MAXIMUM LIKELIHOOD ESTIMATION OF PARAMETERS IN BIRTH AND DEATH PROCESSES BY POISSON SAMPLING

Yutaka Baba

Tokyo Institute of Technology

(Received May 24, 1980; Final October 5, 1981)

Abstract  Maximum likelihood estimates of parameters in continuous time Markov chains are obtained when the observation plan is Poisson sampling. Furthermore, for birth and death processes, variance-covariance matrix of parameters is obtained. Especially, for queueing model M/M/1, the variance-covariance matrix by Poisson sampling is compared with the variance-covariance matrix by complete observation in detail.

1. Introduction

Let \( \{x(t), t \geq 0\} \) be a continuous time Markov chain with transition intensity matrix \( Q = (q_{ij}(\theta)) \), where \( \theta \) is a m-dimensional column vector which represents an unknown parameter ranging over a set \( \Theta \) in m-dimensional Euclidean space. The problem which we shall consider is that of estimating \( \theta \) from observations on the process.

If \( x(t) \) is observed continuously over the time \([0,T]\), we say that this observation plan is complete observation.

In this paper, we shall consider a different observation plan named Poisson sampling. We shall observe the process at a random observation points \( \{v_0 < v_1 < v_2 < \ldots \} \), and count the number of transitions of states occurring in the random observation intervals \( (v_{k-1}, v_k] \) \((k = 1, 2, \ldots)\). Now, assume that \( \xi_k = v_k - v_{k-1} \) \((k = 1, 2, \ldots)\) are independent identically distributed exponential variates with parameter \( \nu \). This sampling is then called Poisson sampling derived from \( \{x(t), t \geq 0\} \). This method was first considered by Kingman [4]. But an application to a statistical problem using Poisson sampling is only seen in Basawa [1].

Here we shall consider the estimation of the parameters of a continuous time Markov chain based on Poisson sampling. We often encounter the case that
the between-event intervals are unobservable. Such situations are not uncom-
mon in practice, and occurs particularly in certain queueing and other
stochastic models. In such a case, Poisson sampling is a powerful method,
because Kingman [4] showed that it determines the original process completely.

In Section 2, we shall construct the likelihood function of a continuous
time Markov chain of general type when the observation is done by Poisson
sampling. And solving the maximum likelihood equations, we shall obtain the
maximum likelihood estimates of parameters. Furthermore we want to get the
variance-covariance matrix of parameters and compare two matrices in the
complete observation case and Poisson sampling case. However, it is very
difficult to obtain the variance-covariance matrix in a general type Markov
chain. So, in Section 3, we shall only consider the birth and death process
which is a very useful but simple structured model. For the birth and death
process, we have the variance-covariance matrices of parameters for two sim-
ple models given by Wolff [6]. But even for the birth and death process, the
variance-covariance matrix by Poisson sampling is not yet represented by
explicit forms. Hence, in Section 4, we shall deal with in detail a simple
but useful queueing model $M/M/1$ which is a special case of birth and death
processes. For this model, we can compare the variance-covariance matrices of
complete observation case over $[0,T]$ and Poisson sampling case of $w$ ob-
ervation points. And the determination of $w$ is done in order to obtain
smaller variance of parameters than the variance of complete observation
case.

2. Maximum Likelihood Estimation of Parameters in Continuous Time Markov
Chains by Poisson Sampling

We observe a continuous time Markov chain with transition intensity
matrix $Q = (q_{ij}(\theta))$ by Poisson sampling until the number of transitions of
states amounts to $n$. So the data obtained by this sampling are successive
states $(x_0, x_1, \ldots, x_n)$ and $l_i$ $(i = 1, \ldots, n)$ which are the number of ob-
servation points between $x_{i-1}$ and $x_i$.

Let $q_i(\theta) = \sum_{j \neq i} q_{ij}(\theta)$ and $\pi_{ij}(\theta) = q_{ij}(\theta)/q_i(\theta)$. Then, from the prop-
erties of a continuous time Markov chain, the following equations hold clearly:

\begin{equation}
Pr(l_{i+1} = k \mid x_i = j) = \int_0^\infty \frac{e^{-vt}(vt)^k}{k!} q_j(\theta) e^{-q_i(\theta)t} dt
\end{equation}
Estimation by Poisson Sampling

\[ E(\mathcal{Z}_{t+1} | x_i = j) = \frac{v}{q_j(\theta)} \]  

(2.2)

Thus, ignoring the distribution of initial state \( x_0 \), the likelihood function based on these data is

\[ L_n(\theta) = \prod_{i=0}^{n-1} \frac{q_{x_i}(\theta)v^{\mathcal{Z}_{i+1}}}{(v + q_{x_i}(\theta))^{\mathcal{Z}_{i+1}}} \]  

(2.3)

Therefore, from the following maximum likelihood equations,

\[ \frac{\partial \ln L_n(\theta)}{\partial \lambda_i} = 0 \quad (i = 1, \ldots, m) \]  

(2.4)

we can get the maximum likelihood estimates (MLE) of \( \theta \).

Especially, if the state space is finite, that is, \( S = \{1, \ldots, s\} \) and \( q_{ij}(\theta) = q_{ij} \), then the log-likelihood function with Lagrange multiplier becomes

\[ \ln L_n = \sum_{k=0}^{n-1} \ln q_{x_k} x_{k+1}^{\mathcal{Z}_{k+1}} + \sum_{k=0}^{n-1} \ln v^{\mathcal{Z}_{k+1}} \]

\[ - \sum_{k=0}^{n-1} (\mathcal{Z}_{k+1} + 1)\ln(v + q_{x_k}) + \sum_{i=1}^{s} \lambda_i (q_{i} - \sum_{j \neq i} q_{ij}) \]

(2.5)

where \( r_i = \sum_{k=0}^{n-1} \delta_{x_k, i}(\mathcal{Z}_{k+1} + 1) \), \( (\delta_{x, i} \) is the Kronecker delta) and \( n_{i,j} \) is the number of transitions such that \( x_k = i \) and \( x_{k+1} = j \).

Hence the maximum likelihood equations are

\[ \frac{\partial \ln L_n}{\partial q_{ij}} = \frac{n_{ij}}{q_{ij}} - \lambda_i = 0 \quad (i = 1, \ldots, s; j = 1, \ldots, s; j \neq i) \]

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.
Solving equations (2.6), we obtain

\[
\hat{q}_{i_j} = \frac{\gamma_{i_j}}{r_i - \sum_{j \neq i} \gamma_{i_j}} \quad (i = 1, \ldots, s; \ j = 1, \ldots, s; \ j \neq i).
\]

This MLE has a simple form. But the denominator \( r_i - \sum_{j \neq i} \gamma_{i_j} \) of (2.7) may be zero. For such \( i \), we cannot estimate \( q_{i_j} \).

3. Determination of Variance-Covariance Matrices for Birth and Death Processes by Poisson Sampling

In Section 2, we have obtained the MLE of \( q_{i_j}^*(\theta) \). It is, however, very difficult for a Markov chain in general type to obtain the variance-covariance matrix of \( q_{i_j}^*(\theta) \). So, in this Section, we shall consider only for ergodic birth and death processes.

We denote by \( \lambda_i^{(0)} \) and \( \mu_i^{(0)} \) \( (i = 0, 1, \ldots) \) the parameters of birth and death processes. In our notations defined in Section 2, \( \lambda_i^{(0)} = q_{i,i+1}^{(0)} \) and \( \mu_i^{(0)} = q_{i,i-1}^{(0)} \).

In constructing the likelihood function, it will be useful to consider the portion of it, \( L_i \), which represents a single transition \( x_i \rightarrow x_{i+1} \).

Given

\[
x_i = j \text{ and } z_{i+1} = k,
\]

it is easy to show the following expression

\[
\ln L_i(\theta) = \ln \lambda_j^{(0)} + k \ln \nu - (k + 1) \ln (\nu + \lambda_j^{(0)} + \mu_j^{(0)})
\]

(3.2)

\[
= \ln \mu_j^{(0)} + k \ln \nu - (k + 1) \ln (\nu + \lambda_j^{(0)} + \mu_j^{(0)})
\]

if \( x_{i+1} = j + 1 \).

Define

\[
G_u = (\partial / \partial \theta_u) \ln L_i(\theta) \quad (u = 1, \ldots, m)
\]

(3.3)

\[
G = (G_1, \ldots, G_m)
\]

\[
G_{uv} = (\partial^2 / \partial \theta_u \partial \theta_v) \ln L_i(\theta) \quad (u = 1, \ldots, m; \ v = 1, \ldots, m).
\]

Then, from Theorem 7.3 of Billingsley [2], \( \sqrt{n}(\hat{\theta} - \theta) \) is asymptotically normal (it converges in law to the normally distributed random variable) with mean 0 and variance-covariance matrix \( \sigma^{-1}(\theta) \). Besides under regularity conditions the following facts are well known

\[
E(G_u) = 0, \quad E(G_{uv}) = -\sigma_{uv}.
\]

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.
From (2.4), the variance-covariance matrix of $G_u$, $E(GG^T)$, is

\begin{equation}
E(GG^T) = \sigma(\theta) = (\sigma_{uv}) = (-E(G_{uv})).
\end{equation}

We shall calculate the asymptotic variance-covariance matrix for two simple models given by Wolff [6].

Model 1:

- $\lambda_j > 0$, $j = 0, 1, \ldots, M-1$, $\lambda_M = 0$
- $\mu_j > 0$, $j = 1, 2, \ldots, M$, $\mu_0 = 0$.

There are $2M$ unknown parameters $\theta^T = (\lambda_0, \ldots, \lambda_{M-1}, \mu_1, \ldots, \mu_M)$, which are required only to be finite and positive. $\lambda_j$ is the $j + 1$ st component and $\mu_j$ is the $M + j$ th component of $\theta$. From equation (3.2), under the condition (3.1), we have

\begin{align}
G_{j+1} &= \frac{1}{\lambda_j} - \frac{k + 1}{\nu + \lambda_j + \mu_j}, \quad G_{M+j} = -\frac{k + 1}{\nu + \lambda_j + \mu_j} \quad \text{for } x_{i+1} = j + 1 \\
G_{j+1} &= -\frac{k + 1}{\nu + \lambda_j + \mu_j}, \quad G_{M+j} = \frac{1}{\mu_j} - \frac{k + 1}{\nu + \lambda_j + \mu_j} \quad \text{for } x_{i+1} = j - 1
\end{align}

otherwise $G_{uv} = 0$.

We continue to condition on (3.1)

\begin{align}
G_{j+1,j+1} &= -\frac{1}{\lambda_j^2} + \frac{k + 1}{(\nu + \lambda_j + \mu_j)^2}, \\
G_{M+j,M+j} &= \frac{k + 1}{(\nu + \lambda_j + \mu_j)^2}, \quad G_{j+1,M+j} = \frac{k + 1}{(\nu + \lambda_j + \mu_j)^2} \quad \text{for } x_{i+1} = j + 1 \\
G_{j+1,j+1} &= \frac{k + 1}{(\nu + \lambda_j + \mu_j)^2}, \quad G_{j+1,M+j} = \frac{k + 1}{(\nu + \lambda_j + \mu_j)^2} \\
G_{M+j,M+j} &= -\frac{1}{\mu_j^2} + \frac{k + 1}{(\nu + \lambda_j + \mu_j)^2} \quad \text{for } x_{i+1} = j - 1 \\
\text{otherwise } G_{uv} &= 0.
\end{align}

Thus we have

\begin{align}
-E(G_{j+1,j+1} | x_i = j, \ell_{i+1} = k) &= \frac{\lambda_j}{\lambda_j + \mu_j} \left(\frac{1}{\lambda_j} - \frac{k + 1}{(\nu + \lambda_j + \mu_j)^2}\right) - \frac{\mu_j}{\lambda_j + \mu_j} \left(\frac{k + 1}{(\nu + \lambda_j + \mu_j)^2}\right) \\
&= \frac{1}{\lambda_j(\lambda_j + \mu_j)} - \frac{k + 1}{(\nu + \lambda_j + \mu_j)^2}.
\end{align}

And similarly

\begin{align}
-E(G_{M+j,M+j} | x_i = j, \ell_{i+1} = k) &= \frac{1}{\mu_j(\lambda_j + \mu_j)} - \frac{k + 1}{(\nu + \lambda_j + \mu_j)^2}
\end{align}
From equations (3.8), (3.9), and (2.2), we have

$$-E(G'_{j+1}, M_{j+1} \mid x_i = j, l_{i+1} = k) = \frac{k + 1}{(\nu + \lambda_j + u_j)}.$$  

(3.10)

Here we define some notations. Let $P_j (j = 0, 1, \ldots, M)$ be the stationary distribution and $y_j (j = 0, 1, \ldots, M)$ the probability that a randomly selected transition was out of state $j$. Then, using the fact by Wolff [6], it is easily shown that

$$(3.11) \quad y_j = (\lambda_j + u_j)P_j / 2R \quad (j = 0, 1, \ldots, M)$$

where $R = \sum_{j=0}^{M-1} \lambda_j P_j = \sum_{j=0}^{M} u_j P_j$ (R is called the average throughput).

Hence, from (3.10), we have

$$(3.12) \quad V(G_{j+1}) = E[-E(G'_{j+1}, M_{j+1} \mid x_i = k)] = \sum_{k=0}^{M} E(G'_{j+1}, M_{j+1} \mid x_i = k) y_k$$

$$V(G_{M+j}) = \frac{y_j (\nu + u_j)}{\lambda_j (\lambda_j + u_j) (\nu + \lambda_j + u_j)}$$

$$(\lambda_j + u_j)(\nu + \lambda_j + u_j)$$

otherwise $\text{Cov}(G_{M+j}) = 0$.

Hence, using equations (3.11) and (3.12), we have

$$V(G_{j+1}) = \frac{P_j (\nu + u_j)}{2R (\nu + \lambda_j + u_j)} \quad (j = 0, 1, \ldots, M-1)$$

$$V(G_{M+j}) = \frac{P_j (\nu + \lambda_j)}{2R (\nu + \lambda_j + u_j)} \quad (j = 1, 2, \ldots, M)$$

$$(3.13) \quad \text{Cov}(G_{j+1}, M_{j+1}) = -\frac{P_j (\nu + u_j)}{2R (\nu + \lambda_j + u_j)} \quad (j = 1, 2, \ldots, M-1)$$

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.
otherwise \( \text{Cov}(G_u, G_v) = 0 \).

Model 2:

\[
\lambda_j = \lambda h(j), \quad \mu_j = \mu g(j), \quad j = 0, 1, \ldots, \quad \text{and} \quad g(0) = 0
\]

where \( \theta^T = (\lambda, \mu) \), and \( h(j), g(j) \) are assumed to be arbitrary but known function of \( j \). Moreover, for ergodicity, we assume that

\[
\sum_{k=0}^{\infty} \frac{h(k)}{g(k+1)} < \infty.
\]

Given (3.1), we have from (3.2)

\[
G_1 = \frac{1}{\lambda} - \frac{(k+1)h(j)}{\nu + \lambda h(j) + \mu g(j)}, \quad G_2 = -\frac{(k+1)g(j)}{\nu + \lambda h(j) + \mu g(j)}
\]

if \( x_{i+1} = j + 1 \)

\[
G_1 = -\frac{(k+1)h(j)}{\nu + \lambda h(j) + \mu g(j)} \quad G_2 = \frac{1}{\mu} - \frac{(k+1)g(j)}{\nu + \lambda h(j) + \mu g(j)}
\]

if \( x_{i+1} = j - 1 \)

\[
G_{11} = -\frac{1}{\lambda^2} + \frac{(k+1)\{h(j)\}^2}{(\nu + \lambda h(j) + \mu g(j))^2}
\]

\[
G_{22} = \frac{(k+1)\{g(j)\}^2}{(\nu + \lambda h(j) + \mu g(j))^2}
\]

\[
G_{12} = \frac{(k+1)h(j)g(j)}{(\nu + \lambda h(j) + \mu g(j))^2}
\]

if \( x_{i+1} = j + 1 \)

\[
G_{11} = \frac{(k+1)\{h(j)\}^2}{(\nu + \lambda h(j) + \mu g(j))^2}
\]

\[
G_{22} = \frac{1}{\mu^2} + \frac{(k+1)\{g(j)\}^2}{(\nu + \lambda h(j) + \mu g(j))^2}
\]

\[
G_{12} = \frac{(k+1)h(j)g(j)}{(\nu + \lambda h(j) + \mu g(j))^2}
\]

if \( x_{i+1} = j - 1 \).

By similar calculations as Model 1, we have

\[
V(G_1) = \frac{1}{2\lambda^2} - \sum_{j=0}^{\infty} \frac{\{h(j)\}^2P_j}{2R(\nu + \lambda h(j) + \mu g(j))}
\]

\[
V(G_2) = \frac{1}{2\mu^2} - \sum_{j=0}^{\infty} \frac{\{g(j)\}^2P_j}{2R(\nu + \lambda h(j) + \mu g(j))}
\]

(3.15)

\[
\text{Cov}(G_1, G_2) = -\sum_{j=0}^{\infty} \frac{h(j)g(j)P_j}{2R(\nu + \lambda h(j) + \mu g(j))}.
\]

In both Model 1 and Model 2, we have obtained the asymptotic variance-covariance matrices of parameters. But, in general, it is difficult to get the explicit forms of \( P_j \) and \( R \). So, \( \sigma(\theta) \) or its inverse \( \sigma^{-1}(\theta) \) is not
represented here in explicit form. In next Section, we shall deal with in
detail a simple but useful model $M/M/1$ which is a special case of Model 2.

4. Further Investigation of Queueing Model $M/M/1$

Let us define

\[ \lambda_0 = \lambda \quad \text{for } i = 0,1, \ldots \]
\[ \mu_i = \mu_i = 0 \quad \text{for } i = 1,2, \ldots \]
\[ \mu_i = 0 \quad \text{for } i = 0 \]

and for ergodicity assume that $\lambda < \mu$. So that, in our notation in Section 3,

\[ h(i) = 1 \quad \text{for } i = 0,1, \ldots \]
\[ g(i) = 1 \quad \text{for } i = 1,2, \ldots \]
\[ h(i) = 0 \quad \text{for } i = 0. \]

On this model, we shall consider the variance-covariance matrix of parameters both when the observation is complete over $[0,T]$, and when the observation is made by Poisson sampling of $\omega$ observation points.

In Section 3, we have fixed the number of transitions of states. But in this Section, we fix the number of observation points. So the number of transitions of states, $n$, becomes a random variable and depends on $\omega$. We denote it as $n(\omega)$.

We shall also consider the determination of $\omega$ so that we get smaller variance than that in the case of complete observation over the interval $[0,T]$.

When the data were obtained by complete observation over $[0,T]$, we define the following notations.

\[ \gamma_i = \text{total sojourn time in state } i \text{ over } [0,T] \]
\[ a_i = \text{the number of transitions } i \rightarrow i+1 \text{ over } [0,T] \]
\[ b_i = \text{the number of transitions } i \rightarrow i-1 \text{ over } [0,T] \]
\[ a = \sum_{i=0}^{\infty} a_i, \quad b = \sum_{i=0}^{\infty} b_i. \]

Then, by using the notations (4.3), it is easily derived that the maximum likelihood estimates of $\lambda$ and $\mu$ are

\[ \hat{\lambda} = \frac{a}{T}, \quad \hat{\mu} = \frac{b}{T - \gamma_0}. \]
Furthermore, from Billingsley [2] and Reynolds [5], it is easily shown that, for large $T$,

$$
\lambda \sim N(\lambda, \frac{\lambda}{T}), \quad \mu \sim N(\mu, \frac{\mu}{\lambda T})
$$

Next we consider the Poisson sampling case. We construct the likelihood function by Poisson sampling when the observation intervals are independently and identically distributed exponential variates with parameter $\nu$, and the number of observation points is $\omega$. We define some notations similarly to (4.3).

$$
u = \text{the number of transitions $i + i + 1$ until $\omega$-th Poisson observation point}
$$

$$
d = \text{the number of transitions $i + i - 1$ until $\omega$-th Poisson observation point}
$$

Then, similarly to (2.5) in Section 2, let

$$
r_i = \sum_{k=0}^{n(\omega) - 1} \delta_{k, i} \delta_{k+1, i} \quad \text{and} \quad r_i = \sum_{i=0}^{\infty} n_i.
$$

Hence, analogously to (2.3), ignoring the initial distribution and end effect, that is, the information contained in the process from the last transition point to the last observation point, it is easily shown that the likelihood function based on these data is

$$
L = \frac{\lambda u^d}{(\nu + \lambda)^{n_0} (\nu + \lambda + \mu)^{r - n_0}} \cdot K
$$

where $K$ is a constant free from $\lambda$ and $\mu$. Taking the natural logarithm of both sides, we have

$$
\ln L = u \ln \lambda + d \ln \nu + \ln K - r_0 \ln (\nu + \lambda) - (r - r_0) \ln (\nu + \lambda + \mu).
$$

Hence, from the following maximum likelihood equations,

$$
\frac{\partial \ln L}{\partial \lambda} = \frac{u}{\lambda} - \frac{r_0}{\nu + \lambda} - \frac{r - r_0}{\nu + \lambda + \mu} = 0
$$

$$
\frac{\partial \ln L}{\partial \mu} = \frac{d}{\mu} - \frac{r - r_0}{\nu + \lambda + \mu} = 0
$$

we obtain the maximum likelihood estimates.
(4.11) \[ \lambda = \frac{uv}{r - u - d} , \quad \beta = \frac{d(r - d)v}{(r - r_0 - d)(r - u - d)} . \]

Besides, from equations (4.10), we have

\[
\frac{\partial^2 \ln L}{\partial \lambda^2} = -\frac{\mu}{\lambda^2} + \frac{r_0}{(\nu + \lambda)} + \frac{r - r_0}{(\nu + \lambda + \mu)^2} .
\]

(4.12) \[
\frac{\partial^2 \ln L}{\partial \mu^2} = -\frac{d}{\mu^2} + \frac{r - r_0}{(\nu + \lambda + \mu)^2} .
\]

\[
\frac{\partial^2 \ln L}{\partial \lambda \partial \mu} = \frac{r - r_0}{(\nu + \lambda + \mu)^2} .
\]

When the number of observation points is \( w \), actual \( w \)-th observation time, \( T_w' \), obeys \( w \)-th Erlang distribution with parameter \( \nu \). Therefore the probability density function of \( T_w' \), \( f(t) \), is

(4.13) \[ f(t) = \frac{w t^{w-1}}{\Gamma(w)} \exp(-\nu t) . \]

Since it is well known that both \( u \) and \( d \) obey Poisson distribution with parameter \( \lambda T_w' \), \( E(u \mid T_w') = E(d \mid T_w') = \lambda T_w' \). Hence, we have

(4.14) \[ E(u) = E(d) = \int_0^\infty t \nu f(t) dt = \frac{\lambda \nu^w}{\Gamma(w)} \int_0^\infty t^{w-1} \exp(-\nu t) dt = \frac{\lambda \nu^w}{\nu} . \]

Next let \( z \) be the number of observation points from the last transition point to the last observation point. Then it is easily shown that

(4.15) \[ r = u + d + w - z . \]

Taking expectations of both sides, we have

(4.16) \[ E(r) = E(u) + E(d) + w - E(z) = \frac{2\lambda \nu}{\nu} + w - E(z) . \]

Since, for large \( w \), it is easily verified that \( E(z) \) is negligible, we have

(4.17) \[ E(r) \approx \frac{2\lambda \nu}{\nu} + w . \]

For \( M/M/1 \), stationary distribution \( P_j \) and average throughput \( R \) defined in Section 3 have simple forms and are represented by

(4.18) \[ P_j = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^j \quad (j = 0, 1, \ldots) \quad R = \lambda . \]

And from (3.11), we have

(4.19) \[ \gamma_0 = \frac{1}{2}(1 - \frac{\lambda}{\mu}) . \]

Hence we have

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.
Estimation by Poisson Sampling

(4.20) \[ E(u_0) = E(u + d)y_0 = \frac{2w}{\nu} \left( 1 - \frac{\lambda}{\mu} \right) = \frac{w\lambda(u - \lambda)}{\mu
u}. \]

Furthermore by taking the expectation of both sides in (4.7), we have

(4.21) \[ E(v_0) = E(u_0) \cdot \frac{v + \lambda}{\lambda} = \frac{w(\mu - \lambda)(v + \lambda)}{\mu
u}. \]

where \( \frac{v + \lambda}{\lambda} \) is the expectation of the number of observation points + 1 when a transition is from 0 to 1. From equations (4.12), (4.14), (4.17), (4.20) and (4.21) we obtain

\[ -E \left( \frac{\partial^2 \ln L}{\partial \lambda^2} \right) \approx \frac{w(\lambda^2 + \nu^2 + \lambda\nu + \mu\nu)}{\lambda \nu(v + \lambda + \mu)}. \]

\[ -E \left( \frac{\partial^2 \ln L}{\partial \mu^2} \right) \approx \frac{w\lambda(v + \lambda)}{\nu^2 \nu(v + \lambda + \mu)}. \]

\[ -E \left( \frac{\partial^2 \ln L}{\partial \lambda \partial \mu} \right) \approx -\frac{w\lambda}{\nu \nu(v + \lambda + \mu)}. \]

Therefore asymptotic variance-covariance matrix of \( \lambda \) and \( \mu \) are

\[
\begin{pmatrix}
-E \left( \frac{\partial^2 \ln L}{\partial \lambda^2} \right) & -E \left( \frac{\partial^2 \ln L}{\partial \lambda \partial \mu} \right) \\
-E \left( \frac{\partial^2 \ln L}{\partial \lambda \partial \mu} \right) & -E \left( \frac{\partial^2 \ln L}{\partial \mu^2} \right)
\end{pmatrix}
\approx
\begin{pmatrix}
\frac{\lambda(v + \lambda)}{\nu} & \frac{\lambda}{\nu} \\
\frac{\lambda}{\nu} & \frac{w^2(\lambda^2 + \nu^2 + \lambda\nu + \mu\nu)}{w\lambda(v + \lambda)}
\end{pmatrix}
\]

Comparing with the complete observation case, in order to obtain smaller variances of \( \lambda \) and \( \mu \) than those of complete observation over \([0,T]\), the number of observation points by Poisson sampling, \( \omega \), will be needed as follows by using (4.5) and (4.23)

\[ \frac{\lambda(v + \lambda)}{\nu} \leq \frac{\lambda}{T}, \quad \frac{w^2(\lambda^2 + \nu^2 + \lambda\nu + \mu\nu)}{w\lambda(v + \lambda)} = \frac{\mu^2}{\lambda T}. \]

Solving the inequalities (4.24), we obtain

\[ \omega \geq \frac{(v + \lambda + \frac{w(\mu - \lambda)}{v + \lambda})}{T}. \]

Now, we consider the problem which is better a complete observation or a Poisson sampling when we need the observation cost. Let \( C_0 \), \( C_1 \) be the observation cost per unit time by complete observation and the observation cost per an observation by Poisson sampling, respectively. Then the cost of observing the process over \([0,T]\) by complete observation is given by \( C_0 T \) and the cost of observing the process until the \( \omega \)-th observation point by Poisson sampling is given by \( C_1 \omega \). From the inequality (4.25), if

Copyright © by ORSJ. Unauthorized reproduction of this article is prohibited.
then the cost for Poisson sampling is less than that for complete observation, in order to get the same variances. In such a case Poisson sampling is better than the complete observation. In Fig. 1, we showed the maximum value of \( C_1/C_0 \) in the case where the cost for Poisson sampling with parameter \( \nu \) is less than that for complete observation, when \( \mu = 1 \) and \( \rho = \lambda/\mu \) is 0.5, 0.7 and 0.9.

**Fig. 1**

**Acknowledgement**

I wish to thank Professor Hidenori Morimura for his invaluable suggestions and helpful comments.
References


Yutaka BABA: Department of Information Sciences, Tokyo Institute of Technology
Oh-okayama, Meguro-ku, Tokyo 152, Japan.