Some Exponential Alternatives to the Normal Model in Time Series Analysis

by

Eric Staf Hodson
Lieutenant, United States Navy
B.A., University of Colorado, 1968

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 1975

Author ______________________________

Approved by: ______________________________
Thesis Advisor

____
Second Reader

____
Chairman, Department of Operations Research and Systems Analysis

____
Academic Dean

ABSTRACT

An alternative to the normal model in time series analysis is presented wherein the sequence of random variables have exponential and mixed-exponential marginal distributions. The moving-average and autoregressive models are analysed with respect to serial correlations and conditional expectations. The mixed-exponential autoregressive model and the mixed autoregressive-moving-average model with exponential marginals are presented. A method of estimating the serial correlation coefficient is examined, and digital computer simulations of several of the models are given.

TABLE OF CONTENTS

I. INTRODUCTION

Gaver and Lewis (1975) have described a first-order autoregressive stationary sequence of random variables having exponentially distributed marginal distributions, and have given extensions to Gamma distributed processes. Lawrance and Lewis (1975) have described similar models for the moving-average process. Motivation for these models includes finding an alternative to the normal theory of time series analysis, and developing a model with correlated random variables with non-normal marginals.

Several extensions of these models are also of interest, including a mixed-exponential autoregressive process, a second-order moving-average process, the kth-order moving-average process, and the mixed autoregressive moving-average process and the estimation of the parameters concerned. This thesis is concerned with detailing the properties of some of these extensions.

Digital computer simulation of these models is also of interest, particularly in the estimation of parameters.

II. THE EXPONENTIAL AUTOREGRESSIVE MODEL (EAR1)

A basic stochastic model often useful in representing time series is the first-order autoregressive model. Unlike the moving-average process, discussed later, the current value of the first-order autoregressive process is a linear combination of the previous value of the process and an independent error term. Although similar in appearance to the moving-average model in that it, too, is a linear combination of random variables, a principle difference is that each value of the EAR1 sequence, regardless of order, is correlated (depending on the correlation coefficient) to every previous value. It can be written as an infinite moving-average over the sequence of 1.1.d. error terms.

The general form of the kth-order autoregressive model is

$$x_i = \rho_1 x_{i-k} + \rho_2 x_{i-k+1} + \dots + \rho_{k-1} x_i + \varepsilon_i .$$

Only the first-order process will be considered here as it has great potential for describing actual series, and higher-order processes, with the conditions that will be imposed, are difficult to analyse. The first-order process, then, is

$$x_i = \rho x_{i-1} + \varepsilon_i \qquad \text{II.1}$$

$$= \sum_{j=0}^{\infty} \rho^j \varepsilon_{i-j}.$$

If the ε_i are normally distributed random variables, the x_i are a normal sequence, and this is the case which is explicitly or implicitly considered in the literature (see e.g., Anderson, 1971; Box & Jenkins, 1970). In this thesis the condition will be imposed that the $\{x_i\}$ will be a stationary sequence having a marginal distribution that is exponential with parameter λ. A sufficient condition for such an AR sequence to be stationary is that $|\rho| < 1$ (Box & Jenkins, 1970).

The determination of the error term ε_i is relatively simple (Gaver & Lewis, 1975). Beginning with the basic AR1 equation, and assuming that the x_i sequence is stationary,

$$x_i = \rho x_{i-1} + \varepsilon_i,$$

and taking the Laplace transforms

$$E\{e^{-sx_i}\} = E\{e^{-s\rho x_{i-1} - s\varepsilon_i}\}$$

$$= E\{e^{-s\rho x_{i-1}}\}E\{e^{-s\varepsilon_i}\}$$

by the independence of (ε_i, x_{i-1}). Therefore,

$$\phi_{\varepsilon_i}(s) = E\{e^{-s\varepsilon_i}\} = \frac{E\{e^{-sx_i}\}}{E\{e^{-s\rho x_{i-1}}\}} .$$

Now the distribution of $\{x_i\}$ is defined to be exponential (λ), i.e.,

$$E\{e^{-sx_i}\} = \frac{\lambda}{\lambda + s} .$$

Therefore

$$\phi_{\varepsilon_i}(s) = \frac{\phi_{x_i}(s)}{\phi_{x_{i-1}}(\rho s)} = \frac{\frac{\lambda}{\lambda + s}}{\frac{\lambda}{\lambda + \rho s}}$$

$$= \frac{\rho s + \lambda}{s + \lambda} = \rho + (1-\rho) \frac{\lambda}{\lambda + s} .$$

It is simple to see that this is in fact the Laplace transform of a positive random variable, by inverting with respect to s; for $0 \leq \rho < 1$

$$\varepsilon_i = \begin{cases} 0, & \text{probability } \rho \\ \varepsilon_i', & \text{probability } (1-\rho), \end{cases} \qquad \text{II.2.}$$

where ε_i' is an exponential random variable with parameter λ. Therefore

$$x_i = \rho x_{i-1}, \qquad \text{probability } \rho \qquad \text{II.3}$$
$$\rho x_{i-1} + \varepsilon_i' \quad \text{probability } (1-\rho).$$

It is simple to see that it is not possible to allow ρ to be hegative; since x_{i-1} is by definition positive and with ρx_i negative it is heuristically clear that it is impossible to find an independent ε_i which makes x_i positive.

Owing to the nature of the model it will be seen that, unlike a mov-ng average process, there is a partial correlation of all orders. In examining the joint distribution of x_i and x_{i-1}, it is easier to analyse the joint Laplace transform. Thus moving

$$x_i = \rho x_{i-1} + \varepsilon_i$$

$$= \rho^2 x_{i-2} + \rho\varepsilon_{i-1} + \varepsilon_i$$

and

$$x_{i-1} = \rho x_{i-2} + \varepsilon_{i-1}.$$

Therefore

$$E\{e^{-s_1 x_i - s_2 x_{i-1}}\} = E\{e^{-s_1(\rho^2 x_{i-2} + \rho\varepsilon_{i-1} + \varepsilon_i) - s_2(\rho x_{i-2} + \varepsilon_{i-1})}\}$$

$$= E\{e^{-(\rho^2 s_1 + \rho s_2)x_{i-2}}\}E\{e^{-(\rho s_1 + s_2)\varepsilon_{i-1}}\}E\{e^{-s_1\varepsilon_i}\},$$

by the independence of x_{i-2}, ε_{i-1}, ε_i.

By definition

$$E\{e^{-sx_i}\} = \frac{\lambda}{\lambda + s}$$

and by derivation

$$E\{e^{-s\varepsilon_i}\} = \frac{\lambda + s\rho}{\lambda + s} ,$$

so that

$$E\{e^{-s_1 x_i - s_2 x_{i-1}}\} = \frac{\lambda}{\lambda+\rho^2 s_1+\rho s_2} \quad \frac{\lambda+\rho(\rho s_1+s_2)}{\lambda+\rho s_1+s_2} \quad \frac{\lambda+\rho s_1}{\lambda+s_1}$$

$$= \frac{\lambda(\lambda+\rho s_1)}{(\lambda+s_1)(\lambda+\rho s_1+s_2)}$$

One aspect of this equation is that if we set $s_1 = s_2 = s$ we get the Laplace transform of $x_i + x_{i-1}$, and it is not Gamma (2) because of the dependence between x_i and x_{i-1}. It is Gamma (2) if $\rho = 0$, however. If either s_1 or s_2 is set to zero we get the marginal exponential distributions of x_i and x_{i-1}, but the transform is not symmetrical in s_1 and s_2.

The first and second moments of x_i and x_{i-1} can be obtained by taking the first and second partial derivatives with respect to each variable, and they are found to be the same as expected for a marginal exponential distribution, i.e., $E\{x_i\} = E\{x_{i-1}\} = 1/\lambda$; $E\{x_i^2\} = E\{x_{i-1}^2\} = 2/\lambda^2$. The variance is then $1/\lambda^2$.

All that remains to finding the correlation is to find the expected value of the joint distribution. We have

$$E\{x_i x_{i-1}\} = \left.\frac{\partial^2 \phi}{\partial s_1 \partial s_2}\right|_{s_1=0+;\, s_2=0+.}$$

$$= \frac{\rho + 1}{\lambda^2} .$$

Thus, since

$$\text{Cov } (x_i x_{i-1}) = \frac{\rho + 1}{\lambda^2} - \frac{1}{\lambda^2} = \frac{\rho}{\lambda^2}$$

and

$$\text{Corr } (x_i x_{i-1}) = \frac{\text{Cov}(x_i x_{i-1})}{\text{Var}(x_i)\text{Var}(x_{i-1})} = \rho .$$

Note that the first-order serial correlation is bounded by 0 and 1 so that the EAR1 model in some respects is a more versatile model than the EMA1 model discussed later. For the EMA1 model the correlation is positive and bounded above by 1/4.

The other serial correlations may be found in a similar way. All that need be computed is the joint expected value of the lagged values of $\{x\}$. For example, ρ_2 may be determined by finding $E\{x_i x_{i-2}\}$ and replacing it in the above formula, although it is known from the theory of Markov processes that $\rho_k = \rho^k$. It can be shown without much difficulty that

$$E\{e^{-s_1 x_i - s_2 x_{i-k}}\} = \frac{\lambda(\lambda + \rho^k s_1)}{(\lambda + \rho^k s_1 + s_2)(\lambda + s_1)}, \qquad \text{II.4}$$

so that

$$E\{x_i x_{i-k}\} = \frac{\rho^k + 1}{\lambda^2},$$

and hence that

$$\mathrm{Corr}(x_i x_{i-k}) = \rho^k, \qquad (k = 0, 1, 2, \ldots)$$

as surmised above.

Nothing has been said about about conditions for stationarity in the EAR1 sequence. It is simple to check, however, that if the sequence starts with $x_o = {}_o$ as an exponentially distributed random variable, and then continues as defined by II.1 for $i = 1, 2, \ldots$ with the error sequence II.2, then the x_i sequence is stationary.

A. CONDITIONAL EXPECTATIONS OF THE EAR1 SEQUENCE

The conditional means and variances of pairs of random variables in the sequence $\{x_i\}$ may be examined for comparison with those of other bivariate exonential distributions and as a potential method for evaluating the sample coefficient of correlation.

In any process, the conditional expectation (mean) of the x_{i+1} term, given the ith is

$$E\{x_{i+1}|x_i = t\} = \int_{-\infty}^{\infty} x_{i+1} f_{x_{i+1}|x_i}(x_{i+1}|t)\, dx_{i+1}$$

$$= \frac{\int_{-\infty}^{\infty} x_{i+1} f(t, x_{i+1})\, dx_{i+1}}{f_{x_i}(t)} \quad .$$

The joint distribution of x_i and x_{i+1} is not available in a simple form. The joint Laplace transform of this joint distribution has, however, already been established (II.4):

$$\Phi \equiv \phi_{x_i x_{i+1}}(s_1, s_2) = \int_0^{\infty} e^{-s_1 x_i - s_2 x_{i+1}} f_{x_i x_{i+1}}(x_i, x_{i+1})\, dx_i dx_{i+1}$$

Therefore, if the first partial derrivative with respect to the second variable of the joint Laplace is taken,

$$\frac{\partial \Phi}{\partial s_2} = \int\int_0^{\infty} -x_{i+1} e^{-s_1 x_i - s_2 x_{i+1}} f_{x_i x_{i+1}}(x_i x_{i+1})\, dx_i dx_{i+1}$$

and then s_2 set to 0+, we have

$$\frac{\partial \Phi}{\partial s_2}\Big|_{s_2=0+} = \int_0^{\infty}\int -x_{i+1} e^{-s_1 x_i} f_{x_i x_{i+1}}(x_i, x_{i+1})\; dx_i dx_{i+1}$$

and it will be found to be the same form as the negative of the Laplace transform of the numerator of the conditional expectation, $E\{x_{i+1}|x_i = t\}$, identified on the previous page, viz.,

$$\text{numerator} = \int_0^{\infty} x_{i+1} f(x_i, x_{i+1})\; dx_{i+1}$$

$$\text{Laplace transform} = \int_0^{\infty}\int x_{i+1} e^{-s_1 x_i} f_{x_i x_{i+1}}(x_i x_{i+1})\; dx_i dx_{i+1}.$$

Therefore, applying this procedure to the Laplace of the joint distribution Φ, to find $E\{x_{i+1}\; x_i = t\}$, first differentiate with respect to s_2, then set s_2 to zero, invert with respect to s_1, multiply by minus one, and finally divide by the marginal distribution of x_i for the desired result.

Applying this to the EAR1 model yields the following,

$$E\{x_{i+1}\; x_i = t\} = \frac{1}{\lambda(1-\rho)}\;(1 - e^{-\lambda t(1-\rho)/\rho})\;.$$

Nearly the same scheme may be used to find the 'backward' conditional expectation, $E\{x_i|x_{i+1} = t\}$. The procedure here is to take the partial with respect to the first variable, set s_1 to zero, invert with respect to s_2, multiply by minus one, and divide by the marginal of x_{i+1}. This yields the following for the EAR1,

$$E\{x_i | x_{i+1} = t\} = \lambda^{-1}\{\rho\lambda t + (1-\rho)\} .$$

Other conditional expectations may also be found in a similar manner. The conditional expectation of any function $\Psi\colon R_{X_2} \to E^1$ given X_1 is (Zehna, 1970)

$$E\{\Psi(x_2) | x_1\} = \int_{-\infty}^{\infty} \Psi(x_2) f(x_2 | x_1)\, dx_2$$

$$= \frac{\int_{-\infty}^{\infty} \Psi(x_2) f(x_1 x_2)\, dx_2}{f_{x_1}(x_1)} .$$

Again, the Laplace transform of the joint distribution is available

$$\Phi \equiv E\{e^{-s_1 x_i - s_2 x_{i+1}}\}$$

$$\int\int_0^{\infty} e^{-s_1 x_i - s_2 x_{i+1}} f_{x_i x_{i+1}}(x_i x_{i+1})\, dx_i dx_{i+1} .$$

The second partial derivative of Φ with respect to s_2 is

$$\frac{\partial^2 \Phi}{(\partial s_2)^2} = \int_0^{\infty}\!\!\int x_{i+1}^2 \, e^{-s_1 x_i - s_2 x_{i+1}} \, f_{x_i x_{i+1}}(x, x_{i+1}) \, dx_i dx_{i+1}.$$

And setting s_2 to zero yields

$$\left.\frac{\partial^2 \Phi}{(\partial s_2)^2}\right|_{s_2=0+} = \int_0^{\infty}\!\!\int x_{i+1}^2 \, e^{-s_1 x_i} \, f_{x_i x_{i+1}}(x_i x_{i+1}) \, dx_i dx_{i+1}$$

This, of course, is identical to the Laplace transform of the numerator of $E\{x_{i+1}^2 | x_i\}$ with respect to x_i, $E\{e^{-s_1 x_i}\}$. Applying this procedure in a manner similar to that for finding the conditional means, to find the conditional second moment of x_{i+1} given x_i: i) take the second partial derivative of Φ with respect to the second variable (s_2); ii) set $s_2 = 0+$; iii) invert with respect to s_1; iv) divide by the marginal of x_i. As with the conditional mean, the procedure for finding the conditional second moment in the 'backward' sense, $E\{x_i^2 | x_{i+1}\}$, is the same with the first and second variables being reversed.

For the EAR1 process, then, the two additional second moments are

$$E\{x_{i+1}^2 | x_i = t\} = 2\lambda^{-2} \left[\frac{1}{(1-\rho)^2} - \frac{\lambda t}{(1-\rho)\rho}\right] e^{-\lambda t(1-\rho)/\rho}$$

and

$$E\{x_i^2|x_{i+1} = t\} = t^2(2\rho^2-\rho) + 4t\lambda^{-1}(\rho-\rho^2) + 2\lambda^{-2}(1-\rho) .$$

These conditional moments (and the variance which follow directly) present a contrast to those of the MA1 process (presented later) in which all forms are exponential. Also, allowing ρ to approach zero or one produces different results, namely that as ρ approaches one (i.e., completely correlated events) the distribution becomes degenerate and all members of the sequence are identical to the first. If ρ goes to zero, it becomes an ordinary Poisson process. Both ideas are summarized in the table below. A partial check on the accuracy of the above is made by first multiplying by the conditioning marginal and then integrating. The result is found to be the remaining marginal's expected value.

Term	$\lim_{\rho\to 0}$ {term}	$\lim_{\rho\to 1}$ {term}
$E\{x_{i+1}\|x_i = t\}$	$1/\lambda$	t
$E\{x_i\|x_{i+1} = t\}$	$1/\lambda$	t
$E\{x_{i+1}^2\|x_i = t\}$	$2/\lambda^2$	t^2
$E\{x_i^2\|x_{i+1} = t\}$	$2/\lambda^2$	t^2

An important property of Markov sequences is that while the correlation between variates k apart is non-zero out to infinity, the correlation between x_i and x_{i+k}, given any value of x_j between x_i and x_{i+k}, i.e. $i < j < i+k$ is zero. The zero partial correlation is a means of identifying a Markov sequence.

III. FIRST ORDER AUTOREGRESSIVE MIXED EXPONENTIAL PROCESS

In this section, the basic autoregressive model will be used, but rather than demanding the marginal distribution of the $\{x_i\}$ sequence be a simple exponential (λ), a marginal distribution of the three-parameter mixed exponential $(\lambda_1, \lambda_2, \pi)$ will be imposed. It will be shown that for certain values of λ_1, λ_2 and π, an ε_i sequence exists which gives this marginal distribution.

This model is far more versatile with respect to fitting observed distributions, particularly very skewed ones. With suitable choices of the parameters, distributions having any desired mean and coefficient of variation between 1 and ∞ may be produced (Cox, 1962).

Using the autoregressive scheme mentioned earlier,

$$x_i = \rho x_{i-1} + \varepsilon_i, \quad 0 \leq \rho < 1, \quad (i = 0, 1, 2, \ldots)$$

and taking the appropriate Laplace transforms,

$$\begin{aligned}\phi_{x_i}(s) = E\{e^{-sx_i}\} &= E\{e^{-\rho x_{i-1}s - \varepsilon_i s}\} \\ &= E\{e^{-\rho x_{i-1}s}\, e^{-\varepsilon_i s}\} \\ &= E\{e^{-\rho x_{i-1}s}\}E\{e^{-\varepsilon_i s}\},\end{aligned}$$

ε_i, x_{i-1} independent.

Therefore

$$\phi_{\varepsilon_i}(s) = E\{e^{-\varepsilon_i s}\} = \frac{\phi_{x_i}(s)}{\phi_{x_{i-1}}(\rho s)} .$$

The probability density function of the three parameter mixed exponential is

$$f(x) = \pi\lambda_1 e^{-\lambda_1 x} + (1-\pi)\lambda_2 e^{-\lambda_2 x} .$$

The Laplace transform of this, and hence the transform of the imposed ma ginal distribution for the x_i's will be

$$\phi_x(s) = \frac{\pi\lambda_1}{\lambda_1 + s} + \frac{(1-\pi)\lambda_2}{\lambda_2 + s}$$

$$= \frac{\lambda_1\lambda_2 + \lambda_1 s + (1-\pi)\lambda_2 s}{(\lambda_1 + s)(\lambda_2 + s)} .$$

Similarly, the transform of ρx_{i-1} will be

$$\phi_{x_{i-1}}(\rho s) = \frac{\lambda_1\lambda_2 + \pi\lambda_1\rho s + (1-\pi)\lambda_2\rho s}{(\lambda_1 + \rho s)(\lambda_2 + \rho s)} .$$

Then

$$\phi_{\varepsilon_i}(s) = \frac{\phi_{x_i}(s)}{\phi_{x_{i-1}}(\rho s)}$$

$$= \frac{\dfrac{\lambda_1\lambda_2 + \pi\lambda_1\rho s + (1-\pi)\lambda_2\rho s}{(\lambda_1 + s)(\lambda_2 + s)}}{\dfrac{\lambda_1\lambda_2 + \pi\lambda_1\rho s + \lambda_2\rho s - \pi\lambda_2\rho s}{(\lambda_1 + \rho s)(\lambda_2 + \rho s)}}$$

$$= \frac{(\lambda_1 + \rho s)(\lambda_2 + \rho s)(\lambda_1\lambda_2 + \pi\lambda_1 s + \lambda_2 s - \pi\lambda_2 s)}{(\lambda_1 + s)(\lambda_2 + s)(\lambda_1\lambda_2 + \pi\lambda_1\rho s + \lambda_2\rho s - \pi\lambda_2\rho s)}$$

$$= \frac{(\lambda_1 + \rho s)(\lambda_2 + \rho s)(\lambda_1\lambda_2 + s(\pi\lambda_1 + \lambda_2 - \pi\lambda_2))}{(\lambda_1 + s)(\lambda_2 + s)(\lambda_1\lambda_2 + s(\pi\lambda_1 + \lambda_2 - \pi\lambda_2))}$$

$$= \frac{(\lambda_1 + \rho s)(\lambda_2 + \rho s)\left(s + \dfrac{\lambda_1\lambda_2}{\pi\lambda_1 + \lambda_2 - \pi\lambda_2}\right)}{(\lambda_1 + s)(\lambda_2 + s)\left(\rho s + \dfrac{\lambda_1\lambda_2}{\pi\lambda_1 + \lambda_2 - \pi\lambda_2}\right)} .$$

Let

$$C \equiv \frac{\lambda_1\lambda_2}{\pi\lambda_1 + \lambda_2 - \pi\lambda_2} = \frac{\lambda_1\lambda_2}{\pi\lambda_1 + (1-\pi)\lambda_2} .$$

Consider C as a function of π. At $\pi = 1$, $C = \lambda_2$, and at $\pi = 0$, $C = \lambda_1$. The derivative of the denominator is $\lambda_1 - \lambda_2$, so that the denominator increases from λ_2 to λ_1

monotonically as π goes from 0 to 1 if $\lambda_1 > \lambda_2$. Therefore C decreases monotonically from λ_1 to λ_2 as π goes from 0 to 1, and also $\lambda_2 \leq C \leq \lambda_1$ and $C > 0$. By extension $\lambda_1 \leq C \leq \lambda_2$ or $\lambda_2 \leq C \leq \lambda_1$, depending on which of the parameters λ_1, λ_2 is larger.

Then

$$\phi_{\varepsilon_i}(s) = \frac{(\lambda_1 + \rho s)(\lambda_2 + \rho\, s)(s + C)}{(\lambda_1 + s)(\lambda_2 + s)(C + \rho s)}$$

$$= \frac{(\rho s + \lambda_1)(\rho s + \lambda_2)(\frac{s}{\rho} + \frac{C}{\rho})}{(\lambda_1 + s)(\lambda_2 + s)(\frac{C}{\rho} + s)}$$

$$= \rho + (1-\rho)(\frac{1}{\rho}) \quad \frac{s\lambda_1\lambda_2+\rho s\lambda_1\lambda_2+\rho s^2\lambda_2+\rho s^2\lambda_1+C\lambda_1\lambda_2-\rho Cs^2}{(\lambda_1 + s)(\lambda_2 + s)(\frac{C}{\rho} + s)}$$

This is not necessarily the Laplace transform of a distribution function; it is shown that this is the Laplace transform of a distribution function for certain values of the three parameters by explicitly inverting it. Churchill (1970) summarizes Heavside's equations in stating

> If f(s) is the quotient p(s)/q(s) of two polynomials such that q(s) has the higher degree and contains the factor <u>s-a</u>, which is not repeated, then the term in F(t)

corresponding to that factor can be written in either of these two forms:

$$\Phi(a)\, e^{at} \quad \text{or} \quad \frac{p(a)}{q'(a)}\, e^{at},$$

where $\Phi(a)$ is the quotient of $p(a)$ divided by the product of all factors of $q(s)$ except $\underline{s-a}$. This Theorem is valid when the constant $\underline{a}$ is any complex number.

Therefore, taking the inverse Laplace transform, term by term, using Heavside's equations and writing it as a probabilistic mixture yields, if $\lambda_1 \neq \lambda_2$, $e\lambda_2 \neq C$ or $\rho\lambda_1 \neq C$,

(term 1) with probability ρ, $\varepsilon_i = 0$

with probability $(1-\rho)$, ε_i has density

$$\text{(term 2)} \quad \left(\frac{1}{\rho}\right) e^{-Ct/\rho} \; \frac{-\frac{C}{\rho}\lambda_1\lambda_2 - \frac{C}{\rho}\lambda_1\lambda_2 + \frac{C}{\rho}\frac{C}{\rho}\lambda_2 + \frac{C}{\rho}\frac{C}{\rho}\lambda_1 + C\lambda_1\lambda_2 - C\frac{C}{\rho}\frac{C}{\rho}}{(\lambda_1 - \frac{C}{\rho})(\lambda_2 - \frac{C}{\rho})}$$

$$\text{(term 3)} \quad +\left(\frac{1}{\rho}\right) e^{-\lambda_1 t} \; \frac{-\lambda_1\lambda_1\lambda_2 - \rho\lambda_1\lambda_1\lambda_2 + \rho\lambda_1\lambda_1\lambda_2 + \rho\lambda_1\lambda_1\lambda_1 + C\lambda_1\lambda_2 - C\lambda_1\lambda_1}{(\lambda_2 - \lambda_1)(\frac{C}{\rho} - \lambda_2)}$$

$$\text{(term 4)} \quad +\left(\frac{1}{\rho}\right) e^{-\lambda_2 t} \; \frac{-\lambda_2\lambda_1\lambda_2 - \rho\lambda_2\lambda_1\lambda_2 + \rho\lambda_2\lambda_2\lambda_2 + \rho\lambda_2\lambda_2\lambda_1 + C\lambda_1\lambda_2 - \rho C\lambda_2\lambda_2}{(\lambda_1 - \lambda_2)(\frac{C}{\rho} - \lambda_2)}$$

This may be rewritten as

$$\varepsilon_i = 0, \quad \text{probability } \rho;$$

Otherwise the density is

$$= \frac{C}{\rho} e^{-Ct/\rho} \frac{(C - \lambda_1)(C - \lambda_2)\rho}{(\rho\lambda_1 - C)(C - \rho\lambda_2)}$$

$$+ \lambda_1 e^{-\lambda_1 t} \frac{(C - \lambda_1)(\rho\lambda_1 - \lambda_2)}{(\lambda_1 - \lambda_2)(C - \rho\lambda_1)}$$

$$+ \lambda_2 e^{-\lambda_2 t} \frac{(C - \lambda_2)(\rho\lambda_2 - \lambda_1)}{(\lambda_1 - \lambda_2)(\rho\lambda_2 - C)}, \quad \text{probability } (1-\rho),$$

where

$$C = \frac{\lambda_1\lambda_2}{\pi\lambda_1 + (1 - \pi)\lambda_2}$$

$\lambda_1, \lambda_2 > 0$, $0 \le \pi \le 1$, $0 \le \rho < 1$, with points of singularity at $\rho\lambda_1 = C$, $\rho\lambda_2 = C$, $\lambda_2 = \lambda_1$.

Again, we are assuming that we have a density function i.e., that the function is positive and its integral equal to one.

For the case $\rho\lambda_1 = C$,

$$\varepsilon_i = 0, \qquad \text{probability } \rho$$

$$= e^{-\lambda_2 t} \frac{\lambda_2(\lambda_2 - \rho\lambda_1)(\rho\lambda_2 - \lambda_1)}{\rho(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_1)}, \quad \text{probability } (1-\rho).$$

For the case $\rho\lambda_2 = C$,

$$\varepsilon_i = 0, \qquad \text{probability } \rho$$

$$= e^{-\lambda_1 t} \frac{\lambda_1(\rho\lambda_2 - \lambda_1)(\rho\lambda_1 - \lambda_2)}{\rho(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_1)}, \quad \text{probability } (1-\rho).$$

For the case $\lambda_1 = \lambda_2$, the necessary conditions for Heavside's equation are not met, and the inverse must be taken after equating λ_1 to λ_2. Recalling from above,

$$\phi_{\varepsilon_i}(s) = \frac{(\lambda_1 + \rho s)(\lambda_2 + \rho s)(\lambda_1\lambda_2 + \pi\lambda_1 s + \lambda_2 s - \pi\lambda_2 s)}{(\lambda_1 + s)(\lambda_2 + s)(\lambda_1\lambda_2 + \pi\lambda_1\rho s + \lambda_2\rho s - \pi\lambda_2\rho s)}.$$

Setting $\lambda_1 = \lambda_2 = \lambda$,

$$= \frac{(\lambda + \rho s)(\lambda + \rho s)(\lambda^2 + \pi\lambda s + \lambda s - \pi\lambda s)}{(\lambda + s)(\lambda + s)(\lambda^2 + \pi\lambda\rho s + \lambda\rho s - \pi\lambda\rho s)}$$

$$= \frac{\lambda + \rho s}{\lambda + s}.$$

Finally inverting with respect to s,

$$\varepsilon_i = 0, \quad \text{probability } \rho$$
$$\lambda e^{-\lambda t}, \quad \text{probability } (1-\rho).$$

Which is, naturally, the same as that for demanding $\{x_i\}$ be distributed exponential, parameter λ, or that $\pi = 0, 1$. The inverse transform, a function of t, may be verified as being a probability density function by showing that it i) integrates over the domain to unity, and ii) it is non-negative over the domain. Establishing the first condition,

$$\int_0^\infty \varepsilon_i \, dt = \rho + (1-\rho) \int_0^\infty e^{-Ct/\rho} \frac{-C(C - \lambda_1)(C - \lambda_2)}{(\rho\lambda_1 - C)(\rho\lambda_2 - C)} dt$$

$$+ (1-\rho) \int_0^\infty e^{-\lambda_1 t} \frac{\lambda_1(C - \lambda_1)(\rho\lambda_1 - \lambda_2)}{(\lambda_1 - \lambda_2)(C - \rho\lambda_1)} dt$$

$$+ (1-\rho) \int_0^\infty e^{-\lambda_2 t} \frac{\lambda_2(C - \lambda_2)(\rho\lambda_2 - \lambda_1)}{(\lambda_1 - \lambda_2)(\rho\lambda_2 - C)} dt$$

$$= \rho + (1-\rho) \frac{-(C - \lambda_1)(C - \lambda_2)}{(\rho\lambda_1 - C)(\rho\lambda_2 - C)} + \frac{(C - \lambda_1)(\rho\lambda_1 - \lambda_2)}{(\lambda_1 - \lambda_2)(C - \rho\lambda_1)}$$

$$+ \frac{(C - \lambda_2)(\rho\lambda_2 - \lambda_1)}{(\lambda_1 - \lambda_2)(\rho\lambda_2 - C)}$$

$$\int_0^\infty \varepsilon_i \, dt = \rho + (1-\rho)[(C-\lambda_1)(C-\lambda_2)(\lambda_1-\lambda_2)+(C-\lambda_1)(\rho\lambda_1-\lambda_2)(\rho\lambda_2-C)$$

$$+ (C-\lambda_2)(\rho\lambda_2-\lambda_1)(C-\rho\lambda_1)] \div (\lambda_1-\lambda_2)(C-\lambda_1)(\rho\lambda_2-C)$$

$$= \rho + (1-\rho)(-C\rho\lambda_1\lambda_2-CC\rho\lambda_2+\rho\lambda_1\lambda_1\lambda_2+C\rho\lambda_1\lambda_2+C\rho\rho\lambda_1\lambda_2$$

$$+ CC\lambda_2-\rho\rho\lambda_1\lambda_1\lambda_2-C\lambda_1\lambda_2+CC\rho\lambda_2-CC\lambda_1+C\rho\lambda_1\lambda_1$$

$$-C\rho\rho\lambda_1\lambda_2-C\rho\lambda_2\lambda_2+C\lambda_1\lambda_2-\rho\lambda_1\lambda_1\lambda_2+\rho\rho\lambda_1\lambda_2\lambda_2)$$

$$\div (\lambda_1 - \lambda_2)(C - \rho\lambda_1)(\rho\lambda_2 - C)$$

$$= \rho + (1-\rho) \frac{(\lambda_1 - \lambda_2)(C - \rho\lambda_1)(\rho\lambda_2 - C)}{(\lambda_1 - \lambda_2)(C - \rho\lambda_1)(\rho\lambda_2 - C)}$$

$$= 1$$

The non-negativity conditions are not as readily established. Analysis is facilitated by forming the following cases:

Case 1	:	$\rho\lambda_2 > C$
Case 2	:	$\rho\lambda_2 < C$
Case A	:	$\rho\lambda_2 > \lambda_1$
Case B	:	$\rho\lambda_2 < \lambda_1$

With the previously stated conditions, $\lambda_1, \lambda_2 > 0$, $0 \leq \rho < 1$, and $\lambda_1 \leq C \leq \lambda_2$, it is clear that case 1A, 2A, and 2B are of concern and 1B is not well-defined.

Beginning with case 2B ($\rho\lambda_2 < C$ and $\rho\lambda_2 < \lambda_1$), it is clear that each of the coefficients of the exponential terms is non-negative and hence, over that parameter range, the function is non-negative.

For case 2A, the first and second terms are non-negative and the third non-positive. If the exponent of the second term is reduced from $-\lambda_1 t$ to $-Ct/\rho$ this will decrease the value of the positive term. Similarly, if the exponent of the third term is increased from $-\lambda_2 t$ to $-Ct/\rho$, the value of this negative term will be decreased. Thus, the resulting function will everywhere be of lesser or equal value to the original form. In terms of equations,

$$\varepsilon_i \geq \varepsilon_i' \equiv e^{-Ct/\rho} \left[\frac{C(C - \lambda_1)(C - \lambda_2)}{(C - \rho\lambda_1)(\rho\lambda_2 - C)} + \frac{\lambda_1(C - \lambda_1)(\rho\lambda_1 - \lambda_2)}{(\lambda_1 - \lambda_2)(C - \rho\lambda_1)} + \frac{\lambda_2(C - \lambda_2)(\rho\lambda_2 - \lambda_1)}{(\lambda_1 - \lambda_2)(\rho\lambda_2 - C)} \right]$$

$$\varepsilon_i \geq \varepsilon_i' = e^{-Ct/\rho} \; (C(C-\lambda_1)(C-\lambda_2)(\lambda_1-\lambda_2)+\lambda_1(C-\lambda_1)(\rho\lambda_1-\lambda_2)(\rho\lambda_2-C)$$

$$+ \lambda_2(C-\lambda_2)(\rho\lambda_2-\lambda_1)(C- \rho\lambda_1))$$

$$\div (\lambda_1 - \lambda_2)(\rho\lambda_2 - C)(C - \rho\lambda_1)$$

$$= e^{-Ct/\rho} \frac{(\lambda_2 - \lambda_1)(C - \rho\lambda_2)(C - \rho\lambda_1)(\lambda_2 + \lambda_1 - C)}{(\lambda_1 - \lambda_2)(\rho\lambda_2 - C)(C - \rho\lambda_1)}$$

$$= e^{-Ct/\rho} (\lambda_2 + \lambda_1 - C) \geq 0$$

The fourth case, $\rho\lambda_2 > C$, cannot explicitly be demonstrated to be everywhere non-negative, although it is so at $t = 0$ and as t approaches infinity. Even without considering this case, the model has great versatility, particularly with small values of ρ.

IV. ESTIMATION OF ρ FOR THE EXPONENTIAL FIRST-ORDER AUTOREGRESSIVE MODEL

Using the exponential AR1 model defined above,

$$x_i \sim \text{Exp}(\lambda) \qquad i = 0, 1, 2, \ldots$$

$$x_i = \rho x_{i-1} + \varepsilon_i'$$

$$= \sum_{j=0}^{\infty} \rho^j \varepsilon_{i-j} ,$$

where ε_i' is 0 with probability ρ, and exponential λ with probability $(1-\rho)$. Now if a new random variable is defined as

$$z_i = \frac{x_i}{x_{i-1}} \qquad i = 1, 2, \ldots$$

then,

$$z_i = \frac{x_{i-1} + \varepsilon_i'}{x_{i-1}} = \rho + \frac{\varepsilon_i'}{x_{i-1}} .$$

Again, if the x_i's are distributed Exp (λ),

$$z_i = \begin{cases} \rho + \dfrac{\varepsilon_i}{x_{i-1}} , & \text{probability } (1-\rho) \\ \rho , & \text{probability } \rho \end{cases}$$

Clearly, if the sample is large enough and/or ρ is large enough, some of the z_i's will "estimate" ρ "exactly."

Now in a sequential estimation situation one only has to observe the minimum of successive z_i's, and as soon as a match is observed between a current value of z_i and the minimum of the previous z_i's, that value is exactly ρ. The time to the first occurrence of the value ρ in the sequence z_i is geometric with parameter ρ and the time to the second match, say M, is the sum of two independent geometric random variables. This is a negative binomial random variable and $E(M) = 2/\rho$.

It is also clear that once ρ is known it is possible to unravel the ε_i' error sequence, and there will be M-2 (a random number) of exponential variates from which to estimate the exponential parameter λ. This can be done in the usual way, and more observations can be taken to achieve any desired precision in the estimate of λ.

For fixed times of observation, say from x_0 to x_i, giving n observed values of z_i, there are three possibilities:

(1) A match occurs in determining the minimum z_i sequence, so that ρ is known exactly, as above. The probability of this is $1 - (1-\rho)^n - n\rho^1(1-\rho)^{n-1}$, which will be very close to one for $\rho \sim 0.5$ and n large.

(2) Either a) for exactly one i, $x_i = \rho x_{i-1}$ (probability $n\rho(1-\rho)^{n-1}$) or

b) for no i do we have $x_i = \rho x_{i-1}$ (probability $(1-\rho)^n$).

In the second case it seems fairly clear that a reasonable estimator of ρ, with a positive bias, will be $\min_{1 \le i \le n} (z_i)$. Note that the ratio ε_i'/x_{i-1} has an F-distribution, $F_{2,2}$, where

$$F_{m,n}(y) = \frac{\left(\frac{n}{m}\right)\left(\frac{m+n}{2}\right) y^{n/2-1}}{\left(\frac{n}{2}\right)\left(\frac{m}{2}\right)(1 + ny/m)^{(n+m)/2}}$$

Since the ratio is independent of λ, so will any estimator of ρ based on the z_i's.

It is possible to get an exact estimator of ρ, with known bias term, by averaging the z_i's. The variance is of order $1/n$, whereas it is conjectured that the variance of the estimator based on the minimum of the z_i's is of order $1/n^2$.

Some simulation studies of the (min z_i) estimator are given in the Appendix. The simulations tend to substantiate the speculation that the distribution of the minimum (z_i) (when it is not an "exact" match of ρ) is exponential (ρ,n), the two parameter exponential distribution with mean = $\rho + 1/n$ and minimum value ρ. (Johnson & Kotz, 1970). It would thus appear that

$$\text{Min}(z_i) = \begin{cases} \rho & \text{, probability } 1 - (1-\rho)^n - n\rho(1-\rho)^{n-1} \\ \text{Exp}(\rho,n) & \text{, probability } (1-\rho)^n + n\rho(1-\rho)^{n-1} \end{cases}$$

The expected value of this is

$$E(\min(z_i)) = E(\rho) = (1-(1-\rho)^n - n\rho(1-\rho)^{n-1})\rho + ((1-\rho)^n + n\rho(1-\rho)^{n-1})(\rho+1/n)$$

with a lower bound of ρ.

V. FIRST-ORDER EXPONENTIAL MOVING AVERAGE PROCESS

The general moving average model of time series analysis (see e.g. Box & Jenkins, 1970) may be described as a sequence of random variables formed as the weighted average of a number of previous observations. It should be noted that it is not a true "average" in that the weighting coefficients do not sum to unity (or any other specified number).

$$x_i = \beta_1 \varepsilon_i + \beta_2 \varepsilon_{i+1} + \ldots + \beta_k \varepsilon_{i+k-1} + \varepsilon_{i+k} \quad (i=0,\pm 1,\pm 2,\ldots),$$
$$(-1 < \beta_i < 1)$$

This is one of the two simplest forms of times series models (the other being the autoregressive model, discussed previously), and hence is relatively straightforward in terms of calculations and correspondingly limited in terms of describing natural phenomena. As the number of terms in the sequence is finite, the sequence may be shown to be stationary (Box & Jenkins, 1970).

Considering, at first, the first-order forward moving average process,

$$x_i = \beta \varepsilon_i + \varepsilon_{i+1} \qquad (i = 0, \pm 1. \pm 2. \ldots; -1 < \beta < 1)$$

it is clear that there is first-order dependence, but nothing further (i.e., x_i may be correlated to x_{i-1}, depending on the correlation parameter, but not to x_{i-2}). This first order dependence is different from the Markovian AR1 model, where x_i and x_{i-2} are conditionally independent. The model is described as "forward" to distinguish it from the "backward" process,

$$x_i = \beta\varepsilon_i + \varepsilon_{i-1}$$

whose properties are similar. Note that if the i.i.d. sequence $\{\varepsilon_i\}$ is normally distributed, x_i will be normally distributed as well.

If the condition is imposed that the marginal distribution of the x_i's be exponential parameter λ, the EAR1 model shows that we may construct a first-order moving average as

$$x_i = \begin{cases} \beta\varepsilon_i\ , & \text{probability } \beta \\ \beta\varepsilon_i + \varepsilon_{i+1}, & \text{probability } (1-\beta) \end{cases}$$

Verifying that the distribution of $\{x_i\}$ is Exponential (λ), the Laplace transform of the distribution is evaluated as

$$L\{f_{x_i}(x_i)\} = f^*_{x_i}(s)$$

$$= E\{e^{-sx_i}\}$$

$$= E\{e^{-s\beta\varepsilon_i}\}\beta + E\{e^{-s\beta\varepsilon_i - s\varepsilon_{i+1}}\}(1-\beta)$$

$$= E\{e^{-s\beta\varepsilon_i}\}\beta + E\{e^{-s\beta\varepsilon_i}\}E\{e^{-s\varepsilon_{i+1}}\}(1-\beta) ,$$

by the independence of the terms of $\{\varepsilon_i\}$. Thus, for any i,

$$f^*_{x_i}(s) = \frac{\beta\lambda}{\lambda + s} + \frac{(1 - \beta)\lambda}{(\lambda + s\beta)(s + \lambda)}$$

$$= \frac{\lambda}{\lambda + s}$$

$$= L\{\lambda e^{-\lambda t}\}, \quad \text{Q.E.D.}$$

The first serial correlation ρ_1 is defined in the conventional manner:

$$\rho_1 = \frac{E\{x_i x_{i+1}\} - E\{x_i\}E\{x_{i+1}\}}{Var\{x_i\}Var\{x_{i+1}\}}$$

The expected value of the joint distribution of $\{x_i x_{i+1}\}$ may be found in a number of ways. Two of these are i) to examine the form produced by the product and ii) to find the first mixed partial derivative of the double Laplace transform of the joint distribution evaluated at s = 0+. Both techniques will be used here as the results will be useful later.

We have

$$x_i = \beta\varepsilon_i, \qquad \text{probability } \beta$$

$$\beta\varepsilon_i + \varepsilon_i + 1 \qquad \text{probability } (1-\beta)$$

and

$$x_{i+1} = \beta\varepsilon_{i+1}, \qquad \text{probability } \beta$$

$$\beta\varepsilon_{i+1} + \varepsilon_{i+2} \qquad \text{probability } (1-\beta)$$

The product, then, is

$$x_i x_{i+1} = \beta^2 \varepsilon_i \varepsilon_{i+1}, \qquad \text{probability } \beta^2$$

$$\beta\varepsilon_i(\beta\varepsilon_{i+1} + \varepsilon_{i+2}), \qquad \text{probability } \beta(1-\beta)$$

$$(\beta\varepsilon_i + \varepsilon_{i+1})\beta\varepsilon_{i+1}, \qquad \text{probability } (1-\beta)\beta$$

$$(\beta\varepsilon_i + \varepsilon_{i+1})(\beta\varepsilon_{i+1} + \varepsilon_{i+2}), \quad \text{probability } (1-\beta)^2$$

$$E\{x_i x_{I+1}\} = \beta^2 E\{\beta^2 \varepsilon_i \varepsilon_{i+1}\}$$

$$+ \beta(1-\beta) E\{\beta^2 \varepsilon_i \varepsilon_{i+1} + \beta \varepsilon_i \varepsilon_{i+2}\}$$

$$+ (1-\beta)\beta E\{\beta^2 \varepsilon_i \varepsilon_{i+1} + \beta \varepsilon_{i+1}^2\}$$

$$+ (1-\beta)^2 E\{\beta^2 \varepsilon_i \varepsilon_{i+1} + \beta \varepsilon_i \varepsilon_{i+2} + \beta \varepsilon_{i+1}^2 + \varepsilon_{i+1} \varepsilon_{i+2}\}$$

Defining $\mu \equiv E\{\varepsilon_i\}$, and $\sigma^2 \equiv Var\{\varepsilon_i\}$, then, by stationarity, $\mu = E\{\varepsilon_{i+1}\}$ and $\sigma^2 = Var\{\varepsilon_{i+1}\}$. Now, $E\{\varepsilon_i \varepsilon_{i+j}\} = E\{\varepsilon_i\} E\{\varepsilon_{i+j}\} = \mu^2$, $j \neq 0$, by the independence of the terms of $\{\varepsilon_i\}$, whereas $E\{\varepsilon_{i+j} \varepsilon_{i+j}\} = \mu^2 + \sigma^2$, by definition. With this in mind,

$$E\{x_i x_{i+1}\} = \beta^2 (\beta^2 \mu^2)$$

$$+ \beta(1-\beta)(\beta^2 \mu^2 + \beta \mu^2)$$

$$+ (1-\beta)\beta(\beta^2 \mu^2 + \beta(\mu^2 + \sigma^2))$$

$$+ (1-\beta)^2 (\beta^2 \mu^2 + \beta \mu^2 + \beta(\mu^2 + \sigma^2) + \mu^2)$$

$$= (1-\beta)\sigma^2 + \mu^2.$$

Therefore,

$$\rho_1 = \frac{E\{x_i x_{i+1}\} - E\{x_i\} E\{x_{i+1}\}}{Var\{x_i\} Var\{x_{i+1}\}}$$

$$\rho_1 = \frac{\beta(1-\beta)\sigma^2 + \mu^2 - \mu^2}{\sigma^2\sigma^2}$$

$$= \beta(1-\beta)$$

It should be noted that $\{\varepsilon_i\}$ is distributed exponential parameter λ, so that $\sigma_{\varepsilon_i} = \mu_{\varepsilon_i} = 1/\lambda$, and this could have been substituted earlier. This substitution will be useful in considering higher-order moving-average processes. The correlation ρ_1 is the most limiting aspect of the model as the correlation of the x_i's is strictly non-negative and bounded above by 1/4, as compared with the conventional (i.e., normal) MA1 model where $-1/2 \leq \rho_1 \leq 1/2$. A second problem associated with this EMA1 model is the symmetry of the functional form about $\beta = 1/2$. For purposes of estimation, differentiation between a coefficient β of, say, .15 and .85 will be impossible if estimat on is based on ρ_1. As will be shown later, higher-order EMA models are similarly limited. By the nature of the model, all higher-order correlations in the EMA1 model are zero.

A. LAPLACE TRANSFORM OF THE JOINT INTERVALS IN THE EXPONENTIAL MOVING AVERAGE MODEL

The basic model is still

$$x_i = \begin{cases} \beta\varepsilon_i, & \text{probability } \beta \quad (0 \leq \beta \leq 1, \\ \beta\varepsilon_i + \varepsilon_{i+1}, & \text{probability } (1-\beta) \quad i=0,\pm1,\pm2,\ldots) \end{cases}$$

and the product random variable is

		probability
$x_i x_{i+1}$	$= \beta\varepsilon_i \beta\varepsilon_{i+1}$	β^2
	$\beta\varepsilon_i(\beta\varepsilon_{i+1} + \varepsilon_{i+2})$	$\beta(1-\beta)$
	$(\beta\varepsilon_i + \varepsilon_{i+1})\beta\varepsilon_{i+1}$	$(1-\beta)\beta$
	$(\beta\varepsilon_i + \varepsilon_{i+1})(\beta\varepsilon_{i+1} + \varepsilon_{i+2})$	$(1-\beta)^2.$

Taking the double Laplace transform, we have

$$\phi_{x_i,x_{i+1}}(s_1,s_2) = E\{e^{-s_1 x_i - s_2 x_{i+1}}\}$$

$$= \beta^2 E\{e^{-s_1\beta\varepsilon_i - s_2\beta\varepsilon_{i+1}}\}$$

$$+ \beta(1-\beta)E\{e^{-s_1\beta\varepsilon_i - s_2(\beta\varepsilon_{i+1} + \beta_{i+2})}\}$$

$$+ (1-\beta)\beta E\{e^{-s_1(\beta\varepsilon_i + \varepsilon_{i+1}) - s_2\beta\varepsilon_{i+1}}\}$$

$$+ (1-\beta)^2 E\{e^{-s_1(\beta\varepsilon_i + \varepsilon_{i+1}) - s_2(\beta\varepsilon_{i+1} + \varepsilon_{i+2})}\}$$

$$\phi_{x_i, x_{i+1}}(s_1, s_2) = \beta^2 E\{e^{-s_1\beta\varepsilon_i - s_2\beta\varepsilon_{i+1}}\}$$

$$+ \beta(1-\beta)E\{e^{-s_1\beta\varepsilon_i - s_2\beta\varepsilon_{i+1} - s_2\varepsilon_{i+2}}\}$$

$$+ (1-\beta)\beta E\{e^{-s_1\beta\varepsilon_i - (s_1 + s_2\beta)\varepsilon_{i+1}}\}$$

$$+ (1-\beta)^2 E\{e^{-s_1\beta\varepsilon_i - (s_1 + s_2\beta)\varepsilon_{i+1} - s_2\varepsilon_{i+2}}\}$$

$$= \beta^2 E\{e^{-s_1\beta\varepsilon_i}\}E\{e^{-s_2\beta\varepsilon_{i+1}}\}$$

$$+ \beta(1-\beta)E\{e^{-s_1\beta\varepsilon_i}\}E\{e^{-s_2\beta\varepsilon_{i+1}}\}E\{e^{-s_2\varepsilon_{i+2}}\}$$

$$+ (1-\beta)\beta E\{e^{-(s_1 + s_2\beta)\varepsilon_{i+1}}\}E\{e^{-(s_1+s_2\beta)\varepsilon_{i+1}}\}$$

$$+ (1-\beta)^2 E\{e^{-s_1\beta\varepsilon_i}\}E\{e^{-(s_1+s_2\beta)\varepsilon_{i+1}}\}E\{e^{-s_2\varepsilon_{i+2}}\},$$

by the independence of $\{\varepsilon_i\}$. With $\varepsilon_i \sim$ exponential (λ),

$$\phi_{x_i x_{i+1}}(s_1, s_2) = \beta^2 [\frac{\lambda}{\lambda + \beta s_1}] \; [\frac{\lambda}{\lambda + \beta s_2}]$$

$$+ \beta(1-\beta) \; [\frac{\lambda}{\lambda + \beta s_1}] \; [\frac{\lambda}{\lambda + \beta s_2}] \; [\frac{\lambda}{\lambda + s_2}]$$

$$+ (1-\beta)\beta [\frac{\lambda}{\lambda + \beta s_1}] \; [\frac{\lambda}{\lambda + s_1 + \beta s_2}]$$

$$+ (1-\beta)^2 [\frac{\lambda}{\lambda + \beta s_1}] [\frac{\lambda}{\lambda + s_1 + \beta s_2}] [\frac{\lambda}{\lambda + s_2}]$$

$$\phi_{x_i x_{i+1}}(s_1,s_2) = (\beta^2\lambda^4 + \beta^2\lambda^3 s_1 + \beta^3\lambda^3 s_2 + \beta^2\lambda^3 s_2 + \beta^2\lambda^2 s_1 s_2$$

$$+ \beta^3\lambda^2 s_2^{\,2} + \beta\lambda^4 + \beta\lambda^3 s_1 + \beta^2\lambda^3 s_2 - \beta^2\lambda^4 - \beta^2\lambda^3 s_1 + \beta^3\lambda^3 s_2$$

$$+ \beta\lambda^4 + \beta^2\lambda^3 s_2 + \beta\lambda^3 s_2 + \beta^2\lambda^2 s_2^{\,2} - \beta^2\lambda^4 - \beta^3\lambda^3 s_2 - \beta^2\lambda^3$$

$$- \beta^3\lambda^2 s_2^{\,2} + \lambda^4 + \beta\lambda^3 s_2 - \beta\lambda^4 - \beta^2\lambda^3 s_2 - \beta\lambda^4$$

$$- \beta^2\lambda^3 s_2 + \beta^2\lambda^4 + \beta^3\lambda^2 s_2)$$

$$\div ((\lambda + \beta s_1)(\lambda + \beta s_2)(\lambda + s_1 + s_2\beta))$$

$$= \frac{\lambda^4 + \beta\lambda^3 s_1 + \beta^2\lambda^2 s_1 s_2 + \beta\lambda^3 s_1 + \beta\lambda^3 s_2 + \beta^2\lambda^2 s_2^{\,2} + \beta\lambda^3 s_2}{(\lambda + \beta s_1)(\lambda + \beta s_2)(\lambda + s_2)(\lambda + s_1 + s_2\beta)}$$

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

This result may be partially verified by setting either of the s's to zero; this yields the Laplace transform of a single exponential (λ) distribution. Further, if β is set to zero or one, the result is the Laplace transform of the sum of two independent exponentials or an Erlang $(2,\lambda)$.

B. CONDITIONAL EXPECTATIONS OF THE EMA1 SEQUENCE

The conditional means and variances are found in the same manner as the EAR1 model, and only the results are provided:

$$E\{x_{i+1}|x_i = t\} = \lambda^{-1}(\beta\lambda t + \frac{1-2\beta}{1-\beta} + \frac{\beta}{1-\beta} e^{-\lambda t(1-\beta)/\beta})$$

$$E\{x_i|x_{i+1} = t\} = \lambda^{-1}(1 + \beta - e^{-\lambda t(1-\beta)/\beta})$$

$$E\{x_{i+1}^2|x_i = t\} = 2\lambda^{-2}[\frac{1}{2}\beta^2\lambda^2 t + \frac{(1-2\beta)}{1-\beta}\lambda t + \frac{1-3\beta+2\beta^2+\beta^3}{(1-\beta)^2}$$

$$+ \frac{\beta(1-\beta-\beta^2)}{(1-\beta)^2} e^{-(1-\beta)\lambda t/\beta}]$$

$$E\{x_i^2|x_{i+1} = t\} = 2\lambda^{-2}[\frac{1+\beta-\beta^2}{1-\beta} - \frac{1+\beta-\beta^3}{1-\beta} e^{-(1-\beta)\lambda t/\beta}$$

$$-\frac{\lambda t}{\beta} e^{-(1-\beta)\lambda t/\beta}]$$

The conditional variances follow naturally, but none of these expectations provides much assistance in the estimation problem as each of the above expressions is relatively insensitive to changes in β.

C. SERIAL DEPENDENCE AND CONDITIONAL CORRELATION IN THE EMA1 MODEL

The conditional correlation shows some very interesting properties of the EMA1 process. In the EAR1 it was shown that there is some dependence in every term of the sequence with every other term. In the EMA1, the dependence is far more limited. The x_{i-1} term is, depending on the value of the correlation coefficient, correlated with the x_i and x_{i-2} terms, but on no others, and the x_{i+1} term is correlated with the x_i and x_{i+2} terms, but no others. However, the x_i term is correlated with both the x_{i-1} and x_{i+1} terms. Thus, in the EMA1 process, there is no 'transitivity law' with respect to serial correlations in that although each x_{i-1} and x_{i+1} depend on x_i, x_{i-1} and x_{i+1} are independent.

In any event, unlike the AR1 process, the EMA1 is not Markovian as $E\{x_{i+1}|x_i x_{i-1}\}$ is not the same as $E\{x_{i+1}|x_i\}$. The reason for this becomes obvious when considering the basic construction of the model.

It appears, then, that a conditional correlation involving x_{i-1}, x_i, and x_{i+1} might be of interest in examining what this dependence is. Choosing one of the three possibilities, then, of $\mathrm{Corr}\{x_{i-1},x_{i+1}|x_i = t\}$ as a measure, we define this as $\rho_2(t)$,

$$\rho_2(t) = \frac{E\{X_{i-1},X_{i+1}|X_i = t\} - E\{X_{i-1}|X_i = t\}E\{X_{i+1}|X_i = t\}}{\{\mathrm{Var}(X_{i-1}|X_i = t)\ \mathrm{Var}(X_{i+1}|X_i = t)\}^{1/2}}$$

The covariance $\mathrm{Cov}\{x_{i-1}x_{i+1}|x_i = t\}$, the numerator of the above expression, is sufficient for a cursory examination. All but $E\{X_{i-1}X_{i+1}|X_i = t\}$ has been previously calcul ted.

The triple Laplace transform of the joint p.d.f. of x_{i-1}, x_i, x_{i+1} is found in a similar manner to the double transform found earlier,

$$f^{***}_{x_{i-1}, x_i, x_{i+1}}(s_1, s_2, s_3) = E\{e^{-s_1 x_{i-1} - s_2 x_i - s_3 x_{i+1}}\}$$

$$= \iiint_0^\infty e^{-s_1 x_{i-1} - s_2 x_i - s_3 x_{i+1}} f_{x_{i-1} x_i x_{i+1}}(x_{i-1}, x_i, x_{i+1})\, dx_{i-1} dx_i dx_{i+1}.$$

The mixed partial derivatives with respect so s_1 and s_2 are taken of this, and s_1 and s_3 are set to zero. This form is then inverted with respect to s_2 and divided by the marginal of x_i. After subtraction of the product of the conditional means, the conditional covariance remains,

$$\text{Cov}\{X_{i-1} X_{i+1} | X_i = t\} = -\frac{\beta^2}{1-\beta} + \{(1-\beta)\lambda t - \beta\} e^{-(1-\beta)\lambda t/\beta}$$

$$- \frac{\beta}{1-\beta} e^{-2(1-\beta)\lambda t/\beta}$$

Although non-zero, the values of the conditional correlation are small and are of no apparent aid in the estimation of the correlation coefficient.

VI. THE KTH-ORDER EXPONENTIAL MOVING AVERAGE MODEL

The basis for the first-order moving average model of the previous section was the solution for the form of the distribution of components of the double infinite error sequence in the autoregressive model. In that model, however, as applied to the exponential case, the unweighted term in the moving average is just exponential. This suggests making it a moving average of two further ε's, i.e., ε_{i+1} and ε_{i+2}. Using this iterative procedure we get a moving average of any desired order.

In observing the form of the first few exponential models, a very clear pattern of progression is noted:

EMA1	$x_i = \beta\varepsilon_i$	probability β
	$\beta\varepsilon_i + \varepsilon_{i+1}$	probability $(1-\beta)$
EMA2	$x_i = \beta_1\varepsilon_i$	probability β_1
	$\beta_1\varepsilon_i + \beta_2\varepsilon_{i+1}$	probability $(1-\beta_1)\beta_2$
	$\beta_1\varepsilon_i + \beta_2\varepsilon_{i+1} + \varepsilon_{i+2}$	probability $(1-\beta_1)(1-\beta_2)$

EMA3 $\quad x_i = \beta_1 \varepsilon_i$ $\qquad$ probability β_1

$\beta_1 \varepsilon_i + \beta_2 \varepsilon_{i+1}$ $\qquad$ probability $(1-\beta_1)\beta_2$

$\beta_1 \varepsilon_i + \beta_2 \varepsilon_{i+1} + \beta_3 \varepsilon_{i+2}$ $\qquad$ probability $(1-\beta_1)(1-\beta_2)\beta_3$

$\beta_1 \varepsilon_i + \beta_2 \varepsilon_{i+1} + \beta_3 \varepsilon_{i+2} + \varepsilon_{i+3}$ $\qquad$ probability $(1-\beta_1)(1-\beta_2)(1-\beta_3)$

And so forth. These may all be verified as forming probability distributions. Moreover, the procedure is not specific to exponential moving averages; all that is necessary is that for given marginal distribution of the x_i's, the ε_i sequence be that form which provides a solution for the first order autoregressive model.

Now, rather than looking at these as probabilities and associated terms, they may be analysed as terms with associated probabilities, that is, the $\beta_1 \varepsilon_i$ term appears in all models with probability 1. Similarly the $\beta_2 \varepsilon_{i+1}$ term appears in all models ($k \geq 2$) with probability $(1 - \beta_1)$, etc. Complete generalization fails only in describing the final term. The condition must be provided that a coefficient is not present in the final (i.e., k+1st) term of the model.

The use of a series of random indicator functions, $I_i^{(n)}$, permits the complete description of the model with complete generality to all orders in a closed form. Thus, for $n = 1, \ldots, k$,

$$I_i^{(n)} = 0, \quad \text{probability } \beta_n$$
$$= 1, \quad \text{probability } (1-\beta_n),$$

where i refers to the i<u>th</u> term of the n<u>th</u> independent sequence of independent Bernoulli random variables. So, $I_i^{(n)}$ ~ i.i.d. Bernoulli (1, probability $1-\beta_n$, 0 otherwise). Define $I_i^{(0)}$ to be identically 1 and β_{k+1} to be identically 1 for all i.

Using this notation, the coefficient of the ε_i term (for the x_i element) is simply β_1. The second term is $\beta_2 I_i^{(1)} \varepsilon_{i+1}$, the third $\beta_3 I_i^{(1)} \varepsilon_{i+2}$, etc. In closed form, then, the k<u>th</u>-order moving average process is given by

$$x_i = \sum_{j=0}^{k} \beta_{j+1} \varepsilon_{i+j} \prod_{n=0}^{j} I_i^{(n)},$$

where i is the serial number of the i<u>th</u> element of the series, k is the order of the process, and j and n are indices. For example, expanding this for the x_i<u>th</u> element for the EMA3, we get

$$x_i = \beta_1\varepsilon_i + \beta_2 I_i^{(1)}\varepsilon_{i+1} + \beta_3 I_i^{(1)} I_i^{(2)}\varepsilon_{i+2}$$

$$+ I_i^{(1)} I_i^{(2)} I_i^{(3)}\varepsilon_{i+2}$$

The general form above can be established through mathematical induction with the expression for k = 3 and k = k above; the expression for k = k+1 is

$$x_i = \sum_{j=0}^{k} \beta_{j+1}\varepsilon_{i+j} \prod_{n=0}^{j} I_i^{(n)} + \beta_{k+2}\varepsilon_{i+k+1} \prod_{n=0}^{k+1} I_i^{(n)}$$

$$= \sum_{j=0}^{k+1} \beta_{j+1}\varepsilon_{i+j} \prod_{n=0}^{j} I_i^{(n)} .$$

The purpose in the creation of this model is to provide models for data with longer dependence than that obtained with the first-order model and to examine any tendencies of the upper bound on the serial correlations to increase. As mentioned earlier, using the standard formula for serial correlation,

$$\rho_j = \frac{E\{X_i X_{i+j}\} - E\{X_i\}E\{X_{i+j}\}}{\mathrm{Var}\{x_i\}\ \mathrm{Var}\{x_{i+j}\}} ,$$

the only non-zero contributions to this correlation will be a term in the joint expected value that is not present in the product of the expected value of the marginals. In terms of the model, this will occur only where there is a product of two identical error terms, e.g., $E\{\varepsilon_i\varepsilon_{i+j}\} = E\{\varepsilon_i\}E\{\varepsilon_{i+j}\} = \mu^2$, $j \neq 0$, by the independence of $\{\varepsilon_i\}$, whereas $E\{\varepsilon_{i+j}\varepsilon_{i+j}\} = \mu^2 + \sigma^2$, be definition, and in the case where $\{\varepsilon_i\} \sim$ Exponential (λ), $E\{\varepsilon_{i+j}\varepsilon_{i+j}\} = 2/\eta^2 = 2 \cdot E\{\varepsilon_i\}E\{\varepsilon_{i+j}\}$. Therefore, if the term $C\varepsilon_{i+j}\varepsilon_{i+j}$ were to appear, its contribution to the covariance would be $C(\mu^2 + \sigma^2) - C\mu^2 = C\sigma^2$, and the contribution to the correlation would be $C\lambda^2/\lambda^2 = C$. The contribution of $C\varepsilon_i\varepsilon_{i+j}$, however, would be zero.

Thus, for example, ρ_1 for the MA2 process is found as follows:

$$x_i = \beta_1\varepsilon_i + \beta_2 I_i^{(1)}\varepsilon_{i+1} + I_i^{(1)}I_i^{(2)}\varepsilon_{i+2}$$

$$x_{i+1} = \beta_1\varepsilon_i + \beta_2 I_{i+1}^{(1)}\varepsilon_{i+2} + I_{i+1}^{(1)}I_{i+1}^{(2)}\varepsilon_{i+3}$$

This gives

$$\rho_1 = \beta_1\beta_2(1-\beta_1) + \beta_2(1-\beta_1)(1-\beta_1)(1-\beta_2).$$

From this example, it is easy to note the pattern for the serial correlations. This pattern may be condensed for all serial correlations of all orders of the exponential moving average process as

$$\rho_{j,k} = \sum_{i=1}^{k-j+1} \beta_i \prod_{n=0}^{i-1} (1-\beta_n) \beta_{i+j} \prod_{m=0}^{i+j-1} (1-\beta_m), \quad 1 \leq j \leq k$$

where k is the order, j is the degree of serial correlation and i, n, m indices. For a moving-average model, $\rho_{j,k} = 0$ for $j > k$.

Examining a few special cases,

$$\rho_{kk} = \sum_{i=1}^{1} \beta_i \prod_{n=0}^{i-1} (1-\beta_n) \beta_{i+k} \prod_{m=0}^{i+k-1} (1-\beta_m)$$

$$= \beta_1 \beta_{k+1} \prod_{m=0}^{k} (1-\beta_m)$$

$$= \beta_1 \beta_{k+1} (1-\beta_1)(1-\beta_2)\ldots(1-\beta_k)$$

$$= \beta_1 (1-\beta_1)(1-\beta_2)\ldots(1-\beta_k) \; ,$$

as β_{k+1} is defined to be 1.

With $0 \leq \beta \leq 1$, this correlation is clearly limited to $0 \leq \rho_{kk} \leq 1/4$. This may be obtained by setting β_1 to 1/2 and all others to zero. For notational purposes, this maximum is achieved with the beta k-tuple of (.5,0,0,...,0). Any change in the values of the other beta values will cause a decrease in the value of ρ_{kk}.

The first serial correlation coefficients of the first two orders are given by

$$\rho_{1,1} = \sum_{i=0}^{1-1+1} \beta_i \prod_{n=0}^{i-1} (1-\beta_n) \beta_{i+1} \prod_{m=0}^{i+1-1} (1-\beta_m)$$

$$= \beta_1 (1 - \beta_1),$$

as previously noted, and

$$\rho_{1,2} = \sum_{i=1}^{2} \beta_i \prod_{n=0}^{i-1} (1-\beta_n) \beta_{i+1} \prod_{m=0}^{i} (1-\beta_m)$$

$$= \beta_1 \beta_2 (1-\beta_1) \qquad (\leq 1/4)$$

$$+ \beta_2 (1-\beta_1)(1-\beta_1)(1-\beta_2) \quad (\leq 1/4)$$

$$= \beta_2 (1-\beta_1) \}1 - \beta_2 (1-\beta_1)\}$$

Letting $A = \beta_2(1-\beta_1)$, then $\rho_{1,2} = A(1-A)$, with a maximum of 1/4 at $A = 1/2 = \beta_2(1-\beta_1)$.

The second serial correlation coefficients of the second and third order processes are given by

$$\rho_{2,2} = \rho_{k,k}\Big|_{k=2} = \beta_1 (1-\beta_1)(1-\beta_2),$$

which has a maximum value of 1/4 at $\overline{\beta} = (.5,0)$, and

$$\rho_{2,3} = \sum_{i=1}^{2} \beta_i \prod_{n=0}^{i-1} (1-\beta_n) \beta_{i+2} \prod_{m=0}^{i+1} (1-\beta_m)$$

$$= \beta_1 \beta_3 (1-\beta_1)(1-\beta_2)$$

$$+ \beta_2 (1-\beta_1)(1-\beta_1)(1-\beta_2)(1-\beta_3)$$

$$= \beta_3 \{\beta_1 (1-\beta_1)(1-\beta_2)\}$$

$$+ (1-\beta_3)\{\beta_2 (1-\beta_1)(1-\beta_1)(1-\beta_2)\} ,$$

which may be regarded as a convex combination (or weighted average) of the two terms in braces, the mixture being determined by the value of β_3. The term β_3 is not present in the terms in braces. The maximum value of each term in braces is 1/4 (although not attained at the same time). The maximum value for $\rho_{2,3}$ is thus no greater than a weighted average of 1/4 and 1/4, or 1/4. This technique will be used to establish the maximum value of ρ for a number of classes of serial correlations.

One such class of correlations is that of the (k-1)st serial correlation,

$$\rho_{k-1,k} = \sum_{i=1}^{k-(k-1)+1} \beta_i \prod_{n=0}^{i-1} (1-\beta_n) \beta_{i+k-1} \prod_{m=0}^{i+k-1-1} (1-\beta_m) \qquad k \geq 2$$

$$\rho_{k-1,k} = \sum_{i=1}^{2} \beta_i \prod_{n=0}^{i-1} (1-\beta_n)\, \beta_{i+k-1} \prod_{m=0}^{i+k-2} (1-\beta_m)$$

$$= \beta_1\beta_k(1-\beta_1)(1-\beta_2)\ldots(1-\beta_{k-1})$$

$$+ \beta_2(1-\beta_1)(1-\beta_1)(1-\beta_2)\ldots(1-\beta_{k-1})(1-\beta_k)$$

$$= \beta_k\{\beta_1(1-\beta_1)(1-\beta_2)\ldots(1-\beta_{k-1})\}$$

$$+ (1-\beta_k)\{\beta_2(1-\beta_1)(1-\beta_1)(1-\beta_2)\ldots(1-\beta_{k-1})\},$$

$$(k \geq 2)$$

which is also a convex combination of terms each of whose maximum is 1/4. Another class of correlations is that of the (k-2)nd serial correlation,

$$\rho_{k-2,k} = \sum_{i=1}^{k-(k-2)+1} \beta_i \prod_{n=0}^{i-1} (1-\beta_n)\, \beta_{i+k-2} \prod_{m=0}^{i+(k-2)-1} (1-\beta_m) \qquad (k \geq 3)$$

$$= \sum_{i=1}^{3} \beta_i \prod_{n=0}^{i-1} (1-\beta_m)\, \beta_{i+k-2} \prod_{m=0}^{i+k-3} (1-\beta_m), \quad k \geq 4$$

$$\rho_{k-2,k} = \beta_1\beta_{k-1}(1-\beta_1)(1-\beta_2)\ldots(1-\beta_{k-2})$$

$$+ \beta_2(1-\beta_1)\beta_k(1-\beta_1)(1-\beta_2)\ldots(1-\beta_{k-2})(1-\beta_{k-1})$$

$$+ \beta_3(1-\beta_1)(1-\beta_2)(1-\beta_1)(1-\beta_2)\ldots(1-\beta_{k-2})(1-\beta_{k-1})(1-\beta_k)$$

$$= \beta_{k-1}\{\beta_1(1-\beta_1)(1-\beta_2)\ldots(1-\beta_{k-2})\}$$

$$+ (1-\beta_{k-1})\beta_k\{\beta_2(1-\beta_1)(1-\beta_1)(1-\beta_2)\ldots(1-\beta_{k-2})\}$$

$$+ (1-\beta_{k-1})(1-\beta_k)\{\beta_3(1-\beta_1)(1-\beta_2)(1-\beta_1)(1-\beta_2)(1-\beta_3)$$

$$\ldots(1-\beta_{k-2})\}.$$

As above, this is a convex combination of terms each of whose maximum is 1/4. In this case, there are three terms, but the "weights" still sum to unity. Therefore, this class of serial correlations is also bounded above by 1/4.

A somewhat broader class of serial correlations is that where $2j \geq k+1$. These correlations may also be specified and a fixed upper bound obtained. Using the general expression for correlations in the EMA model above, the individual terms are

$$\rho_{j,k} = \beta_1\beta_{1+j}(1-\beta_1)(1-\beta_2)\ldots(1-\beta_j)$$

$$+ \beta_2(1-\beta_1)\beta_{2+j}(1-\beta_1)(1-\beta_2)\ldots(1-\beta_j)(1-\beta_{j+1})$$

$$+ \beta_3(1-\beta_1)(1-\beta_2)\beta_{3+j}(1-\beta_1)(1-\beta_2)\ldots(1-\beta_{j+1})(1-\beta_{j+2})$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$+ \beta_{k-j}(1-\beta_1)\ldots(1-\beta_{k-j-1})\beta_k(1-\beta_1)\ldots(1-\beta_j)\ldots(1-\beta_{k-1})$$

$$+ \beta_{k-j+1}(1-\beta_1)\ldots(1-\beta_{k-j})(1-\beta_1)\ldots(1-\beta_j)\ldots(1-\beta_{k-1})(1-\beta_k).$$

With the added restriction imposed that $2j \geq k + 1$, the above may be factored into another convex combination form:

$$\rho_{j,k} = \beta_{1+j}\{\beta_1(1-\beta_1)(1-\beta_2)\ldots(1-\beta_j)\}$$

$$+ (1-\beta_{1+j})\beta_{2+j}\{\beta_2(1-\beta_1)(1-\beta_2)\ldots(1-\beta_j)\}$$

$$+ (1-\beta_{1+j})(1-\beta_{2+j})\beta_{3+j}\{\beta_3(1-\beta_1)(1-\beta_2)\ldots(1-\beta_j)\}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$+ (1-\beta_{1+j})(1-\beta_{2+j})(1-\beta_{3+j})\ldots(1-\beta_{k-1})\beta_k\{\beta_{k-j}(1-\beta_1)\ldots(1-\beta_j)\}$$

$$\rho_{j,k} = (1-\beta_{1+j})(1-\beta_{2+j})(1-\beta_{3+j})\ldots(1-\beta_{k-1})(1-\beta_k)\{\beta_{k-j+1}(1-\beta_1)\ldots(1-\beta_j)\}$$

This form also has a maximum value of 1/4, being the combination of terms each of whose maximum is 1/4.

Analysis of the remaining serial correlations of the various orders of the EMA models is more difficult. The following observations are made about the remaining correlations and a speculation with respect to the implications of these observations follows. Briefly, it appears as though the maximum serial correlation of any order for any of these models is 1/4.

A. Each serial correlation of each exponential moving average model contains k-j+1 terms, each being the product of various elements of β_i and $(1-\beta_i)$, $(i=1,2,\ldots,k)$.

B. Each of the individual terms is non-negative, bounded above by 1/4 and below by 0.

C. When any one term is maximized (i.e., the $\overline{\beta}$ vector is adjusted to yield 1/4), the remaining terms are each 0.

D. The sum of up to k-j terms can be shown to be bounded by 1/4.

E. When terms are one-by-one maximized, the overall serial correlation may be shown to be a local maximum.

With the above established, it is speculated that the maximum value for any of the serial correlations is 1/4.

This is borne out experimentally for EMA3 and EMA4. The scheme indicated below shows how the k-j+1 maxima for each correlation may be obtained. The general formula for serial correlation in the EMA model, again, is

$$\rho_{j,k} = \sum_{i=1}^{k-j+1} \beta_i \prod_{n=0}^{i-1} (1-\beta_n)\, \beta_{i+j} \prod_{m=0}^{i+j-1} (1-\beta_m).$$

It should be noted in this formula that there are exactly two beta elements in each term and that they are not the same. Further, the β_i element, present in all terms, is matched once and only once by a $(1-\beta_i)$ element from the second product. This is the sole restriction of the maximum value on the individual term, as all other elements are mentioned without their complements.

Using the above observations, it is clear to see that each term is maximized in the following manner:

$$\beta_i = .5$$

$$\beta_1 \rightarrow \beta_{i-1} = 0$$

$$\beta_{i+1} \rightarrow \beta_{i+j-1} = 0$$

$$\beta_{i+j} = 1$$

$\beta_{i+j+1} \rightarrow \beta_k =$ (Arbitrary — not mentioned in the term)

Establishing this beta vector for the i*th* term will drive the remaining terms to zero. The i+j-1 terms except the i*th* each has a β_i element that has been set to zero. All terms i+1 through k-j+1 have a $(1-\beta_{i+j})$ element which is zero, owing to the first product in the general form, since the i in this notation is at least one greater than the i representing the i*th* term that has been maximized.

The following table summarizes the establishment of the maximum serial correlations (each maximum is 1/4).

Serial Correlations

		1	2	3	4	5	6	7	8	9	10
	1	X									
	2	X	X								
EMA	3	E	X	X							
Order	4	E	E	X	X						
	5	S	S	S	X	X					
	6	S	S	S	X	X	X				
	7	S	S	S	X	X	X	X			
	8	S	S	S	S	X	X	X	X		
	9	S	S	S	S	X	X	X	X	X	
	10	S	S	S	S	S	X	X	X	X	X

X - An exact analytic solution obtained for maximum correlation

E - Demonstrated experimentally

S - Speculated

APPENDIX

SIMULATIONS

INTRODUCTION

Chapter 4 addresses the question of estimating the correlation parameter in the EAR1 model. Due to the symmetric nature of the expected correlation, the conventional method of estimating the serial correlation is less than adequate. The following simulations apply the procedure outlined in Chapter 4.

The simulations were performed in APL/360 at the Computer Center at the Naval Postgraduate School. A brief description of the functions used follows.

PROCEDURE

Three simulations were made of the test procedure outlined in Chapter 4. The generated samples each contained 500 EAR1 sequences with mean 1. The first two were generated with a correlation coefficient of .1 and the third with one of .05. Sequence lengths are 10, 5, and 20, respectively.

In each case, a new set of random variables is formed as $z_i = x_{i+1}/x_i$. The minimum of these from each sequence is recorded and the remainder discarded. The probability that this minimum estimates the correlation coefficient "exactly" is given in Chapter 4, and only the distribution of the non-exact estimates are of real interest in the simulation. The "exact" estimates are therefore censored, and the remaining sample examined for exponential tendencies.

Moments of these samples are calculated and the values are plotted as an empirical log survivor curve (Cox & Lewis, 1966). The plotted curves are very nearly linear which is indicative of an underlying exponential distribution. Johnson & Kotz (1970) give as maximum likelihood estimates for the two parameters of a two-parameter exponential, $\hat{\theta} = \text{Min}\ (x_i)$ and $\hat{\sigma} = \overline{X} - \hat{\theta}$. The results are summarized in the table below.

No further tests were conducted on these simulations; however, a fourth simulation was made by generating 500 random exponential deviates with mean of .1. These were censored at the .1 point and were examined as above. The results are nearly identical to the first simulation, adding credibility to the speculation that the distribution of the minimums is the two parameter exponential (n,ρ).

EAR1 Simulation Summary

Simulation number	I	II	III	IV
Sample size	500	500	400	500
Number of terms in each EAR1 sequence	10	5	20	---
Marginal exponential parameter	1	1	1	.1
Correlation coefficient	.1	.1	.05	---
Censored sample size	201	306	159	201
Sample mean	.192	.307	.1	.203
Censoring point	.1	.1	.05	.1
Minimum value (= θ)	.1	.1	.05	.1
ML Estimate of σ (= $\overline{X}-\theta$)	.092	.207	.05	.103
Hypothesized value of σ (=1/number in sequence)	.1	.2	.05	(.1)

USER-DEFINED APL FUNCTIONS

The following functions were employed in the simulations conducted, and listed at the end of the output.

A. n EXVAR l - Exponential random number generator. n is the desired sample size of parameter l exponentials.

B. n BERN k - Bernoulli random number generator. n is the desired sample size of random variates of parameter k.

C. n AR111 parm - EAR1 random sequence generator. n is the desired number of random sequences with parameters parm. parm is a three-element vector of parameters where parm(1) is the parameter for the marginal exponential, parm(2) is the number of terms in the sequence, and parm(3) is the correlation coefficient used to generate the sequence.

D. ZQUOT m - Takes a matrix consisting of a series of sequences and returns a matrix such that $z_{i,j} = x_{i+1,j}/x_{i,j}$.

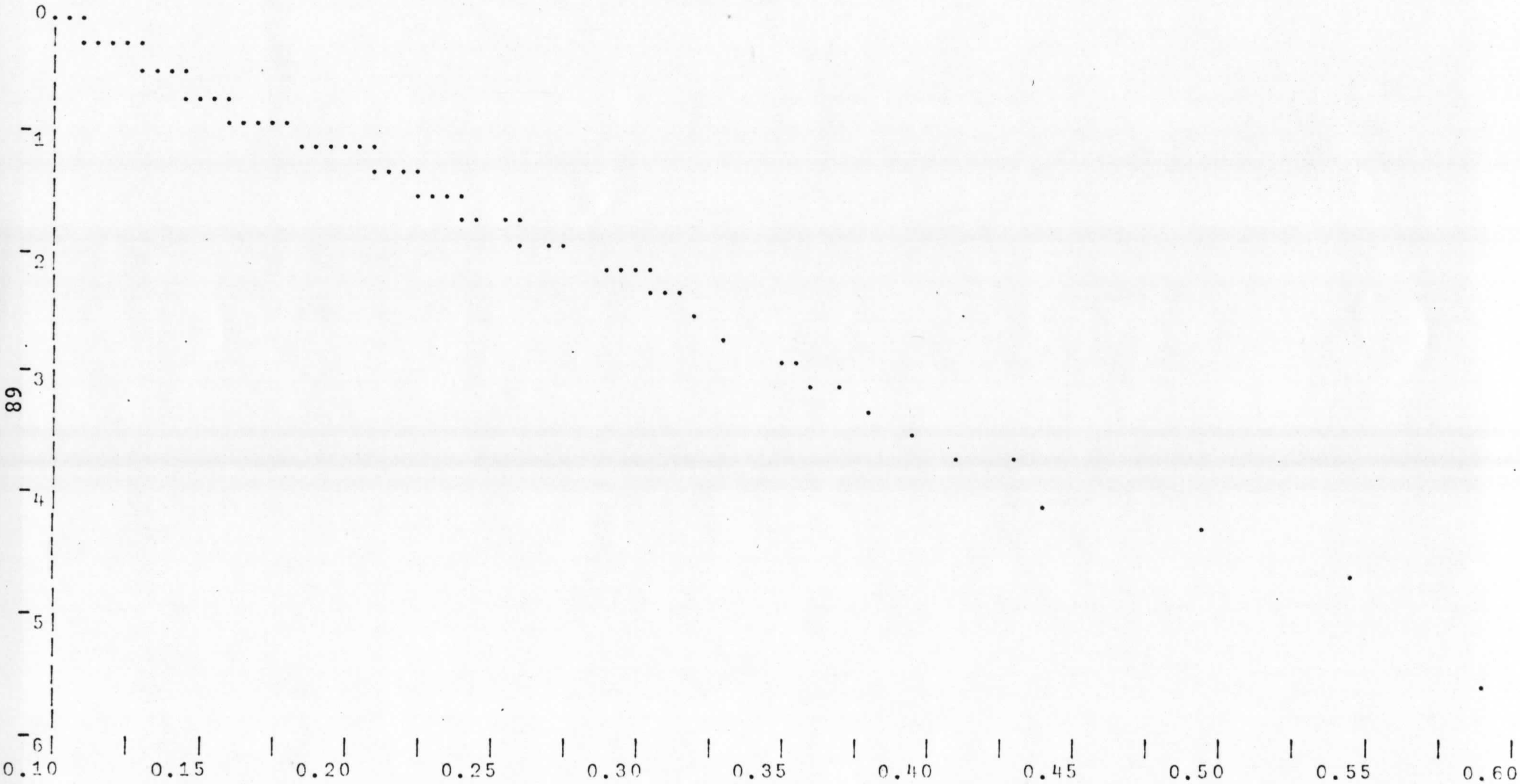
E. LOGSURV n - Generates a vector of length n for use as an axis in plotting the log survivor function. Computes $z = \log\{1-(i/n+1)\}$ for a vector.

The remaining functions used are standard APL/360 library functions and primatives.

```
      Z← 100 AR111 1 10 .1
COMPLETE
      Z←ZQUOT Z
      T←⌊/Z
      Z←100 AR111 1 10 .1
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z
      Z←100 AR111 1 10 .1
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z
      Z←100 AR111 1 10 .1
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z
      Z←100 AR111 1 10 .1
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z


      MOMENTS T

NUMBER OF DATA          :  500
MEAN                    :  0.137
VARIANCE                :  0.00515
STANDARD DEVIATION      :  0.0717
M3                      :  0.000986
M4                      :  0.000302
MEAN DEVIATION          :  0.0495
GEOMETRIC MEAN          :  0.126
HARMONIC MEAN           :  0.119
COEFF OF VARIATION      :  0.523
SKEWNESS                :  2.67
CURTOSIS                :  8.39
BETA 1                  :  0.00098
BETA 2                  :  0.0003
```

```
      +/T=.1
299
      T←T[⍋T]
      T←299↓T

      MOMENTS T

NUMBER OF DATA              :  201
MEAN                        :  0.192
VARIANCE                    :  0.00772
STANDARD DEVIATION          :  0.0878
M3                          :  0.00109
M4                          :  0.000361
MEAN DEVIATION              :  0.0675
GEOMETRIC MEAN              :  0.177
HARMONIC MEAN               :  0.164
COEFF OF VARIATION          :  0.457
SKEWNESS                    :  1.61
CURTOSIS                    :  3.07
BETA 1                      :  0.00107
BETA 2                      :  0.000356
```

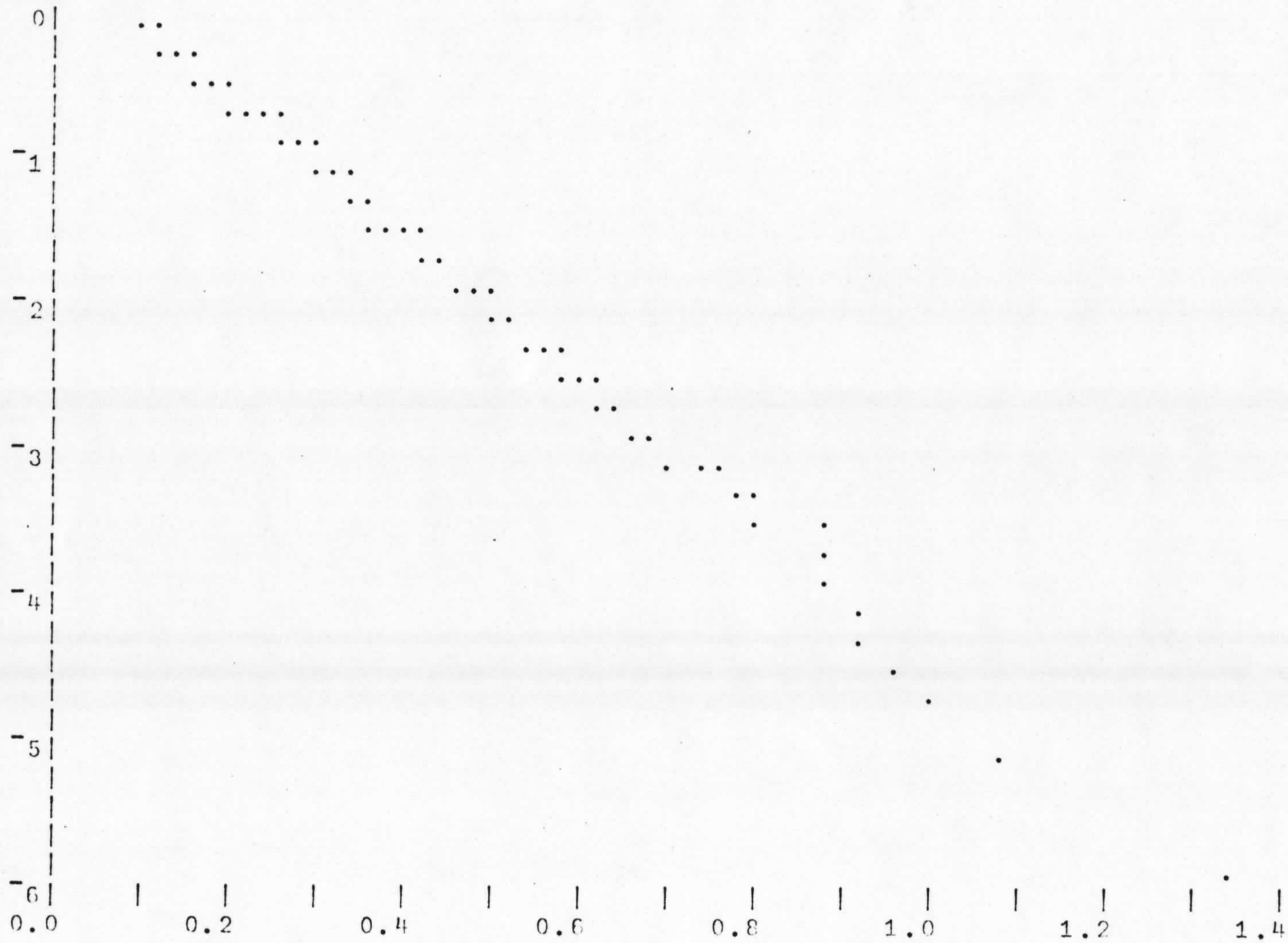
50 100 PLOT T AND LOGSURV 201



```
      Z←100 AR111 1 5 .1
COMPLETE
      Z←ZQUOT Z
      T←⌊/Z
      Z←100 AR111 1 5 .1
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z
      Z←100 AR111 1 5 .1
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z
      Z←100 AR111 1 5 .1
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z
      Z←100 AR111 1 5 .1
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z


      MOMENTS T


NUMBER OF DATA          :  500
MEAN                    :  0.227
VARIANCE                :  0.0349
STANDARD DEVIATION      :  0.187
M3                      :  0.0138
M4                      :  0.0102
MEAN DEVIATION          :  0.139
GEOMETRIC MEAN          :  0.178
HARMONIC MEAN           :  0.15
COEFF OF VARIATION      :  0.824
SKEWNESS                :  2.12
KURTOSIS                :  5.35
BETA 1                  :  0.0137
BETA 2                  :  0.0101
```

```
      +/T=.1
194
      T←T[⍋T]
      T←194↓T

      MOMENTS T

NUMBER OF DATA         :   306
MEAN                   :   0.307
VARIANCE               :   0.0404
STANDARD DEVIATION     :   0.201
M3                     :   0.0136
M4                     :   0.0105
MEAN DEVIATION         :   0.15
GEOMETRIC MEAN         :   0.257
HARMONIC MEAN          :   0.219
COEFF OF VARIATION     :   0.654
SKEWNESS               :   1.68
KURTOSIS               :   3.44
BETA 1                 :   0.0135
BETA 2                 :   0.0104
```

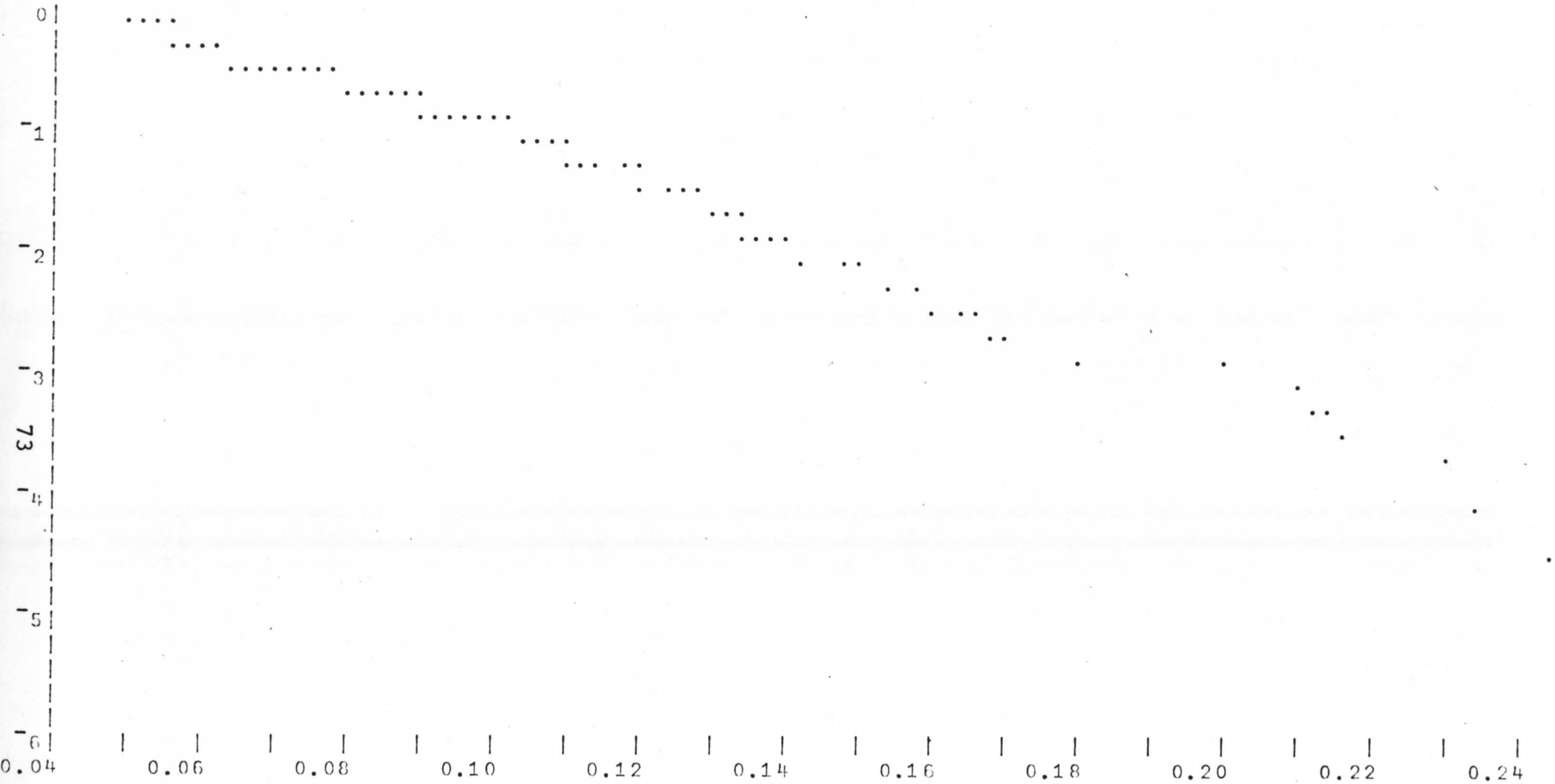
50 100 *PLOT T AND LOGSURV* 306



*

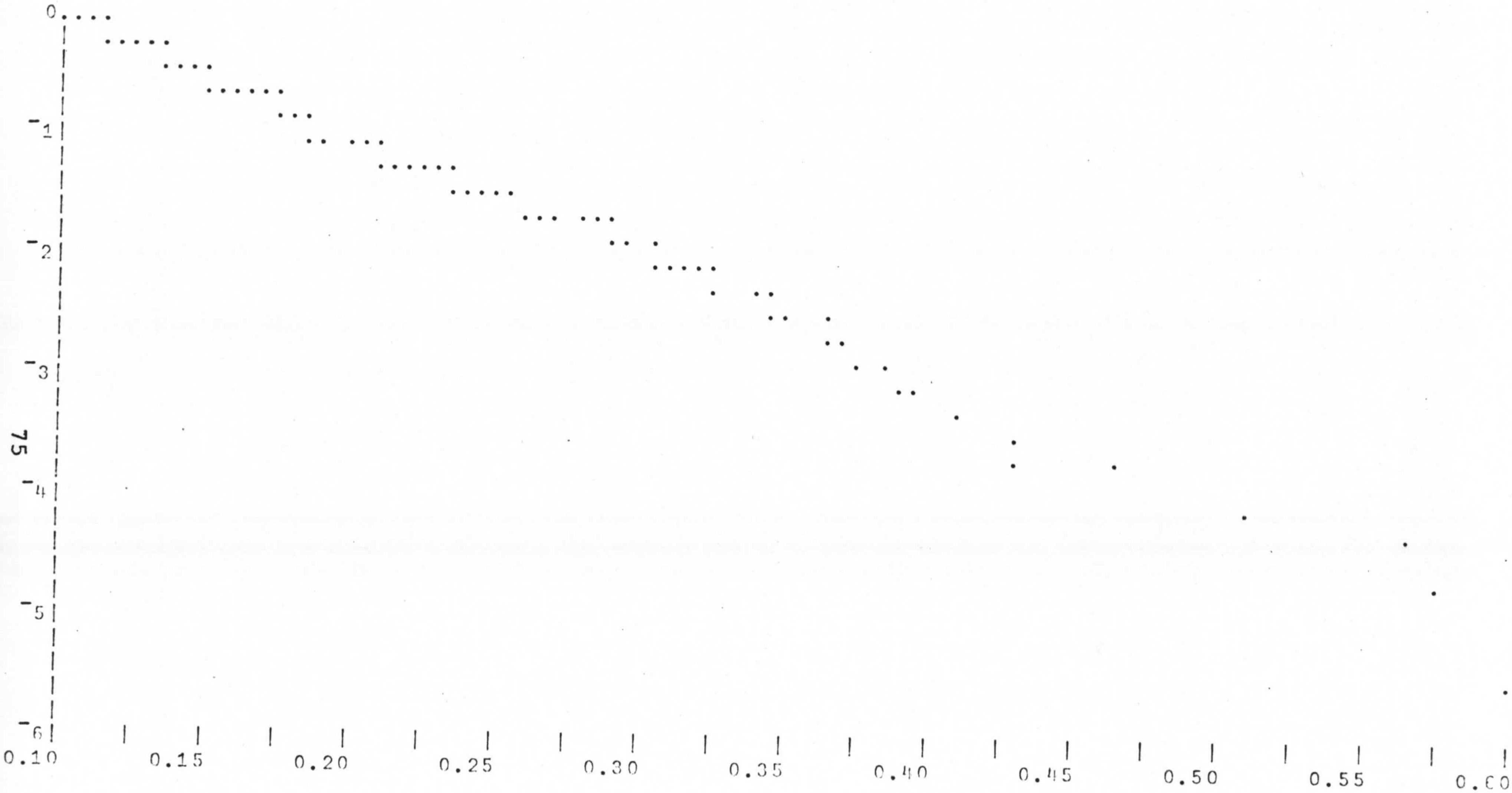
```
      Z←100 AR111 1 20 .05
COMPLETE
      Z←ZQUOT Z
      T←⌊/Z
      Z←100 AR111 1 20 .05
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z
      Z←100 AR111 1 20 .05
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z
      Z←100 AR111 1 20 .05
COMPLETE
      Z←ZQUOT Z
      T←T,⌊/Z

      MOMENTS T


NUMBER OF DATA             :   400
MEAN                       :   0.0699
VARIANCE                   :   0.00138
STANDARD DEVIATION         :   0.0371
M3                         :   0.00012
M4                         :   1.66E¯5
MEAN DEVIATION             :   0.0268
GEOMETRIC MEAN             :   0.0637
HARMONIC MEAN              :   0.0598
COEFF OF VARIATION         :   0.532
SKEWNESS                   :   2.35
CURTOSIS                   :   5.7
BETA 1                     :   0.000119
BETA 2                     :   1.64E¯5

      +/T=.05
241
      T←T[⍋T]
      T←241↓T
      MOMENTS T


NUMBER OF DATA             :   159
MEAN                       :   0.1
VARIANCE                   :   0.00197
STANDARD DEVIATION         :   0.0443
M3                         :   0.00011
M4                         :   1.68E¯5
MEAN DEVIATION             :   0.0345
GEOMETRIC MEAN             :   0.0918
HARMONIC MEAN              :   0.0851
COEFF OF VARIATION         :   0.443
SKEWNESS                   :   1.27
CURTOSIS                   :   1.35
BETA 1                     :   0.000108
BETA 2                     :   1.66E¯5
```

50 100 PLOT T AND LOGSURV 159



```
      TEST←500 EXPVAR .1


      MOMENTS TEST


NUMBER OF DATA            :    500
MEAN                      :    0.108
VARIANCE                  :    0.0104
STANDARD DEVIATION        :    0.102
M3                        :    0.00176
M4                        :    0.000679
MEAN DEVIATION            :    0.0763
GEOMETRIC MEAN            :    0.0634
HARMONIC MEAN             :    0.0191
COEFF OF VARIATION        :    0.939
SKEWNESS                  :    1.67
CURTOSIS                  :    3.34
BETA 1                    :    0.00174
BETA 2                    :    0.000675


      +/TEST≤.1
299

      TEST←TEST[⍋TEST]
      TEST←299↓TEST

      MOMENTS TEST


NUMBER OF DATA            :    201
MEAN                      :    0.203
VARIANCE                  :    0.00945
STANDARD DEVIATION        :    0.0972
M3                        :    0.00146
M4                        :    0.0005
MEAN DEVIATION            :    0.0742
GEOMETRIC MEAN            :    0.185
HARMONIC MEAN             :    0.171
COEFF OF VARIATION        :    0.479
SKEWNESS                  :    1.59
CURTOSIS                  :    2.6
BETA 1                    :    0.00144
BETA 2                    :    0.000493
```

50 100 *PLOT TEST AND LOGSURV* 201



```
      ∇EXPVAR[⎕]∇
    ∇ Z←N EXPVAR L
[1]   Z←-L×⍟(÷100000)×?Nρ100000
    ∇


      ∇BERN[⎕]∇
    ∇ Z←N BERN K
[1]   Z←⌊K+(?Nρ100000)÷100000
    ∇


      ∇AP111[⎕]∇
    ∇ Z←N AP111 PARM
[1]   L←PARM[1]
[2]   MNY←V←N
[3]   N←PARM[2]
[4]   R←PARM[3]
[5]   I←1
[6]   Z←(MNY,N)ρ0.1
[7]   Z[;I]←V EXPVAR L
[8]  L2:→(N<I←I+1)/L1
[9]   Z[;I]←(R×Z[;I-1])+(V EXPVAR L)×V BERN 1-R
[10]  →L2
[11] L1:'COMPLETE'
    ∇


      ∇ZQUOT[⎕]∇
    ∇ Z←ZQUOT M
[1]   Z←(0 1 ↓M)÷ 0 ¯1 ↓M
    ∇


      ∇LOGSURV[⎕]∇
    ∇ Z←LOGSURV N
[1]   Z←⍟1-(⍳N)÷N+1
    ∇
```

```
      ∇MOMENTS[⎕]∇
    ∇ Z←MOMENTS X;N;M1;M2;M3;M4;SD;CV;HM;GM;BETA1;BETA2;SK;CURT;MD;S2;S3;S4;D2;D3;D4
[1]   M1←(+/X)÷N←ρX
[2]   SD←(M2←(D2←((S2←X+.×X)-N×M1×M1))÷N-1)*0.5
[3]   MD←(+/|X-M1)÷N
[4]   HM←0
[5]   →(+/X=0)ρ2+I26
[6]   HM←N÷+/÷X
[7]   GM←10*(+/10⍟X)÷N
[8]   CV←0
[9]   →(M1=0)ρ2+I26
[10]  CV←SD÷|M1
[11]  M3←(D3←(S3←X+.×X×X)-(3×M1×S2)-(2×N×M1*3))×N÷(N-1)×(N-2)
[12]  M4←(((D4←(S4←X+.×X×X×X)-(4×M1×S3)-(6×M1×M1×S2)-3×N×M1*4)×3+(N-2)×N))÷(N-1)×(N-2)×(N-3)
[13]  M4←M4-M2×M2×3×(N-1)×(¯3+2×N)÷N×(N-2)×(N-3)
[14]  SK←M3÷SD×M2
[15]  CURT←¯3+M4÷M2×M2
[16]  BETA1←D3÷N
[17]  BETA2←D4÷N
[18]  →(CC=0)ρ2+I26
[19]  Z←N,M1,M2,SD,M3,M4,MD,GM,HM,CV,SK,CURT,BETA1,BETA2
[20]  →(DD=0)ρ0
[21]  2ρLF
[22]  'NUMBER OF DATA        :  ';N
[23]  'MEAN                  :  ';M1
[24]  'VARIANCE              :  ';M2
[25]  'STANDARD DEVIATION    :  ';SD
[26]  'M3                    :  ';M3
[27]  'M4                    :  ';M4
[28]  'MEAN DEVIATION        :  ';MD
[29]  'GEOMETRIC MEAN        :  ';GM
[30]  'HARMONIC MEAN         :  ';HM
[31]  'COEFF OF VARIATION    :  ';CV
[32]  'SKEWNESS              :  ';SK
[33]  'CURTOSIS              :  ';CURT
[34]  'BETA 1                :  ';BETA1
[35]  'BETA 2                :  ';BETA2
    ∇
```

LIST OF REFERENCES

1. Anderson, T.W., Statistical Analysis of Stationary Time Series, New York: Wiley, 1971.

2. Box, G.E.P., and Jenkins, G.M., Time Series Analysis Forecasting and Control, San Francisco: Holden Day, 1970.

3. Churchill, R.V., Operational Mathematics, Third Edition, New York: McGraw-Hill, 1972.

4. Cox, D.R., Renewal Theory, London: Methuen & Co., 1962.

5. Cox, D.R., and Lewis, P.A.W., The Statistical Analysis of Series of Events, London: Methuen & Co., 1966.

6. Cramer, H., Mathematical Methods of Statistics, Princeton: Princeton University Press, 1946.

7. Cramer, H., The Elements of Probability Theory, Second Edition, New York: Krieger, 1966.

8. Feller, W., An Introduction to Probability Theory and Its Applications, Vol. I, Second Edition, New York: Wiley, 1957.

9. Feller, W., An Introduction to Probability Theory and Its Applications, Vol. II, Second Edition, New York: Wiley, 1971.

10. Gaver, D.P., and Lewis, P.A.W., An Autoregressive Exponential Point Process (EAR1), Naval Postgraduate School, Monterey, 1975.

11. Johnson, N.L., and Kotz, S., Continuous Univariate Distributions -1, New York: Houghton Mifflin, 1970.

12. Lawrance, A.J., and Lewis, P.A.W, A Moving Average Exponential Point Process (EMA1), Naval Postgraduate School, Monterey, 1975.

13. Zehna, P.W., Probability Distributions and Statistics, Boston: Allyn & Bacon, 1970.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
2.	Department Chairman, Code 55 Department of Operations Research and Systems Analysis Naval Postgraduate School Monterey, California 93940	2
3.	Professor P. A. W. Lewis, Code 55 Lw Department of Operations Research and Systems Analysis Naval Postgraduate School Monterey, California 93940	1
4.	LT Eric Staf Hodson, USN 1706 Lorraine Street, Box 37 Colorado Springs, Colorado 80906	1