Math. Meth. Oper. Res. (2006) 63: 371–384
DOI 10.1007/s00186-005-0047-z

ORIGINAL ARTICLE

Tsung-Yin Wang · Jau-Chuan Ke
Kuo-Hsiung Wang · Siu-Chuen Ho

Maximum likelihood estimates and confidence intervals of an M/M/R queue with heterogeneous servers

Received: 15 March 2003 / Accepted: 15 July 2004 / Published online: 17 January 2006

Abstract This paper studies maximum likelihood estimates as well as confidence intervals of an M/M/R queue with heterogeneous servers under steady-state conditions. We derive the maximum likelihood estimates of the mean arrival rate and the three unequal mean service rates for an M/M/3 queue with heterogeneous servers, and then extend the results to an M/M/R queue with heterogeneous servers. We also develop the confidence interval formula for the parameter ρ, the probability of empty system P_0, and the expected number of customers in the system $E[N]$, of an M/M/R queue with heterogeneous servers.

Keywords Confidence interval · Heterogeneous servers · Maximum likelihood estimate · Queue

1 Introduction

In this paper, we study both point estimations and confidence intervals of an M/M/R queue with ordered heterogeneous servers under steady-state conditions. It is assumed that customers arrive following a Poisson process with rate λ and with service times according to an exponential distribution with R unequal mean service rates μ_i, $(i = 1, 2, \ldots R)$, where $\mu_1 > \mu_2 > \cdots > \mu_R$. We assume the following: (i) Arriving customers at the servers form a single waiting line and are served in the order of their arrivals; (ii) Each server may serve only one customer

T.-Y. Wang
Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu 300, Taiwan

J.-C. Ke
Department of statistics, National Taichung Institute of Technology Taichung 404, Taiwan

K.-H. Wang (✉) · S.-C. Ho
Department of Applied Mathematics, National Chung-Hsing University, Taichung, 402, Taiwan
E-mail: khwang@amath.nchu.edu.tw

at a time; (iii) If all servers are idle, the first customer in the waiting line goes to the fastest server; (iv) If part of the servers are idle, the first customer goes to the faster server; (v) If all servers are busy, the first customer waits until any one server becomes free; (vi) The arrival process and the service process are independent.

The statistical analysis of queueing systems are rarely found in the literature and the work of related systems in the past mainly concentrates on only one server or two servers. A landmark paper in parameter estimations for queueing models was first proposed by Clarke (1957), who derived maximum likelihood estimates for the arrival and service parameters of an M/M/1 queue. Lilliefors (1966) investigated the confidence intervals for the M/M/1, M/E_k/1 and M/M/2 queues. Basawa and Prabhu (1981) examined moment estimates as well as maximum likelihood estimates for a G/G/1 queue. An illustration of statistical estimation technique applied to the queueing problems can be found in Rubin and Robson (1990). Jain (1991) obtained maximum likelihood estimates of the parameters for an M/E_k/1 queue. Basawa et al. (1996) studied maximum likelihood estimates of the parameters in the single-server queue using waiting time data. Rodrigues and Leite (1998) used Bayesian analysis to investigate the confidence intervals of an M/M/1 queue. Maximum likelihood estimates and confidence intervals in an M/M/2 queue with heterogeneous servers were derived by Dave and Shah (1980) and Jain and Templeton (1991), respectively. Abou-E1-Ata and Hariri (1995) developed point estimations and confidence intervals of an M/M/2/N queue with balking and heterogeneous servers. Recently, an overview of literature on the statistical analysis of several queueing systems was provided by Dshalalow (1997).

The main purpose of this paper is to derive maximum likelihood estimates and confidence intervals of an M/M/R queue with ordered heterogeneous servers. In section 2, we derive the maximum likelihood estimates of parameters for an M/M/3 queue with heterogeneous servers and consider two special cases. Similar procedure is used and extended to an M/M/R queue with heterogeneous servers and the results are presented in section 3. Two special cases are also considered. Finally, section 4 presents the confidence interval formula for the parameter ρ, the probability of empty system P_0, and the expected number of customers in the system $E[N]$, of an M/M/R queue with heterogeneous servers.

2 M/M/3 queue with heterogeneous servers

In this section, our objective is to develop the maximum likelihood estimates of the mean arrival rate λ and the three unequal mean service rates μ_1, μ_2 and μ_3, where $\mu_1 > \mu_2 > \mu_3$ of the M/M/3 queue with heterogeneous servers.

In steady-state, the following notations are used.

$P_0 \equiv$ probability that there are no customers in the system,

$P_n \equiv$ probability that there are n customers in the system,

where $n = 1, 2, 3, \ldots$.

Steady-state equations for an M/M/3 queue with heterogeneous servers are given by:

$$\lambda P_0 = \mu_1 P_1, \tag{1}$$

$$(\lambda + \mu_1) P_1 = (\mu_1 + \mu_2) P_2 + \lambda P_0, \tag{2}$$

$$(\lambda + \mu_1 + \mu_2)P_2 = (\mu_1 + \mu_2 + \mu_3)P_3 + \lambda P_1, \tag{3}$$

$$(\lambda + \mu_1 + \mu_2 + \mu_3)P_n = (\mu_1 + \mu_2 + \mu_3)P_{n+1} + \lambda P_{n-1}, \quad n \geq 3. \tag{4}$$

Solving recursively, analytic solutions P_n are derived in the following:

$$P_n = \begin{cases} \frac{\lambda}{\mu_1} P_0, & n = 1 \\ \frac{\lambda^2}{\mu_1(\mu_1+\mu_2)} P_0, & n = 2 \\ \frac{\lambda^3}{\mu_1(\mu_1+\mu_2)(\mu_1+\mu_2+\mu_3)} P_0, & n \geq 3 \end{cases} \tag{5}$$

where

$$P_0 = \left[1 + \frac{\lambda}{\mu_1} + \frac{\lambda^2}{\mu_1(\mu_1 + \mu_2)(1 - \rho)}\right]^{-1}, \tag{6}$$

and

$$\rho = \frac{\lambda}{\mu_1 + \mu_2 + \mu_3}.$$

Since the queue is in steady-state, so ρ must be less than 1 or equivalently $\lambda < \mu_1 + \mu_2 + \mu_3$.

2.1 Likelihood function and maximum likelihood estimates

At time $t = 0$, the queue has just started operation with m_0 customers present. Let T denote a fixed sufficiently large interval of time during which the queue is being observed. During T, we assume that there are N_a number of arrivals to the queue and N_d number of departures from the queue. Following Dave and Shah (1980), we observe during T that:

- $T_e \equiv$ amount of time during which three servers are idle;
- $T_{B_1} \equiv$ amount of time during which only the fastest server is busy;
- $T_{B_2} \equiv$ amount of time during which both the fastest server and faster server are busy;
- $T_{B_3} \equiv$ amount of time during which three servers are busy;
- $N_e \equiv$ number of arrivals to an empty queue when three servers are idle (transitions E_0 to E_1);
- $N_{B_1} \equiv$ number of arrivals to a partially busy queue when the fastest server is busy (transitions E_1 to E_2);
- $N_{B_2} \equiv$ number of arrivals to a partially busy queue time when the fastest server and faster server are busy (transitions E_2 to E_3);
- $N_{B_3} \equiv$ number of arrivals to a completely busy queue when three servers are busy (transitions E_i to E_{i+1}, $i \geq 3$);
- $N_{D_1} \equiv$ number of departures from a partially busy queue when the fastest server is busy (transitions E_1 to E_0);
- $N_{D_2} \equiv$ number of departures from a partially busy queue when the fastest server and faster server are busy (transitions E_2 to E_1);

N_{D_3} $\equiv$ number of departures from a completely busy queue when three servers are busy (transitions E_i to $E_{i-1},\ i \geq 3$).

Obviously,

$$T = T_e + T_{B_1} + T_{B_2} + T_{B_3},$$
$$N_a = N_e + N_{B_1} + N_{B_2} + N_{B_3},$$
$$N_d = N_{D_1} + N_{D_2} + N_{D_3}.$$

Following Abou-E1-Ata and Hariri (1995), the corresponding likelihood function can be broken down into the following five basic components:

(i) The probability that there are initial m_0 customers in the system can be obtained from (6) yielding $P_{m_0} = \frac{\lambda^3}{\mu_1(\mu_1+\mu_2)(\mu_1+\mu_2+\mu_3)}\rho^{m_0-3}P_0,\ m_0 \geq 3$;

(ii) The probability density function of N_e transitions (E_0 to E_1) occurring during time T_e is given by $f_1 = \lambda^{N_e}e^{-\lambda T_e}$;

(iii) The probability density function of N_{B_1} transitions (E_1 to E_2) occurring and N_{D_1} transitions (E_1 to E_0) occurring during time T_{B_1} is given by $f_2 = \left(\lambda^{N_{B_1}}e^{-\lambda T_{B_1}}\right)\left(\mu_1^{N_{D_1}}e^{-\mu_1 T_{B_1}}\right)$;

(iv) The probability density function of N_{B_2} transitions (E_2 to E_3) occurring and N_{D_2} transitions (E_2 to E_1) occurring during time T_{B_2} is given by $f_3 = \left(\lambda^{N_{B_2}}e^{-\lambda T_{B_2}}\right)\left[(\mu_1+\mu_2)^{N_{D_2}}e^{-(\mu_1+\mu_2)T_{B_2}}\right]$;

(v) The probability density function of N_{B_3} transitions (E_i to $E_{i+1},\ i \geq 3$) occurring and N_{D_3} transitions (E_i to $E_{i-1},\ i \geq 3$) occurring during time T_{B_3} is given by $f_4 = \left(\lambda^{N_{B_3}}e^{-\lambda T_{B_3}}\right)\left(\mu^{N_{D_3}}e^{-\mu T_{B_3}}\right)$, where $\mu = \mu_1 + \mu_2 + \mu_3$.

Since the random variables m_0, T_{B_i} , N_{B_i} and N_{D_i} $(i = 1, 2, 3)$ are mutually independent, the likelihood function is given by

$$\begin{aligned} L_1(\lambda, \mu_1, \mu_2, \mu_3) = {} & \lambda^{m_0+N_a}e^{-\lambda T}\mu^{N_{D_3}-m_0+3}\mu_1^{N_{D_1}} \\ & \times(\mu_1+\mu_2)^{N_{D_2}}e^{-\mu_1 T_{B_1}-(\mu_1+\mu_2)T_{B_2}-\mu T_{B_3}} \\ & \times\left[\frac{P_0}{\mu_1(\mu_1+\mu_2)\mu}\right]. \end{aligned} \tag{7}$$

Since the queue is in steady-state, the probability P_{m_0} can be neglected. Taking the logarithm of (7), it implies that

$$\begin{aligned} lnL_1 = lnL_1(\lambda, \mu_1, \mu_2, \mu_3) = {} & N_a ln\lambda - \lambda T + N_{D_1} ln\mu_1 + N_{D_2} ln(\mu_1+\mu_2) \\ & + N_{D_3} ln\mu - \mu_1 T_{B_1} - (\mu_1+\mu_2)T_{B_2} - \mu T_{B_3}. \end{aligned} \tag{8}$$

Differentiating (8) with respect to the parameters λ, μ_1, μ_2 and μ_3, respectively, we finally obtain

$$\frac{\partial lnL_1}{\partial\lambda}|_{\lambda=\hat{\lambda}} = \frac{N_a}{\hat{\lambda}} - T = 0, \tag{9}$$

$$\frac{\partial lnL_1}{\partial\mu_1}|_{\mu_i=\hat{\mu}_i} = \frac{N_{D_1}}{\hat{\mu}_1} + \frac{N_{D_2}}{\hat{\mu}_1+\hat{\mu}_2} + \frac{N_{D_3}}{\hat{\mu}} - (T_{B_1}+T_{B_2}+T_{B_3}) = 0, \tag{10}$$

$$\frac{\partial lnL_1}{\partial\mu_2}|_{\mu_i=\hat{\mu}_i} = \frac{N_{D_2}}{\hat{\mu}_1+\hat{\mu}_2} + \frac{N_{D_3}}{\hat{\mu}} - (T_{B_2}+T_{B_3}) = 0, \tag{11}$$

$$\frac{\partial lnL_1}{\partial\mu_3}|_{\mu_i=\hat{\mu}_i} = \frac{N_{D_3}}{\hat{\mu}} - T_{B_3} = 0. \tag{12}$$

From (9), we have

$$\hat{\lambda} = \frac{N_a}{T}. \tag{13}$$

Subtracting (11) from (10), we get

$$\hat{\mu}_1 = \frac{N_{D_1}}{T_{B_1}}. \tag{14}$$

Subtracting (12) from (11) and using (14), we get

$$\hat{\mu}_2 = \frac{N_{D_2}}{T_{B_2}} - \frac{N_{D_1}}{T_{B_1}}. \tag{15}$$

We obtain from (12)

$$\hat{\mu} = \hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3 = \frac{N_{D_3}}{T_{B_3}}. \tag{16}$$

It implies from (14)–(16)

$$\hat{\mu}_3 = \frac{N_{D_3}}{T_{B_3}} - \frac{N_{D_2}}{T_{B_2}}. \tag{17}$$

Thus, the maximum likelihood estimates of λ , μ_1, μ_2 and μ_3 are given in (13), (14), (15), and (17), respectively.

2.2 Special cases

We consider the following two special cases:
Case 1:
Let $\mu_i - \mu_{i+1} = \delta_i,\ (i = 1, 2)$, we have

$$\mu_1 = \mu_2 + \delta_1 = \mu_3 + \delta_1 + \delta_2, \tag{18}$$

$$\mu_2 = \mu_3 + \delta_2. \tag{19}$$

Substituting (18)–(19) into (8) and using $\Omega = T_{B_1} + 2T_{B_2} + 3T_{B_3}$, $\Delta_1 = \delta_1 + \delta_2$ and $\Delta_2 = \delta_1 + 2\delta_2$, yields the following log-likelihood function:

$$\begin{aligned} lnL_2 = lnL_2(\lambda, \mu_3) = N_a ln\lambda - \lambda T + N_{D_1} ln(\mu_3 + \Delta_1) + N_{D_2} ln(2\mu_3 + \Delta_2) + \\ + N_{D_3} ln(3\mu_3 + \Delta_2) - \mu_3 T_{B_1} - 2\mu_3 T_{B_2} - 3\mu_3 T_{B_3} \end{aligned} \tag{20}$$

Using a derivation analogous to that of the previous section, we get the estimates of λ and μ_3 as follows:

$$\hat{\lambda} = \frac{N_a}{T}, \tag{21}$$

and

$$\begin{aligned} \hat{\mu}_3^3 & + \Big(\frac{6\Delta_1 + 5\Delta_2}{6} - \frac{N_d}{\Omega}\Big)\hat{\mu}_3^2 \\ & + \Big(\frac{5\Delta_1\Delta_2 + \Delta_2^2}{6} - \frac{5\Delta_2 N_{D_1} + 2(3\Delta_1 + \Delta_2)N_{D_2} + 3(2\Delta_1 + \Delta_2)N_{D_3}}{6\Omega}\Big)\hat{\mu}_3 \\ & + \frac{\Delta_1\Delta_2^2}{6} - \frac{\Delta_2^2 N_{D_1} + 2\Delta_1\Delta_2 N_{D_2} + 3\Delta_1\Delta_2 N_{D_3}}{6\Omega} = 0. \end{aligned} \tag{22}$$

It should be noted that the positive real value of $\hat{\mu}_3$ should be taken in order to give the maximum log-likelihood function, and then $\hat{\mu}_1$ and $\hat{\mu}_2$ can be obtained from the expressions (18)–(19).

If $\delta_1 = \delta_2 = 0$, the estimates of μ_1, μ_2 and μ_3 are given by

$$\hat{\mu}_1 = \hat{\mu}_2 = \hat{\mu}_3 = \frac{N_d}{T_{B_1} + 2T_{B_2} + 3T_{B_3}}. \tag{23}$$

If $\delta_1 = 0$ and $\delta_2 \neq 0$, we have

$$\begin{aligned} \hat{\mu}_3^3 + \Big(\frac{8\delta_2}{3} - \frac{N_d}{\Omega}\Big)\hat{\mu}_3^2 + \delta_2\Big(\frac{7\delta_2}{3} - \frac{2N_d}{\Omega} + \frac{N_{D_1} + N_{D_2}}{3\Omega}\Big)\hat{\mu}_3 + \\ \delta_2^2\Big(\frac{2\delta_2}{3} - \frac{N_d}{\Omega} + \frac{N_{D_1} + N_{D_2}}{3\Omega}\Big) = 0. \end{aligned} \tag{24}$$

The positive real value of $\hat{\mu}_3$ should be taken in order to give the maximum log-likelihood function, and then $\hat{\mu}_1$ and $\hat{\mu}_2$ can be obtained from the expressions (18)–(19).

If $\delta_1 \neq 0$ and $\delta_2 = 0$, we get

$$\hat{\mu}_3^3 + \Big(\frac{11\delta_1}{6} - \frac{N_d}{\Omega}\Big)\hat{\mu}_3^2 + \delta_1\Big(\delta_1 - \frac{3N_d}{2\Omega} + \frac{4N_{D_1} + N_{D_2}}{6\Omega}\Big)\hat{\mu}_3 + \delta_1^2\Big(\frac{\delta_1}{6} - \frac{N_d}{2\Omega} + \frac{2N_{D_1} + N_{D_2}}{6\Omega}\Big) = 0. \quad (25)$$

The positive real value of $\hat{\mu}_3$ should be taken in order to give the maximum log-likelihood function, and then $\hat{\mu}_1$ and $\hat{\mu}_2$ can be obtained from the expressions (18)–(19).

If $\delta_1 = \delta_2 = \delta \neq 0$, we obtain

$$\hat{\mu}_3^3 + \Big(\frac{9\delta}{2} - \frac{N_d}{\Omega}\Big)\hat{\mu}_3^2 + \delta\Big(\frac{13\delta}{2} - \frac{7N_d}{2\Omega} + \frac{2N_{D_1} + N_{D_2}}{2\Omega}\Big)\hat{\mu}_3 + \delta^2\Big(3\delta - \frac{3N_d}{\Omega} + \frac{3N_{D_1} + 2N_{D_2}}{2\Omega}\Big) = 0. \quad (26)$$

The positive real value of $\hat{\mu}_3$ should be taken in order to give the maximum log-likelihood function, and then $\hat{\mu}_1$ and $\hat{\mu}_2$ can be obtained from the expressions (18)–(19).

Case 2:

Let $\frac{\mu_{i+1}}{\mu_i} = \theta_i < 1,\ (i = 1, 2)$, we have

$$\mu_2 = \theta_1\mu_1,$$

$$\mu_3 = \theta_2\mu_2 = \theta_1\theta_2\mu_1.$$

Using a derivation similar to that of *Case 1*, the estimates of μ_1, μ_2 and μ_3 are given by

$$\hat{\mu}_1 = \frac{N_d}{T_{B_1} + (1+\theta_1)T_{B_2} + (1+\theta_1+\theta_1\theta_2)T_{B_3}}, \quad (27)$$

$$\hat{\mu}_2 = \frac{\theta_1 N_d}{T_{B_1} + (1+\theta_1)T_{B_2} + (1+\theta_1+\theta_1\theta_2)T_{B_3}}, \quad (28)$$

$$\hat{\mu}_3 = \frac{\theta_1\theta_2 N_d}{T_{B_1} + (1+\theta_1)T_{B_2} + (1+\theta_1+\theta_1\theta_2)T_{B_3}}. \quad (29)$$

If $\theta_1 = \theta_2 = 1$, we have

$$\hat{\mu}_i = \frac{N_d}{T_{B_1} + 2T_{B_2} + 3T_{B_3}}, \quad i = 1, 2, 3. \quad (30)$$

If $\theta_1 = \theta_2 = \theta < 1$, we obtain

$$\hat{\mu}_i = \frac{\theta^{i-1} N_d}{T_{B_1} + (1+\theta)T_{B_2} + (1+\theta+\theta^2)T_{B_3}}, \quad i = 1, 2, 3. \quad (31)$$

3 M/M/R queue with heterogeneous servers

In this section, we will derive the maximum likelihood estimates of the mean arrival rate λ and the R unequal mean service rates μ_i, $(i = 1, 2, \ldots R)$, where $\mu_1 > \mu_2 > \cdots > \mu_R$ of the M/M/R queue with heterogeneous servers.

As in the M/M/3 queue, the steady-state equations for an M/M/R queue with heterogeneous servers are as follows:

$$\lambda P_0 = \mu_1 P_1, \tag{32}$$

$$(\lambda + \sum_{j=1}^{n} \mu_j) P_n = \sum_{j=1}^{n+1} \mu_j P_{n+1} + \lambda P_{n-1}, 1 \le n \le R-1 \tag{33}$$

$$(\lambda + \mu) P_n = \mu P_{n+1} + \lambda P_{n-1}, \quad n \ge R \tag{34}$$

where $\mu = \sum_{j=1}^{R} \mu_j$.

Solving recursively for this set of linear equations, we have

$$P_n = \begin{cases} \dfrac{\lambda^n}{\prod_{k=1}^{n} \sum_{j=1}^{k} \mu_j} P_0, & 1 \le n \le R-1 \\ \dfrac{\lambda^R}{\prod_{k=1}^{R} \sum_{j=1}^{k} \mu_j} \rho^{n-R} P_0, & n \ge R \end{cases} \tag{35}$$

and

$$P_0 = \left[1 + \sum_{n=1}^{R-1} \frac{\lambda^n}{\prod_{k=1}^{n} \sum_{j=1}^{k} \mu_j} + \frac{\lambda^R}{(1-\rho) \prod_{k=1}^{R} \sum_{j=1}^{k} \mu_j} \right]^{-1}. \tag{36}$$

Let $E[N]$ denote the expected number of customers in the system. From (34), we finally obtain

$$E[N] = \sum_{n=1}^{\infty} n P_n = \left[\sum_{n=1}^{R-1} \frac{n \lambda^n}{\prod_{k=1}^{n} \sum_{j=1}^{k} \mu_j} + \frac{\lambda^R}{\prod_{k=1}^{R} \sum_{j=1}^{k} \mu_j} \frac{R - R\rho + \rho}{(1-\rho)^2} \right] P_0, \tag{37}$$

where $\rho = \lambda / \sum_{j=1}^{R} \mu_j$.

3.1 Likelihood function and maximum likelihood estimates

At time $t = 0$, the queue has just started operation with m_0 customers present. Let T denote a fixed sufficiently large interval of time during which the queue is being observed. During T, we assume that there are N_a number of arrivals to the queue and N_d number of departures from the queue. During T, we observe the following:

$T_e \equiv$ amount of time during which all servers are idle;

$T_{B_1} \equiv$ amount of time during which only the fastest server is busy;

$T_{B_i} \equiv$ amount of time during which i faster servers are busy where $i = 2, 3, \ldots, R-1$;

$T_{B_R} \equiv$ amount of time during which all servers are busy;
$N_e \equiv$ number of arrivals to an empty queue when all servers are idle (transitions E_0 to E_1);

$N_{B_1} \equiv$ number of arrivals to a partially busy queue when only the fastest server is busy (transitions E_1 to E_2);

$N_{B_i} \equiv$ number of arrivals to a partially busy queue time when i servers are busy (transitions E_i to E_{i+1}) where $i = 2, 3, \ldots, R-1$;

$N_{B_R} \equiv$ number of arrivals to a completely busy queue when all servers are busy (transitions E_i to $E_{i+1},\ i \geq R$);

$N_{D_1} \equiv$ number of departures from a partially busy queue when only the fastest server is busy (transitions E_1 to E_0);

$N_{D_i} \equiv$ number of departures from a partially busy queue when i faster servers are busy (transitions E_i to E_{i-1}) where $i = 2, 3, \ldots, R-1$;

$N_{D_R} \equiv$ number of departures from a completely busy queue when all servers are busy (transitions E_i to $E_{i-1},\ i \geq R$).

It is clear that

$$T = T_e + \sum_{j=1}^{R} T_{B_j},$$

$$N_a = N_e + \sum_{j=1}^{R} N_{B_j},$$

$$N_d = \sum_{j=1}^{R} N_{D_j}.$$

Since the queue is in steady-state, the probability P_{m_0} can be neglected. As in the M/M/3 queue, the corresponding likelihood function can be broken down into the following three basic components:

(i) The probability density function of N_e transitions (E_0 to E_1) occurring during time T_e is given by $\lambda^{N_e} e^{-\lambda T_e}$;

(ii) The probability density function of N_{B_i} transitions (E_i to $E_{i+1},\ 1 \leq i \leq R-1$) occurring and N_{D_i} transitions (E_i to $E_{i-1},\ 1 \leq i \leq R-1$) occurring during time T_{B_i} is given by $\left(\lambda^{N_{B_i}} e^{-\lambda T_{B_i}}\right)\left[\left(\sum_{j=1}^{i} \mu_j\right)^{N_{D_i}} e^{-\left(\sum_{j=1}^{i} \mu_j\right) T_{B_i}}\right]$;

(iii) The probability density function of N_{B_R} transitions (E_i to $E_{i+1},\ i \geq R$) occurring and N_{D_R} transitions (E_i to $E_{i-1},\ i \geq R$) occurring during time T_{B_R} is given by $\left(\lambda^{N_{B_R}} e^{-\lambda T_{B_R}}\right)\left(\mu^{N_{D_R}} e^{-\mu T_{B_R}}\right)$.

Using a derivation similar to that of the previous section, the log-likelihood function is given by

$$
\begin{aligned}
lnL_3 = lnL_3(\lambda, \mu_1, \mu_2, \ldots, \mu_R) = N_a ln\lambda - \lambda T + \sum_{k=1}^{R-1} N_{D_k} ln\Big(\sum_{j=1}^{k} \mu_j\Big) \\
+ N_{D_R} ln\mu - \sum_{k=1}^{R-1}\sum_{j=1}^{k} \mu_j T_{B_k} - \mu T_{B_R}. \quad (38)
\end{aligned}
$$

Using (38) and after some algebraic manipulations, we obtain the maximum likelihood estimates of λ and μ_i ($i = 1, 2, \ldots, R$)

$$\frac{\hat{\lambda} = N_a}{T}, \tag{39}$$

$$\frac{\hat{\mu}_1 = N_{D_1}}{T_{B_1}}, \tag{40}$$

and

$$\hat{\mu}_i = \frac{N_{D_i}}{T_{B_i}} - \frac{N_{D_{i-1}}}{T_{B_{i-1}}}, \quad \text{for } 2 \leq i \leq R. \tag{41}$$

From (40)–(41), we get

$$\hat{\mu} = \sum_{i=1}^{R} \hat{\mu}_i = \frac{N_{D_R}}{T_{B_R}}. \tag{42}$$

Thus, we get the maximum likelihood estimate of ρ

$$\frac{\hat{\rho} = N_a T_{B_R}}{N_{D_R} T}. \tag{43}$$

3.2 Special cases

Let $\frac{\mu_{i+1}}{\mu_i} = \theta_i,\ (i = 1, 2, \ldots, R-1)$, we have

$$\mu_{i+1} = \mu_1 \prod_{k=1}^{i} \theta_k, \quad i = 1, 2, \ldots, R-1.$$

Using the procedure in section 2.2, we get the estimates of μ_i

$$\hat{\mu}_i = \frac{N_d \prod_{k=1}^{i-1} \theta_k}{\sum_{k=1}^{R-1} \left(1 + \sum_{j=2}^{k} \prod_{l=1}^{k-1} \theta_l\right) T_{B_k} + \left(1 + \sum_{j=2}^{R} \prod_{l=1}^{j-1} \theta_l\right) T_{B_R}}, \quad 1 \le i \le R \tag{44}$$

where the $\sum_{n=a}^{b}$ notation indicates the term is 0 when $a > b$ and the $\prod_{n=a}^{b}$ notation indicates the term is 1 when $a > b$.

Two special cases are considered in the following:

Case 1:

$\theta_i = 1$, ($i = 1, 2, \dots, R-1$), (44) can be simplified to

$$\hat{\mu}_1 = \hat{\mu}_2 = \dots = \hat{\mu}_R = \frac{N_d}{\sum_{k=1}^{R} k T_{B_k}}. \tag{45}$$

Case 2:

$\theta_i = \theta < 1$, ($i = 1, 2, \dots, R-1$), (44) can be simplified to

$$\hat{\mu}_i = \frac{(1-\theta)\theta^{i-1} N_d}{\sum_{k=1}^{R} (1-\theta^k) T_{B_k}}, \quad 1 \le i \le R. \tag{46}$$

4 Confidence interval formula for ρ, P_0 and $E[N]$

In this section, we will develop the confidence interval formula for ρ, P_0 and $E[N]$ of an M/M/R queue with heterogeneous servers. To achieve our aim, we first establish the following results.

For a simple birth–death process to an M/M/R queueing system with heterogeneous servers, it follows from Appendix that

$$E[N] = -\rho \frac{\partial ln P_0}{\partial \rho} \ge 0, \tag{47}$$

and the variance

$$Var[N] = \rho \frac{\partial E[N]}{\partial \rho} \ge 0. \tag{48}$$

Next, applying the approach by Lilliefors (1966), we have the $(1-\alpha) \times 100\%$ lower and upper confidence limits L_ρ and U_ρ of ρ as follows;

$$L_\rho = \hat{\rho} F_{1-\alpha/2}(2N_{\mathrm{a}}, 2N_{\mathrm{d}}), \tag{49}$$

and

$$U_\rho = \hat{\rho} F_{\alpha/2}(2N_{\mathrm{a}}, 2N_{\mathrm{d}}), \tag{50}$$

where $\hat{\rho}$ is given by (43).

One observes from (47) that P_0 is a monotonic decreasing function of ρ. Hence, the $(1-\alpha)\times 100\%$ lower and upper confidence limits, L_{P_0} and U_{P_0} of P_0 can be obtained through (36) and (49)–(50). That is,

$$L_{P_0} = P_0|_{\rho=U_\rho}, \tag{51}$$

and

$$U_{P_0} = P_0|_{\rho=L_\rho}. \tag{52}$$

From (48), we observe that $E[N]$ is a monotonic increasing function of ρ. Thus, the $(1-\alpha)\times 100\%$ lower and upper confidence limits, $L_{E[N]}$ and $U_{E[N]}$ of $E[N]$, can be obtained through (37) and (49)–(50). That is,

$$L_{E[N]} = E[N]|_{\rho=L_\rho}, \tag{53}$$

and

$$U_{E[N]} = E[N]|_{\rho=U_\rho}. \tag{54}$$

5 Conclusions

In this paper, we have developed the maximum likelihood estimates for the arrival and service parameters of the M/M/3 queue and M/M/R queue with heterogeneous servers, respectively. We also have demonstrated that both results for the maximum likelihood estimates of the parameters are the functions of the observations only which are consistent with the results of Dave and Shah (1980). The estimates of λ and μ_i $(i = 1, 2, \ldots, R)$ can be easily computed, due to the fact that these observations can easily be made for an M/M/R queue with heterogeneous servers. Next, we have derived the confidence interval formula for ρ, P_0 and $E[N]$ of an M/M/R queue with heterogeneous servers.

Appendix

Derivations of (47) and (48)

Taking the logarithm of (36) and differentiating it with respect to λ, we finally get

$$\frac{\partial ln P_0}{\partial\lambda} = -\Big[\sum_{n=1}^{R-1}\frac{n\lambda^{n-1}}{\prod_{k=1}^{n}\sum_{j=1}^{k}\mu_j} + \frac{\lambda^{R-1}}{\prod_{k=1}^{R}\sum_{j=1}^{k}\mu_j}\cdot\frac{R-R\rho+\rho}{(1-\rho)^2}\Big]P_0. \tag{A–1}$$

Multiplying (A–1) by $-\lambda$ and using (37), we obtain

$$E[N] = -\lambda\frac{\partial ln P_0}{\partial\lambda}. \tag{A–2}$$

Since $\frac{\partial lnP_0}{\partial \rho} = \frac{\partial lnP_0}{\partial \lambda} / \frac{\partial \rho}{\partial \lambda}$, we have $\frac{\partial lnP_0}{\partial \rho} = \mu \frac{\partial lnP_0}{\partial \lambda}$. We obviously get

$$\rho \frac{\partial lnP_0}{\partial \rho} = \lambda \frac{\partial lnP_0}{\partial \lambda}. \tag{A–3}$$

It follows from (A–2) and (A–3) that

$$E[N] = -\rho \frac{\partial lnP_0}{\partial \rho} \geq 0, \tag{A–4}$$

which is the result given in (47).
Since $\frac{\partial lnP_0}{\partial \lambda} = \frac{\partial P_0}{\partial \lambda} \frac{\partial lnP_0}{\partial P_0} = \frac{\partial P_0}{\partial \lambda} \frac{1}{P_0}$, we have $\frac{\partial P_0}{\partial \lambda} = P_0 \frac{\partial lnP_0}{\partial \lambda}$. It yields from (A–2) that

$$\frac{\partial P_0}{\partial \lambda} = -\frac{P_0}{\lambda} E[N]. \tag{A–5}$$

Let

$$Q = \sum_{n=1}^{R-1} \frac{n\lambda^n}{\prod_{k=1}^n \sum_{j=1}^k \mu_j} + \frac{\lambda^R}{\prod_{k=1}^R \sum_{j=1}^k \mu_j} [\frac{R - R\rho + \rho}{(1-\rho)^2}].$$

Thus

$$Q = -\frac{\lambda}{P_0} \frac{\partial lnP_0}{\partial \lambda} = -\frac{\lambda}{P_0^2} \frac{\partial P_0}{\partial \lambda}. \tag{A–6}$$

From (A–5) and (A–6), we obtain $E[N] = QP_0$. Differentiating $E[N]$ with respect to λ yields

$$\frac{\partial E[N]}{\partial \lambda} = P_0 \frac{\partial Q}{\partial \lambda} + Q \frac{\partial P_0}{\partial \lambda}, \tag{A–7}$$

where

$$\frac{\partial Q}{\partial \lambda} = \sum_{n=1}^{R-1} \frac{n^2 \lambda^{n-1}}{\prod_{k=1}^n \sum_{j=1}^k \mu_j} + \frac{\lambda^{R-1}}{\prod_{k=1}^R \sum_{j=1}^k \mu_j} \Big[\frac{R^2}{1-\rho} + \frac{(2R+1)\rho}{(1-\rho)^2} + \frac{2\rho^2}{(1-\rho)^3}\Big].$$

After doing some algebraic manipulations in (35), we obtain

$$\begin{aligned}
E[N^2] &= \sum_{n=1}^{\infty} n^2 P_n \\
&= \Big[\sum_{n=1}^{R-1} \frac{n^2 \lambda^n}{\prod_{k=1}^n \sum_{j=1}^k \mu_j} + \frac{\lambda^R \rho^{-R}}{\prod_{k=1}^R \sum_{j=1}^k \mu_j} \Big(\sum_{n=R}^{\infty} n^2 \rho^n\Big)\Big] P_0 \\
&= \Big[\sum_{n=1}^{R-1} \frac{n^2 \lambda^n}{\prod_{k=1}^n \sum_{j=1}^k \mu_j} + \frac{\lambda^R}{\prod_{k=1}^R \sum_{j=1}^k \mu_j} \big(\frac{R^2}{1-\rho} + \frac{(2R+1)\rho}{(1-\rho)^2} + \frac{2\rho^2}{(1-\rho)^3}\big)\Big] P_0 \\
&= \lambda P_0 \frac{\partial Q}{\partial \lambda}.
\end{aligned} \tag{A–8}$$

Since $\frac{\partial E[N]}{\partial \rho} = \frac{\partial E[N]}{\partial \lambda}\frac{\partial \lambda}{\partial \rho}$, we have $\frac{\partial E[N]}{\partial \rho} = \mu \frac{\partial E[N]}{\partial \lambda}$.
From (A–7), we get

$$\rho \frac{\partial E[N]}{\partial \rho} = \lambda \frac{\partial E[N]}{\partial \lambda} = \lambda P_0 \frac{\partial Q}{\partial \lambda} + \lambda Q \frac{\partial P_0}{\partial \lambda}. \tag{A–9}$$

It follows from (A–5) and (A–8) that

$$\rho \frac{\partial E[N]}{\partial \rho} = E[N^2] - (E[N])^2 = Var[N] \geq 0, \tag{A–10}$$

which is the result given in (48).

References

Abou-E1-Ata MO, Hariri AMA (1995) Point estimation and confidence intervals of the M/M/2/N queue with balking and heterogeneity. Am J. Math. Manage Sci **15** 35–55

Basawa IV, Bhat UN, Lund R (1996) Maximum likelihood estimation for single server queues from waiting time data. Queueing Syst Theory Appl **24** 155–167

Basawa IV, Prabhu NU (1981) Estimation in single server queues. Naval Res. Logist. Quart **28** 475–487

Clarke AB (1957) Maximum likelihood estimates in a simple queue. Ann. Math. Stat. **28** 1036–1040

Dave U, Shah YK (1980) Maximum likelihood estimates in an M/M/2 queue with heterogeneous servers. J Oper Res Soc **31** 423–426

Dshalalow JH (1997) Frontiers in queueing: Models and Applications in Science and Engineering. CRC Press, Inc

Jain S (1991) Estimation in M/E$_k$ /1 queueing systems. Commun Stat Theory Meth **20** 1871–1879

Jain S, Templeton JGC (1991) Confidence interval for M/M/2 queue with heterogeneous servers. Oper Res. Lett **10** 99–101

Lilliefors HW (1966) Some confidence intervals for queues. Oper Res **14** 723–727

Rodrigues J, Leite JG (1998) A note on Bayesian analysis in M/M/1 queues derived from confidence intervals. Statistics **31** 35–42

Rubin G, Robson DS (1990) A single server queue with random arrivals and balking: confidence interval estimation. Queueing Syst Theory Appl **7** 283–306