MAXIMUM LIKELIHOOD ESTIMATION FOR SINGLE SERVER QUEUES FROM WAITING TIME DATA

Ishwar V. Basawa
University of Georgia

U. Narayan Bhat
Southern Methodist University

Robert Lund
University of Georgia

ABSTRACT

Maximum likelihood estimators for the parameters of a GI/G/1 queue are derived based on the information on waiting times $\{W_t\}, t = 1, ..., n$, of n successive customers. The consistency and asymptotic normality of the estimators are established. A simulation study of the M/M/1 and M/E_k/1 queues is presented.

1 Introduction

The literature on parameter estimation in queues is extensive. Basawa and Prabhu (1981, 1988) have discussed moment and maximum likelihood estimation of the model parameters for GI/G/1 queues. Basawa and Bhat (1992) studied sequential inference for the parameters of a GI/G/1 queue. An empirical Bayes approach was used for estimation by Thiruvaiyaru and Basawa (1992). The estimation methods used in the above mentioned papers required full information in the sense that the information on all service times and inter-arrival times was needed. See Bhat and Rao (1987) for a review of the literature on inference for queueing systems.

Often, only a partial information on a queue may be available. It is, therefore, useful to develop inferential methods based on sampling plans which do not require the full information on service and inter-arrival times. One such sampling plan, which is easy to implement in practice, is based on observing only the waiting times of the successive customers in a queue. In this paper, we discuss the maximum likelihood estimation of the model parameters based only on waiting times, and derive their limit distributions. The waiting time process is Markovian with transition densities having a jump at the origin.

This creates special problems in the derivation of the likelihood function and the Fisher information matrix. The problem, therefore, is non-standard, even for the special case of M/M/1 queues.

The likelihood function based on waiting times is derived in Section 2. The consistency and the asymptotic normality of the maximum likelihood (ML) estimators are established in Section 3. Section 4 considers specific applications to the M/M/1 and M/E$_k$/1 queues. Simulation results on the performance of the ML estimators are presented in Section 5. Finally, the Appendix contains regularity conditions and the derivation of the limiting Fisher information matrix for the M/M/1 queue.

2 The likelihood function based on waiting times

Let W_t denote the waiting time of the t^{th} customer in a GI/G/1 queue. The waiting time process $\{W_t\}, t = 1, 2, \ldots$, has the following representation:

$$W_{t+1} = \begin{cases} W_t + X_{t+1}, & \text{if } W_t + X_{t+1} > 0 \\ 0, & \text{if } W_t + X_{t+1} \leq 0, \end{cases} \tag{2.1}$$

where $X_t = V_{t-1} - U_t$, with V_t and U_t denoting, respectively, the service and inter-arrival times corresponding to the t^{th} customer (Prabhu (1980), p2). Note that $\{X_t\}$ is a sequence of independent and identically distributed random variables and that X_{t+1} is independent of W_t. From (2.1), it follows that $\{W_t\}$ is a Markov chain on the state space $S = [0, \infty)$, with the stationary transition distribution function

$$\begin{aligned} P(W_{t+1} \leq y | W_t = w) &= P(W_{t+1} = 0 | W_t = w) + P(0 < W_{t+1} \leq y | W_t = w) \\ &= P(w + X_{t+1} \leq 0) + P(0 < w + X_{t+1} \leq y) \\ &= \begin{cases} F_x(y - w), & y \geq 0 \\ 0, & y < 0, \end{cases} \end{aligned} \tag{2.2}$$

where $F_x(\cdot)$ is the distribution function of X_1.

It is clear from (2.2) that the transition distribution function has a jump discontinuity at $y = 0$, with jump-size equal to $F_x(-w)$. Let ϕ denote the measure which assigns measure 1 to the point $w = 0$, and is the Lebesgue measure on $(0, \infty)$. Then, the transition density of W_{t+1} given $W_t = w$, with respect to the measure ϕ, is given by

$$p(y|w) = \begin{cases} 1 - \alpha(w), & y = 0 \\ f_x(y - w), & y > 0 \\ 0, & y < 0, \end{cases} \tag{2.3}$$

where

$$\alpha(w) = 1 - F_x(-w), \tag{2.4}$$

and $f_x(\cdot)$ denotes the density (assumed to exist) corresponding to $F_x(\cdot)$. Denoting the transition density of W_{t+1} given W_t by $p(W_{t+1}|W_t)$, we have, from (2.3),

$$p(W_{t+1}|W_t) = (1-\alpha(W_t))^{1-Z_{t+1}}(f_x(W_{t+1}-W_t))^{Z_{t+1}}, \tag{2.5}$$

where

$$Z_{t+1} = \begin{cases} 0, & \text{if } W_{t+1} = 0 \\ 1, & \text{if } W_{t+1} > 0. \end{cases} \tag{2.6}$$

The likelihood function, based on the sample $(W_1, W_2, ..., W_n)$, is given by

$$L = p(W_1)\prod_{t=1}^{n-1} p(W_{t+1}|W_t), \tag{2.7}$$

where $p(W_{t+1}|W_t)$ is given by (2.5) and $p(W_1)$ denotes the marginal density of W_1. Now, suppose that $f_x(.)$ depends on an unknown parameter vector $\theta = (\theta_1, ..., \theta_r)'$, and let $f_x(.;\theta)$ denote the density of X_t. Also, we will use the notation $\alpha(W_t;\theta)$ to indicate dependence of $\alpha(.)$ on θ. The likelihood score vector for estimating θ can be obtained from (2.5) - (2.7):

$$\begin{aligned}\frac{dlnL}{d\theta} &= \frac{dlnp(W_1;\theta)}{d\theta} - \sum_{t=1}^{n-1}(1-Z_{t+1})\frac{d\alpha(W_t;\theta)}{d\theta}(1-\alpha(W_t;\theta))^{-1} \\ &+ \sum_{t=1}^{n-1} Z_{t+1}\left(\frac{dlnf_x(W_{t+1}-W_t;\theta)}{d\theta}\right).\end{aligned} \tag{2.8}$$

3 Consistency and asymptotic normality of the maximum likelihood estimator

The maximum likelihood (ML) estimator $\hat{\theta}_n$ of θ is typically obtained as a solution of the likelihood equation $\frac{dlnL}{d\theta} = 0$. Under some regularity conditions, we shall now establish the consistency and the asymptotic normality of the ML estimator. Let $\alpha = EX_t = EV_t - EU_t$. If $\alpha < 0$, it is known that $W_t \xrightarrow{d} W$, as $t \to \infty$, where the form of the distribution of the limiting random variable W is contained in Prabhu (1980) p. 28. In particular we note that $P(W=0) = 1-(EV_t)/(EU_t) > 0$. If we take the initial distribution of W_1 to be the same as the distribution of W, the Markov process $\{W_t\}$ becomes strictly stationary. We shall assume throughout the paper that $\alpha < 0$ (i.e., the system is in equilibrium).

Also, for $\alpha < 0$, it is seen that $\{W_t\}$ is ergodic in a regenerative sense. In order to verify this, it is enough to show that the set $\{n : P(W_n = 0|W_1 = 0) > 0\}$ has unity as its greatest common divisor. Since $\alpha < 0$, we have $P(W_2 = 0|W_1 = 0) > 0$; the Markov property of $\{W_t\}$ now shows that $P(W_n = 0|W_1 = 0) > 0$, for all $n \geq 2$. Hence the greatest common divisor of the above set is unity as required. Notice also that the initial state W_1 does not influence ergodicity. This is because there is an $n > 2$ such that $P(W_n = 0|W_1 = x) > 0$ for all $x \geq 0$; for if this is not the case, $P(W = 0) = 0$.

The first term in (2.8), due to the initial observation W_1, is asymptotically negligible. See, for instance, Billingsley (1961). Consequently, we will define the likelihood score function $S_n(\theta)$ neglecting the first term, i.e.,

$$S_n(\theta) = \sum_{t=1}^{n-1} U_t(\theta), \text{with}$$

$$U_t(\theta) = (Z_{t+1} - 1)\frac{d\alpha(W_t;\theta)}{d\theta}(1 - \alpha(W_t;\theta))^{-1} + Z_{t+1}\frac{d\ln f_x(W_{t+1} - W_t;\theta)}{d\theta}. \quad (3.1)$$

We shall first show that $\{S_n(\theta)\}$ is a zero-mean martingale with respect to the σ-field, $\sigma(W_{n-1}, ..., W_1)$. It is enough to verify that

$$E(U_{t+1}(\theta)|W_t) = 0, \text{a.s.} \quad (3.2)$$

We have

$$E(U_{t+1}(\theta)|W_t) = -\frac{d\alpha(W_t;\theta)}{d\theta} + \int_0^\infty (\frac{df_x(W_{t+1} - W_t;\theta)}{d\theta} dW_{t+1}, \quad (3.3)$$

since

$$E(Z_{t+1}|W_t) = \alpha(W_t;\theta). \quad (3.4)$$

Differentiating both sides of (3.4) with respect to θ, and assuming differentiability under the integral sign, we have

$$\int_0^\infty (\frac{df_x(W_{t+1} - W_t;\theta)}{d\theta})dW_{t+1} = \frac{d\alpha(W_t;\theta)}{d\theta}. \quad (3.5)$$

The result in (3.2) now follows from (3.3) and (3.5).

Consider the conditional information matrix $\xi_t(\theta)$ defined by

$$\xi_t(\theta) = E(U_{t+1}(\theta)U'_{t+1}(\theta)|W_t).$$

Using (3.1) and the fact that $(Z_{t+1} - 1)Z_t = 0$ gives

$$\begin{aligned}
\xi_t(\theta) &= E[(Z_{t+1} - 1)^2(1 - \alpha(W_t;\theta))^{-2}(\frac{d\alpha(W_t;\theta)}{d\theta})(\frac{d\alpha(W_t;\theta)}{d\theta})'|W_t] \\
&+ E[Z_{t+1}^2(\frac{d\ln f_x(W_{t+1} - W_t;\theta)}{d\theta})(\frac{d\ln f_x(W_{t+1} - W_t;\theta)}{d\theta})'|W_t] \\
&= (1 - \alpha(W_t;\theta))^{-1}(\frac{d\alpha(W_t;\theta)}{d\theta})(\frac{d\alpha(W_t;\theta)}{d\theta})' \\
&+ \int_0^\infty f_x^{-1}(W_{t+1} - W_t;\theta)(\frac{df_x(W_{t+1} - W_t;\theta)}{d\theta})(\frac{df_x(W_{t+1} - W_t;\theta)}{d\theta})' dW_{t+1} \quad (3.6)
\end{aligned}$$

Define

$$J_n(\theta) = \sum_{t=1}^{n-1} \xi_t(\theta). \quad (3.7)$$

From the ergodic theorem, it follows that,

$$\frac{1}{n}J_n(\theta) \overset{a.s.}{\to} J(\theta), \tag{3.8}$$

where

$$J(\theta) = E(\xi_t(\theta)), \tag{3.9}$$

the expectation in (3.9) being with respect to the stationary distribution. It will be assumed that $J(\theta)$ is positive definite. It follows, from the central limit theorem for martingales (Billingsley (1961)), that

$$n^{-1/2}S_n(\theta) \overset{d}{\to} N_r(0, J(\theta)). \tag{3.10}$$

Assuming $J_n(\theta)$ to be positive definite (a.s.) for every n, an alternative form of the result in (3.10) is

$$J_n^{-1/2}(\theta)S_n(\theta) \overset{d}{\to} N_r(0, I_r), \tag{3.11}$$

where I_r is the $r \times r$ identity matrix. Under regularity conditions, such as those in Billingsley (1961), (see Appendix A.1)) we have the following limit results.

Theorem 3.1. There exists a consistent solution $\hat{\theta}_n$ of the equation $S_n(\theta) = 0$, with probability tending to 1 as $n \to \infty$. Moreover, $\hat{\theta}_n$ is a local maximum of $L_n(\theta)$ with probability tending to 1 as $n \to \infty$.

Theorem 3.2. If $\hat{\theta}_n$ is any consistent solution of the likelihood equation $S_n(\theta) = 0$, then

$$(i) \;\; \sqrt{n}(\hat{\theta}_n - \theta) \overset{d}{\to} N_r(0, J^{-1}(\theta)),$$

$$(ii) \;\; 2(lnL_n(\hat{\theta}_n) - lnL_n(\theta)) \overset{d}{\to} \chi^2(r), \text{as } n \to \infty.$$

Remarks: An alternative form of the result in Theorem 3.2 (i) is:

$$(i)^* \;\; J_N^{1/2}(\theta)(\hat{\theta}_n - \theta) \overset{d}{\to} N_r(0, I_r).$$

The result in Theorem 3.2 (ii) establishes the limit distribution of the log-likelihood ratio statistic.

4 Applications

In this section, we consider two special cases, viz., M/M/1 and M/E_k/1 queues. Note that the likelihood function in (2.7) involves the functions $f_x(x)$, and $\alpha(w)$, where $f_x(x)$ is the density function of $X_t = V_{t-1} - U_t$, and $\alpha(w) = \int_{-w}^{\infty} f_x(y)dy$. We comment that, for simplicity of notation, we have suppressed dependence on θ.

M/M/1 Queue

Let λ and μ denote the arrival and service rates respectively. The stationarity condition is $\lambda/\mu < 1$. It is readily seen that

$$f_x(x) = \begin{cases} \frac{\lambda\mu}{(\lambda+\mu)}e^{-\mu x}, & 0 \le x < \infty \\ \frac{\lambda\mu}{(\lambda+\mu)}e^{\lambda x}, & -\infty < x < 0, \end{cases} \tag{4.1}$$

and

$$\alpha(w) = 1 - (\frac{\mu}{\lambda+\mu})e^{-\lambda w}. \tag{4.2}$$

Here the parameter to be estimated is $\theta = (\lambda, \mu)'$. The score vector in (3.1) is $S_n(\theta) = (S_1(\theta), S_2(\theta))'$, with

$$S_1(\theta) = \sum_{t=1}^{n-1}[(Z_{t+1}-1)(1-\alpha(W_t))^{-1}\frac{\partial}{\partial\lambda}(\alpha(W_t)) + Z_{t+1}f_x^{-1}(W_{t+1}-W_t)\frac{\partial}{\partial\lambda}f_x(W_{t+1}-W_t)],$$

and

$$S_2(\theta) = \sum_{t=1}^{n-1}[(Z_{t+1}-1)(1-\alpha(W_t))^{-1}\frac{\partial}{\partial\mu}(\alpha(W_t)) + Z_{t+1}f_x^{-1}(W_{t+1}-W_t)\frac{\partial}{\partial\mu}f_x(W_{t+1}-W_t)].$$

Equations (4.1) and (4.2) give

$$\begin{aligned} \frac{\partial}{\partial\lambda}(\alpha(W_t)) &= (\frac{\mu e^{-\lambda W_t}}{\lambda+\mu})(W_t + \frac{1}{\lambda+\mu}), \\ \frac{\partial}{\partial\mu}(\alpha(W_t)) &= -\frac{\lambda}{(\lambda+\mu)^2}e^{-\lambda W_t}, \\ \frac{\partial}{\partial\lambda}f_x(x) &= \begin{cases} \frac{\mu^2 e^{-\mu x}}{(\lambda+\mu)^2}, & \text{if } x \ge 0 \\ \frac{\mu e^{\lambda x}}{(\lambda+\mu)}(x\lambda + \frac{\mu}{\lambda+\mu}), & \text{if } x < 0, \end{cases} \end{aligned}$$

and

$$\frac{\partial}{\partial\mu}f_x(x) = \begin{cases} \frac{\lambda e^{-\mu x}}{(\lambda+\mu)}(-x\mu + \frac{\lambda}{\lambda+\mu}), & \text{if } x \ge 0 \\ \frac{\lambda^2}{(\lambda+\mu)^2}e^{\lambda x}, & \text{if } x < 0. \end{cases}$$

This completes the specification of the score vector. The conditional information matrix $\xi_t(\theta)$ in (3.6) involves $\alpha(w)$, $f_x(x)$, and their partial derivatives with respect to λ and μ, all of which are given above.

In order to compute the unconditional Fisher information matrix $J(\theta)$ in (3.9), we need the stationary distribution of the process $\{W_t\}$. The stationary distribution function of W_t is given by (see Prabhu (1980), p. 33)

$$F_W(w) = \begin{cases} 1 - \frac{\lambda}{\mu}e^{-(\mu-\lambda)w}, & w \ge 0 \\ 0, & w < 0, \end{cases}$$

with the density given by

$$f_W(w) = \begin{cases} 1-\frac{\lambda}{\mu}, & w=0 \\ (\frac{\lambda}{\mu})(\mu-\lambda)e^{-(\mu-\lambda)w}, & w>0 \\ 0, & w<0. \end{cases} \quad (4.3)$$

M/E_k/1 Queue

Let λ denote the arrival rate as before. The service time density is given by

$$g(v) = \frac{e^{-\mu v}(\mu v)^{k-1}\mu}{(k-1)!}, 0 < v < \infty.$$

The stationarity condition is: $\rho = (k\lambda/\mu) < 1$. The relevant quantities needed in the derivation of the score vector and the conditional information matrix are given below.

$$f_x(x) = \begin{cases} (\frac{\mu}{\lambda+\mu})^k \lambda e^{-\mu x} \sum_{r=0}^{k-1} \frac{[(\lambda+\mu)x]^r}{r!}, & x \geq 0, \\ (\frac{\mu}{\lambda+\mu})^k \lambda e^{\lambda x}, & x < 0. \end{cases} \quad (4.4)$$

$$\alpha(w) = 1 - (\frac{\mu}{\lambda+\mu})^k e^{-\lambda w}. \quad (4.5)$$

The partial derivatives of $f_x(x)$ and $\alpha(w)$ with respect to λ and μ are easily found, and we omit the details.

Finally, the stationary distribution of $\{W_t\}$ has the Laplace transform given by

$$\int_{0-}^{\infty} e^{-\theta w} dF_W(w) = \frac{(1-\rho)\theta}{\theta - \lambda + \lambda(\frac{k\mu}{k\mu+\theta})^k}, \quad \theta > 0. \quad (4.6)$$

See, for instance, Prabhu (1980), p. 38.

The Laplace transform in (4.6) can be inverted to show that, for $0 < w < \infty$, the density of W can be expressed as a weighted sum of k exponential densities, and that the distribution of W has a jump at $W = 0$, with jump size

$$P(W = 0) = 1 - \rho.$$

See Prabhu (1965), p 28, for the details.

5 Simulation results

Table 1 summarizes simulation results on the performance of the ML estimators for various M/M/1 queues. Three pairs of values for (λ, μ) were used corresponding to the traffic intensities $\rho = 1/2, 2/3$, and $5/6$ respectively. One hundred simulations of waiting times were performed for each parameter pair and each of the sample size, $n = 50$, 100, 250, 500 and 1000. We set $W_1 = 0$ in these simulations. The sample means, standard deviations, and root mean squared errors were computed for the simulation estimators. The asymptotic standard deviations quoted in Table 1, are $\sigma_1/\sqrt{n}$ and $\sigma_2/\sqrt{n}$, where σ_1^2 and σ_2^2 are the diagonal elements of the inverse of the limiting Fisher information matrix J derived in Appendix A.2.

Tables 2 and 3 present the simulation results for $M/E_k/1$ for $k = 2$ and 4 respectively. Set of values for (λ, μ) were selected to yield the traffic intensities $1/2, 2/3$ and $5/6$. The asymptotic standard deviation was not computed since the expressions for the limiting Fisher information were too unwieldy. The sample standard deviation should be a good approximation, however, as in the M/M/1 case.

One possible source of bias may be due to the partial informations, $\{W_i\}$, used for the MLE's. It is to be noted that the bias in the ML estimates appears large even for $n = 1000$. This is, in any case, a typical phenomenon for MLE's based on dependent data. However, both the standard deviation and the mean squared error (MSE) decrease as n becomes large.

Table 1: Maximum Likelihood Simulation Study of the M/M/1 queue

True Parameters: $\lambda = 1, \mu = 2$

Order of entries: λ, μ

	n=50	n=100	n=250	n=500	n=1000
Sample Mean	1.481, 2.486	1.439, 2.311	1.423, 2.309	1.401, 2.285	1.407, 2.269
Sample Stan. Dev	0.450, 0.696	0.303, 0.347	0.192, 0.234	0.110, 0.188	0.088, 0.104
Sample Root MSE	0.659, 0.848	0.533, 0.466	0.464, 0.388	0.416, 0.341	0.417, 0.289
Asymptotic Stan. Dev.	0.258, 0.432	0.183, 0.306	0.115, 0.193	0.082, 0.137	0.058, 0.097

(Table 1) True Parameters: $\lambda = 2, \mu = 3$

Order of entries: λ, μ

	n=50	n=100	n=250	n=500	n=1000
Sample Mean	2.723, 3.408	2.671, 3.488	2.609, 3.368	2.625, 3.330	2.550, 3.285
Sample Stan. Dev.	0.692, 0.731	0.464, 0.570	0.278, 0.334	0.170, 0.251	0.120, 0.168
Sample Root MSE	1.001, 0.837	0.816, 0.751	0.669, 0.497	0.648, 0.414	0.563, 0.331
Asymptotic Stan. Dev.	0.438, 0.585	0.310, 0.414	0.196, 0.262	0.139, 0.185	0.098, 0.131

(Table 1) True Parameters: $\lambda = 5, \mu = 6$

Order of entries: λ, μ

	n=50	n=100	n=250	n=500	n=1000
Sample Mean	6.214, 6.609	6.035, 6.435	5.790, 6.386	5.706, 6.309	5.711, 6.290
Sample Stan. Dev.	1.290, 1.342	0.790, 0.985	0.475, 0.552	0.295, 0.389	0.199, 0.265
Sample Root MSE	1.772, 1.474	1.302, 1.077	0.922, 0.674	0.765, 0.497	0.739, 0.393
Asymptotic Stan. Dev.	0.963, 1.091	0.681, 0.772	0.431, 0.488	0.305, 0.345	0.215, 0.244

Table 2: Maximum Likelihood Simulation Study of the M/E_2/1 queue

True Parameters: $\lambda = 1, \mu = 4$

Order of entries: λ, μ

	n=50	n=100	n=250	n=500	n=1000
Sample Mean	1.417, 4.590	1.386, 4.447	1.411, 4.513	1.393, 4.516	1.337, 4.504
Sample Stan. Dev.	0.403, 0.921	0.246, 0.579	0.183, 0.418	0.114, 0.309	0.079, 0.216
Sample Root MSE	0.580, 1.094	0.457, 0.731	0.450, 0.661	0.410, 0.602	0.375, 0.548

(Table 2) True Parameters: $\lambda = 2, \mu = 6$

Order of entries: λ, μ

	n=50	n=100	n=250	n=500	n=1000
Sample Mean	2.663, 6.916	2.675, 6.608	2.633, 6.640	2.596, 6.628	2.579, 6.552
Sample Stan. Dev.	0.596, 1.156	0.381, 0.894	0.213, 0.545	0.177, 0.387	0.125, 0.273
Sample Root MSE	0.892, 1.475	0.775, 1.205	0.667, 0.841	0.622, 0.738	0.592, 0.616

(Table 2) True Parameters: $\lambda = 5, \mu = 12$

Order of entries: λ, μ

	n=50	n=100	n=250	n=500	n=1000
Sample Mean	6.021, 13.079,	6.046, 12.915	5.864, 12.917	5.779, 12.643	5.760, 12.616
Sample Stan. Dev.	1.147, 2.182	0.739, 1.405	0.466, 1.056	0.297, 0.668	0.219, 0.483
Sample Root MSE	1.535, 2.434	1.281, 1.677	0.981, 1.398	0.833, 0.927	0.791, 0.783

Table 3: Maximum Likelihood Simulation Study of the $M/E_4/1$ queue

True Parameters: $\lambda = 1, \mu = 8$

Order of entries: λ, μ

	n=50	n=100	n=250	n=500	n=1000
Sample Mean	1.382, 9.094	1.362, 8.903	1.324, 8.758	1.338, 8.840	1.326, 8.795
Sample Stan. Dev.	0.303, 1.735	0.247, 0.872	0.146, 0.646	0.106, 0.443	0.728, 0.299
Sample Root MSE	0.487, 2.051	0.438, 1.255	0.356, 0.996	0.355, 0.950	0.334, 0.849

(Table 3) True Parameters: $\lambda = 2, \mu = 12$

Order of entries: λ, μ

	n=50	n=100	n=250	n=500	n=1000
Sample Mean	2.649, 13.531	2.540, 13.100	2.588, 13.049	2.544, 12.969	2.538, 12.907
Sample Stan. Dev.	0.522, 2.224	0.353, 1.261	0.233, 0.866	0.170, 0.605	0.109, 0.357
Sample Root MSE	0.832, 2.670	0.645, 1.673	0.633, 1.360	0.570, 1.143	0.549, 0.974

(Table 3) True Parameters: $\lambda = 5, \mu = 24$

Order of entries: λ, μ

	n=50	n=100	n=250	n=500	n=1000
Sample Mean	6.052, 26.094	5.926, 25.412	5.800, 25.561	5.782, 25.290	5.785, 25.145
Sample Stan. Dev.	1.043, 3.230	0.606, 2.430	0.402, 1.567	0.263, 1.103	0.190, 0.794
Sample Root MSE	1.481, 3.908	1.106, 2.811	0.895, 2.012	0.825, 1.697	0.808, 1.393

Appendix

A.1. Regularity conditions for the consistency and asymptotic normality of the ML estimates

(Theorems 3.1 and 3.2)

(C.1). The set $A = \{x : f_x(x;\theta) > 0\}$ does not depend on θ.

(C.2) The partial derivatives with respect to θ, up to the third order of $f_x(x;\theta)$ exist and are continuous in θ. The partial derivatives with respect to θ, up to the third order of $\alpha(w;\theta)$, exist and are continuous in θ.

(C.3) The Fisher information matrix $J(\theta)$, defined in (3.9), is positive definite.

(C.3) For each $w \geq 0$, and $i, j = 1, 2, ..., r$, there exists a neighborhood N of θ, such that

$$\int_0^\infty \sup_{\theta \in N} \left| \frac{\partial f_x(y-w)}{\partial \theta_i} \right| dy < \infty,$$

$$\int_0^\infty \sup_{\theta \in N} \left| \frac{\partial^2 f_x(y-w)}{\partial \theta_i \partial \theta_j} \right| dy < \infty.$$

(C.4) For each $w \geq 0$, and $i, j, k = 1, 2, ..., r$,

$$K(w) = \int_0^\infty \sup_{\theta \in N} \left| \frac{\partial^3 ln f_x(y-w)}{\partial \theta_i \partial \theta_j \partial \theta_k} \right| dy < \infty,$$

and $E(K(W)) < \infty$, where W is the random variable corresponding to the stationary distribution of $\{W_t\}$.

It is easy to check that (C.1) to (C.4) above imply regularity conditions 1.1 and 1.2 of Billingsley (1961). Theorems 3.1 and 3.2 therefore follow readily from Billingsley's (1961) Theorems 2.1 and 2.2 respectively. It can be verified that the above conditions are satisfied for the M/M/1 and M/E_k/1 queues.

A.2 Derivation of the limiting Fisher information matrix for M/M/1 queue

We shall derive explicit expressions for the conditional and the limiting Fisher information matrices $\xi_t(\theta)$ and $J(\theta)$ defined in (3.6) and (3.9) respectively.

After tedious but straightforward computations, starting with (3.6), (4.1) and (4.2), we find that the conditional Fisher information matrix is given by

$$E(U_{t+1}U'_{t+1}|W_t) = \xi_t(\theta) = M_1 - M_2 e^{-\lambda W_t},$$

where

$$M_1 = \begin{pmatrix} \frac{\mu(\lambda\mu+\lambda^2+(\lambda+\mu)^2)}{\lambda^2(\lambda+\mu)^3} & -\frac{\mu+\lambda}{(\lambda+\mu)^3} \\ -\frac{\mu+\lambda}{(\lambda+\mu)^3} & \frac{\lambda(\lambda\mu+\mu^2+(\lambda+\mu)^2)}{\mu^3(\lambda+\mu)^3} \end{pmatrix},$$

and

$$M_2 = \begin{pmatrix} \frac{\mu}{\lambda^2(\lambda+\mu)} & 0 \\ 0 & 0 \end{pmatrix}.$$

The limiting Fisher informaiton matrix can now be obtained by integrating $\xi_t(\theta)$ against the limiting distribution in (4.3):

$$\begin{aligned} J(\theta) = E(\xi_t(\theta)) &= M_1 - M_2[(1 - \frac{\lambda}{\mu}) + \int_0^\infty e^{-\lambda w}(\frac{\lambda}{\mu})(\mu - \lambda)e^{-(\mu-\lambda)w} dw] \\ &= M_1 - M_2[1 - \frac{\lambda}{\mu} + (\frac{\lambda}{\mu})(\mu - \lambda) \int_0^\infty e^{-\mu w} dw] \\ &= M_1 - M_2(1 - \rho^2), \text{where } \rho = \lambda/\mu. \end{aligned}$$

Acknowledgements

I. V. Basawa's work was supported partially by the National Science Foundation and the Office of Naval Research.

References

Basawa, I. V. and B. R. Bhat (1992). Sequential inference for single server queues. In *Queueing and Related Models*, Eds. U. N. Bhat and I. V. Basawa, Oxford University Press, Oxford, pp 325-336.

Basawa, I. V. and N. U. Prabhu (1981). Estimation in single server queues. *Nav. Res. Logist. Quart.* 28, 475-487.

Basawa, I. V. and N. U. Prabhu (1988). Large sample inference from single server queues. *Queueing Systems* 3, 289-306.

Bhat, U. N. and S. S. Rao (1987). Statistical analysis of queueing systems. *Queueing Systems* 1, 217-247.

Billingsley, P. (1961). *Statistical Inference for Markov Processes.* Chicago Univ. Press, Chicago.

Prabhu, N. U. (1965). *Queues and Inventories.* Wiley, New York.

Prabhu, N. U. (1980). *Stochastic Storage Processes.* Springer-Verlag, New York.

Thiruvaiyaru, D. and I. V. Basawa (1992). Empirical Bayes estimation for queueing systems and networks. *Queueing Systems* 11, 179-202.