \psi(s_2)] \tag{5.2}$$

$$= \frac{\lambda^2 (\lambda + \beta s_1 + \beta s_2)}{(\lambda + \beta s_1) (\lambda + s_2) (\lambda + s_1 + \beta s_2)}. \tag{5.3}$$

We note that (5.3) is not symmetrical in  $s_1$  and  $s_2$ , and this is to be expected since the process is not time reversible; this is one feature which distinguishes it from earlier bivariate exponentials. The backward moving average model (1.3) corresponding to (1.1) has the joint interval distribution which is specified by (4.3) with  $s_1$  and  $s_2$  interchanged.



An explicit form of the joint distribution (5.3) can be obtained directly, rather than by inversion of the transform which is less informative. By the structure of the model the joint distribution of  $(X_i, X_{i+1})$  is a mixture of the joint distributions of  $(\beta\epsilon_i, \beta\epsilon_{i+1})$ ,  $(\beta\epsilon_i, \beta\epsilon_{i+1} + \epsilon_{i+2})$ ,  $(\beta\epsilon_i + \epsilon_{i+1}, \beta\epsilon_{i+1})$  and  $(\beta\epsilon_i + \epsilon_{i+1}, \beta\epsilon_{i+1} + \epsilon_{i+2})$  with the corresponding probabilities  $\beta^2$ ,  $\beta(1-\beta)$ ,  $\beta(1-\beta)$  and  $(1-\beta)^2$ . These joint pdf's can be listed in an obvious notation as follows:

$$f_{\beta\epsilon_i, \beta\epsilon_{i+1}}(x, y) = (\lambda/\beta)e^{-(\lambda/\beta)x}(\lambda/\beta)e^{-(\lambda/\beta)y}, \quad (x, y > 0)$$

$$f_{\beta\epsilon_i, \beta\epsilon_{i+1} + \epsilon_{i+2}}(x, y) = \lambda e^{-\lambda x} (1-\beta)^{-1} [\lambda e^{-\lambda y} - \lambda e^{-(\lambda/\beta)y}], \quad (x, y > 0)$$

$$f_{\beta\epsilon_i + \epsilon_{i+1}, \beta\epsilon_{i+1}}(x, y) = (\lambda/\beta)e^{-(\lambda/\beta)(x-y/\beta)} (\lambda/\beta)e^{-(\lambda/\beta)y}, \quad (\beta x > y > 0)$$

$$f_{\beta\epsilon_i + \epsilon_{i+1}, \beta\epsilon_{i+1} + \epsilon_{i+2}}(x, y) = \begin{cases} \lambda^2 e^{-(\lambda/\beta)x} [e^{\lambda(1-\beta)y/\beta^2} - e^{-\lambda y}] / (1-\beta+\beta^2), & (\beta x > y > 0) \\ \lambda^2 [e^{-\lambda(1-\beta)x} - e^{-(\lambda/\beta)x}] e^{-\lambda y} / (1-\beta+\beta^2). & (y > \beta x > 0). \end{cases} \quad (5.4)$$

We thus see that the joint pdf of  $X_i, X_{i+1}$  will be continuous in both variables but will have different analytical expressions over the regions  $\beta x > y$  and  $\beta x \leq y$ ; there appears to be no compact analytical form for  $f_{X_i, X_{i+1}}(x, y)$ . This is unfortunate because it makes it difficult to derive maximum likelihood estimates of the parameters  $\lambda$  and  $\beta$  in the model.

Different bivariate exponentials also can be compared through their conditional properties and so we will derive these for the present distribution. Conditional pdf's are not succinct enough, and so we concentrate on conditional moments. These may be obtained from (5.3). For instance, to obtain  $E(X_1 | X_{i-1} = t)$  we differentiate with respect to  $s_2$ , set  $s_2 = 0+$ , invert with respect to  $s_1$  and then divide by the marginal (exponential) density of  $X_{i-1}$ . The two conditional means are in this way found to be

$$E(X_1 | X_{i-1} = t) = \lambda^{-1} \left[ \beta \lambda t + \frac{1-2\beta}{1-\beta} + \frac{\beta}{1-\beta} e^{-\lambda(1-\beta)t/\beta} \right] \quad (5.5)$$

and

$$E(X_1 | X_{i+1} = t) = \lambda^{-1} [1 + \beta e^{-(1-\beta)\lambda t/\beta}]. \quad (5.6)$$

Thus, both regressions have exponential components; this property is shared by the Marshall and Olkin bivariate distribution, although that distribution has a singular component along  $X_1 = X_{i+1}$ . For the continuous distribution treated by Downton both the conditional means are linear, as are the conditional variances.

Examining these regression functions more closely we see that  $E(X_1 | X_{i+1} = t)$  is equal to  $\lambda^{-1}$  for  $\beta = 0$  or  $\beta = 1$ ; otherwise it increases exponentially from  $\beta \lambda^{-1}$  to the constant value  $(1+\beta)\lambda^{-1}$  as  $t$  increases. The transient is long for  $\beta$  close to 1, but very short when  $\beta$  is close to 0. Thus unlike the serial correlation coefficient  $\rho_1$  there is differentiation in this conditional mean between the cases where the parameter is  $\beta$  and the case when it has value  $1 - \beta$ .

The conditional mean (5.5) is more complex. It starts at  $t = 0$  with value  $\lambda^{-1}$  and negative slope  $\beta - 1$ . There is a unique minimum at  $t = -\beta \ln \beta / \{\lambda(1-\beta)\}$  and the function eventually increases linearly with  $t$ . Since we have for large  $t$  that

$$E[X_1 | X_{i-1}=t] \sim \lambda^{-1} \left[ \frac{1-2}{1-\beta} \right] + \beta t,$$

the rate of increase depends only on  $\beta$ , not on  $\lambda$ .

The conditional variances for the present bivariate exponential are also exponential functions, and their explicit forms are given by

$$\text{Var}(X_i | X_{i-1}=t) = \lambda^{-2} \left[ \frac{1-2\beta+2\beta^3}{(1-\beta)^2} + \frac{2\beta^2(1+\lambda t)}{1-\beta} e^{-(1-\beta)\lambda t/\beta} - \frac{\beta^2}{(1-\beta)^2} e^{-2(1-\beta)\lambda t/\beta} \right] \quad (5.7)$$

and

$$\text{Var}(X_i | X_{i+1}=t) = \lambda^{-2} \left[ \frac{1+\beta+\beta^2-\beta^3}{1-\beta} - 2 \left\{ \frac{\beta}{1-\beta} + \frac{\lambda t}{\beta} \right\} e^{-\lambda(1-\beta)t/\beta} - e^{-2(1-\beta)\lambda t/\beta} \right]. \quad (5.8)$$

These conditional variances are quite different forms as shown in Figures 3 and 4. In practice it is clear that conditional means and variance could only be calculated for  $t$  in the more central regions of the marginal distributions. In these situations  $\text{Var}(X_i | X_{i-1})$  is fairly constant, while  $\text{Var}(X_i | X_{i+1})$  is reasonably linear in  $t$ . In all cases the asymptotic values are reached much quicker for the lower value of  $\beta$ .

6. The Conditional Correlation of  $X_{i-1}$ ,  $X_{i+1}$  given  $X_i$

We now wish to carry the study of dependence in the sequence of intervals  $\{X_i\}$  a step further, in particular to trivariate distributions. The dependence in the EMAL process has a very particular structure:  $X_i$  is dependent on  $X_{i-1}$  and  $X_{i+1}$  but not on  $X_{i-2}$ ,  $X_{i+2}$ ,  $X_{i-3}$ ,  $X_{i+3}$ , and so on. It thus appears that the joint distribution of  $X_{i-1}$ ,  $X_i$ ,  $X_{i+1}$  has some natural significance for this process, and it will be a trivariate exponential distribution; we should note however that in view of the coupling effect of the dependence, this trivariate distribution is not enough to describe completely the dependence in the sequence  $\{X_i\}$ . In particular the sequence is certainly not Markovian since the distribution of  $X_{i+1} | X_i, X_{i-1}$  will depend on the value of  $X_{i-1}$ .

The process, by its structure, has the somewhat strange feature that although  $X_{i-1}$  and  $X_{i+1}$  both depend on  $X_i$ , the variables  $X_{i-1}$  and  $X_{i+1}$  are independent. For this reason, it is felt that the joint distribution of  $X_{i-1}$  and  $X_{i+1}$  conditional on  $X_i$  is of interest, and we shall give calculations of the conditional correlation of  $X_{i-1}$  and  $X_{i+1}$  given  $X_i = t$ . The other two pairwise conditional joint distributions may also of course be used, but the corresponding unconditional joint distributions show that the intervals concerned are not independent. We think of the conditional correlation, written  $\text{Corr}(X_{i-1}, X_{i+1} | X_i = t)$ , as a descriptive function of the higher order dependence, with the thought that it may be used comparatively with other trivariate exponentials. The general properties of conditional correlations are not well understood, but Lawrance (1975) has shown that it is equal to the corresponding partial correlation only in very special cases, one of which is the trivariate normal, and the present distribution is not one of these cases.

The triple Laplace transform of the joint p.d.f. of  $X_{i-1}$ ,  $X_i$ ,  $X_{i+1}$  is calculated by a straightforward extension of the procedure used to obtain the

bivariate Laplace transform at (5.2). The result is the sum of eight expectation terms with their associated binomial probabilities, and can be cast in the form

$$E\{e^{-s_1 X_{i-1} - s_2 X_i - s_3 X_{i+1}}\} = f_{X_{i-1}, X_i, X_{i+1}}^{***}(s_1, s_2, s_3) \\ = \psi(\beta s_1) \{\beta \psi(\beta s_2) + (1-\beta) \psi(s_1 + \beta s_2)\} \{\beta \psi(\beta s_3) + (1-\beta) \psi(s_2 + \beta s_3)\} \{\beta + (1-\beta) \psi(s_3)\}. \quad (6.1)$$

This reduces to the appropriate bivariate distributions where one  $s$  is set to zero. Before passing to the conditional moments, we may note that the generalization of (6.1) to  $r$  adjacent intervals is

$$E\{\exp[-\sum_{i=1}^r s_i X_i]\} = \psi(\beta s_1) \prod_{j=2}^r [\beta \psi(\beta s_j) + (1-\beta) \psi(s_{j-1} + s_j)] [\beta + (1-\beta) \psi(s_r)]. \quad (6.2)$$

When  $s_1 = s_2 = \dots = s_r$  we recover the result for  $X_1 + X_2 + \dots + X_r$  given at (3.5).

We now return to  $\text{Corr}(X_{i-1}, X_{i+1} | X_i = t)$  which we shall denote as  $\rho_2(t)$  the conditional correlation of  $X_{i-1}$  and  $X_{i+1}$  given  $X_i = t$ . This has the explicit expression

$$\rho_2(t) = \frac{E(X_{i-1}, X_{i+1} | X_i = t) - E(X_{i-1} | X_i = t) E(X_{i+1} | X_i = t)}{[\text{Var}(X_{i-1} | X_i = t) \text{Var}(X_{i+1} | X_i = t)]^{1/2}}. \quad (6.3)$$

In view of the results (5.5)-(5.8), there only remains to calculate

$E(X_{i-1}, X_{i+1} | X_i = t)$ . This is obtained from (6.1) by inverting

$$(\lambda e^{-\lambda t})^{-1} \frac{\partial^2}{\partial s_1 \partial s_2} [f_{X_{i-1}, X_i, X_{i+1}}(s_1, s_2, s_3)]_{(s_1=s_3=0)} \quad (6.4)$$

as a function of  $s_2$ , to recover the variable  $t$ . After subtraction of the product of the conditional means, we have for the conditional covariance

$$\begin{aligned} & \text{Cov}(X_{i-1}, X_{i+1} | X_i = t) \\ &= -\frac{\beta^2}{1-\beta} + \{(1-\beta)\lambda t - \beta\} e^{-(1-\beta)\lambda t/\beta} - \frac{\beta}{1-\beta} e^{-2(1-\beta)\lambda t/\beta} \quad (6.5) \end{aligned}$$

Hence the expression for  $\rho_2(t)$  and the graphs given in Fig. 5. The conditional correlation is far from constant in  $t$ , although in the range  $(0, 2\lambda)$ , within which it would be possible to estimate it in practice, the values are positive and small.

## 7. Stationary Initial Conditions

Up to this point we have dwelt on aspects of the process which involve the intervals between the events, we have emphasized that these are a correlated but stationary sequence of exponential variables. This situation is typified by the choice of an arbitrary event for the initial point of a sequence of intervals. We now consider the corresponding problem when the initial point is chosen without knowledge of the event times; this is usually called an arbitrary time and is of interest when stationarity in the counts of events is suggested (Cox and Lewis, 1966, Chapter 4), as opposed to stationarity in the intervals between events. However, for stationarity in counts of events, the initial point of the interval of the counting must be chosen in a particular probabilistic way. We shall now obtain the appropriate initial conditions, using the approach and definition discussed in Lawrance (1972) in which the process is considered at time  $t$  and  $t$  is then allowed to tend to infinity. The sequence of intervals between events beginning with the arbitrary time, usually called the asynchronous sequence, is not exponential or stationary, but the counting variable of this sequence has stationary increments, although not Poisson distributed.

At time  $t$  in the process (after a start in any convenient way) it is apparent that for the process to continue, we must specify:

- (i) the time to the next event in the  $\{X_i\}$  sequence, and
- (ii) the random variable  $\epsilon_{i+1}$  which is associated with the end of the  $X_i$  interval covering  $t$ . The first of these will be denoted by  $\chi$  and is just the forward recurrence time of the EMAI process, and this is bound asymptotically to be exponential, but it will be dependent on the second, denoted by  $e$  which will not be exponential, even asymptotically. It is their joint distribution as  $t \rightarrow \infty$  which gives the required initial conditions.

Suppose the process starts at  $t = 0$  in the synchronous mode, and suppose that in  $(0, t]$  there are  $r-1$  events. Let the joint pdf of  $T_{r-1}$  and  $\beta\epsilon_r$  be  $f_{T_{r-1}, \beta\epsilon_r}(x, y)$ . When the  $r^{\text{th}}$  interval is of the  $\beta\epsilon_r$  form, then the joint pdf for  $(\chi = w, e = z)$  is

$$\int_{x=0}^t f_{T_{r-1}, \beta\epsilon_r}(x, t-x+w) dx \psi_\epsilon(z), \quad (7.1)$$

where  $\psi_\epsilon(z)$  is the pdf of  $\epsilon_{r+1}$ . If the  $r^{\text{th}}$  interval is of the  $\beta\epsilon_r + \epsilon_{r+1}$  form, there are two similar expressions according as  $z < w$  or  $z > w$ ; these are

$$\int_{x=0}^t f_{T_{r-1}, \beta\epsilon_r}(x, t-x+w-z) dx \psi_\epsilon(z) \quad (z < w) \quad (7.2)$$

and

$$\int_{x=0}^{t-(z-w)} f_{T_{r-1}, \beta\epsilon_r}(x, t-x+w-z) dx \psi_\epsilon(z). \quad (z > w) \quad (7.3)$$

The expressions become evident on considering the configuration of events.

The joint pdf of  $\chi$  and  $e$  at time  $t$  may thus be written

$$f_{\chi, e}(w, z; t) = \left\{ \begin{array}{l} \sum_{r=0}^{\infty} \int_{x=0}^t f_{T_{r-1}, \beta\epsilon_r}(x, t-x+w) dx \psi_\epsilon(z) \quad \text{with probability } \beta \\ \sum_{r=0}^{\infty} \int_{x=0}^t f_{T_{r-1}, \beta\epsilon_r}(x, t-x+w-z) dx \psi_\epsilon(z) \quad (z < w) \\ \sum_{r=0}^{\infty} \int_{x=0}^{t-(z-w)} f_{T_{r-1}, \beta\epsilon_r}(x, t-x+w-z) dx \psi_\epsilon(z) \quad (z > w) \end{array} \right\} \begin{array}{l} (7.4) \\ \text{with} \\ \text{probability} \\ 1 - \beta. \end{array}$$

The  $r = 0$  and  $1$  terms here are really special cases, but will not contribute as  $t \rightarrow \infty$  and do not need to be obtained explicitly. We shall now use the result that

$$\lim_{t \rightarrow \infty} f_{\chi, e}(w, z; t) = \lim_{s \rightarrow 0} sf^*(w, z; s) = f_{\chi, e}(w, z) \quad (7.5)$$



to obtain the limit distribution at an arbitrary time. Now for the Laplace transform with respect to  $t$  of (7.4) we need the joint pdf of  $T_{r-1}$  and  $\beta\epsilon_r$ , which by (3.4) is

$$f_{T_{r-1}, \beta\epsilon_r}(x, y) = \int_{u=0}^x C_{r-1}(x-u)k(u, y)du, \quad (7.6)$$

and in terms of Laplace and double Laplace transforms

$$C_{r-1}^*(s) = \psi(\beta s)[\psi(s, \beta s)]^{r-2}, \quad \text{and} \quad k^{**}(s_1, s_2) = \beta\psi(\beta s_2) + (1-\beta)\psi(s_1 + \beta s_2). \quad (7.7)$$

Hence the Laplace transform with respect to  $t$  of the first line of (7.4) after ignoring the  $r = 0$  and  $r = 1$  terms is

$$\frac{\psi(\beta s)}{1-\psi(s, \beta s)} \int_{u=0}^{\infty} \int_{a=0}^{\infty} e^{-su-sa} k(u, a+w) da \psi_{\epsilon}(z). \quad (7.8)$$

Taking the limit as in (7.5) then gives

$$v\psi_{\epsilon}(z) \int_{u=0}^{\infty} \int_{a=0}^{\infty} k(u, a+w) da = v\psi_{\epsilon}(z) \int_{a=0}^{\infty} f_{\beta\epsilon_r}(a+w) da = v\psi_{\epsilon}(w/\beta)\psi_{\epsilon}(z), \quad (7.9)$$

where  $v$  is the mean of the  $\epsilon$  distribution and  $\psi_{\epsilon}(z)$  is its survivor function. The limits of the other terms in (7.4) can similarly be obtained, and give the final result as

$$f_{\chi, \epsilon}(w, z) = \left\{ \begin{array}{ll} v^{-1} \psi_{\epsilon}(w/\beta)\psi_{\epsilon}(z) & z, w > 0 \\ v^{-1} \psi_{\epsilon}\{(w-z)/\beta\}\psi_{\epsilon}(z) & 0 < z < w \\ v^{-1} \psi_{\epsilon}(z) & 0 < w < z \end{array} \right\} \begin{array}{l} \text{with probability} \\ \beta \\ \\ \text{with probability} \\ 1 - \beta \end{array} \quad (7.10)$$

The marginal distribution of  $\epsilon$  has pdf

$$f_{\epsilon}(z) = \beta\psi_{\epsilon}(z) + (1-\beta)z\psi_{\epsilon}(z)/v. \quad (7.11)$$

The marginal distribution of  $\chi$  is in general rather complicated, but in the EMAI case is exponential with parameter  $\lambda$ . From (7.11) we see that in the EMAI case the distribution of the first  $\epsilon$  variable after an arbitrary time (e) is the weighted sum of exponential and Erlang 2 distributions. This result implies that the second asynchronous interval does not have the exponential distribution, although all the following intervals do; the non-stationarity of the asynchronous sequence of intervals is thus caused only by the second interval.

The distribution for the number of events in  $(0,t]$  when  $t = 0$  is an arbitrary time, that is in the stationary situation, may now be obtained directly. As in the synchronous case of section 3 we need the distribution of the time to the  $r^{\text{th}}$  event for  $r \geq 1$ . The function  $\phi_1(s_1, s_2)$  of section 3 is now the double Laplace transform of (7.10), and so

$$\phi_1(s_1, s_2) = \frac{1}{vs_1} [\psi(s_2) - \psi(\beta s_1) \{ \beta \psi(s_2) + (1-\beta) \psi(s_1 + s_2) \}]. \quad (7.12)$$

Generally, for the double Laplace transform of the pdf of  $T_r$  and  $\epsilon_{r+1}$  measured from an arbitrary event, we have as at (3.2),

$$\phi_r(s_1, s_2) = \psi(s_1, s_2) [\psi(s_1, \beta s_1)]^{r-2} \phi_1(s_1, \beta s_1) \quad (r \geq 2). \quad (7.13)$$

This leads, using (3.9) to the Laplace transform of the pgf of  $N(t)$  as

$$\phi^*(z; s) = \frac{\beta(1+\beta)s^2 + [-\beta(1-\beta)z + 2\beta + 1]\lambda s + [1 + \beta(1-\beta)z(1-z)]\lambda^2}{(s+\lambda) \{ \beta(1+\beta)s^2 + (1+2\beta-2\beta z)\lambda s + (1-z)\lambda^2 \}}. \quad (7.14)$$

Setting  $\beta = 0$  or  $1$  reduces this to the Poisson process result and reminds us that the distribution of  $N(t)$  here can be considered as a generalization of the Poisson distribution appropriate to counting events in a correlated exponential sequence. The customary differentiations and inversions of (7.14) give

$$E\{N(t)\} = \lambda t$$

and

$$\text{Var}\{N(t)\} = [1+2\beta(1-\beta)]\lambda t - 2\beta(1-\beta)(1+\beta+\beta^2) - \frac{2\beta(1-\beta)}{\beta^2+\beta-1} [(\beta^2+\beta)^2 e^{-\lambda t/(\beta^2+\beta)} - e^{-\lambda t}] \quad (7.15)$$

when  $\beta^2 + \beta \neq 1$ ; there is an individual expression for (7.15) when  $\beta^2 + \beta = 1$ .

We notice that the distribution is asymptotically over dispersed as compared to the Poisson distribution. The results (7.14), (7.15) may also be obtained from general theory and the previous synchronous results, but the initial conditions have much wider applicability.

We have then been able to explicitly obtain the main probabilistic properties of the EMAl process in respect of stationary intervals and stationary counts; the process is thus unusually tractable, and this is of considerable merit as compared with many other models.

## 8. Conclusions and Extensions

There are several extensions to both the first order autoregressive and moving average point processes and sequences which will be considered subsequently:

- (i) By replacing  $\epsilon_{i+1}$  in (1.1) with  $\gamma\epsilon_{i+1}$  with probability  $\gamma$  and with  $\gamma\epsilon_{i+1} + \epsilon_{i+2}$  we obtain a second order moving average process. This may be extended to any order; like the present model the serial correlations are restricted to lie between 0 and 1/4.
- (ii) The autoregressive and moving average structures can be combined to give what appears to be a much richer class of processes.
- (iii) In Gaver and Lewis (1975) it is shown that if the  $X_i$  is taken to be Gamma distributed  $(K, \lambda)$ , then the solution to (1.2) shows that  $\epsilon_i'$  has Laplace transform  $\{(\rho\lambda+s)/(\lambda+s)\}^k$  and this is the Laplace transform of an infinitely divisible distribution. Thus autoregressive, moving average and mixed Gamma processes can be constructed. Their properties are much more complex than the corresponding exponential processes, but are tractable.

The EMAL and EMap processes are easily simulated, as are the Gamma processes for integer  $k$ . Estimation problems remain to be considered; they are treated for the first-order autoregressive processes in Gaver and Lewis (1975). The use of the EMAL sequence and point process in cluster processes, congestion models and computer systems models will be discussed elsewhere.

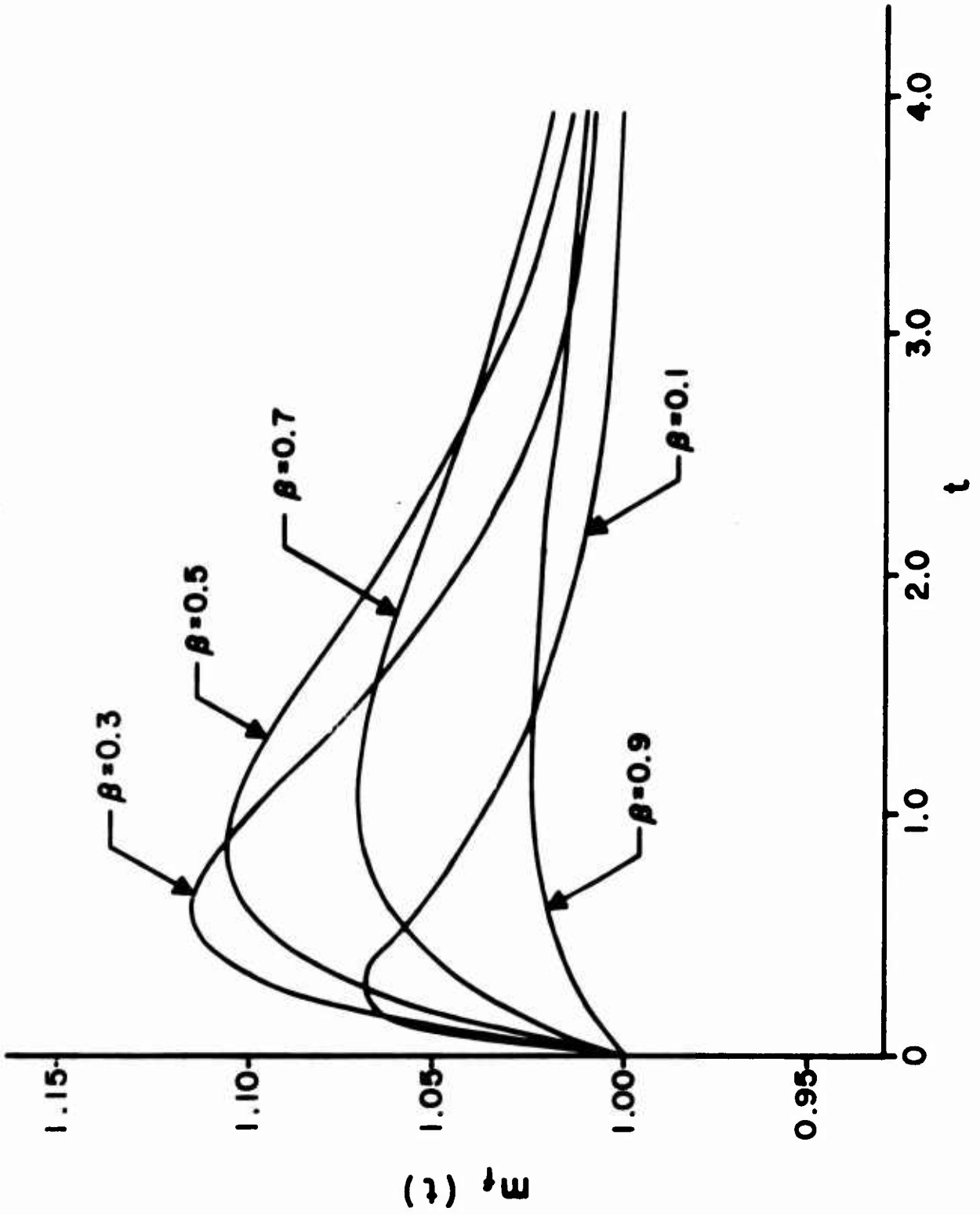


Figure 1

(z) + w

24a

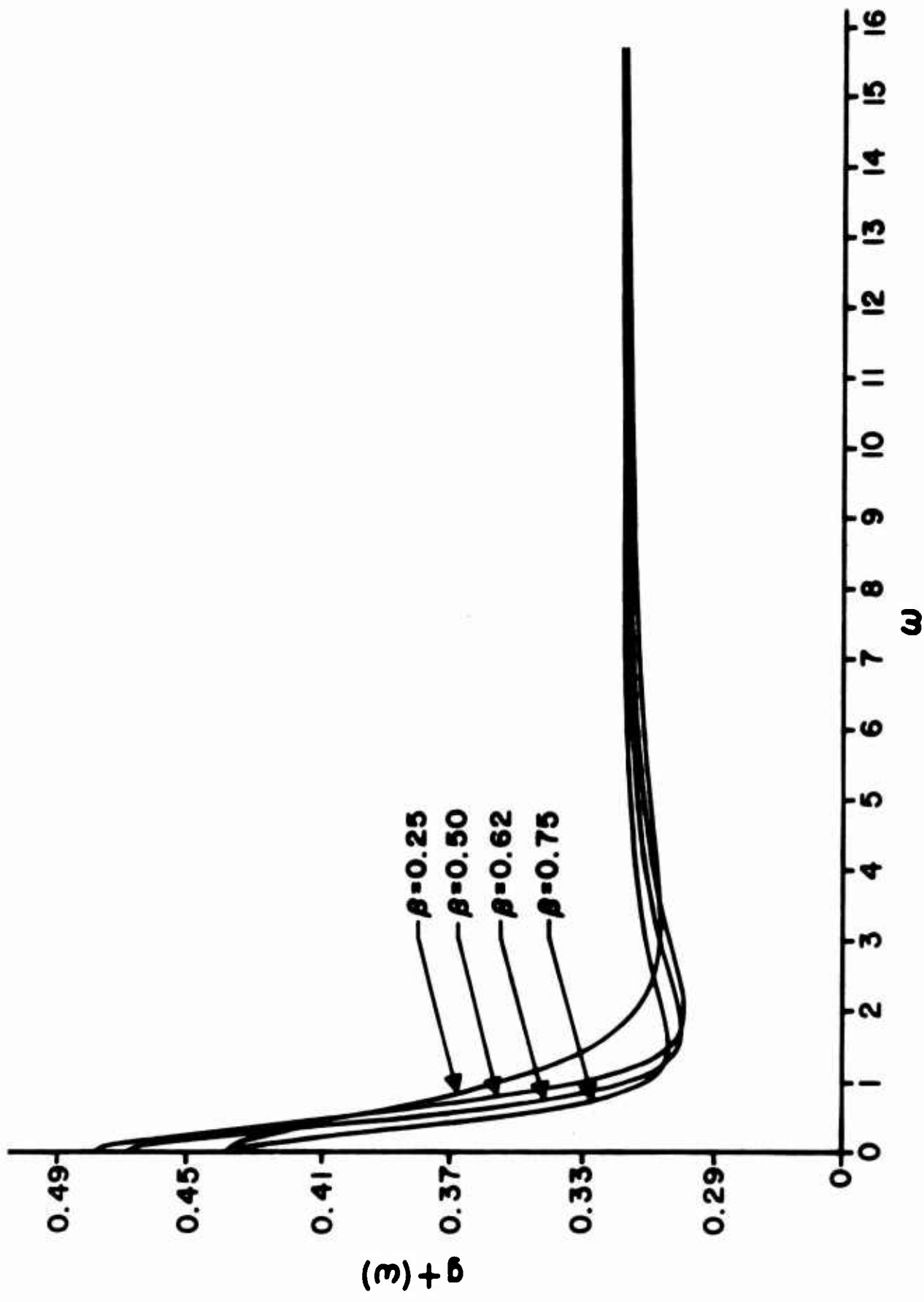


Figure 2

(3) + b

24b

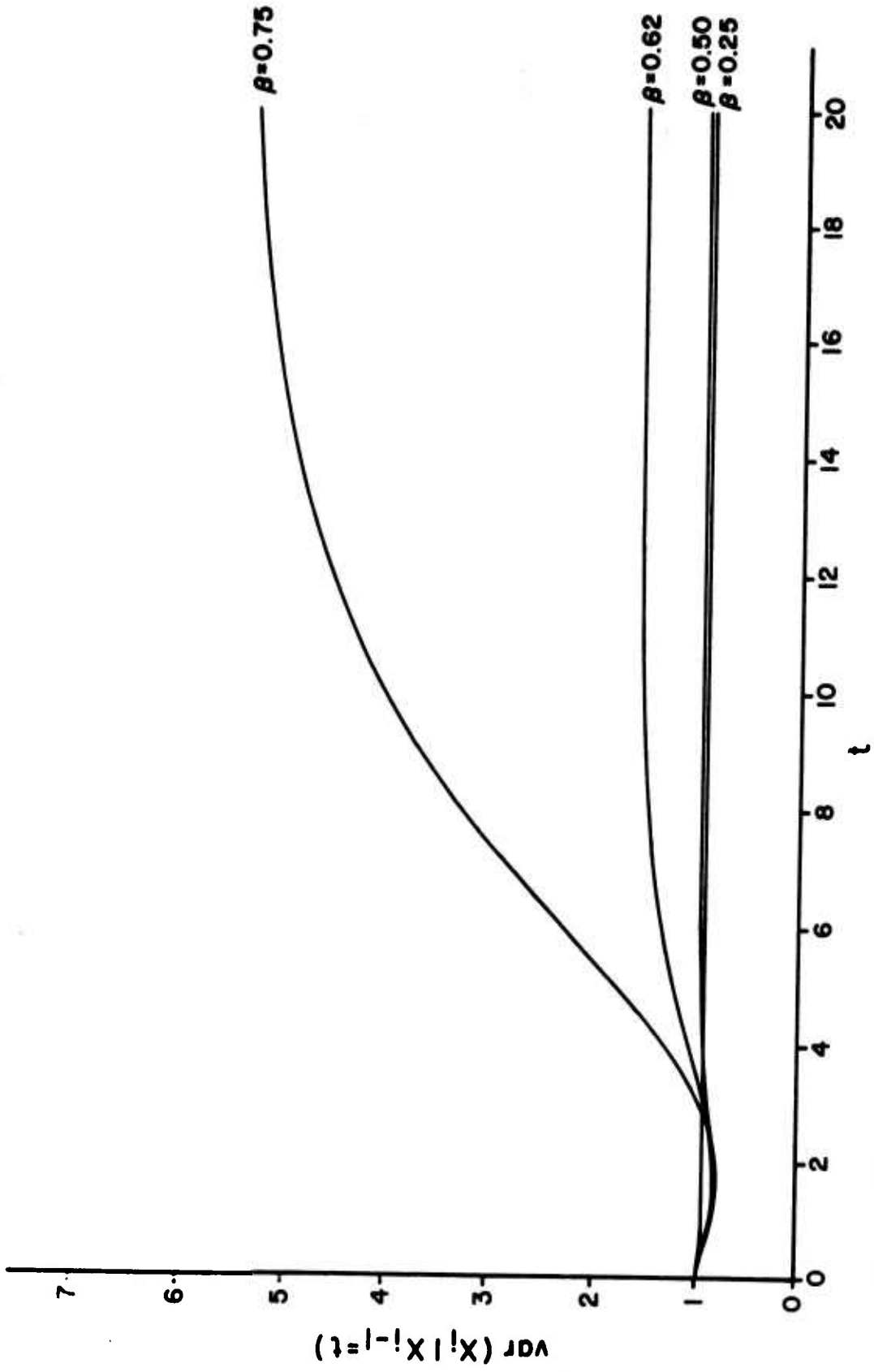


Figure 3

24c

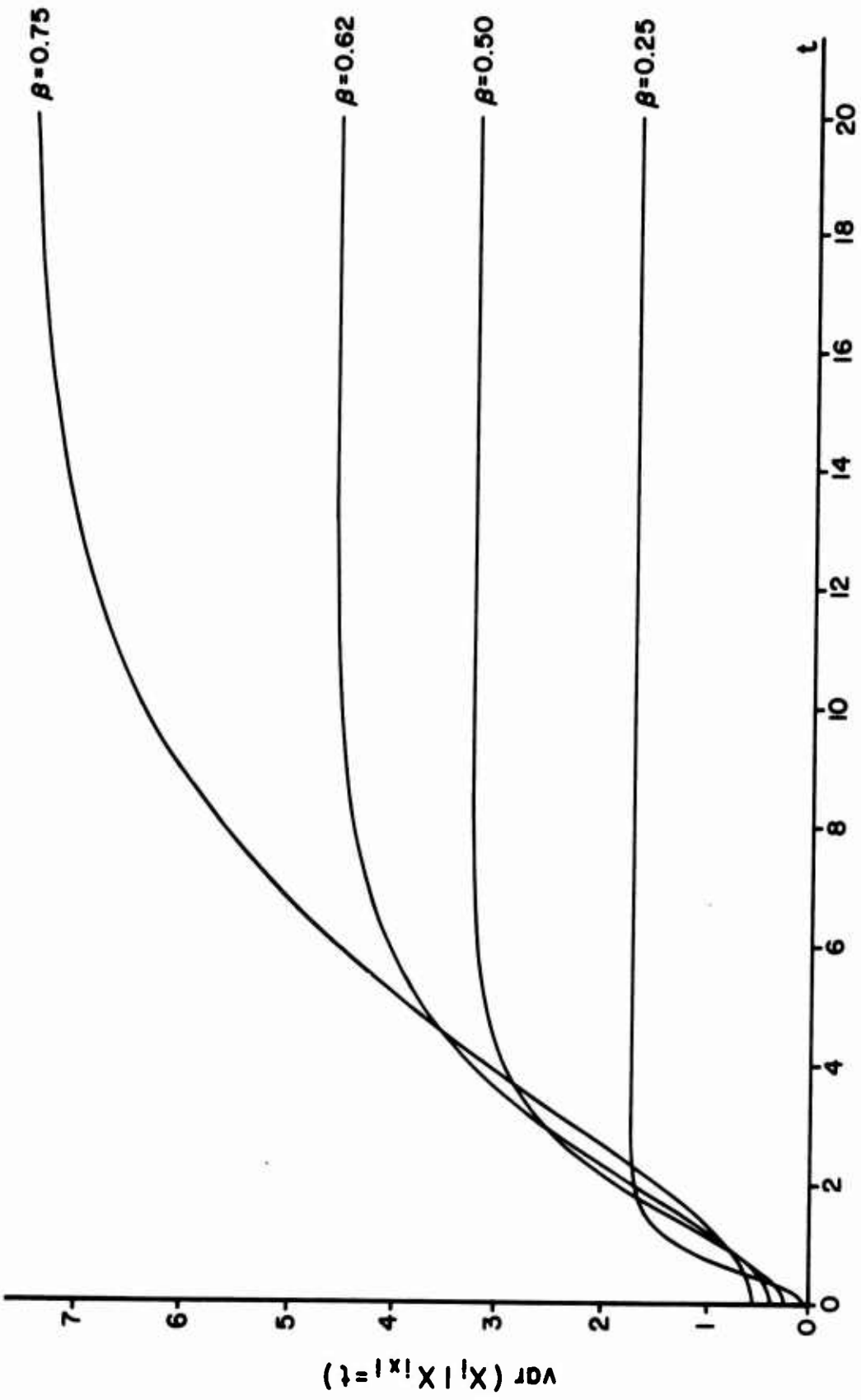


Figure 4

24 d



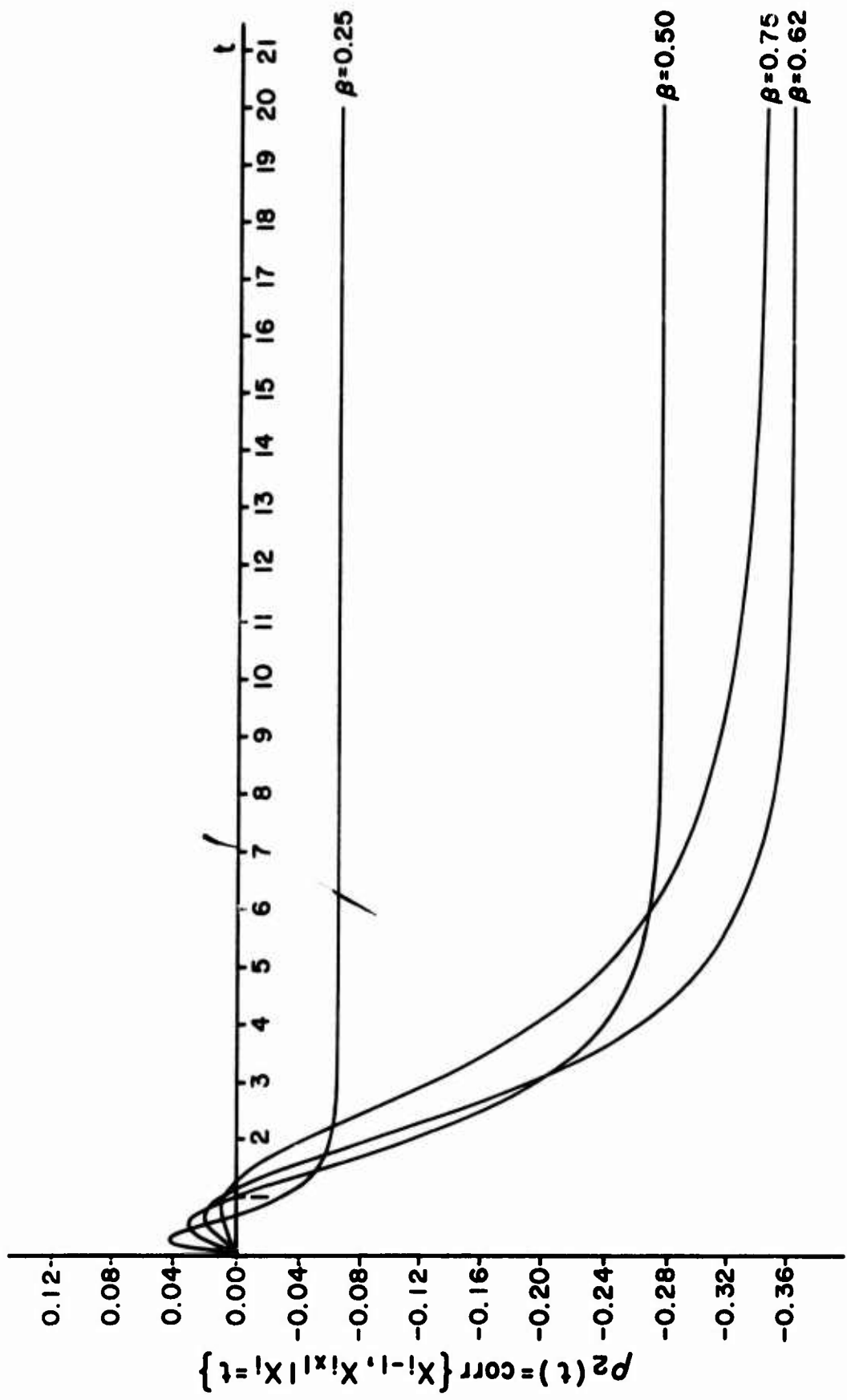


Figure 5

242

## BIBLIOGRAPHY

- Downton, F. (1970). Bivariate exponential distributions of reliability theory. *J. R. Statist. Soc. B* 32, 408-417.
- Cox, D. R. and Lewis, P. A. W. (1966). The Statistical Analysis of Series of Events. Methuen, London and Wiley, New York.
- Gaver, D. P. and Lewis, P. A. W. (1975). First order autoregressive Gamma sequences and point processes. To appear.
- Lawrance, A. J. (1972). Some models for stationary series of univariate events. In Stochastic Point Processes (P. A. W. Lewis, ed.) Wiley, New York, 199-256.
- Lawrance, A. J. (1975). On conditional and partial correlation. To appear.

### Figure Captions

Figure 1. The intensity function  $m_f(t)$  for the EMAl process. The functions is plotted for values  $\beta = 0.1, 0.3, 0.5, 0.7$  and  $0.9$  and  $\lambda = 1$ . The deviation from the constant, Poisson process value  $\lambda = 1$  is small. Unlike the serial correlations for intervals this function does discriminate between the cases  $\beta$  and  $1-\beta$ .

Figure 2. The spectrum of counts  $g_+(w)$  for the EMAl process. The spectrum is flat with value  $1/\pi$  for the Poisson process ( $\beta = 1$  or  $\beta = 0$ ). Unlike the spectrum of intervals it does discriminate between the cases  $\beta$  and  $1-\beta$ .

Figure 3. The conditional variance of  $X_1$ , given  $X_{i-1} = t$ , for the bivariate exponential distribution ( $\lambda = 1$ ) arising in the EMAl process.

Figure 4. The conditional variance of  $X_1$ , given  $X_{i+1} = t$ , for the bivariate exponential distribution ( $\lambda = 1$ ) arising in the EMAl process.

Figure 5. The conditional correlation  $\rho_2(t)$  for intervals  $X_{i-1}$  and  $X_{i+1}$ , given  $X_i = t$ , for the EMAl process. The joint distribution of  $X_{i-1}, X_i, X_{i+1}$  is a trivariate exponential distribution. Again there is differentiation between the cases  $\beta$  and  $(1-\beta)$ .