The Annals of Mathematical Statistics
1968, Vol. 39, No. 1, 183–192

ESTIMATION OF THE PARAMETER IN THE STOCHASTIC MODEL FOR PHAGE ATTACHMENT TO BACTERIA[1]

BY R. C. SRIVASTAVA

The Ohio State University and Banaras Hindu University

0. Summary. Recently Gani [3] has considered a stochastic model for the attachment of phages to bacteria. In this paper we describe a simple method of estimating the parameter α which occurs in this model and study the asymptotic properties of the estimate.

1. Introduction. The stochastic model for phage attachment to bacteria gives rise to a multivariate stochastic process $N(t) = (n_0(t), \cdots, n_r(t))'$, depending on a single unknown parameter α. The process $N(t)$ is Markovian and its transition probabilities are functions of α.

Let n_{00} be the number of bacteria in a nutrient medium and suppose that v_{00} be the number of phages released into it. Then in a random fashion phages attach themselves to bacteria. Also let $m = v_{00}/n_{00}$ be the multiplicity of phages and r the saturation capacity of a bacterium. Further let $n_i(t)$ $(i = 0, 1, \cdots, r)$ be the number of bacteria with exactly i phages attached to them and $v_0(t)$ be the number of free phages at time t, $0 < t < t_0$. The duration of the experiment is taken to be small, less than t_0, so that the number of bacteria or phages neither increases nor decreases in this period. Let $P(n_0, \cdots, n_r; t)$ denote the probability that there are $n_0, \cdots, n_r$ bacteria with $0, \cdots, r$ phages attached to them respectively at time $t \geqq 0$. Now if the probability of attachment during interval $(t, t + \delta t)$ of a phage to a bacterium already having exactly i phages be $\lambda_i n_i v_0 \delta t + o(\delta t)$, $i = 0, 1, \cdots, r$, where $\lambda_i = (r - i)\alpha$, $\lambda_r = 0$, $\alpha > 0$, then it is shown by Gani [3] that

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

that is, at any fixed time t, the distribution of $N(t) = [n_0(t), \cdots, n_r(t)]'$ is multinomial with parameter n_{00} and probabilities $a_{00}(t), \cdots, a_{0r}(t)$. The probabilities $a_{0j}(t)$ are functions of a single parameter α, defined by

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

where

$$\rho(t) = \alpha^{-1} \log \{[r - m \exp(-\mu\alpha t)]/(r - m)\}; \qquad \mu = n_{00}(r - m).$$

It is easily seen that expected value of $N(t)$ is $n_{00}[a_{00}(t), \cdots, a_{0r}(t)]'$ and the

Received 27 May 1966; revised 13 June 1967.

[1] This paper is based on a portion of the author's doctoral dissertation, accepted by Michigan State University. This research was supported by the National Science Foundation under contracts G-18976, GP-2496 and GP-4307.

covariance matrix is $\Sigma(t) = ||\Sigma_{ij}||$ where

$$\Sigma_{ii}(t) = n_{00}a_{0i}(t)(1 - a_{0i}(t)); \qquad \Sigma_{ij}(t) = -n_{00}a_{0i}(t)a_{0j}(t).$$

Some of the basic properties of this process were studied by the author [6]. It is shown that for $t_1 < t_2$ (equation (2.8) in [6])

$$\text{Cov}\,\{n_i(t_1)n_j(t_2)\} = n_{00}a_{0i}(t_1)\{a_{ij}(t_1, t_2) - a_{0j}(t_2)\}$$

where

$$\begin{aligned} a_{ij}(t_1, t_2) &= \tbinom{r-i}{j-i} \exp\left(-(r-i)\alpha\rho(t_1, t_2)\right)[\exp\left(\alpha\rho(t_1, t_2)\right) - 1]^{j-i}, & j \geqq i, \\ &= 0, & j < i, \end{aligned}$$

and

$$\rho(t_1, t_2) = \alpha^{-1} \log\{[r - m \exp(-\mu\alpha t_2)]/[r - m \exp(-\mu\alpha t_1)]\}.$$

Let $\xi_j(t) = n_{00}^{-\frac{1}{2}}[n_j(t) - n_{00}a_{0j}(t)]$ and $\xi(t) = (\xi_0(t), \cdots, \xi_r(t))'$. It is proved in [6], Theorem 4.3, that if m, the multiplicity of phages is such that $m \to r$ with $m < r$ and n_{00} tends to infinity such that $n_{00}(r - m) = \mu_0$ where $\mu_0 > 0$ is a fixed constant, then the asymptotic joint distribution of a finite number of observations $\xi(t_1), \cdots, \xi(t_k)(t_1 < t_2 \cdots < t_k)$ on $\xi(t)$ is a multivariate normal distribution. We will need this result to prove the asymptotic normality of our estimate $\hat{\alpha}$ of α.

In Section 2 we describe a simple method, of the type originated by Ruben [5], for estimating the parameter α. The estimate is based on the consecutive differences of the relative frequencies $n_i(t_j)/n_{00}$ observed at discrete time points $t_1, \cdots, t_k$; $t_j = j\tau$ ($\tau > 0, j = 1, \cdots, k$). Following Ruben we call such an estimate the MSCF (Mean Square Consecutive Fluctuation) estimate. In Section 3 we prove an extension of the implicit function theorem which is used to prove the existence of the estimate and its asymptotic properties in Section 4. In Section 5 we derive a lower limit to the asymptotic variance of a consistent estimate satisfying certain conditions and use our results to derive an expression for the asymptotic efficiency of the estimate.

2. Derivation of the estimation equation. Let $N(t_1), \cdots, N(t_k)$ be k observations on the process $N(t)$ at times $t_j = j\tau$ ($\tau > 0, j = 1, \cdots, k$).

Let

$$\begin{aligned} \mathbf{d}(i) &= [d_0(i), \cdots, d_{r-1}(i)]' \\ &= n_{00}^{-1}[n_0(t_i) - n_0(t_{i-1}), \cdots, n_{r-1}(t_i) - n_{r-1}(t_{i-1})]' \end{aligned}$$

and let

$$R_i = E(\mathbf{d}(i)\mathbf{d}'(i)).$$

Thus $R_i = \|R_{ipq}\|$ where $\|a_{pq}\|$ denotes a matrix whose (p, q)th element is a_{pq} and

$$\begin{aligned} R_{ipq} &= E(d_p(i)\, d_q(i)) \\ &= n_{00}^{-2}[n_{00}^2 a_{0p}(t_i)a_{0q}(t_i) - n_{00}a_{0p}(t_i)a_{0q}(t_i) \\ &\quad - n_{00}a_{0q}(t_{i-1})a_{qp}(t_{i-1}, t_i) - n_{00}(n_{00} - 1)a_{0q}(t_{i-1})a_{0p}(t_i) \\ &\quad - n_{00}a_{0p}(t_{i-1})a_{pq}(t_{i-1}, t_i) - n_{00}(n_{00} - 1)a_{0p}(t_{i-1})a_{0q}(t_i) \\ &\quad + n_{00}^2 a_{0p}(t_{i-1})a_{0q}(t_{i-1}) - n_{00}a_{0p}(t_{i-1})a_{0q}(t_{i-1})]. \end{aligned}$$

Now we have

$$E[r^{-1}\mathbf{d}_i'R_i^{-1}\mathbf{d}_i] = 1. \tag{2.1}$$

The equation (2.1) is true for all values of i; $i = 1, \cdots, k$. This suggests that we use

$$(rk)^{-1} \textstyle\sum_{i=1}^{k} \mathbf{d}'(i)R_i^{-1}\mathbf{d}(i) = 1 \tag{2.2}$$

as an estimation equation; this may be rewritten as

$$(rk)^{-1} \textstyle\sum_{i=1}^{k} \sum_{p,q=0}^{r-1} R_i^{pq}\, d_p(i)\, d_q(i) = 1 \tag{2.3}$$

where R_i^{pq} denotes the (p, q)th element of R_i^{-1}. Any solution of (2.2) or (2.3) which effectively depends on the relative frequencies $q_i(t_j) = n_i(t_j)/n_{00}$ 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. An extension of the implicit function theorem. In this section we prove an extension of the implicit function theorem which may be found in Taylor [8], page 1052. The proof of our lemma is essentially similar to that of a lemma due to Ferguson on page 1052 in [2]. First we state the implicit function theorem.

IMPLICIT FUNCTION THEOREM. *Let $x = (x_1, \cdots, x_n)$ 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*

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

Then, there exists a neighbourhood

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

such that the following are true:

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

$$f(a) = c. \tag{3.1}$$

LEMMA 3.1. *Let* $x = (x_1, \cdots, x_n)$ *and let* $F(x, z)$ *be a function defined on the open set*

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

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

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

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* $x_1, \cdots, x_n$ *and* z.

(iii) $F(p(z), z) = 0$ *and* $(\partial F/\partial z)_{(p(z),z)} \neq 0$ *for all* $z \varepsilon D$.

Then, there exists a neighbourhood N *of the set* $S = \{p(z) \mid z \varepsilon D\}$ *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* $z \varepsilon D$,

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

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

PROOF. From the implicit function theorem, for any $z \varepsilon D$, there is a neighbourhood $N(p(z)) = \{x \mid |x_i - p_i(z)| < A_i, i = 1, \cdots, n\}$ of the point $p(z) = (p_1(z), \cdots, p_n(z))$ 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,

$$\text{(3.2)} \quad \text{(ii)} \qquad f_z(p(z)) = z,$$

and

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

$$\text{(3.3)} \qquad F(x, f_z(x)) = 0 \quad \text{and} \quad |f_z(x) - z| < C_z .$$

That is, for any point $x \varepsilon N(p(z)), f_z(x) \varepsilon N_z$ 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 $N^*(p(z)) \subset f_z^{-1}(N_z) \cap N(p(z))$ and $p^{-1}(N^*(p(z))) \subset N_z$ because p^{-1} is a continuous function. Now if $p(z_1) \varepsilon N^*(p(z_2))$ for any z_1, z_2 in D, then $p(z_1) \varepsilon p(N_{z_2})$ but then $z_1 \varepsilon N_{z_2}$. 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 $p(z_1) \varepsilon N^*(p(z_2))$ for any $z_1, z_2 \varepsilon D$, the $z_1 \varepsilon N_{z_2}$.

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

We will show that if $x^0 \varepsilon N^{**}(p(z_1)) \cap N^{**}(p(z_2))$, 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

$$p(z_1) \varepsilon N^*(p(z_2)) \tag{3.4}$$

or

$$p(z_2) \varepsilon N^*(p(z_1)). \tag{3.5}$$

Suppose (3.4) 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 $z_1 \varepsilon N_{z_2}$, hence

$$f_{z_2}(p(z_1)) = f_{z_1}(p(z_1)) = z_1; \tag{3.6}$$

If $x \varepsilon N^{**}(p(z_1)) \cap N^*(p(z_2))$, then f_{z_2} is continuous and satisfies $F(x, f_{z_2}(z)) = 0$. Also $f_{z_1}(x) \varepsilon N_{z_1}$ for $x \varepsilon N^*(p(z_1))$ 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

$$f_{z_1}(p(z_1)) = z_1 \quad \text{and} \quad F(x, f_{z_1}(x)) = 0. \tag{3.7}$$

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

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

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 $x \varepsilon B^*$, then $f_{z_1}(x) = g(x)$. Since $f_{z_1}(x) \varepsilon N_{z_1}$, 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

$$f_{z_1}(p(z_1)) = z_1 \quad \text{and} \quad F(x, f_{z_1}(x)) = 0. \tag{3.8}$$

Hence

$$f_{z_1}(x^0) = f_{z_2}(x^0). \tag{3.9}$$

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

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

It may be remarked here that in view of (3.9), we may take any z such that $x \varepsilon N^{**}(p(z))$. 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 $\bigcup_{z \varepsilon D} (N^{**}(p(z)) \times N_z)$.

4. Properties of the MSCF estimate. In this section we consider the question of existence of a root (or roots) of the estimating equation and study the asymptotic properties of the MSCF estimate $\hat{\alpha}$ of α. In Theorem 4.1 we prove that there exists a unique root $\hat{\alpha}$ of the estimation equation which is a continuous function of $d_p(i)d_q(i)$ $(i = 1, \cdots, k; p, q = 0, \cdots, r - 1)$ possessing continuous first partial derivative with respect to each $d_p(i)d_q(i)$. Then we deduce the consistency and asymptotic normality of the estimate $\hat{\alpha}$ of α. We also obtain an expression of the asymptotic variance of $\hat{\alpha}$.

THEOREM 4.1. (i) *As n_{00} tends to infinity, there exists, with probability tending to one, one and only one function $\hat{\alpha}$ of $d_p(i)d_q(i)$ $(i = 1, \cdots, k; p, q = 0, \cdots, r-1)$ which satisfies the estimation equation* (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 $\alpha \varepsilon D$ (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* (4.6).

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 as can be seen by observing the form of $a_{00}(t)$ and checking that its derivative with respect to α is positive for all positive α.

REMARK 2. Let d be a vector defined by

$$(d_0(1)d_0(1), \cdots, d_{r-1}(1)d_{r-1}(1), \cdots, d_{r-1}(k)d_{r-1}(k))$$

and let $R(\alpha) = E(d)$. Then $R(\alpha)$ is a one-to one continuous function of α.

PROOF. Clearly $R(\alpha)$ is a continuous function of α because $R_{ipq}(\alpha)$ $(i = 1, \cdots, k; p, q = 0, \cdots, r - 1)$ 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. By definition

$$\begin{aligned} R_{100}(\alpha) &= n_{00}^{-2}E(n_0(t_1) - n_{00})^2 \\ &= (1 - a_{00}(t_1))^2 + a_{00}(t_1)(1 - a_{00}(t_1))/n_{00}. \end{aligned}$$

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 $f(x) = (f_1(x), \cdots, f_k(x))$ 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 $f_1(x), \cdots, f_k(x)$ is one-to-one and inversely continuous.

PROOF. Suppose $f_1(x)$ is one-to-one and inversely continuous. Let $f^n(x_0)$; $n = 0, 1, \cdots$ be a sequence of points in the range space of $f(x)$ such that $f^n(x_0)$ tends to $f^0(x_0)$ 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 $f_1^{-1}(f_1^0(x_0)) = x_0$. 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 4.1, (i) AND (ii). The estimation equation is

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

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 $\alpha \varepsilon (0, \infty)$. 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 $\partial F/\partial\alpha$ exists and is given by

$$\partial F/\partial \alpha = (rk)^{-1} \sum_i \sum_{p,q} (\partial R_i^{pq}/\partial\alpha) d_p(i) d_q(i) \tag{4.1}$$

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

$$\sum_i \sum_{p,q} R_i^{pq} R_{ipq} = rk. \tag{4.2}$$

Differentiating (4.2) with respect to α, we get

$$\begin{aligned}\sum_i \sum_{p,q} (\partial R_i^{pq}/\partial\alpha) R_{ipq} &= -\sum_i \sum_{p,q} R_i^{pq} (\partial R_{ipq}/\partial\alpha) \\ &= -\sum_i |R_i|^{-1} (\partial/\partial\alpha)|R_i| \neq 0.\end{aligned}$$

The last sum is not equal to zero because R_i is a non-singular matrix and is not equal to a constant for all values of the parameter α. Thus we see that all the conditions of Lemma 3.1 are satisfied. Hence there exists a neighbourhood N of the set $S = \{R(\alpha) | \alpha \varepsilon D\}$ 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 $\alpha \varepsilon D$,

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

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

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

$$\begin{aligned} R_{ipq} &= E(d_p(i) d_q(i)) \\ &= n_{00}^{-2} E(n_p(t_i) - n_p(t_{i-1}))(n_q(t_i) - n_q(t_{i-1})) \\ &\to (a_{0p}(t_i) - a_{0p}(t_{i-1}))(a_{0q}(t_i) - a_{0q}(t_{i-1})) \end{aligned} \tag{4.3}$$

as n_{00} tends to infinity. Also

$$\begin{aligned} d_p(i) d_q(i) &= n_{00}^{-2} (n_p(t_i) - n_p(t_{i-1}))(n_q(t_i) - n_q(t_{i-1})) \\ &\to_P (a_{0p}(t_i) - a_{0p}(t_{i-1}))(a_{0q}(t_i) - a_{0q}(t_{i-1})) \end{aligned} \tag{4.4}$$

as n_{00} tends to infinity. Hence from (4.3) and (4.4), we have

$$d_p(i) d_q(i) - R_{ipq} \to_P 0 \tag{4.5}$$

as n_{00} tends to infinity. Therefore $d - R \to_P 0$ as n_{00} tends to infinity. 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). Since $\hat{\alpha}(d)$ is a continuous function of d, and $d_p(i)d_q(i) - R_{ipq} \rightarrow_P 0$ as n_{00} tends to infinity, it follows from (b) that $\alpha(d) \rightarrow_P \alpha$ as n_{00} tends to infinity. Here $\rightarrow_P$ denotes convergence in probability. That is $\hat{\alpha}(d)$ is consistent.

PROOF OF (iv). Let

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

and

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

By Taylor's formula, we have

$$n_{00}^{\frac{1}{2}}(\hat{\alpha}(q) - \alpha) = n_{00}^{\frac{1}{2}} \sum_i \sum_j (q_i(t_j) - a_{0i}(t_j))[\partial\hat{\alpha}/\partial q_i(t_j)]_{q^*}$$

where q^* is a point on the line segment joining q and V. Since $[\partial\hat{\alpha}/\partial q_i(t_j)]_{q^*} \rightarrow_P \partial\hat{\alpha}/\partial a_{0i}(j)$; it follows from Rao (4.5e) that $n_{00}^{\frac{1}{2}}(\hat{\alpha}(q) - \alpha)$ is asymptotically normally distributed with mean zero and variance $\sigma^2(\alpha)$. Here

$$(4.6) \quad \sigma^2(\alpha) = \sum_{i,i'} \sum_{j,j'} [(\partial\alpha/\partial a_{0i}(t_j))(\partial\alpha/\partial a_{0i'}(t_{j'}))] \cdot a_{0i}(t_j)(a_{ii'}(t_j, t_{j'}) - a_{0i'}(t_{j'})).$$

5. Efficiency of the MSCF estimate. We now discuss the asymptotic efficiency of the MSCF estimate $\hat{\alpha}$. First we prove that consistent estimates satisfying Assumption A are asymptotically normally distributed and obtain a sufficient condition for such an estimate to have a minimum asymptotic variance. We also obtain a lower limit to the variance of consistent estimates satisfying Assumption A and use this result to obtain an expression for the asymptotic efficiency of $\hat{\alpha}$.

Let us recall that

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

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

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

and

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

Further, let

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

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

THEOREM 5.1. *If $T(q)$ is a consistent estimate of α satisfying Assumption A, 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*

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

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

$$(5.1) \qquad T(q) = T(V) + \sum_{i,j} (q_i(t_j) - a_{0i}(t_j))[\partial T/\partial q_i(t_j)]_{q_i^*(t_j)}$$

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

Part (i) of the theorem follows from (5.1) 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 $\partial T/\partial q_i(t_j)$ is a bounded function of q.

It follows from Rao [3], Section 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 A, and let

$$\tau_0 = [\partial T_0/\partial a_{00}(t_1), \cdots, \partial T_0/\partial a_{0,r-1}(t_1), \cdots, \partial T_0/\partial a_{0,r-1}(t_k)]'.$$

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

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

This proves part (iii) of the theorem.

It is easy to check that the minimum variance is C_0^{-1}. Thus we have obtained a lower limit to the asymptotic variance of a consistent estimate satisfying Assumption A. Hence the efficiency of the MSCF estimate is given by $(n_{00} C_0 \sigma^2(\alpha))^{-1}$ where $\sigma^2(\alpha)$ is defined by equation (4.6).

6. Remarks. The method of estimation used in this paper is based on the general technique of estimation by the method of moments. Since in the classical setting, the maximum likelihood method is generally preferable to the method of moments from the point of view of asymptotic properties of the estimates, it seems natural for anyone to ask whether one would not prefer the maximum likelihood approach in this problem.

The maximum likelihood estimate of the parameter α based on k observations $N(t_1), \cdots, N(t_k)$ can be obtained in a number of ways.

(a) We may obtain the maximum likelihood estimate of α by maximising the joint probability of $N(t_1), \cdots, N(t_k)$. As is shown in [6], the conditional distribution of $N(t_2)$ given $N(t_1)$ is rather formidable and the algebraic details are hopelessly complicated. This can be done for $k = 1$.

(b) It is shown in [6], Theorem 4.3, that under certain conditions the joint distribution of $N(t_1), \cdots, N(t_k)$ converges to a multivariate normal distribution whose mean and covariance matrix are rather complicated functions of the parameter α. In this case also the algebraic details are intractable.

(c) One can observe the points $t_1, t_2, \cdots$ at which a transition occurs, that is, the points at which a phage attaches itself to a bacterium and then can construct a density on the sample functions following a general approach due to Albert [1] and obtain the maximum likelihood estimate of the parameter α. Since the number of bacteria and phages in an experiment are of the order of 10^6 or 10^7, such a method will not be useful in application.

Thus we see that except for $k = 1$ the method of maximum likelihood estimation does not lead to any simple solution. Hence for simplicity, it is decided to apply the present technique for estimating the parameter α. Even in this case, the formula (2.3) for computing the MSCF estimate $\hat{\alpha}$ of α and also the formula (5.1) for calculating the (asymptotic) relative efficiency of the MSCF estimate must be programmed for any application. Numerical examples illustrating the application of this method and also the method of maximum likelihood for the case $k = 1$, together with the programs will be considered in a subsequent paper.

Acknowledgment. I wish to express my thanks to Dr. J. Gani for suggesting the problem and for many useful discussions and suggestions. My thanks are also due to Dr. Leo Katz under whose guidance this research was conducted. I am also indebted to Dr. J. F. Hannan for some useful suggestions and discussions that led to the present form of Theorem 4.1.

REFERENCES

[1] ALBERT, ARTHUR (1962). Estimating the infinitesimal generator of a continuous time finite state Markov process. *Ann. Math. Statist.* **33** 727–753.

[2] FERGUSON, T. (1958). A method of generating best asymptotically normal estimates with applications to the estimation of bacterial density. *Ann. Math. Statist.* **29** 1046–1062.

[3] GANI, J. (1965). Stochastic phage attachment to bacteria. *Biometrics* **21** 134–139.

[4] RAO, C. R. (1952). *Advanced Statistical Methods in Biometric Research.* Wiley, New York.

[5] RUBEN, HAROLD. (1963). The estimation of a fundamental interaction parameter in an emigration-immigration process. *Ann. Math. Statist.* **24** 238–259.

[6] SRIVASTAVA, R. C. (1967). Some aspects of the stochastic model for the attachment of phages to bacteria. *J. Appl. Prob.* **4** 9–18.

[7] TAYLOR, A. E. (1955). *Advanced Calculus.* Ginn, New York.