LARGE SAMPLE INFERENCE FROM SINGLE SERVER QUEUES *

I.V. BASAWA

Department of Statistics, University of Georgia, Athens, GA 30602, U.S.A.

and

N.U. PRABHU

Mathematical Sciences Institute, Cornell University, Caldwell Hall, Ithaca, NY 14853, U.S.A.

Received: 16 April 1987
Revised: 17 March 1988

Abstract

Problems of large sample estimation and tests for the parameters in a single server queue are discussed. The service time and the interarrival time densities are assumed to belong to (positive) exponential families. The queueing system is observed over a continuous time interval $(0, T]$ where T is determined by a suitable stopping rule. The limit distributions of the estimates are obtained in a unified setting, and without imposing the ergodicity condition on the queue length process. Generalized linear models, in particular, log-linear models are considered when several independent queues are observed. The mean service times and the mean interarrival times after appropriate transformations are assumed to satisfy a linear model involving unknown parameters of interest, and known covariates. These models enhance the scope and the usefulness of the standard queueing systems.

Keywords: Single server queues, maximum likelihood, stopping times, exponential families, generalized linear models, tests of fit, asymptotic inference.

1. Introduction

In a previous paper by Basawa and Prabhu [1] moment and maximum likelihood estimates were obtained for single server queueing systems observed until the epoch of the dth departure. The consistency and the asymptotic normality of the estimates were established without imposing any ergodicity requirements on the queue-length process. In sections 2 and 3 of this paper we present a unified framework in which the system is observed over a time interval $(0, T]$ where T is a suitable stopping time. Four different stopping rules are considered. If a random norming is used it is shown that the limit distribution does not depend on the particular stopping time used. On the other hand, the use

* Partially supported by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University.

of a non-random norming leads to a limit distribution whose covariance matrix does depend on the stopping rule. This is illustrated for the M/M/1 queue in section 4. We assume that the interarrival time and the service time densities belong to the class of (non-negative) exponential families. This class includes exponential, the Erlangian, the beta, and the inverse Gaussian among others.

In section 5 we consider a generalized linear model using several independent (but not necessarily identical) queueing systems. It is assumed that the mean interarrival times and the mean service times, after appropriate transformations, satisfy a linear model involving unknown parameters and known covariates. This model differs from the usual generalized linear model (see McCullagh and Nelder [7]) in that (a) we have a random number of observations from each system, (b) observations are over a continuous time interval, and (c) asymptotics is discussed as the time of observation becomes large (rather than as the number of independent realizations increases). The generalized linear model presented here enhances the scope and usefulness of the standard queueing models. Finally, in section 6 large sample tests for testing goodness of fit and for homogeneity are considered briefly.

Notation: We shall use $\Rightarrow$ to denote convergence in distribution. Convergence in probability is denoted as $\xrightarrow{p}$ or p lim. The random vector having the k-variate normal density with mean vector μ and covariance matrix Σ is denoted by $N_k\{\mu, \Sigma\}$.

References to early work on inference from queueing systems can be found in Basawa and Prabhu [1]. More recent references are given in a comprehensive survey by Bhat and Rao [2]. For Bayesian approach to inference from queues see McGrath, Gross and Singpurwalla [8], and McGrath and Singpurwalla [9].

2. The likelihood function and the stopping rules

Consider a GI/G/1 queueing system in which the interarrival times $\{u_k, k \geqslant 1\}$ and the service times $\{v_k, k \geqslant 1\}$ are two independent sequences of independent and identically distributed non-negative random variables with densities $f(u; \theta)$ and $g(v; \phi)$ respectively, where θ and ϕ are unknown parameters. We assume that f and g belong to the continuous exponential families given by

$$f(u; \theta) = a_1(u)\, e^{\{\theta h_1(u) - k_1(\theta)\}}, \tag{1}$$

$$g(v; \phi) = a_2(v)\, e^{\{\phi h_2(v) - k_2(\phi)\}}. \tag{2}$$

It is further assumed that the densities in (1) and (2) are equal to zero on $(-\infty, 0)$.

For simplicity we assume that the initial customer arrives at time $t = 0$. Our sampling scheme is to observe the system over a continuous time interval $(0, T]$ where T is a suitable stopping time. The sample data consist of

$$\{A(T), D(T), u_1, u_2, \ldots, u_{A(T)}, v_1, v_2, \ldots, v_{D(T)}\} \tag{3}$$

where $A(T)$ is the number of arrivals, and $D(T)$ the number of departures (service completions) during $(0, T]$. Note that no arrivals occur during

$$\left(\sum_1^{A(T)} u_i,\ T\right] \text{ and no departures during } \left(\gamma(T) + \sum_1^{D(T)} v_k,\ T\right],$$

where $\gamma(T)$ is the total idle period during $(0, T]$.

The data in (3) can be represented equivalently in terms of the queue-length process $\{Q(t),\ t \geqslant 0\}$ observed over $(0, T]$, which in turn is determined by the total number of transitions (arrivals or departures) during $(0, T]$, the intervals of time between successive transitions, and the queue length at the epochs of the transitions.

Some possible stopping rules to determine T are given below:

Rule 1. Observe the system until a fixed time t. Here $T = t$ with probability one, and $A(t)$ and $D(t)$ are both random variables.

Rule 2. Observe the system until d departures have occurred. Here, $T = \gamma(T) + v_1 + v_2 + \ldots + v_d$, $D(T) = d$, and $A(T)$ is a random variable. This rule was used previously by Basawa and Prabhu [1].

Rule 3. Observe the system until m arrivals take place. Thus, $T = u_1 + u_2 + \ldots u_m$, $A(T) = m$, and T and $D(T)$ are random variables.

Rule 4. Stop at the nth transition epoch. With this rule, T, $A(T)$, and $D(T)$ are all random variables, and $A(T) + D(T) = n$.

In the general case we observe the system over a time interval $(0, T)$, where $T \equiv T_c$ is a stopping time depending upon a parameter $c (0 < c < \infty)$ in such a way that as $c \to \infty$, $T_c \to \infty$ a.s. For example, suppose that the cost of observation over $(0, t]$ is given by $C(t)$, where $C(t)$ is calculated from the data (3) with T replaced by t, and is assumed to be nondecreasing function of t. Clearly $\{C(t),\ t \geqslant 0\}$ is a stochastic process. For the stopping time $T \equiv T_c$ we choose the one determined by

$$T = \inf\{t > 0:\ C(t) > c\}. \tag{4}$$

For uniformity we write $T \to \infty$ a.s. in the limit relations derived in the paper, it being understood that $T \equiv T_c \to \infty$ a.s. as $c \to \infty$.

The likelihood function based on the data (3) is given by

$$\begin{aligned} L_T(\theta, \phi) = {} & \prod_{k=1}^{A(T)} f(u_k;\ \theta) \prod_{k=1}^{D(T)} g(v_k;\ \phi) \\ & \times \left[1 - F_\theta\left(T - \sum_1^{A(T)} u_k\right)\right]\left[1 - G_\phi\left(T - \gamma(T) - \sum_1^{D(T)} v_k\right)\right], \end{aligned} \tag{5}$$

where F and G are the distribution functions corresponding to the densities f and g respectively. The likelihood in (5) remains valid under all the stopping rule described above.

We define the approximate likelihood $L_T^a(\theta, \phi)$ as the expression in (5) omitting the last two factors, i.e.,

$$L_T^a(\theta, \phi) = \prod_{k=1}^{A(T)} f(u_k;\, \theta) \prod_{k=1}^{D(T)} g(v_k;\, \phi). \tag{6}$$

The maximum likelihood estimates obtained from (6) are asymptotically equivalent to those obtained from (5) provided the following two conditions are satisfied as $T \to \infty$:

$$\{A(T)\}^{-1/2} \frac{\partial}{\partial\theta} \ln\left[1 - F_\theta\left(T - \sum_1^{A(T)} u_k\right)\right] \xrightarrow{P} 0 \tag{7}$$

$$\{D(T)\}^{-1/2} \frac{\partial}{\partial\phi} \ln\left[1 - G_\phi\left(T - \gamma(T) - \sum_1^{D(T)} v_k\right)\right] \xrightarrow{P} 0. \tag{8}$$

In order to understand the implications of these conditions note that

$$T - \sum_1^{A(T)} u_k = u'_{A(T)+1}, \quad T - \gamma(T) - \sum_1^{D(T)} v_k = v'_{D(T)+1}$$

where $u'_{A(T)+1}$ is the last (partially observed) interarrival time, and $v'_{D(T)+1}$ is the residual service time of the customer (if any) still being served at time T. If $T = t$ with probability one (stopping rule 1) it is known from renewal theory that $u'_{A(T)+1}$ has a limit distribution as $t \to \infty$. Since

$$\frac{\partial}{\partial\theta} \ln\left[1 - F_\theta(u'_{A(t)+1})\right]$$

is a continuous function of $u'_{A(t)+1}$ it follows that the former also has a limit distribution as $t \to \infty$. Consequently, the condition (7) is satisfied under rule 1. A similar argument can be used to verify (8). For rule 2, Basawa and Prabhu [1] have verified (7) for M/G/1 and E_k/G/1; (8) is satisfied trivially since the last factor on the right side of (5) is unity. Similarly (7) and (8) are satisfied under rule 3 for the system GI/M/1 and GI/E_k/1. We have not verified these conditions for rule 4, but conjecture that they are satisfied at least for some special cases.

It is seen from the above discussion that the class of stopping rules satisfying (7) and (8) is nonempty, and that it indeed contains important special cases. Accordingly, we shall use (6) as a basis of inference and study the asymptotic properties of the estimates so obtained.

In order to derive the limit distributions of our estimates we shall impose the following stability conditions on our stopping times:

$$\frac{A(T)}{EA(T)} \xrightarrow{P} 1, \quad \frac{D(T)}{ED(T)} \xrightarrow{P} 1 \quad \text{as } T \to \infty \text{ a.s.} \tag{9}$$

In practice these conditions can be replaced by equivalent, but simpler conditions, as for example, for the system M/M/1 (section 4).

3. Approximate maximum likelihood estimates

We now use the fact that the interarrival time density $f(u;\ \theta)$, and the service time density $g(v,\ \phi)$ belong to exponential families given by (1) and (2). It is easily verified that the moment generating function of the random variable $h_1(u)$ is given by $M(t)=\exp[k_1(t+\theta)-k_1(\theta)]$. Consequently

$$\eta_1(\theta)=E_\theta[h_1(u)]=k_1'(\theta),\quad \sigma_1^2(\theta)=\mathrm{Var}_\theta[h_1(u)]=k_1''(\theta). \tag{10}$$

Similarly,

$$\eta_2(\phi)=E_\phi[h_2(v)]=k_2'(\phi),\quad \sigma_2^2(\phi)=\mathrm{Var}_\phi[h_2(v)]=k_2''(\phi). \tag{11}$$

The approximate likelihood function (6) reduces to

$$L_T^{\mathrm{a}}(\theta,\ \phi)=a_1(u)a_2(v)\exp\left\{\sum_{i=1}^{A(T)}[h_1(u_i)-k_1(\theta)] + \sum_{i=1}^{D(T)}[\phi h_2(v_i)-k_2(\phi)]\right\}. \tag{12}$$

From (12) the likelihood equations are found to be

$$\frac{\partial}{\partial\theta}\ln L_T^{\mathrm{a}}=\sum_1^{A(T)}h_1(u_i)-A(T)k_1'(\theta)=0,$$

and

$$\frac{\partial}{\partial\phi}\ln L_T^{\mathrm{a}}=\sum_1^{D(T)}h_2(v_i)-D(T)k_2'(\phi)=0.$$

From now on we shall write L for L_T^{a}. The estimating equations reduce to the moment estimation equations, namely,

$$\frac{1}{A(T)}\sum_1^{A(T)}h_1(u_i)=\eta_1(\theta),\quad \frac{1}{D(T)}\sum_1^{D(T)}h_2(v_i)=\eta_2(\phi). \tag{13}$$

The solutions for θ and ϕ from (13) are given by

$$\hat\theta=\eta_1^{-1}\left[(A(T))^{-1}\sum_1^{A(T)}h_1(u_i)\right],\quad \hat\phi=\eta_2^{-1}\left[(D(T))^{-1}\sum_1^{D(T)}h_2(v_i)\right], \tag{14}$$

where $\eta_i^{-1}(\cdot)$ denote the inverse functions of $\eta_i(\cdot)$ $(i=1,2,)$. The existence of these inverses follows from properties of the exponential family (viz., $\eta_i'=k_i''=\sigma_i^2>0$). The sample Fisher information is given by

$$F(\theta,\ \phi)=-\begin{bmatrix}\dfrac{\partial^2\ln L}{\partial\theta^2} & \dfrac{\partial^2\ln L}{\partial\theta\partial\phi}\\[2ex] \dfrac{\partial^2\ln L}{\partial\phi\partial\theta} & \dfrac{\partial^2\ln L}{\partial\phi^2}\end{bmatrix}=\begin{bmatrix}A(T)\sigma_1^2 & 0\\ 0 & D(T)\sigma_2^2\end{bmatrix}, \tag{15}$$

and hence the Fisher information matrix is

$$I(\theta, \phi) = EF(\theta, \phi) = \begin{bmatrix} \sigma_1^2 EA(T) & 0 \\ 0 & \sigma_2^2 ED(T) \end{bmatrix}. \tag{16}$$

We can now state the following:

PROPOSITION 1

If the stopping time T satisfies the following stability conditions (9), then

$$F^{-1/2}(\theta, \phi) \begin{pmatrix} \dfrac{\partial}{\partial\theta} \ln L \\ \dfrac{\partial}{\partial\phi} \ln L \end{pmatrix} \Rightarrow N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \quad \text{as } T \to \infty \text{ a.s.} \tag{17}$$

Proof

We have

$$F^{-1/2}(\theta, \phi) \begin{pmatrix} \dfrac{\partial}{\partial\theta} \ln L \\ \dfrac{\partial}{\partial\phi} \ln L \end{pmatrix} = \begin{pmatrix} \left(A(T)\sigma_1^2(\theta)\right)^{-1/2} \sum_1^{A(T)} \{h_1(u_i) - \eta_1(\theta)\} \\ \left(D(T)\sigma_2^2(\phi)\right)^{-1/2} \sum_1^{D(T)} \{h_2(v_i) - \eta_2(\phi)\} \end{pmatrix}.$$

The desired result follows from the random sum central limit theorem (see Billingsley [4]) and the Cramer-Wold argument, in view of the assumptions (9). □

COROLLARY 1

Under the stability conditions (9),

$$I^{-\frac{1}{2}}(\theta, \phi) \begin{pmatrix} \dfrac{\partial \ln L}{\partial\theta} \\ \dfrac{\partial \ln L}{\partial\phi} \end{pmatrix} \Rightarrow N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \quad \text{as } T \to \infty \text{ a.s.} \tag{18}$$

Proof

The result follows from proposition 1 and eq. (16). □

The limit properties of the maximum likelihood estimates $\hat{\theta}$ and $\hat{\phi}$ are given by the following two theorems.

THEOREM 1 (consistency)

We have

$$\hat{\theta} \xrightarrow{P} \theta, \quad \hat{\phi} \xrightarrow{P} \phi \quad \text{as } T \to \infty \text{ a.s.} \tag{19}$$

Proof

We can write (14) as

$$\hat{\theta}=\alpha_1(\bar{h}_1), \quad \hat{\phi}=\alpha_2(\bar{h}_2) \tag{20}$$

where $\alpha_1=\eta_1^{-1}$, $\alpha_2=\eta_2^{-1}$ and $\bar{h}_1$, $\bar{h}_2$ denote the sample averages with random number of observations of $\{h_1(u_i),\ i=1,2,\ldots,A(T)\}$, and $\{h_2(v_i),\ i=1,2,\ldots,D(T)\}$. By the law of large numbers we find that

$$\bar{h}_1 \xrightarrow{p} E_\theta(h_1(u)), \quad \bar{h}_2 \xrightarrow{p} E_\phi(h_2(v)).$$

The desired results follow since α_1, α_2 are continuous functions of θ, ϕ. □

THEOREM 2

Under the stability conditions (9), we have

$$F^{1/2}(\theta,\phi)\begin{pmatrix}\hat{\theta}-\theta\\ \hat{\phi}-\phi\end{pmatrix} \Rightarrow N_2\left[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1&0\\0&1\end{pmatrix}\right] \quad \text{as } T\to\infty \text{ a.s.} \tag{22}$$

Proof

From (14) and (15) we have

$$F^{1/2}(\theta,\phi)\begin{pmatrix}\hat{\theta}-\theta\\ \hat{\phi}-\phi\end{pmatrix}=\begin{pmatrix}\left(A(T)\sigma_1^2(\theta)\right)^{1/2}\left[\alpha_1(\bar{h}_1)-\alpha_1(\eta_1(\theta))\right]\\ \left(D(T)\sigma_2^2(\phi)\right)^{1/2}\left[\alpha_2(\bar{h}_2)-\alpha_2(\eta_2(\phi))\right]\end{pmatrix}. \tag{23}$$

From the random sum central limit theorem and the Cramer-Wold device we obtain

$$\begin{pmatrix}\left(A(T)\sigma_1^{-2}(\theta)\right)^{1/2}\left(\bar{h}_1-\eta_1(\theta)\right)\\ \left(D(T)\sigma_2^{-2}(\phi)\right)^{1/2}\left(\bar{h}_2-\eta_2(\phi)\right)\end{pmatrix} \Rightarrow N_2\left[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1&0\\0&1\end{pmatrix}\right], \tag{24}$$

in distribution. Since α_1 and α_2 are continuous functions the desired result follows from (23) and (24) upon noting that $\alpha_1'(\theta)=\sigma_1^{-2}(\theta)$ and $\alpha_2'(\phi)=\sigma_2^{-2}(\phi)$, and using a well known convergence theorem (see Rao [10], p. 387). □

COROLLARY 2

Under the stability conditions (9), we have

$$I^{1/2}(\theta,\phi)\begin{pmatrix}\hat{\theta}-\theta\\ \hat{\phi}-\phi\end{pmatrix} \Rightarrow N_2\left[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1&0\\0&1\end{pmatrix}\right] \quad \text{as } T\to\infty. \tag{25}$$

Proof

The result follows directly from theorem 2 and the stability conditions (9). □

REMARKS

1. The use of random normings in theorem 2 simplifies the derivation of the limit distributions as well as its application in practice. Moreover, since the sample Fisher information F remains formally the same for any stopping rule (rules 1 to 4 described in section 2) the limit distribution also remains the same. It is to be noted that no ergodicity assumptions regarding the queue-length process are made in the above derivations.

2. If we choose to use the nonrandom norming $I(\theta, \phi)$ instead of $F(\theta, \phi)$ the expression for $I(\theta, \phi)$ depends on the stopping rule used since $EA(T)$ and $ED(T)$ depend on the distribution of the stopping rule. This is illustrated in the following section for the M/M/1 system, where the stability conditions (9) are replaced by simpler conditions.

4. Computation of limiting Fisher information for M/M/1 under stopping rules

If we wish to use a non-random norming to obtain the limit distribution of the maximum likelihood estimates we need to find the limiting Fisher information under the specific stopping rule used. The limiting covariance matrix for the estimates is then given by the inverse of the limiting Fisher information. In this section we derive the limiting Fisher information for the M/M/1 queue under the stopping rules 1 to 4 described in section 2. Here $f(u;\, \theta) = \theta e^{-\theta u}$ and $g(v;\, \phi) = \phi e^{-\phi v}$. Denote $\rho = \theta\phi^{-1}$.

RULE 1: T FIXED AT t

Here we replace the stability conditions (9) by

$$\text{(a)}\quad \frac{A(T)}{t} \xrightarrow{p} C_1(\theta), \quad \text{and (b)}\ \frac{D(t)}{t} \xrightarrow{p} C_2(\theta, \phi), \tag{26}$$

as $t \to \infty$. The limiting Fisher information $J(\theta, \phi)$ is then given by

$$J(\theta, \phi) = p \lim_{t\to\infty} \left\{ \frac{F(\theta, \phi)}{t} \right\} = \begin{pmatrix} C_1(\theta)\theta^{-2} & 0 \\ 0 & C_2(\theta, \phi)\phi^{-2} \end{pmatrix}, \tag{27}$$

using (15) and the fact that $\mathrm{Var}(u) = \theta^{-2}$, $\mathrm{Var}(v) = \phi^{-2}$. We know that $C_1(\theta) = \theta$ by the renewal theorem and $C_2(\theta, \phi) = \phi\xi$ by lemma 2 of Basawa and Prabhu [1]; in fact, both conditions hold almosts surely. Here $\xi = \min(1, \rho)$. Thus, finally we get

$$J(\theta, \phi) = \begin{pmatrix} \theta^{-1} & 0 \\ 0 & \phi^{-1}\xi \end{pmatrix}. \tag{28}$$

We can then conclude from theorem 2 that as $t \to \infty$,

$$\sqrt{t}\begin{pmatrix} (\hat\theta - \theta) \\ (\hat\phi - \phi) \end{pmatrix} \Rightarrow N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \theta & 0 \\ 0 & \phi\xi^{-1} \end{pmatrix} \right). \tag{29}$$

Wolff [11] used the (random) norming $N(t)$, the total number of transitions, instead of t. Note, however, that he assumes steady state (i.e., $\rho < 1$) to derive his results and invokes Billingsley's [3] limit results for ergodic Markov processes. Our technique is much simpler and at the same time we obtain results without imposing the steady state restriction. Our result in (29) is valid for any ρ.

In order to compare (29) with the corresponding result of Wolff [11] for the steady state we now find the limiting Fisher information using $N(t)$ as the norming. We have

$$\frac{N(t)}{t} = \frac{A(t)}{t} + \frac{D(t)}{t} \xrightarrow{p} \theta + \phi\xi,$$

and therefore using (28) above, we obtain

$$\frac{F}{N(t)} = \left(\frac{F}{T}\right)\left(\frac{t}{N(t)}\right) \xrightarrow{p} \begin{pmatrix} \{\theta(\theta+\phi\xi)\}^{-1} & 0 \\ 0 & \{\phi(\theta+\phi\xi)\}^{-1}\xi \end{pmatrix} \tag{30}$$

as $t \to \infty$. We can now state the following result as $t \to \infty$:

$$\sqrt{N(t)}\begin{pmatrix} \hat{\theta}-\theta \\ \hat{\phi}-\theta \end{pmatrix} \Rightarrow N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \theta(\theta+\phi\xi) & 0 \\ 0 & \phi(\theta+\phi\xi)\xi^{-1} \end{pmatrix}\right). \tag{31}$$

Now, for $\rho < 1$, $\xi = \rho$, and (31) reduces to

$$\sqrt{N(t)}\begin{pmatrix} \hat{\theta}-\theta \\ \hat{\phi}-\phi \end{pmatrix} \Rightarrow N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2\theta^2 & 0 \\ 0 & 2\phi^2 \end{pmatrix}\right), \tag{32}$$

which coincides with Wolff's [11] result for this case. Note however that (31) is valid for all ρ while (32) is a special case for $\rho < 1$.

RULE 2. OBSERVE UNTIL d DEPARTURES

This rule was used by Basawa and Prabhu [1] for the GI/G/1 queue. The limiting Fisher information is seen to be

$$J(\theta, \phi) = p \lim_{d\to\infty}\left(\frac{F}{d}\right) = \begin{pmatrix} \eta\theta^{-2} & 0 \\ 0 & \phi^{-2} \end{pmatrix}$$

where

$$p \lim_{d\to\infty}\left(\frac{A(T)}{d}\right) = \eta = \max(1, \rho)$$

as shown in lemma 1 of Basawa and Prabhu [1]. Consequently, we conclude that, as $d \to \infty$,

$$\sqrt{d}\begin{pmatrix} \hat{\theta}-\theta \\ \hat{\phi}-\phi \end{pmatrix} \Rightarrow N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \eta^{-1}\theta^2 & 0 \\ 0 & \phi^2 \end{pmatrix}\right). \tag{33}$$

RULE 3. OBSERVE UNTIL m ARRIVALS

We have

$$F=\begin{pmatrix} m\theta^{-2} & 0 \\ 0 & D(T)\phi^{-2} \end{pmatrix}.$$

Note that $T=\sum_1^m u_i \to \infty$ as $m\to\infty$. Now

$$\frac{D(T)}{m}=\frac{D(u_1+u_2+\cdots+u_m)}{u_1+u_2+\ldots u_m}\cdot\frac{u_1+u_2+\ldots u_m}{m}$$

$$\to \phi\xi\theta^{-1}=\rho^{-1}\xi$$

by (26b), which holds almost surely, and the law of large numbers. Hence the limiting Fisher information is

$$J(\theta,\phi)=p\lim_{m\to\infty}\left(\frac{F}{m}\right)=\begin{pmatrix} \theta^{-2} & 0 \\ 0 & \phi^{-2}\phi^{-1}\xi \end{pmatrix}. \tag{34}$$

Finally, the limit distribution of the maximum likelihood estimates is given by

$$\sqrt{m}\begin{pmatrix} \hat\theta-\theta \\ \hat\phi-\phi \end{pmatrix} \Rightarrow N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \theta^2 & 0 \\ 0 & \phi^2\rho\xi^{-2} \end{pmatrix}\right), \tag{35}$$

as $m\to\infty$.

RULE 4. OBSERVE UNTIL n TRANSITIONS

Note that here $T=\sum_1^n T_j$, where $\{T_j\}$ are the intervals between successive transitions. We have

$$\frac{T}{n}=\frac{1}{n}\left\{\sum_1^n T_j\mathscr{I}(X_{j-1}=0)+\sum_1^n T_j\mathscr{I}(X_{j-1}>0)\right\}$$

where $\mathscr{I}(A)$ denotes the indicator function of the event A, and $X_j=Q(T_1+T_2+\ldots+T_{j-1})=$ queue-length at the $(j-1)$th transition epoch. It is easily verified that, as $n\to\infty$,

$$\frac{T}{n}\xrightarrow{p}(1-\xi)\theta^{-1}+\xi(\theta+\phi)^{-1}.$$

Furthermore, we have

$$\frac{A(T)}{T}\xrightarrow{p}\theta \quad \text{as } n\to\infty \text{ (and hence } T\to\infty \text{ a.s.).}$$

Finally, we have, as in $\to\infty$,

$$\frac{A(T)}{n}=\left(\frac{A(T)}{T}\right)\left(\frac{T}{n}\right)\xrightarrow{p}\theta\{(1-\xi)\theta^{-1}+\xi(\theta+\phi)^{-1}\}$$

$$=1-(1+\rho)^{-1}\xi.$$

Also,

$$\frac{D(T)}{n} = 1 - \frac{A(T)}{n} \to (1+\rho)^{-1}\xi.$$

The limiting Fisher information is given by

$$J(\theta, \phi) = p \lim_{n\to\infty} \left(\frac{F}{n}\right) = \begin{pmatrix} \theta^{-2}\{1-(1+\rho)^{-1}\xi\} & 0 \\ 0 & \phi^{-2}(1+\rho)^{-1}\xi \end{pmatrix}, \tag{36}$$

where $\xi = \min(1, \rho)$ as before. Thus,

$$\sqrt{n}\begin{pmatrix} \hat{\theta}-\theta \\ \hat{\phi}-\phi \end{pmatrix} \Rightarrow N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, J^{-1}(\theta, \phi)\right), \quad \text{as } n \to \infty. \tag{37}$$

It is interesting to compare (37) with (31) where t was fixed and $N(t)$ was random while in (37) T is random and n is fixed.

REMARKS

We have demonstrated above the fact that the use of a non-random norm necessitates one to compute the limiting Fisher information which in turn varies with the stopping rule. On the otherhand, if we use the random norming theorem 2 gives directly:

$$\begin{pmatrix} (A(T))^{1/2}(\hat{\theta}-\theta) \\ (D(T))^{1/2}(\hat{\phi}-\phi) \end{pmatrix} \Rightarrow N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \theta^2 & 0 \\ 0 & \phi^2 \end{pmatrix}\right), \tag{38}$$

as $T \to \infty$ a.s. Note that (38) remains formally the same under all the stopping rules 1 to 4. For this reason it is simpler to use (38) no matter what stopping rule is used. Moreover, steady-state assumption is *not* required for the validity of (38). Nor is it required for the computation of the limiting Fisher information using non-random normings.

5. Generalized linear model

Suppose we observe k independent GI/G/1 queues over $(0, T_i]$ where T_i is a suitable stopping time , $(i = 1, \ldots, k)$. The sample data consist of $\{(A_i(T_i), u_{ij})$, $i = 1, \ldots, k$, $j = 1, \ldots, A_i(T_i)$, and $(D_i(T_i), v_{il})$, $l = 1, \ldots, D_i(T_i)\}$ where $A_i(T_i)$ is the number of arrivals for the ith queue, $D_i(T_i)$ the number of departures, u_{ij} denotes the ith interarrival time, and v_{il} is the service time of the lth customer. The densities of u_{ij} and v_{il} are assumed respectively to be

$$f_i(u_{ij};\, \theta_i) = a_{1i}(u_{ij}) \exp\{\theta_i h_{1i}(u_{ij}) - k_{1i}(\theta_i)\} \tag{39}$$

$$g_i(v_{il};\, \phi_i) = a_{2i}(v_{il}) \exp\{\phi_i h_{2i}(v_{il}) - k_{2i}(\phi_i)\}. \tag{40}$$

The log-likelihood function in this case is given by (ignoring terms free of parameters)

$$\ln L(\theta, \phi) = \sum_{i=1}^{k} \left[\left\{ \theta_i \sum_{j=1}^{A_i} h_{1i}(u_{ij}) - A_i k_{1i}(\theta_i) \right\} + \left\{ \phi_i \sum_{l=1}^{D_i} h_{2i}(v_{il}) - D_i k_{2i}(\phi_i) \right\} \right]. \tag{41}$$

Let $\mu_{1i}(\theta) = k'_{1i}(\theta) = E_\theta(h_1(u_{ij}))$, and $\mu_{2i}(\phi) = k'_{2i}(\phi) = E_\phi(h_2(v_{il}))$. Since μ_1 and μ_2 are positive quantities we consider the log-linear model

$$\ln \mu_{1i}(\theta) = \sum_{j=1}^{p} a_{ij}\alpha_j, \quad \ln \mu_{2i}(\phi) = \sum_{j=1}^{q} b_{ij}\beta_j, \tag{42}$$

where $((a_{ij}))$ and $((b_{ij}))$ are known covariates, and the vectors $\alpha^{\mathrm{T}} = (\alpha_1, \ldots, \alpha_p)$, $\beta^{\mathrm{T}} = (\beta_1, \ldots, \beta_q)$ are the unknown parameters of interest. The likelihood estimating equations for α and β are given by

$$\frac{\partial \ln L}{\partial \alpha} = \sum_{i=1}^{k} \left\{ (S_{1i} - \mu_{1i}(\alpha) A_i) \sigma_{1i}^{-2}(\alpha) \right\} \left(\frac{\partial \mu_{1i}}{\partial \alpha} \right) = 0$$

and

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^{k} \left\{ (S_{2i} - \mu_{2i}(\beta) D_i) \sigma_{2i}^{-2}(\beta) \right\} \left(\frac{\partial \mu_{2i}}{\partial \beta} \right) = 0$$

where

$$S_{1i} = \sum_{j=1}^{A_i} h_{1i}(u_{ij}), \quad S_{2i} = \sum_{l=1}^{D_i} h_{2i}(v_{il}), \tag{43}$$

$$\sigma_{1i}^2(\alpha) = \operatorname{Var}(h_1(u_{ij})), \quad \sigma_{2i}^2(\beta) = \operatorname{Var}(h_2(v_{il})). \tag{44}$$

The sample Fisher information for α and β is given by

$$F = \begin{pmatrix} F_{11} & 0 \\ 0 & F_{22} \end{pmatrix},$$

where

$$F_{11} = -\frac{\partial^2 \ln L}{\partial \alpha^2}$$

$$= -\sum_{i=1}^{k} \left[\left(\frac{S_{1i} - A_i \mu_{1i}(\alpha)}{\sigma_{1i}^2(\alpha)} \right) \left\{ \left(\frac{\partial^2 \mu_{1i}}{\partial \alpha^2} \right) - \frac{\left(\frac{\partial \mu_{1i}}{\partial \alpha} \right) \left(\partial \sigma_{1i}^2 \quad \partial \alpha \right)^{\mathrm{T}}}{\sigma_{1i}^2(\alpha)} \right\} \right]$$

$$+ \sum_{i=1}^{k} \frac{A_i}{\sigma_{1i}^2(\alpha)} \left(\frac{\partial \mu_{1i}}{\partial \alpha} \right) \left(\frac{\partial \mu_{1i}}{\partial \alpha} \right)^{\mathrm{T}},$$

and a similar expression for F_{22} Since by random sum central limit theorem we have

$$A_i^{-1/2}(S_{1i} - A_i\mu_{1i}) \Rightarrow N(0, \sigma_{1i}^2), \quad \text{as } A_i \to \infty \text{ a.s.}$$

for each $i = 1, \ldots, k$, it follows that

$$F_{11} = \sum_{i=1}^{k} \left\{ \frac{A_i}{\sigma_{1i}^2(\alpha)} \left(\frac{\partial \mu_{1i}}{\partial \alpha} \right) \left(\frac{\partial \mu_{1i}}{\partial \alpha} \right)^{\mathrm{T}} [1 + o_p(1)] \right\}.$$

We therefore consider an approximate form for the sample Fisher information given by

$$F^* = \begin{pmatrix} F_{11}^* & 0 \\ 0 & F_{22}^* \end{pmatrix} \tag{45}$$

where

$$F_{11}^* = \sum_{i=1}^{k} \frac{A_i}{\sigma_{1i}^2(\alpha)} \left(\frac{\partial \mu_{1i}}{\partial \alpha} \right) \left(\frac{\partial \mu_{1i}}{\partial \alpha} \right)^{\mathrm{T}},$$

and

$$F_{22}^* = \sum_{i=1}^{k} \frac{D_i}{\sigma_{2i}^2(\alpha)} \left(\frac{\partial \mu_{2i}}{\partial \beta} \right) \left(\frac{\partial \mu_{2i}}{\partial \beta} \right)^{\mathrm{T}}.$$

Consider now the following stability conditions:

$$\frac{A_i}{E(A_i)} \xrightarrow{p} 1, \quad \frac{D_i}{E(D_i)} \xrightarrow{p} 1 \quad \text{as } T_i \to \infty \text{ a.s.,} \quad \text{for each } i = 1, \ldots, k. \tag{46}$$

One can show using essentially the same arguments as in section 3 that the following result holds irrespective of which of the stopping rules 1–4 is used.

THEOREM 3

We have

$$F^{*1/2} \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix} \Rightarrow N_{p+q}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, I \right),$$

as $T_i \to \infty$ a.s. $(i = 1, \ldots, k)$, where I denotes the $(p+q) \times (p+q)$ identity matrix.

REMARKS

(i) It is interesting to note that F^* has essentially the same form as the exact (expected) Fisher information in the classical application of the generalized linear model, where the sample sizes A_i and D_i are treated as non-random. Here F^* is an approximate sample Fisher information, and A_i, D_i are of course random variables.

(ii) In the classical application of the generalized linear model one typically lets the number of samples $k \to \infty$, while each sample size (A_i and D_i) is treated fixed. In the above situation however, we have treated k as fixed and let the sample sizes A_i, D_i increase. If required, however, we could let A_i and D_i be finite (random variables), and carry out asymptotics as $k \to \infty$. The stability conditions required for this approach are a minor modification of those given by Fahrmeir and Kaufmann ([5], [6]) for the classical generalized linear model. We shall not pursue this point further here.

EXAMPLE: M/M/1

We now consider the special case for M/M/1. The log-likelihood function ignoring terms free from parameters is given by

$$\ln L(\theta, \phi) = \sum_{i=1}^{k} \left\{ \left(\theta_i S_{1i} - A_i \ln \theta_i^{-1} \right) + \left(\phi_i S_{2i} - D_i \ln \phi_i^{-1} \right) \right\} \tag{47}$$

where $S_{1i} = -\sum_{j=1}^{A_i} u_{ij}$, and $S_{2i} = -\sum_{j=1}^{D_i} v_{il}$. Consider the log-linear model:

$$\ln \mu_{1i} = \ln \theta_i^{-1} = \alpha_1 + \alpha_2 a_i,$$

and

$$\ln \mu_{2i} = \ln \phi_i^{-1} = \beta_1 + \beta_2 b_i, \; (i = 1, \dots, k), \tag{48}$$

where $\{a_i\}$ and $\{b_i\}$ are known covariates, and $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ are the unknown parameters of interest. The estimating equations are:

$$\frac{\partial \ln L}{\partial \alpha_1} = \sum_{i=1}^{k} \{ S_{1i} - A_i \exp(\alpha_1 + \alpha_2 a_i) \} \exp\{ -(\alpha_1 + \alpha_2 a_i) \} = 0$$

$$\frac{\partial \ln L}{\partial \alpha_2} = \sum_{i=1}^{k} \{ S_{1i} - A_i \exp(\alpha_1 + \alpha_2 a_i) \} a_i \exp\{ -(\alpha_1 + \alpha_2 a_i) \} = 0$$

$$\frac{\partial \ln L}{\partial \beta_1} = \sum_{i=1}^{k} \{ S_{2i} - D_i \exp(\beta_1 + \beta_2 b_i) \} \exp\{ -(\beta_1 + \beta_2 b_2) \} = 0$$

$$\frac{\partial \ln L}{\partial \beta_2} = \sum_{i=1}^{k} \{ S_{2i} - D_i \exp(\beta_1 + \beta_2 b_i) \} b_i \exp\{ -(\beta_1 + \beta_2 b_i) \} = 0.$$

The approximate sample Fisher information is given by F^* in (41) with

$$F_{11}^* = \begin{pmatrix} \sum_1^k A_i & \sum_1^k a_i A_i \\ \sum a_i A_i & \sum a_i^2 A_i \end{pmatrix}, \quad \text{and } F_{22}^* = \begin{pmatrix} \sum_1^k D_i & \sum_1^k b_i D_i \\ \sum b_i D_i & \sum b_i^2 D_i \end{pmatrix}. \tag{49}$$

Note that F_{11}^* and F_{22}^* do not depend on the parameters. Theorem 3 gives the limit distribution of $(\hat{\alpha}_1, \hat{\alpha}_2)$, $(\hat{\beta}_1, \hat{\beta}_2)$, using the normalizing factors (random) given in (49). If the non-random normings are to be used one can utilize the results of the previous section to compute limiting Fisher information matrix under the appropriate stopping rules.

6. Large sample tests

In this section we briefly mention two large sample tests and their limiting null distributions. Similar tests can be constructed for other parametric hypotheses.

A TEST OF FIT

We now wish to test the hypothesis of log linearity of the interarrival time and service time means within the exponential family with arbitrary means. Specifically, we test the composite null hypothesis (42), namely:

$$H\colon \ln \mu_{1i}(\theta) = \sum_{j=1}^{p} a_{ij}\alpha_j \quad \text{and} \ln \mu_{2i}(\phi) = \sum_{j=1}^{q} b_{ij}\beta_j \tag{50}$$

where $\{\alpha_j\}$ and $\{\beta_j\}$ are unknown parameters.
The unrestricted maximum likelihood estimates of θ_i and ϕ_i under the model in (41) are obtained by solving

$$S_{1i} = A_i\mu_{1i}(\theta_i), \quad \text{and } S_{2i} = D_i\mu_{2i}(\phi_i) \quad (i = 1, \ldots, k), \tag{51}$$

where S_{1i} and S_{2i} are given by (43). Let $\hat{\theta}_i$ and $\hat{\phi}_i$ be the solutions of (51). The likelihood ratio statistic for testing (50) within the model (41) is given by

$$Q_n = -2 \ln\{L_H(\hat{\alpha}, \hat{\beta})/L(\hat{\theta}, \hat{\phi})\}, \tag{52}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the maximum likelihood estimates under H obtained in section 5. It can be shown via standard asymptotics that under conditions (46) of section 5, Q_n is asymptotically $\chi^2(2k - p - q)$ where we assume that $2k > (p + q)$.

A TEST OF HOMOGENEITY

Consider the M/M/1 model discussed in section 5. It is assumed that the log-linear model is a good fit. Suppose we wish to test the hypothesis H_0: $\alpha_2 = 0$, $\beta_2 = 0$. This is equivalent to testing the hypothesis that the k queues have the same arrival and service rates, i.e.,

$$\theta_1 = \ldots = \theta_k = e^{-\alpha_1}, \quad \text{and } \phi_1 = \ldots = \phi_k = e^{-\beta_1}, \tag{53}$$

when the parameters α_1 and β_1 are unknown. The log likelihood function under H_0 is given by

$$\ln L_{H_0} = \left\{ e^{-\alpha_1} \sum_{i=1}^{k} S_{1i} - \alpha_1 \sum_{i=1}^{k} A_i \right\} + \left\{ e^{-\beta_1} \sum_{1}^{k} S_{2i} - \beta_1 \sum_{1}^{k} D_i \right\}. \tag{54}$$

The maximum likelihood estimates of α_1 and β_1 under H_0 are obtained as

$$\hat{\alpha}_{10} = \ln\left\{\frac{\sum_{i=1}^{k}\sum_{j=1}^{A_i} u_{ij}}{\sum_{1}^{k} A_i}\right\}, \quad \text{and } \hat{\beta}_{10} = \ln\left\{\frac{\sum_{i=1}^{k}\sum_{l=1}^{D_i} v_{il}}{\sum_{1}^{k} D_i}\right\}. \tag{55}$$

The unrestricted maximum likelihood estimates of α_1, α_2, β_1 and β_2 are obtained as in section 5. The likelihood ratio statistic for testing H_0 within H is given by

$$Q_n = -2\ln\left\{L_{H_0}\left(\hat{\alpha}_{10}, \hat{\beta}_{10}, \alpha_2 = 0, \beta_2 = 0\right)/L_H\left(\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2\right)\right\}. \tag{56}$$

The statistic Q_n has a limiting χ^2 distribution under H_0.

Acknowledgement

The authors are grateful to a referee and Professor S. Zacks for many helpful comments. The first author's work was partially supported by a grant from the Office of the Naval Research.

References

[1] I.V. Basawa and N.U. Prabhu, Estimation in single server queues. Nav. Res. logist. Quart. 28 (1981) 475–487.

[2] U.N. Bhat and S.S. Rao, Statistical analysis of queueing systems, Queueing Systems Theory Appl. 1 (1987) 217–247.

[3] P. Billingsley, *Statistical Inference for Markov Processes* (Chicago University Press, Chicago, 1961).

[4] P. Billingsley, *Convergence of Probability Measures* (John Wiley, New York, 1968).

[5] L. Fahrmeir and H. Kaufmann, Consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models, Ann. Statist. 13 (1985) 342–368.

[6] L. Fahrmeir and H. Kaufmann, Correction to consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models, Ann. Statist. 14 (1986) 1643.

[7] P. McCullagh and J.A. Nelder, *Generalized Linear Models* (Chapman and Hall, London, 1983).

[8] M.F. McGrath, D. Gross and N.D. Singpurwalla, A subjective approach to the theory of queues I-modeling, Queueing Systems Theory Appl. 1 (1987) 317–333.

[9] M.F. McGrath and N.D. Singpurwalla, A subjective approach to the theory of queues II-inference and information in M/M/-queues. Queueing Systems Theory Appl. 1 (1987) 335–353.

[10] C.R. Rao, *Linear Statistical Inference and its Applications*, 2nd ed. (John Wiley, New York, 1973).

[11] R.W. Wolff, Problems of statistical inference for birth and death queueing models, Oper. Res. 13 (1965) 343–357.