ESTIMATION OF THE PARAMETER IN THE STOCHASTIC MODEL FOR PHAGE ATTACHMENT TO BACTERIA

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY

Ramesh C. Srivastava

1965

This is to certify that the

thesis entitled

ESTIMATION OF THE PARAMETER IN THE
STOCHASTIC MODEL FOR PHAGE
ATTACHMENT TO BACTERIA

presented by

Ramesh C. Srivastava

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Statistics

Leo Katz
Major professor

Date May 24, 1965

O-169

ABSTRACT

ESTIMATION OF THE PARAMETER IN THE STOCHASTIC MODEL FOR PHAGE ATTACHMENT TO BACTERIA

by Ramesh C. Srivastava

Recently Gani has considered the following stochastic model for phage attachment to bacteria in suspension. Let n_{00} be the number of bacteria and ν_{00} be the number of phages at time t = 0. Also let $m = \frac{\nu_{00}}{n_{00}}$ be the multiplicity of phages and r be the saturation capacity of a bacterium. Further let $n_i(t)$ $(i = 0,\ldots,r)$ be the number of bacteria with exactly i phages attached to them at time $t \geq 0$. If $P(n_0,\ldots,n_r;t)$ denotes the probability that there are $n_0,\ldots,n_r$ bacteria with $0,\ldots r$ phages attached to them respectively at time $t \geq 0$, then, under certain assumptions, it is shown by Gani that

$$P(n_0,\ldots,n_r;t) = \frac{n_{00}!}{n_0!\ldots n_r!} \prod_{i=0}^{r} (a_{0i}(t))^{n_i(t)}$$

that is, at any fixed time t, $0 \leq t \leq t_0$, the distribution of $\underset{\sim}{n}(t) = (n_0(t),\ldots,n_r(t))$ is multinomial. The probabilities $a_{0j}(t)$ are functions of a single parameter α, defined by

$$a_{0j}(t) = \sum_{i=0}^{j} (-1)^{j-i} \binom{r}{i} \binom{r-i}{j-i} e^{-(r-i)\alpha\rho(t)}$$

where $$p(t) = \frac{1}{2} \log \left(\frac{r-m \exp(-\mu\alpha t)}{r-m} \right) .$$

In this thesis we investigate some of the basic properties of this model, describe a method of estimating the parameter α, and study the asymptotic properties of the estimate.

In Chapter 1, we describe the deterministic and the stochastic models for phage attachment to bacteria and review different methods of estimation for Markov chains with continuous time parameter.

In Chapter 2, the joint probability generating function of $\underset{\sim}{n}(t_1)$ and $\underset{\sim}{n}(t_2)$ $(t_1 < t_2)$ is derived and is used for calculating the mixed moments of the process. The rest of the chapter is devoted to the study of the asymptotic distribution of $\underset{\sim}{n}(t)$ and the limiting joint distribution of $\underset{\sim}{n}(t_1),\ldots,\underset{\sim}{n}(t_k)$ $(t_1 < t_k)$ as n_{00} tends to infinity.

The problem of estimating the parameter α in the stochastic model is considered in Chapter 3. In section 3.2 a method of estimation is described and is shown to yield a consistent estimate satisfying certain conditions. Then we obtain a lower bound to the variance of a consistent estimate satisfying certain conditions and use our result to obtain the asymptotic efficiency of the estimate. Finally we indicate a simpler method of estimating the parameter α. The modified method of estimation yields an estimate with the same properties as that obtained by the original procedure.

ESTIMATION OF THE PARAMETER IN THE STOCHASTIC MODEL FOR PHAGE ATTACHMENT TO BACTERIA

by

Ramesh C. Srivastava

handra

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics

1965

To My Parents

ACKNOWLEDGMENTS

I wish to express my thanks to Professor Leo Katz for his inspiring guidance throughout this investigation. His comments and suggestions were very helpful.

Many thanks are also due to Professor Joseph Gani for suggesting the problem. His suggestions and many illuminating discussions were extremely helpful. I am also indebted to Professor J. F. Hannan for some useful suggestions and discussions that led to the present form of theorems 2.4.1 and 3.5.1.

The financial support and creative atmosphere provided by the Department of Statistics at the Michigan State University were invaluable. I am also grateful to the National Science Foundation for their financial support under contracts G-18976, GP-2496, and GP-4307, under the direction of Professor Leo Katz.

TABLE OF CONTENTS

CHAPTER 1

REVIEW OF PREVIOUS WORK

1.0. Introduction: Recently Gani [9]* has considered a stochastic model for the attachment of phages to bacteria. In the following we study some of the basic properties of this model and describe a method for estimating the parameter and studying the asymptotic properties of the estimate.

This chapter consists of two parts. In the first part, which includes sections 1.1 and 1.2, we describe the deterministic model due to Yassky [20] and the stochastic model due to Gani for the attachment of phages to bacteria. In the second part, which includes sections 1.3, 1.4 and 1.5, we review different methods of estimation for finite Markov chains with continuous time parameter t.

Before describing mathematical models for the attachment of phages to bacteria, it may be useful to give a brief account of some essential facts about bacteriophage. It is known from plaque tests on bacterial suspensions, that phages attack and destroy bacterial colonies, after first attaching themselves to some of the bacteria. A variety of other experiments indicate that one or more phages may attach themselves to a bacterium.

In practice, a suspension of n_{00} bacteria in a nutrient medium are infected with ν_{00} phages at time $t = 0$. Within a short time (4 or 5 minutes) one or more phages attach themselves to a

*Numbers in the square bracket refer to the bibliography.

bacterium and infect it. Inside an infected bacterium phage particles reproduce following a complex reproduction cycle and then the bacterial cell bursts out, releasing some 200 to 300 phage particles. The disintegration of the bacterium and the scattering of new phage particles is called lysis. The entire reproductive cycle takes place in about 20-25 minutes. Meanwhile, the uninfected bacteria continue to reproduce. The new phages in turn attach themselves to other bacteria and infect them. This cycle goes on for awhile. Since the reproduction of phages is faster than that of bacteria, after some time all bacteria are killed. The probability of extinction of a bacterial colony by phages has been calculated by Gani [8].

Here we are mainly concerned with the phenomenon of phage attachment to bacteria. The mathematical models which are described in the sections 1.1 and 1.2 are derived under the following assumptions.

Assumptions.

A_1. The duration of experiment is taken to be very small, so that the number of bacteria or phages neither increases nor decreases in this period.

A_2. No attachment occurs between like particles.

A_3. Collision of particles is due only to Brownian motion, both bacteria and phages being non-motile.

A_4. On the basis of Brenner's work [5], we assume that there exists a maximum number of phage particles that can become attached to a bacterium. This number is called the saturation capacity of a bacterium.

1.1. A Deterministic Model for the Attachment of Phages to Bacteria:

Recently Yassky [20] has considered a deterministic model for the attachment of phages to bacteria in suspension. Let $n_i(t)$ be the number of bacteria with exactly i $(i=0,1,\ldots,r)$ phage particles attached to them at time $t \geq 0$ and let $\nu_0(t)$ be the concentration of free phage particles. Let $n_{00} = n_0(0)$ and $\nu_{00} = \nu_0(0)$, then

$$\sum_{i=0}^{r} n_i(t) = n_{00} \qquad \text{and} \qquad \sum_{i=0}^{r} i\, n_i(t) = \nu_{00} - \nu_0(t).$$

Assumption B_1: The rate of attachment of a phage particle to a bacterium which is already attacked by i phage particles is proportional to the product of $n_i(t)$ and $\nu_0(t)$.

Under assumptions A_1 through A_4 and assumption B_1, it is shown by Yassky that

$$\frac{dn_0}{dt} = -\lambda_0 n_0 \nu_0$$

$$\frac{dn_i}{dt} = (\lambda_{i-1} n_{i-1} - \lambda_i n_i)\, \nu_0 \qquad (i=1,\ldots,r-1) \qquad (1.1.1)$$

$$\frac{dn_r}{dt} = (\lambda_{r-1} n_{r-1})\, \nu_0$$

and

$$\frac{d\nu_0}{dt} = -\nu_0 \sum_{i=0}^{r} \lambda_i n_i \qquad (1.1.2)$$

where $\lambda_i > 0$ is a constant of proportionality, and as a first approximation $n_i(t)$ are regarded as continuous and differentialbe functions of t.

Assumption B_2: Assume that $\lambda_i = (r - i)\,\alpha \qquad 0 \le i < r$

$\qquad = 0 \qquad\qquad i \ge r$

where $\alpha > 0$.

Solving (1.1.2), we get

$$\nu_0(t) = \frac{\nu_{00}(r-m)}{r\exp(\mu\alpha t)-m} \tag{1.1.3}$$

where $m = \frac{\nu_{00}}{n_{00}}$ is the multiplicity of phages, and $\mu = n_{00}(r-m)$.

Next, solving (1.1.1), we get

$$n_0(t) = n_{00}[e^{-\mu\alpha t} + \frac{r}{r-m}(1 - e^{-\mu\alpha t})]^{-r},$$

$$n_i(t) = n_{00}\binom{-r+i-1}{i}\left[\frac{n_0(t)}{n_{00}}\right]\left[1-\left(\frac{n_0(t)}{n_{00}}\right)^{-\frac{1}{r}}\right]^i$$

$$n_r(t) = n_{00}\left(1 - \sum_{j=0}^{r-1}\frac{n_i(t)}{n_{00}}\right). \tag{1.1.4}$$

1.2. A Stochastic Model for the Attachment of Phages to Bacteria:

Let n_{00} be the number of bacteria at time $t = 0$ in suspension and $n_i(t)$ $(i=0,1,\ldots,r)$ denote the number of bacteria with exactly i phages attached to them at time $t \ge 0$. Let $P = P(n_0,\ldots,n_r;t)$ be the probability that there are $n_0,n_1,\ldots,n_r$ bacteria with $0,1,\ldots,r$ phages attached to them at time $t \ge 0$ and $\nu_0(t)$ denote the number of unattached phages. In this model, the deterministic value of $\nu_0(t)$ given by (4) in [9] is taken.

Assumption C: In addition to the assumptions A and B_1, we make the following assumption.

C. The probability of attachment during interval $(t,t + dt)$

of a phage to a bacterium already having i phages is

$$\lambda_i n_i \nu_0 dt + o(dt) \qquad (i=0,1,\ldots,r-1)$$

and the probability of attachment of more than one phage is o(dt).

Then we obtain in the usual way

$$\frac{dP}{dt} = -\sum_{i=0}^{r-1} \lambda_i n_i \nu_0 P + \sum_{i=0}^{r-1} \lambda_i (n_i + 1)\nu_0 \, P(n_0,\ldots,n_i+1,$$

$$n_{i+1} - 1,\ldots,n_r;t).$$

If $\varphi(u_0,\ldots,u_r;t)$ denotes the probability generating function (p.g.f.) of these probabilities, then it follows that

$$\frac{\partial\varphi}{\partial t} = \sum_{i=0}^{r-1} \lambda_i \nu_0 \, (u_{i+1} - u_i) \, \frac{\partial\varphi}{\partial u_i}. \tag{1.2.1}$$

This is a particular case of the p.g.f. for the multivariate Markov process first considered by Bartlett [3], and has been solved for this particular case by Gani [9]. To solve the first order linear differential equation (1.2.1), we consider the auxiliary equations

$$\frac{dt}{-1} = \frac{du_r}{0} = \frac{d\varphi}{0} = \frac{du_i}{\lambda_i \nu_0 (u_{i+1} - u_i)} \qquad (i=0,1,\ldots,r-1) \tag{1.2.2}$$

These can be rewritten as

$$\frac{d}{dt}\begin{bmatrix} u_0 \\ \cdot \\ \cdot \\ u_{r-1} \\ u_r \end{bmatrix} = \nu_0 \begin{bmatrix} \lambda_0 & -\lambda_0 & & & \\ & \lambda_1 & -\lambda_1 & & \\ & & \cdot\ \cdot\ \cdot & & \\ & & & \lambda_{r-1} & -\lambda_{r-1} \\ & & & & \lambda_r \end{bmatrix} \begin{bmatrix} u_0 \\ \cdot \\ \cdot \\ u_{r-1} \\ u_r \end{bmatrix}$$

$$= \nu_0 L \underset{\sim}{u}\,. \tag{1.2.3}$$

The solution of this is of the form

$$e^{-L\rho(t)}\underset{\sim}{u} = c \qquad (1.2.4)$$

where $\underset{\sim}{u}$ is the column vector $(u_0, u_1, \ldots, u_r)'$, c a constant vector, L the matrix array of λ_i and $\rho(t) = \int_0^t \nu_0(\tau)d\tau$.

Thus we have

$$\varphi(\underset{\sim}{u};t) = \varphi_0(e^{-L\rho(t)}\underset{\sim}{u}) \qquad (1.2.5)$$

where φ_0 is some suitable function of the new variables.

To determine the form of φ_0, we take into account the initial conditions and the fact that

$$\varphi(\underset{\sim}{u};0) = u_0^{n_{00}} \quad . \qquad (1.2.6)$$

Then

$$\varphi_0(\underset{\sim}{u};0) = ([e^{-L\rho(t)}\underset{\sim}{u}]_0)^{n_{00}} \qquad (1.2.7)$$

where $[e^{-L\rho(t)}\underset{\sim}{u}]_0$ is the 0th element of the column vector $[e^{-L\rho(t)}\underset{\sim}{u}]$. We now proceed to calculate this element. Since by assumption the λ_i's are distinct and non-negative, the matrix L can be written in the canonical form

$$L = N\Lambda M = N \begin{bmatrix} 0 & & & & \\ & \lambda_1 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \lambda_r \end{bmatrix} M \qquad (1.2.8)$$

where $N = M^{-1}$ and M is the matrix whose rows are eigenvectors corresponding to the eigenvalues $\lambda_0, \lambda_1, \ldots, \lambda_r$ respectively.

It is easy to see that

$$M_{ii} = 1 \qquad M_{ij} = \prod_{s=i+1}^{j} \frac{\lambda_{s-1}}{\lambda_s - \lambda_i} \qquad (r \geq j > i)$$

$$N_{ii} = 1 \qquad N_{ij} = \prod_{s=i}^{j-1} \frac{\lambda_s}{\lambda_s - \lambda_j} \qquad (r \geq j > i) \qquad (1.2.9)$$

It follows from (1.2.8) that

$$e^{-L\rho(t)} \underset{\sim}{u} = N\, e^{-\Lambda\, \rho(t)} M\, \underset{\sim}{u} \qquad (1.2.10)$$

Let $e^{-L\rho(t)} = N\, e^{-\Lambda\, \rho(t)}\, M = \| \, a_{ij}(t) \, \|$ (1.2.11)

where $a_{ij}(t) = 0$ for $i > j$ and $\sum_{j=0}^{r} a_{0j}(t) = 1$. Then we see from (1.2.7) and (1.2.11), that

$$\varphi(\underset{\sim}{u};t) = \left(\sum_{j=0}^{r} a_{0j}(t)\, u_j \right)^{n_{00}}$$

gives the p.g.f.ofa multinomial distribution and so

$$P(n_0,\ldots,n_r;t) = \frac{n_{00}\,!}{n_0! \ldots . n_r!} \prod_{i=0}^{r} a_{0i}(t)^{n_i(t)} \qquad (1.2.12)$$

where $$a_{0j}(t) = \sum_{i=0}^{j} N_{0i}\, M_{ij}\, e^{-\lambda_i \rho(t)} \quad . \qquad (1.2.13)$$

<u>The Particular Case $\lambda_i = (r - i)\,\alpha$:</u> The above result holds generally for any λ_i; now we discuss the particular case obtained by setting

$$\lambda_i = (r - i)\alpha \qquad (i=0,1,\ldots,r-1) \text{ and } \lambda_r = 0.$$

It also follows from (1.2.9) that

$$M_{ij} = 0 \qquad \text{for } i > j$$

$$M_{ii} = 1$$

$$M_{ij} = (-1)^{j-i} \binom{r-i}{j-i} \quad \text{for } j = i+1,\ldots,r. \tag{1.2.14}$$

and

$$N_{ij} = 0 \quad \text{for } i > j$$

$$N_{ii} = 1 \tag{1.2.15}$$

$$N_{ij} = \binom{r-i}{j-i} \qquad \text{for } j = i+1,\ldots,r.$$

Hence, from (1.2.13), (1.2.14) and (1.2.15) we have

$$a_{0j}(t) = \sum_{i=0}^{j} (-1)^{j-i} \binom{r}{i} \binom{r-i}{j-i} e^{-(r-i)\alpha\rho(t)} \tag{1.2.16}$$

where

$$\rho(t) = \frac{1}{\alpha} \log\left(\frac{r-m \exp(-\mu\alpha t)}{r-m} \right) \quad \text{and } \mu = n_{00}(r-m).$$

In the present investigation this particular multivariate stochastic process is discussed in some detail, and the problem of estimating the parameter α is considered.

1.3. Estimation Methods for Evolutive Processes:

In an early paper Kendall [10] considered the problem of estimating the birth rate for a purely reproductive process. Let n_0 be the number of individuals at time $t = 0$ and suppose each individual is capable of giving birth to a new individual in accordance with the following:

(a) The sub-populations generated by two co-existing individuals develop in complete independence of one another,

(b) the probability that an individual existing at time t will reproduce a new individual during $(t,t + dt)$ is

$$\lambda dt + o(dt)$$

and the probability of more than one birth is $o(dt)$.

Let $P_n(t)$ be the probability that at time t, there are n individuals. If $n_0 = 1$, then as shown in Kendall [10]

$$P_n(t) = e^{-\lambda t} (1-e^{-\lambda t})^{n-1} \qquad (n \geq 1). \qquad (1.3.1)$$

If $n_0 = a > 1$, then it follows from (1.3.1), that

$$P_n(t) = \binom{n-1}{a-1} e^{-\lambda ta} (1 - e^{-\lambda t})^{n-a}. \qquad (1.3.2)$$

Now we consider a pure birth process and suppose observations are taken at times

$$0 \quad \tau \quad 2\tau \quad . \quad . \quad . \quad k\tau = T$$

and the observed sizes of the population are respectively

$$n_0 \quad n_1 \quad n_2 \quad . \quad . \quad . \quad n_k.$$

The conditional probability $P(n_1,\ldots,n_k \mid n_0)$ is

$$P(n_1,\ldots,n_k \mid n_0) = \prod_{i=0}^{k-1} P(n_{i+1} \mid n_i)$$

and by (1.3.2),

$$P(n_{i+1} \mid n_i) = \binom{n_{i+1}-1}{n_i \quad -1} e^{-n_i\lambda\tau} (1 - e^{-\lambda\tau})^{n_{i+1}-n_i}.$$

Therefore the log likelihood function is

$$L = \text{constant} + (n_k-n_0) \log(1-e^{-\lambda\tau}) - \lambda\tau \sum_{i=0}^{k-1} n_i. \qquad (1.3.3)$$

Differentiating (1.3.3) with respect to λ, we get

$$\frac{\partial L}{\partial \lambda} = \frac{(n_k-n_0)}{1 - e^{-\lambda\tau}} \tau e^{-\lambda\tau} - \tau \sum_{i=0}^{k-1} n_i$$

$$= \frac{(n_k - n_0)}{e^{\lambda\tau} - 1}\,\tau - \tau \sum_{i=0}^{k-1} n_i \ .$$

Thus the maximum likelihood estimate $\hat{\lambda}$ of λ is given by

$$e^{\hat{\lambda}\tau} = \frac{n_1 + \ldots + n_k}{n_0 + \ldots + n_{k-1}} \ .$$

Kendall also suggested another rough procedure for estimating λ and obtained an expression for its asymptotic variance.

In [11] and [12] Moran investigated the problem of estimation for some simple processes; for example, the Poisson process, the pure birth process, birth and death processes, etc.

For a pure birth process, Moran suggested the following alternative procedure:

Let $N(t)$ denote the size of the population at time t; clearly $N(t)$ is a non-decreasing function of t. The number of births during an interval $(0,T]$ is then given by $k=N(T)-N(0)$. Suppose these births occurred at times $t_1,\ldots,t_k$.

Now we consider a sample function of $N(t)$ with k jumps. Starting from time t_s, the time $(t_{s+1} - t_s)$ to the next jump is such that

$$2\lambda(N(0) + s)\,(t_{s+1} - t_s)$$

is distributed as χ^2 with 2 degrees of freedom. Then the log likelihood function is

$$L = \sum_{s=0}^{k-1} \log[\lambda(N(0) + s)\, e^{-\lambda(N(0)+s)(t_{s+1}-t_s)}] \qquad (t_0 = 0)$$

$$= \sum_{s=0}^{k-1} [\log \lambda + \log (N(0) + s) - \lambda(N(0) + s) \quad (t_{s+1} - t_s)].$$

Equating the derivative of L to zero, we get

$$\frac{\partial L}{\partial \lambda} = k\lambda^{-1} - \sum_{s=0}^{k-1} (N(0) + s)(t_{s+1} - t_s) = 0. \tag{1.3.4}$$

Therefore the maximum likelihood estimate $\hat{\lambda}$ of λ is

$$\hat{\lambda} = \frac{k}{\sum_{s=0}^{k-1} (N(0) + s)(t_{s+1} - t_s)} \; . \tag{1.3.5}$$

The sum $\sum_{s=0}^{k-1} (N(0) + s)(t_{s+1} - t_s)$ is equal to the area under the curve $N(t)$ and so

$$\hat{\lambda} = \frac{k}{\int_0^T N(t)dt} = \frac{N(T) - N(0)}{\int_0^T N(t)dt} \; .$$

Also $\frac{1}{k}\int_0^T N(t)dt$ is an unbiased estimate of λ^{-1} and its sampling distribution is $(2k\lambda)^{-1} \chi^2$ with $2k$ degrees of freedom and so its variance is $(k\lambda^2)^{-1}$.

Further, Moran considered the problem of estimation for the parameters in a birth and death process. Anscombe [2] can also be consulted for sequential estimation in a birth and death process.

<u>1.4. Maximum Likelihood Estimation for a Finite Markov Chain with Continuous Time Parameter</u>: Let $\{X(t), t \geq 0\}$ be a separable* finite Markov chain with continuous time parameter and let $P(t) = ||p_{ij}(t)||$ be the stationary transition matrix function. Then under certain conditions $P(t)$ can be written in the form $P(t) = \exp(tQ)$ where $Q = ||q_{ij}||$ is an MXM matrix known as the "infinitesimal generator" of the process.

For a finite Markov chain with continuous time parameter, two types of estimation problems have been studied. In [1], Albert has considered the problem of estimating $Q = ||q_{ij}||$, the "infinitesimal generator"of the process and in [4], Billingsley has considered the problem of estimating the parameter θ when $Q = ||q_{ij}(\theta)||$ (or equivalently when $P(t) = ||p_{ij}(t,\theta)||$) is a function of θ.

Assume that $p_{ij}(t)$ has a derivative $p_{ij}'(t)$ with respect to t for all $t \geq 0$ and let

$$q_i = \lim_{t \to 0} \frac{1-p_{ii}(t)}{t}$$

and

$$q_{ij} = \lim_{t \to 0} \frac{p_{ij}(t)}{t} \qquad i \neq j$$

and let $Q = ||q_{ij}||$ be the matrix where $q_{ii} = -q_i$.

It is well known (see for example Doob [6]) that the probabilities $p_{ij}(t)$ satisfy Kolmogorov's forward and backward equations.

* For definition and other questions regarding separability see Doob [6].

Let (Ω, S, P) be a probability space and $\{X(t),\ t \geq 0\}$ be a separable Markov process defined on this space and taking values in a finite set $\chi_0 = \{1,\ldots,M\}$. Then it is shown in Doob [6], chapter vi, that if $||p_{ij}(t)||$ is stationary transition matrix function satisfying the condition

$$\begin{aligned} \lim_{t\to 0} p_{ij}(t) &= 1 \qquad \text{for } i = j \\ &= 0 \qquad \text{for } i \neq j \end{aligned} \tag{1.4.1}$$

then the limits q_i and q_{ij} exist for all i and j. Further it follows from theorem 1.4. page 248 of [6], that almost all sample functions are step functions with a finite number of jumps in any finite interval.

(a) <u>The Estimation of the Infinitesimal Generator</u>:

For estimating the "infinitesimal generator" of a Markov chain, first Albert [1] constructs a density on the set of all realizations of the process; then the likelihood equation is defined and finally large sample properties of maximum likelihood estimates are studied.

Suppose observations are made on the process $\{X(t), 0 \leq t < T\}$ where T is finite. Due to theorem 1.4., page 248 of [6], the sample function is a step function. Let X_i denote the state after the ith jump and T_i be the time the system stays in state X_i. Then with probability one, a sample function

$X(\omega,\cdot)$, $0 \leq t < T$ can be written as an ordered sequence

$$\{(X_0,T_0),\ldots,(X_{\nu(T)-1},T_{\nu(T)-1}),\ X_{\nu(T)}\} \tag{1.4.2}$$

where $\nu(T)$ is the maximum n such that

$$T_0 + \ldots + T_{n-1} \le T < T_1 + \ldots + T_n .$$

Now we can obtain the probability distribution on the space of sample functions.

Theorem A(Albert): Let

$$q'(i,j) = \begin{cases} 0 & \text{if } i = j \\ q_{i,j} & \text{if } i \ne j \end{cases}$$

and

$$q(i) = q_i .$$

Then

$$P[\nu(T) = \nu;\ X_0 = x_0,\ T_0 \le \alpha_0; \ldots;\ X_{\nu-1} = x_{\nu-1},\ T_{\nu-1} \le \alpha_{\nu-1};$$

$$X_\nu = x_\nu] \quad = \quad P[X_0 = x_0]\ e^{-q(x_\nu)T} \quad \times$$

$$\int_{S_\nu} \prod_{j=0}^{\nu-1} dt_j\ q'(x_j, x_{j+1})\ e^{-(q(x_j) - q(x_\nu))t_j} \tag{1.4.3}$$

where $S_\nu = \{(t_0, \ldots, t_{\nu-1}) : \sum_{j=0}^{\nu-1} t_j < T \text{ and } 0 \le t_j \le \alpha_j\}$ if $\nu > 0$,

and

$$P[\nu(T) = 0,\ X_0 = x_0] \quad = \quad P[X_0 = x_0]\ e^{-q(x_0)T} .$$

Next we proceed to construct the density on the space of all sample functions.

Any sample function with ν jumps in $[0,T)$ can be represented as a point in

$$\chi_\nu = \prod_{j=0}^{\nu-1} (\chi_0 \times R^+) \times \chi_0$$

where $\chi_0 = \{1,\ldots,M\}$ and $R^+ = (0,\infty)$.

Let ℓ be the Lebesgue measure on R^+ and let c be the counting measure on χ_0. Let $\sigma^{(\nu)}$ be the product on χ_ν, defined by

$$\sigma^{(\nu)} = \{\prod_{j=0}^{\nu-1} (c \times \ell)\} \times c.$$

Then almost all sample functions of the process $\{X(t),\ 0 \le t < T\}$ can be represented as a point in

$$\chi = \bigcup_{\nu=0}^{\infty} \chi_\nu.$$

For any subset A of χ for which $A\cap\chi_\nu$ is $\sigma^{(\nu)}$ measurable, we define

$$\sigma^*(A) = \sum_{\nu=0}^{\infty} \sigma^{(\nu)}(A\cap\chi_\nu).$$

Thus σ^* is defined on the σ-field B^*, which is the smallest σ-field containing all subsets A whose projection on

$\{\prod_{j=0}^{\nu-1} (\chi_0 \times R^+)\} \times \chi_0$ is a measurable set for each ν. Let σ be

a measure on the space of all sample functions, defined for all subsets B such that $B\cap\chi \in B^*$,

$$\sigma(B) = \sigma^* (B\cap\chi).$$

Now we obtain the density on the set of all realizations of the process with respect to σ which is a σ-finite measure.

<u>Theorem B (Albert)</u>: If B is a subset of the space of all sample functions over $[0,T)$ which is measurable with respect to σ, then

$$P[B] = \int_B f_Q(v)\,d\sigma(v)$$

where

$$f_Q(v) = \begin{cases} P[X(0) = x_0]\, e^{-q(x_0)T} & \text{if } v = (x_0) \\[2ex] P[X(0) = x_0]\, e^{-q(x_\nu)T} \prod_{j=0}^{k-1} q'(x_j, x_{j+1}) e^{-[q(x_j)-q(x_\nu)]t_j} & \\[2ex] \text{if } v = ((x_0,t_0),\ldots,(x_{\nu-1},t_{\nu-1}),\ x_\nu) & \\ \text{with } \nu > 0,\ x_j \in \chi_0,\ t_j \geq 0\ (j=0,\ldots,\nu-1) & \\ \text{and } \sum_{j=0}^{\nu-1} t_j < T & \\[2ex] 0 & \text{otherwise.} \end{cases} \tag{1.4.4}$$

Now we define the likelihood function and obtain the maximum likelihood estimate. If k independent realizations $v_1,\ldots,v_k$ of $\{X(t),\ 0 \leq t < T\}$ are observed, then the likelihood function is defined by the equation

$$L_Q^k = \prod_{j=1}^{k} f_Q(v_j). \tag{1.4.5}$$

Let $N_T^k\,(i,j)$ be the total number of transitions from state i to j observed in k trials and $A_T^k(i)$ be the total length of time that state i is occupied during k trials. Then we can write

$$\log L_Q^k = c_k + \sum_i \sum_{j \neq i} N_T^{(k)}(i,j) \log q(i,j) - \sum_i A_T^{(k)}(i) q(i) \tag{1.4.6}$$

where c_k is finite with probability one and does not depend on Q.

From (1.4.6), we conclude that

(1) $\{N_T^{(k)}(i,j),\ A_T^{(k)}(i)\}_{j\neq i}$ is a sufficient statistic for Q.

(2) The maximum likelihood estimate $\hat{q}_T^k(i,j)$ of q_{ij} is given by

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

If $A_T^{(k)}(i) = 0$, then the maximum likelihood estimate does not exist and we define

$$\hat{q}_T^{(k)}(i,j) = 0 \quad \text{if} \quad i\neq j \text{ and } A_T^{(k)}(i) = 0. \tag{1.4.7}$$

In this context, the term 'large sample' can be interpreted in two ways. Many independent realizations of the process $\{X(t),\ 0 \leq t < T\}$, keeping T fixed could be observed or a single (finite k) realization of the process over long period of time may be observed. In both cases, Albert [1] proved that under certain conditions the maximum likelihood estimates are strongly consistent and asymptotically normally distributed.

(b) <u>Estimating the Parameter θ When the Infinitesimal Generator is a Function of θ</u>: Consider again a separable Markov process $\{X(t), t \geq 0\}$ defined on the probability space (Ω, S, P_θ) and taking values in a finite set $\chi_0 = \{1,\ldots;M\}$ where $\theta = (\theta_1,\ldots,\theta_r)$ is a parameter, taking values in an open subset S in r-dimensional Euclidean space R_r. Let $P(t,\theta)=||p_{ij}(t,\theta)||$ denote the stationary transition probability matrix function. In this case the condition expressed in (1.4.1) becomes

$$\begin{aligned} \lim_{t\to 0} p_{ij}(t,\theta) &= 1 \quad \text{if} \quad i = j \\ &= 0 \quad \text{if} \quad i \neq j. \end{aligned} \tag{1.4.8}$$

Then it follows from Doob [6, theorems 1.2 and 1.3 of chapter vi], that under certain conditions the limits

$$\lim_{t\to 0} \frac{1-p_{ii}(t,\theta)}{t} = q_i(\theta) < \infty$$

and (1.4.9)

$$\lim_{t\to 0} \frac{p_{ij}(t,\theta)}{t} = q_{ij}(\theta) \qquad i \neq j$$

exist for all i and j.

Now we state the assumptions which are needed for proving the asymptotic properties of the maximum-likelihood estimate.

<u>Assumptions</u>:

D_1. For each ω, the sample function is a right continuous step function. The limits in (1.4.8) exist for all i and j and so there exist functions $q_i(\theta)$ and $q_{ij}(\theta)(i \neq j)$ satisfying (1.4.9). For all θ and i, $q_i(\theta) > 0$.

Since by assumption D_1, the sample function is a right continuous step function, the system starts at time $t = 0$ in state x_0, remains there for time t_0, then it jumps to some other state say x_1, stays there for time t_1 and so on. If $\nu(T)$ denotes the number of jumps in time $[0,T)$, then it follows that $\{x_n\}$ is a Markov process and

$$P_\theta\{t_{n+1} > \alpha \mid t_0,\ldots,t_n;x_0,\ldots,x_{n+1}\} = e^{-q(x_{n+1},\theta)\alpha}, \qquad \alpha \geq 0 \qquad (1.4.10)$$

and

$$P_\theta\{x_{n+1}=j \mid t_0,\ldots,t_n;x_0,\ldots,x_n\} = \frac{q(x_n,j,\theta)}{q(x_n,\theta)} . \qquad (1.4.11)$$

From (1.4.10) and (1.4.11) it follows that $\{(x_n,t_n), n=1,2,\ldots\}$ is a Markov process on the Cartesian product $\chi_0 \times R^+$ where $R^+ = (0,\infty)$. This process is called the imbedded process and as remarked by Billingsley (page 38, [4]) the information contained in the sample $\{X(t), 0 \le t < T\}$ is essentially the same as that in the sample $\{(x_k,t_k); k=0,\ldots,\nu(T)-1\}$. The complete sample contains only slightly more information than the sample from the imbedded process. The additional information is the length for which the system is in state $x_{\nu(T)}$.

Let $\pi_{ij}(\theta) = P_\theta(x_{n+1}=j \mid x_n = i) \qquad i \neq j$.

Then $\qquad q_{ij}(\theta) = \pi_{ij}(\theta)\; q_i(\theta)$.

We assume that the set D of pairs (i,j) for which $\pi_{ij}(\theta) > 0$ (or equivalently, for which $q_{ij}(\theta) > 0$) is independent of θ. (Note that $(i,i) \notin D$ by construction). Hence, $d \le M(M - 1)$ where d is the number of elements in D.

Assumption D_2: The set D of pairs (i,j) such that $q_{ij}(\theta) > 0$ is independent of θ and the functions $q_{ij}(\theta)$ have continuous third order derivatives throughout S. The $d \times r$ matrix

$\left\| \frac{\partial q_{ij}(\theta)}{\partial \theta_u} \right\|$ has rank r for all $\theta \in S$.

Assumption D_3: For each $\theta \in S$, the Markov chain $\{x_n\}$ has only one ergodic set and there are no transient states.

Now given a sample $\{(x_k,t_k), 0 \le k \le \nu(T)\}$ we can write the log likelihood function from (1.4.6) by setting $k = 1$,

$$\log L_T^{(1)} = \sum_D N_T^{(1)}(i,j) \log q_{ij}(\theta) - \sum_i A_T^{(1)}(i)\, q_i(\theta) \qquad (1.4.12)$$

where $A_T^{(1)}(i)$ denotes total time for which the system is in state i and $N_T^{(1)}(i,j)$ denotes the total number of direct transitions from state i to state j. In writing the expression for the log likelihood function we have, following Billingsley, dropped the term $P_\theta(X(0) = x_0)$ because the information contained in $P_\theta(X(0) = x_0)$ does not increase as $\nu(T) \to \infty$.

Thus the likelihood equations are:

$$\frac{\partial}{\partial\theta_u} \log L_T^{(1)} = \sum_D N_T^{(1)}(i,j) \frac{\partial}{\partial\theta_u} \log q_{ij}(\theta) - \sum_i A_T^{(1)}(i) \frac{\partial q(i,\theta)}{\partial\theta_u} \qquad (1.4.13)$$

$$u = 1,\ldots,r.$$

Finally we state the asymptotic properties of the maximum-likelihood estimate. For a proof see Billingsley [4].

Theorem (Billingsley): If $(X(t), q_{ij}(\theta), S)$ satisfies conditions D_1, D_2 and D_3 and $\theta^0 \varepsilon S$ is the true value of the parameter, then there exists a consistent solution $\hat{\theta}$ of (1.4.13). Moreover,

$$\lim \nu(T)^{\frac{1}{3}} (\hat{\theta} - \theta^0) \underset{\rightarrow}{P} 0, \quad \text{and}$$

$$\lim \nu(T)^{\frac{1}{3}} (q_{ij}(\hat{\theta}) - q_{ij}(\theta^0)) \underset{\rightarrow}{P} 0.$$

1.5. An Estimation Procedure in the Emigration-immigration Process: While a great deal of general theory of maximum likelihood estimation for Markov chains with continuous time parameter has been developed, only a few particular problems for evolutive processes have been investigated thus far. In most of the cases the maximum likelihood estimation of the parameter is intractable.

In [17] Ruben suggested an interesting method of estimating the interaction parameter in an emigration-immigration process. We use this method of estimation for estimating the parameter α in the stochastic model for phage attachment to bacteria. Before describing Ruben's method of estimation, we describe the emigration-immigration process.

An emigration-immigration or Poisson Markov process [3] is a multivariate stochastic process. Suppose $R_1, \dots, R_m$ are m disjoint non-empty subsets of a space and $R^* = (R_1 \cup \dots \cup R_m)'$ where A' denotes the complement of a set A. Consider a system of particles performing some type of randcm motion.

Let $N(t) = \begin{bmatrix} N_1(t) \\ \cdot \\ \cdot \\ N_m(t) \end{bmatrix}$ be the number of particles in the regions $R_1, \dots, R_m$ respectively at time $t \geq 0$.

<u>Assumptions</u>:

E_1. The particles are moving independently of each other.

E_2. The probability that a particle moves from R_r to R_s $(r \neq s = 1, \dots, m)$ in time dt is

$$N_r(t)\, \lambda_{rs} dt + o(dt).$$

E_3. The probability that a particle moves from R_r to R^* in time dt is

$$N_r(t)\,\lambda^*_r dt + o(dt),\ r = 1,\dots,m.$$

E_4. The probability that a particle moves from R^* to R_r in time dt is

$$\mu_r dt + o(dt),\ r = 1,\dots,m.$$

E_5. In time dt, the probability of movement of more than one particle is $o(dt)$.

Let $$\Delta = \begin{bmatrix} \lambda_1^* + \sum_{j\neq 1}\lambda_{1j} & -\lambda_{12} & -\lambda_{13} & \cdot & \cdot & -\lambda_{1m} \\ -\lambda_{21} & \lambda_2^* + \sum_{j\neq 2}\lambda_{2j} & -\lambda_{23} & \cdot & \cdot & -\lambda_{2m} \\ \cdot & \cdot & & \cdot & \cdot & \cdot \\ -\lambda_{m1} & -\lambda_{m2} & & \cdot & \cdot & \lambda_m^* + \sum_{j\neq m}\lambda_{mj} \end{bmatrix}$$

and define a column vector $\nu = (\nu_1,\dots,\nu_m)'$ by $\mu' = \nu'\Delta$. Then it is shown in [3] that at any time t, the distribution of $N(t)$ is

$$\prod_{r=1}^{m} e^{-\nu_r} \frac{(\nu_r)^{N_r(t)}}{(N_r(t))\,!} . \tag{1.5.1}$$

That is, the distribution of $N(t)$ is the same as that of m independent Poisson random variables with parameters $\nu_1,\dots,\nu_m$.

In [17] Ruben considers the problem of estimating the fundamental interaction parameter θ when the parameters λ_{rs},λ^* and μ_r are known functions of a single parameter θ. The estimate

E_3. The probability that a particle moves from R_r to R^* in time dt is

$$N_r(t)\,\lambda^*_r dt + o(dt),\ r = 1,\dots,m.$$

E_4. The probability that a particle moves from R^* to R_r in time dt is

$$\mu_r dt + o(dt),\ r = 1,\dots,m.$$

E_5. In time dt, the probability of movement of more than one particle is o(dt).

$$\text{Let } \Delta = \begin{bmatrix} \lambda_1^* + \sum_{j\neq 1}\lambda_{1j} & -\lambda_{12} & -\lambda_{13} & \cdot & \cdot & -\lambda_{1m} \\ -\lambda_{21} & \lambda_2^* + \sum_{j\neq 2}\lambda_{2j} & -\lambda_{23} & \cdot & \cdot & -\lambda_{2m} \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ -\lambda_{m1} & -\lambda_{m2} & & \cdot & \cdot & \lambda_m^* + \sum_{j\neq m}\lambda_{mj} \end{bmatrix}$$

and define a column vector $\nu = (\nu_1,\dots,\nu_m)'$ by $\mu' = \nu'\Delta$. Then it is shown in [3] that at any time t, the distribution of N(t) is

$$\prod_{r=1}^{m} e^{-\nu_r}\,\frac{(\nu_r)^{N_r(t)}}{(N_r(t))\,!}\,. \tag{1.5.1}$$

That is, the distribution of N(t) is the same as that of m independent Poisson random variables with parameters $\nu_1,\dots,\nu_m$.

In [17] Ruben considers the problem of estimating the fundamental interaction parameter θ when the parameters λ_{rs}, λ^* and μ_r are known functions of a single parameter θ. The estimate is based on the difference of consecutive observations on the process $N(t)$ taken at equal intervals of time and is called Mean Square Consecutive Fluctuation (M.S.C.F.) estimate.

Let $N(\tau)$, $N(2\tau),\ldots,N(k\tau)$ be k observations on the process $N(t)$ at times τ, $2\tau,\ldots,k\tau$ respectively and let

$$E(N(t) - \nu)\ (N(t) - \nu)' = \Sigma$$

and (1.5.2)

$$E(N(t) - \nu)\ (N(t+\tau) - \nu)' = P(\tau)\ \Sigma.$$

Here $P(t) = e^{-\Delta t} \qquad \text{for } t \geq 0$

$$= G^{-1}K(t)G$$

where K is the diagonal matrix with diagonal elements $e^{-k_1 t},\ldots,e^{-k_m t}$; k_r denoting the real positive eigenvalues of Δ and G is the matrix of row eigenvectors of Δ.

Consider the difference between consecutive observations $\underset{\sim}{\delta}(i)$ defined by:

$$\underset{\sim}{\delta}(i) = N(i\tau) - N((i-1)\tau) \qquad i=2,\ldots,k.$$

Define

$$D_i^2 = m^{-1}\ \underset{\sim}{\delta}'(i)\ Q^{-1}\ \underset{\sim}{\delta}(i)$$

where Q is the dispersion matrix of $\underset{\sim}{\delta}(i)$.

Then

$$E\ D_i^2 = 1; \qquad \text{for all} \quad i = 1,\ldots,k. \tag{1.5.3}$$

The equation (1.5.3) suggests that we use

$$\frac{1}{k-1} \sum_{i=2}^{k} D_i^2 = 1 \tag{1.5.4}$$

as an estimation equation for θ.

The equation (1.5.4) can be rewritten as

$$\frac{1}{m(k-1)} \sum_{i=2}^{k} \sum_{p,q=0}^{m} Q^{pq} \, \delta_p(i) \, \delta_q(i) = 1 \tag{1.5.5}$$

where Q^{pq} is the (p,q)th element of Q^{-1}.

If the quantities ν_r entering in the specification of Q are unknown, then their unbiased estimates are given by

$$\hat{\nu} = \frac{1}{k} \sum_{i=1}^{k} N(i\tau). \tag{1.5.6}$$

The estimation equation (1.5.4) or (1.5.5) is a transcendental equation in θ (the Q^{pq} are functions of θ) and cannot therefore be solved explicitly. However, a numerical solution can be found.

Ruben also derived the large sample variance of the M.S.C.F. estimate and considered three particular cases; namely, (a) known ratio of ν_r, (b) the symmetric model, and (c) the symmetric linear model. Finally Ruben discussed the large sample efficiency in model (c) in the limiting case

$$\nu_j \to \infty, \qquad \text{for} \quad j = 1,\ldots,m.$$

CHAPTER 2

SOME ASPECTS OF THE STOCHASTIC MODEL FOR THE ATTACHMENT OF PHAGES TO BACTERIA

2.0. Introduction: In this chapter we study some of the basic properties of the stochastic model for phage attachment to bacteria, described in section 1.2. In section 2.1 the joint probability generating function (p.g.f.) of $\underset{\sim}{n}(t_1)$ and $\underset{\sim}{n}(t_2)$ $(t_1 < t_2)$ is derived; this is useful in calculating the mixed moments of the process used in subsequent work. In section 2.2 we investigate the limiting distribution of $\underset{\sim}{n}(t)$ under different conditions. In section 2.3 we prove a convergence theorem which is used to obtain the limiting joint distribution of

$$\underset{\sim}{n}(t_1),\ \underset{\sim}{n}(t_2),\ldots,\underset{\sim}{n}(t_k) \quad (t_1 < t_2 < \ldots < t_k)\ .$$

2.1. The Joint p.g.f. and Moments of $\underset{\sim}{n}(t_1)$ and $\underset{\sim}{n}(t_2)$ $(t_1 < t_2)$:

Let $\Psi(t_1,t_2,\underset{\sim}{u}_1,\underset{\sim}{u}_2)$ denote the joint p.g.f. of $\underset{\sim}{n}(t_1)$ and $\underset{\sim}{n}(t_2)$

where $$\underset{\sim}{u}_i = \begin{bmatrix} u_{i0} \\ \cdot \\ \cdot \\ \cdot \\ u_{ir} \end{bmatrix} \quad (i = 1,2)$$

and $\varphi^*(t_2,\underset{\sim}{u}_2 \mid \underset{\sim}{n}(t_1))$ be the conditional p.g.f. of $\underset{\sim}{n}(t_2)$ given $\underset{\sim}{n}(t_1)$.

If $P(n_0,\ldots,n_r;t)$ denotes the probability that there are $n_0,\ldots,n_r$ bacteria with $0,\ldots,r$ phages attached to them respectively at time $t \geq 0$, then we obtain in the usual way by assumption C,

$$\frac{dP}{dt} = -\sum_{i=0}^{r-1} \lambda_i n_i \nu_0 P + \sum_{i=0}^{r-1} \lambda_i (n_i+1)\nu_0 P(n_0,\ldots,n_i+1,n_{i+1}-1,\ldots,n_r;t)$$

and

$$\frac{\partial\varphi^*}{\partial t_2} = \sum_{i=0}^{r-1} \lambda_i \nu_0 \, (u_{2,i+1} - u_{2i}) \frac{\partial\varphi^*}{\partial u_{2i}} \,. \tag{2.1.2}$$

This is a particular case of a p.g.f. for the multivariate Markov process first considered by Bartlett [3] and can be easily solved.

We now proceed to solve (2.1.2). The auxiliary equations are

$$\frac{dt_2}{-1} = \frac{du_{2r}}{0} = \frac{d\varphi^*}{0} = \frac{du_{2i}}{\lambda_i \nu_0 (u_{2,i+1} - u_{2i})} \qquad (i=0,\ldots,r-1) \tag{2.1.3}$$

These can be rewritten as

$$\frac{d}{dt_2}\begin{bmatrix} u_{20} \\ \cdot \\ \cdot \\ u_{2,r-1} \\ u_{2r} \end{bmatrix} = \nu_0 \begin{bmatrix} \lambda_0 & -\lambda_0 & & & & \\ & \lambda_1 & -\lambda_1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & & & \lambda_{r-1} & -\lambda_{r-1} \\ & & & & & \lambda_r \end{bmatrix} \begin{bmatrix} u_{20} \\ \cdot \\ \cdot \\ u_{2,r-1} \\ u_{2r} \end{bmatrix}$$

$$= \nu_0 L \underset{\sim}{u}_2 \tag{2.1.4}$$

The solution of (2.1.4) is of the form

$$e^{-L\rho(t_1,t_2)} \underset{\sim}{u}_2 = c \tag{2.1.5}$$

where c is a constant column vector, L the matrix array of λ_i and

$$\rho(t_1,t_2) = \int_{t_1}^{t_2} \nu_0(\tau)\,d\tau .$$

Thus we have

$$\varphi^{*}(t_2, \underset{\sim}{u}_2 | \underset{\sim}{n}(t_1)) = \varphi_0^{*}\,(e^{-L\rho(t_1,t_2)}\underset{\sim}{u}_2) \tag{2.1.6}$$

where φ_0^{*} is some suitable function of the new variables.

To determine the form of φ_0^{*} we take into account the fact that

$$\varphi^{*}(t_1, \underset{\sim}{u}_2 | \underset{\sim}{n}(t_1)) = \prod_{i=0}^{r} u_{2i}^{\,n_i(t_1)}\,, \tag{2.1.7}$$

then we have

$$\varphi_0^{*}\,(e^{-L\rho(t_1,t_2)}\underset{\sim}{u}_2) = \prod_{i=0}^{r} (e^{-L\rho(t_1,t_2)}\underset{\sim}{u}_2)_i^{\,n_i(t_1)} \tag{2.1.8}$$

where $(e^{-L\rho(t_1,t_2)}\underset{\sim}{u}_2)_i$ is the ith element of the column vector $e^{-L\rho(t_1,t_2)}\underset{\sim}{u}_2$. We now proceed to calculate it. Since by assumption λ_i's are distinct and non-negative, the matrix L can be written in the form

$$L = N\Lambda M = N \begin{bmatrix} \lambda_0 & & & & \\ & \lambda_1 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \lambda_r \end{bmatrix} M \tag{2.1.9}$$

where $N = M^{-1}$ and M is the matrix whose rows are eigenvectors of L corresponding to the eigenvalues $\lambda_0, \ldots, \lambda_r$ respectively. It is easy to see that

$$N_{ii} = 1 \qquad N_{ij} = \prod_{s=i}^{j-1} \frac{\lambda_s}{\lambda_s - \lambda_j} \qquad (r \geq j > i)$$

$$M_{ii} = 1 \qquad M_{ij} = \prod_{s=i+1}^{j} \frac{\lambda_{s-1}}{\lambda_s - \lambda_i} \qquad (r \geq j > i). \tag{2.1.10}$$

In particular, $N_{ir} = 1$ for all $i = 0,\ldots,r$ and $M_{ir} = -\frac{\lambda_{r-1}}{\lambda_i} M_{ir-1}$ for $i = 0,\ldots,r-1$. It follows from (2.1.9) that

$$e^{-L\rho(t_1,t_2)} \underset{\sim}{u}_2 = N\, e^{-\Lambda\rho(t_1,t_2)} M \underset{\sim}{u}_2 . \tag{2.1.11}$$

Let $e^{-L\rho(t_1,t_2)} \doteq N\, e^{-\Lambda\rho(t_1,t_2)} M = ||a_{ij}(t_1,t_2)||$ (2.1.12)

where $a_{ij}(t_1,t_2) = 0$ for $i > j$ and $\sum_{j=0}^{r} a_{ij}(t_1,t_2) = 1$ for all $i = 0,\ldots,r$. Hence we see from (2.1.7) and (2.1.11), that

$$\varphi^*(t_2;\underset{\sim}{u}_2 \mid \underset{\sim}{n}(t_1)) = \prod_{i=0}^{r} \left(\sum_{j=0}^{r} a_{ij}(t_1,t_2)u_{2j} \right)^{n_i(t_1)}, \tag{2.1.13}$$

which gives

$$\Psi(t_1,t_2 \,;\, \underset{\sim}{u}_1,\underset{\sim}{u}_2) = \left[\sum_{i=0}^{r} a_{0i}(t_1) \left(\sum_{j=0}^{r} a_{ij}(t_1,t_2)u_{2j}\right) u_{1i} \right]^{n_{00}} . \tag{2.1.14}$$

It is clear from (2.1.13) that the conditional distribution of $\underset{\sim}{n}(t_2)$ given $\underset{\sim}{n}(t_1)$ is the same as that of the sum of $(r+1)$ independent random variables such that the jth random variable has a multinomial distribution with parameter $n_j(t_1)$ and probabilities $a_{j0}(t_1,t_2),\ldots,a_{jr}(t_1,t_2)$.

Next we proceed to calculate the second order mixed moments which will be needed in subsequent work.

Evaluation of elements of $R(t_1,t_2) = E\, \underset{\sim}{n}(t_1)\, \underset{\sim}{n}'(t_2)$

$$= ||\, En_\alpha(t_1) n_\beta(t_2) ||.$$

Differentiating $\Psi(t_1,t_2,\underset{\sim}{u}_1,\underset{\sim}{u}_2)$ with respect to $u_{1\alpha}$ and $u_{2\beta}$ and putting $\underset{\sim}{u}_1 = \underset{\sim}{u}_2 = \underset{\sim}{1}$ where $\underset{\sim}{1}$ is the column vector, we get

$$E\, n_\alpha(t_1)\, n_\beta(t_2) = n_{00}(n_{00}-1)\, a_{0\alpha}(t_1)\Big[\sum_{i=\beta}^{r} a_{0i}(t_1) a_{i\beta}(t_1,t_2)\Big] \text{if } \beta < \alpha$$

$$= n_{00}\, a_{0\alpha}(t_1)\, a_{\alpha\beta}(t_1,t_2)$$

$$+ n_{00}(n_{00}-1)\, a_{0\alpha}(t_1)\, \Big[\sum_{i=\beta}^{r} a_{0i}(t_1) a_{i\beta}(t_1,t_2)\Big]$$

$$\text{if } \beta \geq \alpha \qquad (2.1.15)$$

Since $N\, e^{-\Lambda\rho(t_2)} M = N\, e^{-\Lambda\rho(t_1)\, -\Lambda\rho(t_1,t_2)} M$

$$= N\, e^{-\ \rho(t_1)} M\, N\, e^{-\ \rho(t_1,t_2)} M\ ,$$

that is,

$$||a_{ij}(t_2)|| = ||a_{ij}(t_1)||\ \ ||a_{ij}(t_1,t_2)||\ ,$$

then we have

$$a_{0\lambda}(t_2) = \sum_{j=0}^{r} a_{0j}(t_1)\, a_{j\lambda}(t_1,t_2) \qquad \text{for } \lambda = 0,\ldots,r\ .$$

Therefore (2.1.15) can be written as:

$$E\, n_\alpha(t_1)\, n_\beta(t_2) = n_{00}(n_{00}-1)\, a_{0\alpha}(t_1)\, a_{0\beta}(t_2) \quad \text{if } \beta < \alpha.$$

$$= n_{00}\, a_{0\alpha}(t_1)\, a_{\alpha\beta}\, (t_1,t_2)$$

$$+ n_{00}(n_{00}-1)\, a_{0\alpha}(t_1)\, a_{0\beta}(t_2) \quad \text{if } \beta \geq \alpha. \ (2.1.16)$$

This gives

$$\text{Cov } n_\alpha(t_1)n_\beta(t_2) = - n_{00}\, a_{0\alpha}(t_1)\, a_{0\beta}(t_2) \qquad \text{if } \beta < \alpha.$$

$$= n_{00}\, a_{0\alpha}(t_1)\, \{a_{\alpha\beta}(t_1,t_2) - a_{0\beta}(t_2)\}$$

$$\text{if } \beta \geq \alpha. \qquad (2.1.17)$$

<u>2.2. General Discussion of the Model</u>: We have shown that for fixed t, $0 \leq t \leq t_0$, $\underset{\sim}{n}(t)$ has a multinomial distribution with parameter n_{00} and probabilities $a_{00}(t),\ldots,a_{0r}(t)$ where

$$a_{0j}(t) = \sum_{i=0}^{j} (-1)^{j-i} \binom{r}{i} \binom{r-i}{j-i} e^{-(r-i)\alpha\rho(t)} \quad .$$

These probabilities are functions of n_{00},m, and t; that is, they depend on the initial concentration of bacteria, the multiplicity of phages and the duration of the experiment. In this section we confirm mathematically certain experimental facts. For example, we consider the limiting behavior of the probability distribution of $\underset{\sim}{n}(t)$ when m, the multiplicity of phages, is large, and prove that in a short time all bacteria are saturated with the maximum of r phages, as is known experimentally. The following cases are considered:

(a) Let $m \to \infty$, keeping t fixed. Then $\mu = n_{00}(r-m) \to -\infty$ as $m \to \infty$.

Since $\qquad \rho(t) = \frac{1}{\alpha} \log \left(\frac{r-m\, e^{-\mu\alpha t}}{r-m} \right) \quad ,$

$$a_{0j}(t) = \sum_{i=0}^{j} (-1)^{j-i} \binom{r}{i} \binom{r-i}{j-i} \left(\frac{r-m}{r-m\,\exp(-\mu\alpha t)} \right)^{r-i}$$

$$\to 0 \qquad\qquad \text{for } j = 0,\ldots,r-1$$

and

$$a_{0r}(t) \to 1$$

as m tends to infinity, that is, if the multiplicity of phages is large, then in a short time all bacteria are saturated by the maximum of r phages.

(b) If m is small and fixed, then m $\exp(-\mu\alpha t)$ is very small and

$$\rho(t) \simeq \frac{1}{\alpha} \log \frac{r}{r-m}$$

(the sign $\simeq$ means approximate equality). In this case we get a fixed distribution after some time. In fact in a short time all the phages attach themselves to bacteria and no more phages are left. Therefore further observations do not give us any more information.

(c) Thus we see that if m is very large, then in a short time all the bacteria are saturated by the maximum of r phages, and if m is small, then in a short time all phages attach themselves to bacteria and no more phages are left. To avoid both these extreme cases, we choose m less than r in such a way that the product $n_{00}(r-m)$ is constant; this proves to be an interesting case. We shall assume throughout that m (the multiplicity of phages) is such that

$$n_{00}(r-m) = \mu_0 > 0 \tag{2.2.1}$$

is constant.

If (2.2.1) is satisfied, it is clear from (1.2.12) that the distribution of $\underset{\sim}{n}(t)$ tends to a multivariate normal distribution.

More precisely, if

$$\xi_j(t) = \frac{n_j(t) - n_{00}a_{0j}(t)}{(n_{00}a_{0j}(t)(1-a_{0j}(t)))^{\frac{1}{2}}}$$

then as n_{00} tends to infinity, the joint distribution of $\xi_0(t),\ldots,\xi_r(t)$ tends to a multivariate normal distribution with means zero and covariance matrix

$$W_{11} = || \Sigma_{ij} || \quad \text{where}$$

$$\Sigma_{ii} = 1$$

and

$$\Sigma_{ij} = -\left(\frac{a_{0i}(t)a_{0j}(t)}{(1-a_{0i}(t))(1-a_{0j}(t))}\right)^{\frac{1}{2}}, \quad i \neq j.$$

2.3. A Useful Convergence Theorem: In this section we prove a convergence theorem which is used in the next section to derive the limiting joint distribution of $\underset{\sim}{n}(t_1),\ldots,\underset{\sim}{n}(t_k), t_1 < ,\ldots, < t_k$.

For the formulation of our theorem, we require the notion of UC^* convergence. The notion of the UC^* convergence has been introduced by Parzen in [13]; we use a slightly different definition of it, already given by Sethuraman [18]. Let $\nu_n(\theta,\cdot)$, $n = 0,1,\ldots$ be a family of sequences of probability measures on R_k, the Euclidean space of k dimensions. Assume that θ takes values in a compact metric space I. Let $\Psi_n(\underset{\sim}{u},\theta)$ denote the characteristic function (c.f.) of $\nu_n(\theta,\cdot)$, that is,

$$\Psi_n(\underset{\sim}{u},\theta) = \int_{R_k} \exp(i\,\underset{\sim}{u}'\underset{\sim}{m})\nu_n(\theta,dm) \qquad n = 0,1,\ldots \quad (2.3.1)$$

Definition (Sethuraman): The family of sequences $\nu_n(\theta,\cdot)$ is said to converge in the UC^* sense to $\nu_0(\theta,\cdot)$ (denoted by $\underset{\rightarrow}{UC}^*$) relative to $\theta \in I$ as n tends to infinity if

(a) $\sup_{\theta\in I} \left| \Psi_n(\underset{\sim}{u},\theta) - \Psi_0(\underset{\sim}{u},\theta) \right| \to 0$ as $n_{00} \to \infty$,

(b) $\Psi_0(\underset{\sim}{u},\theta)$ is equicontinuous in θ at $\underset{\sim}{u} = 0$;

and

(c) $\Psi_0(\underset{\sim}{u},\theta)$ is a continuous function of θ for each $\underset{\sim}{u}$.

We shall also require the notion of weak convergence of distribution functions (d.f.'s). Let $F_n(x)(n = 0,1,\ldots)$ be a sequence of d.f.'s. We say that $F_n(x)$ converges weakly (denoted by $F_n(x) \underset{\rightarrow}{w} F_0(x)$) to $F_0(x)$ if $\int f(x)dF_n(x)$ tends to $\int f(x)dF_0(x)$ as n tends to infinity for all bounded continuous functions.

Now let $H_n(x_1,\ldots,x_k)$ $(n = 0,1,\ldots)$ be a sequence of distribution functions (d.f.'s) in $m_1 + \ldots m_k = m$ dimensional Euclidean space R_m where $x_i \in R_{m_i}$. Also let $H_n^1(x_1)$ $(n=0,1,\ldots)$ be the corresponding sequence of the marginal d.f.'s and $H_n^i(x_i|x_1\ldots x_{i-1})$ $(n = 0,1,\ldots)$ be the sequences of the conditional d.f.'s $(i = 2,\ldots,k)$. Then we have the following theorem

Theorem 2.3.1: If (a) $H_n^1(x_1) \underset{\rightarrow}{w} H_0^1(x_1)$, and

(b) $H_n^i(x_i|x_1,\ldots,x_{i-1}) \underset{\rightarrow}{UC}^* H_0^i(x_i|x_1,\ldots,x_{i-1})$

relative to $(x_1,\ldots,x_{i-1}) \in I$ where I is any compact subset of $R_{m_1 + \ldots + m_{i-1}}$, and $i = 2,\ldots,k$, then

$$H_n(x_1,\ldots,x_k) \xrightarrow{w} H_0(x_1,\ldots,x_k)$$
$$= \int_0^{x_1} H_0^1(dx_1) \cdots\cdots \int_0^{x_k} H_0^k(dx_k|x_1,\ldots,x_{k-1}).$$

The proof of this theorem depends on the following result due to Sethuraman [18]. Let $F_n(x_1,x_2)$ $(n = 0,1,\ldots)$ be a sequence of d.f.'s. Also let $F_n^1(x_1)$ and $F_n^2(x_2|x_1)$ $(n=0,1,\ldots)$ be the corresponding sequences of the marginal and the conditional d.f.'s, respectively.

Theorem (Sethuraman): If (a) $F_n^1(x_1) \xrightarrow{w} F_0^1(x_1)$, and

(b) $F_n^2(x_2|x_1) \xrightarrow{UC^*} F_0^2(x_2|x_1)$

relative to $x_1 \in I$ where I is any compact subset of R_p and $x_2 \epsilon R_q$, then

$$F_n(x_1,x_2) \xrightarrow{w} F_0(x_1,x_2)$$
$$= \int_0^{x_1} F_0(dx_1) \int_0^{x_2} F_0^2(dx_2|x_1).$$

Now we prove the theorem 2.3.1.

Proof of theorem 2.3.1: We prove this theorem by induction. The result is true for k = 2 by Sethuraman's theorem. Suppose it is true for k, then again by Sethuraman's theorem it is true for k + 1.

2.4. <u>The Limiting Joint Distribution of $\underset{\sim}{n}(t_1),\ldots,\underset{\sim}{n}(t_k)$</u>: Let $\underset{\sim}{n}(t_1),\ldots,\underset{\sim}{n}(t_k)$ be k observations on the process $\underset{\sim}{n}(t)$ and let

$$\xi_i(t_j) = \frac{n_i(t_j) - n_{00}a_{0i}(t_j)}{\{n_{00}a_{0i}(t_j)(1-a_{0i}(t_j))\}^{\frac{1}{2}}}$$

and $$\underset{\sim}{\xi}(t_j) = \begin{bmatrix} \xi_0(t_j) \\ \xi_1(t_j) \\ \cdot \\ \cdot \\ \xi_r(t_j) \end{bmatrix} \qquad \text{and} \qquad \underset{\sim}{\xi} = \begin{bmatrix} \xi_0(t_1) \\ \cdot \\ \cdot \\ \xi_r(t_1) \\ \cdot \\ \cdot \\ \xi_0(t_k) \\ \cdot \\ \cdot \\ \xi_r(t_k) \end{bmatrix}$$

Also let $W = E(\underset{\sim}{\xi}\underset{\sim}{\xi}')$, that is, W is a square matrix of order k(r+1). It is the covariance matrix of the random vector $\underset{\sim}{\xi}$. Further $\sum_{i=0}^{r} \xi_i(t_j) = 0$ for $j = 1,\ldots,k$; that is, the random vector $\underset{\sim}{\xi}$ takes values in rk-dimensional subspace and W is a singular matrix of rank rk. The matrix W consists of k^2 submatrices W_{ij} $(i,j = 1,\ldots,k)$ each of order (r+1) where

$$W_{jj} = E(\underset{\sim}{\xi}(t_j)\underset{\sim}{\xi}'(t_j))$$

and

$$W_{ij} = E(\underset{\sim}{\xi}(t_i)\underset{\sim}{\xi}'(t_j)) \qquad (i \neq j = 1,\ldots,k)$$

Before deriving the limiting joint distribution of $\underset{\sim}{n}(t_1),\ldots,\underset{\sim}{n}(t_k)$, we prove several lemmas which are used for proving that the conditional distribution of $\underset{\sim}{\xi}(t_2)$ given $\underset{\sim}{\xi}(t_1) = \underset{\sim}{x}(t_1)$ converges in the UC* sense relative to any compact set I to a multivariate normal distribution ($t_1 < t_2$).

Lemma 2.4.1: The c.f. of $\underset{\sim}{\xi}(t_2)$ given $\underset{\sim}{\xi}(t_1) = \underset{\sim}{x}(t_1)$ is given by

$$\Phi_{n_{00}}(\underset{\sim}{u}_2;t_2 \mid \underset{\sim}{x}(t_1))$$

$$= \exp[-i\, n_{00}^{\frac{1}{2}} \sum_{j=0}^{r} \left(\frac{a_{0j}(t_2)}{1-a_{0j}(t_2)}\right)^{\frac{1}{2}} u_{2j}$$

$$+ \sum_{i=0}^{r} \{n_{00}a_{0i}(t_1) + x_i(t_1)\, n_{00}^{\frac{1}{2}}\, a_{0i}^{\frac{1}{2}}(t_1)\, (1-a_{0i}(t_1))^{\frac{1}{2}}\} \times$$

$$\log\{\sum_{j=0}^{r} a_{ij}(t_1,t_2) \exp(\frac{i\, u_{2j}}{(n_{00}a_{0j}(t_2)(1-a_{0j}(t_2)))^{\frac{1}{2}}})\}]\cdot \quad (2.4.1)$$

Proof: As remarked in section 2.1, it follows from (2.1.13) that the conditional distribution of $\underset{\sim}{n}(t_2)$ given $\underset{\sim}{n}(t_1)$ is the same as that of the sum of (r+1) independent random variables such that the jth random variable has a multinomial distribution with parameter $n_j(t_1)$ and probabilities $a_{j0}(t_1,t_2),\ldots,a_{jr}(t_1,t_2)$.
So we have

$$\Phi_{n_{00}}(\underset{\sim}{u}\,;t_2|\underset{\sim}{x}(t_1))$$

$$= E(\exp[i \sum_{j=0}^{r} \frac{n_j(t_2)-n_{00}a_{0j}(t_2)}{(n_{00}a_{0j}(t_2)(1-a_{0j}(t_2)))^{\frac{1}{2}}}] \mid \underset{\sim}{x}(t_1))$$

$$= \exp[-i\, n_{00}^{\frac{1}{2}} \sum_{j=0}^{r} (\frac{a_{0j}(t_2)}{1-a_{0j}(t_2)})^{\frac{1}{2}} u_{2j}] \times$$

$$\prod_{i=0}^{r} [\sum_{j=0}^{r} a_{ij}(t_1,t_2)\exp(\frac{i\, u_{2j}}{(n_{00}a_{0j}(t_2)(1-a_{0j}(t_2)))^{\frac{1}{2}}})]^{n_i(t_1)}$$

where $n_i(t_1) = n_{00}a_{0i}(t_1)+x_i(t_1)(n_{00}a_{0i}(t_1)(1-a_{0i}(t_1)))^{\frac{1}{2}}$.

$$= \exp[-i\, n_{00}^{\frac{1}{2}} \sum_{j=0}^{r} (\frac{a_{0j}(t_2)}{1-a_{0j}(t_2)})^{\frac{1}{2}} u_{2j}$$

$$+ \sum_{i=0}^{r} \{n_{00}a_{0i}(t_1)+x_i(t_1)n_{00}^{\frac{1}{2}}\, a_{0i}^{\frac{1}{2}}(t_1)(1-a_{0i}(t_1))^{\frac{1}{2}}\} \times$$

$$\log\{ \sum_{j=0}^{r} a_{ij}(t_1,t_2) \exp(\frac{i\, u_{2j}}{(n_{00}\, a_{0j}(t_2)(1-a_{0j}(t_2)))^{\frac{1}{2}}})\}].$$

The equation (2.4.1) may be compared with (2.1.13) which gives the conditional p.g.f. of $\underset{\sim}{n}(t_2)$ given $\underset{\sim}{n}(t_1)$.

<u>Lemma 2.4.2</u>: Let C be any compact subset of R_{r+1}, the Euclidean space of (r+1) dimensions. Then for any fixed $\underset{\sim}{u}_2$,

$$\lim_{n_{00}\to\infty} \Phi_{n_{00}}(\underset{\sim}{u}_2;t_2|\underset{\sim}{x}(t_1)) = \Phi_0(\underset{\sim}{u}_2;t_2|\underset{\sim}{x}(t_1))$$

unifromly in $\underset{\sim}{x}(t_1)$ on C.

Here

$$\Phi_0(\underset{\sim}{u}_2;t_2|\underset{\sim}{x}(t_1))$$

$$= \exp[i\, \{ \sum_{i=0}^{r} x_i(t_1)\, a_{0i}(t_1)^{\frac{1}{2}}\, (1-a_{0i}(t_1))^{\frac{1}{2}} \sum_{j=0}^{r} \frac{a_{ij}(t_1,t_2)u_{2j}}{(a_{0j}(t_2)(1-a_{0j}(t_2))^{\frac{1}{2}}} \}$$

$$+ \sum_{i=0}^{r} a_{0i}(t_1) \{ \sum_{j \neq j'} \frac{a_{ij}(t_1,t_2)\, a_{ij'}(t_1,t_2) u_{2j} u_{2j'}}{(a_{0j}(t_2) a_{0j'}(t_2)(1-a_{0j}(t_2))(1-a_{0j'}(t_2)))^{\frac{1}{2}}}$$

$$- \frac{1}{2} \sum_{j} \frac{a_{ij}(t_1,t_2)(1-a_{ij}(t_1,t_2)) u_{2j}^2}{a_{0j}(t_2)(1-a_{0j}(t_2))} \}]$$

Proof: It is sufficient to prove the theorem for any bounded closed rectangle $C = \{\underset{\sim}{x}(t_1) \mid x_i' \leq x_i(t_1) \leq x_i'' \; ; \; i = 0,\dots,r\}$ because any compact set can be enclosed by a bounded closed rectangle and uniform convergence on a set implies uniform convergence on all subsets of the set.

From (2.4.1), we have

$$\Phi_{n_{00}}(\underset{\sim}{u}_2; t_2 | \underset{\sim}{x}(t_1))$$

$$= \exp[-i\, n_{00}^{\frac{1}{2}} \sum_{j=0}^{r} (\frac{a_{0j}(t_2)}{1-a_{0j}(t_2)})^{\frac{1}{2}} u_{2j}$$

$$+ \sum_{i=0}^{r} \{ n_{00}\, a_{0i}(t_1) + x_i(t_1)(n_{00} a_{0i}(t_1)(1-a_{0i}(t_1)))^{\frac{1}{2}} \} \times$$

$$\log(\sum_{j=0}^{r} a_{ij}(t_1,t_2) \exp(\frac{i\, u_{2j}}{(n_{00} a_{0j}(t_2)(1-a_{0j}(t_2)))^{\frac{1}{2}}})]$$

$$= \exp[-i\, n_{00}^{\frac{1}{2}} \sum_{j=0}^{r} (\frac{a_{0j}(t_2)}{1-a_{0j}(t_2)})^{\frac{1}{2}} u_{2j}$$

$$+ \sum_{i=0}^{r} \{n_{00}a_{0i}(t_1) + x_i(t_1)(n_{00}a_{0i}(t_1)(1-a_{0i}(t_1)))^{\frac{1}{2}}\} \times$$

$$\log\{1 + i \sum_{j=0}^{r} \frac{a_{ij}(t_1,t_2)u_{2j}}{(n_{00}a_{0j}(t_2)(1-a_{oj}(t_2))^{\frac{1}{2}}}$$

$$- \frac{1}{2} \sum_{j=0}^{r} \frac{a_{ij}^2(t_1,t_2)u_{2j}^2}{n_{00}a_{0j}(t_2)(1-a_{0j}(t_2))} + \ldots \}] .$$

$$= \exp[i\{ \sum_{i} x_i(t_1)\ a_{0i}^{\frac{1}{2}}(t_1)(1-a_{0i}(t_1))^{\frac{1}{2}} \sum_{j=0}^{r} \frac{a_{ij}(t_1,t_2)u_{2j}}{(a_{0j}(t_2)(1-a_{0j}(t_2)))^{\frac{1}{2}}}\}$$

$$+ \sum_{i=0}^{r} (a_{0i}(t_1) + n_{00}^{-\frac{1}{2}}\ x_i(t_1)\ a_{0i}^{\frac{1}{2}}(t_1)(1-a_{0i}(t_1)^{\frac{1}{2}}) \times$$

$$\{ \sum_{j\neq j'} \frac{a_{ij}(t_1,t_2)a_{ij'}(t_1,t_2)u_{2j}u_{2j'}}{(a_{0j}(t_2)a_{0j'}(t_2)(1-a_{0j}(t_2))(1-a_{0j'}(t_2)))^{\frac{1}{2}}}$$

$$- \frac{1}{2} \sum_{j} \frac{a_{ij}(t_1,t_2)(1-a_{ij}(t_1,t_2)}{a_{0j}(t_2)(1-a_{0j}(t_2))} u_{2j}^2 + \ldots \} ,$$

since $\quad i\, n_{00}^{\frac{1}{2}} \sum_{j=0}^{r} \left(\frac{a_{0j}(t_2)}{1-a_{0j}(t_2)} \right)^{\frac{1}{2}} u_{2j}$

$$= i\, n_{00}^{\frac{1}{2}} \sum_{j=0}^{r} \sum_{i=0}^{r} \frac{a_{0i}(t_1)a_{ij}(t_1,t_2)\, u_{2j}}{(a_{0j}(t_2)(1-a_{0j}(t_2)))^{\frac{1}{2}}}$$

$$= i\, n_{00}^{\frac{1}{2}} \sum_{j=0}^{r} a_{0i}(t_1) \sum_{j=0}^{r} \frac{a_{ij}(t_1,t_2)u_{2j}}{(a_{0j}(t_2)(1-a_{0j}(t_2)))^{\frac{1}{2}}} \quad .$$

But $[n_{00}^{-\frac{1}{2}}\, x_i(t_1)\, a_{0i}^{\frac{1}{2}}(t_1)(1-a_{0i}(t_1))^{\frac{1}{2}}\,] \times$

$$\sum_{j\neq j'} \frac{a_{ij}(t_1,t_2)a_{ij'}(t_1,t_2)u_{2j}u_{2j'}}{(a_{0j}(t_2)a_{0j'}(t_2)(1-a_{0j}(t_2))(1-a_{0j'}(t_2))\,)^{\frac{1}{2}}}$$

$$- \frac{1}{2} \sum_{j} \frac{a_{ij}(t_1,t_2)(1-a_{ij}(t_1,t_2)u_{2j}^{\;2}}{a_{0j}(t_2)(1-a_{0j}(t_2))} \}\,]$$

tends to zero uniformily in $x_i(t_1)$ for $x_i' \le x_i(t_1) \le x_i''$ as n_{00} tends to infinity for $i = 0,\ldots r$, therefore

$\Phi_{n_{00}}(\underset{\sim}{u}_2;t_2|\underset{\sim}{x}(t_1))$ coverges to

$$\exp[i\{ \sum_{i} x_i(t_1)a_{0i}^{\frac{1}{2}}(t_1)(1-a_{0i}(t_1))^{\frac{1}{2}} \sum_{j=0}^{r} \frac{a_{ij}(t_1,t_2)u_{2j}}{(a_{0j}(t_2)(1-a_{0j}(t_2)))^{\frac{1}{2}}} \}$$

$$+ \sum_{i=0}^{r} a_{0i}(t_1)\, \{ \sum_{j\neq j'} \frac{a_{ij}(t_1,t_2)a_{ij'}(t_1,t_2)u_{2j}u_{2j'}}{(a_{0j}(t_2)a_{0j'}(t_2)(1-a_{0j}(t_2))(1-a_{0j'}(t_2)))^{\frac{1}{2}}}$$

$$- \frac{1}{2} \sum_{j} \frac{a_{ij}(t_1,t_2)(1-a_{ij}(t_1,t_2)u_{2j}^{2}}{a_{0j}(t_2)(1-a_{0j}(t_2))} \}\,]$$

uniformly on C as n_{00} tends to infinity.

Lemma 2.4.3: $\Phi_0(\underset{\sim}{u}_2;t_2|\underset{\sim}{x}(t_1))$ is a continuous function of $\underset{\sim}{x}(t_1)$ for any fixed $\underset{\sim}{u}_2$.

The proof of this lemma is immediate.

Lemma 2.4.4: If $f(x,y)$ is a continuous function on $[a,b] \times [c,d]$, then the family of functions $f(x,\cdot)$, for $x \in [a,b]$, is equicontinuous in x at $y = y_0 \in [c,d]$.

Proof: First we note that the function $f(x,y)$ is uniformly continuous since $[a,b] \times [c,d]$ is a compact set. The family of functions $f(x,\cdot)$, $x \in [a,b]$, is equicontinuous in x at $y = y_0$ if given $\epsilon > 0$, there exists $\delta(\epsilon)$ (depending only on ϵ) such that $|f(x,y) - f(x,y_0)| < \epsilon$ if $|y - y_0| < \delta(\epsilon)$. But such a $\delta(\epsilon)$ exists since $f(x,y)$ is a uniformly continuous function. Hence the family $f(x,\cdot)$ is equicontinuous at $y = y_0$.

Lemma 2.4.5: $\Phi_0(\underset{\sim}{u}_2;t_2|\underset{\sim}{x}(t_1))$ is equicontinuous in $\underset{\sim}{x}(t_1) \in I$ at $\underset{\sim}{u}_2 = 0$ where I is any compact subset of R_{r+1}.

Proof: It is sufficient to prove the theorem for closed rectangles $I = \{\underset{\sim}{x}(t_1)|x_i'(t_1) \le x_i(t_1) \le x_i''(t_1), i = 0,\dots,r\}$. The function $\Phi_0(\underset{\sim}{u}_2;t_2|\underset{\sim}{x}(t_1))$ is continuous in $(\underset{\sim}{x}(t_1),\underset{\sim}{u}_2)$ on $I \times J$ where J is any closed subset of R_{r+1} containing the point $\underset{\sim}{u}_2 = 0$. Hence by lemma 2.4.4, the function $\Phi_0(\underset{\sim}{u}_2;t_2|\underset{\sim}{x}(t_1))$ is equicontinuous in $\underset{\sim}{x}(t_1)$ at $\underset{\sim}{u}_2 = 0$.

Theorem 2.4.1: The conditional distribution of $\underset{\sim}{\xi}(t_2)$ given $\underset{\sim}{\xi}(t_1) = \underset{\sim}{x}(t_1)$ converges in the UC* sense relative to any compact subset I of R_{r+1} to a multivariate normal d.f.

Proof: Since $\Phi_0(\underset{\sim}{u}_2;t_2|\underset{\sim}{x}(t_1))$ is a c.f. of a multivariate normal d.f., we conclude from the lemmas 2.4.2, 2.4.3, and 2.4.5 that the conditional d.f. of $\underset{\sim}{\xi}(t_2)$ given $\underset{\sim}{\xi}(t_1) = \underset{\sim}{x}(t_1)$ converges in the UC^* sense relative to any compact subset I of R_{r+1} to a multivariate normal d.f.

Lemma 2.4.6: If F(x,y) is a d.f. such that the marginal d.f. G(x) has the c.f. $\exp[-\frac{1}{2}\sum_{i,j}\alpha_{ij}u_{1i}u_{1j}]$ and the conditional d.f. of Y given X = x has the c.f. $\exp[i\sum_j x_j u_{2j} - \frac{1}{2}\sum_{i,j}\beta_{ij}u_{2i}u_{2j}$, then the joint d.f. F(x,y) is multivariate normal d.f.

Proof: The c.f. of F(x,y) is

$$\exp[-\frac{1}{2}\sum_{i,j}\beta_{ij}u_{2i}u_{2j}].\quad E(\exp[i\sum_j x_j(u_{1j}+u_{2j})]\,).$$

$$= \exp[-\frac{1}{2}\sum_{i,j}\alpha_{ij}(u_{1i}+u_{2i})(u_{1j}+u_{2j}) - \frac{1}{2}\sum_{i,j}\beta_{ij}u_{2i}u_{2j}].$$

But this is the c.f. of a multivariate normal d.f.

Corollary: If $F(x_1,\ldots,x_k)$ is a d.f. such that (i) the marginal d.f. of X_1 has the c.f. $\exp[-\frac{1}{2}\sum_{i,j}\alpha_{1ij}u_{1i}u_{1j}]$ and (ii) the conditional d.f. of X_i given $X_1 = x_1,\ldots,X_{i-1} = x_{i-1}$ has the c.f.

$$\exp[i\sum_j x_{i-1,j}u_{i,j} - \frac{1}{2}\sum_{j,j'}\alpha_{i,j,j'},u_{ij}u_{ij'}]$$

for $i = 2,\ldots,k$, then the joint d.f. $F(x_1,\ldots,x_k)$ is a multivariate normal d.f.

Proof: We prove this result by induction. This is true for k = 2 by the above lemma. Suppose it is true for k, then again by the above lemma, it is true for k+1.

Theorem 2.4.2: The joint d.f. of $\underset{\sim}{\xi}(t_1)$ and $\underset{\sim}{\xi}(t_2)$ converges weakly to the multivariate normal d.f. $N(0,\left\|\begin{matrix} W_{11} & W_{12} \\ W_{21} & W_{22}\end{matrix}\right\|)$ as n_{00} tends to infinity.

Proof: It follows from theorem 2.3.1, that the joint d.f. of $\underset{\sim}{\xi}(t_1)$ and $\underset{\sim}{\xi}(t_2)$ converges weakly to a joint d.f. as n_{00} tends to infinity. The limiting d.f. is a multivariate normal d.f. $N(0,\left\|\begin{matrix} W_{11} & W_{12} \\ W_{21} & W_{22}\end{matrix}\right\|)$ by lemma 2.4.6 since the marginal d.f. of $\underset{\sim}{\xi}(t_1)$ converges weakly to a multivariate normal d.f. $N(0,W_{11})$ and by theorem 2.4.1 the conditional d.f. of $\underset{\sim}{\xi}(t_2)$ given $\underset{\sim}{\xi}(t_1) = \underset{\sim}{x}(t_1)$ converges in the UC^* sence relative to any compact subset I of R_{r+1} to a multivariate normal d.f. whose c.f. is $\Phi_0(\underset{\sim}{u}_2;t_2|\underset{\sim}{x}(t_1))$.

Now we state and prove the main theorem of this section.

Theorem 2.4.3: The joint d.f. of $\underset{\sim}{\xi}(t_j)(j = 1,\ldots,k)$ converges weakly to a multivariate normal d.f. $N(0,W)$ as n_{00} tends to infinity.

Proof: (i) It is clear from (1.2.12) that the marginal d.f. of $\underset{\sim}{\xi}(t_1)$ converges weakly to a multivariate normal d.f. $N(0,W_{11})$.

(ii) From theorem 2.4.1, we know that the conditional d.f. of $\underset{\sim}{\xi}(t_i)$ given $\underset{\sim}{\xi}(t_{i-1}) = \underset{\sim}{x}(t_{i-1})$ converges in the UC^* sense relative to any compact subset to a multivariate normal d.f. whose c.f. is $\Phi_0(\underset{\sim}{u}_i;t_i|\underset{\sim}{x}(t_{i-1}))$ as n_{00} tends to infinity. This is true for $i = 2,\ldots,k$.

(iii) Fr m the corollary to lemma 2.4.6 and theorem 2.3.1, we conclude that the joint d.f. of $\underset{\sim}{\xi}(t_1),\ldots,\underset{\sim}{\xi}(t_k)$ converges weakly to a multivariate normal d.f. $N(0,W)$ as n_{00} tends to infinity.

CHAPTER 3

ESTIMATION OF THE PARAMETER IN THE STOCHASTIC MODEL FOR PHAGE ATTACHMENT TO BACTERIA.

3.0. Summary: In this chapter we consider the problem of estimating the parameter α in the stochastic model for the attachment of phages to bacteria described in section 1.3. A simple method, of the type originated by Ruben [17], for estimating the parameter α in this model is described in section 3.2. The estimate is based on k observations $\underset{\sim}{n}(t_1)$, $\ldots,\underset{\sim}{n}(t_k)$ at times $t_j = j\tau$ ($\tau > o, j=1,\ldots k$) and is shown to be consistent and asymptotically normally distributed. In section 3.6 we study the efficiency of the estimate.

3.1. Introduction: The stochastic model for phage attachment to bacteria gives rise to a multivariate stochastic process $\underset{\sim}{n}(t)$, depending on a single unknown parameter α. This process $\underset{\sim}{n}(t)$ is Markovian and its transition probabilities are functions of α. For a Markov process in general it does not seem unrealistic to expect that relatively efficient estimates for the transition probabilities, hence for the parameter α, may sometimes be obtained from the consecutive differences of the relative frequencies observed at discrete points in time. Following Ruben we call such an estimate the Mean Square Consecutive Fluctuation(M.S.C.F.) estimate.

Consider the differences

$$\begin{bmatrix} d_0(i) \\ \cdot \\ \cdot \\ d_r(i) \end{bmatrix} = \frac{1}{n_{00}} \begin{bmatrix} n_0(t_i) - n_0(t_{i-1}) \\ \cdot \\ \cdot \\ n_r(t_i) - n_r(t_{i-1}) \end{bmatrix}.$$

Since $\sum_{j=0}^{r} d_j(i) = 0$, the covariance matrix of $(d_0(i),\ldots,d_r(i))$

is singular. Hence we base our method on the differences

$$\underset{\sim}{d}_i = \begin{bmatrix} d_0(i) \\ \cdot \\ \cdot \\ d_{r-1}(i) \end{bmatrix} = \frac{1}{n_{00}} \begin{bmatrix} n_0(t_i) - n_0(t_{i-1}) \\ \cdot \\ \cdot \\ n_{r-1}(t_i) - n_{r-1}(t_{i-1}) \end{bmatrix}.$$

We remark here that the covariance matrix $R_i = E(\underset{\sim}{d}_i\underset{\sim}{d}_i')$ of $\underset{\sim}{d}_i$ is non-singular. For, suppose the covariance matrix of $\underset{\sim}{d}_i$ is singular, then there exists a linear relation

$$\sum_{j=0}^{r-1} \alpha_j d_j(i) = \beta_0 \tag{3.1.1}$$

with probability one. That is,

$$\sum_{j=0}^{r-1} \alpha_j n_j(t_i) = \sum_{j=0}^{r-1} \alpha_j n_j(t_{i-1}) + \beta_0 n_{00}. \tag{3.1.2}$$

But this means that given $n_0(t_{i-1}),\ldots,n_r(t_{i-1})$, the random variables $n_0(t_i),\ldots,n_{r-1}(t_i)$ can take only such values as satisfy (3.1.2); but this is not true, as is shown below.

From (3.1.2), for some $\underset{\sim}{n}^*(t) = (n_0^*(t_i),\ldots,n_r^*(t_i))'$, we have

$$\sum_{j=0}^{r-1} \alpha_j n_j^*(t_i) = \sum_{j=0}^{r-1} \alpha_j n_j(t_{i-1}) + \beta_0 n_{00} = \text{constant}. \qquad (3.1.3)$$

Also $\underset{\sim}{n}(t_i) = (n_0^*(t_i) + 1,\ n_1^*(t_i) - 1,\ n_2^*(t_i),\ldots,n^*(t_i))'$

satisfies (3.1.2), hence we have from (3.1.3)

$$\alpha_0 - \alpha_1 = 0$$

i.e. $\alpha_0 = \alpha_1$.

Similarly we can show that

$$\alpha_0 = \alpha_1 = \ldots. = \alpha_{r-1}.$$

But this implies that with probability one $\sum_{j=0}^{r-1} n_j(t_i)$ is a constant, and hence $n_r(t_i)$ is also a constant, given $\underset{\sim}{n}(t_{i-1})$. This is a contradiction. This completes our proof that the covariance matrix R_i of $\underset{\sim}{d}_i$ is non-singular.

Also $R_{ipq}(\alpha) = E(d_p(i)d_q(i))$, is a linear function of $a_{0j}(t_i)$, $a_{0j}(t_{i-1})$ and $a_{jk}(t_{i-1},t_i)$. But the a_{jk}'s are transcendental functions of α. Hence R_{ipq} is not constant for all values of α.

In section 3.2 we describe an estimation procedure based on the differences $\underset{\sim}{d}_i(i = 1,\ldots k)$. This procedure is originally due to Ruben [17] and has been used by Ruben for estimating the interaction parameter in the emigration-immigration process. In section 3.5 we prove that this method of estimation yields a consistent estimate satisfying certain conditions. Then we obtain a lower limit to the asymptotic variance of a consistent estimate satisfying certain conditions, and use our result to obtain the asymptotic efficiency of the estimate. Finally we indicate a simpler method of estimating the parameter α. The modified method of estimation yields an estimate with the same properties as that obtained by the original procedure.

3.2. Derivation of the Estimation Equation: Let $\underset{\sim}{n}(t_1),\ldots,\underset{\sim}{n}(t_k)$ be k observations on the process $\underset{\sim}{n}(t)$ at times $t_j = j\tau (j=1,\ldots,k)$. Let

$$\underset{\sim}{d}_i = \begin{bmatrix} d_0(i) \\ \cdot \\ \cdot \\ d_{r-1}(i) \end{bmatrix} = \frac{1}{n_{00}} \begin{bmatrix} n_0(t_i) - n_0(t_{i-1}) \\ \cdot \\ \cdot \\ n_{r-1}(t_i) - n_{r-1}(t_{i-1}) \end{bmatrix}$$

and let $R_i = E(\underset{\sim}{d}_i \underset{\sim}{d}_i')$. Thus, $R_i = ||R_{ipq}|| = ||\ Ed_p(i)d_q(i)\ ||$

where

$$E(d_p(i)\ d_q(i)) = E\{ \frac{1}{n^2_{00}} (n_p(t_i) - n_p(t_{i-1}))(n_q(t_i) - n_q(t_{i-1}))\}$$

$$= \frac{1}{n^2_{00}} [n^2_{00}\ a_{0p}(t_i)\ a_{0q}(t_i) - n_{00}\ a_{0p}(t_i)\ a_{0q}(t_i)$$

$$- n_{00}\ a_{0q}(t_{i-1})\ a_{qp}(t_{i-1},t_i) - n_{00}(n_{00}-1) a_{0q}(t_{i-1}) a_{0p}(t_i)$$

$$- n_{00}\ a_{0p}(t_{i-1})\ a_{pq}(t_{i-1},t_i) - n_{00}(n_{00}-1)\ a_{0p}(t_{i-1}) a_{0q}(t_i)$$

$$+ n^2_{00}\ a_{0p}(t_{i-1})\ a_{0q}(t_{i-1}) - n_{00}\ a_{0p}(t_{i-1})\ a_{0q}(t_{i-1})]$$

and, we have

$$E[\frac{1}{r}\ \underset{\sim}{d}_i'\ R_i^{-1} \underset{\sim}{d}_i] = 1. \qquad (3.2.1)$$

The equation (3.2.1) is true for all values of $i, i=1,\ldots,k$. This suggests that we use

$$\frac{1}{rk} \sum_{i=1}^{k} \underset{\sim}{d}_i' \, R_i^{-1} \, \underset{\sim}{d}_i = 1 \tag{3.2.2}$$

as an estimation equation; this may be rewritten as

$$\frac{1}{rk} \sum_{i=1}^{k} \sum_{p,q=0}^{r-1} R_i^{pq} \; d_p(i) d_q(i) = 1 \tag{3.2.3}$$

where R_i^{pq} denotes the (p,q)th element of R_i^{-1}.

Any solution of (3.2.2) or (3.2.3) which effectively depends on the relative frequencies $q_i(t_j) = \dfrac{n_i(t_j)}{n_{00}}$ may be taken as an estimate of α. It may be noted that the estimation equation is a transcendental equation in α and therefore in general cannot be solved explicitly. However, a numerical solution may be found.

3.3. Preliminary Results: The purpose of this section is to prove that consistent estimates satisfying assumption F are asymptotically normally distributed, and to obtain a sufficient condition for such an estimate to have a minimum asymptotic variance. Let

$$V = [a_{00}(t_1),\ldots,a_{0,r-1}(t_1),\ldots,a_{0,r-1}(t_k)]'$$

and $T = T(V)$ be a function of V. Also

let

$$q = [q_0(t_1),\ldots,q_{r-1}(t_1),\ldots,q_{r-1}(t_k)]',$$

and

$$V_0 = \frac{\partial V}{\partial \alpha} = [\frac{\partial a_{00}(t_1)}{\partial \alpha}, \ldots, \frac{\partial a_{0,r-1}(t_1)}{\partial \alpha}, \ldots, \frac{\partial a_{0,r-1}(t_k)}{\partial \alpha}]'.$$

Further, let

$$W_0 = n_{00} E((q-V)(q-V)').$$

Assumption F: Assume that T admits continuous first partial derivatives, with respect to all $q_i(t_j)$ $(i=0,\ldots,r-1;\ j=1,\ldots,k)$.

Theorem 3.3.1: If T(q) is a consistent estimate of α satisfying assumption F, then we have the following:

(i) $T(V) = \alpha$,

(ii) $n_{00}^{\frac{1}{2}}\ (T(q) - \alpha)$ is asymptotically normally distributed with mean zero and variance $\tau' W_0 \tau$, where

$$\tau = [\frac{\partial T}{\partial a_{00}(t_1)}, \ldots, \frac{\partial T}{\partial a_{0,r-1}(t_1)}, \ldots, \frac{\partial T}{\partial a_{0,r-1}(t_k)}]'$$

and

(iii) a sufficient condition for T(q) to have a minimum variance is that $\tau' = \frac{1}{C_0} V_0' W_0^{-1}$; moreover, the minimum variance is $\frac{1}{C_0}$ where

$$C_0 = V_0' W_0^{-1} V_0 .$$

Proof: Expanding T(q) by Taylor's series about the point V up to first order terms, we have

$$T(q) = T(V) + \sum_{i,j} (q_i(t_j) - a_{0i}(t_j)) \left[\frac{\partial T}{\partial q_i(t_j)}\right]_{q_i^*(t_j)} \tag{3.3.1}$$

where $q_i^*(t_j) \in (q_i(t_j), a_{0i}(t_j))$.

Part (i) of the theorem follows from (3.3.1) since T(q) is a consistent estimate of α, $q_i(t_j)$ converges in probability to $a_{0i}(t_j)$ as n_{00} tends to infinity and $\frac{\partial T}{\partial q_i(t_j)}$ is a bounded function of q.

It follows from Rao [14,§5e] that $n_{00}^{\frac{1}{2}}$ $(T(q)-\alpha)$ is asymptotically normally distributed with mean zero and variance $\tau' W_0 \tau$ since the asymptotic distribution of q is a multivariate normal. This proves part (ii) of the theorem.

Differentiating $T(V) = \alpha$ with respect to α, we get $\tau' V_0 = 1$. Let $T_0(q)$ be any estimate of α which is consistent and satisfies assumption F, and let

$$\tau_0 = \left[\frac{\partial T_0}{\partial a_{00}(t_1)}, \ldots, \frac{\partial T_0}{\partial a_{0,r-1}(t_1)}, \ldots, \frac{\partial T_0}{\partial a_{0,r-1}(t_k)}\right]'.$$

Then $(\tau_0 - \tau)' W_0(\tau_0 - \tau)$ is non-negative and

$$\begin{aligned}(\tau_0-\tau)'W_0(\tau_0-\tau) &= \tau_0' W_0\tau_0 - \tau_0' W_0\tau - \tau' W_0\tau_0 + \tau' W_0\tau. \\ &= \tau_0' W_0\tau_0 - \frac{1}{C_0}\tau_0' W_0 W_0^{-1} V_0 - \frac{1}{C_0} V_0' W_0^{-1} W_0\tau_0 + \tau' W_0\tau \\ &= \tau_0' W_0\tau_0 - \tau' W_0\tau.\end{aligned}$$

This proves part (iii) of the theorem.

It is easy to check that the minimum variance is $\frac{1}{C_0}$.

<u>Remark</u>: Thus we have obtained a lower limit to the asymptotic variance of a consistent estimate satisfying assumption F. In section 3.5 we show that the estimation equation (3.2.2) has a root satisfying assumption F, and use the result of our theorem to study the efficiency of the M.S.C.F. estimate.

<u>3.4. An Extension of the Implicit Function Theorem</u>: In this section we prove a lemma which is an extension of the implicit function theorem. A similar extension is given by Ferguson in [7]. Proof of our lemma is essentially the same as that of a lemma due to Ferguson on page 1052 in [7]. First we state the implicit function theorem which may be found in [19], page 244 from which the lemma will follow.

<u>Implicit Function Theorem</u>: Let $x = (x_1,\dots,x_n)$ and let $F(x,z)$ be defined on an open set B containing the point (a,c). Suppose that F has continuous partial derivatives in B. Also assume that

$$F(a,c) = 0 \qquad\qquad \left(\frac{\partial F}{\partial z}\right)_{(a,c)} \neq 0.$$

Then, there exists a neighbourhood

$$A(a,c) = \{(x,z) \mid |x_i - a_i| < A_i,\ i=1,\dots,n;\ |z-c| < C\}$$

such that the following are true:

Let $N(a) = \{x \mid |x_i - a_i| < A_i,\ i=1,.,n\}$, then

(i) for any $x \in N(a)$, there is a unique z such that $|z-c| < C$ and $F(x,z) = 0$.

Let us express this dependence of z on x by $z = f(x)$.

(ii) The function f is continuous in N.

(iii) The function f has continuous first partial derivatives.

Remark: It follows from (i), that

$$f(a) = c. \qquad (3.4.1)$$

Lemma 3.4.1: Let $x = (x_1,..,x_n)$ and let $F(x,z)$ be a function defined on the open set

$$B = \{x \mid -1 < x_i < 1,\ i = 1,\dots,n;\ z \in D = (0,\infty)\}.$$

Also let $p(z)$ be a function from D into the set

$$A = \{x \mid -1 < x_i < 1,\ i = 1,\dots,n\}.$$

Assume that

(i) $p(z)$ is one-to-one and inversely continuous.

(ii) $F(x,z)$ is continuous and has continuous first partial derivatives with respect to $x_1,\dots,x_n$ and z.

(iii) $F(p(z),z) = 0$ and $\left(\frac{\partial F}{\partial z}\right)_{(p(z),z)} \neq 0$ for all $z \in D$.

Then, there exists a neighbourhood N of the set $S = \{p(z) \mid z \in D\}$ and a unique function f from the set A into the set D such that

(a) f is continuous and has continuous first partial derivatives on N,

(b) $f(p(z)) = z$ for all $z \in D$,

(c) $F(x,f(x)) = 0$ for all $x \in N$,

(d) there exists a neighbourhood of the curve $\{(p(z),z) \mid z \in D\}$ in which the only zeros of the function $F(x,z)$ are the points $(x,f(x))$.

<u>Proof</u>: From the implicit function theorem, for any $z \in D$, there is a neighbourhood $N(p(z)) = \{x | \ |x_i - p_i(z)| < A_i, i=1,\dots,n\}$ of the point $p(z) = (p_1(z),\dots,p_n(z))$ and the unique function f_z (which may, in general, depend on z) from the set $N(p(z))$ into the set D such that

(i) f_z is continuous on $N(p(z))$ and has continuous first partial derivatives,

(ii) $f_z(p(z)) = z$, (3.4.2)

and

(iii) for any point $x \in N(p(z))$,

$$F(x, f_z(x)) = 0 \qquad \text{and} \qquad |f_z(x) - z| < C_z. \tag{3.4.3}$$

That is, for any point $x \in N(p(z))$, $f_z(x) \in N_z$ where

$$N_z = (z - C_z, z + C_z) .$$

Since f_z is a continuous function, the set $f_z^{-1}(N_z)$ is an open set and contains $p(z)$. Also $f_z^{-1}(N_z) \cap N(p(z))$ is an open set containing $p(z)$. So we can choose a spherical neighbourhood $N^*(p(z))$ of $p(z)$ such that

$$N^*(p(z)) \subset f_z^{-1}(N_z) \cap N(p(z)) \qquad \text{and} \qquad p^{-1}(N^*(p(z))) \subset N_z$$

because p^{-1} is a continuous function. Now if $p(z_1) \in N^*(p(z_2))$ for any z_1, z_2 in D, then $p(z_1) \in p(N_{z_2})$ but then $z_1 \in N_{z_2}$. That is, due to inverse continuity of p and continuity of f_z we can replace the neighbourhood $N(p(z))$ by the spherical neighbourhood $N^*(p(z))$ with the additional property

(iv) if $p(z_1) \in N^*(p(z_2))$ for any $z_1, z_2 \in D$, then $z_1 \in N_{z_2}$.

Now consider the spherical neighbourhoods $N^{**}(p(z))$ with radii equal to 1/3 that of $N^*(p(z))$ with centre at $p(z)$. Let $N = \bigcup_{z \in D} N^{**}(p(z))$. The set N is clearly a neighbourhood of the set $S = \{p(z) \mid z \in D\}$.

We will show that if $x^0 \in N^{**}(p(z_1) \cap N^{**}(p(z_2))$, then $f_{z_1}(x^0) = f_{z_2}(x^0)$. Since $N^{**}(p(z_1)) \cap N^{**}(p(z_2)) \neq \emptyset$ (where $\emptyset$ denotes the null set) we have, either

$$p(z_1) \in N^*(p(z_2)) \tag{3.4.4}$$

or

$$p(z_2) \in N^*(p(z_1)). \tag{3.4.5}$$

Suppose (3.4.4) is true. Then

$$F(p(z_1), f_{z_2}(p(z_1))) = 0.$$

But $F(p(z_1), f_{z_1}(p(z_1))) = 0$ and $z_1 \in N_{z_2}$, hence

$$f_{z_2}(p(z_1)) = f_{z_1}(p(z_1)) = z_1. \tag{3.4.6}$$

If $x \in N^{**}(p(z_1)) \cap N^*(p(z_2))$, then f_{z_2} is continuous and satisfies $F(x, f_{z_2}(z)) = 0$. Also $f_{z_1}(x) \in N_{z_1}$ for $x \in N^*(p(z_1))$ and this implies that f_{z_1} is the unique function, as is shown below, which is continuous and has continuous first partial derivatives in $N^{**}(p(z_1))$ such that

$$f_{z_1}(p(z_1)) = z_1 \qquad \text{and} \qquad F(x, f_{z_1}(x)) = 0. \qquad (3.4.7)$$

Suppose g is any other continuous function , having continuous first partial derivatives in $N^{**}(p(z_1))$ such that

$$g(p(z_1)) = z_1 \qquad \text{and} \qquad F(x, g(x)) = 0.$$

Let $B = \{x \mid f_{z_1}(x) - g(x) = 0\}$. Then B is a closed set since $f_{z_1}(x) - g(x)$ is a continuous function. Let $x \in B$, then $f_{z_1}(x) = g(x)$. Since $f_{z_1}(x) \in N_{z_1}$, there exists an open set G, containing $f_{z_1}(x)$ (hence also $g(x)$) such that $f_{z_1}^{-1}(G)$ and $g^{-1}(G)$ are both contained in $N(p(z_1))$. Hence by the implicit function theorem, $f_{z_1}(x) = g(x)$ on $g^{-1}(G)$, that is, x is an interior point of B. So B is open. Therefore either B is the null set or the whole set $N^{**}(p(z_1))$. Since B is not null, B is $N^{**}(p(z_1))$. So f_{z_1} is a unique continuous function, having continuous first partial derivatives in $N^{**}(p(z_1))$ such that

$$f_{z_1}(p(z_1)) = z_1 \qquad \text{and} \qquad F(x, f_{z_1}(x)) = 0. \qquad (3.4.8)$$

Hence $$f_{z_1}(x^0) = f_{z_2}(x^0). \qquad (3.4.9)$$

If $x \in \dot{N} = \bigcup_{z \in D} N^{**}(p(z))$, then $x \in N^{**}(p(z))$ for some z.

Define

$$f(x) = f_z(x).$$

It may be remarked here that in view of (3.4.9), we may take any z such that $x \in N^{**}(p(z))$. Thus we have defined a function on N. Clearly this function has the properties (a),(b) and (c) of the lemma. For (d), the neighbourhood can be taken to be $\bigcup_{z \in D} (N^{**}(p(z)) \times N_z)$.

<u>3.5. Properties of the M.S.C.F. Estimate</u>: In this section we consider the question of existence of a root (or roots) of the estimation equation (3.2.2) and study the analytic properties of such a root. The idea of studying the analytic properties of an estimate was initiated by Rao [15] and has been found to be very useful in the theory of maximum likelihood estimation. As noted by Rao [15], "In fact, many probability statements concerning the maximum likelihood estimate are direct consequences of the continuity and differentiability properties of the maximum likelihood estimate as a function of the observed relative frequencies." This is also true for the M.S.C.F. estimate. First we prove that the M.S.C.F. estimate $\hat{\alpha}$ is a continuous function of $d_p(i)\ d_q(i)$ $(i = 1,\ldots,k;\ p,q = 0,\ldots r-1)$ possessing continuous first partial derivatives with respect to each $d_p(i)\ d_q(i)$. Then we deduce the consistency and asymptotic normality of the M.S.C.F. estimate $\hat{\alpha}$.

We need the notion of Fisher Consistency (F.C.).

<u>Definition:</u> A statistic T which is a function of $d_p(i)\ d_q(i)$ $(i = 1,\ldots,k;\ p,q = 0,\ldots,r-1)$ only is F.C. if when the expected values of $d_p(i)\ d_q(i)$, are substituted in T, the function T identically reduces to the value of the parameter.

If T is F.C. and is also a continuous function of $d_p(i)d_q(i)$, then T converges in probability to α as n_{00} tends to infinity, since as shown in (3.5.4), $d_p(i)d_q(i) - R_{ipq}$ converges in Probability to zero as n_{00} tends to infinity. That is, in this case, F.C. implies the usual consistency (convergence in probability)

Theorem 3.5.1: (i) As n_{00} tends to infinity, there exists, with probability tending to one, one and only one function $\hat{\alpha}$ of $d_p(i)d_q(i)$ ($i = 1,\ldots,k$; $p,q = 0,\ldots,r-1$) which satisfies the estimation equation (3.2.2) and has the following properties,

(ii) $\hat{\alpha}$ possesses continuous first partial derivatives with respect to all $d_p(i)d_q(i)$, (iii) $\hat{\alpha}(R(\alpha)) = \alpha$ for all $\alpha \in D$ (which implies that $\hat{\alpha}(d)$ is a consistent estimate of α),

(iv) $n_{00}^{\frac{1}{2}}(\hat{\alpha} - \alpha)$ is asymptotically normally distributed with mean zero and variance $\sigma^2(\alpha)$ as n_{00} tends to infinity, where $\sigma^2(\alpha)$ is defined in (3.5.6).

Before we present the proof of this theorem, we make a few remarks.

Remark 1: $a_{00}(t)$ is a one-to-one continuous function of α for any fixed t.

Proof: We have from (1.2.6)

$$a_{00}(t) = \left(\frac{r-m}{r-m\exp(-n_{00}(r-m)\alpha t)} \right)^r.$$

Clearly $a_{00}(t)$ is a continuous function of α. Suppose $\alpha_1 \neq \alpha_2$ but

$$\left(\frac{r-m}{r-m\exp(-n_{00}(r-m)\alpha_1 t)} \right)^r = \left(\frac{r-m}{r-m\exp(-n_{00}(r-m)\alpha_2 t)} \right)^r$$

or $m\exp(-n_{00}(r-m)\alpha_1 t) = m\exp(-n_{00}(r-m)\alpha_2 t)$. (3.5.1)

But (3.5.1) implies that $\alpha_1 = \alpha_2$. This proves the one-to-one continuity of $a_{00}(t)$.

Remark 2: Let d be a vector whose elements are $d_p(i)d_q(i)$ $(i = 1,\ldots,k;\ p,q = 0,\ldots,r-1)$ and $R(\alpha) = E(d)$. Then $R(\alpha)$ is a one-to-one continuous function of α.

Proof: Clearly $R(\alpha)$ is a continuous function of α because $R_{ipq}(\alpha)$ $(i = 1,\ldots,k;\ p,q = 0,\ldots,r-1)$ is a continuous function of α. Now $R(\alpha)$ is a one-to-one function of α if one element of $R(\alpha)$ is a one-to-one function of α. We show that $R_{100}(\alpha)$ is a one-to-one function of α. Be definition

$$R_{100}(\alpha) = n_{00}^{-2}\ E(n_0(t_1) - n_{00})^2$$
$$= (1 - a_{00}(t_1))^2 + \frac{a_{00}(t_1)\ (1 - a_{00}(t_1))}{n_{00}}\ .$$

Clearly $R_{100}(\alpha)$ is a one-to-one function of $a_{00}(t_1)$ which is a one-to-one function of α. Hence $R_{100}(\alpha)$ is a one-to-one function of α.

Remark 3: If $f(x) = (f_1(x),\ldots,f_k(x))$ is a one-to-one continuous vector valued function of a real variable x, then f(x) is inversely continuous if one of the functions $f_1(x),\ldots,f_k(x)$ is one-to-one and inversely continuous.

Proof: Suppose $f_1(x)$ is one-to-one and inversely continuous. Let $f^n(x_0)$; $n = 0,1,\ldots$ be a sequence of points in the range space of f(x) such that $f^n(x_0)$ tends $f^0(x_0)$ as n tends to infinity. Then $f_1^n(x_0)$ tends to $f_1^0(x_0)$. Due to the inverse continuity of $f_1(x)$, $f_1^{-1}(f_1^n(x_0)$ tends to $f_0^{-1}(f_1^0(x_0)) = x_0$. But

$(f\ )^{-1}(f^n(x_0)) = f_1^{-1}(f_1^n(x_0)$, hence $(f_n)^{-1}(f^n(x_0))$ tends to x_0 as n tends to infinity. This proves remark 3.

Remark 4: In view of remarks 2 and 3, $R(\alpha)$ is a one-to-one and bicontinuous function of α.

Proof of Theorem 3.5.1, (i) and (ii): The estimation equation is

$$\frac{1}{rk} \sum_i \sum_{p,q} R_i^{pq}\, d_p(i) d_q(i) - 1 = 0 .$$

The expression on the left hand side of the above equation is a function of d and α. Denote this function by $F(d,\alpha)$.

By assumption B_2, $\alpha \in (0,\infty)$. The function $R(\alpha)$ is a one-to-one and inversely continuous function of α as shown above. Also the function $F(d,\alpha)$ is continuous in d and α. Clearly $\frac{\partial F}{\partial \alpha}$ exists and is given by

$$\frac{\partial F}{\partial \alpha} = \frac{1}{rk} \sum_i \sum_{p,q} \frac{\partial R_i{}^{pq}}{\partial \alpha}\, d_p(i) d_q(i)$$

which is a continuous function of d and α. Also the derivatives $\frac{\partial F}{\partial (d_p(i) d_q(i))}$ exist and are continuous. We have

$$\sum_i \sum_{p,q} R_i{}^{pq} R_{ipq} = rk. \tag{3.5.2}$$

Differentiating (3.5.2) with respect to α, we get

$$\sum_i \sum_{p,q} \frac{\partial R_i{}^{pq}}{\partial \alpha} R_{ipq} = - \sum_i \sum_{p,q} R_i^{pq} \frac{\partial R_{ipq}}{\partial \alpha}$$

$$= - \sum_i |R_i| \frac{\partial}{\partial \alpha} |R_i| \neq 0.$$

Thus we see that all the conditions of lemma 3.4.1 are satisfied. Hence there exists a neighbourhood N of the set $S = \{R(\alpha) | \alpha \epsilon D\}$ and a unique function $\hat{\alpha}(d)$ from N to D such that

(a) $\hat{\alpha}(d)$ is continuous and has continuous first partial derivatives,

(b) $\hat{\alpha}(R(\alpha)) = \alpha$ for all $\alpha \in D$,

(c) $F(R(\alpha),\alpha) - 1 = 0$ for all $R(\alpha) \in N$,

(d) there exists a neighbourhood of the curve $\{(R(\alpha),\alpha)|\alpha\epsilon D\}$ in which the only zeros of the function $F(x,z)$ are the points $(d,\hat{\alpha}(d))$.

Thus we see that for $d \in N$, the estimation equation has one and only one root $\hat{\alpha}(d)$ which possesses the properties mentioned in (a) through (d). By definition $R_{ipq} = E(d_p(i)d_q(i))$

$$= n_{00}^{-2} E(n_p(t_i) - n_p(t_{i-1}))\ (n_q(t_i) - n_q(t_{i-1}))$$

$$= n_{00}^{-2} E[n_p(t_i)n_q(t_i) - n_p(t_{i-1})n_q(t_i) - n_p(t_i)n_q(t_{i-1}) + n_p(t_{i-1})n_q(t_{i-1})]$$

$$= n_{00}^{-2}[n_{00}^2\ a_{0p}(t_i)a_{0q}(t_i) - n_{00}\ a_{0p}(t_i)\ a_{0q}(t_i)$$

$$- n_{00}\ a_{0q}(t_{i-1})a_{qp}(t_{i-1},t_i) - n_{00}(n_{00}-1)a_{0q}(t_{i-1})a_{0p}(t_i)$$

$$- n_{00}\ a_{0p}(t_i)a_{pq}(t_{i-1},t_i) - n_{00}(n_{00}-1)a_{0p}(t_i)a_{0q}(t_i)$$

$$+ n_{00}^2\ a_{0p}(t_{i-1})a_{0q}(t_{i-1}) - n_{00}a_{0p}(t_{i-1})a_{0q}(t_{i-1})]$$

$$\rightarrow (a_{0p}(t_i) - a_{0p}(t_{i-1}))(a_{0q}(t_i) - a_{0q}(t_{i-1})) \qquad (3.5.3)$$

as n_{00} tends to infinity.

Also

$$d_p(i)d_q(i) = n_{00}^{-2}(n_p(t_i)-n_p(t_{i-1}))(n_q(t_i)-n_q(t_{i-1}))$$

$$\xrightarrow{P} (a_{0p}(t_i)-a_{0p}(t_{i-1}))(a_{0q}(t_i)-a_{0q}(t_{i-1})) \quad (3.5.4)$$

as n_{00} tends to infinity.

Hence from (3.5.3) and (3.5.4), we have

$$d_p(i)d_q(i) - R_{ipq} \xrightarrow{P} 0 \quad (3.5.5)$$

as n_{00} tends to infinity.

Therefore given ε,η positive, there exists $n(\varepsilon,\eta)$ such that for $n_{00} > n(\varepsilon,\eta)$

$$P(|d-R| < \varepsilon) > 1-\eta.$$

Hence with probability tending to one, as n_{00} tends to infinity the estimation equation has one and only one root which possesses the properties mentioned in the theorem.

Proof of (iii): If follows from (b) that $\hat{\alpha}(d)$ is F.C. Since $\hat{\alpha}(d)$ is a continuous function of d, and $d_p(i)d_q(i) - R_{ipq} \xrightarrow{P} 0$ as n_{00} tends to infinity, $\hat{\alpha}(d) \xrightarrow{P} \alpha$ as n_{00} tends to infinity. That is, $\hat{\alpha}(d)$ is consistent.

Proof of (iv): To prove this part of the theorem, we need the following lemma:

Lemma 3.5.1: If $X_n - Y_n \xrightarrow{P} 0$ and $F_n(x) \xrightarrow{W} F(x)$, then $G_n(x) \xrightarrow{W} F(x)$ where $F_n(x) = P(X_n \le x)$ and $G_n(x) = P(Y_n \le x)$.

This is a well known result.

By Taylor's formula, we have

$$n_{00}^{\frac{1}{2}} (\hat{\alpha}(d) - \alpha) = n_{00}^{\frac{1}{2}} \sum_i \sum_{p,q} (d_p(i)d_q(i) - R_{ipq}) \left[\frac{\partial\hat{\alpha}}{\partial(d_p(i)d_q(i))}\right]_{d*}$$

where d_1^* is a point on the line segment joining d and R. Let

$$X_{n_{00}} = n_{00}^{-\frac{1}{2}} \sum_i \sum_{p,q} (a_{0q}(t_i)-a_{0q}(t_{i-1}))[n_p(t_i)-n_p(t_{i-1})-a_{0p}(t_i)-a_{0p}(t_{i-1})].$$

But

(i) $n_{00}^{-1}(n_q(t_i)-n_q(t_i \overset{P}{\to} (a_{0q}(t_i)-a_{0q}(t_{i-1}))$ as $n_{00} \to \infty$;

(ii) we know from (3.5.3), that

$$R_{ipq} \to (a_{0p}(t_i) - a_{0p}(t_{i-1}))\ (a_{0q}(t_i) - a_{0q}(t_{i-1}))$$

as n_{00} tends to infinity; and

(iii) $\left[\frac{\partial\hat{\alpha}}{\partial(d_p(i)d_q(i))} \right]_{d*} \overset{P}{\to} \frac{\partial\hat{\alpha}}{\partial R_{ipq}}$

as n_{00} tends to infinity, since $\hat{\alpha}(d)$ has continuous partial derivatives. Clearly (i),(ii) and (iii) imply that $n_{00}^{\frac{1}{2}} (\hat{\alpha} - \alpha) - X_{n_{00}}$ converges in probability to zero as n_{00} tends to infinity. But it follows from Rao[14,§ 5e] that $X_{n_{00}}$ is asymptotically normally distributed with mean zero and variance $\sigma^2(\alpha)$ since $X_{n_{00}}$ is a linear function of normally distributed random variables. Here

$$\sigma^2(\alpha) = \frac{\sum\limits_{i,i'} \sum\limits_{p,q,p'q'} [R_i^{pq}\ R_{i'}^{p'q'}]\ A_i^q\ A_{i'}^{q'}\ E(X_p(i)X_{p'}(i'))}{\left(\sum\limits_i \sum\limits_{pq} \frac{\partial R_i^{pq}}{\partial\alpha}\ R_{ipq} \right)^2} \tag{3.5.6}$$

where $X_p(i) = n_{00}^{-\frac{1}{2}} (n_p(t_i)-n_p(t_{i-1})-a_{0p}(t_i)-a_{0p}(t_{i-1}))$ and

$A_i^q = (a_{0q}(t_i) - a_{0q}(t_{i-1}))$.

Hence $n_{00}^{\frac{1}{2}} (\hat{\alpha} - \alpha)$ is asymptotically normally distributed with mean zero and variance $\sigma^2(\alpha)$.

3.6. Efficiency of the M.S.C.F. Estimate: We now discuss the asymptotic efficiency of the M.S.C.F. estimate $\hat{\alpha}$. In the theorem 3.5.1. we proved that the M.S.C.F. estimate $\hat{\alpha}$ is asymptotically normally distributed with mean α and variance $\frac{1}{n_{00}} \sigma^2(\alpha)$. In section 3.3 we considered the whole class of consistent estimates satisfying assymption F and proved that the variance of a consistent estimate is greater than or equal to $\frac{1}{C_0}$.

Hence the efficiency of the M.S.C.F. estimate $\hat{\alpha}$ of α is given by

$$\frac{1}{n_{00}C_0\sigma^2(\alpha)} .$$

3.7. A Modified Estimation Procedure: The estimation procedure described in section 3.2. is somewhat lengthy and is not suitable for numerical computation. In practice the value of r (the maximum number of phages that can become attached to a bacterium) is usually quite large (for example, 130 or 140), and this requires the inversion of a 140 × 140 matrix. In this section we describe a simpler method which is essentially a modification of the method described in 3.2.

The modification consists in pooling the data in p classes. Let $r_1, \ldots, r_p$ be p positive integers such that $r_1 + \ldots + r_p = r-1$ and let

$$m_0(t) = \sum_{j=0}^{r_1} n_j(t)$$

$$m_1(t) = \sum_{j=r_1+1}^{r_2} n_j(t)$$

$$\vdots$$

$$m_p(t) = \sum_{j=r_1+.+r_{p-1}+1}^{r_1+.+r_p} n_j(t).$$

In practice, we may take p = 4 from the point of view of numerical calculation.

<u>Modified Estimation Equation</u>:

Let $$\underset{\sim}{\delta}_{(i)} = \frac{1}{n_{00}} \begin{bmatrix} m_0(t_i) - m_0(t_{i-1}) \\ \vdots \\ m_p(t_i) - m_p(t_{i-1}) \end{bmatrix}$$

and $\eta_i = E(\underset{\sim}{\delta}_{(i)}\underset{\sim}{\delta}_{(i)}')$, then we have

$$\frac{1}{p} E(\underset{\sim}{\delta}_{(i)}' \eta_i^{-1} \underset{\sim}{\delta}_{(i)}) = 1. \qquad (3.7.1)$$

The equation (3.7.1) is true for all values of i, i = 1,...,k. This suggest that we use

$$\frac{1}{pk} \sum_{i=1}^{k} (\underset{\sim}{\delta}_{(i)}' \eta_i^{-1} \underset{\sim}{\delta}_{(i)}) = 1, \qquad (3.7.2)$$

as an estimation equation. The estimation equation (3.7.2) may be rewritten as

$$\frac{1}{pk} \sum_{i=1}^{k} \sum_{p,q=0}^{p} \eta_i^{pq} \delta_{p(i)} \delta_{q(i)} = 1 \tag{3.7.3}$$

where η_i^{pq} denotes the (p,q)th element of η_i^{-1}. It may be noted that the estimation equation (3.7.3) connot be solved explicitly; however, numerical solutions can be obtained fairly readily. It is clear that the modified procedure of estimation will yield several estimates, one of these with the same asymptotic properties as that obtained by the original procedure described in 3.2.

In general (3.7.3) will have many roots and the problem of selecting the root satisfying the conditions of theorem 3.5.1 does not seem to have a simple solution and needs some further investigation.

Suppose $\alpha_1,\ldots,\alpha_p$ are p roots of (3.7.3), then sometimes it may be possible to determine p functions $\hat{\alpha}_1,\ldots,\hat{\alpha}_p$ of d such that $\hat{\alpha}_1(d) = \alpha_1,\ldots,\hat{\alpha}_p(d) = \alpha_p$. Now according to theorem 3.5.1 there is one and only one root which possessses the properties (ii) and (iii) of theorem 3.5.1 and we can select this root by checking these conditions. This is one possible solution and it is proposed to study this method in some detail.

BIBLIOGRAPHY

[1] ALBERT, ARTHUR (1962). Estimating the Infinitesimal Generator of a Continuous Time Finite State Markov Process. Ann. Math. Statist. 33, pp. 727-753.

[2] ANSCOMBE, F. J. (1953). Sequential Estimation. Jour.Roy.Statist.Soc., B 15, pp.1-21.

[3] BARTLETT, M. S. (1949). Some Evolutionary Stochastic Processes. Jour.Roy.Statist.Soc., B 11, pp.211-229.

[4] BILLINGSLEY, P. (1961). Statistical Inference for Markov Processes. Institute of Mathematical Statistics, Univ. of Chicago. Statistical Research Monographs, Univ. of Chicago Press.

[5] BRENNER, S. (1955). The Adsorption of Bacteriophages by Sensitive and Resistant Cell of Escherichia Coli Strain B. Proc.Roy.Soc. 3.144, pp. 93-99.

[6] DOOB, J. L. (1953). Stochastic Processes. John Wiley, New York.

[7] FERGUSON, T. (1958). A Method of Generating Best Asymptotically Normal Estimate with Application to the Estimation of Bacterial Densities. Ann.Math. Statist. 29, pp.1046-1062.

[8] GANI, J. (1963). Models for a Bacterial Growth Process with Removals. Jour.Roy.Statist.Soc.,B 25, pp.140-149.

[9] GANI, J. (1965). Stochastic Phage Attachment to Bacteria. To appear in Biometrics, March, 1965.

[10] KENDALL, D. G. (1949). Stochastic Processes and Population Growth. Jour.Roy.Statist.Soc., B 11, pp. 230-264.

[11] MORAN, P.A.P. (1951). Estimation Methods for Evolutive Processes. Jour.Roy.Statist.Soc., B 13, pp. 141-145.

[12] MORAN, P.A.P. (1953). The Estimation of the Parameter of a Birth and Death Process. Jour.Roy.Statist.Soc., B 15, pp. 241-245.

[13] PARZEN, E. (1954). On Uniform Convergence of Families of Sequences of Random Variables, University of Calif. publications is Statistics, 2, No. 2, pp. 23-54.

[14] RAO, C. R. (1952). Advanced Statistical Methods in Biometric Research.

[15] RAO, C. R. (1957). Maximum Likelihood Estimation for the Multinomial Distribution. Sankhyá 18, pp. 139-148.

[16] RUBEN, HAROLD (1962). Some Aspects of the Emigration-immigration Process. Ann.Math.Statist. 33, pp. 111-129.

[17] RUBEN, HAROLD (1963). The Estimation of a Fundamental Interaction Parameter in an Emigration-immigration Process. Ann.Math.Statist. 84, pp. 238-259.

[18] SETHURAMAN, J. (1961). Some Limit Theorems for Joint Distributions. Sankhyá 23, A pp. 379-386.

[19] TAYLOR, A. E. (1955). Advanced Calculus. Ginn and Company.

[20] YASSKY, D. (1962). A Model for the Kinetics of Phage Attachment to Bacteria in Suspension. Biometrics 18, pp. 185-191.