STATISTICAL INFERENCE FOR MARKOV RENEWAL PROCESSES

by

Dwight B. Brock

Technical Report No. 103
Department of Statistics ONR Contract

April 15, 1971

Research sponsored by the Office of Naval Research
Contract N00014-68-A-0515
Project NR 042-260

Reproduction in whole or in part is permitted
for any purpose of the United States Government.

This document has been approved for public release
and sale; its distribution is unlimited.

DEPARTMENT OF STATISTICS
Southern Methodist University

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

1. Background

A Markov Renewal Process (M.R.P.) with $m(< \infty)$ states is one which records at each time t the number of times a particular system may visit each of the m states in the time interval $(0,t)$. The system moves from state to state according to a Markov chain of m states with transition probability matrix $P_0 = [p_{ij}]$, and the sojourn time in each state before the next transition is a random variable (r.v.) whose distribution function (d.f.), $F_{ij}(x)$, may, in general, depend on the two states between which the move or transition is being made.

Let Z_t represent the state of the system at time t; that is, $Z_t=j$ for $j = 1, \ldots, m$. The process $\{Z_t; t > 0\}$ is called a semi-Markov process (S.-M.P.). This process was introduced simultaneously and independently by Lévy [21] and Smith [32]. Also at the same time Takács [33] introduced essentially the same kind of stochastic process to study some problems in counter theory. Pyke was the first to use the term Markov Renewal Process in a series of papers: [26] and [27]; [28] and [29] with Schaufele; and [23] with Moore, among others. The present paper is more related to Pyke's work, and his notation will be followed as far as possible.

Markov Renewal Processes and semi-Markov processes have been extensively studied in both theory and applications following the catalytic work of Pyke. They have been used as models in counter theory (see, for example, Takács [33]), queueing theory (by Çinlar [7], among others), and in problems of inventory, reliability, maintenance and others. Papers on the subject now number well above 100 (see the bibliography of Neuts [24] and the expository paper of Çinlar [7]), and no attempt will be made here to present a complete historical account of the development of the theory, but rather to summarize the basic results that will be needed for use in the later chapters.

Before moving ahead with more detail, let us review a few notions about M.R.P.'s and S.-M.P.'s discussed by Pyke ([26] and [27]). First, an S.-M.P. is Markovian only at the instants of transition, so one might think of it as a Markov chain for which the time axis has been randomly transformed. Therefore, an S.-M.P. with $F_{ij}(x)$ degenerate and equal to unity for all i and j is a Markov chain. Also, a semi-Markov process with all $F_{ij}(x)$ exponential and independent of j is a continuous-time countable-state Markov process. A third special case is that of a two-state S.-M.P. with a special P_0 matrix, which is also called an alternating Renewal process (see, for example, Cox [8], p. 81). Finally, a one-state M.R.P. is an ordinary Renewal process, a fact which has received strong mention in Pyke's papers. In fact, in [26] he calls M.R.P.'s a blending or marriage of Markov chains and Renewal processes. Some of these similarities will be brought out, and a few comparisons will be made in the ensuing discussion.

2. Definitions and notation

The observations of an M.R.P. consist of the observed states of the system J_0, J_1, J_2, ... together with the observed sojourn times

X_0 (assumed to be zero), X_1, X_2, ... in the successive states. For example, the M.R.P. makes transition from J_{k-1} to J_k after spending time X_k in state J_{k-1}. Since the J_n's are realizations of the imbedded Markov chain, we have

$$P\{J_n = j | J_0, J_1, \ldots, J_{n-1} = i\} = P\{J_n = j | J_{n-1} = i\} = p_{ij}, \qquad (1.2.1)$$

which are the transition probabilities, with $\sum_{j=1}^{m} p_{ij} = 1$. Also,

$$P\{X_n \leq x | J_0, J_1, \ldots, J_{n-1} = i, J_n = j\} = P\{X_n \leq x | J_{n-1} = i, J_n = j\} = F_{ij}(x), \quad (1.2.2)$$

and combining these two, we have

$$P\{J_n = j, X_n \leq x | J_0, J_1, \ldots, J_{n-1} = i\} = P\{J_n = j, X_n \leq x | J_{n-1} = i\}$$

$$= p_{ij} F_{ij}(x) = Q_{ij}(x) , \qquad (1.2.3)$$

which is called the transition d.f. of the M.R.P., and is considered to be a basic quantity thereof. From (1.2.3), it is seen that

$$Q_{ij}(\infty) = p_{ij} ,$$

and if we let $H_i(x) = \sum_{j=1}^{m} Q_{ij}(x)$, then $H_i(\infty) = 1$, since <u>the underlying Markov chain is assumed to be an irreducible aperiodic recurrent class</u>, an assumption made throughout the paper. If we let $S_0 = X_0$, $S_n = X_0 + X_1 + \ldots + X_n$ be the time to the n^{th} transition, then the process $\{J_n, S_n, n \geq 0\}$ is equivalent to $\{Z_t; t \geq 0\} = J_n$, for $S_n < t < S_{n+1}$. Now, the vector of initial probabilities $\underline{a}$ for the imbedded Markov chain is given by

$$P\{J_0 = i\} = a_i, \quad i = 1, \ldots, m , \qquad (1.2.4)$$

and the M.R.P. is then determined by $(m, \underline{a}, Q)$, where Q is the matrix of transition d.f.'s $Q_{ij}(x)$.

Another set of observations related to the J_n's is $N_j(t)$, the number of times the system visits state j in time (0,t), and $N_{ij}(t)$, the number of transitions from state i to state j in (0,t). Now,

$$N(t) = \sum_{j=1}^{m} N_j(t) = \sup \{n \geq 0 \mid \sum_{i=0}^{n} X_i \leq t\} \tag{1.2.5}$$

is the total number of transitions, and we let the vector

$$\underline{N}(t) = [N_1(t), N_2(t), \ldots, N_m(t)]' \quad .$$

Then the process $\{\underline{N}(t); t \geq 0\}$ is the M.R.P. we are interested in. This is a natural analogue to the counting process in Renewal theory.

Throughout this paper random variables and their distribution functions will be denoted by capital letters, whereas the corresponding Laplace-Stieltjes Transforms (L.-S.T.) will be denoted by lower case letters. For example, the transition distribution function $Q_{ij}(x)$ will have as its L.-S.T.

$$q_{ij}(s) = \int_0^\infty e^{-sx}\, dQ_{ij}(x) \;, \qquad \text{for } s \geq 0 \;, \tag{1.2.6}$$

with q(s) representing the matrix of quantities $q_{ij}(s)$, when they exist. A similar notation will be used for the distributions $F_{ij}(x)$, especially since $Q_{ij}(x)$ will always be taken to be equal to $p_{ij}F_{ij}(x)$. The same is also true for the distributions $H_i(x)$.

Considerable mention will be made of the moments of these distributions from time to time, so the following terms will be used to designate these moments:

$$\mu_{ij} = \int_0^\infty t\,dF_{ij}(t) \;, \qquad \eta_i = \int_0^\infty t\,dH_i(t) \tag{1.2.7}$$

$$\sigma_{ij}^2 = \int_0^\infty (t-\mu_{ij})^2 dF_{ij}(t) \;, \qquad \sigma_i^{\,2} = \int_0^\infty (t-\eta_i)^2 dH_i(t)$$

$$\mu_{3ij} = \int_0^\infty (t-\mu_{ij})^3 \, dF_{ij}(t) \quad .$$

Another technique we employ which will require additional notation is that of matrix convolution. The idea is that of matrix multiplication, except that the entries in the resulting matrix of convolutions are sums of convolutions instead of products; that is, for matrix-valued functions Q and R,

$$Q*R = \sum_{k=1}^{m} Q_{ik}*R_{ik} \ , \qquad \text{where}$$
$$Q_{ik}*R_{ik} = \int_0^t Q_{ik}(t-u)d_u R_{ik}(u) \ . \tag{1.2.8}$$

For n-fold convolutions we will use the symbol $Q^{(n)}$, with $Q^{(0)} = I$, $Q^{(n)} = Q*Q^{(n-1)}$, and

$$\sum_{n=0}^{\infty} Q^{(n)} = (I-Q)^{(-1)} \quad \text{, if the series converges.}$$

When L.-S.T.'s are taken, the convolutions are replaced by products.

Also, with regard to the matrices used throughout the paper, the symbol ${}_dA$ will mean a diagonal matrix whose elements are the diagonal elements of A, and whose off-diagonal elements are zero; that is,

$${}_dA = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & a_{mm} \end{bmatrix}$$

A special case is the identity matrix, denoted by I. The zero or null matrix is represented by 0, $\underline{e}$ will denote a column vector all of whose elements are unity, and E will signify the matrix $\underline{e}\,\underline{e}'$. Finally, δ_{ij} will be the usual Kronecker delta.

As mentioned in section 1 of this chapter, it is our intention to present those previous results which will be necessary for use in the later chapters. Our ultimate aim is twofold: first, we wish to develop a χ^2 goodness-of-fit test for a specified M.R.P. model, an extension of the results of Bartlett [1] and Patankar [25]; secondly, we wish to extend the results of Martin [22] in Bayesian analysis of Markov chains to M.R.P.'s, using the results of Kshirsagar and Wysocki [20]. To do these things we need the distribution and moments of $N_j(t)$. These were found in terms of generating functions by Pyke [27] and extended by Kshirsagar and Gupta [18] to include $\underline{N}(t)$. In [17] and [19] they derived asymptotic expressions for the first two moments of $\underline{N}(t)$, using series expansions of the L.-S.T.'s of these moments. We will refine these results by calculating the variances and covariances from the raw moments using the new expansion [19]. All this is basic for the goodness of fit work. For the Bayesian analysis we will need the distribution and moments of $N(t)=[N_{ij}(t)]$, found by Kshirsagar and Wysocki [20]. After this, some special cases are considered for both types of analyses, and a summary will be given.

3. Renewal Theory

The relationship between M.R.P.'s and Renewal processes was pointed out in section 1. It will be instructive at times to compare the results obtained for M.R.P.'s with those in Renewal processes. Therefore, we will stop at this point and review some of the basic work in Renewal processes. This account will necessarily be brief, but a more detailed account may be found in Cox [8].

Consider a sequence of nonnegative independent random variables with common continuous distribution function $F(x)$. In most examples the r.v.'s represent times to failure of a component, and a "renewal" occurs at each failure. Let $N(t)$ denote the number of renewals in $(0,t)$, and let S_r be the time to the r^{th} renewal. Now,

$$P\{N(t) < r\} = P\{S_r > t\} = 1 - F^{(r)}(t) , \qquad (1.3.1)$$

where $F^{(r)}(t)$ is the r-fold convolution of $F(t)$ with itself. Then

$$P\{N(t) = r\} = P\{N(t) < r+1\} - P\{N(t) < r\} = F^{(r+1)}(t) - F^{(r)}(t). \quad (1.3.2)$$

To make it easier to discuss statistical properties of $N(t)$, define the probability generating function (p.g.f.) of $N(t)$

$$G(t,\xi) = \sum_{r=0}^{\infty} \xi^r P\{N(t) = r\} = 1 + \sum_{r=1}^{\infty} \xi^{r-1}(\xi-1) F^{(r)}(t) . \qquad (1.3.3)$$

Now, let

$$f(s) = \int_0^{\infty} e^{-st} dF(t) ,$$

the L.-S.T. of $F(t)$, and

$$f^*(s) = \int_0^{\infty} e^{-st} f(t)dt,$$

which is Cox's notation for the L.-S.T. of $f(t)$, the density corresponding to $F(t)$. Since

$$\int_0^{\infty} e^{-st} F(t)dt = \frac{1}{s} \int_0^{\infty} e^{-st} f(t)dt = \frac{1}{s} f^*(s) , \qquad (1.3.4)$$

then we make the substitution $f(s) = \frac{1}{s} f^*(s)$ to make the notation consistent with the earlier definition of L.-S.T. Now, the L.-S.T. of G, using (1.3.3), is

$$g_0(s,\xi) = \frac{1}{s} + \frac{1}{s} \sum_{r=1}^{\infty} \xi^{r-1} (\xi-1) [f^*(s)]^r$$

$$= \frac{1 - f^*(s)}{s\{1 - \xi f^*(s)\}}$$

$$= [1 - f(s)] [1 - \xi f(s)]^{-1} \qquad (1.3.5)$$

for an ordinary renewal process. For a modified renewal process - one in which the first failure time has a different distribution from all the others, say $f_1(x_1)$ - the expression corresponding to (1.3.5) is

$$g_m(s,\xi) = \frac{1 - \xi f^*(s) + \xi f_1{}^*(s) - f_1{}^*(s)}{s\{1 - \xi f^*(s)\}}$$

$$= 1 - (1-\xi)\, f_1(s)[1 - \xi f(s)]^{-1} . \qquad (1.3.6)$$

From expressions (1.3.5) and (1.3.6) we can find the L.-S.T. of the renewal function $H(t) = E\{N(t)\}$ for the ordinary and modified renewal processes. The L.-S.T. of $H_0(t)$ is

$$h_0(s) = \left.\frac{\partial g_0(s,\xi)}{\partial \xi}\right|_{\xi = 1} = f(s)\ [1 - f(s)]^{-1} \qquad (1.3.7)$$

and the L.-S.T. of $H_m(t)$ is

$$h_m(s) = \left.\frac{\partial g_m(s,\xi)}{\partial \xi}\right|_{\xi = 1} = f_1(s)[1 - f(s)]^{-1} . \qquad (1.3.8)$$

It is practically impossible to invert these L.-S.T.'s in general and even in many specific cases. Instead it is necessary to expand them in series about $s = 0$ and use Tauberian arguments as outlined by Cox [8]. Since for any continuous function $k(x)$

$$k(s) = \int_0^\infty e^{-st}\, dk(t) ,$$

small values of s (always > 0) will correspond to large values of t. Therefore, the expressions obtained in this manner are valid only for large values of t. If we let

$$\mu = \int_0^\infty x dF(x) \quad \text{and} \quad \sigma^2 = \int_0^\infty x^2 dF(x) - \mu^2 ,$$

then we may expand

$$f(s) = 1 - s\mu + \frac{s^2}{2!}(\mu^2+\sigma^2) + o(s^2) , \qquad (1.3.9)$$

and we get

$$h_0(s) = \frac{1}{s\mu} + \frac{\sigma^2-\mu^2}{2\mu^2} + o(1) \quad , \tag{1.3.10}$$

whose inverse for large t is

$$H_0(t) = \frac{t}{\mu} + \frac{\sigma^2-\mu^2}{2\mu^2} + o(1) \quad . \tag{1.3.11}$$

Another special case which will be useful to us later on is what is known as an equilibrium renewal process. This is a modified renewal process in which $f_1(x)$ has the special form $\frac{1}{\mu}[1 - F(x)]$. This case arises when one begins observing a renewal process at some time $t_0 > 0$. When t_0 is large, Cox [8] shows that the first failure time distribution has the above form. Substituting into (1.3.8) $f_1(s) = \frac{1-f(s)}{\mu s}$, we get

$$h_e(s) = \frac{1}{s\mu} \quad , \quad \text{whence}$$

$$H_e(t) = \frac{t}{\mu} \text{ for large } t. \tag{1.3.12}$$

The variance of the number of renewals is also a quantity which will be needed for use later in the paper. It can be calculated using principles of elementary probability theory thusly:

$$\text{Var}\{N(t)\} = E\{N(t)[N(t)-1]\} + E\{N(t)\} - E^2\{N(t)\} \quad . \tag{1.3.13}$$

Now again from elementary theory, for an ordinary renewal process,

$$\text{L.-S.T. } E\{N(t)[N(t)-1]\} = \left.\frac{\partial^2 g_0(s,\xi)}{\partial \xi^2}\right|_{\xi = 1}$$

$$= 2\{[f(s)]\,[1-f(s)]^{-1}\}^2 \quad . \tag{1.3.14}$$

Then, expanding, inverting, and using (1.3.13), we get

$$\text{Var}\{N_0(t)\} = \frac{\sigma^2 t}{\mu^3} + \left(\frac{1}{12} + \frac{5\sigma^4}{4\mu^4} - \frac{2\mu_3}{\mu^3}\right) + o(1) \quad , \tag{1.3.15}$$

where μ_3 is the third central moment of $f(x)$. For the equilibrium renewal process a similar argument, using (1.3.6) and (1.3.12), yields

$$\mathrm{Var}\{N_e(t)\} = \frac{\sigma^2 t}{\mu^3} + \left(\frac{1}{6} + \frac{\sigma^4}{2\mu^4} - \frac{\mu_3}{3\mu^3}\right) + o(1) \ . \qquad (1.3.16)$$

One final comment should be made with regard to **ordinary** renewal processes. This concerns how large t should be for the asymptotic results to hold. Although no mathematically rigorous proof has been given, Cox argues that roughly when $\sigma << \mu$ for the failure-time distribution, then a minimum requirement for application of the asymptotic formulas is

$$t > \frac{\mu^3}{\sigma^2} \quad .$$

These are some of the main results from renewal theory necessary for use in this paper. As mentioned earlier, some similarities will be pointed out later between renewal processes and M.R.P.'s.

4. Basic Results in M.R.P.'s

The basic results given here include the distribution and moments of $N_j(t)$, found by Pyke [27], and the distribution and moments of the first passage time, also found by Pyke [27], all in terms of generating functions.

Pyke has shown that $P\{N_j(t) = k | Z_0 = i\}$, for $i,j = 1, \ldots, m$, $k > 0$, $t > 0$, has L.-S.T. p.g.f. matrix given by

$$\Psi_\xi = E - (1-\xi)[I-q(s)]^{-1} q(s)\{\xi I + (1-\xi)_d \ [(I-q(s))^{-1}]\} \quad . \qquad (1.4.1)$$

The moments of $N_j(t)$, conditional on $Z_0 = i$, may be found by differentiating (1.4.1) with respect to ξ the appropriate number of times and evaluating at $\xi = 1$. Let $M(t) = [M_{ij}(t)]$, and $R(t) = [R_{ij}(t)]$, where

$$M_{ij}(t) = E\{N_j(t)|Z_0 = i\} \quad , \tag{1.4.2}$$

and

$$R_{ij}(t) = E\{N_j(t)[N_j(t)-1]|Z_0 = i\} \quad , \tag{1.4.3}$$

with L.-S.T. matrices $m(s) = [m_{ij}(s)]$ and $r(s) = [r_{ij}(s)]$. Upon taking the derivative of (1.4.1), we have

$$\frac{\partial}{\partial\xi}\Psi_\xi = (I-q(s))^{-1}q(s)[\xi I + (1-\xi)_d\{(I-q(s))^{-1}\}]^{-1}$$

$$+ (1-\xi)(I-q(s))^{-1}q(s)[\xi I + (1-\xi)_d\{(I-q(s))^{-1}]^{-1}$$

$$\times [I - {}_d\{(I-q(s))^{-1}\}][\xi I + (1-\xi)_d\{(I-q(s))^{-1}\}]^{-1} .$$

Evaluation at $\xi = 1$ gives

$$m(s) = [I-q(s)]^{-1}q(s) \quad , \tag{1.4.4}$$

also given by Pyke [27]. To simplify some of our results, we show

$$[I-q(s)]^{-1}q(s) = q(s)[I-q(s)]^{-1} = [I-q(s)]^{-1} - I . \tag{1.4.5}$$

Starting with the identity $q(s) = q(s)$, we have

$$q(s) = q(s)[I-q(s)][I-q(s)]^{-1} = [I-q(s)]q(s)[I-q(s)]^{-1}, \text{ or}$$

$$[I-q(s)]^{-1}q(s) = q(s)\,[I - q(s)]^{-1} .$$

Also $\quad [I-q(s)]\,q(s)\,[I-q(s)]^{-1} = I - [I-q(s)]$, from the second step above, or

$$q(s)[I-q(s)]^{-1} = [I-q(s)]^{-1} - I .$$

Therefore m(s) can be written in any of these three forms. It is worthwhile mentioning here the analogous expression for the case of the ordinary renewal process, namely

$$h_0(s) = f(s)[1-f(s)]^{-1} ,$$

which is the same as m(s) would be with $m = 1$; that is, replace the matrices with scalars and we have $h_0(s)$.

Pyke has pointed out that the form of m(s) is important in that a knowledge of m(s) is equivalent to a knowledge of the basic quantity q(s). Thus, an M.R.P. is equally as well determined by (m, $\underline{a}$, M(t)) as it is by (m, $\underline{a}$, Q).

Now, to calculate the second factorial moment, differentiate (1.4.1) twice and evaluate at $\xi = 1$. Omitting the algebra, we have the L.-S.T.

$$\begin{aligned} r(s) &= 2[(I-q(s))^{-1}q(s)]_d\{(I-q(s))^{-1} - I\} \\ &= 2m(s)\,{}_dm(s) \quad . \end{aligned} \tag{1.4.6}$$

Calculation of the variance from this formula will be discussed in a later section.

Another result of Pyke's for which we will have need is the distribution of the so-called first passage time from state i into state j. Let

$$G_{ij}(t) = P[N_j(t) > 0|Z_0 = i] \quad . \tag{1.4.7}$$

If one considers transitions from state i into state i, $G_{ii}(t)$ is very similar to an ordinary renewal process except that the times spent before re-entering state i are not identically distributed for each i (i=1,...,m). Thus, the $G_{ij}(t)$ will be useful to us for purposes of comparison.

Pyke [27] has shown that $G_{ij}(t)$ has the L.-S.T.

$$g(s) = m(s)[I - {}_dm(s)(I + {}_dm(s))^{-1}] \quad , \tag{1.4.8}$$

or
$$g(s) = m(s)\,[I + {}_d m(s)]^{-1} , \qquad (1.4.9)$$

so that
$$g_{ij}(s) = \frac{m_{ij}(s)}{1+m_{jj}(s)} . \qquad (1.4.10)$$

The mean recurrence times for $G_{ii}(t)$, say b_{ii}, can be found using (1.4.9). If we let $B = [b_{ij}]$, then

$$B = \lim_{s \to 0} \{\frac{1}{s}[E - g(s)]\} \text{ , whose diagonal elements}$$

$$_dB = \lim_{s \to 0} \{_d[\frac{1}{s}(I - q(s))^{-1}]\} , \qquad (1.4.11)$$

are the desired mean recurrence times. The variance can be calculated by taking

$$\lim_{s \to 0} \{\frac{1}{s^2}[E - B - g(s)]\}$$

and subtracting the squares of the appropriate elements of B.

5. Further Results in M.R.P.'s

Çinlar [6] and Kshirsagar and Gupta [18] have extended Pyke's results to include the joint distribution of all the $N_j(t)$'s and Z_t. We give here the expressions for $\underline{N}(t)$ as derived by Kshirsagar and Gupta [18]. These will be necessary for later use. The L.-S.T. p.g.f. of this distribution is given by

$$\underline{\Psi} = [I - q(s)\xi]^{-1}\,[I - q(s)]\underline{e} , \qquad (1.5.1)$$

where the argument of the p.g.f. $\xi = \text{diag}(\xi_1, \ldots, \xi_m)$, and $|\xi_i| < 1$ for each i. The moments are found by differentiation as in the case of a single $N_j(t)$, and for this purpose we let

$$\underline{M}(t) = E\{\underline{N}(t)|Z_0 = i\} \quad \text{with L.-S.T. vector } \underline{m}(s),$$

$$R_{ij}(t) = E\{N_j(t)[N_j(t)-1]|Z_0 = i\} \quad \text{with L.-S.T. matrix } r(s),$$

$$C^i_{jk}(t) = E\{N_j(t)N_k(t)|Z_0 = i\} \quad \text{with L.-S.T. } c^i_{jk}(s).$$

Taking the appropriate derivatives and evaluating at $\xi = I$, we obtain the L.-S.T.'s

$$\underline{m}(s) = [I - q(s)]^{-1} q(s)\underline{e}$$

$$r(s) = 2[I-q(s)]^{-1} q(s)_d \{[I-q(s)]^{-1} - I\} \qquad (1.5.2)$$

$$c_{jk}^{i}(s) = m_{jk}(s)\, m_{ij}(s) + m_{kj}(s)\, m_{ik}(s),$$

the last two terms having been obtained by taking respectively $\left.\frac{\partial^2 \underline{\Psi}}{\partial \xi_j^2}\right|_{\xi = I}$ and $\left.\frac{\partial^2 \underline{\Psi}}{\partial \xi_j \partial \xi_k}\right|_{\xi = I}$. From (1.5.2) we can obtain the moments required for later use.

A result required for the discussion of Bayesian analysis is the distribution and moments of $N_{ij}(t)$, the number of one-step transitions in $(0,t)$. These quantities were derived by Kshirsagar and Wysocki [20], generalizing the results to M.R.P.'s that Whittle [34] obtained for Markov Chains.

When an M.R.P. is observed, it is seen that transition into a state k must be followed by an exit from that state, and that transition out of a state was necessarily preceded by entrance into it, with the exception of the final state j in the former case and the initial state i in the latter. These conditions, then, define a set of constraints on the observed $N(t) = [N_{ij}(t)]$, namely,

$$N_{k\cdot}(t) - N_{\cdot k}(t) = \delta_{ik} - \delta_{jk}, \quad (k = 1, \ldots, m), \qquad (1.5.3)$$

where $N_{k\cdot}(t) = \sum_{\ell=1}^{m} N_{k\ell}(t)$, $N_{\cdot k}(t) = \sum_{\ell=1}^{m} N_{\ell k}(t)$. Let $N = [n_{ij}]$, a matrix of fixed nonnegative integers. Then we define

$$W_{ij}(N,t) = P[N(t) = N, Z_t = j \mid Z_0 = i] \qquad (1.5.4)$$

as the joint distribution of the transition count matrix and the final state of the system at time t, given the initial state. If N does not satisfy (1.5.3), the probability is zero. Likewise, if N = 0, there are no transitions in (0,t), so the initial and final states are the same. Thus,

$$W_{ij}(0,t) = \begin{cases} 0, & j \neq i \\ 1 - H_i(t), & \text{otherwise} \end{cases} , \qquad (1.5.5)$$

noting that N = 0 satisfies (1.5.3) only when i = j, so we have zero probability again if (1.5.3) does not hold. To summarize, let us define

$$\Phi_m(i,j,P_0) = \{N | n_{ij} > 0, \text{ an integer}; n_{k\cdot} - n_{\cdot k} = \delta_{ik} - \delta_{jk};$$
$$n_{\alpha\beta} = 0 \text{ if } p_{\alpha\beta} = 0; \text{ i and j are initial and final states respectively}\} .$$

Now, for $N \neq 0$ and $N \varepsilon \Phi_m(i,j,P_0)$, considering transitions first to state k and then to j, we have

$$W_{ij}(N,t) = \sum_{k=1}^{m} Q_{ik}(t) * W_{kj}[N(i,k),t] , \qquad (1.5.6)$$

where N(i,k) denotes the N matrix with $(i,k)^{th}$ element reduced by one. Then (1.5.3) - (1.5.6) may be assembled as

$$W_{ij}(N,t) = \begin{cases} \sum_{k=1}^{m} Q_{ik}(t) * W_{kj}[N(i,k),t], & N \varepsilon \Phi_m(i,j,P_0) \\ 1 - H_i(t) , & N = 0 \text{ and } i = j \\ 0 , & \text{otherwise} . \end{cases} \qquad (1.5.7)$$

The L.-S.T. p.g.f. of $W_{ij}(N,t)$ is then given by

$$w_{ij}(\Xi,s) = \sum_{N\varepsilon\Phi_m(i,j,P_0)} \int_0^\infty e^{-st} \, dW_{ij}(N,t) \prod_{i,j} \xi_{ij}^{n_{ij}} , \qquad (1.5.8)$$

where $\Xi = [\xi_{ij}]$, $|\xi_{ij}| \leq 1$, $i,j = 1, \ldots, m$. Using (1.5.8) in (1.5.7)

$$w_{ij}(\Xi,s) = \sum_{k=1}^{m} q_{ik}(s)\, \xi_{ik} w_{kj}(\Xi,s) + \delta_{ij}\{1 - h_i(s)\} , \qquad (1.5.9)$$

or in matrix notation

$$w(\Xi,s) = q(s) \,\square\, \Xi\; w(\Xi,s) + I - h(s) , \qquad (1.5.10)$$

where $q(s) \,\square\, \Xi = [q_{ij}(s)\xi_{ij}]$. Solving (1.5.10) yields

$$w(\Xi,s) = [I - q(s) \,\square\, \Xi\,]^{-1} [I - h(s)]. \qquad (1.5.11)$$

From Whittle [34] we have that the coefficient of $\prod_{k,\ell=1}^{m} a_{k\ell}^{n_{k\ell}}$ in the $(i,j)^{th}$ element of $(I - A)^{-1}$, where $A = [a_{ij}]$, is

$$\begin{cases} N_{ji}^* \dfrac{\prod_{k=1}^{m} n_{k.}!}{\prod_{k,\ell=1}^{m} n_{k\ell}!} , & N \in \Phi_m(i,j,P_0) \\ 0, & \text{otherwise} , \end{cases}$$

where N_{ji}^* is the cofactor of the $(j,i)^{th}$ element of $N^* = [n_{ij}^*]$, and

$$n_{ij}^* = \begin{cases} \delta_{ij} - \dfrac{n_{ij}}{n_{i.}} , & n_{i.} > 0 \\ \delta_{ij} , & n_{i.} = 0 . \end{cases}$$

Hence the L.-S.T. of $W_{ij}(N,t)$ is

$$\begin{cases} N_{ji}^* \dfrac{\prod_{k=1}^{m} N_{k.}!}{\prod_{k,\ell=1}^{m} N_{k\ell}!} \prod_{k,\ell=1}^{m} [q_{k\ell}(s)]^{N_{k\ell}} [1-h_k(s)] , & N \in \Phi_m(i,j,P_0) \\ 0, & \text{otherwise} . \end{cases} \qquad (1.5.12)$$

To find the conditional distribution of $N(t)$ alone given $Z_0=i$, sum with respect to the final state j using (1.5.11) and (1.5.12). The value set for this is

$$\Phi_m(i,P_0) = \bigcup_{j=1}^{m} \Phi_m(i,j,P_0) ,$$

the set of all possible transition count matrices N which can occur in an M.R.P. with transition matrix P_0 and initial state i. These are matrices $N = [n_{\alpha\beta}]$ such that for <u>some</u> j $n_{.k} - n_{k.} = \delta_{ik} - \delta_{jk}$ $(k = 1, \ldots, m)$. If

$$W_i(N,t) = P\{N(t) = N|Z_0 = i\} \quad , \quad (i = 1, \ldots, m)$$

the L.-S.T. p.g.f. of $W_i(N,t)$ is given by $w_i(\Xi,s)$, the i^{th} element of

$$\underline{w}(\Xi,s) = w(\Xi,s)\underline{e}$$

$$= [I-q(s) \,\square\, \Xi]^{-1} [I-h(s)]\underline{e}$$

$$= [I-q(s) \,\square\, \Xi]^{-1} [I-q(s)]\underline{e} \, . \qquad (1.5.13)$$

Using (1.5.12), the L.-S.T. of $W_i(N,t)$ is

$$\begin{cases} \sum_{j=1}^{m} N_{ji}^* \cdot \dfrac{\prod_{\alpha=1}^{m} N_{\alpha.}!}{\prod_{\alpha,\beta=1}^{m} N_{\alpha\beta}!} \cdot \prod_{\alpha,\beta=1}^{m} \{q_{\alpha\beta}(s)\}^{N_{\alpha\beta}}\{1-h_\alpha(s)\}, & N \in \Phi_m(i,P_0) \\ 0, \quad \text{otherwise} \; . \end{cases} \qquad (1.5.14)$$

To calculate the moments of $W_i(N,t)$, we use the L.-S.T. and a technique Martin [22] uses for the case of Markov chains. Let

$$m_{\alpha\beta}(i,s) = \text{L.-S.T. of } \; E\{N_{\alpha\beta}(t)|Z_0 = i\}$$

$$c_{\alpha\beta\gamma\delta}(i,s) = \text{L.-S.T. of } \; E\{N_{\alpha\beta}(t)N_{\gamma\delta}(t)|Z_0 = i\} \; , \qquad (1.5.15)$$

and let $\underline{m}_{\alpha\beta}(s)$ and $\underline{c}_{\alpha\beta\gamma\delta}(s)$ be the respective column vectors. To employ Martin's method, first note that if $N(t|Z_0 = i)$ denotes a transition count in $(0,t)$ given initial state i, then

$$N_{\alpha\beta}(t|Z_0 = i) = \begin{cases} \delta_{i\alpha}\delta_{k\beta} + N_{\alpha\beta}(t-x_1|Z_0 = k), & \text{if } J_1 = k \\ 0, & \text{if there is no transition in } (0,t) . \end{cases} \quad (1.5.16)$$

Taking the L.-S.T. of the expectation on both sides,

$$\begin{aligned} m_{\alpha\beta}(i,s) &= \sum_{k=1}^{m} q_{ik}(s)\delta_{i\alpha}\delta_{k\beta} + \sum_{k=1}^{m} q_{ik}(s)m_{\alpha\beta}(k,s) \\ &= q_{i\beta}(s)\delta_{i\alpha} + \sum_{k=1}^{m} q_{ik}(s)m_{\alpha\beta}(k,s) \quad . \end{aligned}$$

Now

$$\underline{m}_{\alpha\beta}(s) = q_{\alpha\beta}(s)\underline{e}_{\alpha} + q(s)\underline{m}_{\alpha\beta}(s) \quad , \quad (1.5.17)$$

where $\underline{e}_{\alpha}$ is a vector with unity in the α^{th} position and zeros elsewhere. Solving for $\underline{m}_{\alpha\beta}(s)$, we have

$$\begin{aligned} \underline{m}_{\alpha\beta}(s) &= q_{\alpha\beta}(s)\ [I - q(s)]^{-1}\ \underline{e}_{\alpha} \\ &= q_{\alpha\beta}(s) \cdot \alpha^{th} \text{ column of } [I-q(s)]^{-1} \text{ , or} \end{aligned}$$

$$m_{\alpha\beta}(i,s) = q_{\alpha\beta}(s)\ \{[I - q(s)]^{-1}\}_{i\alpha} \quad , \quad (1.5.18)$$

where $\{Y\}_{i\alpha}$ is the $(i,\alpha)^{th}$ element of the matrix Y. By similar reasoning

$$N_{\alpha\beta}(t|Z_0=i)N_{\gamma\delta}(t|Z_0=i) = \begin{cases} \delta_{\alpha\gamma}\delta_{\beta\delta}\delta_{\alpha i}\delta_{\beta k}+\delta_{i\gamma}\delta_{k\delta}N_{\alpha\beta}(t-x_1|Z_0=i) \\ +\delta_{i\alpha}\delta_{k\beta}N_{\gamma\delta}(t-x_1|Z_0=i)+N_{\alpha\beta}(t-x_1|Z_0=i)N_{\gamma\delta}(t-x_1|Z_0=i), \\ \text{if } J_1 = k,\ (k = 1, \ldots, m) \\ 0, \text{ if there is no transition in } (0,t). \end{cases} \quad (1.5.19)$$

Taking L.-S.T.'s and combining previous results yields

$$c_{\alpha\beta\gamma\delta}(i,s) = \delta_{\alpha\gamma}\delta_{\beta\delta}m_{\alpha\beta}(i,s) + m_{\alpha\beta}(\delta,s)m_{\gamma\delta}(i,s) + m_{\gamma\delta}(\beta,s)m_{\alpha\beta}(i,s). \quad (1.5.20)$$

6. Asymptotic Formulas

All the results given up to this point have been expressed in terms of the L.-S.T.'s of the quantities involved. As mentioned in section 3 of this chapter, except for the simplest of cases, inversion of these L.-S.T.'s is not possible. It then becomes necessary to consider inversion of power series expansions about $s = 0$ of the quantities in which one is interested. The first successful attempt at this procedure was made by Kshirsagar and Gupta [17], who derived asymptotic results in terms of the basic quantities (m, $\underline{a}$, Q). It should be noted that the expression $[I - q(s)]^{-1}$ appears often in the results of the previous sections, so the necessity of an expansion for this quantity is obvious. The motivation for the first expansion was to find the means $M_{ij}(t)$ and variances $V_{ij}(t)$ of the $N_j(t)$'s. In [17], Kshirsagar and Gupta pointed out that the only things needed were the renewal process results (1.3.11) and (1.3.15), because one could consider "the instants at which the system enters the j^{th} state only ... an ordinary renewal process where the d.f. of the time interval between any two such instants is $G_{jj}(t)$." However, use of the first three moments of that distribution requires the expression (1.4.9), which involves $[I - q(s)]^{-1}$. Hence, "... the whole procedure boils down to expanding $[I - q(s)]^{-1}$ in powers of s," substituting it into the appropriate expression, examining the behavior of the quantity of interest for small s, and applying Tauberian arguments as in the case of an ordinary renewal process.

In their paper, Kshirsagar and Gupta obtained an expansion from

$$I - q(s) = I - P_0 + sP_1 - \frac{s^2}{2!} P_2 + \frac{s^3}{3!} P_3 + \dots , \qquad (1.6.1)$$

where

$$P_k = \int_0^\infty x^k \, dQ(x) , \quad (k = 0, 1, 2, \dots) \qquad (1.6.2)$$

when these moments exist. The trouble here, of course, is that since $(I-P_0)\underline{e} = 0$, then $(I-P_0)$ is singular. Therefore the expansion was given in terms of certain matrices H_r $(r = 0, 1, 2, \dots)$ obtainable from the adjoint of the matrix $I - P_0 + sP_1$.

Two recent papers by Hunter [12] and Keilson [14] have given different and independent solutions to the same problem. Hunter's approach involves the use of a generalized inverse (Rao [31]) of $I - P_0$ and the relationship between the moments of G_{ij} and Q_{ij}. Keilson, on the other hand, finds the first two derivatives of $[I-q(s)]^{-1} - I$ with respect to s at $s = 0$, using the spectral decomposition of $q(s)$. These derivatives give the coefficients in powers of s of the expansion of $[I-q(s)]^{-1} - I$.

A still more recent paper by Kshirsagar and Gupta [19] uses the generalized inverse method directly on $[I-q(s)]^{-1}$. This method saves the unnecessary complications and manual labor involved in obtaining limits from the spectral decomposition, and since it works directly on $[I-q(s)]^{-1}$, it is more desirable than the method of indirectly going through the moments of G_{ij}.

Let $\underline{U}' = [U_1, \dots, U_m]$ be the vector of stationary state probabilities for the Markov chain. Then it is well known from Markov chain theory that

$$P_0\underline{e} = \underline{e} \quad \text{and} \quad \underline{U}'P_0 = \underline{U}' . \qquad (1.6.3)$$

Let $L = \underline{e}\,\underline{U}'$. The matrix $Z = (I - P_0 + L)^{-1}$ is known as the fundamental matrix of the Markov chain (see Kemeny and Snell [15]). If we let $k_r = \underline{U}'P_r\underline{e}$, $(r = 1, 2, \ldots)$, it can be shown (Hunter [12]) that $LP_rL = k_rL$.

To expand $[I-q(s)]^{-1}$, we use a well-known result from ordinary renewal processes (see, for example, Keilson [14]) that if the Markov chain is ergodic and the P_k's exist, then

$$[I-q(s)]^{-1} = \frac{1}{s} A_{-1} + A_0 + sA_1 + s^2A_2 + \ldots \quad . \tag{1.6.4}$$

By equating coefficients of s^r in the identity

$$[I - P_0 + sP_1 - \frac{s^2}{2!} P_2 + \ldots][\frac{1}{s} A_{-1} + A_0 + sA_1 + s^2A_2 + \ldots] \equiv I, \tag{1.6.5}$$

Kshirsagar and Gupta [19] have determined the A_i's as

$$A_r = \sum_{\alpha=1}^{r+1} \frac{(-1)^\alpha}{\alpha!} \{ZP_\alpha - \frac{1}{k_1} LP_1ZP_\alpha - \frac{1}{k_1(\alpha+1)} LP_{\alpha+1}\} A_{r-\alpha} +$$
$$+ \delta_{r0} (I - \frac{1}{k_1} LP_1) Z + \frac{1}{k_1} \delta_{r+1,0} L , \quad (r = -1, 0, 1, \ldots). \tag{1.6.6}$$

Specifically,

$$A_{-1} = \frac{1}{k_1} L$$
$$A_0 = (I - \frac{1}{k_1} LP_1) Z(I - \frac{1}{k_1} P_1L) + \frac{k_2}{2k_1^2} L \tag{1.6.7}$$
$$A_1 = \{-ZP_1 + \frac{1}{k_1}LP_1ZP_1 + \frac{1}{2k_1}LP_2\}A_0 + \frac{1}{2k_1}\{ZP_2 - \frac{1}{k_1}LP_1ZP_2 - \frac{1}{3k_1}LP_3\}L.$$

From this expansion we can obtain immediately an expression for the L.-S.T. of the renewal function m(s), namely,

$$m(s) = [I-q(s)]^{-1} - I$$
$$= [\frac{1}{s} A_{-1} + A_0 + sA_1 + o(s)] - I$$
$$= \frac{1}{s} A_{-1} + (A_0-I) + sA_1 + o(s) \tag{1.6.8}$$

whose inverse is, for large t,

$$M(t) = t \cdot A_{-1} + (A_0 - I) + o(1) \quad . \qquad (1.6.9)$$

Note that $\lim_{t \to \infty} \frac{1}{t} M(t) = A_{-1} = \frac{1}{k_1} L$, which is independent of the initial state. To calculate the variance of the number of visits to state j given $Z_0 = i$, $V(t) = [V_{ij}(t)]$, use the expansion, (1.4.6), and the technique described in section three (1.3.13) to obtain

$$V(t) = [V_{ij}(t)] = \mathrm{Var}[N_j(t) | Z_0 = i] = t[2A_{-1}\, {}_dA_0 - A_{-1}] \; +$$

$$+ [A_0\, {}_dA_0 - A_0 \square A_0 + {}_dA_0 - A_0 + A_{-1}\, {}_dA_1 + A_1\, {}_dA_{-1}] + o(1) . \qquad (1.6.10)$$

Likewise, the covariances may be calculated from (1.5.2) in the following way. Let

$$\Gamma^i_{jk}(t) = \mathrm{Cov}[N_j(t), N_k(t) | Z_0 = i] \; . \qquad (1.6.11)$$

Then

$$\Gamma^i_{jk}(t) = C^i_{jk}(t) - M_{ij}(t)\, M_{ik}(t) \quad . \qquad (1.6.12)$$

From (1.5.2) we get the L.-S.T. $c^i_{jk}(s)$, which, when we expand, invert, and use (1.6.12), yields

$$\Gamma^i_{jk}(t) = t \cdot \frac{1}{k_1}[U_j a^{(0)}_{jk} + U_k a^{(0)}_{kj}] + [\frac{1}{k_1}\{U_j(a^{(1)}_{jk} + a^{(1)}_{ik}) + U_k(a^{(1)}_{ij} + a^{(1)}_{kj})\} +$$

$$+ a^{(0)}_{jk}(a^{(0)}_{ij} + a^{(0)}_{ik} - \delta_{ij} - \delta_{ik}) - (a^{(0)}_{ij} - \delta_{ij})(a^{(0)}_{ik} - \delta_{ik})] + o(1) \; , \quad (1.6.13)$$

where $a^{(0)}_{ij}$ and $a^{(1)}_{ij}$ are the $(i,j)^{th}$ elements of A_0 and A_1 (1.6.7), respectively. These are all the applications of the expansion of $[I-q(s)]^{-1}$ to be considered for the present. However, in later chapters more use will be made of this equation.

This ends our discussion of previous results necessary for later use. In the ensuing chapters we shall discuss a χ^2 goodness-of-fit test for a specified M.R.P. model, some Bayesian analysis of M.R.P.'s, some special cases of both these procedures, and a summary and outline of some avenues of further research in the field.

CHAPTER II

THE χ^2 GOODNESS-OF-FIT TEST

1. Introduction

Bartlett [1] has considered a χ^2 test for testing the goodness of fit of a hypothetical matrix of transition probabilities in the case of a Markov chain, a realization of which is available. His procedure consisted of observing a Markov chain with m states and computing the classical χ^2 statistic, namely, $\sum_{i=1}^{m} \frac{(n_i-m_i)^2}{m_i}$, where n_i denotes the number of visits to the i^{th} state $(i = 1, \ldots, m)$ of the Markov chain, and m_i denotes the expected number of visits to the i^{th} state, assuming the hypothetical matrix of the transition probabilities to be the true one. The hypothesis is then rejected if the χ^2 exceeds a given percentage point of the standard χ^2 distribution with $m - 1$ degrees of freedom. The validity of this approach lies in the asymptotic normality of the n_i, which Bartlett proved in [1].

Patankar [25] modified Bartlett's procedure by calculating the expectation and variance of the χ^2, say A and 2B, respectively, thus taking $\frac{A\chi^2}{B}$ to have an asymptotic χ^2 distribution with $\frac{A^2}{B}$ degrees of freedom. This is a better approximation to the standard χ^2 in that now the first two moments of the modified statistic agree exactly with those of a χ^2. These results were applied to Markov chain models in [25].

Billingsley [4], Bhat [3], Whittle [34], Good[11], and others have also dealt with statistical inference for Markov chains. However, not much has been done in the area of statistical inference for Markov Renewal processes. Moore and Pyke [23] have derived estimators for the p_{ij} and $F_{ij}(x)$ of an M.R.P. The estimators of p_{ij} were proved in [23] to be asymptotically normally distributed. This fact will be used in this chapter.

It is our intention to extend the procedures of Bartlett [1] and Patankar [25] to M.R.P.'s, using the results of Moore and Pyke [23], and the results given in the previous chapter. At this point some differences between the properties of the Markov chain χ^2 and that of the M.R.P. should be noted. First and most important, in a Markov chain model a fixed (total) number of transitions occurs in a fixed length of time, whereas if we observe the M.R.P. for a fixed length of time $(0,t)$, as done by Moore and Pyke [23], the number of transitions will be random. In a Markov chain a transition will always occur after every unit length of time, but this is not necessarily so in an M.R.P. Furthermore, since for the Markov chain $\sum_{i=1}^{m} n_i = n$, a fixed value, there is a linear constraint on the variables and the resulting χ^2 can be proved to have $m - 1$ degrees of freedom. No such restriction exists for the M.R.P. Also, the expectations m_i of the n_i in the Markov chain are explicit functions of the transition probabilities p_{ij} and the total number of transitions n. However, in an M.R.P., the matrix of d.f.'s is also involved, and is assumed to be known.

2. The χ^2 Statistic for an M.R.P. of m States

In section I.2 we stated that the imbedded Markov chain of an M.R.P. is assumed to be irreducible aperiodic. We will continue to assume this throughout the paper. From this and from Moore and Pyke [23] we know that the $N_j(t)$ defined in the previous chapter are asymptotically normally distributed with means $M_{ij}(t)$, variances $V_{ij}(t)$ and covariances $\Gamma^i_{jk}(t)$, when the initial state is i.

The χ^2 goodness-of-fit statistic (similar to Bartlett's in the Markov chain case) is

$$\chi^2 = \sum_{j=1}^{m} \frac{[N_j(t) - M_{ij}(t)]^2}{M_{ij}(t)} \quad . \tag{2.2.1}$$

We shall approximate the distribution of this quantity by a standard χ^2 distribution. The $M_{ij}(t)$ are available only in terms of asymptotic expansions given in section I.6.

Now, if we write

$$\underline{Y}' = [N_1(t)-M_{i1}(t), \ldots, N_m(t)-M_{im}(t)] \ , \tag{2.2.2}$$

then asymptotically $\underline{Y}' \sim N(0,V)$, where the elements of V are variances and covariances of the $N_j(t)$, and

$$\chi^2 = \underline{Y}'Q\underline{Y} \ , \tag{2.2.3}$$

where

$$Q = \text{diag}\left[\frac{1}{M_{i1}(t)}, \ldots, \frac{1}{M_{im}(t)}\right] \quad .$$

We now prove the following lemma about the first two moments of the χ^2.

LEMMA. (1) $E(\chi^2) = E(\underline{Y}'Q\underline{Y}) = \text{trace } (VQ)$. (2.2.4)

(2) $\text{Var}(\chi^2) = \text{Var}(\underline{Y}'Q\underline{Y}) = 2 \cdot \text{trace}\{(VQ)^2\}$ (2.2.5)

PROOF. (1) $\underline{Y} \sim N(0,V)$, so there exists a nonsingular matrix C such that $C'VC = I$; i.e., $V = (C'C)^{-1}$. Then $\underline{W} = C\underline{Y} \sim N(0,I)$, and we have $\underline{Y}'C'C'^{-1}QC^{-1}C\underline{Y} = \underline{W}C'^{-1}QC^{-1}\underline{W} = \underline{W}'K\underline{W}$, say. Now, there exists an orthogonal matrix P such that $PKP' = \mathrm{diag}(\lambda_1,\ldots,\lambda_m)$, where $\lambda_1, \ldots, \lambda_m$ are the latent roots of K. Then $\underline{Z}=P\underline{W}\sim N(0,I)$, and $\underline{W}'P'PKP'P\underline{W} = \underline{Z}'\mathrm{diag}(\lambda_1,\ldots,\lambda_m)\underline{Z} = \sum_{k=1}^{m} \lambda_k Z_k^2$. Since $Z_k \sim NID(0,1)$, then $Z_k^2 \sim \chi^2(1)$, and $E[\underline{Z}'\, \mathrm{diag}(\lambda_1,\ldots,\lambda_m)\underline{Z}] = \sum_{k=1}^{m} \lambda_k \cdot 1 = \mathrm{tr}K$, the trace of K. But, $\mathrm{tr}K = \mathrm{tr}\{C'^{-1}QC^{-1}\} = \mathrm{tr}\{C^{-1}C'^{-1}Q\} = \mathrm{tr}\{(C'C)^{-1}Q\} = \mathrm{tr}VQ$.

(2) Likewise,
$$\begin{aligned}\mathrm{Var}\{\underline{Z}'\,\mathrm{diag}(\lambda_1,\ldots,\lambda_m)\underline{Z}\} &= \sum_{k=1}^{m} 2\lambda_k^2 \\ &= 2\,\mathrm{tr}(K^2) \\ &= 2\,\mathrm{tr}\{(VQ)^2\}.\end{aligned}$$

The elements of the variance-covariance matrix V are designated by

$$\begin{cases} \Gamma^{i}_{jk}(t), & \text{if } j \neq k \\ V_{ij}(t), & \text{if } j = k \end{cases},$$

and we are assuming a fixed initial state i, at least for this discussion. Then we have

$$VQ = \begin{bmatrix} V_{i1}(t) & \Gamma^{i}_{12}(t) & \cdots & \Gamma^{i}_{1m}(t) \\ & \vdots & & \\ \Gamma^{i}_{m1}(t) & \Gamma^{i}_{m2}(t) & \cdots & V_{im}(t) \end{bmatrix} \begin{bmatrix} \frac{1}{M_{i1}(t)} & & 0 \\ & \frac{1}{M_{i2}(t)} & \\ 0 & & \ddots \; \frac{1}{M_{im}(t)} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{V_{i1}(t)}{M_{i1}(t)} & \frac{\Gamma^{i}_{12}(t)}{M_{i2}(t)} & \cdots & \frac{\Gamma^{i}_{1m}(t)}{M_{im}(t)} \\ & \vdots & & \\ \frac{\Gamma^{i}_{m1}(t)}{M_{i1}(t)} & \frac{\Gamma^{i}_{m2}(t)}{M_{i2}(t)} & \cdots & \frac{V_{im}(t)}{M_{im}(t)} \end{bmatrix} . \qquad (2.2.6)$$

The trace of a matrix is known to be the sum of the diagonal elements, so, using the lemma, we have

$$E(\chi^2) = \sum_{j=1}^{m} \frac{V_{ij}(t)}{M_{ij}(t)} \quad . \qquad (2.2.7)$$

Substituting the appropriate expansions from section I.6, dividing, and simplifying gives us the expansion, for large t, of the expectation. That is,

$$A = E(\chi^2) = \sum_{j=1}^{m} \frac{t\cdot\left[\frac{1}{k_1}U_j(2a_{jj}^{(0)}-1)\right]+\left[\delta_{ij}\left\{\frac{2}{k_1}U_j a_{jj}^{(1)}\right\}-(1-\delta_{ij})\left\{(a_{ij}^{(0)})^2+a_{ij}^{(0)}\right\}\right]+o(1)}{t\cdot\frac{1}{k_1}U_j + (a_{ij}^{(0)}-\delta_{ij}) + o(1)}$$

$$= \sum_{j=1}^{m} (2a_{jj}^{(0)}-1) + o(1) \quad , \qquad (2.2.8)$$

where, again, $a_{jj}^{(r)}$ is the $(j,j)^{th}$ element of A_r defined in (1.6.7).

To find the variance of the χ^2, we use the second part of the lemma, $2tr\{(VQ)^2\}$. Since the elements of V are variances and covariances, $\Gamma^i_{k\ell}(t) = \Gamma^i_{\ell k}(t)$ for specified k and ℓ, because $Cov[N_\ell(t),N_k(t)|Z_0=i] = Cov[N_k(t),N_\ell(t)|Z_0=i]$. Also, since we are interested in the trace of $(VQ)^2$, we need write down only the diagonal elements of $(VQ)^2$. After squaring VQ and making the symmetry simplifications we get the following diagonal elements:

$${}_d\{(VQ)^2\} = \text{diag}\left[\frac{(V_{i1}(t))^2}{(M_{i1}(t))^2} + \frac{(\Gamma^i_{12}(t))^2}{M_{i1}(t)M_{i2}(t)} +\dots+ \frac{(\Gamma^i_{1m}(t))^2}{M_{i1}(t)M_{im}(t)}, \dots, \frac{(\Gamma^i_{m1}(t))^2}{M_{im}(t)M_{i1}(t)} +\dots+ \frac{(V_{im}(t))^2}{(M_{im}(t))^2}\right] \quad . \qquad (2.2.9)$$

Then, taking $2tr\{(VQ)^2\}$ yields

$$Var(\chi^2) = 2\cdot\sum_{j,k=1}^{m}(1-\delta_{jk})\frac{(\Gamma^i_{jk}(t))^2}{M_{ij}(t)M_{ik}(t)} + 2\cdot\sum_{j=1}^{m}\left[\frac{V_{ij}(t)}{M_{ij}(t)}\right]^2 . \qquad (2.2.10)$$

We wish, as before, to substitute in the proper expansions from section I.6, divide and simplify the resulting expressions. The algebra is complicated by the squares and the double summation in the first term of (2.2.10), but after some simplification, we get

$$\mathrm{Var}(\chi^2) = 2 \cdot \sum_{j,k=1}^{m} (1-\delta_{jk}) \frac{t^2 \cdot \frac{1}{k_1^2}\left[U_j^2 (\overset{(0)}{a}_{jk})^2 + 2U_jU_k \overset{(0)}{a}_{kj} \overset{(0)}{a}_{jk} + U_k^2 (\overset{(0)}{a}_{kj})^2\right] + t \cdot c_1 + c_2 + o(1)}{t^2 \cdot \frac{1}{k_1^2} U_jU_k + c_3 + o(1)}$$

$$+ 2 \cdot \sum_{j=1}^{m} [(2\overset{(0)}{a}_{jj} - 1) + o(1)]^2 \text{ , where } c_1, c_2, c_3 \text{ are constants.}$$

Thus,

$$\mathrm{Var}(\chi^2) = 2 \cdot \sum_{j,k=1}^{m} (1-\delta_{jk}) \left\{ \frac{U_j(\overset{(0)}{a}_{jk})^2}{U_k} + 2\overset{(0)}{a}_{jk}\overset{(0)}{a}_{kj} + \frac{U_k(\overset{(0)}{a}_{kj})^2}{U_j} \right\}$$

$$+ 2 \cdot \sum_{j=1}^{m} \{4(\overset{(0)}{a}_{jj})^2 - 4\overset{(0)}{a}_{jj} + 1\} + o(1) \; . \tag{2.2.11}$$

Then, taking the mean and variance to be A and 2B, respectively, we can get a better approximation that $\chi'^2 = A\chi^2/B$ is distributed as a χ^2 with approximately A^2/B degrees of freedom, if the hypothetical probabilities p_{ij} are the true ones. The moments of χ'^2, then, exactly fit the data. In a later chapter, we will discuss the calculation of these quantities for a specific two-state M.R.P.

3. The χ^2 Statistic for an Equilibrium M.R.P.

In the previous chapter, the discussion of one-state renewal processes included the analysis of an equilibrium renewal process, which was defined and described in section I.3. The mean and variance of the number of renewals are given by (1.3.12) and (1.3.16), respectively, and, as can be seen from those equations, the moments are different from those of the ordinary renewal process. The reason for this, of course, is that in general the time to the first renewal has a different distribution from the subsequent life distributions. The same is true in the case of an M.R.P. Suppose that, instead of observing the process from its start, we begin observing it only after a sufficiently long time has elapsed. Then the first state and transition have different distributions from the subsequent ones. The process is called an Equilibrium Markov Renewal Process (E.M.R.P.), and Pyke [27] has proved that this process has a set of initial probabilities given by

$$\tilde{a}_i = \frac{\eta_i}{b_{ii}} \;, \quad (i = 1, \ldots, m) \qquad (2.3.1)$$

and the transition d.f. for the first state and transition is

$$\tilde{Q}_{ij}(t) = \frac{p_{ij}}{\eta_i} \int_0^t [1 - F_{ij}(y)]dy \;, \quad (i,j = 1, \ldots, m), \qquad (2.3.2)$$

where η_i, b_{ii}, p_{ij} and $F_{ij}(\cdot)$ are as defined in section I.2. After the first transition occurs, the transition d.f. becomes $Q_{ij}(x)$ as in the case of an ordinary M.R.P.

To calculate the goodness-of-fit statistic for this type of process, we need the distribution and moments of the $N_j(t)$ for this set of assumptions. Kshirsagar and Gupta [18] have provided us with most of these quantities, and we shall extend and improve their work using the expansion (1.6.4). First, they have shown that for the E.M.R.P.

$$b_{ii} = \frac{k_1}{U_i} \;, \text{ and furthermore, that}$$

$$\tilde{q}(s) = \left[\frac{p_{ij} - q_{ij}(s)}{s\eta_i}\right] = \frac{1}{s}\,\mathrm{diag}\left(\frac{1}{\eta_1}, \ldots, \frac{1}{\eta_m}\right)\left(P_0 - q(s)\right) . \qquad (2.3.3)$$

Now, since we are not concerned about the initial state of the system, we find the L.-S.T. of the absolute means of $N_1(t), N_2(t), \ldots, N_m(t)$ instead of the means conditional on the initial state as in the ordinary M.R.P. These L.-S.T.'s were also found by Kshirsagar and Gupta to be

$$[\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m]\tilde{q}(s)\,[I - q(s)]^{-1}$$

$$= \left[\frac{\eta_1}{b_{11}}, \ldots, \frac{\eta_m}{b_{mm}}\right] \frac{1}{s} \operatorname{diag}\left[\frac{1}{\eta_1}, \ldots, \frac{1}{\eta_m}\right] [P_0 - q(s)][I - q(s)]^{-1}$$

$$= \frac{1}{s} \cdot \frac{1}{k_1}\;[\underline{U}' - \underline{U}'q(s)]\,[I - q(s)]^{-1}, \text{ since } \underline{U}'P_0 = \underline{U}'$$

$$= \frac{1}{s} \cdot \frac{1}{k_1}\,\underline{U}' \quad . \qquad (2.3.4)$$

Therefore

$$E[N_1(t), \ldots, N_m(t)] = t\left[\frac{1}{b_{11}}, \ldots, \frac{1}{b_{mm}}\right], \text{ or}$$

$$E[\underline{N}'(t)] = t \cdot \frac{1}{k_1}\,\underline{U}' \quad , \qquad (2.3.5)$$

a result which compares with (1.3.12) for the one-state case. We might also point out that this result corresponds to the idea of treating each state separately and considering transitions into that state only as the renewals using the $G_{ii}(x)$ as the life distributions (see section I.6). Another thing this shows is that the means are independent of the d.f.'s $F_{ij}(x)$.

The above means were derived using the Kshirsagar and Gupta result that $\tilde{m}(s) = \tilde{q}(s)[I-q(s)]^{-1}$, where $\tilde{q}(s)$ was defined in (2.3.3), and multiplying by $\underline{\tilde{a}}$, the vector of initial probabilities. To find the unconditional variances and covariances of the $N_j(t)$, we use the formulas of [18] for the L.-S.T.'s of the second factorial and cross-product moments, namely

$$\tilde{r}(s) = 2\tilde{m}(s)_d[m(s)] \quad \text{and} \quad \tilde{c}^{\,i}_{jk}(s) = \tilde{m}_{ij}(s)m_{jk}(s) + \tilde{m}_{ik}(s)m_{kj}(s), \quad (2.3.6)$$

and multiply by the vector of initial probabilities. First,

$$\underset{\sim}{a}'\tilde{r}(s) = \underset{\sim}{a}'2\tilde{m}(s)_d m(s)$$

$$= \frac{2}{s}\ \frac{1}{k_1}\ \underset{\sim}{U}'_d\{[I-q(s)]^{-1} - I\}\ , \qquad (2.3.7)$$

and, using the expansion (1.6.4), we have

$$\underset{\sim}{a}'\tilde{r}(s) = \frac{2}{s^2}\left[\frac{1}{k_1^2}(U_1^{\,2},\ldots,U_m^{\,2})\right] + \frac{1}{s}\left[\frac{2}{k_1}\ \underset{\sim}{U}'(_dA_0-I)\right] + \frac{2}{k_1}\ \underset{\sim}{U}'\ _dA_1 + o(1)\ , \qquad (2.3.8)$$

and, inverting and applying the same technique as (1.3.13) to the $(i,j)^{th}$ element of these unconditional expectations, we have

$$\mathrm{Var}[N_j(t)] = t^2\left[\frac{1}{k_1^2}U_j^{\,2}\right]+2t\cdot\frac{1}{k_1}(U_j a_{jj}^{(0)}-U_j)+2\cdot\frac{1}{k_1}U_j a_{jj}^{(1)}+t\cdot\frac{1}{k_1}U_j-t^2\cdot\frac{1}{k_1^2}U_j^{\,2}+o(1)$$

$$= t\left[\frac{1}{k_1}U_j(2a_{jj}^{(0)}-1)\right] + 2\cdot\frac{1}{k_1}\ U_j a_{jj}^{(1)} + o(1)\ . \qquad (2.3.9)$$

Again this satisfies our intuition with regard to comparison with the one-state renewal process, in that the coefficient of t in the expansion (2.3.9) is identical to that of (1.6.10), considering individual elements. Notice, too, that the constant terms are not equal, also the case with renewal processes as seen in (1.3.15) and (1.3.16).

To find the unconditional covariances, we apply the same arguments as above to the $\tilde{c}^{\,i}_{jk}(s)$. Now

$$\sum_{i=1}^{m}\tilde{a}_i\tilde{c}^{\,i}_{jk}(s) = \sum_{i=1}^{m}\tilde{a}_i[\tilde{m}_{ij}(s)m_{jk}(s) + \tilde{m}_{ik}(s)m_{kj}(s)]$$

$$= (\tilde{a}_1,\ldots,\tilde{a}_m)[j^{th}\text{ column of }\tilde{m}(s)]m_{jk}(s)+(\tilde{a}_1,\ldots,\tilde{a}_m) \times [k^{th}\text{ column of }\tilde{m}(s)]m_{kj}(s)$$

$$= \frac{1}{s}\cdot\frac{1}{k_1}\ U_j\cdot m_{jk}(s) + \frac{1}{s}\cdot\frac{1}{k_1}\ U_k\cdot m_{kj}(s)$$

$$= \frac{2}{s^2}\left[\frac{1}{k_1^2}\ U_jU_k\right]+\frac{1}{s}\cdot\frac{1}{k_1}\left[U_j(a_{jk}^{(0)}-\delta_{jk})+U_k(a_{kj}^{(0)}-\delta_{kj})\right]+\left[\frac{1}{k_1}(U_j a_{jk}^{(1)}+U_k a_{kj}^{(1)})\right]+o(1)\ . \qquad (2.3.10)$$

Then, as above, inverting and subtracting the product of the unconditional expectations, we get

$$\mathrm{Cov}\{N_j(t),N_k(t)\} = t\cdot\frac{1}{k_1}[U_j(a^{(0)}_{jk}-\delta_{jk})+U_k(a^{(0)}_{kj}-\delta_{kj})]+[\frac{1}{k_1}(U_j a^{(1)}_{jk}+U_k a^{(1)}_{kj})]+o(1). \tag{2.3.11}$$

If we denote the unconditional means, variances, and covariances above by $M_j(t)$, $V_j(t)$, and $\Gamma_{jk}(t)$, then the χ^2 statistic for this process becomes $\sum_{j=1}^{m} \frac{[N_j(t)-M_j(t)]^2}{M_j(t)}$ with mean

$$E(\chi^2) = \sum_{j=1}^{m} \frac{V_j(t)}{M_j(t)}$$

$$= \sum_{j=1}^{m} (2a^{(0)}_{jj}-1) + o(1) \quad . \tag{2.3.12}$$

The variance of the χ^2 for the E.M.R.P. becomes

$$\mathrm{Var}(\chi^2) = 2\sum_{j,k=1}^{m} (1-\delta_{jk}) \frac{[\Gamma_{jk}(t)]^2}{M_j(t)M_k(t)} + 2\sum_{j=1}^{m} \left[\frac{V_j(t)}{M_j(t)}\right]^2$$

$$= 2\sum_{j,k=1}^{m} (1-\delta_{jk})\left\{\frac{U_j}{U_k}(a^{(0)}_{jk}-\delta_{jk})^2+2(a^{(0)}_{jk}-\delta_{jk})(a^{(0)}_{kj}-\delta_{kj})+\frac{U_k}{U_j}(a^{(0)}_{kj}-\delta_{kj})^2\right\}$$

$$+ 2\sum_{j=1}^{m} \{4(a^{(0)}_{jj}) - 4a^{(0)}_{jj} + 1\} + o(1) \quad . \tag{2.3.13}$$

Considering the first summation, we see that in order to get a nonzero expression out of it, it is necessary that $j \neq k$. Since this is true, the δ_{jk} and δ_{kj} are zero, and (2.3.13) is identical to (2.2.11), a result we would intuitively anticipate since we are observing the process for a long time, and hopefully it reaches equilibrium sometime in our interval of observation.

As before we modify the computed χ^2 by multiplying by the mean and dividing by half the variance to obtain a new approximation with first two moments fitted to the data. Again, this procedure will be discussed in a later chapter for a specific two-state M.R.P.

CHAPTER III

BAYESIAN ANALYSIS FOR M.R.P.'S

1. Introduction

J. J. Martin [22] has given a quite thorough treatment of Bayesian analysis of Markov chains, using some of the results of Whittle [34] and others, as well as his own work. Judge, Lee, and Zellner [13] have also done some Bayesian estimation for Markov chain models. In these works the procedure is to assume a multivariate or matrix beta prior distribution for the transition probabilities p_{ij}, and find the posterior distribution given a sample of n observations from the Markov chain, where the sample size n is fixed in advance. The required estimates are then computed from this distribution and whatever loss function is required.

The observations consist of the states of the Markov chain over n transitions as in the case of the χ^2 analysis in the previous chapter, and the likelihood of the sample is written as

$$\prod_{i,j=1}^{m} p_{ij}^{n_{ij}} , \tag{3.1.1}$$

where the p_{ij} are the transition probabilities and n_{ij} are the observed number of transitions from states i to j in the observed n transitions. The distribution of the sample was first found by Whittle [34] and bears his name. Martin takes this distribution and uses the matrix beta distribution as a prior to find the posterior distribution he seeks.

Our purpose in this chapter is to extend the above procedure to Markov Renewal processes. We will use the matrix beta prior as Martin does, and we will identify and describe it in a later section. The sample distribution needed for this discussion is that of the $N_{ij}(t)$, found by Kshirsagar and Wysocki [20] and described in section I.5 above. Given these quantities, we find expressions for the posterior distributions with known and unknown initial states.

2. The Transition Count and Initial State Distributions

In section I.5 the L.-S.T. p.g.f. of the distribution of $N_{ij}(t)$ given the initial state was derived (as done by Kshirsagar and Wysocki [20]). The L.-S.T.'s of the first two moments were also given. By utilizing the expansion of $[I-q(s)]^{-1}$, (1.6.4), we may simplify those results somewhat and obtain expansions for them when t is large, as we have done for other quantities. Applying the expansion to (1.5.18), we have

$$m_{\alpha\beta}(i,s) = \frac{1}{s}\left(\frac{1}{k_1}p_{\alpha\beta}U_\alpha\right)+\left(p_{\alpha\beta}\overset{(0)}{a_{i\alpha}}-\mu_{\alpha\beta}p_{\alpha\beta}\cdot\frac{1}{k_1}U_\alpha\right)+s\left(p_{\alpha\beta}\overset{(1)}{a_{i\alpha}}-\mu_{\alpha\beta}p_{\alpha\beta}\overset{(0)}{a_{i\alpha}}\right)+o(1), \quad (3.2.1)$$

and $c_{\alpha\beta\gamma\delta}(i,s)$ can be found using $m_{\alpha\beta}(i,s)$ and (1.5.20). We now turn to the joint and marginal distributions of the $N_{ij}(t)$ and the initial state.

Up to now it has been assumed that N(t) is an r.v. with P_0, i, and j (transition probabilities, initial state, and final state, respectively) fixed. Since j was summed out in (1.5.13) and (1.5.14), we have only N(t) and Z_0, the initial state. Now, suppose Z_0 is an r.v. with values 1,...,m and corresponding probabilities $a_1, a_2, \ldots, a_m$ such that $\sum_{i=1}^{m} a_i = 1$. Let us find the joint distribution of N(t) and Z_0.

We may write

$$T_i(N,t) = P\{N(t)=N, Z_0=i\} = P\{N(t)=N \mid Z_0=i\} \cdot P\{Z_0 = i\}$$

$$= \begin{cases} a_i W_i(N,t), \quad i = 1,\ldots,m, \quad N\varepsilon\Phi_m(i,P_0) \\ 0, \qquad \text{otherwise} \end{cases} \quad , \qquad (3.2.2)$$

where W_i and Φ_m are as defined in section I.5. Since a_i is functionally independent of N and t, then the L.-S.T. p.g.f. of $T_i(N,t)$ is simply

$$t_i(\Xi,s) = a_i w_i(\Xi,s) \ , \qquad (3.2.3)$$

where $w_i(\Xi,s)$ is the i^{th} element of $\underline{w}(\Xi,s)$ in (1.5.13). Now, from (1.5.14), we have that the L.-S.T. of $T_i(N,t)$ is

$$\begin{cases} a_i \sum_{j=1}^{m} N_{ji}{}^{*} \dfrac{\prod_{\alpha=1}^{m} N_{\alpha\cdot}!}{\prod_{\alpha,\beta=1}^{m} N_{\alpha\beta}!} \cdot \prod_{\alpha,\beta=1}^{m} \{q_{\alpha\beta}(s)\}^{N_{\alpha\beta}} \{1-h_\alpha(s)\}, \quad N\varepsilon\Phi_m(i,P_0) \\ 0, \qquad \text{otherwise.} \end{cases} \qquad (3.2.4)$$

It is natural to consider here the marginal distributions of Z_0 and N(t). First, we know, since $T_i(N,t)$ is a d.f. for $N\varepsilon\Phi_m(i,P_0)$ that $\sum_{N\varepsilon\Phi_m(i,P_0)} T_i(N,t) = 1$; therefore,

$$\sum_{N\varepsilon\Phi_m(i,P_0)} a_i T_i(N,t) = a_i \ , \quad i = 1, \ldots, m \ , \qquad (3.2.5)$$

and, since $a_i > 0$ for every i, and $\sum_{i=1}^{m} a_i = 1$, then $\underline{a}' = (a_1,\ldots,a_m)$ is the marginal distribution of Z_0, as we would have expected.

To find the marginal distribution of N(t), we must first consider the following sets. Let

$$\Phi_m(P_0) = \bigcup_{i=1}^{m} \Phi_m(i,P_0) \ ,$$

and partition it into sets

$$\Phi_{m1}(P_0) = \{N|N\epsilon\Phi_m(P_0);n_{i.}=n_{.i}\} \text{ and } \Phi_{m2}(P_0) = \Phi_m(P_0) - \Phi_{m1}(P_0).$$

It is clear that $\Phi_m(P_0)$ is the set of all possible transition counts arising from an M.R.P. with transition matrix P_0. $\Phi_{m1}(P_0)$ consists of all such sets except that the system starts and ends in the same state. $\Phi_{m2}(P_0)$ is obviously that set of counts in $\Phi_m(P_0)$ which do not fall into $\Phi_{m1}(P_0)$.

From Martin's Lemma 6.1.5 (p. 128 of [22]) there exist m pairs $(x,y) = (i,i)$ which satisfy $n_{k.}-n_{.k} = \delta_{kx}-\delta_{ky}$, for $N\epsilon\Phi_{m1}(P_0)$. Also, there exists only one pair $(x,y) = (i,j)$ which satisfies the above equation for $N\epsilon\Phi_{m2}(P_0)$. So,

$$T(N,t) = \sum_{i=1}^{m} T_i(N,t) = \begin{cases} \sum_{\ell=1}^{m} a_\ell W_\ell(N,t), & N\epsilon\Phi_{m1}(P_0), \text{ (i.e., } i=j) \\ a_i W_i(N,t), & N\epsilon\Phi_{m2}(P_0) \\ 0, & \text{otherwise} . \end{cases} \tag{3.2.6}$$

The L.-S.T. p.g.f. of T(N,t) is

$$t(\Xi,s) = \delta_{ij} \sum_{\ell=1}^{m} a_\ell w_i(\Xi,s) + (1-\delta_{ij})a_i w_i(\Xi,s) . \tag{3.2.7}$$

Then from (3.2.4), the L.-S.T. of T(N,t) is

$$\delta_{ij} \sum_{\ell=1}^{m} a_\ell \sum_{j=1}^{m} N^*_{j\ell} \frac{\prod_{\alpha=1}^{m} N_{\alpha.}!}{\prod_{\alpha,\beta=1}^{m} N_{\alpha\beta}!} \cdot \prod_{\alpha,\beta=1}^{m} \{q_{\alpha\beta}(s)\}^{N_{\alpha\beta}}\{1-h_\alpha(s)\} + $$
$$+ (1-\delta_{ij})a_i \sum_{j=1}^{m} N^*_{ji} \frac{\prod_{\alpha=1}^{m} N_{\alpha.}!}{\prod_{\alpha,\beta=1}^{m} N_{\alpha\beta}!} \prod_{\alpha,\beta=1}^{m} \{q_{\alpha\beta}(s)\}^{N_{\alpha\beta}}\{1-h_\alpha(s)\} . \tag{3.2.8}$$

The moments of T(N,t) may be computed as

$$E[N_{\alpha\beta}(t)|P_0] = \sum_{i=1}^{m} a_i \, E[N_{\alpha\beta}(t)|Z_0=i] \ , \quad \text{and}$$

$$\text{L.-S.T.}\{E[N_{\alpha\beta}(t)|P_0]\} = \sum_{i=1}^{m} a_i \, m_{\alpha\beta}(i,s) \ , \qquad (3.2.9)$$

from (1.5.15) and (3.2.6). Furthermore, using (1.5.20), we have

$$\text{L.-S.T.}\{E[N_{\alpha\beta}(t)N_{\gamma\delta}(t)|P_0] = \sum_{i=1}^{m} a_i \, c_{\alpha\beta\gamma\delta}(i,s) \ . \qquad (3.2.10)$$

As before these quantities may be expanded using (3.2.1) and inverted to give asymptotic values of these moments for large t.

3. The Multivariate and Matrix Beta Distributions and Extended Natural Conjugates

The material discussed in this section is contained in section 6.2 of Martin [22] and will be condensed here for the sake of brevity. Only those results required for later use will be included, and the proofs of theorems, all contained in Martin, will be omitted.

The random stochastic vector $\underline{p} = (p_1,\ldots,p_m)$ has the multivariate beta distribution with parameter $\underline{\nu}$ if $\underline{p}$ has the joint density function

$$f_B^{(m)}(\underline{p}|\underline{\nu}) = \begin{cases} B_m(\underline{\nu}) \prod_{i=1}^{m-1} p_i^{\nu_i - 1} \left(1 - \sum_{k=1}^{m-1} p_k\right)^{\nu_m - 1} , & \text{for } \underline{p} \text{ stochastic} \\ 0, & \text{otherwise} . \end{cases} \qquad (3.3.1)$$

The parameter vector $\underline{\nu}' = (\nu_1,\ldots,\nu_m)$ has $\nu_i > 0$, $(i=1,\ldots,m)$, and the constant $B_m(\underline{\nu})$ is a multivariate generalization of the beta function such that

$$B_m(\underline{\nu}) = \frac{\Gamma\left(\sum_{i=1}^{m} \nu_i\right)}{\prod_{i=1}^{m} \Gamma(\nu_i)} \ .$$

Martin's theorem 6.2.2 states that (3.3.1) is a proper density function, and his theorem 6.2.3 gives a partitioning of $\underline{p}$ into $\underline{q} = (p_1,\ldots,p_\mu,1-\sum_{k=1}^{\mu} p_k)$ and $\underline{v}' = (v_1,\ldots,v_{\gamma-1}, 1-\sum_{i=1}^{\gamma-1} v_i)$ with $\mu + \gamma = m$ and $v_i = \dfrac{p_{\mu+i}}{1-\sum_{k=1}^{\mu} p_k}$ such that

$$D(\underline{q},\underline{v}|\underline{\nu}_1,\underline{\nu}_2) = f_\beta^{(\mu+1)}(\underline{q}|\underline{\nu}_1)f_\beta^\gamma(\underline{v}|\underline{\nu}_2) , \qquad (3.3.2)$$

where $\underline{\nu}'_1 = (\nu_1,\ldots, \nu_\mu, \sum_{k=1}^{\gamma} \nu_{\mu+k})$, and $\underline{\nu}'_2 = (\nu_{\mu+1},\ldots,\nu_m)$.

The nonstandard multivariate beta distribution has density

$$f_\beta^*(\underline{q}|c,\underline{\nu}) = \begin{cases} \dfrac{B_m(\underline{\nu})}{c^{(\Sigma\nu_i)-1}} \prod_{i=1}^{m-1} q_i^{\nu_i-1}\left[c - \sum_{k=1}^{m-1} q_k\right]^{\nu_m-1} , & \text{if } \underline{q}\varepsilon R_m(c) \\ 0, & \text{otherwise} , \end{cases} \qquad (3.3.3)$$

where $R_m(c) = \{\underline{q}|0 \leq q_i \leq c, \Sigma q_i = c\}$. (3.3.3) is obtained by substituting $\underline{q} = c\underline{p}$ into (3.3.1). Again, we may partition $\underline{p}$ into components as before and consider certain marginal distributions, but the method is the same, and the details will be omitted for the present.

The moments of the multivariate beta distribution may be summarized as follows:

$$E[\underline{p}] = \frac{\underline{\nu}}{\sum_{i=1}^{m} \nu_i} , \quad \text{and}$$

$$\mathrm{Var}\ [\underline{p}] = \frac{1}{\left[\sum_{i=1}^{m} \nu_i\right]^2 \left[\sum_{i=1}^{m} (\nu_i)+1\right]} \left[\sum_{i=1}^{m} \nu_i\ \Lambda + \underline{\nu}\ \underline{\nu}'\right] , \qquad (3.3.4)$$

where $\Lambda = \mathrm{diag}[\nu_i]$. For the nonstandard multivariate beta, the moments are given by

$$E[\underline{q}] = \frac{c}{\Sigma\nu_i} \cdot \underline{\nu} ,$$

and

$$\mathrm{Var}[\underline{q}] = \frac{c^2}{[\Sigma\nu_i]^2[\Sigma(\nu_i)+1]} [\Sigma\nu_i\Lambda - \underline{\nu}\ \underline{\nu}'] . \qquad (3.3.5)$$

The matrix beta distribution may be characterized by the density function

$$f_{M\beta}^{(K,m)}(P|N) = \begin{cases} \prod_{i,j=1}^{m} \prod_{k=1}^{K_i} B_m(\underline{\nu}_i^k)(p_{ij}^k)^{\nu_{ij}^k - 1} , & \text{if } P \text{ is a } K \times m \text{ stochastic matrix} \\ 0, & \text{otherwise} \quad . \end{cases} \tag{3.3.6}$$

The $K \times m$ random generalized stochastic matrix $P = [p_{ij}^k]$ is the type of matrix encountered as a transition matrix in an M.R.P., as P_0 in previous sections. The parameter N is a $K \times m$ matrix such that $\nu_{ij}^k > 0$, $k=1,\ldots,K_i$, $i,j=1,\ldots,m$, $B_m(\underline{\nu}_i^k)$ is the generalized beta function defined above, $\underline{\nu}_i^k$ is the generic row of N, and the total number of rows of both P and N is $K = \sum_{i=1}^{m} K_i$. It should be noted here that for the most part K_i will be unity and K will be m when an M.R.P. of m states is being analyzed. Incidentally, Martin shows that (3.3.6) is the product of K multivariate beta density functions, and from this it follows that (3.3.6) is a proper normalized density.

The moments of the matrix beta distribution (Martin, theorem 6.3.1) are

$$E[p_{ij}^k] = \frac{\nu_{ij}^k}{\sum_j \nu_{ij}^k} , \quad k = 1, \ldots, m; \; i,j = 1, \ldots, m ,$$

$$\mathrm{Var}[p_{ij}^k] = \frac{\nu_{ij}^k \left[\sum_j (\nu_{ij}^k) - \nu_{ij}^k\right]}{\left[\sum_j \nu_{ij}^k\right]^2 \left[\sum_j (\nu_{ij}^k) + 1\right]} , \; k=1, \ldots, m; \; i,j=1, \ldots, m , \tag{3.3.7}$$

$$\mathrm{Cov}[p_{\alpha\beta}^k, p_{\delta\varepsilon}^j] = \begin{cases} \dfrac{-\nu_{\alpha\beta}^k \nu_{\alpha\varepsilon}^k}{\left[\sum_{j=1}^{m} \nu_{\alpha j}^k\right]^2 \left[\sum_{j=1}^{m} (\nu_{\alpha j}^k) + 1\right]} , & j=k=1, \ldots, K_\alpha; \; \alpha=\delta=1, \ldots, m; \; \beta,\varepsilon = 1, \ldots, m, \; \beta \neq \varepsilon \\ 0, & j \neq k \text{ or } \alpha \neq \delta \; . \end{cases}$$

Higher moments may be computed using Martin's theorem 6.3.2 (p. 144). An analogue to the nonstandard multivariate beta distribution exists for the matrix beta distribution, but it will not be discussed here. It is useful in finding marginal and conditional distributions of submatrices of P, and, if needed for that purpose, it will be introduced later.

Martin shows in his section 2.2 that the matrix beta distribution is the natural conjugate prior distribution for the matrix P_0 of transition probabilities. This means that if one assumes that P_0 follows a matrix beta distribution prior to observing a Markov chain, then the distribution of P_0, given the observations, is also matrix beta. This property will be applied to M.R.P.'s as well. Another concept Martin utilizes is that of extended natural conjugate distributions for enlarged sets of prior parameters to give additional flexibility such as correlated rows in P_0. The basic ideas of his are contained in his theorem 6.4.1. Let S_{Km} be the set of all $K \times m$ stochastic matrices, and let

$$h(P|N,\omega) = \begin{cases} C(N,\omega) \prod_{i,j=1}^{m} \prod_{k=1}^{K_i} (p_{ij}^{k})^{\nu_{ij}^{k}-1} g(P|\omega) , & P \in S_{Km} \\ 0, & \text{otherwise} , \end{cases} \tag{3.3.8}$$

where $g(P|\omega)$ is a nonnegative measurable function positive on some subset of S_{Km}, and $C(N,\omega)$ is the normalizing constant for the density. Let $R_{ij}^{k}(N)$ be the matrix N with element ν_{ij}^{k} increased by unity. Then,

$$E[p_{ij}^{k}|N,\omega] = \frac{C(N,\omega)}{C(R_{ij}^{k}(N),\omega)} = \bar{p}_{ij}^{k}(N,\omega)$$

and (3.3.9)

$$\text{Cov}[p_{ij}^{\alpha},p_{k\ell}^{\beta}] = \bar{p}_{ij}^{\alpha}(N,\omega)[\bar{p}_{k\ell}^{\beta}\{R_{ij}^{\alpha}(N),\omega\}- \bar{p}_{k\ell}^{\beta}(N,\omega)]$$

$$\alpha,\beta = 1, \ldots, K_i; \quad i,j,k, = 1, \ldots, m .$$

Martin also discusses the case of a stationary probability vector

$$\underline{\pi}(P_0) = [\pi_1(P_0), \ldots, \pi_m(P_0)] .$$

Let $\underline{\gamma}' = (\gamma_1,\ldots,\gamma_m)$ be a vector of nonnegative integers. Then an extended natural conjugate distribution for the Markov chain case may be formed by letting

$$g(P_0|\underline{\gamma}) = \begin{cases} \prod_{i=1}^{m} [\pi_i(P_0)]^{\gamma_i} , & \text{for } P_0 \varepsilon S_m^* \\ 0, & \text{otherwise} , \end{cases}$$

where S_m^* is the set of $m \times m$ stochastic matrices possessing stationary probability vectors. Using the above function in (3.3.8), we get a matrix beta -1 distribution, namely,

$$f_{M\beta 1}^{(m)}(P_0|N,\underline{\gamma}) = W(N,\underline{\gamma}) \prod_{i,j=1}^{m} B_m(\nu_i)[\pi_i(P_0)]^{\gamma_i} p_{ij}^{\nu_{ij}-1} , \qquad (3.3.10)$$

where $\frac{1}{W(N,\underline{\gamma})} = \int_{S_m} \prod_{i=1}^{m} [\pi_i(P_0)]^{\gamma_i} f_{M\beta}^{(m,m)}(P_0|N)dP_0$. The moments of (3.3.10) are

$$E[p_{ij}|N,\underline{\gamma}] = \frac{\nu_{ij}}{\sum_j \nu_{ij}} \cdot \frac{W(N,\underline{\gamma})}{W[R_{ij}(N),\underline{\gamma}]}$$

and

$$E[p_{\alpha\beta}p_{\delta\varepsilon}|N,\underline{\gamma}] = \begin{cases} \frac{\nu_{\alpha\beta}(\nu_{\alpha\beta}+1)}{(\sum_\beta \nu_{\alpha\beta})(\sum_\beta \nu_{\alpha\beta}+1)} \cdot \frac{W(N,\underline{\gamma})}{W[T_{\alpha\beta}\{T_{\alpha\beta}(N,\underline{\gamma})\},\underline{\gamma}]} ; & \alpha=\delta, \ \beta=\varepsilon \\ \frac{\nu_{\alpha\beta}\nu_{\alpha\varepsilon}}{(\sum_\beta \nu_{\alpha\beta})(\sum_\beta \nu_{\alpha\beta}+1)} \cdot \frac{W(N,\underline{\gamma})}{W[T_{\alpha\varepsilon}\{T_{\alpha\beta}(N,\underline{\gamma})\},\underline{\gamma}]} ; & \beta\neq\varepsilon \\ \frac{\nu_{\alpha\beta}\nu_{\delta\varepsilon}}{(\sum_\beta \nu_{\alpha\beta})(\sum_\varepsilon \nu_{\delta\varepsilon})} \cdot \frac{W(N,\underline{\gamma})}{W[T_{\delta\varepsilon}\{T_{\alpha\beta}(N,\underline{\gamma})\},\underline{\gamma}]} ; & \alpha\neq\delta . \end{cases} \qquad (3.3.11)$$

4. Unconditional Distributions of the Transition Counts

In this section expressions will be given for the unconditional distribution of the transition count of an M.R.P., assuming that the transition matrix P_0 obeys a matrix beta distribution for a fixed initial state. Then we will assume that the initial state follows the distribution

$\underline{a}' = (a_1,\ldots,a_m)$ as in section III.2, and we will give an expression for the distribution of N(t) unconditional of the initial state.

For fixed i and j let

$$\Phi_m(i,j) = \{N | n_{ij} \geq 0, \text{ an integer}; \quad n_{k\cdot} - n_{\cdot k} = \delta_{ki} - \delta_{kj}\} \quad , \qquad (3.4.1)$$

and let

$$\Phi_m(i) = \bigcup_{j=1}^{m} \Phi_m(i,j) \ .$$

Clearly, $\Phi_m(i)$ is the set of all possible transition counts arising from an M.R.P. with positive transition probability matrix, starting from state i. Then the desired distribution is

$$P\{N(t)=N | i,\mathcal{N}\} = D_i(N,t,\mathcal{N}) = \begin{cases} \int_{S_m} W_i(N,t) f_{M\beta}^{(m,m)}(P_0|\mathcal{N})\, d P_0 \ , & N \varepsilon \Phi_m(i) \\ 0 \ , & \text{otherwise} \ , \end{cases} \qquad (3.4.2)$$

where i = 1, ..., m and $\mathcal{N} = [\nu_{ij}]$ is an m × m matrix with all $\nu_{ij} > 0$.

By summing first over $N \varepsilon \Phi_m(i,P_0)$ and then integrating, we see that $D_i(N,t,\mathcal{N})$ is a proper probability mass function. To find the L.-S.T. p.g.f. of $D_i(N,t,\mathcal{N})$ we must take first

$$D_i(\Xi,t,\mathcal{N}) = \sum_{N \varepsilon \Phi_m(i)} \prod_{k,\ell} \xi_{k\ell}^{n_{k\ell}} D_i(N,t,\mathcal{N})$$

and then

$$d_i(\Xi,s,\mathcal{N}) = \int_0^\infty e^{-st}\, d_t D_i(\Xi,t,\mathcal{N}) \quad ,$$

which is obviously quite difficult to calculate in general. An attempt will be made in a later chapter to carry this out for a specific two-state M.R.P.

The assumption of a prior distribution for the initial state further complicates the above problem, since the above expression will have to be multiplied by the initial state distribution and summed. This will also be attempted in a later chapter.

CHAPTER IV

THE χ^2 STATISTIC FOR A TWO-STATE M.R.P.

In this chapter we specialize the results of Chapter II to consider the χ^2 goodness-of-fit statistic for a two-state M.R.P. We begin by discussing some basic results for two-state M.R.P.'s. Then the χ^2 statistic is given, and its mean and variance are computed for the ordinary M.R.P. Since the results turn out to be the same for the equilibrium M.R.P., as shown in Chapter II, we do not include them here. Finally, a numerical illustration is given for a realization of a particular two-state M.R.P.

1. Basic Results

For the two-state case we have

$$
\begin{aligned}
P_0 &= \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} a & 1-a \\ b & 1-b \end{bmatrix}, \text{ say,} \\
P_1 &= [c_{ij}] = [p_{ij}\mu_{ij}], \\
P_2 &= [d_{ij}] = [p_{ij}(\mu_{ij}^2 + \sigma_{ij}^2)],
\end{aligned}
\tag{4.1.1}
$$

where μ_{ij} and σ_{ij}^2 are the means and variances of the $F_{ij}(x)$, respectively. Then, following Kshirsagar and Gupta [17], we may write

$$
I-q(s) = \begin{bmatrix} 1 - a + sc_{11} - \frac{1}{2} s^2 d_{11} & a - 1 + sc_{12} - \frac{1}{2} s^2 d_{12} \\ -b + sc_{21} - \frac{1}{2} s^2 d_{21} & b + sc_{22} - \frac{1}{2} s^2 d_{22} \end{bmatrix}. \tag{4.1.2}
$$

By direct evaluation

$$\det[I-q(s)] = \frac{1}{\alpha} s + \beta s^2 \quad , \tag{4.1.3}$$

where $$\frac{1}{\alpha} = (1-a)\ (c_{21} + c_{22}) + b(c_{11} + c_{12}) \ ,$$

and $$\beta = \det(P_1) - \frac{1}{2}\{(1-a)(d_{21}+d_{22}) + b(d_{11}+d_{12})\}. \tag{4.1.4}$$

Expanding $\{1/\det[I-q(s)]\}$ in powers of s, evaluating $[I-q(s)]^{-1} - I$ and inverting, we get

$$M(t) = [M_{ij}(t)] = \begin{bmatrix} b\alpha t+(c_{22}\alpha-\alpha^2\beta b-1)+o(1) & (1-a)\alpha t-\alpha c_{21}-\alpha^2\beta(1-a)+o(1) \\ b\alpha t-(c_{12}\alpha+b\alpha^2\beta) + o(1) & (1-a)\alpha t+c_{11}\alpha-(1-a)\alpha^2\beta-1+o(1) \end{bmatrix} . \tag{4.1.5}$$

The second factorial moment of $N_j(t)$ for the two-state case was given by Kshirsagar and Gupta [17] as

$$R(t) = [R_{ij}(t)] = \begin{bmatrix} A_{11}t^2+2A_{12}t+2A_{13}+o(1) & A_{21}t^2+2A_{22}t+2A_{23}+o(1) \\ A_{31}t^2+2A_{32}t+2A_{33}+o(1) & A_{41}t^2+2A_{42}t+2A_{43}+o(1) \end{bmatrix}, \tag{4.1.6}$$

where the A_{ij} are functions of a, b, c_{ij}, d_{ij}, α, and β, also given in [17]. Using this and the technique described in section I.3, we may find the variances of the $N_j(t)$'s for m = 2, namely,

$$V_{11}(t) = t[2\alpha b(c_{22}\alpha - \alpha^2\beta b-1) + \alpha b] + o(t)$$

$$V_{12}(t) = t(1-a)\alpha[2\alpha(c_{11}-c_{12}) - 2\alpha^2\beta(1-a) + 2\alpha c_{21} -1] + o(t)$$

$$V_{21}(t) = t(\alpha b)[2\alpha(c_{22}-c_{21}) - 2\alpha^2 b\beta + 2c_{12}\alpha -1] + o(t) \tag{4.1.7}$$

$$V_{22}(t) = t(1-a)\alpha[2\{c_{11}\alpha - \alpha^2\beta(1-a) -1\} +1] + o(t) \ .$$

Now, as in section I.6, we may find the covariances of the $N_j(t)$ for the two-state case. First, since $\Gamma^i_{12}(t) = \Gamma^i_{21}(t)$ for fixed i, then we need find only the two quantities $\Gamma^1_{12}(t)$ and $\Gamma^2_{12}(t)$, say. Using the fact that $c^i_{jk}(s) = m_{ik}(s)m_{kj}(s) + m_{ij}(s)m_{jk}(s)$, we have

$$\Gamma^1_{12}(t) = \Gamma^1_{21}(t) = t[\alpha(1-a)\{-\alpha c_{12} - \alpha^2\beta b\} - \alpha b\{\alpha(2c_{12}-c_{21})-\alpha^2\beta(1-a)\}]+o(t)$$
$$\Gamma^2_{12}(t) = \Gamma^2_{21}(t) = t[\alpha b\{-\alpha c_{12} - \alpha^2\beta(1-a)-1\} - \alpha(1-a)\{\alpha c_{12}+\alpha^2 b\beta\}] + o(t). \quad (4.1.8)$$

These quantities may now be used to compute the χ^2 statistic, its mean and its variance.

2. The χ^2 Statistic and its First Two Moments

The statistic of Chapter II for the m-state M.R.P. reduces to

$$\chi^2 = \frac{[N_1(t)-M_{i1}(t)]^2}{M_{i1}(t)} + \frac{[N_2(t)-M_{i2}(t)]^2}{M_{i2}(t)} \quad (4.2.1)$$

with

$$E(\chi^2) = \frac{V_{i1}(t)}{M_{i1}(t)} + \frac{V_{i2}(t)}{M_{i2}(t)} \quad (4.2.2)$$

and

$$Var(\chi^2) = 2\left\{\left[\frac{V_{i1}(t)}{M_{i1}(t)}\right]^2 + \left[\frac{V_{i2}(t)}{M_{i2}(t)}\right]^2 + 2\,\frac{[\Gamma^i_{12}(t)]^2}{M_{i1}(t)M_{i2}(t)}\right\} \quad (4.2.3)$$

for a given initial state i. Substituting the quantities from (4.1.5), (4.1.7), and (4.1.8) into (4.2.2) and (4.2.3) we have

$$E(\chi^2) = 2\alpha(c_{11}-c_{12}+c_{21}+c_{22}) - 2\alpha^2\beta[(1-a)+b] - 2 + o(1), \quad (4.2.4)$$

$$Var(\chi^2) = [2c_{22}\alpha-2\alpha^2\beta b-1]^2 + [2c_{11}\alpha-2\alpha^2\beta(1-a)-1]^2 + 2[\tfrac{1-a}{b}\{-\alpha c_{12}-\alpha^2\beta b\}^2 +$$
$$+ 2\{\alpha c_{12}+\alpha^2\beta b\}\{\alpha(2c_{12}-c_{21})+\alpha^2\beta(1-a)\} + \frac{b}{1-a}\{\alpha(2c_{12}-c_{21})+\alpha^2\beta(1-a)\}^2]$$
$$+ o(1), \quad (4.2.5)$$

for i, the initial state, equal to one, and

$$E(\chi^2) = 2\alpha(c_{11}+c_{12}-c_{21}+c_{22}) - 2\alpha^2\beta[(1-a)+b] - 2 + o(1), \quad (4.2.6)$$

$$\begin{aligned}\mathrm{Var}(\chi^2) = {} & [2\alpha(c_{22}-c_{21}+c_{12}) - 2\alpha^2\beta b-1]^2 + [2\alpha c_{11}-2\alpha^2\beta(1-a)-1]^2 + \\ & + 2\Big[\frac{b}{1-a}\{-\alpha c_{12}-\alpha^2\beta(1-a)-1\}^2 + 2\{\alpha c_{12}+\alpha^2\beta(1-a)+1\}\{\alpha c_{12}+\alpha^2\beta b\} + \\ & + \frac{1-a}{b}\{\alpha c_{12}+\alpha^2\beta b\}^2\Big] + o(1) \quad , \qquad (4.2.7)\end{aligned}$$

when the initial state is state two.

3. A Numerical Illustration

In order to apply our test for a hypothetical P_0 matrix, we generated a sample of data from a two-state M.R.P. We wish to test the matrix

$$P_0 = \begin{bmatrix} .3 & .7 \\ .6 & .4 \end{bmatrix} ,$$

using a matrix of distribution functions of the form

$$\begin{bmatrix} 1 - e^{-\frac{x}{2}} & 1 - e^{-\left(\frac{x}{2}\right)^2} \\ 1 - e^{-\frac{x}{2}} & 1 - e^{-\left(\frac{x}{2}\right)^2} \end{bmatrix} .$$

For this case we have

$$P_1 = \begin{bmatrix} 0.600 & 1.239 \\ 1.200 & 0.708 \end{bmatrix} \text{ and } P_2 = \begin{bmatrix} 2.400 & 2.800 \\ 4.800 & 1.600 \end{bmatrix} ,$$

with a, b, c_{ij}, and d_{ij} as the appropriate elements of P_0, P_1, and P_2 as defined in section 1 above. Using (4.1.4), we have $\frac{1}{\alpha} = 2.439$ and $\beta = -4.862$.

The process was initially in state 1 and was observed for $t = 80$ units of time. The observed states of the system and times spent in the states are:

(1,2.693), (2,1.954), (1,1.073), (2,0.914), (2,2.506), (1,4.229),
(1,3.212), (1,0.309), (2,1.582), (1,2.090), (2,2.512), (1,1.347),
(2,3.987), (1,2.412), (2,2.138), (2,3.108), (1,11.800), (2,1.487),
(2,1.963), (1,7.017), (2,2.906), (1,5.594), (2,1.023), (2,2.190),
(1,0.965), (2,2.068), (1,2.817), (2,0.810), (1,0.438), (1,3.226).

From this we see that $N_1(t) = 15$ and $N_2(t) = 15$. Now, using (4.2.1) and (4.1.5) to calculate the statistic; and (4.2.4) and (4.2.5) to calculate its moments, we obtain the following values for our special case:

Observed χ^2	$A = E(\chi^2)$	$2B = Var(\chi^2)$	Modified $\chi'^2 = A\chi^2/B$	D.F. $= A^2/B$
3.828	1.165	1.368	6.52	1.98

Compared to a standard χ^2 with two degrees of freedom at the .01 level, this statistic is not significant and would indicate as far as one may rely on this test that there is no reason to doubt the fit of the hypothetical P_0 matrix.

CHAPTER V

BAYESIAN ANALYSIS OF A TWO-STATE M.R.P.

In this chapter we will consider some prior-posterior and preposterior analysis of a two-state M.R.P. in much the same manner as Martin has done for two-state Markov chains. We begin by giving some preliminary remarks on the 2×2 P_0 matrix, its properties, and its prior distribution, the 2×2 matrix beta distribution. Secondly, we discuss prior-posterior and preposterior analysis of a two-state M.R.P. with known initial state. Finally, we consider the case in which the initial state is unknown but has a prior distribution itself.

1. Preliminaries

We assume that the transition probability matrix P_0 is an unknown, fixed matrix of the form

$$P_0 = \begin{bmatrix} 1-x & x \\ y & 1-y \end{bmatrix}, \quad \text{for } 0 \leq x,\ y \leq 1 . \qquad (5.1.1)$$

It can easily be shown that the latent roots of P_0 are $\lambda_1 = 1$ and $\lambda_2' = 1-x-y$, and thus P_0 has the spectral decomposition:

$$P_0 = \begin{bmatrix} \frac{y}{x+y} & \frac{x}{x+y} \\ \frac{y}{x+y} & \frac{x}{x+y} \end{bmatrix} + (1-x-y) \begin{bmatrix} \frac{x}{x+y} & \frac{-x}{x+y} \\ \frac{-y}{x+y} & \frac{y}{x+y} \end{bmatrix}, \quad x,y \neq 0 . \qquad (5.1.2)$$

These two equations immediately yield the stationary state probability vector

$$\underline{U}' = \begin{bmatrix} \frac{y}{x+y}, & \frac{x}{x+y} \end{bmatrix} , \quad x,y \neq 0 . \tag{5.1.3}$$

For the prior-posterior and preposterior analysis we wish to carry out in this chapter we again assume that P_0 follows a matrix beta prior distribution. If we denote the parameter matrix by

$$N = \begin{bmatrix} \nu_{11} & \nu_{12} \\ \nu_{21} & \nu_{22} \end{bmatrix} , \tag{5.1.4}$$

then we may write the 2 × 2 matrix beta density function as

$$f_{M\beta}^{(2,2)}(P_0|N) = \frac{1}{B(\nu_{11},\nu_{12})B(\nu_{21},\nu_{22})} \; x^{\nu_{12}-1}(1-x)^{\nu_{11}-1} y^{\nu_{21}-1}(1-y)^{\nu_{22}-1} . \tag{5.1.5}$$

Thus x and y are independent random variables, each distributed according to a univariate beta distribution. The set S_2 of 2 × 2 stochastic matrices is now

$$S_2 = \{x,y \mid 0 \leq x, y \leq 1\} , \tag{5.1.6}$$

and this will form the set over which we evaluate the integrals in the following sections.

2. Initial State Known

In this section and the rest of this chapter the notation of Moore and Pyke [23] will be used to denote the density of the two-state M.R.P. being considered. As before, the M.R.P. is observed for a fixed interval of time (0,t), and two things may happen. First, the system may not make any transition at all, and then it will stay in the initial state, say i.

This occurence will be denoted by the symbols $v = i$, where v represents the sample of observations. The other possibility is that there will be one or more transitions between the two states. This sample will be denoted by $v = (j_0, j_1, \ldots, j_n, x_1, x_2, \ldots, x_n)$, where j_k is the state of the system immediately after the k^{th} transition, x_k is the holding time in state j_{k-1}, and n is the number of transitions observed in $(0,t)$. Finally, following Moore and Pyke [23], we assume that the $F_{ij}(t)$ are of the form $F_i(t)$ with density $f_i(t)$, since Pyke and Schaufele [28] point out that this assumption incurs no loss of generality.

To consider prior-posterior analysis, suppose that a sample from an M.R.P. is available, for which we may write (as in Moore and Pyke [23]) the conditional probability, given P_0, of observing the sample as

$$L(v|P_0) = \begin{cases} 1 - F_i(t), & \text{if } v = i \\ [1-F_j(u_t)]\left[\prod_{k=0}^{n-1} f_{j_k}(x_{k+1})\right]\left[x^{N_{12}(t)}(1-x)^{N_{11}(t)}y^{N_{21}(t)}(1-y)^{N_{22}(t)}\right], & \text{if } v = (j_0,\ldots,j_n, x_1,\ldots,x_n) \\ 0, & \text{otherwise,} \end{cases} \tag{5.2.1}$$

where $u_t = t-x_1-x_2-\ldots-x_n$, j is the final state of the system at time t, and the $N_{k\ell}(t)$ are the observed number of transitions from state k to ℓ as defined in section I.5. Now, denoting the unconditional distribution of v by $L_1(v)$, we have

$$L_1(v) = [1-F_i(t)] \int_0^1\int_0^1 \frac{1}{B(\nu_{11},\nu_{12})B(\nu_{21},\nu_{22})} x^{\nu_{12}-1}(1-x)^{\nu_{11}-1}y^{\nu_{21}-1}(1-y)^{\nu_{22}-1}\,dxdy$$

$$= 1 - F_i(t), \quad \text{for } v = i. \tag{5.2.2}$$

Now, if $v = (j_0, \ldots, j_n, x_1, \ldots, x_n)$,

$$L_1(v) = \frac{[1-F_j(u_t)]\prod_{k=0}^{n-1} f_{j_k}(x_{k+1})}{B(\nu_{11},\nu_{12})B(\nu_{21},\nu_{22})} \int_0^1\int_0^1 x^{N_{12}(t)+\nu_{12}-1}(1-x)^{N_{11}(t)+\nu_{11}-1}$$
$$\times\, y^{N_{21}(t)+\nu_{21}-1}(1-y)^{N_{22}(t)+\nu_{22}-1}\,dxdy$$

$$= \frac{[1-F_j(u_t)]\prod_{k=0}^{n-1} f_{j_k}(x_{k+1})}{B(\nu_{11},\nu_{12})B(\nu_{21},\nu_{22})} \cdot B(N_{11}(t)+\nu_{11},N_{12}(t)+\nu_{12})$$
$$\times\, B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}). \tag{5.2.3}$$

Then, the posterior distribution of P_0, given the sample, may be found by calculating

$$L_2(P_0|v) = \frac{L(v|P_0)f_{M\beta}^{(2,2)}(P_0|N)}{L_1(v)} \tag{5.2.4}$$

$$= \begin{cases} \dfrac{1}{B(\nu_{11},\nu_{12})B(\nu_{21},\nu_{22})}\, x^{\nu_{12}-1}(1-x)^{\nu_{11}-1} y^{\nu_{21}-1}(1-y)^{\nu_{22}-1}, & \text{if } v=i \\ \dfrac{1}{B(N_{11}(t)+\nu_{11},N_{12}(t)+\nu_{12})B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22})}\, x^{N_{12}(t)+\nu_{12}-1} & \\ \quad\times\,(1-x)^{N_{11}(t)+\nu_{11}-1} y^{N_{21}(t)+\nu_{21}-1}(1-y)^{N_{22}(t)+\nu_{22}-1}, & \text{if} \\ & v = (j_0,\ldots,j_n,\ x_1,\ldots,x_n). \end{cases}$$

Both of these posterior distributions are matrix beta, but with different parameters. Thus the matrix beta is the conjugate prior for a sample from an M.R.P., and it may be used to find any required Bayes rules.

To perform preposterior analysis for this case we note that prior to observing the M.R.P., the transition count matrix $N(t)$ is a random matrix with distribution (1.5.4), given the initial state i, the final state j, t, and P_0. Since we cannot solve explicitly for $W_{ij}(N,t)$, we must consider its L.-S.T., and the L.-S.T. of $W_i(N,t)$, summing out the final state. The

L.-S.T. of the unconditional distribution of the transition count is given by

$$\frac{1}{B(\nu_{11},\nu_{12})B(\nu_{21},\nu_{22})}\int_0^\infty e^{-st}\left[\int_0^1\int_0^1 W_i(N,t)x^{\nu_{12}-1}(1-x)^{\nu_{11}-1}y^{\nu_{21}-1}(1-y)^{\nu_{22}-1}dxdy\right]dt,$$

$$= \frac{1}{B(\nu_{11},\nu_{12})B(\nu_{21},\nu_{22})}\int_0^1\int_0^1 \text{L.-S.T.}\{W_i(N,t)\}x^{\nu_{12}-1}(1-x)^{\nu_{11}-1}y^{\nu_{21}-1}(1-y)^{\nu_{22}-1}dxdy, \quad (5.2.5)$$

since $W_i(N,t)$ is the only function of t involved. To evaluate (5.2.5), we substitute from (1.5.14) the L.-S.T. of $W_i(N,t)$. First note, however, that Martin ([22], p. 119) shows that the only contribution from the sum (1.5.14) comes when $j = k$, where k is the solution to (1.5.3). Therefore, if $j = k = 1$, (5.2.5) becomes

$$N^*_{1i}K_w\int_0^1\int_0^1 x^{N_{12}(t)+\nu_{12}-1}(1-x)^{N_{11}(t)+\nu_{11}-1}y^{N_{21}(t)+\nu_{21}-1}(1-y)^{N_{22}(t)+\nu_{22}-1}dxdy -$$

$$-N^*_{1i}K_w f_{11}(s)\int_0^1\int_0^1 x^{N_{12}(t)+\nu_{12}-1}(1-x)^{N_{11}(t)+\nu_{11}}y^{N_{21}(t)+\nu_{21}-1}(1-y)^{N_{22}(t)+\nu_{22}-1}dxdy-$$

$$-N^*_{1i}K_w f_{12}(s)\int_0^1\int_0^1 x^{N_{12}(t)+\nu_{12}}(1-x)^{N_{11}(t)+\nu_{11}-1}y^{N_{21}(t)+\nu_{21}-1}(1-y)^{N_{22}(t)+\nu_{22}-1}dxdy$$

$$= N^*_{1i}K_w B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}) -$$

$$- N^*_{1i}K_w f_{11}(s)B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11}+1)B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}) -$$

$$- N^*_{1i}K_w f_{12}(s)B(N_{12}(t)+\nu_{12}+1,N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}), \quad (5.2.6)$$

where $K_w = \dfrac{1}{B(\nu_{11},\nu_{12})B(\nu_{21},\nu_{22})}\;\dfrac{N_1.!N_2.!}{N_{11}!N_{12}!N_{21}!N_{22}!}\prod_{\alpha,\beta=1}^{2}[f_{\alpha\beta}(s)]^{N_{\alpha\beta}}$, and N^*_{1i} is as defined in section I.5. Now, if $j = k = 2$, then, using the same technique, (5.2.5) becomes

$$N^*_{2i}K_w B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}) -$$

$$- N^*_{2i}K_w f_{21}(s)B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21}+1,N_{22}(t)+\nu_{22}) - \quad (5.2.7)$$

$$- N^*_{2i}K_w f_{22}(s)B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}+1) ,$$

where N^*_{2i} is also as in section I.5. These distributions may then be used to carry out preposterior analysis for many types of utility functions.

3. Initial State Unknown

We now assume that the initial state i is unknown but has a prior probability distribution $\underline{a}' = (a_1,a_2)$, where $a_k = P\{J_0=k\}$, k = 1, 2, which is functionally independent of P_0. For prior-posterior analysis we have that the distribution of a sample from an M.R.P., given P_0, corresponding to (5.2.1) (again from Moore and Pyke [23]), is

$$L(v|P_0) = \begin{cases} a_i[1-F_i(t)] , & \text{if } v = i \\ a_i[1-F_j(u_t)]\left[\prod_{k=0}^{n-1} f_{j_k}(x_{k+1})\right]\left[x^{N_{12}(t)}(1-x)^{N_{11}(t)}y^{N_{21}(t)}(1-y)^{N_{22}(t)}\right], & \text{if } v = (j_0,\ldots,j_n, x_1,\ldots,x_n) \\ 0, & \text{otherwise.} \end{cases} \quad (5.3.1)$$

Then the marginal distribution of the sample is

$$L_1(v) = \begin{cases} a_i[1 - F_i(t)] , & \text{if } v = i \\ \dfrac{a_i[1-F_j(u_t)] \prod_{k=0}^{n-1} f_{j_k}(x_{k+1})}{B(\nu_{11},\nu_{12})B(\nu_{21},\nu_{22})} \cdot B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11}) \\ \quad \times B(N_{21}(t)+\nu_{21}, N_{22}(t)+\nu_{22}), & \text{if } v=(j_0,\ldots,j_n,x_1,\ldots,x_n). \end{cases} \quad (5.3.2)$$

Thus, the posterior distribution of P_0, given the sample is

$$L_2(P_0|v) = \begin{cases} \frac{1}{B(\nu_{11},\nu_{12})B(\nu_{21},\nu_{22})} x^{\nu_{12}-1}(1-x)^{\nu_{11}-1} y^{\nu_{21}-1}(1-y)^{\nu_{22}-1}, & \text{if } v=i \\ \frac{1}{B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22})} x^{N_{12}(t)+\nu_{12}-1} & \\ \quad \times (1-x)^{N_{11}(t)+\nu_{11}-1} y^{N_{21}(t)+\nu_{21}-1} (1-y)^{N_{22}(t)+\nu_{22}-1}, & \\ & \text{if } v = (j_0,\ldots,j_n,x_1,\ldots,x_n). \end{cases} \quad (5.3.3)$$

These are again both matrix beta posterior distributions as in the case of a known initial state, and they may be used to carry out many types of prior-posterior analysis.

The results for preposterior analysis are the same as those in section 2 except that each distribution is multiplied by the initial probability a_i. They correspond to the results given in Chapter III for the m-state case.

Specifically, in this case the L.-S.T. of the unconditional distribution of the transition count and the initial state is given by

$$a_i N^*_{1i} K_w B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}) -$$
$$- a_i N^*_{1i} K_w f_{11}(s) B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11}+1)B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}) -$$
$$- a_i N^*_{1i} K_w f_{12}(s) B(N_{12}(t)+\nu_{12}+1,N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}) \; , \quad (5.3.4)$$

if state 1 is the solution to (1.5.3). Now, if state 2 is the solution to (1.5.3), we have for the L.-S.T. of the unconditional distribution of the transition count and the initial state

$$a_i N^*_{2i} K_w B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}) -$$
$$- a_i N^*_{2i} K_w f_{21}(s) B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21}+1,N_{22}(t)+\nu_{22}) -$$
$$- a_i N^*_{2i} K_w f_{22}(s) B(N_{12}(t)+\nu_{12},N_{11}(t)+\nu_{11})B(N_{21}(t)+\nu_{21},N_{22}(t)+\nu_{22}+1) \quad (5.3.5)$$

Again these distributions may be used to perform preposterior analysis for many types of utility functions.

CHAPTER VI

SUMMARY AND FURTHER RESEARCH

This paper is intended to present some new results in statistical inference for Markov Renewal Processes. To do this in a manner which makes the presentation self-contained, we have given in the first chapter those previous results which were required for use in the later chapters. These include the renewal process results of Cox [8], and the M.R.P. results of Pyke [27], Kshirsagar and Gupta([17], [18], and [19]), Çinlar [6], and Kshirsagar and Wysocki [20].

In Chapter II we develop a χ^2 goodness-of-fit test for a hypothetical M.R.P. model, using the $N_j(t)$'s as observations. After calculating the first two moments of our test statistic, we modify it to make its first two moments correspond exactly to those of a standard χ^2, both for an ordinary M.R.P. and for an equilibrium M.R.P.

We discuss Bayesian analysis of an M.R.P. in Chapter III, when the transition probability matrix P_0 is assumed to have a matrix beta prior distribution. This includes new expressions for the joint distribution of the transition count and the initial state, and also some integral expressions for the unconditional distribution of the transition count.

Chapter IV contains the χ^2 goodness-of-fit statistic for a two-state M.R.P. The moments are calculated, and a numerical illustration is given for a particular two-state M.R.P.

In Chapter V we give a detailed discussion of the Bayesian analysis for a two-state M.R.P. This includes prior-posterior and preposterior analysis when the initial state of the system is known and when it is unknown but has a specified prior distribution. The results here correspond to what Martin [22] has obtained for Markov chains.

To give some ideas of the avenues of further research, let us mention here that few papers have been written on statistical inference for M.R.P.'s. It seems, then, that much more work could be done in this area. Specifically, one could develop a χ^2 statistic using the $N_{ij}(t)$'s as the observations much as we have done using the $N_j(t)$'s. Another thing to be considered might be the effect on the χ^2 of estimating the moments of the life distributions $F_{ij}(x)$. A third generalization might be to test several realizations of an M.R.P., developing results similar to those of Darwin [10] in Markov chains.

A possible extension of the Bayesian analysis given here might be to perform prior-posterior and preposterior analysis for an M.R.P. operating in equilibrium with a stationary state probability vector which is a function of P_0. Another problem would be to extend the results of Chapter V to systems with 3, 4, or more states. Finally, one might assume prior distributions for the parameters of the life distributions $F_{ij}(x)$ and carry out Bayesian analysis under those circumstances.

The list of unsolved problems in M.R.P.'s could go on to include such topics as statistical analysis of strategies of replacement, renewals in arbitrary intervals and their Poisson counts, and the superposition of several M.R.P.'s. One might consider more work in space Markov Renewal processes (see Çinlar [6]), limit theorems, or any of several other topics, perhaps even the extension of Markov chain theory to M.R.P.'s.

This list of future research problems is by no means complete, but it does indicate some of the new directions that may be taken by those interested in Markov Renewal processes.

LIST OF REFERENCES

[1] Bartlett, M. S. (1951). "The Frequency Goodness of Fit Test for Probability Chains," Proceedings of the Cambridge Philosophical Society, 47: 86-95.

[2] Bartlett, M. S. (1955). An Introduction to Stochastic Processes. Cambridge University Press, Cambridge, England.

[3] Bhat, B. R. (1961). "On the Asymptotic Distribution of the 'Psi-squared' Goodness of Fit Criteria for Markov Chains and Markov Sequences," Annals of Mathematical Statistics, 32: 49-58.

[4] Billingsley, P. (1961). "Statistical Methods in Markov Chains," Annals of Mathematical Statistics, 32: 12-40.

[5] Billingsley, P. (1961). "Corrections to 'Statistical Methods in Markov Chains'," Annals of Mathematical Statistics, 32: 1343.

[6] Çinlar, E. (1968). "Some Joint Distributions for Markov Renewal Processes," Australian Journal of Statistics, 10: 8-20.

[7] Çinlar, E. (1969). "Markov Renewal Theory," Advances in Applied Probability, 1: 123-187.

[8] Cox, D. R. (1962). Renewal Theory. Methuen and Co., Ltd., London, (page references correspond to Science Paperback edition, 1970).

[9] Cox, D. R., and Smith, W. L. (1954). "On the Superposition of Several Renewal Processes," Biometrika 41: 91-99.

[10] Darwin, J. H. (1959). "Note on the Comparison of Several Realizations of a Markoff Chain," Biometrika, 46: 412-419.

[11] Good, I. J. (1955). "The Likelihood Ratio Test for Markov Chains," Biometrika, 42: 531-533.

[12] Hunter, J. J. (1969). "On the Moments of Markov Renewal Processes," Advances in Applied Probability, 1: 188-210.

[13] Judge, G. G., Lee, T. C., and Zellner, A. (1968). "Maximum Likelihood and Bayesian Estimation of Transition Probabilities," Journal of the American Statistical Association, 63: 1162-1178.

[14] Keilson, J. (1969). "On the Matrix Renewal Function for Markov Renewal Processes," Annals of Mathematical Statistics, 40: 1901-1907.

[15] Kemeny, J. G. and Snell, J. L. (1960). Finite Markov Chains. D. Van Nostrand and Co., Ltd., New York.

[16] Kshirsagar, A. M. (1970). "Poisson Counts of a Markovian Renewal Process," South African Statistical Journal, 4: 67-72.

[17] Kshirsagar, A. M. and Gupta, Y. P. (1967). "Asymptotic Values of the First Two Moments in Markov Renewal Processes," Biometrika, 54: 597-603.

[18] Kshirsagar, A. M. and Gupta, Y. P. (1969). "Some Results in Markov Renewal Processes," Bulletin of the Calcutta Statistical Association, 18: 61-72.

[19] Kshirsagar, A. M. and Gupta, Y. P. (1971). "A Note on the Matrix Renewal Function," Journal of the Australian Mathematical Society, (to appear).

[20] Kshirsagar, A. M. and Wysocki, R. (1970). "Some Distribution and Moment Formulae for the Markov Renewal Process," Proceedings of the Cambridge Philosophical Society, 68: 159-166.

[21] Lévy, P. (1954). "Processus Semi-Markoviens," Proceedings of the International Congress on Mathematics, 3: 416-426.

[22] Martin, J. J. (1967). Bayesian Decision Problems and Markov Chains. John Wiley and Sons, New York.

[23] Moore, E. H. and Pyke, R. (1968). "Estimation of the Transition Distributions of a Markov Renewal Process," Annals of the Institute of Statistical Mathematics, 20: 411-424.

[24] Neuts, M. F. (1968). "A Working Bibliography on Markov Renewal Processes and their Applications," Mimeograph Series No. 140, Dept. of Statistics, Purdue University.

[25] Patankar, V. N. (1954). "The Goodness of Fit of Frequency Distributions Obtained from Stochastic Processes," Biometrika, 41: 450-462.

[26] Pyke, R. (1961). "Markov Renewal Processes: Definition and Preliminary Properties," Annals of Mathematical Statistics, 32: 1231-1242.

[27] Pyke, R. (1961). "Markov Renewal Processes with Finitely Many States," Annals of Mathematical Statistics, 32: 1243-1259.

[28] Pyke, R. and Schaufele, R. (1964). "Limit Theorems for Markov Renewal Processes," Annals of Mathematical Statistics, 35: 1746-1764.

[29] Pyke, R. and Schaufele, R. (1966). "The Existence and Uniqueness of Stationary Measures for Markov Renewal Processes," Annals of Mathematical Statistics, 37: 1439-1462.

[30] Raiffa, H. and Schlaifer, R. (1961). Applied Statistical Decision Theory, Graduate School of Business Administration, Harvard University, Boston, Massachusetts.

[31] Rao, C. R. (1966). "Generalized Inverse for Matrices and its Application in Mathematical Statistics," Research Papers in Statistics, Festschrift for J. Neyman. John Wiley and Sons, New York.

[32] Smith, W. L. (1955). "Regenerative Stochastic Processes," Proceedings Royal Society of London, Series A, 232: 6-31.

[33] Takács, L. (1954). "Some Investigations Concerning Recurrent Stochastic Processes of a Certain Type," Magyar Tud. Akad. Mat. Kutatô Int. Közl, 3: 115-128.

[34] Whittle, P. (1955). "Some Distribution and Moment Formulae for the Markov Chain," Journal of the Royal Statistical Society, Series B, 17: 235-242.

[35] Wysocki, R. K. and Kshirsagar, A. M. (1971). "Distribution of Transition Frequencies of a Markov Renewal Process, Over an Arbitrary Interval of Time," Metron (to appear).

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) SOUTHERN METHODIST UNIVERSITY	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP UNCLASSIFIED

3. REPORT TITLE

"Statistical Inference for Markov Renewal Processes"

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (First name, middle initial, last name)

Dwight B. Brock

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 15, 1971	66	35

8a. CONTRACT OR GRANT NO. N00041-68-A-0515 b. PROJECT NO. NR 042-260 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 103 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research

13. ABSTRACT

A Markov Renewal Process is one which records at each time t the number of times a system visits each of a finite number (m) of states up to time t. The system moves from state to state according to a Markov chain, and the time required for each move (sojourn time) is a random variable whose distribution function may depend on the two states between which the move is made. In this paper we develop a test for the goodness of fit of a hypothetical transition probability matrix for a Markov Renewal Process. We illustrate this procedure numerically by applying it to a realization of a two-state Markov Renewal Process artificially generated on a computer. In addition, we consider some Bayesian analysis for Markov Renewal Processes by assuming a matrix beta prior distribution for the transition probability matrix. We also discuss a special case of this topic and give an illustration for a two-state Markov Renewal Process. In the final chapter we give a summary of results and indicate some possible future research problems.