Confidence limits for expected waiting time of two queuing models

VSS Yadavalli[*]   K Adendorff[†]   G Erasmus[†]   P Chandrasekhar[‡]   SP Deepa[‡]

Abstract

A maximum likelihood estimator (MLE), a consistent asymptotically normal (CAN) estimator and asymptotic confidence limits for the expected waiting time per customer in the queues of $M|M|1|\infty$ and $M|M|1|N$ are obtained.

Key words: Multivariate central limit theorem, CAN estimator, Slutsky theorem

1   Introduction

Parametric estimation is one of the essential tools to understand the random phenomena when using stochastic models. Whenever the systems are fully observable in terms of their basic random components such as inter-arrival times and service times, standard parametric estimation techniques of statistical theory are quite appropriate. Most of the studies on several queueing models are confined to only obtaining expressions for transient or stationary (steady state) solutions and do not consider the associated inference problems. Recently, Bhat (2003) has provided an overview of methods available for estimation, when the information is restricted to the number of customers in the system at some discrete points in time. Narayan Bhat has also described how maximum likelihood estimation is applied directly to the underlying Markov chain in the queue length process in $M|G|1$ and $GI|M|1$. The MLE, CAN and asymptotic confidence limits for the expected waiting time per customer in the queues of $M|M|1|\infty$ and $M|M|1|N$, are obtained in this paper. In the following section, these two models and the expected waiting time per customer for each model are explained briefly.

---

[*]Corresponding author: Department of Industrial and Systems Engineering, University of Pretoria, Pretoria, 0002, South Africa, yadavalli@postino.up.ac.za

[†]Department of Industrial and Systems Engineering, University of Pretoria, Pretoria, 0002, South Africa

[‡]Department of Statistics, Loyola College, Chennai 600 034, India

2   System Descriptions

Model I $(M|M|1) : (FCFS|\infty|\infty)$ queue
It can be readily seen [1] that the difference-differential equations governing $M|M|1$ are given by

$$
\begin{aligned}
p'_n(t) &= \lambda p_{n-1}(t) - (\lambda + \mu)p_n(t) + \mu p_{n+1}(t), \quad n = 1, 2, 3, \ldots \\
p'_0(t) &= -\lambda p_0(t) + \mu p_1(t), \qquad\qquad\qquad (n = 0).
\end{aligned}
$$

As $t \to \infty$, the steady state solution can be proved to exist, when $\lambda < \mu$. Assuming that $p'_n(t) \to 0$ and $p_n(t) \to p_n$ as $t \to \infty$, for $n = 0, 1, 2, \ldots$, we have

$$
\begin{aligned}
-\lambda p_0 + \mu p_1 &= 0, \qquad (n = 0) \\
\lambda p_{n-1} - (\lambda + \mu)p_n + \mu p_{n+1} &= 0, \qquad n = 1, 2, 3, \ldots
\end{aligned}
$$

Solving these difference-differential equations, we have

$$
p_n = (1 - \rho)\rho^n, \qquad n = 0, 1, 2, \ldots \tag{1}
$$

where $\rho = \frac{\lambda}{\mu} < 1$. Clearly (1) corresponds to the probability mass function of the geometric distribution and it can easily be shown that the expected waiting time per customer in the queue is given by

$$
{}_1W_Q = \frac{\lambda}{\mu(\mu - \lambda)}. \tag{2}
$$

Model II $(M|M|1) : (GD|N|\infty)$ queue
This model is essentially the same as Model I, except that the maximum number of customers in the system is limited to N (maximum queue length is $N-1$) [1]. The steady state equations for this model are given by

$$
\begin{aligned}
-\rho p_0 + p_1 &= 0, \qquad (n = 0) \\
\rho p_{n-1} - (\rho + 1)p_n + p_{n+1} &= 0, \qquad n = 1, 2, 3, \ldots, N-1 \\
\rho p_{N-1} - p_N &= 0, \qquad (n = N).
\end{aligned}
$$

The solution of these difference-differential equations is given by

$$
p_n = \frac{(1 - \rho)}{(1 - \rho^{N+1})}\rho^n, \qquad n = 0, 1, 2, \ldots, N.
$$

The expected number in the system is given by

$$
L_s = \frac{\rho\left\{1 - (N+1)\rho^N + N\rho^{N+1}\right\}}{(1 - \rho)(1 - \rho^{N+1})}, \qquad \rho \neq 1.
$$

Since there is a limit on the queue length and some customers are lost, it is necessary to compute the effective arrival rate λ_{eff}, which is given by $\lambda_{\text{eff}} = \lambda(1 - p_N)$. Further, it can be shown that the expected number of customers in the queue is

$$
L_Q = L_s - \frac{\lambda_{\text{eff}}}{\mu} = \frac{\rho^2\left[1 - N\rho^{N-1} + (N-1)\rho^N\right]}{(1 - \rho)(1 - \rho^{N+1})}.
$$

Hence, the expected waiting time per customer in the queue is given by

$$_2W_Q = \frac{L_Q}{\lambda_{\text{eff}}} = \frac{\lambda\left[(\mu^N - \lambda^N) - N\lambda^{N-1}(\mu - \lambda)\right]}{\mu(\mu - \lambda)(\mu^N - \lambda^N)}. \tag{3}$$

3 MLE and CAN estimator for the expected waiting time

3.1 ML Estimator

Let $X_{i1}, X_{i2}, \ldots, X_{in}$ (with $i = 1, 2$ representing Models I and II) be random samples of size n, each randomly drawn from different exponential inter-arrival time populations with the parameter λ. Also, let $Y_{i1}, Y_{i2}, \ldots, Y_{in}$ (with $i = 1, 2$ representing Models I and II) be random samples of size n, each drawn from different exponential service time populations with the parameter μ. It is clear that $E(\bar{X}_i) = \frac{1}{\lambda}$ and $E(\bar{Y}_i) = \frac{1}{\mu}$, where $\bar{X}_i$ and $\bar{Y}_i$, $i = 1, 2$, are the sample means of inter-arrival times and service times respectively corresponding to Models I and II. It can be shown that $\bar{X}_i$ and $\bar{Y}_i$ (with $i = 1, 2$ representing Models I and II) are the MLEs of $\frac{1}{\lambda}$ and $\frac{1}{\mu}$ respectively. Let $\theta_1 = \frac{1}{\lambda}$ and $\theta_2 = \frac{1}{\mu}$ respectively.

Model I
The average waiting time per customer in the queue, given in (2), reduces to

$$_1W_Q = \frac{\theta_2^2}{(\theta_1 - \theta_2)}$$

and hence the MLE of W_Q is given by

$$_1\hat{W}_Q = \frac{\bar{Y}_1^2}{(\bar{X}_1 - \bar{Y}_1)}. \tag{4}$$

Model II
The average waiting time per customer in the queue, given in (3), reduces to

$$_2W_Q = \frac{\theta_2^2[(\theta_1^N - \theta_2^N) + N\theta_2^{N-1}(\theta_2 - \theta_1)]}{(\theta_2 - \theta_1)(\theta_2^N - \theta_1^N)}. \tag{5}$$

and hence the MLE of W_Q is given by

$$_2\hat{W}_Q = \frac{\bar{Y}_2^2[(\bar{X}_2^N - \bar{Y}_2^N) + N\bar{Y}_2^{N-1}(\bar{Y}_2 - \bar{X}_2)]}{(\bar{Y}_2 - \bar{X}_2)(\bar{Y}_2^N - \bar{X}_2^N)} \tag{6}$$

It may be noted that $_i\hat{W}_Q$, given in (4) and (6), are real valued functions in $\bar{X}_i$ and $\bar{Y}_i$, $i = 1, 2$, which are also differentiable. Consider the following application of multivariate central limit theorem [3].

3.2 Application of multivariate central limit theorem

Suppose $T_1', T_2', T_3', \ldots$ are independent and identically distributed k-dimensional random variables such that $T_n' = (T_{1n}, T_{2n}, T_{3n}, \ldots, T_{kn})$, $n = 1, 2, 3, \ldots$ having the first and second order moments $E(T_n) = \mu$ and $\text{Var}(T_n) = \Sigma$. Define the sequence of random variables $\bar{T}'_n = (\bar{T}_{1n}, \bar{T}_{2n}, \bar{T}_{3n}, \ldots, \bar{T}_{kn})$, $n = 1, 2, 3, \ldots$ where $\bar{T}_{in} = \frac{\sum_{j=1}^{n} T_{ij}}{n}$, $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, n$. Then, $\sqrt{n}(\bar{T}_n - \mu) \xrightarrow{d} N(0, \Sigma)$ as $n \to \infty$.

3.3   CAN Estimator

Model I

By applying the multivariate central limit theorem [4], it readily follows that $\sqrt{n}[(\bar{X}_1, \bar{Y}_1) - (\theta_1, \theta_2)] \xrightarrow{d} N(0, \Sigma)$ as $n \to \infty$, where the dispersion matrix $\Sigma = ((\sigma_{ij}))$ is given by $\Sigma = diag(\theta_1^2, \theta_2^2)$. Again, from Rao [3], we have $\sqrt{n}(_1\hat{W}_Q - {}_1W_Q) \xrightarrow{d} N(0, {}_1\sigma^2(\theta))$, as $n \to \infty$, where $\theta = (\theta_1, \theta_2)$ and

$$_1\sigma^2(\theta) = \sum_{i=1}^2 \left(\frac{\partial_1 W_Q}{\partial \theta_i}\right)^2 \sigma_{ii} = \frac{\theta_2^2[\theta_1^2 + \theta_2^2(2\theta_1 - \theta_2)^2]}{(\theta_1 - \theta_2)^4}. \tag{7}$$

Hence, $_1\hat{W}_Q$ is a CAN estimator of $_1W_Q$. There are several methods for generating CAN estimators — the Method of Moments and the Method of Maximum likelihood are commonly used to generate such estimators [4].

Model II

As in Model I, here too, we have $\sqrt{n}(_2\hat{W}_Q - {}_2W_Q) \xrightarrow{d} N(0, {}_2\sigma^2(\theta))$, as $n \to \infty$, where $\theta = (\theta_1, \theta_2)$, and where $_2W_Q$ and $_2\hat{W}_Q$ are given by (5) and (6) respectively. Further, $_2\sigma^2(\theta)$ is computed from the partial derivatives $\left(\frac{\partial_2 W_Q}{\partial \theta_i}\right)$, $i = 1, 2$ as in Model I. Thus $_2\hat{W}_Q$ is a CAN estimator of $_2W_Q$.

4   Confidence limits for expected waiting time

Let $_i\sigma^2(\hat{\theta})$ be the estimator of $_i\sigma^2(\theta)$ (with $i = 1, 2$ representing Models I and II) obtained by replacing θ by a consistent estimator $_i\hat{\theta} = (\bar{X}_i, \bar{Y}_i)$, $i = 1, 2$. Let $_i\hat{\sigma}^2 = {}_i\sigma^2(\hat{\theta})$. Since $_i\sigma^2(\theta)$ is a continuous function of θ, $_i\hat{\sigma}^2$ is a consistent estimator of $_i\sigma^2(\theta)$, i.e., $_i\hat{\sigma}^2 \xrightarrow{P} {}_i\sigma^2(\theta)$ as $n \to \infty$, $i = 1, 2$. By the Slutsky theorem, we have

$$\frac{\sqrt{n}(_i\hat{W}_Q - W_Q)}{_i\hat{\sigma}} \xrightarrow{d} N(0, 1),$$

i.e., $\Pr\left[-k_{\frac{\alpha}{2}} < \frac{\sqrt{n}(_i\hat{W}_Q - {}_iW_Q)}{_i\hat{\sigma}} < k_{\frac{\alpha}{2}}\right] = (1 - \alpha)$, where $k_{\frac{\alpha}{2}}$ is obtained from normal tables. Hence, a $100(1 - \alpha)\%$ asymptotic confidence interval for $_iW_Q$ is given by

$$_i\hat{W}_Q \pm k_{\frac{\alpha}{2}} \frac{_i\hat{\sigma}}{\sqrt{n}}, \qquad i = 1, 2.$$

5   Numerical Results

As is to be expected, W_q is an increasing function of λ, and a decreasing function of μ, for both $M|M|1|\infty$ and $M|M|1|N$ queuing systems [See Tables 1 and 2].

$\mu \setminus \lambda$	0.04	0.06	0.08	0.1
0.01	(8.289842877 : 8.376823789)	(3.316926722 : 3.349739944)	(1.777150118 : 1.794278453)	(1.105858754 : 1.116363468)
0.02	(24.83223528 : 25.16776472)	(8.286718407 : 8.37994826)	(4.144904413 : 4.18842892)	(2.487402722 : 2.512597278)
0.03	(74.16104529 : 75.83895471)	(16.55477386 : 16.77855947)	(7.456343866 : 7.543656134)	(4.262400695 : 4.309027877)

Table 1: $M|M|1|\infty : FCFS$ with 99% confidence interval and sample size of 20.

	$\mu \setminus \lambda$	0.04	0.06	0.08	0.1
	0.01	(8.331269275 : 8.334920554)	(3.331916366 : 3.334744788)	(1.784518941 : 1.786909398)	(1.110057009 : 1.112165194)
N=10	**0.02**	(24.75248955 : 24.7587519)	(8.328275966 : 8.332745576)	(4.164721818 : 4.168373097)	(2.498408635 : 2.501570885)
	0.03	(60.07542472 : 60.08759256)	(16.50061597 : 16.50687832)	(7.490679322 : 7.495571505)	(4.283054105 : 4.287193479)
	$\mu \setminus \lambda$	0.04	0.06	0.08	0.1
	0.01	(8.331507591 : 8.335159075)	(3.33191912 : 3.334747547)	(1.784519057 : 1.786909514)	(1.110057019 : 1.112165204)
N=20	**0.02**	(24.99636102 : 25.0026853)	(8.33109717 : 8.335569306)	(4.164840925 : 4.168492408)	(2.498418861 : 2.501581139)
	0.03	(73.40387298 : 73.41482651)	(16.66318663 : 16.66951092)	(7.497549755 : 7.502448733)	(4.283644082 : 4.287784475)
	$\mu \setminus \lambda$	0.04	0.06	0.08	0.1
	0.01	(8.331507591 : 8.335159075)	(3.33191912 : 3.334747547)	(1.784519057 : 1.786909514)	(1.110057019 : 1.112165204)
N=40	**0.02**	(24.99683772 : 25.00316228)	(8.331097265 : 8.335569401)	(4.164840925 : 4.168492409)	(2.498418861 : 2.501581139)
	0.03	(74.98446701 : 74.99541962)	(16.66350439 : 16.66982894)	(7.49755051 : 7.50244949)	(4.283644089 : 4.287784482)

Table 2: $M|M|1|N : FCFS$ with 99% confidence interval and sample size of 20.

6 Acknowledgements

The authors wish to thank the referees for their valuable comments.

References

[1] TAHA HA, 1976, *Operations research*, Mac Millan Publishing Co. Inc., New York (NY).

[2] BHAT NU, 2003, *Parameter estimation in $M|G|1$ and $GI|M|1$ queues using queue length data*, pp 96-107 in SRINIVASAN SK & VIJAYAKUMAR A (EDS.), *Stochastic point processes*, Narosa Publishing House, New Delhi.

[3] RAO RC, 1974, *Linear statistical inference and its applications*, Wiley Eastern Pvt. Ltd, New Delhi.

[4] SINHA SK, 1986, *Reliability and life testing*, Wiley Eastern Ltd., New Delhi.