DOCUMENT RESUME

ED 395 962 TM 025 100

AUTHOR	Longford, Nicholas T.
TITLE	Classes of Multivariate Exponential and Multivariate Geometric Distributions Derived from Markov Processes. Program Statistics Research Technical Report No. 89-87.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RR-89-13
PUB DATE	Apr 89
NOTE	20p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Markov Processes; *Maximum Likelihood Statistics; *Multivariate Analysis; *Statistical Distributions
IDENTIFIERS	Exponents (Mathematics); *Time Series Analysis

ABSTRACT

A class of multivariate exponential distributions is defined as the distributions of occupancy times in upwards skip-free Markov processes in continuous time. These distributions are infinitely divisible, and the multivariate gamma class defined by convolutions and fractions is a substantial generalization of the class defined by N. L. Johnson and S. Kotz (1972). Parallel classes of multivariate geometric and multivariate negative binomial distributions are constructed from occupancy times in "instant" upwards skip-free Markov chains. Maximum likelihood estimation and time series applications are discussed. An appendix demonstrates the density of trivariate exponential distribution. (Contains 17 references.) (Author/SLD)

RR-89-13

Classes of Multivariate Exponential and Multivariate Geometric Distributions Derived from Markov Processes

Nicholas T. Longford

PROGRAM STATISTICS RESEARCH

TECHNICAL REPORT NO. 89-87

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY 08541

CLASSES OF MULTIVARIATE EXPONENTIAL AND MULTIVARIATE GEOMETRIC DISTRIBUTIONS DERIVED FROM MARKOV PROCESSES

Nicholas T. Longford

Program Statistics Research
Technical Report No. 89-87

Research Report No. 89-13

Educational Testing Service
Princeton, New Jersey 08541-0001

April 1989

The Program Statistics Research Technical Report Series is designed to make the working papers of the Research Statistics Group at Educational Testing Service generally available. The series consists of reports by the members of the Research Statistics Group as well as their external and visiting statistical consultants.

TABLE OF CONTENTS

ABSTRACT.

We define a class of multivariate exponential distributions as the distributions of occupancy times in upwards skip-free Markov processes in continuous time. These distributions are infinitely divisible, and the multivariate gamma class defined by convolutions and fractions is a substantial generalization of the class defined by Johnson & Kotz (1972). Parallel classes of multivariate geometric and multivariate negative binomial distributions are constructed from occupancy times in 'instant' upwards skip-free Markov chains. Maximum likelihood estimation and times series applications are discussed.

Key words: multivariate exponential, multivariate geometric, skip-free Markov chain.

1. INTRODUCTION.

The exponential distribution plays a central role in several fields of probability and statistics, and ranks in overall importance next to the normal distribution. While for the normal case we have a well established multivariate normal distribution, in the exponential case the situation is far from clear-cut. A variety of bivariate exponential distributions (**BVE**) have been defined in the past, some of them extendable to higher dimensions and to gamma distributions.

The univariate exponential distribution has a number of important characterizations. Multivariate extensions of some of these characterizations were used for construction of multivariate exponential distributions (**MVE**) by Marshall & Olkin (1967) (lack of memory property) and Paulson (1973) (a stochastic difference equation). Standard transformation techniques were exploited by Kibble (1941) using a χ^2-type derivation, and Moran (1969) using the distribution-function transformation and then the '**-log**' transformation to obtain a **MVE** from the multivariate normal via a multivariate uniform distribution. Gumbel (1960) explored possibilities of defining **BVE** and **MVE** classes based on the form of the joint distribution and density functions. Arnold (1975) constructed nested classes of **BVE**'s by repeated application of geometric compounding. His constructions are equally applicable to multivariate geometric distributions (**MVG**). Wang Zi Kun (1980) derived the distributions of occupancy times (sojourn times) in birth-death processes; these distributions are **MVE**. This brief review of related research is not exhaustive; our research was motivated by these references. We extend the results of Wang Zi Kun to all upwards skip-free Markov processes on $\{0, 1, 2, \ldots\}$, and define a parallel **MVG** class.

In Section 2 we give the definition of occupancy times and related notation, and construct a new class of **MVE** distributions. We will work mainly with moment generating functions (**mgf**), and we derive for them recursive formulae which involve the matrix of transition intensities of the underlying Markov process. Owing to infinite divisibility we have also a class of multivariate gamma distributions (**MVΓ**).

In Section 3 we construct a new class of **MVG** distributions from occupancy times in '**instant**' Markov chains. The constructions in Sections 2 and 3 are completely analogous, and there is a one-to-one correspondence between our **MVE** and **MVG** classes, which is also a one-to-one correspondence between the underlying stochastic processes. This one-to-one correspondence is an extension of the well-known relationship between the **mgf**'s of univariate exponentials and probability generating functions (**pgf**) of geometric distributions:

$$\varphi_\Theta(s) = P_p(s+1) \quad \text{for} \quad \Theta = p^{-1}-1,$$

where φ_Θ is the **mgf** of the exponential with parameter Θ, $\Theta/(\Theta-s)$, and P_p is the **pgf** of the geometric distribution with parameter p, (1-p)/(1-pz).

In Section 4 we discuss maximum likelihood estimation with our **MVE** class and indicate some time series applications. We propose estimation methods based on the **mgf** because it has a much more tractable form than the density function. Our comments in the Section are equally applicable to the **MVG** class, although there the scope of applications is probably limited.

In Section 5 we discuss an alternative definition of an **MVE** class based on the generalization of the χ_2^2 distribution, and state a conjecture that the occupancy times and this generalization define the same class of **MVE**. The support for this conjecture is the equivalence between a pair of subclasses of these distributions, proved by Kent (1982).

2. OCCUPANCY TIMES.

Let $\{Z_t\}_{t\geq 0}$ be an upwards skip-free Markov process on the state space of the non-negative integers $\{0, 1, 2, \ldots\}$ in continuous time $t\geq 0$, given by the matrix of transition intensities

$$Q = \begin{pmatrix} -\nu_0 & \lambda_0 & 0 & \cdot & \cdot & \cdot & & \\ \mu_1 & -\nu_1 & \lambda_1 & 0 & \cdot & \cdot & \cdot & \\ \xi_{2,0} & \mu_2 & -\nu_2 & \lambda_2 & 0 & \cdot & \cdot & \cdot \\ \xi_{3,0} & \xi_{3,1} & \mu_3 & -\nu_3 & \lambda_3 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & \cdot & \cdot & \cdot \end{pmatrix}, \tag{2.1}$$

where $\lambda_i > 0$ $(i\geq 0)$, $\mu_i \geq 0$ $(i\geq 1)$, $\xi_{i,j} \geq 0$ $(i\geq j+2\geq 2)$, and all the row-totals of Q are equal to zero. We denote by Q_n the nxn upper left-hand corner submatrix of Q, corresponding to the states 0, 1, ..., n-1. Let $S_n = \mathbf{diag}\ \{s_0, s_1, \ldots, s_{n-1}\}$ be the diagonal matrix with real numbers s_i on the diagonal; they will be subsequently used as the

arguments of a joint **mgf**

$$\varphi(s_n) = \int \ldots \int \exp(\ 's_n)\ f(x)dx\ , \tag{2.2}$$

where $s_n = (s_0, s_1, \ldots, s_{n-1})$. The index n will be omitted whenever its value is obvious from the context.

The first hitting time from a state k to a state $n \geq k$ is formally defined as

$$\tau_{k,n} = \min\ \{t;\ Z_t = n \mid Z_0 = k\}\ , \tag{2.3}$$

and the vector of occupancy times during passage from k to $n \geq k$ (denoted by $T_{k,n}$) is the decomposition of the first hitting time $\tau_{k,n}$ into the sum of the times that the Markov process Z has spent in each of the states $0, 1, \ldots, n-1$:

$$T_{k,n} = (\tau_{k,n}^{(0)}, \tau_{k,n}^{(1)}, \ldots, \tau_{k,n}^{(n-1)})\ ,$$

where

$$\tau_{k,n}^{(h)} = \int I\ \{Z_t = h \mid Z_0 = k\}\ dt$$

($I\{A\}$ is the indicator function for the event A, and the integral is over the interval $[0, \tau_{k,n}]$).

The joint **mgf** $\varphi_{k,n}(s)$ for the vector of occupancy times during the passage from k to $n \geq k$ can be derived using a backwards equations argument; the derivations below are similar to those for the first hitting times (**fht**) given by Rosenlund (1977). Firstly, owing to the strong Markov property of the process Z we have

$$T_{k,n} = T_{k,k+1} + T_{k+1,k+2} + \ldots + T_{n-1,n}\ , \tag{2.4}$$

with mutually independent summands, and trivially $T_{k,k} = (0, 0, \ldots, 0)$. Correspondingly for the **mgf**'s we have

$$\varphi_{k,n}(s) = \varphi_{k,k+1}(s)\,\varphi_{k+1,k+2}(s)\,\ldots\,\varphi_{n-1,n}(s)\,, \tag{2.5}$$

and $\varphi_{k,k}(s) \equiv 1$. The backwards equations for the 'one-step' **mgf** $\varphi_k{}^+(s) = \varphi_{k,k+1}(s)$ can be expressed in the form

$$\varphi_k{}^+(s) = (\nu_k - s_k)^{-1}\,[\lambda_k + \mu_k\,\varphi_{k-1,k+1}(s) + \xi_{k,k-2}\,\varphi_{k-2,k+1}(s) +$$

$$\ldots + \xi_{k,0}\,\varphi_{0,k+1}(s)\,], \tag{2.6}$$

where $\varphi_0{}^+(s) = \nu_0/(\nu_0 - s_0)$, and $\xi_{1,0} = 0$. The equation (2.6) can be reexpressed as

$$\varphi_k{}^+(s) = \lambda_k/(\nu_k - s_k)\,[1 - \mu_k/(\nu_k - s_k)\;\varphi_{k-1}{}^+(s)$$

$$- \xi_{k,k-2}/(\nu_k - s_k)\,\varphi_{k-2,k}(s) - \ldots - \xi_{k,0}/(\nu_k - s_k)\varphi_{0,k}(s)\,]^{-1}, \tag{2.7}$$

which implies that the occupancy times vector $T_k{}^+ = T_{k,k+1}$ is a convolution of a univariate exponential distribution and a geometric compound distribution. Hence $T_k{}^+$ $(\varphi_k{}^+)$ is infinitely divisible, and owing to (2.5) so are all the occupancy times.

It is easy to show by induction, using (2.5) and (2.7), that $\varphi_k{}^+(s)$ is a ratio of polynomials

$$\varphi_k{}^+(s) = \lambda_k\,R_k(s)\,/\,R_{k+1}(s), \tag{2.8}$$

where

$R_0(s) \equiv 1,$

$R_1(s) = \nu_0 - s_0\,,$

$$R_2(s) = (\nu_0 - s_0)(\nu_1 - s_1) - \lambda_0 \mu_1 ,$$

and generally,

$$R_{k+1}(s) = (\nu_k - s_k)R_k(s) - \mu_k \lambda_{k-1} R_{k-1}(s) - \xi_{k,k-2} \lambda_{k-1} \lambda_{k-2} R_{k-2}(s)$$

$$- \ldots - \xi_{k,1} \lambda_{k-1} \ldots \lambda_2 \lambda_1 R_1(s) - \xi_{k,0} \lambda_{k-1} \ldots \lambda_1 \lambda_0 , \tag{2.9}$$

which is exactly the expansion for **det** $(Q_{k+1}-S_{k+1})$ with respect to the bottom row. Hence

$$R_n(s) = \mathbf{det}\,(Q_n - S_n) \tag{2.10}$$

for $n \geq 1$. As a by-product we have the identity $\mathbf{det}(-Q_n) = \lambda_0 \lambda_1 \ldots \lambda_{n-1}$.

The vectors of occupancy times during passage from 0 to n define our class of n-variate exponential distributions. Their **mgf**'s have the form

$$\varphi_{0,n}(s) = \lambda_0 \lambda_1 \ldots \lambda_{n-1}/R_n(s) , \tag{2.11}$$

where the polynomials R_n, linear in the variables $s_0, s_1, \ldots, s_{n-1}$, are generated recursively by (2.9).

The versions of the identities (2.8) and (2.9) for the birth- death process (all $\xi_{i,j}$ equal to 0) were obtained by Wang Zi Kun (1980).

The bivariate exponential distribution generated by occupancy times during passage from 0 to 2 has the **mgf**

$$\lambda_0 \lambda_1 / R_2(s) = \lambda_0 \lambda_1 / [(\nu_0 - s_0)(\nu_1 - s_1) - \lambda_0 \mu_1] ,$$

and the joint density

$$\lambda_0 \lambda_1 \,\mathbf{exp}\,(-\nu_0 x_0 - \nu_1 x_1)\, L_0(\lambda_0 \mu_h x_0 x_1) \tag{2.12}$$

where

$$L_0(x) = \Sigma_k x^k / (k!)^2 .$$

The distribution (2.12) has been previously defined by Downton (1970), and in a more general context by Kibble (1941).

We define for $h > 0$

$$L_h(x) = \Sigma_k x^k / k!\, \Gamma(k+h+1) , \tag{2.13}$$

which is an analytic version of the Besel function of order h. The bivariate gamma distribution (**BV**Γ) with scale $\sigma > 0$ corresponding to (2.11) has the density

$$(\lambda_0 \lambda_1)^\sigma (x_0 x_1)^{\sigma-1} \mathbf{exp}\, (-\nu_0 x_0 - \nu_1 x_1)\, L_\sigma(\lambda_0 \mu_1 x_0 x_1) . \tag{2.14}$$

The **BVE** distribution (2.11) has the mean $\{(1 + \mu_1/\lambda_1)/\lambda_0,\ 1/\lambda_1\}$ and correlation $\mu_1/\nu_1 \in [0,1)$. No correlation corresponds to independence. The conditional exponential distributions defined from (2.11) by conditioning on x_0 or x_1 have linear regressions:

$$E(X_0 \mid X_1 = x_1) = \lambda_0^{-1}(1 + \mu_1 x_1)$$

and

$$E(X_1 \mid X_0 = x_0) = \nu_1^{-1}(1 + \lambda_0 \mu_1 x_1/\nu_1) .$$

The occupancy times during passage from 0 to $n>2$ in birth-death processes form a conditionally independent sequence and their joint density can be partitioned into a product of conditional exponential densities, see Johnson & Kotz (1972) or Longford (1982). Such a sequence can be used to model an **AR**(1) times series, although the innovation distribution is difficult to describe.

The general trivariate exponential density has the form

$$\lambda_0 \lambda_1 \lambda_2\, \mathbf{exp}\, (-\nu_0 x_0 - \nu_1 x_1 - \nu_2 x_2)$$

$$\times \sum_k L_k(\lambda_0\mu_1 x_1 x_2)\, L_k(\lambda_1\mu_2 x_1 x_2)\, (\lambda_0\lambda_1\xi_{2,0} x_0 x_1 x_2)^k/k! \,. \qquad (2.15)$$

The proof is given in the Appendix. Clearly this and densities for higher dimensions are not suitable for direct maximum likelihood estimation. An alternative approach to **MLE** is discussed in Section 4.

All bivariate marginals of the trivariate exponential (2.13) belong to the **BVE** class. The correlation matrix for the trivariate exponential is

$$\begin{pmatrix} 1 & & \\ 1 - \lambda_1\lambda_2\tau & 1 & \\ 1 - \nu_1\lambda_2\tau & 1 - \lambda_1/\nu_2 & 1 \end{pmatrix}$$

and the means are $\{ 1/(\lambda_0\lambda_1\lambda_2\tau), \nu_2/\lambda_1\lambda_2, 1/\lambda_2 \}$, where $\tau = 1/(\nu_1\nu_2 - \lambda_1\mu_2)$.

3. MULTIVARIATE GEOMETRIC CLASS.

The constructions of the **MVE** class from occupancy times in Markov processes have their obvious analogues for discrete distributions in occupancy times in Markov chains. For example, the birth-death process has its analogue in the Markov chain which allows jumps only one step up or down (discrete random walk, skip-free in both directions). However, the distributions of the occupancy times in such Markov chains have a more complex structure than their birth-death analogues. Rather than give an example we return to this point in the conclusion of this Section.

Let

$$A = \begin{pmatrix} a_0 & u_0 & 0 & \cdot & \cdot & \cdot & & \\ d_1 & a_1 & u_1 & 0 & \cdot & \cdot & \cdot & \\ r_{2,0} & d_2 & a_2 & u_2 & 0 & \cdot & \cdot & \cdot \\ r_{3,0} & r_{3,1} & d_3 & a_3 & u_3 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

be a matrix of transition probabilities of an upwards skip-free Markov chain (all entries non-negative, $u_i > 0$ for all i, row totals equal to 1). In analogy with Section 2 we denote by A_n the nxn upper left-hand corner submatrix of A, and define the first hitting times and occupancy times vectors.

The occupancy time in a state k during passage from 0 to n $(0 \le k \le n)$ is a compound distribution of the individual waiting times in the state which are geometric starting at 1. To obtain a geometric distribution we could either subtract the constant 1 from the occupancy time, or subtract 1 from every waiting time. We choose the latter option, and refer to the underlying Markov process as the **'instant'** Markov chain. In this Markov chain waiting times can be equal to zero with positive probability, and so a sequence of states can be visited within the same time-instant. Our main motivation for this definition of a Markov process is to construct a class of multivariate geometric distributions (**MVG**) with analogous structure to the **MVE** class defined in Section 2.

For the first hitting times and the occupancy times in instant Markov chains given by the probability transition matrix A we use the notation identical to that introduced in Section 2., $\tau_{k,n}$ or τ_k^+, and $T_{k,n}$ or T_k^+, respectively. The vector $\mathbf{z} = (z_0, z_1, z_2, \ldots)$ will be used as the argument in the probability generating functions (**pgf**) $P_{k,n}$ for the occupancy times vectors:

$$P_{k,n}(\mathbf{z}) = E\,(\mathbf{z}^{T_{k,n}}) .$$

The formula (2.5) has a direct analogue in

$$P_{k,n}(\mathbf{z}) = P_k^+(\mathbf{z})\, P_{k+1}^+(\mathbf{z}) \ldots P_{n-1}^+(\mathbf{z}) , \tag{3.1}$$

where $P_h^+(\mathbf{z}) = P_{h,h+1}(\mathbf{z})$, and the backwards equations yield

$$P_k^+(\mathbf{z}) =$$

$$(1 - a_k z_k)^{-1} \, [u_k + d_k P_{k-1,k+1}(\mathbf{z}) + r_{k,k-2} P_{k-2,k+1}(\mathbf{z}) + \ldots + r_{k,0} P_{0,k+1}(\mathbf{z})] \tag{3.2}$$

with the solution in a recursive form

$$P_k^+(\mathbf{z}) =$$

$$u_k/(1 - a_k z_k)\ [1 - d_k/(1 - a_k z_k)\, P_{k-1}^{+}(\mathbf{z}) - r_{k,k-2}/(1 - a_k z_k)\, P_{k-2,k}(\mathbf{z}) +$$

$$\ldots + r_{k,0}/(1 - a_k z_k)\, P_{0,k}(\mathbf{z})]^{-1}, \tag{3.3}$$

$$P_0^{+}(\mathbf{z}) = (1 - a_0)/(1 - a_0 z_0)\,.$$

The formula (3.3) is a convolution of a univariate geometric and a geometric compound distribution. This, together with (3.1), implies infinite divisibility of all occupancy times distributions. We declare the class of all distributions generated by occupancy times during passage from 0 to n as our **MVG** class. This definition can be extended to the class of multivariate negative binomial distributions (**MVNb**) in the obvious way.

The identities (3.2) and (3.3), compared with (2.6) and (2.7) indicate a one-to-one correspondence between occupancy times in continuous and discrete processes. Moreover, we have a one-to-one correspondence between the underlying processes:

If $\quad u_k = \lambda_k/\nu_k \qquad (k \geq 0)$

$\quad d_k = \mu_k/\nu_k \qquad (k \geq 1)$

$\quad r_{k,h} = \xi_{k,h}/\nu_k \qquad (k \geq h+2 \geq 2)$

then

$$\varphi_{k,n}(s) = P_{k,n}(s+\mathbf{1})\ , \tag{3.4}$$

where $\mathbf{1} = (1, 1, \ldots)$. This one-to-one correspondence between the **MVE** and **MVG** classes is the natural extension of the one-to-one correspondence for the univariate exponential and geometric distributions.

In complete analogy with the continuous case we obtain the identity

$$P_{0,n}(\mathbf{z}) = u_0\, u_1\ \ldots\ u_{n-1}\,/\,T_n(\mathbf{z})\,, \tag{3.5}$$

where T_n are polynomials linear in $z_0, z_1, \ldots, z_{n-1}$, generated by the recursive formula

$$T_{n+1}(\mathbf{z}) = (1 - a_n z_n)T_n(\mathbf{z}) - d_n u_{n-1} T_{n-1}(\mathbf{z}) - r_{n,n-2} u_{n-2} u_{n-1} T_{n-2}(\mathbf{z}) -$$

$$\ldots - r_{n,0}\, u_0 u_1\ \ldots\ u_{n-1}\,, \tag{3.6}$$

with $T_0(\mathbf{z}) = 1$ and $T_1(\mathbf{z}) = 1 - a_0 z_0$. It is easy to show by induction that

$$T_n(\mathbf{z}) = \mathbf{det}\,\{I_n - A_n(\mathbf{z})\}\,, \tag{3.7}$$

where $A_n(\mathbf{z})$ is formed from the matrix A_n by replacing its diagonal elements a_k by $a_k z_k$ $(0 \le k < n)$, and is I_n the nxn unit matrix.

For the bivariate and trivariate geometric distributions the joint probabilities and moments (correlations) can be obtained by standard methods, in complete analogy with the exponential case. For higher dimensions the formulae are not tractable.

The backwards equations for the occupancy times in classical Markov chains also define a class of **MVG** distributions (and are infinitely divisible), but the one-to-one correspondence with our **MVE** has a substantially more complex and less natural form. Even the distributions of the first hitting times in Markov chains have a substantially more complex structure than the **fht**'s in continuous time; for details see Kent & Longford (1983).

4. MAXIMUM LIKELIHOOD ESTIMATION AND TIME SERIES APPLICATIONS.

Maximum likelihood estimation for the **BVE** and **BVT** given by the densities (2.12) and (2.14), respectively, can be efficiently carried out by application of standard numerical methods using some well-known recursive formulae for computation of Bessel functions and ratios of Bessel functions.

Since $dL_k(x)/dx = L_{k+1}(x)$, the derivatives of the log-likelihood involving the bivariate densities of the form (2.14) involve ratios of Bessel functions, $L_{k+1}(x)/L_k(x)$. Efficient recursive algorithms for calculation of such ratios were derived by Amos (1974); other useful identities are given in Abramowicz & Stegun (1972). The natural parameter space for the **BVE** is not an open space because of the boundary $\mu_1 = 0$. It is easy to show that the maximum likelihood estimate of μ_1 is positive if and only if the sample covariance $N^{-1} \sum_i x_{0i}x_{1i} - \bar{x}_0\bar{x}_1$ is positive, see Longford (1982) where other numerical details are discussed.

The **MVE** class generated as the occupancy times vectors from birth-death processes have conditionally independent components, and they can be used for modelling of exponential **AR**(1) time series. Since the likelihood for such a time series factors into univariate conditional exponential densities, direct maximum likelihood is feasible.

The form of the density of the general trivariate exponential distribution renders standard maximum likelihood methods impossible, even though the corresponding **mgf**'s have a very

simple structure. Feuerverger and McDonough (1981) have developed procedures for maximum likelihood estimation based on the empirical **mgf** and proved that these procedures can be 'fine-tuned' to arbitrarily high relative efficiency, given some information about the estimated parameters. Their methods appear to be tailor-made for our classes of multivariate distributions (**MVE**, **MVΓ**, **MVG**, and **MVNb**) because they offer a unified approach to estimation in all these classes with generating functions of similar functional form. The main practical point in application of the methods of Feuerverger & McDonough is in determining the number and location of the points in which the **mgf/pgf** would be approximated. These issues could be explored in the special case of **BVE** where direct maximum likelihood estimation is available. It is not clear though to what extent these results could be generalized to **MVE**. Of course, the moment method of estimation is another tractable option, owing to the simple form of the **mgf/pgf**.

The **MVE** class of n-variate distributions (n>2) has the subclass of n independent univariate exponentials and the larger subclass of the distributions with conditionally independent components (**AR**(1), generated from birth-death processes). It appears natural to define a whole set of nested classes of **MVE** distributions by allowing the generating Markov process to have the first 2, 3, ..., n-1 non-zero subdiagonals in the transition intensities matrix Q. If Q has only the first subdiagonal non-zero, we have an **AR**(1) time series. We conjecture that if the first k subdiagonals are non-zero then the resulting **MVE** has an **AR**(k) structure, i.e., it forms a k-step conditionally independent sequence: Z_h and Z_{k+h+1} are conditionally independent, given $Z_{h+1}, Z_{h+2}, \ldots, Z_{k+h}$. Definition of these subclasses of **MVE** imposes a structure upon the entire **MVE** class that could be used for description of the complexity of the correlation structure of an exponential time series or a multivariate sample from **MVE**.

5. MVE AS A GENERALIZED CHI-SQUARE DISTRIBUTION.

Let $\mathbf{X}_1 = (X_{11}, X_{12}, \ldots, X_{1n})$ and $\mathbf{X}_2 = (X_{21}, X_{22}, \ldots, X_{2n})$ be a pair of independent and identically distributed normal random vectors with mean **0** and variance matrix Ω. Then the random vector $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ given by $Y_k = X_{1k}^2 + X_{2k}^2$ defines an n-variate exponential distribution. The original idea for this definition is due to Kibble (1941). We will refer to this derivation as the generalized χ_2^2. It is easy to show that the **mgf** for **Y** is

$$\mathbf{det}\,(1/2\,\Omega^{-1})\,/\,\mathbf{det}\,(1/2\,\Omega^{-1} - S_n)\,,$$

which closely resembles the functional form of our **MVE**, see (2.11) and (2.10). Kent (1982) has in fact proved that the subclass of our **MVE** class arising from birth-death processes coincides

with the subclass of the generalized χ_2^2 distributions derived from variance matrices Ω for which Ω^{-1} is tridiagonal.

An obvious extension of this identity is the following conjecture: The distributions of the **MVE** with **AR**(k) structure (as defined in Section 4) coincide with the generalized χ_2^2 distributions derived from variance matrices Ω such that Ω^{-1} have k non-zero rows below and above the main diagonal. The proof of Kent (1982) cannot be extended for this general proposition, and we do not have an alternative method of proof.

REFERENCES.

Abramowicz, M., and Stegun, I. A., (1972) Handbook of Mathematical Functions. Dover New York.

Amos, D. E., (1974) Computation of modified Bessel functions and their ratios. Math. Comp. 28, 239 - 251.

Arnold, B. C., (1975) Multivariate geometric compounding. Sankhya A 37, 164 - 173.

Block, H. W., (1977) A characterization of a bivariate exponential distribution. Ann. Statist. 5, 808 - 812.

Downton, F., (1970) Bivariate exponential distributions in reliability theory. JRSS B 34, 129 - 131.

Gumbel, E.J., (1960) Bivariate exponential distributions. Biometrika 44, 265 - 268.

Johnson, N. L., and Kotz, S., (1972) Distributions in Statistics: Continuous Multivariate Distributions. Wiley Series in Probability and Mathematical Statistics - Applied John Wiley and Sons, Inc. New York London Sydney Toronto.

Kent, J. T., (1982) The appearance of a multivariate exponential distribution in sojourn times for birth-death and diffusion processes. In Probability, Statistics and Analysis, ed. J.F.C. Kingman and G.E.H. Reuter. Cambridge University Press Cambridge.

Kent, J. T., and Longford, N.T., (1983) An eigenvalue decomposition for first hitting times in random walks. Z. Wahrscheinlichkeitstheorie verw. Gebiete 63, 71 - 84.

Kibble, W. F., (1941) A two-variate gamma-type distribution. Sankhya 5, 137 - 150.

Lawrance A.J., and Lewis, P.A.W., (1985) Modelling and residual analysis of non-linear autoregressive time series in exponential variables. JRSS B 47, 165 - 202.

Longford, N. T., (1982) Hitting times in stochastic processes. Ph.D. thesis, Leeds University, U.K.

Marshall, A. W., and Olkin, I., (1967) A multivariate exponential distribution. JASA 62, 30 - 44.

Moran, P.A.P., (1969) Statistical inference with bivariate gamma distributions. Biometrika 56, 627 - 634.

Paulson, A. S., (1973) A characterization of the exponential distribution and a bivariate exponential distribution. Sankhya A 35, 69 - 78.

Rosenlund, S.I., (1977) Upwards passage times in the non-negative birth-death process. Scand. J. Statist. 4, 90 - 92.

Wang, Zi Kun, (1980) Sojourn times and first passage times for birth and death processes. Sci. Sinica 23, 269 - 279.

APPENDIX

The density of trivariate exponential distribution

The mgf of the trivariate exponential distribution is

$$\frac{\lambda_0\lambda_1\lambda_2}{(\gamma_0 - s_0)(\gamma_1 - s_1)(\gamma_2 - s_2) - \lambda_0\mu_1(\gamma_2 - s_2) - \lambda_1\mu_2(\gamma_0 - s_0) - \xi_{2,0}\lambda_0\lambda_1}$$

$$= \frac{\lambda_0\lambda_1\lambda_2}{(\gamma_0 - s_0)(\gamma_1 - s_1)(\gamma_2 - s_2)} \cdot \sum_{k=0}^{\infty} \frac{1}{(\gamma_1 - s_1)^k} \left[\frac{\lambda_0\mu_1}{\gamma_0-s_0} + \frac{\lambda_1\mu_2}{\gamma_2-s_2} + \frac{\xi_{2,0}\lambda_0\lambda_1}{(\gamma_0-s_0)(\gamma_2-s_2)}\right]^k$$

$$= \lambda_0\lambda_1\lambda_2 \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \sum_{k_3=0}^{\infty} \frac{(k_1+k_2+k_3)!}{k_1!k_2!k_3!} \frac{1}{(\gamma_1-s_1)^{k_1+k_2+k_3+1}} \cdot \frac{1}{(\gamma_0-s_0)^{k_1+k_3+1}}$$

$$\cdot \frac{1}{(\gamma_2-s_2)^{k_2+k_3+1}} \cdot (\lambda_0\mu_1)^{k_1}(\lambda_1\mu_2)^{k_2}(\xi_{2,0}\lambda_0\lambda_1)^{k_3}$$

The summands above are independent Gammas, and the corresponding density is

$$\lambda_0\lambda_1\lambda_2 \exp(-\gamma_0x_0-\gamma_1x_1-\gamma_2x_2) \cdot \sum_{k_1}\sum_{k_2}\sum_{k_3} \frac{1}{k_1!k_2!k_3!(k_1+k_3)!(k_2+k_3)!}$$

$$\cdot (\lambda_0\mu_1x_0x_1)^{k_1} (\lambda_1\mu_2x_1x_2)^{k_2} (\lambda_0\lambda_1\xi_{2,0}x_0x_1x_2)^{k_3}$$

$$= \lambda_0\lambda_1\lambda_2 \exp(-\gamma_0x_0-\gamma_1x_1-\gamma_2x_2) \sum_{k=0}^{\infty} \frac{(\lambda_0\lambda_1\xi_{2,0}x_0x_1x_2)^k}{k!} L_k(\lambda_0\mu_1x_0x_1)L_k(\lambda_1\mu_2x_1x_2)$$

Note that if $\xi_{2,0} = 0$ (conditional independence), this density collaspes to

$$\lambda_0\lambda_1\lambda_2 \exp(-\gamma_0x_0-\gamma_1x_1-\gamma_2x_2) \cdot L_0(\lambda_0\mu_1x_0x_1)\, L_0(\lambda_1\mu_2x_1x_2)$$