APPLIED MATHEMATICS AND STATISTICS LABORATORIES

STANFORD UNIVERSITY
CALIFORNIA

ESTIMATING THE INFINITESIMAL GENERATOR OF A CONTINUOUS TIME, FINITE STATE MARKOV PROCESS

By

ARThUR ALBERT

TECHNICAL REPORT NO. 60
August 31, 1960

PREPARED UNDER CONTRACT Nonr-225(52)
(NR-342-022)
FOR
OFFICE OF NAVAL RESEARCH
ESTIMATING THE INFINITESIMAL GENERATOR OF A CONTINUOUS TIME,

FINITE STATE MARKOV PROCESS

by

Arthur Albert

TECHNICAL REPORT NO. 60
August 31, 1960

PREPARED UNDER CONTRACT Nonr-225(52)
(NR 342-022)
OFFICE OF NAVAL RESEARCH

This work was sponsored by the Army, Navy, and Air Force through the Joint Services Advisory Committee for Research Groups in Applied Mathematics and Statistics by Contract No. Nonr-225(52)(NR 342-022).

Reproduction in Whole or in Part is Permitted for any Purpose of the United States Government

APPLIED MATHEMATICS AND STATISTICS LABORATORIES
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
ESTIMATING THE INFINITESIMAL GENERATOR OF A CONTINUOUS TIME, FINITE STATE MARKOV PROCESS

by

Arthur Albert

Abstract.

Let \( \{Z(t), t > 0\} \) be a separable, continuous time Markov Process with stationary transition probabilities \( P_{ij}(t), i, j = 1, 2, \ldots, M \). Under suitable regularity conditions, the matrix of transition probabilities, \( P(t) \), can be expressed in the form \( P(t) = \exp tQ \), where \( Q \) is an \( M \times M \) matrix and is called the "infinitesimal generator" for the process.

In this paper, a density on the space of sample functions over \([0,t)\) is constructed. This density depends upon \( Q \). If \( Q \) is unknown, the maximum likelihood estimate \( \hat{Q}(k,t) = \|\hat{a}_{ij}(k,t)\| \), based upon \( k \) independent realizations of the process over \([0,t)\) can be derived.

If each state has positive probability of being occupied during \([0,t)\) and if the number of independent observations, \( k \), grows large (\( t \) held fixed), then \( \hat{a}_{ij} \) is strongly consistent and the joint distribution of the set \( \{\sqrt{k} (\hat{a}_{ij} - a_{ij})\}_{i \neq j} \) (suitably normalized), is asymptotically normal with zero mean and covariance equal to the identity matrix.

If \( k \) is held fixed (at one, say) and if \( t \) grows large, then \( \hat{a}_{ij} \) is again strongly consistent and the joint distribution of the set \( \{\sqrt{t} (\hat{a}_{ij} - a_{ij})\}_{i \neq j} \) (suitably normalized), is asymptotically normal with zero mean and covariance equal to the identity matrix, provided that the process \( \{Z(t), t > 0\} \) is positively regular.

The asymptotic variances of the \( \hat{a}_{ij} \) are computed in both cases.
0. Introduction.

Markov Processes enjoy a wide and diverse field of application in the natural and social sciences. Loosely speaking, a Markov model is a reasonable description of any phenomenon whose future (probabilistic) behavior can be assumed dependent upon its past history, only through its most recent past.

The following model will serve as a prototype for a finite state, temporally homogeneous, continuous time Markov process:

A physical system can be in one of $M$ possible states (which can be taken as the integers $1, 2, \ldots, M$ without loss of generality). At any time, the system can change its state. The times at which changes of state occur are random variables, and the state that the system will go to when a change occurs is a random variable.

We assume that the conditional probability that the system is in a particular state at time $t$, given the entire past of the system up to time $t', t' \leq t$, is the same as the conditional probability that the system is in that state at time $t$, given the state of the system at time $t'$. (Symbolically, $\Pr[Z(t) = j|Z(s), 0 \leq s \leq t'] = \Pr[Z(t) = j|Z(t')]$. It is further assumed that $\Pr[Z(t+h) = j|Z(t) = i]$ depends only on the time difference $h, (h \geq 0)$:

$$\Pr[Z(t+h) = j|Z(t) = i] = P_{ij}(h),$$

and that for each $j, (j \neq i)$, $P_{ij}(h)$ is approximately proportional to $h$ for small, non-negative values of $h$.

If we write $P_{ij}(h) = q(i,j)h + o(h)$ as $h \rightarrow 0$ for all $i$ and $j$, $(i \neq j)$, it turns out that a knowledge of the $M^2 - M$ constants $q(i,j)$, along with a knowledge of the initial probability distribution of the system, $\{\Pr[Z(0) = i]\}$ completely determines the probabilistic behavior of the system. In many practical situations, the initial distribution of the system is known (it may be deterministic for example), but the $q(i,j)$ are not known.

It is the purpose of the present study to suggest an estimate for the $q(i,j)$'s, and to study some of the asymptotic properties of this estimate.
1. **Mathematical Formulation.**

The probabilistic behavior of a finite state, continuous time Markov process \([Z(t), t \geq 0]\) is determined by a knowledge of the **transition probability matrix** \(P(t,s)\), whose entries are

\[
P_{ij}(t,s) = \Pr[Z(s) = j | Z(t) = i]
\]

(where \(t \leq s\) and \(i\) and \(j\) range over the possible states of the process). The process is said to have **stationary transition probabilities** if \(P(t,s)\) depends only on the difference between \(s\) and \(t\):

\[P(t,s) = P(s-t).\]

If certain regularity conditions on the behavior of the (stationary) transition probabilities are met, the so-called **forward and backward systems** of differential equations can be derived:

\[P'(t) = QP(t) = P(t)Q; \quad P(0) = I,\]

where \(Q\) (called the **infinitesimal generator** for the process) is a square matrix whose value is constant in time.

The (unique bounded) solution to the forward and backward equations is

\[P(t) = \exp tQ = \sum_{n=0}^{\infty} \frac{(tQ)^n}{n!}.\]

In this paper, we will discuss the problem of estimating \(Q\) when it is not known. In particular, a maximum likelihood estimate for \(Q\) will be derived and its large-sample properties investigated. The results
obtained here are the continuous time analogues of the results stated by Anderson and Goodman in [1], which describe the asymptotic behavior of the maximum likelihood estimate of the transition probability matrix of a (discrete time) Markov chain.

The question of defining a maximum likelihood estimate (hereafter denoted by M.L.E.) is not an altogether trivial one. Customarily, an M.L.E. is defined relative to a density (which depends upon the parameter to be estimated) over the sample space of the experiment. In the case at hand, the sample space will be a function space (though happily a simple one), and so, one of the first major tasks to be attended to is the construction of a density over the set of realizations for the process. Once this is accomplished, we will see that the M.L.E. is quite simply expressed as a function of the observations, and the remainder of the paper will concern itself with questions of consistency and asymptotic distribution theory.

2. General Properties of Finite-State Markov Processes.

Let \((\Omega, \mathcal{F}, P)\) be a probability space and let \(\{Z(t, \cdot); t \geq 0\}\) be a separable Markov process on that space, which takes its values in a finite set (which for convenience will be taken to be the set \(\{1, 2, \ldots, M\}\)). (Occasionally, the second argument of \(Z(\cdot, \cdot)\) will be suppressed as a notational convenience. It is hoped that no confusion will arise.) The value of \(Z(\cdot, \cdot)\) at time \(t\) is a random variable which will be called the state of the process (S.O.P.) at time \(t\).
It is assumed that the transition probability function is homogeneous in time (stationary):

For any \( s, t \geq 0 \),

\[ P[Z(t + s) = j | Z(s) = i] = P_{ij}(t) \]

depends only upon the time difference \( t \). It is further assumed that

\[ \lim_{h \to 0} \frac{1 - P_{ii}(h)}{h} (= q(i)) \]

exists for all \( i \),

and

\[ \lim_{h \to 0} \frac{P_{ij}(h)}{h} (= q(i,j)) \]

exists for all \( i \) and \( j \),

\( j \neq i \).

The following theorem can be established on the basis of these assumptions, and will serve as the basis for constructing a density on the space of sample functions. For a proof, the reader is referred to Chapter VI of [6].

Theorem 2.1. a) Let

\[ q(i,j) = \begin{cases} 
-q(i) = -\sum_{j \neq i} q(i,j) & \text{if } i = j, \\
q(i,j) & \text{if } i \neq j,
\end{cases} \]

and let \( Q \) be the \((M \times M)\) matrix whose \((i,j)^{th}\) entry is \( q(i,j) \). The matrix of transition probability functions is given by

\[ P(t) = \exp tQ. \]
b) \[ P[Z(t) = i, t_o \leq t \leq t_o + \alpha | Z(t_o) = i] = \exp(-q(i) \alpha) \]

for all non-negative \( t_o \) and \( \alpha \).

c) If \( Z(t_o) = i \) and \( q(i) > 0 \), there is, with probability one, a sample function discontinuity for some \( t > t_o \), and in fact, a first discontinuity which is a jump. If \( 0 < \alpha \leq \infty \), the conditional probability that the first discontinuity in \( [t_o, t_o + \alpha) \) is a jump to \( j \), given that \( Z(t_o) = i \) and that there is a discontinuity in \( [t_o, t_o + \alpha) \), is \( q(i,j)/q(i) \).

d) Almost all sample functions are step functions with a finite number of jumps in any finite time interval.

3. The Space of Sample Functions.

Suppose observations are made on the process \( \{Z(t); 0 \leq t < T\} \) (where \( T \) is finite). By virtue of theorem 2.1, a sample function can be specified by a knowledge of the number of jumps made in \( [0,T) \), the (ordered) lengths of time between jumps, and the succession of values taken on by the process in \( [0,T) \).

To be more precise, suppose \( \Omega' \) is that subset of the underlying probability space for which sample functions of the process \( \{Z(t, \cdot), t \geq 0\} \) are step functions with a finite number of jumps in any finite interval.

By theorem 2.1, this set has probability one. Now, let us define the following random variables (\( \mathbf{x}, \mathbf{y}, \mathbf{z} \)'s):
\[ \tau_0(\omega) = 0, \quad \tau_i(\omega) = \begin{cases} 
\text{The time at which the } i^{\text{th}} \text{ jump occurs} \\
\text{if } \omega \in \Omega' \\
+\infty \text{ otherwise,}
\end{cases} \]

\[ T_i(\omega) = \begin{cases} 
\tau_{i+1}(\omega) - \tau_i(\omega) \quad \text{if } \tau_i(\omega) < \infty \\
0 \quad \text{otherwise,}
\end{cases} \]

\[ N(T,\omega) = \text{The largest integer, } n, \text{ for which } \tau_n(\omega) < T, \]

\[ Z_i(\omega) = Z(\tau_i(\omega), \omega) \quad i = 0, 1, 2, \ldots. \]

(With probabilities one, \( T_1(\omega) \) is the time spent in the \( i^{\text{th}} \) s.o.p., \( N(T,\omega) \) is the number of jumps made by the process in \([0,T]\), and \( Z_i(\omega) \) is the s.o.p. immediately after the \( i^{\text{th}} \) jump.)

In fact, with probability one, a sample function of 
\( \{Z(t,\omega), 0 \leq t < T\} \) can be represented as an ordered sequence:

\[ w.p. \ 1 \]

\[ \{Z(t,\omega), 0 \leq t < T\} = \]

\[ ((Z_0(\omega), T_0(\omega)), \ldots, (Z_{N(T,\omega)-1}(\omega), T_{N(T,\omega)-1}(\omega)), Z_{N(T,\omega)}). \]

By this we mean:

If

\[ \{Z(t,\omega), 0 \leq t < T\} = ((z_0, t_0), \ldots, (z_{n-1}, t_{n-1}), z_n), \]
then the path function starts at $z_0$ at time zero, remains in $z_0$ for $t_0$ units of time, makes a jump to $z_1$, remains in $z_1$ for $t_1$ units of time, ..., jumps to $z_{n-1}$, remains there for $t_{n-1}$ units of time and then makes the final jump to $z_n$, and remains there at least until time $T$. (Note that $n$ jumps, in all, have been made.)

We can write down the probability distribution on the space of sample functions quite easily now:

**Theorem 3.1:**

Let $q'(i,j) = \begin{cases} 0 & \text{if } i = j \\ q(i,j) & \text{if } i \neq j. \end{cases}$

Then

$$\Pr[N(T) = n \& Z_0 = z_0 \& T_0 \leq \alpha_0 \& \ldots \& Z_{n-1} = z_{n-1} \& T_{n-1} \leq \alpha_{n-1} \& Z_n = z_n] = \Pr[Z(0) = z_0] \exp -q(z_n)T \int_{\mathcal{S}_n} \prod_{j=0}^{n-1} dt_j q'(z_j, z_{j+1}) \exp -[q(z_j) - q(z_n)]t_j,$$

where

$$\mathcal{S}_n = \{(t_0, t_1, \ldots, t_{n-1}); \sum_{j=0}^{n-1} t_j < T \& 0 \leq t_j \leq \alpha_j \} \text{ if } n > 0.$$

$$\Pr[N(T) = 0 \& Z_0 = z_0] = \Pr[Z(0) = z_0] \exp -q(z_0)T.$$

**Proof:** The second assertion follows directly from theorem 2.1(b).
If \( n > 0 \), then

\[
\Pr[N(T) = n \land \alpha_o \leq \alpha \land \ldots \land \alpha_{n-1} \leq \alpha_{n-1} \land Z_n = z_n] = \Pr[[Z_o = z_o \land \ldots \land Z_n = z_n \land \sum_{j=0}^{n} T_j \geq T] \cap \tilde{S}_n],
\]

where

\[
\tilde{S}_n = \{(T_o, T_1, \ldots, T_{n-1}) \in S_n\}.
\]

By theorem 2.1(b) and (c)

\[
\Pr[Z_n = z_n \mid Z_o = z_o, \ldots, Z_{n-1} = z_{n-1}, T_o, \ldots, T_{n-1}] = \frac{q'(z_{n-1}, z_n)}{q(z_{n-1})}
\]

and

\[
\Pr[T_n \leq \alpha_n \mid Z_o = z_o, \ldots, Z_n = z_n, T_o, \ldots, T_{n-1}] = 1 - e^{-q(z_n)\alpha_n}.
\]

By induction,

\[
\Pr[Z_o = z_o, \ldots, Z_n = z_n, T_o \leq \alpha_o, \ldots, T_n \leq \alpha_n] = \Pr[Z(0) = z_o] \left\{ \int_{S_n} \left[ \frac{q'(z_j, z_{j+1})}{q(z_j)} \left[1 - e^{-q(z_j)\alpha_j}\right]\right] \left[1 - e^{-q(z_n)\alpha_n}\right] dt_j \right\}
\]

It follows that

\[
\Pr[[Z_o = z_o, \ldots, Z_n = z_n \land \sum_{j=0}^{n} T_j \geq T] \cap \tilde{S}_n]
\]

\[
= \Pr[Z(0) = z_o] \left[ \int_{S_n} \left[ \frac{q'(z_j, z_{j+1})}{q(z_j)} e^{-q(z_n)T_n} dt_n \right] \right] \left[ \int_{\sum_{j=0}^{n-1} t_j}^{\infty} q(z_n) e^{-q(z_j)t_n} dt_n \right] \left[1 - e^{-q(z_n)\alpha_n}\right] dt_j
\]

\[
\times \prod_{j=0}^{n-1} q'(z_j, z_{j+1}) e^{-q(z_j)T_j},
\]
The conclusion is obtained by performing the inner integration.*

It is now (conceptually) quite simple to construct the desired density. Unfortunately, if rigor is to be maintained, a terminology must be built up before stating the desired conclusion:

Every sample function which makes \( n \) jumps in \([0,T]\) can be represented as a point in

\[
\mathcal{W}_n = \left[ \prod_{j=1}^{n} (\mathcal{W}_o \otimes \mathbb{R}^1) \right] \otimes \mathcal{W}_o,
\]

where

\[
\mathcal{W}_o = \{1, 2, \ldots, M\} \quad \text{(the state space of the process)}
\]

and

\[
\mathbb{R}^1 \text{ is the real line.}
\]

Let \( \mathcal{L} \) be the Lebesgue measure on \( \mathbb{R}^1 \) and let \( C \) be the counting measure on \( \mathcal{W}_o \):

\[
C([z]) = 1 \quad \text{if} \quad z \in \mathcal{W}_o.
\]

Let \( \sigma^{(n)} \) be the (sigma-finite) product measure on \( \mathcal{W}_n \), defined by the relation:

\[
\sigma^{(n)} = \left[ \prod_{j=1}^{n} (C \times \mathcal{L}) \right] \times C.
\]

By theorem 2.1(d), almost every sample function of the process \( \{Z(t), 0 \leq t < T\} \) can be represented as a point in

\[
\mathcal{W} = \bigcup_{n=0}^{\infty} \mathcal{W}_n.
\]
For each set \( W \subseteq \mathcal{W} \) for which \( W \cap \mathcal{W}_n \) is \( \sigma^{(n)} \) measurable define
\[
\sigma^*(W) = \sum_{n=0}^{\infty} \sigma^{(n)}(W \cap \mathcal{W}_n).
\]
(\( \sigma^* \) is defined on the Borel-field \( \mathcal{B}^* \), which is the smallest Borel-field containing all sets \( W \subseteq \mathcal{W} \) whose projection on \( \prod_{j=1}^{n} \mathbb{R}^1 \) is a Borel set for each \( n \).

Let \( \sigma \) be a measure on the space of all sample functions, defined for all subsets \( B \) whose intersection with \( \mathcal{W} \) is in \( \mathcal{B}^* \):
\[
\sigma(B) = \sigma^*(B \cap \mathcal{W}). \quad (\sigma \text{ is a sigma-finite measure}.)
\]

The desired density function can now be exhibited. (It is a density with respect to \( \sigma \).)

**Theorem 3.2:**

If \( B \) is a subset of the space of all sample functions over \([0, T]\) which is measurable with respect to \( \sigma \), then
\[
\Pr[B] = \int_B f_Q(v) \, d\sigma(v),
\]
where
\[
f_Q(v) = \begin{cases} 
-\langle z_0, z_n \rangle T & \text{if } v = (z_0); \\
\exp[-q(z_j, z_{j+1})] & \text{if } v = ((z_0, t_0), \ldots, (z_{n-1}, t_{n-1}), z_n) \\
\end{cases}
\]
with \( n > 0, z_j \in \mathcal{W}_j, t_j \geq 0 \) (\( j = 0, 1, \ldots, n-1 \)),
and
\[
\sum_{j=0}^{n-1} t_j < T;
\]
otherwise.
Proof:

By theorem 3.1, the conditional distribution of

\(((Z_0, T_0), \ldots, (Z_{n-1}, T_{n-1}), Z_n)\) given that \(N(T) = n\) is, for each \(n\), absolutely continuous with respect to \(\sigma^{(n)}\). If the derivative exists,

\[
\frac{\partial^n}{\partial t_0 \cdots \partial t_{n-1}} \frac{Pr[N(T) = n \& Z_0 = z_0 \& \cdots \& Z_n = z_n \& T_0 < t_0 \& \cdots \& T_{n-1} < t_{n-1}}{Pr[N(T) = n]}
\]

is the (joint) conditional density of

\(((Z_0, T_0), \ldots, (Z_{n-1}, T_{n-1}), Z_n)\) with respect to \(\sigma^{(n)}\),

given that \(N(T) = n\) (for \(n > 0\)).

In this case,

\[
Pr[B | N(T) = n] = \int_B \frac{\partial^n}{\partial t_0 \cdots \partial t_{n-1}} N(T) = n \sigma^{(n)}
\]

for all sets \(B\) whose intersection with \(\mathcal{N}_n\) is \(\sigma^{(n)}\) measurable, so that if \(B\) is a subset of the space of all possible sample functions, and \(B \cap \mathcal{N}_n\) is \(\sigma^{(n)}\) measurable for each \(n\), then

\[
Pr[B] = \sum_{n=0}^{\infty} \int_B \frac{\partial^n}{\partial t_0 \cdots \partial t_{n-1}} N(T) = n \sigma^{(n)}
\]

A routine computation (in connection with theorem 3.1) shows that the derivative in question does indeed exist and that \(g_n Pr[N(T) = n]\) corresponds to \(f_Q\) on \(\mathcal{N}_n\).
Thus,
\[
Pr[B] = \sum_{n=0}^{\infty} \int_{B \cap \mathcal{F}_n} f_Q^{(n)} \, d\sigma = \int_{B \cap \mathcal{F}} f_Q^{*} \, d\sigma = \int_B f_Q \, d\sigma
\]

Although we shall not discuss the hypothesis testing problem here, the following result is quite important in that context and is included here since it follows so readily from the results established thus far.

**Theorem 3.3:**

Let \( Q_0 = ||q_0(i,j)|| \) and \( Q_1 = ||q_1(i,j)|| \) \( i,j = 1,2, \ldots, M \) be two values of the parameter \( Q \).

a) If \( q_0(i,j) \) vanishes whenever \( q_1(i,j) \) does, then the probability measure on \((\Omega, \mathcal{F}, P)\) under \( Q_0 \) is absolutely continuous with respect to the one under \( Q_1 \).

b) If for some \( i \) and \( j(j \neq i) \), \( q_1(i,j) = 0 \) and \( q_0(i,j) > 0 \) and \( P_{Q_0}[Z(t_0) = i] > 0 \) for some \( t_0 > 0 \), then \( P_{Q_0} \) is not absolutely continuous with respect to \( P_{Q_1} \).

**Proof:**

a) Under stated conditions, \( f_{Q_0} \) vanishes with \( f_{Q_1} \).

b) Let \( B \) be the set of sample functions whose value at time \( t_0 \) is \( i \) and whose next (distinct) value is \( j \). By theorem 2.1(c), \( P_{Q_1}[B] = 0 \) and \( P_{Q_0}[B] > 0 \).

It can be shown that \( P_{1j}(t) \) never vanishes for \( t > 0 \) unless it vanishes identically (see [6], Ch. VI). It follows then, that \( Pr[Z(t) = i] \) never vanishes for \( t > 0 \) unless it too vanishes identically.
If the parameter space is restricted to those values of $Q$ for which every state $i$ has positive probability of being occupied eventually, then a necessary and sufficient condition that $P_{Q_{\infty}}$ be absolutely continuous with respect to $P_{Q_{\infty}}$ is that $q_{Q}(i,j)$ vanish whenever $q_{\infty}(i,j)$ does.


Suppose $k$ independent realizations $v_{1}, v_{2}, \ldots, v_{k}$ of $\{Z(t), 0 \leq t < T\}$ are observed. The likelihood function, $L_{Q}^{(k)}$, has been traditionally defined by the equation

$$L_{Q}^{(k)} = \prod_{j=1}^{k} p_{Q}(v_{j}) .$$

If we let

$$N_{T}^{(k)}(i,j) = \text{the total number of transitions from state } i \text{ to state } j \text{ observed during the } k \text{ trials}$$

and

$$A_{T}^{(k)}(i) = \text{the total length of time that state } i \text{ is occupied during the } k \text{ trials},$$

we may write

$$\log L_{Q}^{(k)} = C_{k} + \sum_{i} \sum_{j \neq 1} N_{T}^{(k)}(i,j) \log q(i,j) - \sum_{i} A_{T}^{(k)}(i)q(i) ,$$

where $C_{k}$ is finite with probability one and does not depend upon $Q$.

The Halmos-Savage factorization theorem can be applied to the last expression, and by inspection we see that the set
is a sufficient statistic for \( Q \).

The maximum likelihood estimates (m.l.e) for \( q(i,j) \) (i.e., those values of \( q(i,j) \) which maximize \( \log \mathcal{L}_Q^{(k)} \)) are seen to be

\[
\hat{q}_T^{(k)}(i,j) = \frac{N_T^{(k)}(i,j)}{A_T^{(k)}(i)} \quad \text{if} \quad i \neq j \quad \text{and} \quad A_T^{(k)}(i) > 0.
\]

If \( A_T^{(k)} = 0 \), the m.l.e. does not exist and so we adopt the convention that

\[
\hat{q}_T^{(k)}(i,j) = 0 \quad \text{if} \quad i \neq j \quad \text{and} \quad A_T^{(k)}(i) = 0.
\]

5. Moments.

In the sections to follow, we will investigate the large sample properties of the set \( \{\hat{q}_T^{(k)}(i,j)\}_{j \neq i} \) both as \( T \) approaches infinity while \( k \) is fixed, and as \( k \) approaches infinity with \( T \) fixed. A knowledge of certain moments of \( N_T^{(k)}(i,j) \) and \( A_T^{(k)}(i) \) proves to be of use, and the following lemma reduces the determination of these moments to a fairly routine computational task:

For notational convenience, we will let \( N_T(i,j) = N_T^{(1)}(i,j) \) and \( A_T(i) = A_T^{(1)}(i) \) for all \( i \) and \( j \). \( N(T) \) is as defined in section 3.

**Lemma 5.1:**

a) There are constants \( \alpha \) and \( \beta \) such that:
\[ \Pr[N(h) \geq n] \leq \beta \int_0^h \frac{t^{n-1} e^{\alpha t}}{1-t} \, dt \]

for all \( h \) between zero and one. Consequently,

\[ \sum_{n=2} \Pr[N(h) \geq n] \leq \beta \int_0^h \frac{te^{\alpha t} \, dt}{(1-t)^2} = o(h), \text{ as } h \to 0. \]

b) \( \Pr[N_h(i,j)] = 1 = p(i) q(i,j) h + o(h) \text{ as } h \to 0 \) and

\[ E N_h(i,j) = p(i) q(i,j) h + o(h) \text{ as } h \to 0. \]

(Here, \( p(i) = \Pr[Z(0) = i] \).

Proof:

a) By theorem 3.1, there are constants \( K, \alpha, \) and \( \gamma \), such that:

\[ \Pr[N(h) = n] \leq K \gamma^n \int \cdots \cdots \int \prod_{j=1}^{n} e^{\alpha_j t_j} \, dt_j \]

\[ 0 \leq \sum_{j=1}^{n} t_j \leq h \]

\[ = K' \left( \frac{\gamma'}{(n-1)!} \right) \int_0^h t^{n-1} e^{\alpha t} \, dt, \text{ (where } \gamma' > 0), \]

\[ \leq \beta \int_0^h t^{n-1} e^{\alpha t} \, dt \text{ (where } \beta = K' e^{\gamma'}). \]

b) By theorem 2.1,

\[ \Pr[N_h(i,j) = 1 | N(h) = 1 \& Z(0) = i] = q(i,j)/q(i). \]

By theorem 3.1,

\[ \Pr[N(h) = 1 | Z(0) = i] = q(i) h + o(h). \]
Whence,
\[ \Pr[N_h(i,j) = 1] = p(i) \, q(i,j) \, h + o(h) . \]

Since
\[
\mathbb{E}_h(i,j) = \sum_{n=1}^{\infty} \left[ \Pr[N_h(i,j) \geq n] = \Pr[N_h(i,j) = 1] + \Pr[N_h(i,j) \geq 2] \\
+ \sum_{n=2}^{\infty} \Pr[N_h(i,j) \geq n] . \right.
\]

From part a)
\[ \Pr[N_h(i,j) \geq 2] \leq \Pr[N(h) \geq 2] = o(h) \]
and
\[ \sum_{n=2}^{\infty} \Pr[N_h(i,j) \geq n] \leq \sum_{n=2}^{\infty} \Pr[N(h) \geq n] = o(h) . \]

Theorem 5.2:
\[ a) \ E_n(i,j) = q(i,j) \int_{0}^{T} \Pr[Z(t) = 1] \, dt . \]
\[ b) \ E_{n}(i) = \int_{0}^{T} \Pr[Z(t) = 1] \, dt . \]
\[ c) \ E_n(i,j) N_T^{(r,s)} = q(i,j) q(r,s) \int_{0}^{T} \left[ \sum_{x=1}^{\infty} p_{r}^{(x-t)} \Pr[Z(t) = r] \right] \]
\[ + p_{j}^{(x-t)} \Pr[Z(t) = 1] \, dt \right] dx + \delta(i,j,r,s) q(i,j) \int_{0}^{T} \Pr[Z(t) = 1] \, dt , \]

where
\[ \delta(i,j;r,s) = \begin{cases} 
1 & \text{if } r=1 \text{ and } j=s \\
0 & \text{otherwise} 
\end{cases} . \]
d) \[ E \ A_T(i) A_T(r) = \int_0^T \int_0^X \left[ P_{ir}(x-t)Pr[Z(t) = r] + P_{ir}(x-t)Pr[Z(t) = i] \right] \, dt \, dx. \]

e) \[ E \ N_T(r,s)A_T(i) = q(r,s) \int_0^T \int_0^X \left[ P_{ir}(x-t)Pr[Z(t) = i] + P_{sr}(x-t)Pr[Z(t) = r] \right] \, dt \, dx. \]
f) \[
E \left[ \left( N_T(r,s) - q(r,s) A_T(r) \right) \left( N_T(i,j) - q(i,j) A_T(i) \right) \right]
= \delta(i,j;r,s) q(i,j) \int_0^T Pr[Z(t) = i] \, dt.
\]

Part (f) of the theorem is rather surprising: Since the r.v.'s \( A_T(i) \) are constrained by the relation \( \sum_{i=1}^M A_T(i) = T \), it is unexpected to find that \( (N_T(i,j) - q(i,j) A_T(i)) \) and \( (N_T(i^1,j^1) - q(i^1,j^1) A_T(i^1)) \) are uncorrelated, (even if \( i = i^1 \)) when \( j \neq j^1 \).

Proof:

We will prove (a), (b) and (f) here and relegate the proofs of (c), (d) and (e) to Appendix I.

a) Divide \([0,T]\) into \( n+1 \) equal parts of length \( h = T/n+1 \), and let \( X_k(i,j) \) be the number of transitions from \( i \) to \( j \) during the (time) interval \([k-1,h,(k+1)]\), \( k = 0,1,2,\ldots,n \). Then

\[
E \ N_T(i,j) = \sum_{k=0}^n E X_k(i,j) = \sum_{k=0}^n Pr[Z(k+1)h = j & Z(k\cdot h) = i] + o(h)
\]

\[
= \sum_{k=0}^n Pr[Z(k\cdot h) = i] q(i,j)h + o(h) \rightarrow q(i,j) \int_0^T Pr[Z(t) = i] \, dt ,
\]
as \( n \rightarrow \infty \).
b) Let \( Y_i(t) = \begin{cases} \ 1 & \text{if } Z(t) = 1, \\ \ 0 & \text{otherwise} \end{cases} \).

Then

\[ A_T(i) = \int_0^T Y_i(t) \, dt, \] so that

\[ E A_T(i) = \int_0^T E Y_i(t) \, dt = \int_0^T \Pr[Z(t) = 1] \, dt. \]

c) \[
E[\left( N_T(i,j) - q(i,j) A_T(1) \right) \left( N_T(r,s) - q(r,s) A_T(r) \right)]
\]

\[ = \ E N_T(i,j) N_T(r,s) + q(r,s) q(i,j) E A_T(i) A_T(r) \]

\[- (q(i,j) E N_T(r,s) A_T(i) + q(r,s) E N_T(i,j) A_T(r)), \]

and the result follows from comparison with (c), (d) and (e). ■

6. **Large Sample Properties of the M.L.E.**

The term "large sample" can be interpreted in two ways in relation to the problem at hand: Many independent realizations of the process \( [Z(t), 0 \leq t < T] \) could be observed. (This corresponds to an investigation of the behavior of \( \Lambda^{(k)}_{\alpha_T}(i,j) \) as \( k \to \infty \).) On the other hand, a single realization of the process can be observed over a long period of time. (This corresponds to an investigation of \( \Lambda^{(1)}_{\alpha_T}(i,j) \) as \( T \to \infty \).) We will obtain results pertaining to the consistency and asymptotic normality of these estimates in both cases.

First, we will study the (easier) problem of holding \( T \) fixed and letting \( k \) grow large. The reader will recall that if \( i \neq j \),
\[ a^{(k)}_{q_T}(i,j) = \begin{cases} \frac{N_T^{(k)}(i,j)}{A_T^{(k)}(i)} & \text{if } A_T^{(k)}(i) > 0 \\ 0 & \text{if } A_T^{(k)}(i) = 0 \end{cases} \]

(For convenience's sake, we will sometimes suppress the "T" in this part of the discussion.)

If \( \Pr[A_T(i) = 0] = 1 \), then \( a^{(k)}(i,j) = 0 \) with probability one for every \( k \), even if \( q(i,j) > 0 \). In this case, the estimate is a bad one. This situation only occurs when we try to estimate parameters associated with transitions out of a state that can never be reached (for \( \Pr[A_T(i) = 0] = 1 \) iff \( \Pr[Z(t) = 1] = 0 \) for every \( t \)). This case must be excluded from consideration.

If \( \Pr[Z(t) = 1] > 0 \), then \( E A_T(i) > 0 \), so that by theorem 5.2,

\[ \frac{E N_T(i,j)}{E A_T(i)} = q(i,j). \]

Since \( N_T^{(k)}(i,j) \) and \( A_T^{(k)}(i) \) are sums of independent, identically distributed variables,

\[ \lim_{k \to \infty} a^{(k)}(i,j) = \lim_{k \to \infty} \frac{N_T^{(k)}(i,j)/k}{A_T^{(k)}(i)/k} = q(i,j) \]

with probability one.

By applying a theorem of Cramér (c.f. [4], theorem 2, or [5], page 254), the set of r.v.'s

\[ \{ \sqrt{k} (a^{(k)}(i,j) - q(i,j)) \} \]

\[ i \neq j \]
has the same asymptotic distribution as the set

$$\left\{ \frac{1}{E A_T(i)} \left( \frac{N_T^k(i,j) - q(i,j) A_T^{k}(i)}{\sqrt{k}} \right) \right\}_{i \neq j}$$

and by the multivariate central limit theorem, the last is asymptotically normal, with mean zero and covariances

$$C(i,j;k,l)$$

$$= \frac{1}{E A_T(i) E A_T(k)} E[N_T(i,j) - q(i,j) A_T(i)][N_T(k,l) - q(k,l) A_T(k)]$$

$$= 8(i,j;k,l) q(i,j) \int_0^T \Pr[Z(t) = 1] \, dt .$$

This completes the proof of

**Theorem 6.1:**

a) If there is positive probability of the \(i^{th}\) state being occupied at some time \(t \geq 0\), then

$$\lim_{k \to \infty} \hat{q}^k(i,j) = q(i,j) \text{ with probability one.}$$

b) If every state has positive probability of being occupied, then the set of r.v.'s

$$\left\{ \sqrt{k} (\hat{q}^k(i,j) - q(i,j)) \right\}_{j \neq i}$$

are asymptotically normal and independent with zero mean and variances.
\[ q(i,j) / \int_0^T \Pr [Z(t) = i] \, dt. \]

(Notice that, \( q^{(k)}(i,j) = 0 \) with probability one for every \( k \) if \( q(i,j) = 0 \), since, by theorem 5.1, \( N_T(i,j) \) has zero mean in this case.)

Now let us turn our attention to the more interesting problems of consistency and asymptotic normality when \( k \) is held fixed (\( k = 1 \)) and \( T \) grows large.

First we will derive a series of results which lead to a demonstration of the (joint) asymptotic normality of the set of r.v.'s

\[
\left\{ \frac{N_T(i,j) - q(i,j) A_T(i)}{\sqrt{T}} \right\}_{j \neq i}.
\]

From this, we can conclude that

\[
\frac{1}{T} \left[ N_T(i,j) - q(i,j) A_T(i) \right]
\]

converges to zero in probability as \( T \to \infty \).

Then, we will show that \( \frac{1}{T} N_T(i,j) \) and \( \frac{1}{T} A_T(i) \) converge with probability one as \( T \to \infty \), and, in fact, converge to constants with probability one. If \( \lim_{T \to \infty} \frac{1}{T} A_T(i) \) is positive, Cramer's theorem can be used to show that the set of r.v.'s

\[
\left\{ \sqrt{T} (\hat{q}_T(i,j) - q(i,j)) \right\}_{j \neq i}
\]

have a joint distribution which is asymptotically normal.
Furthermore, $\hat{q}_T(i,j)$ is consistent in the strong sense, since

$$\frac{N_T(i,j)}{A_T(i)} - q(i,j)$$

converges to zero with probability one as $T \to \infty$.

**Definition:**

a) If $\hat{\alpha} = \{\alpha_{i,j}\}_{i \neq j}$ and $\hat{\beta} = \{\beta_{i,j}\}_{i \neq j}$ are two sequences indexed by double subscripts which run over the integers $1,2,\ldots,M$, we define

$$\langle \hat{\alpha}, \hat{\beta} \rangle = \sum_{i} \sum_{j \neq i} \alpha_{i,j} \beta_{i,j}.$$  

b) If

$$\hat{\delta} = \{\delta_{i}\}_{i=1}^{M} \quad \text{and} \quad \hat{\epsilon} = \{\epsilon_{i}\}_{i=1}^{M}$$

we define

$$\langle \hat{\delta}, \hat{\epsilon} \rangle = \sum_{i=1}^{M} \delta_{i} \epsilon_{i}.$$  

c) For any matrix $A$, let $A^\ast$ be the transpose of $A$.

**Theorem 6.2:**

Let $\hat{\xi}_T = (\hat{\xi}_T(i,j))_{i \neq j}$, where

$$\hat{\xi}_T(i,j) = \frac{1}{\sqrt{T}} [N_T(i,j) - q(i,j) A_T(i)].$$

Let $\hat{\omega} = \{\omega(i,j)\}_{i \neq j}$, and $\varphi(\hat{\omega} ; T) = E \exp - \langle \hat{\xi}_T , \hat{\omega} \rangle$ be the joint moment generating function of the set $(\hat{\xi}_T(i,j))_{i \neq j}$.  

Then

\[ \varphi(\overrightarrow{\omega}; t) = \left( e^{t \overrightarrow{R}(\overrightarrow{\omega}, t) \cdot \overrightarrow{J}}, \overrightarrow{\eta} \right), \]

where \( R(\overrightarrow{\omega}, t) \) is an \( M \times M \) matrix whose \( (i, j) \)th entry is

\[ r_{ij}(\overrightarrow{\omega}, t) = \begin{cases} -Q(i) + \frac{1}{\sqrt{t}} \sum_{k \neq 1} q(i, k) \omega(i, k) & \text{if } i = j, \\ q(i, j) \exp -\frac{\omega(i, j)}{\sqrt{t}} & \text{if } i \neq j, \end{cases} \]

and \( \overrightarrow{\eta} \) and \( \overrightarrow{\zeta} \) are \( M \)-vectors:

\[ \overrightarrow{\zeta} = \begin{pmatrix} \Pr [Z(0) = 1] \\ \Pr [Z(0) = 2] \\ \vdots \\ \Pr [Z(0) = M] \end{pmatrix}, \quad \overrightarrow{\eta} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \]

**Proof:**

Let \( g_k(\overrightarrow{\omega}, t) = \mathbb{E}(\exp - < \overrightarrow{\xi}_t, \overrightarrow{\omega} > | Z(t) = k) \) be the conditional moment generating function of \( \overrightarrow{\xi}_t \), given that \( Z(t) = k \). Since

\[ A_t(k) = t - \sum_{i \neq k} A_t(i), \]

\[ < \overrightarrow{\xi}_t, \overrightarrow{\omega} >= -\sqrt{t} \alpha_k + \frac{1}{\sqrt{t}} \left[ \sum_i \sum_{j \neq i} N_t(i, j) \omega(i, j) + \sum_{i \neq k} (\alpha_k - \alpha_i) A_t(i) \right], \]

where

\[ \alpha_i = \sum_{j \neq i} q(i, j) \omega(i, j), \quad i = 1, 2, \ldots, M. \]
Let 

\[ \hat{n} = (n_{12}, n_{13}, \ldots, n_{M,M-1}) \]

be a generic point whose \( M^2 - M \) components \( \{n_{ij}\}_{i \neq j} \) are non-negative integers, and let \( \hat{a} = (a_1, a_2, \ldots, a_m) \) be a generic point whose components are non-negative real numbers. Let \( G_k(\hat{n}, \hat{a}; t) \) be the joint conditional distributive function of the r.v.'s \( \{N_t(i,j)\}_{i \neq j} \) and \( \{A_t(1)\}_{i \neq k} \) given that \( Z(t) = k \):

\[
G_k(\hat{n}, \hat{a}; t) = \Pr \left( \left( \bigcap_i \bigcap_{j \neq i} [N_t(i,j) \leq n_{ij}] \right) \cap \left( \bigcap_{i \neq k} [A_t(i) \leq a_i] \right) | Z(t) = k \right),
\]

and let

\[ \mathcal{N} = \{(n_{12}, n_{13}, \ldots, n_{M,M-1}) : n_{12} \geq 0, \ldots, n_{M,M-1} \geq 0 \}, \]

\[ \mathcal{A}_k = \{\hat{a} : a_1 > 0, \ldots, a_{k-1} > 0, a_{k+1} > 0, \ldots, a_m > 0\} , \]

and

\[ \mathcal{B}_k = \bigcup_{i \neq k} \{\hat{a} : a_i = 0\} . \]

We can write

\[
ge_k(\hat{\omega}, t) = e^{\alpha_k \sqrt{t}} \left\{ \int_{\mathcal{N}} \int_{\mathcal{A}_k} \exp \left( -\frac{1}{\sqrt{t}} \left[ \frac{1}{2} < \hat{n}, \hat{\omega} > + \sum_{i \neq k} (\alpha_k - \alpha_i) a_i \right] \right) \, dG_k \\
+ \int_{\mathcal{N}} \int_{\mathcal{B}_k} \exp \left( -\frac{1}{\sqrt{t}} \left[ \frac{1}{2} < \hat{n}, \hat{\omega} > + \sum_{i \neq k} (\alpha_k - \alpha_i) a_i \right] \right) \, dG_k \right\} .
\]
The inner integral of the first term can be integrated by parts
(see Appendix II) and we find that:

\[(6.2.1) \quad \int \int \exp - \frac{1}{\sqrt{t}} [<\hat{n}, \hat{\omega}> + \sum_{i \neq k} (\alpha_k - \alpha_i) a_i] \, \mathcal{d}c_k \]

\[= \left[ \frac{\alpha_k - \alpha_i}{\sqrt{t}} \right] \sum_{n_1 > 0} \ldots \sum_{n_M, M-1 > 0} \int_0^\infty \int_0^\infty \int_0^\infty \left( \prod_{i \neq k} a_i \right) F_k(\hat{n}, \hat{\omega}; t) \exp - \frac{1}{\sqrt{t}} [<\hat{n}, \hat{\omega}> + \sum_{i \neq k} (\alpha_k - \alpha_i) a_i] \, \mathcal{d}c_k , \]

where

\[F_k(\hat{n}, \hat{\omega}; t) = \Pr[(\bigcap_{i \neq 1} [N_i(1, j) = n_{i, j}) \bigcap (\bigcap_{i \neq k} [A_i(1) \leq a_i]) \bigcap [Z(t) = k])] . \]

Hence,

\[(6.2.2) \quad \varphi(\hat{\omega}; t) = \sum_{k=1}^{M} \Pr[Z(t) = k] \tilde{c}_k(\hat{\omega}; t) = \sum_{k=1}^{M} \tilde{c}_k^\prime(\hat{\omega}; t) , \]

where

\[\tilde{c}_k(\hat{\omega}; t) = \]

\[= e_k \left( \alpha_k - \frac{\alpha_i}{\sqrt{t}} \right) \sum_{n_1 > 0} \ldots \sum_{n_M, M-1 > 0} \int_0^\infty \int_0^\infty \int_0^\infty \left( \prod_{i \neq k} a_i \right) F_k(\hat{n}, \hat{\omega}; t) \exp - \frac{1}{\sqrt{t}} [<\hat{n}, \hat{\omega}> + \sum_{i \neq k} (\alpha_k - \alpha_i) a_i] , \]
The last expression is closely akin to the Laplace transform of \( F_k \).

If we let

\[
f_k(\vec{u}, \vec{v}; t) = \sum_{n_1 \geq 0} \cdots \sum_{n_{M-1} \geq 0} \int_0^\infty \cdots \int_0^\infty (\prod_{i=1}^M d\alpha_i) F_k \exp\left[\langle \vec{n}, \vec{u} \rangle + \langle \vec{a}, \vec{v} \rangle\right],
\]

then since \( F_k \) is constant as a function of \( a_k \), we see that

\[
f_k(\vec{u}, \vec{v}; t) = \frac{1}{v_k} \sum_{n_1 \geq 0} \cdots \sum_{n_{M-1} \geq 0} \int_0^\infty \cdots \int_0^\infty (\prod_{i \neq k} d\alpha_i) F_k \exp\left[\langle \vec{n}, \vec{u} \rangle + \sum_{i \neq k} a_i v_i\right],
\]

so that

\[(6.2.3) \quad f_k(\vec{u}, \vec{v}; t) = \tilde{g}_k(\vec{\omega}, t),\]

if we set:

\[(6.2.4) \quad u_{i,j} = \omega(i,j)/\sqrt{t} \quad i,j = 1,2,\ldots,M, \quad i \neq j,\]

\[v_i = (\alpha_k - \alpha_i)/\sqrt{t} \quad \text{if} \quad i \neq k\]

and

\[v_k = \left[ e^{k \sqrt{t}} \prod_{i \neq k} \left( \frac{\alpha_k - \alpha_i}{\sqrt{t}} \right) \right]^{-1}.\]

It now remains to exhibit \( f_k(\vec{u}, \vec{v}; t) \):

In Appendix III, it is shown that the set of functions \( F_k, k=1,2,\ldots,M \), satisfy a system of first order linear differential-difference equations:

\[
\frac{\partial F_k}{\partial t} = -q(k) F_k + \sum_{\nu \neq k} q(\nu, k) \Delta_{\nu k} F_\nu.
\]
\[ F_k = \Pr[Z(0) = k] \quad \text{if } t = 0 \text{ and } n_{ij} = 0 \text{ for all } i, j, \]
\[ F_k = 0 \quad \text{if } n_{ij} < 0 \text{ for some } i \text{ and } j \]
\[ F_k = 0 \quad \text{if } t = 0 \text{ and } n_{ij} > 0 \text{ for some } i \text{ and } j. \]

(Here, \( \Delta_{uk} \) is the first order difference operator:)

\[ \Delta_{uk} F_k (\vec{n}, \vec{\alpha}; t) = F_k (D_{uk} \vec{n}, \vec{\alpha}; t) \text{ where} \]
\[ D_{uk} (n_{12}, \ldots, n_{uk}, \ldots, n_{M,M-1}) = (n_{12}, \ldots, n_{uk} - 1, \ldots, n_{M,M-1}). \]

It is easily verified that

\[ \frac{\partial f_k}{\partial t} = -q(k)f_k + \sum_{\nu \neq k} q(\nu, k) e^{-u_{\nu k}} f_\nu , \quad f_k (\vec{u}, \vec{v}; 0) = \left( \prod_{i=1}^{M} v_i \right)^{-1} \Pr[Z(0) = k]. \]

If we let

\[ \vec{f}(\vec{u}, \vec{v}; t) = \begin{pmatrix} f_1 (\vec{u}, \vec{v}; t) \\ f_2 (\vec{u}, \vec{v}; t) \\ \vdots \\ f_M (\vec{u}, \vec{v}; t) \end{pmatrix}, \]

and adopt a matrix notation, we see that

\[ \frac{\partial \vec{f}(\vec{u}, \vec{v}; t)}{\partial t} = W(\vec{u}) \vec{f}(\vec{u}, \vec{v}; t) , \quad \vec{f}(\vec{u}, \vec{v}; 0) = \left( \prod_{i=1}^{M} v_i \right)^{-1} \vec{v}, \]

where \( W(\vec{u}) \) is an \( M \times M \) matrix whose \( (i,j) \)-th entry is

\[ v_{i,j}(\vec{u}) = \begin{cases} -q(i) & \text{if } i = j \\ q(i,j) e^{-u_{ij}} & \text{if } i \neq j \end{cases}. \]
The only bounded solution to this set of equations is known to be

\[(6.2.5) \quad \hat{P}(\tilde{u}, \tilde{v}; t) = (\prod_{j=1}^{M} v_j)^{-1} e^{t\tilde{W}^*}\hat{\tilde{u}}, \hat{\tilde{v}}, \hat{\tilde{\eta}}\]

where

\[e^{t\tilde{W}^*} = \sum_{n=0}^{\infty} \frac{(t\tilde{W}^*)^n}{n!} .\]

If we let \(\hat{e}_k\) be the M-vector whose components are all zero except for the \(k^{th}\) which is unity, then

\[r_k(\tilde{u}, \tilde{v}; t) = (\prod_{i=1}^{M} v_i)^{-1} (e^{t\tilde{W}^*(\tilde{u})}\hat{\tilde{u}}, e^{t\tilde{W}^*(\tilde{v})}\hat{\tilde{v}}, \hat{e}_k) .\]

Referring to equations 6.6.3 and 6.6.4, we see that

\[\tilde{g}_k(\tilde{\omega}, t) = e^{\alpha_k \frac{\sqrt{t}}{v_f}} (e^{t\tilde{W}^*(\tilde{\omega})}\hat{\tilde{\omega}}, \hat{e}_k) = (e^{t\tilde{W}^*(\tilde{\omega})}\hat{\tilde{\omega}}, e^{\alpha_k \frac{\sqrt{t}}{v_f}} \hat{e}_k) , \quad \text{so that} \]

\[\varphi(\tilde{\omega}, t) = \sum_{k=1}^{M} \tilde{g}_k(\tilde{\omega}, t) = (e^{t\tilde{W}^*(\tilde{\omega})}\hat{\tilde{\omega}}, \sum_{k=1}^{M} e^{\alpha_k \frac{\sqrt{t}}{v_f}} \hat{e}_k) .\]

But

\[\sum_{k=1}^{M} e^{\alpha_k \frac{\sqrt{t}}{v_f}} \hat{e}_k = \begin{pmatrix} e^{\frac{\alpha_1 \sqrt{t}}{v_f}} & 0 & \cdots & 0 \\ 0 & e^{\frac{\alpha_2 \sqrt{t}}{v_f}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\frac{\alpha_M \sqrt{t}}{v_f}} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = e^{\frac{tA(\tilde{\alpha})}{v_f}} \hat{\eta} ,\]
where,
\[ A(\vec{\alpha}) = \begin{pmatrix}
\alpha_1 & 0 & \ldots & 0 \\
0 & \alpha_2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \alpha_M
\end{pmatrix}. \]

Since \( A = A^* \) commutes with \( W \), we find that
\[ \varphi(\vec{w}, t) = \left( e^{tW(\vec{w}/\sqrt{t})}, e^{tA(\vec{w}/\sqrt{t})} \right) = \left( e^{tW^*(\vec{w}/\sqrt{t}) + A^*(\vec{w}/\sqrt{t})} \right), \]

Since \( R^*(\vec{w}, t) = W^*(\vec{w}/\sqrt{t}) + A^*(\vec{w}/\sqrt{t}) \), the theorem follows. ■

The asymptotic behavior of \( \varphi(\vec{w}, t) \) is determined by the next lemma:

**Definition:**

a) For any matrix \( A \), \( \text{adj} A \) is the transposed matrix of cofactors of \( A \).  
b) If \( B(t) \) and \( A \) are \( M \times M \) matrices, we say that \( B(t) = A + o(1) \) as \( t \to t_0 \) if each entry of \( B(t) \) approaches the corresponding entry of \( A \) as \( t \to t_0 \).

**Lemma 6.3:**

Let \( R(t) \) be an \( M \times M \) matrix such that \( R(t) = Q + o(1) \) as \( t \to \infty \). Let \( r(t) = \det R(t) \).

If zero is a simple eigenvalue of \( Q \) and \( \gamma = \lim_{t \to \infty} r(t) \) exists and is finite, then
\[ \lim_{t \to \infty} e^{tR(t)} = \frac{1}{\rho} e^{\gamma/\rho} \text{adj} Q, \]
where \( \rho \) is the product of the non-zero eigenvalue of \( Q \).

Proof:
Let \( \mu_1, \mu_2, \ldots, \mu_M \) be the (not necessarily distinct) eigenvalues of \( Q \) arranged in lexicographical order: \( \Re \mu_i \leq \Re \mu_{i+1} \), and
\[
\Im \mu_i \leq \Im \mu_{i+1} \quad \text{if} \quad \Re \mu_i = \Re \mu_{i+1}.
\]
It is known that zero is always an eigenvalue of \( Q \) (the row sums of \( Q \) are all zero), and that the non-zero eigenvalues of \( Q \) have negative real parts.
(c.f. [2], page 52.) Hence, \( \mu_M = 0 \). Let \( \mu_1(t), \ldots, \mu_M(t) \) be the eigenvalues of \( R(t) \) similarly arranged. Since,
\[
R(t) = Q + o(1) \quad \text{as} \quad t \to \infty,
\]
\[
\mu_k(t) = \mu_k + o(1) \quad \text{as} \quad t \to \infty, \quad k = 1, 2, \ldots, M.
\]
(In particular, \( \mu_M(t) = o(1) \) as \( t \to \infty \).

If \( A \) is any matrix possessing eigenvalues \( \nu_1, \ldots, \nu_r \), with multiplicities \( m_1, \ldots, m_r \), then
\[
e^A = \sum_{i=1}^{r} \frac{1}{(m_i-1)!} \left[ \frac{d^{m_i-1}}{dt^{m_i-1}} \frac{\nu_i \adj(\nu_i A)}{(\nu - \nu_j)^{m_j}} \right] \ nu = \nu_i
\]
(Sylvester's theorem, c.f. [2], page 32.)

If \( m_r = 1 \),
\[
e^A = \frac{\nu_r \adj(\nu_r I - A)}{(\nu - \nu_j)^{m_j}} \sum_{j \neq r} \nu_i B_i(\nu_i) + \sum_{j \neq r} \nu_i B_i(\nu_i),
\]
where \( B_i \) is a matrix whose entries are rational functions of \( \nu_i \).
Since $\mu_1(t), \ldots, \mu_{M-1}(t)$ all have negative real parts for all $t$ sufficiently large,

\[
e^{tR(t)} = \frac{e^{t\mu_M(t)}}{\prod_{j \neq M}(\mu_M(t) - \mu_j(t))} \left[ \text{adj} (\mu_M(t) I - R(t)) \right] + o(1) \quad \text{as } t \to \infty,
\]

\[
= \frac{e^{t\mu_M(t)}}{\rho + o(1)} \left[ \text{adj} R(t) + o(1) \right] + o(1)
\]

\[
= e^{t\mu_M(t)} \left[ \frac{\text{adj} Q + o(1)}{\rho + o(1)} \right] \quad \text{as } t \to \infty.
\]

Since $\det R(t) = \prod_{j=1}^{M} \mu_j(t)$,

\[
\mu_M(t) = \frac{r(t)}{\rho + o(1)} \quad \text{as } t \to \infty.
\]

\[
\therefore \quad \lim_{t \to \infty} e^{tR(t)} = \lim_{t \to \infty} \frac{1}{\rho} e^{t r(t)/\rho} \text{adj } Q,
\]

if the right hand limit exists.

The application of lemma 6.3 to theorem 6.2 is facilitated by

**Lemma 6.4:**

Let $B$ be a square matrix whose row sums all vanish. Let $B(i,j)$ be the $(i,j)^{th}$ cofactor of $B$. Then for every $i$ and $j$,

\[
B(i,j) = B(i,i).
\]
Proof:

Let $D(i,j)$ be the determinant of the matrix obtained from $B$ by replacing the $i^{th}$ entry of the $i^{th}$ row of $B$ by 1, the $j^{th}$ entry of the $i^{th}$ row by $-1$, and all other elements of the $i^{th}$ row by zero. Expand $D(i,j)$ by cofactors of the $i^{th}$ row and we find that

$$D(i,j) = B^{(i,i)} - B^{(i,j)}.$$  

Since the row sums of the modified matrix are all zero,

$$D(i,j) = 0 . \quad \blacksquare$$

Theorem 6.5:

If zero is a simple eigenvalue of $Q$,

$$\lim_{t \to \infty} q(\hat{\omega}, t) = \exp \frac{1}{2} \sum_{i} \sum_{j \neq 1} \frac{q(i,j) Q^{(i,i)}}{\rho} \omega^{2}(i,j).$$

Proof:

Let $r(t) = \det R(\hat{\omega}, t)$, where $R(\hat{\omega}, t)$ is as defined in theorem 6.2. Clearly,

$$R(\hat{\omega}, t) = Q + o(1) \quad \text{as} \quad t \to \infty ,$$

so if $\gamma = \lim_{t \to \infty} r(t)$ exists, lemma 6.3 applies:

$$\lim_{t \to \infty} e^{t} R^{*}(\hat{\omega}, t) = \frac{1}{\rho} e^{\gamma/\rho} \text{adj Q}^{*}.$$
so that
\[ \lim_{t \to \infty} \varphi(\omega, t) = \frac{1}{\rho} e^{\gamma/\rho} \left( \text{adj } Q^* \right) (\eta). \]

Let \( q^*(i,j) \) be the \((i,j)\)th entry of \( \text{adj } Q^* \). Then
\[ q^*(i,j) = q(i,j) \]

and by lemma 6.4,
\[ q(i,j) = q(i,i). \]

Whence, since \( \eta \) is a probability measure,
\[ (\text{adj } Q^* \eta, \eta) = \sum_{i=1}^{M} q(i,i). \]

Let \( \Phi(\mu) = \det (\mu I - Q) = \prod_{j=1}^{M} (\mu - \mu_j). \)

Since \( \mu_M = 0 \) is a simple eigenvalue of \( Q \),
\[ \Phi(\mu) = \mu \prod_{j \neq M} (\mu - \mu_j) \]

and by the rule for differentiating determinants,
\[ \frac{d}{d\mu} \Phi(\mu) \bigg|_{\mu = 0} = (-1)^{M-1} \rho = (-1)^{M-1} \sum_{i=1}^{M} q(i,i) \]

(c.f. [3]).
Hence, if \( \lim_{t \to \infty} t \cdot r(t) = \gamma \) exists and is finite,

\[
\lim_{t \to \infty} \varphi(\pi, t) = e^{e^{\gamma}}.
\]

It remains to evaluate \( \gamma \): If either exists,

\[
\lim_{t \to \infty} t \cdot r(t) = \lim_{y \to 0} \frac{1}{y} \cdot \frac{1}{y} \cdot r\left(\frac{1}{y}\right).
\]

Since the right-hand side is indeterminate \( (\lim_{y \to 0} r(\frac{1}{y}) = \det Q = 0) \),
we resort to L'Hospital's rule:

If \( \lim_{y \to 0} \frac{\frac{d}{dy} \cdot r\left(\frac{1}{y}\right)}{y} \) exists and is finite, then

\[
\lim_{t \to \infty} t \cdot r(t) = \lim_{y \to 0} \frac{\frac{d}{dy} \cdot r\left(\frac{1}{y}\right)}{y}.
\]

By the rule for differentiating determinants (c.f. [3] again), \( \frac{\frac{d}{dy} \cdot r\left(\frac{1}{y}\right)}{y} \)
exists for every positive \( y \). In fact, 

\[
\frac{\frac{d}{dy} \cdot r\left(\frac{1}{y}\right)}{y} = \sum_{i} \sum_{j} R^{(i, j), \frac{1}{y}} \cdot \frac{\frac{d}{dy} r_{ij}(\pi, \frac{1}{y})}{y},
\]

where \( R^{(i, j), \frac{1}{y}} \) is the \( (i, j)\text{th} \) cofactor of \( R(\pi ; \frac{1}{y}) \): Explicitly:

\[
\frac{\frac{d}{dy} r_{ij}(\pi, \frac{1}{y})}{y} = \begin{cases} 
\frac{1}{2} \sum_{k \neq i} q(i, k) \cdot \omega(i, k) & \text{if } i = j \\
\frac{1}{2} \sum_{k \neq i} q(i, j) \cdot \omega(i, j) e^{-\omega(i, j)\sqrt{y}} & \text{if } i \neq j,
\end{cases}
\]

Since \( R^{(i, j), \frac{1}{y}} \to Q^{(i, j)} = Q^{(i, i)} \) as \( y \to 0 \), and since
\[ e^{-\omega(i,j)\sqrt{y}} = 1 - \omega(i,j)\sqrt{y} + O(y) \quad \text{as} \quad y \to 0, \]

we can write

\[
\frac{\partial}{\partial y} r\left(\frac{1}{y}\right) = \frac{1}{2\sqrt{y}} \sum_{i=1}^{M} \left[ \sum_{k \neq i} q(i,k) \omega(i,k)(R(i,i)(\frac{1}{y}) - R(i,k)(\frac{1}{y})) \right]
\]

\[
= \frac{1}{2} \sum_{i=1}^{M} \sum_{k \neq i} q(i,k) Q(i,i) \omega(\frac{1}{y}) + o(1) \quad \text{as} \quad y \to 0.
\]

If we can show that

\[
\lim_{y \to 0} \frac{1}{2\sqrt{y}} \sum_{i=1}^{M} \left[ \sum_{k \neq i} q(i,k) \omega(i,k)(R(i,i)(\frac{1}{y}) - R(i,k)(\frac{1}{y})) \right] = 0,
\]

we are done.

In particular, we can accomplish this by showing that

\[
\lim_{x \to 0} \frac{1}{x} \sum_{k \neq i} q(i,k) \omega(i,k) \left( R(i,i)(\frac{1}{x^2}) - R(i,k)(\frac{1}{x^2}) \right) = 0
\]

for every \( i \):

Let \( i \) be fixed, and let

\[
s_{n,m}(x) = \begin{cases} 
\sum_{k \neq i} q(i,k) \omega(i,k) & \text{if } n = m = i, \\
- q(i,m) \omega(i,m) & \text{if } n = i \text{ and } m \neq i, \\
r_{n,m}(\frac{1}{x^2}) & \text{if } n \neq i.
\end{cases}
\]

\( r_{n,m} \) is as defined in theorem 6.2.)
The matrix \( S(x) = \| s_{n,m}(x) \| \) is obtained from \( R(\tilde{\omega}, \frac{1}{x^2}) \) by replacing the \( i^{th} \) row of \( R \) by the row whose \( i^{th} \) element is \( \sum_{k \neq i} q(i,k) \omega(i,k) \), and whose \( j^{th} \) element is \( -q(i,j) \omega(i,j) \) if \( j \neq i \).

Let \( s(x) = \det S(x) \). If we expand \( s(x) \) by cofactors of its \( i^{th} \) row, we see that

\[
s(x) = \sum_{k \neq i} q(i,k) \omega(i,k)(R^{(i,1)}(\frac{1}{x^2}) - R^{(i,k)}(\frac{1}{x^2})) ,
\]

so that we need only show that

\[
\lim_{x \to 0} \frac{1}{x} s(x) = 0 .
\]

The entries of \( S(x) \) are all analytic functions of \( x \). So then, is \( s(x) \), and it has a power series expansion:

\[
s(x) = s(0) + x s'(0) + o(x) \quad \text{as} \quad x \to 0 ,
\]

\[
s(0) = \det S(0) = 0 \quad \text{since}
\]

the row sums of \( S(0) \) are all zero. Therefore,

\[
\lim_{x \to 0} \frac{s(x)}{x} = s'(0) .
\]

Since

\[
s'(x) = \sum_{i} \sum_{j} s^{(i,j)}(x) s'_{ij}(x) ,
\]

we can carry out the necessary differentiation:
\[ s'(x) = \left\{ \sum_{r \neq 1} \sum_{k \neq r} q(r,k) \omega(r,k)[S^{(r,r)}(x) - S^{(r,k)}(x)] \right\} + o(1) \]

\[ = \sum_{r \neq 1} \sum_{k \neq r} q(r,k) \omega(r,k)[S^{(r,r)}(0) - S^{(r,k)}(0)] + c(1). \]

Since the row sums of \( S(0) \) are all zero, lemma 6.4 applies:

\[ S^{(r,r)}(0) = S^{(r,k)}(0). \]

Whence,

\[ s'(0) = 0. \]

We have, as an immediate

**Corollary 6.6:**

For each \( i \) and \( j, j \neq i \), \( \frac{1}{T} \left[ N_T(i,j) - q(i,j) A_T(i) \right] \to 0 \)
in probability as \( T \to \infty \).

**Proof:**

Each \( \xi_T(i,j) \) converges in law to a normal distribution. Hence \( \frac{1}{\sqrt{T}} \xi_T(i,j) \) converges in law (and hence in probability) to zero. \( \blacksquare \)

Actually, we need a stronger result than this in order to carry out the proposed agenda. An investigation of the almost-sure convergence of \( \frac{1}{T} A_T(i) \) and \( \frac{1}{T} N_T(i,j) \) must be made and, as one might expect, such an investigation utilizes such concepts as metric transitivity and stationarity. A brief digression along these lines is now appropriate and the results of this discussion will point the way to our main theorem.
We say that the process \( \{Z(t), t \geq 0\} \) is strictly stationary (s.s.), if \( \Pr[Z(t) = i] \) does not depend upon \( t \) for any \( i \). Let
\[
\vec{\Lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_M \end{pmatrix}, \quad \text{where} \quad \lambda_i = \Pr[Z(0) = i].
\]

Since,
\[
\Pr[Z(t) = j] = \sum_j P_{ij}(t) \lambda_i,
\]
it seems that \( \vec{\Lambda} \) must possess some mysterious power if the left-hand side is not to depend upon \( t \). Specifically, \( \vec{\Lambda} \) must make the matrix equation,
\[
P^*(t)\vec{\Lambda} = \vec{\Lambda},
\]
work. \( \vec{\Lambda} \) can be characterized in the following way:

**Lemma 6.7:**

a) The process \( \{Z(t), t \geq 0\} \) is s.s. if and only if the initial distribution \( \vec{\Lambda} \) satisfies the equation \( Q^*\vec{\Lambda} = 0 \).

b) If zero is a simple eigenvalue of \( Q \), there is exactly one such \( \vec{\Lambda} \):
\[
\vec{\Lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_M \end{pmatrix}, \quad \text{with} \quad \lambda_i = Q^{(i,i)} / \rho.
\]
c) If zero is a simple eigenvalue of $Q$, then $\lim_{t \to \infty} \Pr[Z(t) = i] = \lambda_i$, ($i = 1, 2, \ldots, M$), independent of the initial distribution of the process.

Proof:

a) The process is s.s. if and only if

$$P^*(t) \vec{\lambda} = \vec{\lambda}.$$

Since

$$P'(t) = Q P(t) = P(t) Q; P(0) = I,$$

the conclusion follows.

b) Under the hypothesis of the lemma, the null space of $Q$ is one dimensional. The vector

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{pmatrix},$$

where $x_i = Q^{(i,i)}$, is always a solution of $Q^* \vec{x} = 0$, since

$$\sum_i q(i, j) x_i = \sum_i q(i, j) Q^{(i,i)}$$

$$= \sum_i q(i, j) Q^{(i,j)} = \det Q = 0.$$

All solutions of $Q^* \vec{\lambda} = 0$ are therefore multiples of $\vec{x}$. 
Let
\[ P = \lim_{t \to \infty} P(t). \]
(It is well known that the matrix limit exists when the limit is taken term-by-term.)

Since \( P(t) \) satisfies the Chapman-Kolmogorov condition:
\[ P(t+s) = P(t) P(s), \]

it follows that
\[ P^* = P^*(t_o) P^* \text{ for every } t_o \geq 0. \]

Since \( P \) is a stochastic matrix, the columns
\[ \tilde{p}(1), \tilde{p}(2), \ldots, \tilde{p}(M) \]
are probability distributions and satisfy
\[ P^*(t) \tilde{p}(k) = \tilde{p}(k), \quad k = 1, 2, \ldots, M, \]
and hence
\[ \tilde{q}^* \tilde{p}(k) = 0. \]

Whence,
\[ \tilde{p}(j) = c \tilde{x} \quad \text{for } j = 1, 2, \ldots, M. \]

Since \( \sum_{i=1}^{M} q_{(i,i)} = \rho \) (c.f. theorem 6.5), \( c = 1/\rho \).
c) For any other initial distribution, $\mathbf{\lambda}'$, 

$$\lim_{t \to \infty} P^*(t)\mathbf{\lambda}' = P^*\mathbf{\lambda} = \mathbf{\lambda}. \quad \square$$

**Lemma 6.8:**

The process $\{Z(t), t \geq 0\}$ is metrically transitive if and only if there is exactly one initial distribution $\mathbf{\lambda}$ which satisfies $Q^*\mathbf{\lambda} = 0$.

**Proof:**

It is well known that $\{Z(t), t \geq 0\}$ is metrically transitive if and only if there is exactly one $\mathbf{\lambda}$ satisfying $P^*(t)\mathbf{\lambda} = \mathbf{\lambda}$ (c.f. [6], page 238, in conjunction with page 511). \quad \square

The last results permit us to state

**Theorem 6.9:**

If zero is a simple eigenvalue of $Q$, and if $Q^{(k,k)} > 0$ for $k = 1, 2, \ldots, M$,

$$Pr\left[ \lim_{T \to \infty} \frac{1}{T} A_T(i) = \lambda_i \right] = 1$$

and

$$Pr\left[ \lim_{T \to \infty} \frac{1}{T} N_T(i,j) = q(i,j) \lambda_i \right] = 1$$

(Here, $\lambda_i = Q^{(i,i)}/\rho$. )
Proof:

If zero is a simple eigenvalue of $Q$, the process $\{Z(t), t \geq 0\}$ is metrically transitive. So then, are the processes $\{Y_i(t), t \geq 0\}$, where,

$$Y_i(t) = \begin{cases} 
1 & \text{if } Z(t) = i \\
0 & \text{otherwise, } i = 1, 2, \ldots, M .
\end{cases}$$

If the initial distribution on $\{Z(t), t \geq 0\}$ is

$$\vec{\lambda} = \begin{pmatrix} 
\lambda_1 \\
\lambda_2 \\
\vdots \\
\lambda_M
\end{pmatrix}, (\lambda_i = Q^{(1,1)}/\rho > 0), i = 1, 2, \ldots, M ,$$

then, by the ergodic theorem,

$$\lim_{T \to \infty} \frac{A_T(i)}{T} = \lim_{T \to \infty} \frac{1}{T} \int_0^T Y_i(t) \, dt = E Y_i(0) = \lambda_i$$

with probability one.

Since $\lambda_k > 0$ for each $k$,

$$\Pr[ \lim_{T \to \infty} \frac{1}{T} A_T(i) = \lambda_i | Z(0) = k] = 1 \text{ for each } k .$$

Hence, no matter what the initial distribution,

$$\Pr[ \lim_{T \to \infty} \frac{1}{T} A_T(i) = \lambda_i ] = 1 .$$

Let $h > 0$ be given and let $t_n = nh$. Let $X_k(i, j)$ be the number of transitions from $i$ to $j$ made in the interval $[(k-1)h, k \cdot h)$, $k = 1, 2, \ldots$.
If the $Z(t)$ process is strictly stationary, so then is the process

$$\{X_k(i,j), k = 1, 2, \ldots \}$$

Since

$$n \sum_{k=1}^{n} X_k(i,j)$$

the ergodic theorem tells us that

$$y(h, i, j) = \lim_{n \to \infty} \frac{1}{n} N^T_n(i,j) = \frac{1}{n} \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} X_k(i,j)$$

exists with probability one for every positive $h$.

By theorem 5.2,

$$\mathbb{E} N^T_n(i,j) = q(i,j) \lambda T$$

and

$$\text{Var} N^T_n(i,j) = 2q^2(i,j) \lambda T \int_0^T \int_x P_{ji}(x-t) \, dt \, dx + q(i,j) \lambda T^2 - q^2(i,j) \lambda T^2$$

if the process has the initial distribution $\hat{\lambda}$.

Since

$$\lim_{T \to \infty} \frac{1}{T^2} \text{Var} N^T_n(i,j) = q^2(i,j) \lambda \lim_{T \to \infty} (P_{ji}(T) - \lambda T)$$
and since (by lemma 6.7) \( \lim_{T \to \infty} P_{ji}(T) = \frac{Q^{(i,j)}}{\rho} \), Tchebycheff's inequality reveals that \( \frac{N_T(i,j)}{T} \to q(i,j) \lambda_1 \) in probability as \( T \to \infty \). Hence \( N_T(i,j) \)

\[
\frac{N_T(i,j)}{t_n} \to q(i,j) \lambda_1 \text{ in probability, so that } y(h,i,j) = q(i,j) \lambda_1
\]

identically in \( h \).

If \( \lambda \) is the initial distribution, then, \( \lim_{T \to \infty} \frac{N_T(i,j)}{T} = q(i,j) \lambda_1 \) with probability one. By proceeding as in the first part of this theorem, it follows that

\[
\frac{N_T(i,j)}{\Pr[\frac{N_T(i,j)}{T} \to q(i,j) \lambda_1]} = 1
\]

independent of the initial distribution. \( \blacksquare \)

The conditions that zero be a simple eigenvalue of \( Q \) and that \( Q^{(i,i)} \) be positive for every \( i \) are sometimes abbreviated by the phrase: "\( \{Z(t), t \geq 0\} \) is positively regular."

The main result can be stated as:

**Theorem 6.10:**

If \( \{Z(t); t \geq 0\} \) is positively regular, then

a) \( \lim_{T \to \infty} \hat{q}_T(i,j) = q(i,j) \)

with probability one, and

b) The joint distribution of the set of r.v.'s

\[
\left\{ \sqrt{T} \left[ q_T(i,j) - q(i,j) \right] \right\}^{M}_{i,j=1, \ i \neq j}
\]
is asymptotically normal and independent with zero mean and variances 
$q(i,j) \rho/q^{(i,i)}$, where $\rho$ is the product of the non-zero-eigenvalues of 
$Q$ and $Q^{(i,i)}$ is the $(i,i)^{th}$ cofactor of $Q$.

Proof:

a) From theorem 6.9,

$$\lim_{T \to \infty} N_T(i,j)/A_T(i) = q(i,j)$$

with probability one. Since $A_T(i)/T$ tends to a positive limit,

$$\lim_{T \to \infty} q_T(i,j) = \lim_{T \to \infty} N_T(i,j)/A_T(i) \text{ with probability one.}$$

b) By Cramér's theorem ([5], page 254), the (joint) asymptotic distribution of the set $\left\{ \sqrt{T}[q_T(i,j) - q(i,j)] \right\}_{j \neq i}$ is the same as that of the set

$$\left\{ \sqrt{T} \left[ \frac{N_T(i,j) - q(i,j) A_T(i)}{A_T(i)} \right] \right\}_{j \neq i}$$

and this, in term, has the same limiting distribution as

$$\left\{ \frac{\xi_T(i,j)}{\lambda_i} \right\}_{j \neq i}.$$

From theorem 6.5, the joint characteristic function of this set tends to

$$\exp \frac{1}{2} \sum_i \sum_{j \neq i} \frac{q(i,j) \rho q^{(i,i)}}{\lambda_i^2 \rho} \omega^2(i,j) \text{ as } T \to \infty.$$
But, \( \lambda_i = \frac{q(i,i)}{\rho} \).

As an afterthought, we point out that \( \hat{q}_T(i,j) = q(i,j) \) with probability one if \( q(i,j) = 0 \). For then, \( E_N_T(i,j) = 0 \) (by theorem 5.2) so that \( \Pr[N_T(0,i,j) = 0] = 1 \).

7. Acknowledgments:

I wish to thank Prof. Ulf Grenander for suggesting this investigation and for offering many fruitful ideas whenever things threatened to come to a grinding halt. I also want to thank the National Science Foundation for their financial support during the initial stages of this investigation.
Bibliography


Appendix I.

Proof of Theorem 5.2 (c,d,e):

Throughout this discussion we will let

\[ Y_i(t) = \begin{cases} 1 & \text{if } Z(t) = 1 \\ 0 & \text{otherwise} \end{cases} \quad i = 1, 2, \ldots, M, \]

\[ X_k(i,j) = \text{the number of jumps from } i \text{ to } j \text{ in} \]

\[ [(k-1)h, k h), \; k = 1, 2, \ldots \]

and

\[ \pi_i(t) = \Pr[Z(t) = i] \quad i = 1, 2, \ldots, M. \]

c) \( E \ A_n(i) A_T(r) = \int_o^T \int_o^T E Y_i(x) Y_r(t) \, dt \, dx \)

\[ = \int_o^T \int_o^T \Pr[Z(x) = 1 \& Z(t) = r] \, dt \, dx \]

\[ = \int_o^T \int_o^T P_{ri}(x-t) \pi_i(t) \, dt \, dx + \int_o^T \int_o^T P_{ir}(t-x) \pi_i(x) \, dx \, dt. \]

d) Divide the interval \([0, T]\) into \( n \) equal parts of length \( h = T/n \).

Then,

\[ N_T(i,j) = \sum_{k=1}^{n} X_k(i,j), \]

so that,

\[ E A_n(r) N_T(i,j) = \sum_{k=1}^{n} \int_o^T E [Y_r(t) X_k(i,j)] \, dt. \]
By lemma 5.1,

\[ E[Y_r(t)X_k(i,j)] = Pr[Z(t) = r \text{ and } X_k(i,j) = 1] + o(h) \]

as \( h \to 0 \), so that

\[
\begin{align*}
E A_T(r) N_T(i,j) &= \sum_{k=1}^{n} \left\{ \int_{0}^{kh} Pr[Z(t) = r \text{ and } Z(kh) = i \text{ and } Z((k+1)h) = j] \, dt \\
&\quad + \int_{kh}^{(k+1)h} Pr[Z(kh) = i \text{ and } Z(t) = r \text{ and } Z((k+1)h) = j] \, dt \\
&\quad + \int_{(k+1)h}^{T} Pr[Z(kh) = i \text{ and } Z((k+1)h) = j \text{ and } Z(t) = r] \, dt \right\} \\
&\quad + o(1) \text{ as } h \to 0, \\
\end{align*}
\]

\[
= \sum_{k=0}^{n} \left[ \int_{0}^{kh} \pi_r(t) P_{ri}(kh-t) q(i,j) \, dt \right] h \\
+ \sum_{k=0}^{n} \left[ \int_{(k+1)h}^{T} \pi_i(kh) q(i,j) P_{jr}(t-(k+1)h) \, dt \right] h \\
+ o(1) \text{ as } h \to 0, \\
\rightarrow q(i,j) \int_{0}^{X} \left\{ \int_{0}^{T} \pi_r(t) P_{ri}(x-t) \, dt \, dx + \int_{0}^{x} \pi_i(x) P_{jr}(t-x) \, dt \, dx \right\} \\
= q(i,j) \int_{0}^{T} \int_{0}^{x} \left\{ \pi_r(t) P_{ri}(x-t) + \pi_i(t) P_{jr}(x-t) \right\} \, dt \, dx
\]

as \( n \to \infty \). (We have used lemma 5.1 repeatedly.)
e) Again, divide \([0, T]\) into \(n\) parts of length \(T/n = h\).

\[
E N_T(i,j) N_T(r,s) = \sum_{k=1}^{n} \sum_{m=1}^{n} E X_k(i,j) X_m(r,s)
\]

\[
= \sum_{k=1}^{n} \sum_{m=1}^{n} \text{Pr}[X_k(i,j) = 1 \& X_m(r,s) = 1] + o(1) \text{ as } h \to 0,
\]

\[
= \sum_{k=1}^{n} \sum_{m=1}^{n} \text{Pr}[X_k(i,j) = 1 \& X_m(r,s) = 1] + \sum_{k=2}^{n} \text{Pr}[X_k(i,j) = 1 \& X_{k-1}(r,s) = 1]
\]

\[
+ \sum_{k=1}^{n} \text{Pr}[X_k(i,j) = 1 \& X_k(r,s) = 1]
\]

\[
+ \sum_{m=3}^{n} \sum_{k=1}^{m-2} \text{Pr}[X_k(i,j) = 1 \& X_m(r,s) = 1] + \sum_{m=2}^{n} \text{Pr}[X_{m-1}(i,j) = 1 \& X_m(r,s) = 1]
\]

\[
+ o(1) \text{ as } h \to 0,
\]

\[
= \sum_{k=1}^{n} \sum_{m=1}^{k-2} \pi_r(m,h) q(r,s) h \ p_{s1}(k \cdot h - (m+1)h) q(i,j) h
\]

\[
+ \delta(i,j;r,s) \sum_{k=0}^{n-1} \pi_i(k \cdot h) q(i,j) h
\]

\[
+ \sum_{m=3}^{n} \sum_{k=1}^{m-2} \pi_i(k \cdot h) q(i,j) h \ p_{jr}(m \cdot h - (k+1)h) q(r,s) h + o(1) \text{ as } h \to 0.
\]

As \(h \to 0\), this expression tends to
\[ q(1,j) q(r,s) \int_0^T \int_0^x \left\{ \pi_i(t) \, P_{si}(x-t) + \pi_i(t) \, P_{ji}(x-t) \right\} \, dt \, dx \]
\[ + \delta(i,j;r,s) \, q(i,j) \int_0^T \pi_i(t) \, dt . \]

**Appendix II:**

Justification of the Integration-by-Parts in Equation 6.2.1

of Theorem 6.2:

It will suffice to show that the set function \( J_k \), where

\[(A2.1) \quad J_k(E) \]
\[= \Pr\left\{ \bigcap_i \bigcap_j \left[ N_T(i,j)=n_{ij} \right] \cap \left[ (A_T(1), \ldots, A_T(k-1), A_T(k+1), \ldots, A_T(M)) \in E \right] \mid Z(T)=k \right\} \]

is absolutely continuous with respect to the \( M-1 \) dimensional Lebesgue measure over \( A_k \) for each \( k = l, 2, \ldots, M \), each \( T > 0 \), and each choice of \( \vec{n} = (n_{12}, \ldots, n_{M,M-1}) \). It will then follow, that the mass function \( G_k(\vec{n}, \cdot, T) \) has no singular component with respect to Lebesgue measure over \( A_k \), and this, in turn, justifies integration by parts over \( A_k \).

Let \( G_k(E, \vec{n}) \) be the subset of sample functions of

\( \{Z(t,\cdot), 0 \leq t < T\} \) for which

\[ Z(T) = k, (A_T(1), \ldots, A_T(k-1), A_T(k+1), \ldots, A_T(M)) \in E , \]

and

\[ N_T(1,2) = n_{12}, N_T(1,3) = n_{13}, \ldots, N_T(M,M-1) = n_{M,M-1} . \]
In the notation of theorem 3.2,

\[(A2.2) \quad J_k(E) = \frac{1}{Pr[Z(T) = k]} \int_{S_k(E, \theta)} f(v) \, dv,\]

if \(E\) is a Borel set and

\[Pr[Z(T) = k] > 0. \quad (\text{If } Pr[Z(T) = k] = 0, \text{ } G_k(\theta, \cdot, T) \text{ vanishes identically and there is nothing to prove.})\]

Any sample function in \(S_k(E, \theta)\) must make a total of

\[r = r(\theta) = \sum_i \sum_{j \neq i} n_{i,j}\]

jumps in \([0, T]\), so that any sample function in \(S_k(E, \theta)\) is representable by a sequence

\[(z_0, z_1, \ldots, z_r, t_0, \ldots, t_{r-1}),\]

where the \(z\)'s are integers from 1, 2, \ldots, \(M\) (the succession of values taken on by \(Z(t)\)), and the \(t\)'s are positive real numbers whose sum is less than \(T\) (the successive occupancy times).

Let \(C_k(\theta)\) be the collection of all sequences of integers of length \(r+1\) which end in \(k\), for which the number of times the integer \(i\) precedes the integer \(j\) is \(n_{i,j}(i, j = 1, 2, \ldots, M)\). (If \(v = (z_0, z_1, \ldots, z_r, t_0, \ldots, t_{r-1}) \in S_k(E, \theta)\), then \((z_0, z_1, \ldots, z_r) \in C_k(\theta)\).)
Let $\hat{x} = (x_0, x_1, \ldots, x_{r-1}, k)$ be a generic element of $C_k(\mathbb{R})$, and let

$$(A2.3) \quad S_x = \{v : z_0 = x_0, z_1 = x_1, \ldots, z_{r-1} = x_{r-1}, z_r = k\}.$$ 

Let $I_x(j)$ be the set of indices, $v$, for which $x_v = j$.

If $v$ is a sample function in $S_x$, and $j \neq k$, the total time spent in the $j^{th}$ state is exactly

$$A_T(j) = \sum_{v \in I_x(j)} t_{v},$$

so that

$$(A2.4) \quad \int_{S_k(\mathbb{E}, \mathbb{R})} f_{Q}(v) \, d\sigma(v) = \sum_{x \in C_k(\mathbb{R})} \int_{S_x \cap D_x(E)} f_{Q}(v) \, d\sigma(v),$$

where

$$D_x(E) = \left\{v : \left(\sum_{v \in I_x(1)} t_{v}, \ldots, \sum_{v \in I_x(k-1)} t_{v}, \sum_{v \in I_x(k+1)} t_{v}, \ldots, \sum_{v \in I_x(M)} t_{v}\right) \in E \cup 0, \sum_{v \notin I_x} t_{v} \leq T \right\}.$$ 

If, for some $j$, $(j \neq k)$, $\sum_{v \neq j} n_{j,v} = 0$, then $J_k(*)$ vanishes identically over $A_k$. For if $\sum_{v \neq j} n_{j,v} = 0$ and $E \subseteq A_k = \bigcap_{v \neq k} [a_v > 0]$, $n_{j,v} = 0$, $z_v = \sum_{v \neq j} n_{j,v}$, $w = (\sum_{v \neq j} n_{j,v})$, $z_v = 0$.

$$\Pr[N_T = \hat{n} \& (A_T(1), \ldots, A_T(k-1), A_T(k+1), \ldots, A_T(M)) \in E \mid Z(T) = k]$$

$$\leq \Pr[\sum_{v \neq j} N_T(j,v) = 0 \& A_T(j) > 0 \mid Z(T) = k]$$

$$= \Pr[A_T(j) > 0 \& Z(t) = j \quad 0 \leq t < T \mid Z(T) = k] = 0.$$
We can therefore assume without loss of generality, that \( \hat{\mathbf{n}} \) is such that \( \sum_{\mathbf{y} \neq \mathbf{n}} n_{\mathbf{y} \mathbf{j}} > 0 \) for all \( j \neq k \).

In this case, \( I_{x}(j) \) is non-empty for all \( \hat{x} \in C_{k}(\hat{\mathbf{n}}) \) and all \( j \neq k \).

Referring to theorem 3.2,

\[
J_{k}(E) = \frac{1}{\Pr[Z(T) = k]} \sum_{\hat{x} \in C_{k}(\hat{\mathbf{n}})} \int_{S_{x} \cap D_{x}(E)} \tilde{f}_{Q}(v) \, d\sigma(v)
\]

\[
= \frac{1}{\Pr[Z(T) = k]} \sum_{(z_{0}, \ldots, z_{r-1}, k) \in C_{k}(\hat{\mathbf{n}})} \int_{D_{x}(E)} \frac{1}{T^{1}} \frac{1}{t_{j}} q(z_{j}, z_{j+1}) e^{-(q(z_{j}) - q(k)) t_{j}},
\]

where

\[
D_{x}(E) = \left\{ (t_{0}, \ldots, t_{r-1}) : \sum_{u \in I_{x}(k-1)} t_{u} \leq T \right\}.
\]

Let \( \alpha_{j}^{k} = q(z_{j}) - q(z_{k}) \). We are done if we can show that

\[
\int_{D_{x}(E)} \frac{1}{T^{1}} \frac{1}{t_{j}} e^{-\alpha_{j}^{k} t_{j}} dt_{j}
\]

vanishes whenever \( E \) has Lebesgue measure zero.

If \( I_{x}(k) \) is void, then

\[
\int_{D_{x}(E)} \frac{1}{T^{1}} \frac{1}{t_{j}} e^{-\alpha_{j}^{k} t_{j}} dt_{j} = \int_{D_{x}(E)} \frac{1}{T^{1}} \frac{1}{t_{j}} e^{-\alpha_{j}^{k} t_{j}} dt_{j},
\]

and by a simple change of variable, the last is equal to
\[ K_x \sum_{E \in E^*} \prod_{j \neq k} (y_j)^{c_x(j)-1} e^{-\alpha_{jk}y_j} dy_j \]

where \( c_x(j) \) is the number of elements in \( I_x(j) \), and \( E^* = \left[ 0 < \sum_{j \neq k} y_j < T \right] \).

If \( I_x(k) \) is not empty (\( c_x(k) > 0 \)), then, since \( \alpha_{kk} = 0 \),

\[ \int_{D_x'(E)} \prod_{j=0}^{r-1} e^{-\alpha_{jk}y_j} dt_j \]

\[ = \prod_{j \neq k} \prod_{i \in I_x^*(j)} e^{-\alpha_{jk}t_i} \int_{j \in I_x^*(k)} dt_j \int_{j \neq k} \prod_{i \in I_x^*(j)} dt_i , \]

\[ \left[ (y_1, \ldots, y_{k-1}, y_{k+1}, \ldots, y_M) \in E \right] \left[ 0 < y_k < T - \sum_{i \neq k} y_i \right] \]

where \( y_j = \sum_{i \in I_x^*(j)} t_i, j = 1, 2, \ldots, M \). The last is equal to:

\[ K'_x \sum_{E \in E^*} \prod_{j \neq k} (y_j)^{c_x(k)} (y_j)^{c_x(j)-1} e^{-\alpha_{jk}y_j} dy_j \]

In general then, if \( \hat{n} \) is such that \( \sum_{i \neq j} n_{ij} > 0 \), and if \( \Pr[Z(T) = k] > 0 \),

\[ J_k(E) = \int_{E} h_k(y_1, \ldots, y_{M-1}) dy_1, \ldots, dy_{M-1} \]

**Appendix III**:

\[ \frac{\partial}{\partial t} F_k = -q(k) F_k + \sum_{i \neq k} q(i,k) \Delta_{bk} F_i \]
Proof:

\[ F_k(n_{12}, \ldots, n_{M,M-1}, a_1, \ldots, a_{k-1}, a_{k+1}, \ldots, a_M; t + h) = \sum_{m=0}^{1} \Pr \left\{ \bigcap_{i \neq 1} \left[ N_{t+h}(i, j) = n_{ij} \right] \cap \left( \bigcap_{i \neq k} \left[ A_t(i) \leq a_i \right] \cap [Z(t+h) = k] \cap B_m \right) \right\} + O(h^2) \text{ as } h \to 0 \], (where \( B_m \) is the event: "\( m \) changes in \([t, t+h)\)."

Thus,

\[ F_k(n_{12}, \ldots, n_{M,M-1}, a_1, \ldots, a_{k-1}, a_{k+1}, \ldots, a_M; t)(1 - q(k)h) + \sum_{v \neq k} F_{v}(n_{12}, \ldots, n_{v,k-1}, \ldots, n_{M,M-1}, a_1, \ldots, a_{v-1}, a_{v+1}, \ldots, a_M; t) q(v,k)h \]

+ o(h) as \( h \to 0 \).

Thus,

\[ \lim_{h \to 0} \frac{1}{h} \left[ F_k(n, a; t+h) - F_k(n, a; t) \right] = -q(k) F_k + \sum_{v \neq k} q(v,k) \Delta_{vk} F_v . \]

The initial conditions are obvious:

If \( n_{ij} > 0 \) for some \( i \) and \( j \),

\[ F_k(n_{12}, \ldots, n_{M,M-1}, a_1, \ldots, a_M; 0) = 0 . \]

Otherwise,

\[ F_k(0, \ldots, 0, a_1; 0) = \Pr[Z(0) = k] \cdot * \]
STAFFORD UNIVERSITY
TECHNICAL REPORT DISTRIBUTION LIST
CONTRACT No. 69-23140
(SE 342-GSB)

Commander, Western Development Division, WHRT
P.O. Box 250
Inglewood, California

Chief, Research Division
Office of Research & Development
Office of Chief of Staff
U.S. Army
Washington 25, D.C.

Director
National Security Agency
Attn: NSBC
Fort George G. Meade, Maryland

Director
Operations Analysis Div., APOSEP
No. U.S. Air Force
Washington 25, D.C.

Director
Snow Ice & Permafrost Research Establishment
Corps of Engineers
1215 Washington Avenue
Wilmette, Illinois

Director
Lincoln Laboratory
Lexington, Massachusetts

Department of Mathematics
Michigan State College
East Lansing, Michigan

Department of Mathematics
University College
Rutgers University
New Brunswick, New Jersey

Document Library
U.S. Atomic Energy Commission
19th & Constitution Avenues N.W.
Washington 25, D.C.

Headquarters
Oklahoma City Air Material Area
United States Air Force
Tinker Air Force Base, Oklahoma

Institute of Statistics
North Carolina State College of A & S
Raleigh, North Carolina

Jet Propulsion Laboratory
California Institute of Technology
Attn: A.J. Eshleman
4800 Oak Grove Drive
Pasadena 3, California

Silverman
The RAND Corporation
1700 Main Street
Santa Monica, California

NASA
Attn: M.B. Jackson, Office of Aeronautics
1736 P Street, N.W.
Washington 25, D.C.

National Applied Mathematics Labs.
National Bureau of Standards
Washington 25, D.C.

Naval Inspector of Ordnance
U.S. Naval Ordnance Factory
Washington 25, D.C.

Attn: Mrs. C.D. Rock
Office, Asst. Chief of Staff, G-4
Research Branch, N & D Division
Department of the Army
Washington 25, D.C.

Office of Technical Services
Department of Commerce
Washington 25, D.C.

Technical Information Officer
Naval Research Laboratory
Washington 25, D.C.

Technical Director
Combat Development Department
Army Electronics Proving Ground
Fort Belvoir, Virginia

Technical Information Service
Attn: Reference Branch
P.O. Box 60
Oak Ridge, Tennessee

Dr. Forrester S. Acton
Department of Mathematics
Princeton University
Princeton, New Jersey

Mr. Irving B. Altman
Inspection & QC Division
Office, Asst. Secretary of Defense
Room 20870, The Pentagon
Washington 25, D.C.

Professor T.W. Anderson
Department of Statistics
Columbia University
New York 7, New York

Dr. Merle M. Andrey
Chief, Mathematics Division
Air Force Office of Scientific Research
Washington 25, D.C.

Professor Fred C. Andrews
Department of Mathematics
University of Oregon
Burgon, Oregon

Professor E.H. Kirshman
Department of Mathematics
University of Washington
Seattle 5, Washington

Dr. David Blackburn
Department of Mathematical Sciences
University of California
Beverly Hills, California

Dr. Charles Boll
173 Partridge Avenue
Menlo Park, California

Professor William G. Cochran
Department of Statistics
Harvard University
2 Divinity Avenue, Room 311
Cambridge 38, Massachusetts

Dr. Joseph Daly
Bureau of the Census
Washington 25, D.C.

Miss Besse B. Dev
Bureau of Ships, Code 310-B
Department of the Navy
Washington 25, D.C.

Dr. Walter L. Demms, Jr.
Operations Analysis Div., 15H/0
HQ, U.S. Air Force
Washington 25, D.C.

Mr. James J. Fleming
Head
Research Branch
U.S. Naval Research Laboratory
Washington 25, D.C.

Dr. Lincoln E. Gephart
Scientific Advisor
G.E. Army Research & Development
Lincoln Group, 460 No. 757
New York, New York

Mrs. Dorothy M.Gilford
Head
Statistics Branch, Code 433
Office of Naval Research
Washington 25, D.C.

Mr. Joseph A. Groves
Director
Bureau of Aeronautics (Q-11)
Department of the Navy
Washington 25, D.C.
Mr. Harold Guelch
Technical Program Analysis Office
U.S. Naval Air Materiel Test Center
Point Magu, California

Professor Donald Guthrie, Jr.
U.S. Naval Postgraduate School
Dept. of Mathematics & Mechanics
Monterey, California

Dr. Virgil Hinde
Office, Chief of Research & Dev.
U.S. Army, Research Division 3930B
Washington 25, D.C.

Professor Harold Kelling
Department of Statistics
University of North Carolina
Chapel Hill, N.C.

Professor Stanley Taeuson
4731 Pleasant Street
Des Moines, Iowa

Professor Carl K. Kossack
Department of Mathematics
Purdue University
Lafayette, Indiana

Professor Eugene Lukacs
Department of Mathematics
Catholic University
Washington 16, D.C.

Dr. Craig Maguire
Department of Mathematics
U.S. Naval Postgraduate School
Monterey, California

Dr. Clifford Meloney
Chief, Statistics Branch
Chemical Corps Biological Warfare Lab.
Fort Detrick, Maryland

Professor Paul Meyer
Department of Mathematics
Washington State College
Pullman, Washington

Dr. Knox T. Millspaugh
Chief Scientist
Air Force Missle Dev. Center
Holloman Air Force Base, New Mexico

Professor J. Neyman
Dept. of Mathematical Statistics
University of California
Berkeley, California

Mr. Monroe Morse
College of Engineering
New York University
401 W. 20th St.
New York 3, N.Y.

Mr. P. H. Myers
B Iverson Division
2000 Ave. M
Fort Monmouth, New Jersey

Professor Edwin G. Olds
Department of Mathematics
College of Engineering & Sciences
Dartmouth Institute of Technology
Pittsburgh 13, Pennsylvania

Dr. William R. Pabst
Quality Control Division
Bureau of Ordnance
Department of the Navy
Washington 25, D.C.

H. Walter Price, Chief
Sensitivity Branch, 750
Second Ordnance Pune Laboratory
Room 105, Building 83
Washington 25, D.C.

Professor Ronald Pyke
Department of Mathematical Stat.
Columbia University
New York 27, New York

Professor Stanley Reiter
Department of Mathematics
Purdue University
Lafayette, Indiana

Dr. Paul Rider
Weight Development Center, WSRB
Weight-Patterson A.F.B., Ohio

Professor Herbert Robbins
Dept. of Mathematical Statistics
Columbia University
New York 27, N.Y.

Professor Herman Rubin
Department of Mathematics
University of Oregon
Eugene, Oregon

Major Oliver A. Shaw
Air Force Office of Scientific Research
Attn: BMM, Room 218
Temporary "T" Building
Washington 25, D.C.

Dr. Milton Seibl
Bell Telephone Laboratories
705 Union Blvd.
Allentown, Pennsylvania 1

Mr. G.F. Steck
Division 7Q1
Sandia Corp., Sandia Base
Albuquerque, New Mexico

Dr. Bruce H. Trest
Head, Applied Mathematics Branch
Naval Research Laboratory
Washington 25, D.C.

Professor Donald Dwyer
Department of Mathematics
University of Kansas
Lawrence, Kansas

Professor John W. Tukey
Department of Mathematics
Princeton University
Princeton, New Jersey

Mr. Harry Weinberg
Special Project Office
Bureau of Ordnance
Department of the Navy
Washington 25, D.C.

Mr. Joseph Weinsteins
Applied Physics Branch
Rivas Signal Laboratory
Fort Monmouth, New Jersey

Dr. Irving Weiss
Bell Telephone Laboratory
1500 Agassiz Street
Andover, Massachusetts

Dr. F. J. Welsh, Director
Mathematical Sciences Division
Office of Naval Research
Washington 25, D.C.

Dr. John Wilkes
Office of Naval Research
Code 200
Washington 25, D.C.

Mr. R. W. Williams
Office, DOT/US for Logistics
Department of the Army
Washington 25, D.C.

Professor Jacob Wolfowitz
Department of Mathematics
Cornell University
Ithaca, New York

Additional copies for project
leader and assistants and reserve
for future requirements

50