International Journal of Computational and Theoretical Statistics
ISSN (2210-1519)
Int. J. Comp. Theo. Stat. 5, No. 2 (Nov-2018)
http://dx.doi.org/10.12785/ijcds/070606

# Estimation of Traffic Intensity Vector of a Two Stage Open Queueing Network Models with Feedback

**Suresh B. Pathare[1] and Vinayak K. Gedam[2]**

[1] *MITCOM, MIT Art, Design and Technology University, Pune, India*
[2] *Department of Statistics & Centre for Advanced Studies, S. P. Pune University, Pune, India*

*Received: March 11, 2018; Revised July 10, 2018 Accepted Sept 7, 2018; Published: November 1, 2018*

**Abstract**: We propose a consistent and asymptotically normal estimator for traffic intensity vector $\underline{\rho}$ of a two stage open queueing network model with feedback with no assumption of arrival and service time distribution. Using this estimator and its estimated covariance matrix $A$, a $100(1-\alpha)\%$ confidence region for traffic intensity vector is constructed. Standard bootstrap, Bayesian bootstrap and percentile bootstrap are applied to develop the confidence regions. Simulation study was undertaken to evaluate performance of the four confidence regions. The performances are assessed in terms of their coverage area percentage, average area and relative coverage area. Further calibration technique is used to improve the coverage area percentages of confidence regions.

**Keywords:** Traffic intensity vector, Coverage percentage area, Relative coverage area, Relative average area, Calibration, Feedback.

## 1 INTRODUCTION

Consider the network model of a computer system with feedback in which a job may return to previously visited nodes as shown in Figure 1.

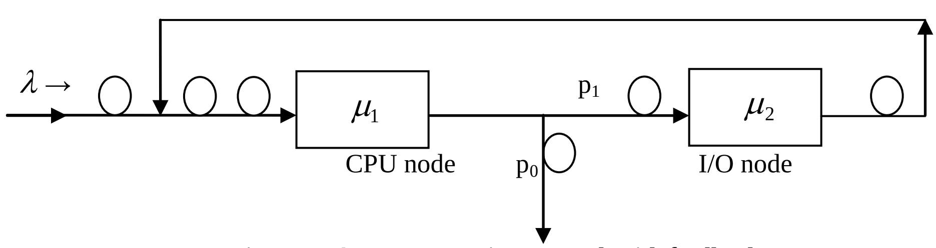


**Figure 1. An open queueing network with feedback.**

The system consists of two nodes i.e. CPU node and I/O node with service rates $\mu_1$ and $\mu_2$ respectively. Arrivals to the CPU node occur either from the outside at the rate $\lambda$ or from the I/O node at rate $\lambda_1$. The total arrival rate to the CPU node is, therefore:

$$\lambda_0 = \lambda + \lambda_1. \tag{1}$$

After service completion at CPU node, the job proceeds to the I/O node with probability $p_1$ and departs from the system with probability $p_0$ where $p_0 = 1 - p_1$. Therefore the average arrival rate to I/O node is given by:

$$\lambda_1 = \lambda_0 p_1. \tag{2}$$

*E-mail address: sureshpathare23@gmail.com, vinaykagedam90901@gmail.com*

Using equation (2) in equation (1) we get

$$\lambda_0 = \lambda + \lambda_0 p_1$$

$$\lambda_0 = \frac{\lambda}{1 - p_1} = \frac{\lambda}{p_0} \tag{3}$$

Now using equation (3) in equation (2) we have

$$\lambda_1 = \frac{\lambda p_1}{p_0} \tag{4}$$

Thus the traffic intensity vector of the CPU node and I/O node is given by

$$\underline{\rho}^F = (\rho_1^F, \rho_2^F)' = \left( \frac{\lambda}{p_0 \mu_1}, \frac{p_1 \mu_1}{p_0 \mu_2} \right) \tag{5}$$

where $\rho_1^F$ and $\rho_2^F$ can be interpreted as expected number of arrivals per mean service time. The condition for stability of the system is both are less than unity. An equivalent queueing network model without feedback shown in Figure 2 is given in [1].

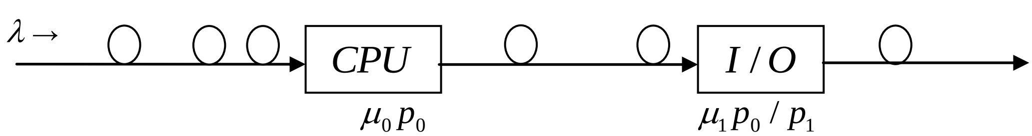


**Figure 2. An equivalent queueing network without feedback.**

The service rates of the CPU and I/O node of an equivalent network in Figure 2 are $p_0 \mu_1$ and $\frac{p_0 \mu_2}{p_1}$ respectively.

Let $S_0$ and $S_1$ denote the total service time required on CPU node and I/O node. Then $E(S_0) = \frac{1}{p_0 \mu_1}$ and $E(S_1) = \frac{p_1}{p_0 \mu_2}$, the expected value of the total service time required on the CPU node and I/O node.

The product form solution also applies to open network of Markovian queues with feedback and Jackson's theorem states that each node behaves like an independent queue is given in [2]. Queueing networks with arrival process that can depend on the state of the system and closed queueing networks with exponential servers is shown in [3]. Basic properties of queueing networks introduced in [4]. Open queueing networks are useful in studying the behavior of computer communication networks [5]. Reference [6] considered the problem of Maximum likelihood estimation for Jackson networks with Poisson arrival and exponential service time at each node. The problem of maximum likelihood estimation for the parameters in a Jackson type queueing network with the arrival at each node following renewal process and service time distribution being arbitrary is discussed in [7]. Reference [8] stated that the departure time distribution from an $M/M/c/\infty$ queue is identical to the interarrival time distribution ,namely , exponential with mean $1/\lambda$ ; hence all stations are independent.

Reference [9] developed and proposed the bootstrap to estimate the sampling distribution of any statistic. For necessary background on bootstrap technique, we refer ([9]-[13]). Besides the standard bootstrap technique, [14] presented the Bayesian bootstrap technique of resampling. A nonparametric approach of intensity for a queueing system with distribution free inter-arrival and service times is proposed by [15]. The statistical inference of parameters in queueing network problems are rarely found in the literature and the work of related problems in the past mainly concentrated on only parametric statistical inference, in which the distribution of population has a known form. However, in practice the functional form of the distribution is seldom known. Consistent and asymptotically normal(CAN) estimator for intensities of two stage queueing network model with feedback with distribution-free inter-arrival and service times is proposed in [16]. Using this estimator and its estimated variance asymptotic confidence interval of intensities is constructed. Also bootstrap approaches such as Standard bootstrap, Bayesian bootstrap, Percentile bootstrap and Bias-corrected and accelerated bootstrap are applied to develop the confidence intervals of intensities. Reference [17] constructed an approximate calibrated CAN, Exact-$t$, Variance-stabilized Bootstrap-$t$, and some bootstrap confidence intervals for intensity parameters of a two stage open queueing network with feedback with distribution-free interval and service times. Calibration technique is used to construct confidence intervals for intensity parameters of a two-stage open queueing network with distribution-free interval and service times in [18]. Numerical simulation study is conducted to demonstrate performances of the calibrated confidence intervals.

Consistent and asymptotically normal estimator for intensity parameters for a queueing network with distribution-free inter-arrival and service times is proposed in [19]. Using this estimator and its estimated variance, asymptotic confidence interval for intensities is constructed. Bootstrap approaches are applied to develop the confidence intervals for intensity parameters. Data based recurrence relation is used to compute a sequence of response time in [20]. The sample means from those response times, denoted by $\hat{r}_1$ and $\hat{r}_2$ are used to estimate true mean response time $r_1$ and $r_2$. Confidence intervals for mean response times $r_1$ and $r_2$ are constructed. Various confidence intervals for mean response times of an open queueing network model with feedback using the calibration approach are constructed in [21]. Data-based recurrence relation is used to compute a sequence of response times. Sample means from those response times are used to estimate true mean response times.

The paper is organized as follows: The calibration technique is given in section 2. Statistical inference and estimation of traffic intensity vector is discussed in section 3. In section 4 different confidence regions for traffic intensity vector are constructed. Section 5 is devoted to evaluate the performance of four confidence regions in terms of simulation analysis. The performances of the confidence regions are assessed in terms of their coverage area percentage, average area and relative coverage area. Calibration technique is used to improve the coverage percentage area of confidence regions. Finally some concluding remarks are given in section 6.

## 2 CALIBRATION TECHNIQUE

The actual coverage of confidence region is rarely equal to the desired level. Hence to improve the coverage accuracy of confidence region we use calibration technique. First use bootstrap to estimate the true coverage of confidence region and the region is then adjusted by comparing with the target nominal level. The general theory of calibration is reviewed in [22], following ideas of ([23]- [ 26]). The bootstrap calibration technique was introduced by [27]. To illustrate, first find $\hat{\gamma}$ for the confidence region for $\underline{\rho}$ with $\gamma$. Then set

$$\begin{aligned} \gamma_1 &= \frac{\gamma^2}{\hat{\gamma}}, \qquad\qquad \textit{if} \quad \hat{\gamma} \geq \gamma \\ &= \gamma + \frac{(1-\gamma)(\gamma-\hat{\gamma})}{(1-\hat{\gamma})}, \quad \textit{if} \quad \hat{\gamma} < \gamma \end{aligned} \tag{6}$$

That is we get the point $(\gamma_1, \gamma)$ by linearly interpolating between

$(i)$ $(0,0)$ *and* $(\gamma, \hat{\gamma})$ *if* $\hat{\gamma} \geq \gamma$

$(ii)$ $(\gamma, \hat{\gamma})$ *and* $(1,1)$ *if* $\hat{\gamma} < \gamma$

Therefore the calibrated confidence region for $\underline{\rho}$ is with $\gamma = \gamma_1$ .

### 3. STATISTICAL INFERENCE OF TRAFFIC INTENSITY VECTOR

Let $(X,Y)$ be nonnegative random variables representing the inter-arrival and service time of CPU node and $(Y,Z)$ be nonnegative random variables representing the inter-arrival and service time of I/O node. Once a job complete service at CPU node, it will proceed to I/O node for further service with probability $p_1$ and departs from the system with probability $p_0$ where $p_0 = 1 - p_1$. The successive service time at both nodes are assumed to be mutually independent and independent of the state of the system. Then the traffic intensity vector $\underline{\rho}$ of CPU and I/O node is defined as follows:

$$\underline{\rho}^F = (\rho_1^F, \rho_2^F)' = \left( \frac{p_0 \mu_Y}{\mu_X} \ , \ \frac{p_0 \mu_Z}{p_1 \mu_Y} \right) \tag{7}$$

where $\mu_X$ and $\mu_Y$ denote the mean inter-arrival time of CPU node and I/O node respectively. Similarly $\mu_Y$ and $\mu_Z$ denote the mean service times of CPU node and I/O node.

### 3.1 ESTIMATION OF TRAFFIC INTENSITY VECTOR

Assume that $(X_1, X_2, \cdots, X_n)$ and $(Y_1, Y_2, \cdots, Y_n)$ are random samples drawn from $X$ and $Y$ respectively. We use $(X_i, Y_i), i = 1,2 \cdots, n$ to represent inter-arrival time and service time for the $i^{th}$ customer of CPU node. Similarly $(Y_1, Y_2, \cdots, Y_n)$ and $(Z_1, Z_2, \cdots, Z_n)$ are random samples drawn from $Y$ and $Z$. We use $(Y_i, Z_i), i = 1,2 \cdots, n$ to represent inter-arrival time and service time for the $i^{th}$ customer of I/O node. Define $\overline{X}$, $\overline{Y}$ and $\overline{Z}$ be the sample means of $X$, $Y$ and $Z$ respectively. According to the Strong Law of Large Numbers, we know that $\overline{X}$, $\overline{Y}$ and $\overline{Z}$ are strongly consistent estimator of $(\mu_X, \mu_Y, \mu_Z)$. Thus strongly consistent estimator of $\underline{\rho}^F$ is given by

$$\underline{\hat{\rho}}^F = (\hat{\rho}_1^F, \hat{\rho}_2^F)' = \left( \frac{p_0 \overline{Y}}{\overline{X}} \ , \ \frac{p_0 \overline{Z}}{p_1 \overline{Y}} \right) \tag{8}$$

The true distributions of $X$, $Y$ and $Z$ are not often known in practical queueing network models with feedback, so the exact distributions of $\underline{\rho}^F$ cannot be derived. Under the assumption that interarrival times and service times are independent, the asymptotic distributions of $\underline{\rho}^F$ can be developed as follows:

**Theorem 3.1** *Suppose* $T_m \xrightarrow{D} N_m(\theta, \Sigma)$. *Let* $g : R^m \to R^k$ *be such that* $\underline{g}(u_1, u_2, \cdots u_m) = (g_1(u_1, u_2, \cdots u_m), g_2(u_1, u_2, \cdots u_m), \cdots, g_n(u_1, u_2, \cdots u_m))$. *Assume* $g_1, g_2, \cdots g_k$ *are totally differentiable functions then* $g(T_m) \xrightarrow{D} N_k(g(\theta), M \Sigma M')$ *[28] where*

$$M = \begin{bmatrix} \frac{\partial g_1}{\partial \theta_1} & \frac{\partial g_1}{\partial \theta_2} & \cdots & \frac{\partial g_1}{\partial \theta_m} \\ \cdots & & & \\ \frac{\partial g_k}{\partial \theta_1} & \frac{\partial g_k}{\partial \theta_2} & \cdots & \frac{\partial g_k}{\partial \theta_m} \end{bmatrix}$$

**Theorem 3.2** *If* $\underline{X}_1, \underline{X}_2, \cdots, \underline{X}_k$ *are independent and identically distributed random vectors with mean* $\underline{\mu} \in R^k$ *and covariance matrix* $\Sigma$ *where* $\Sigma$ *is positive definite and has finite elements, then* $\sqrt{n}(\underline{\overline{X}}_n - \underline{\mu}) \xrightarrow{D} N_k(\underline{0}, \Sigma)$ *where* $\xrightarrow{D}$ *denotes convergence in distribution [28].*

By Theorem 3.2 we have,

$$\sqrt{n}\begin{bmatrix} \overline{X} - \mu_X \\ \overline{Y} - \mu_Y \\ \overline{Z} - \mu_Z \end{bmatrix} \xrightarrow{D} N_3\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \ \sum\right]$$

where,

$$\sum = \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{bmatrix}.$$

$\sigma_x^2$ , $\sigma_y^2$ and $\sigma_z^2$ are variances of $X$ ,$Y$ and $Z$.

Now consider $g : R^3 \to R^2$

$$g(U_1, U_2, U_3) = (g_1(U_1, U_2, U_3), \ \ g_2(U_1, U_2, U_3))$$

$$= \left(\frac{p_0 U_2}{U_1}, \ \frac{p_0 U_3}{p_1 U_2}\right)$$

$$= \left(\hat{\rho}_1^F, \hat{\rho}_2^F\right)$$

By Theorem 3.1 we have,

$$\sqrt{n}\begin{pmatrix} \hat{\rho}_1^F - \rho_1^F \\ \hat{\rho}_2^F - \rho_2^F \end{pmatrix} \xrightarrow{D} N_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ M\Sigma M'\right]$$

where

$$M\Sigma M' = \begin{bmatrix} \frac{p_0^2 U_2^2}{U_1^4}\sigma_X^2 + \frac{p_0^2}{U_1^2}\sigma_Y^2 & -\frac{p_0^2 U_3}{p_1 U_1 U_2^2}\sigma_Y^2 \\ -\frac{p_0^2 U_3}{p_1 U_1 U_2^2}\sigma_Y^2 & \frac{p_0^2 U_3^2}{p_1^2 U_2^4}\sigma_Y^2 + \frac{p_0^2}{p_1^2 U_2^2}\sigma_Z^2 \end{bmatrix}$$

and

$$M = \begin{bmatrix} -\frac{p_0 U_2}{U_1^2} & \frac{p_0}{U_1} & 0 \\ 0 & -\frac{p_0 U_3}{p_1 U_2^2} & \frac{p_0}{p_1 U_2} \end{bmatrix}.$$

$M'$ is transpose of $M$. Let $A^F = M\Sigma M'$. Again by Theorem 3.2 we have,

$$\sqrt{n}\left[\begin{pmatrix}\hat{\rho}_1^F \\ \hat{\rho}_2^F\end{pmatrix} - \begin{pmatrix}\rho_1^F \\ \rho_2^F\end{pmatrix}\right] \xrightarrow{D} N_2\left[\begin{pmatrix}0 \\ 0\end{pmatrix}, \quad A^F\right]$$

That is

$$\sqrt{n}\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right) \xrightarrow{D} N_2\left(\underline{0}\ ,\ A^F\right)$$

where

$$A^F = \begin{bmatrix} \frac{p_0^2 \mu_Y^2}{\mu_X^4}\sigma_X^2 + \frac{p_0^2}{\mu_X^2}\sigma_Y^2 & -\frac{p_0^2 \mu_Z}{p_1 \mu_X \mu_Y^2}\sigma_Y^2 \\ -\frac{p_0^2 \mu_Z}{p_1 \mu_X \mu_Y^2}\sigma_Y^2 & \frac{p_0^2 \mu_Z^2}{p_1^2 \mu_Y^4}\sigma_Y^2 + \frac{p_0^2}{p_1^2 \mu_Y^2}\sigma_Z^2 \end{bmatrix}$$

It follows that $n\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right)' A^{F^{-1}}\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right)$ has a $\chi^2$ - distribution with two degrees of freedom [29].

If $A^F$ is unknown then using the sample estimates of $\mu_X, \mu_Y, \mu_Z, \sigma_X^2, \sigma_Y^2$ and $\sigma_Z^2$ we get estimator $\hat{A}^F$ of $A^F$ as follows:

$$\hat{A}^F = \begin{bmatrix} \frac{p_0^2 \overline{Y}^2}{\overline{X}^4} S_X^2 + \frac{p_0^2}{\overline{X}^2} S_Y^2 & -\frac{p_0^2 \overline{Z}}{p_1 \overline{X}\,\overline{Y}^2} S_Y^2 \\ -\frac{p_0^2 \overline{Z}}{p_1 \overline{X}\,\overline{Y}^2} S_Y^2 & \frac{p_0^2 \overline{Z}^2}{p_1^2 \overline{Y}^4} S_Y^2 + \frac{p_0^2}{p_1^2 \overline{Y}^2} S_Z^2 \end{bmatrix}$$

where

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2, \quad S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2 \quad S_Z^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \overline{Z})^2$$

**Theorem 3.3** *If covariance matrix* $A^F$ *is unknown then* $n\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right)' \hat{A}^{F^{-1}}\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right) \xrightarrow{D} \chi_2^2$.

**Proof:** Consider $T_n = n\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right)' A^{F^{-1}}\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right)$

Let $S_n = n\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right)' \hat{A}^{F^{-1}}\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right)$.

Consider $T_n - S_n = n\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right)'(A^{F^{-1}} - \hat{A}^{F^{-1}})\left(\underline{\hat{\rho}}^F - \underline{\rho}^F\right)$.

Also $\overline{X} \xrightarrow{P} \mu_X$, $\overline{Y} \xrightarrow{P} \mu_Y$, $\overline{Z} \xrightarrow{P} \mu_Z$, $S_X^2 \xrightarrow{P} \sigma_X^2$, $S_Y^2 \xrightarrow{P} \sigma_Y^2$ and $S_Z^2 \xrightarrow{P} \sigma_Z^2$.

Then we have,

$$\frac{\overline{Y}^2}{\overline{X}^4} S_X^2 + \frac{1}{\overline{X}^2} S_Y^2 \xrightarrow{P} \frac{\mu_Y^2}{\mu_X^4}\sigma_X^2 + \frac{1}{\mu_X^2}\sigma_Y^2 \ ;$$

$$\frac{\overline{Z}}{\overline{X}\,\overline{Y}^2} S_Y^2 \xrightarrow{P} \frac{\mu_Z}{\mu_X \mu_Y^2}\sigma_Y^2 \ ;$$

$$\frac{\overline{Z}^2}{\overline{Y}^4}S_Y^2+\frac{1}{\overline{Y}^2}S_Z^2 \xrightarrow{P} \frac{\mu_Z^2}{\mu_Y^4}\sigma_Y^2+\frac{1}{\mu_Y^2}\sigma_Z^2.$$

Thus $\hat{A}^F$ converges component wise in probability to $A^F$ and $\hat{A}^{F^{-1}} \xrightarrow{P} A^{F^{-1}}$.

Hence $T_n - S_n \xrightarrow{P} 0$. But we know that $T_n \xrightarrow{D} \chi_2^2$. Therefore $S_n \xrightarrow{D} \chi_2^2$.

## 4 DIFFERENT CONFIDENCE REGIONS

In this section we construct different confidence regions for traffic intensity vector.

### 4.1 Consistent and Asymptotically Normal Confidence Region

If $A^F$ is unknown, then replace it by $\hat{A}^F$ [29]. By using Theorem 3.3, $100(1-\alpha)\%$ CAN confidence region (CR) for $\underline{\rho}$ is given by,

$$CR=\left\{\underline{\rho}^F \mid n\left(\underline{\rho}^F-\underline{\rho}^F\right)\hat{A}^{F^{-1}}\left(\underline{\rho}^F-\underline{\rho}^F\right)\le \chi_{2,\alpha}^2\right\} \tag{9}$$

### 4.2 Standard Bootstrap Confidence Region

Using standard bootstrap sampling procedure we get estimator of traffic intensity vector are as follows: $\underline{\rho}^{F^*}=\left(\hat{\rho}_1^{F^*},\hat{\rho}_2^{F^*}\right)=\left(\frac{p_0\overline{y}^*}{\overline{x}^*},\frac{p_0\overline{z}^*}{p_1\overline{y}^*}\right)$ where $\overline{x}^*$, $\overline{y}^*$ and $\overline{z}^*$ be the sample means of bootstrap samples $x^*=(x_1^*,x_2^*,\cdots,x_n^*)'$, $y^*=(y_1^*,y_2^*,\cdots,y_n^*)'$ and $z=(z_1,z_2,\cdots,z_n)'$ respectively. Let $\underline{\rho}^{F^*}=\left(\hat{\rho}_1^{F^*},\hat{\rho}_2^{F^*}\right)$ be called a bootstrap estimator of $\underline{\rho}^F$. The above resampling process can be repeated $N$ times. The $N$ bootstrap estimates

$$\underline{\rho}_1^{F^*}=\begin{pmatrix}\hat{\rho}_{11}^{F^*}\\ \hat{\rho}_{21}^{F^*}\end{pmatrix},\underline{\rho}_2^{F^*}=\begin{pmatrix}\hat{\rho}_{12}^{F^*}\\ \hat{\rho}_{22}^{F^*}\end{pmatrix},\cdots,\underline{\rho}_N^{F^*}=\begin{pmatrix}\hat{\rho}_{1N}^{F^*}\\ \hat{\rho}_{2N}^{F^*}\end{pmatrix}$$

can be computed from the bootstrap resamples. Averaging the $N$ bootstrap estimates we get $\underline{\rho}_N^F=\left(\hat{\rho}_{N1}^F,\hat{\rho}_{N2}^F\right)=\left(\frac{1}{N}\sum_{i=1}^N\hat{\rho}_{1i}^{F^*},\frac{1}{N}\sum_{i=1}^N\hat{\rho}_{2i}^{F^*}\right)$ called bootstrap estimate of $\underline{\rho}^F$. And the covariance matrix of $\underline{\rho}^F$ using standard bootstrap can be estimated by

$$\tilde{A}^{F^*}=\begin{bmatrix}\frac{1}{N-1}\sum_{j=1}^n(\hat{\rho}_{1j}^{F^*}-\hat{\rho}_{N1}^F)^2 & \frac{1}{N-1}\sum_{j=1}^n(\hat{\rho}_{1j}^{F^*}-\hat{\rho}_{N1}^F)(\hat{\rho}_{2j}^{F^*}-\hat{\rho}_{N2}^F)\\ \frac{1}{N-1}\sum_{j=1}^n(\hat{\rho}_{1j}^{F^*}-\hat{\rho}_{N1}^F)(\hat{\rho}_{2j}^{F^*}-\hat{\rho}_{N2}^F) & \frac{1}{N-1}\sum_{j=1}^n(\hat{\rho}_{2j}^{F^*}-\hat{\rho}_{N2}^F)^2\end{bmatrix}$$

Using the estimator of $A^F$ as $\tilde{A}^{F^*}$, we construct a $100(1-\alpha)\%$ SB confidence region for $\underline{\rho}$ is as follows:

$$CR=\left\{\underline{\rho}^F \mid n\left(\underline{\rho}^F-\underline{\rho}^F\right)\tilde{A}^{F^{*-1}}\left(\underline{\rho}^F-\underline{\rho}^F\right)\le \chi_{2,\alpha}^2\right\} \tag{10}$$

### 4.3 Bayesian Bootstrap Confidence Region

We use Bayesian bootstrap procedure to construct confidence region for $\underline{\rho}^F$. One BB replication is generated by drawing $n-1$ uniform $(0,1)$ random numbers $r_1, r_2, \cdots, r_{n-1}$, ordering them, and calculating the gaps $u_j = r_{(j)} - r_{(j-1)}, j = 1,2,\cdots,n$, where $r_{(0)} = 0$ and $r_{(n)} = 1$. Then $u' = (u_1, u_2, \cdots, u_n)$ is the vector of probabilities attached to the inter-arrival data $x_1, x_2, \cdots, x_n$. Each BB replication generates a posterior probability for each $x_i$. Considering all BB replications gives the BB distribution of the distribution of $X$ and thus of any parameter of this distribution (Rubin (1981)). Hence for $\mu_x$ (the mean of $X$) in each BB replication we calculate $\mu_x$ as if $u_i$ were the probability that $X = x_i$ that is, we calculate $\bar{x}^{**} = \sum_{i=1}^{n} u_i x_i$. The distribution of the values of $\bar{x}^{**}$ over all BB replications is the BB distribution of $\mu_x$. Similarly we generate a vector of probabilities $v' = (v_1, v_2, \cdots, v_n)$ and $w' = (w_1, w_2, \cdots, w_n)$. These are attached to the data values $y_1, y_2, \cdots, y_n$ and $z_1, z_2, \cdots, z_n$ respectively in a BB replication. We calculate $\bar{y}^{**} = \sum_{i=1}^{n} v_i y_i$ for $\mu_y$ (the mean of $Y$) and $\bar{z}^{**} = \sum_{i=1}^{n} w_i z_i$ for $\mu_z$ (the mean of $Z$).

An estimate of traffic intensity vector is denoted by $\underline{\rho}^{F^{**}} = \left(\hat{\rho}_1^{F^{**}}, \hat{\rho}_2^{F^{**}}\right) = \left(\frac{p_0 \bar{y}^{**}}{\bar{x}^{**}}, \frac{p_0 \bar{z}^{**}}{p_1 \bar{y}^{**}}\right)$ and can be calculated from BB replications. Let $\underline{\rho}^{F^{**}} = \left(\hat{\rho}_1^{F^{**}}, \hat{\rho}_2^{F^{**}}\right)$ be called a Bayesian bootstrap estimator of $\underline{\rho}^F$. The above BB process can be repeated $N$ times. The $N$ BB bootstrap estimates $\underline{\rho}_1^{F^{**}} = \begin{pmatrix} \hat{\rho}_{11}^{F^{**}} \\ \hat{\rho}_{21}^{F^{**}} \end{pmatrix}, \underline{\rho}_2^{F^{**}} = \begin{pmatrix} \hat{\rho}_{12}^{F^{**}} \\ \hat{\rho}_{22}^{F^{**}} \end{pmatrix}, \cdots, \underline{\rho}_N^{F^{**}} = \begin{pmatrix} \hat{\rho}_{1N}^{F^{**}} \\ \hat{\rho}_{2N}^{F^{**}} \end{pmatrix}$ can be computed from the BB replications. Averaging the $N$ BB estimates we get $\underline{\rho}_{BB}^F = \left(\hat{\rho}_{BB_1}^F, \hat{\rho}_{BB_2}^F\right) = \left(\frac{1}{N}\sum_{i=1}^{N} \hat{\rho}_{1i}^{F^{**}}, \frac{1}{N}\sum_{i=1}^{N} \hat{\rho}_{2i}^{F^{**}}\right)$ called BB estimate of $\underline{\rho}^F$ and the covariance matrix of $\underline{\rho}^F$ using Bayesian bootstrap can be estimated by

$$\tilde{A}^{F^{**}} = \begin{bmatrix} \frac{1}{N-1}\sum_{j=1}^{n}(\hat{\rho}_{1j}^{F^{**}} - \hat{\rho}_{BB_1}^F)^2 & \frac{1}{N-1}\sum_{j=1}^{n}(\hat{\rho}_{1j}^{F^{**}} - \hat{\rho}_{BB_1}^F)(\hat{\rho}_{2j}^{F^{**}} - \hat{\rho}_{BB_2}^F) \\ \frac{1}{N-1}\sum_{j=1}^{n}(\hat{\rho}_{1j}^{F^{**}} - \hat{\rho}_{BB_1}^F)(\hat{\rho}_{2j}^{F^{**}} - \hat{\rho}_{BB_2}^F) & \frac{1}{N-1}\sum_{j=1}^{n}(\hat{\rho}_{2j}^{F^{**}} - \hat{\rho}_{BB_2}^F)^2 \end{bmatrix}$$

Using the estimator of $A^F$ as $\tilde{A}^{F^{**}}$ we construct $100(1-\alpha)\%$ BB confidence region for traffic intensity vector $\underline{\rho}^F$ is as follows:

$$CR = \left\{ \underline{\rho}^F \mid n\left(\underline{\rho}^F - \underline{\rho}\right)\tilde{A}^{F^{**-1}}\left(\underline{\rho}^F - \underline{\rho}^F\right) \le \chi_{2,\alpha}^2 \right\} \tag{11}$$

### 4.4 Percentile Bootstrap Confidence Region

Consider $\underline{\rho}_1^{F^*} = \begin{pmatrix} \hat{\rho}_{11}^{F^*} \\ \hat{\rho}_{21}^{F^*} \end{pmatrix}, \underline{\rho}_2^{F^*} = \begin{pmatrix} \hat{\rho}_{12}^{F^*} \\ \hat{\rho}_{22}^{F^*} \end{pmatrix}, \cdots, \underline{\rho}_N^{F^*} = \begin{pmatrix} \hat{\rho}_{1N}^{F^*} \\ \hat{\rho}_{2N}^{F^*} \end{pmatrix}$ is the bootstrap distribution of $\underline{\rho}^F$. To arrange $\underline{\rho}_1^{F^*}, \underline{\rho}_2^{F^*}, \underline{\rho}_3^{F^*}, \cdots, \underline{\rho}_N^{F^*}$ we use Euclidean distance. Hence $\underline{\rho}_1^{F^*}(1)$, $\underline{\rho}_2^{F^*}(2)$, $\underline{\rho}_3^{F^*}(3), \cdots, \underline{\rho}_N^{F^*}(N)$ is the arrangement of $\underline{\rho}_1^{F^*}, \underline{\rho}_2^{F^*}, \underline{\rho}_3^{F^*}, \cdots, \underline{\rho}_N^{F^*}$. Then utilizing the $100(1-\alpha)^{th}$ percentage point of the bootstrap distribution, $100(1-\alpha)\%$ PB confidence region for $\underline{\rho}^F$ is given by

$$CR = \left\{ \underline{\rho}_j^F \mid d_j \le \left( \left[ N(1-\alpha) \right] \right) \right\} \tag{12}$$

where $[x]$ denotes the greatest integer less than or equal to $x$.

## 5 SIMULATION STUDY

A simulation study was performed to examine the performances of confidence regions constructed in equations (9) to (12) for traffic intensity vector $\underline{\rho}^F$. The performances of the confidence regions are assessed in terms of their coverage area percentage,average area and relative coverage area. Relative coverage area is defined as the ratio of coverage area percentage to average area of confidence region. We have simulated $M/E_4/1$ to $E_4/H_4^{Pe}/1$, $M/H_4^{Pe}/1$ to $H_4^{Pe}/E_4/1$, $E_4/H_4^{Pe}/1$ to $H_4^{Pe}/M/1$ and $E_4/H_4^{Po}/1$ to $H_4^{Po}/H_4^{Pe}/1$ queueing network models, where $M$: exponential distribution, $E_4$: 4-stage Erlang distribution, $H_4^{Pe}$: 4-stage hyperexponential distribution and $H_4^{Po}$: 4-stage hypoexponential distribution.

The level of $(\lambda, \mu_1, \mu_2, p_0, p_1)$ are set to $(0.2, 1, 15, 0.25, 0.75)$ so that $(\rho_1^F = 0.8, \rho_2^F = 0.2)$. For $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ random samples of arrival time and service time are drawn. Next $N = 1000$ bootstrap resamples each of size $n(= 5,10,20,30)$ are drawn from the original samples, as well as $N = 1000$ BB replications are simulated for the original samples. The above simulation process is replicated $N = 1000$ times. We compute coverage area percentage, average area and relative coverage area of confidence regions. Further calibration technique is used to improve the coverage percentage area of confidence regions obtained in (9) to (12). Simulation results are shown in Tables 1 to 4. According to equations (9) to (12) we obtain the CAN, the SB, the BB and the PB confidence regions for traffic intensity vector $\underline{\rho}^F$ with confidence level 90%. Confidence regions for $\underline{\rho}^F$ of $M/H_4^{Pe}/1$ to $H_4^{Pe}/E_4/1$ model for different values of $n = 5,10,20,30$ is shown in Figure-3 as an example.

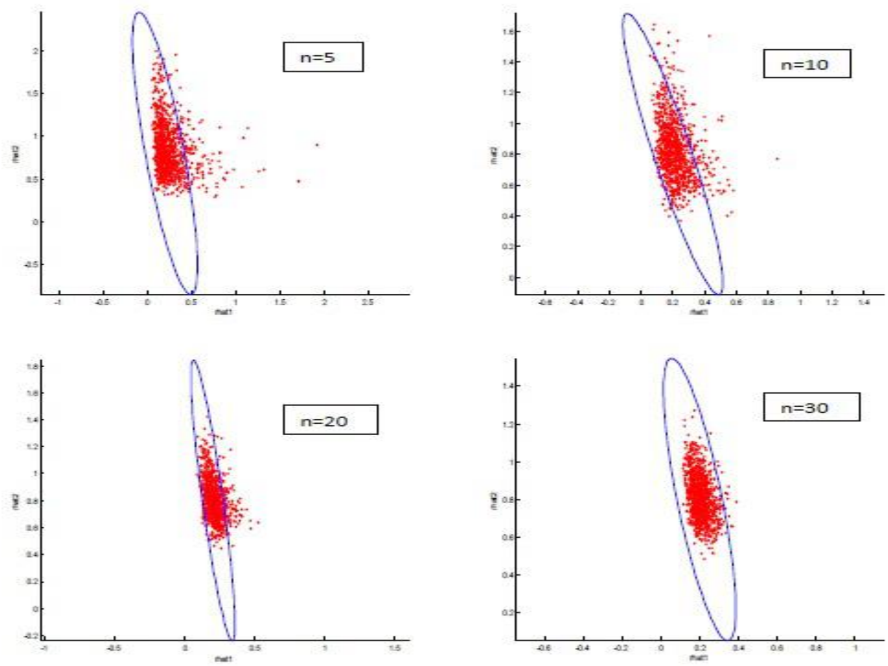


**Figure 3. Confidence Regions for $\underline{\rho}^F$ of $M/H_4^{Pe}/1$ to $H_4^{Pe}/E_4/1$.**

## 6 CONCLUSIONS

Different estimation approaches CAN, SB, BB and PB are applied to construct various confidence regions for $\underline{\rho}^F$. In Tables 1 to 4 we observed that average areas decrease but both coverage area percentages and relative coverage areas increase with n before as well as after calibration. Also we observed that coverage area percentages are approaches to 90 % when n increases to 30. Among all queueing network models the estimation approach **Bayesian Bootstrap** has the greatest relative coverage area before as well as after calibration. Table 5 shows that, due to calibration technique maximum increase in coverage area percentage of confidence regions is 12% for $E_4/H_4^{Po}/1$ to $H_4^{Po}/H_4^{Pe}/1$. After calibration relative coverage area is comparatively more than before calibration for all estimation approaches. These approaches are successfully and efficiently applied to practical queueing network models. Calibration technique is used to improve the coverage area percentages of confidence regions.

**Table 1. Simulation results for confidence regions of $M/E_4/1$ to $E_4/H_4^{Pe}/1$**

| Coverage Area Percentages for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Estimation Approaches | Before Calibration | | | | After Calibration | | | |
| | $n=5$ | $n=10$ | $n=20$ | $n=30$ | $n=5$ | $n=10$ | $n=20$ | $n=30$ |
| CAN | 0.772 | 0.845 | 0.865 | 0.887 | 0.807 | 0.866 | 0.895 | 0.888 |
| SB | 0.814 | 0.873 | 0.891 | 0.893 | 0.816 | 0.870 | 0.897 | 0.894 |
| BB | 0.735 | 0.837 | 0.863 | 0.879 | 0.789 | 0.858 | 0.893 | 0.885 |
| PB | 0.854 | 0.885 | 0.877 | 0.890 | 0.876 | 0.885 | 0.901 | 0.897 |
| Average Area for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
| CAN | 2.058 | 1.926 | 1.887 | 1.938 | 1.904 | 1.908 | 1.887 | 1.925 |
| SB | 0.897 | 0.241 | 0.103 | 0.068 | 0.780 | 0.238 | 0.103 | 0.068 |
| BB | 0.408 | 0.188 | 0.092 | 0.063 | 0.374 | 0.185 | 0.092 | 0.063 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| PB | 0.878 | 0.502 | 0.331 | 0.274 | 0.817 | 0.506 | 0.335 | 0.272 |
| Relative Coverage Area for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
| CAN | 0.375 | 0.439 | 0.458 | 0.458 | 0.424 | 0.454 | 0.474 | 0.461 |
| SB | 0.907 | 3.626 | 8.630 | 13.062 | 1.047 | 3.661 | 8.689 | 13.150 |
| BB | **1.802** | **4.447** | **9.397** | **13.962** | **2.112** | **4.648** | **9.715** | **14.135** |
| PB | 0.973 | 1.761 | 2.647 | 3.244 | 1.073 | 1.750 | 2.691 | 3.297 |

**Table 2. Simulation results for confidence regions of** $M/H_4^{Pe}/1$ **to** $H_4^{Pe}/E_4/1$

| Coverage Area Percentages for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Estimation Approaches | Before Calibration | | | | After Calibration | | | |
| | $n=5$ | $n=10$ | $n=20$ | $n=30$ | $n=5$ | $n=10$ | $n=20$ | $n=30$ |
| CAN | 0.782 | 0.818 | 0.864 | 0.869 | 0.867 | 0.886 | 0.876 | 0.909 |
| SB | 0.835 | 0.860 | 0.887 | 0.879 | 0.884 | 0.894 | 0.883 | 0.914 |
| BB | 0.753 | 0.813 | 0.859 | 0.858 | 0.855 | 0.883 | 0.879 | 0.909 |
| PB | 0.857 | 0.863 | 0.889 | 0.886 | 0.905 | 0.907 | 0.893 | 0.910 |
| Average Area for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
| CAN | 1.949 | 1.896 | 1.887 | 1.900 | 2.023 | 1.890 | 1.896 | 1.900 |
| SB | 1.648 | 0.245 | 0.105 | 0.068 | 1.008 | 0.239 | 0.106 | 0.068 |
| BB | 0.399 | 0.187 | 0.092 | 0.062 | 0.417 | 0.186 | 0.093 | 0.062 |
| PB | 1.016 | 0.518 | 0.344 | 0.281 | 0.898 | 0.511 | 0.349 | 0.280 |
| Relative Coverage Area for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
| CAN | 0.401 | 0.431 | 0.458 | 0.457 | 0.429 | 0.469 | 0.462 | 0.478 |
| SB | 0.507 | 3.517 | 8.463 | 12.995 | 0.877 | 3.734 | 8.373 | 13.521 |
| BB | **1.886** | **4.340** | **9.307** | **13.862** | **2.049** | **4.737** | **9.484** | **14.678** |
| PB | 0.844 | 1.666 | 2.583 | 3.155 | 1.008 | 1.774 | 2.557 | 3.249 |

**Table 3. Simulation results for confidence regions of** $E_4/H_4^{Pe}/1$ **to** $H_4^{Pe}/M/1$

| Coverage Area Percentages for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Estimation Approaches | Before Calibration | | | | After Calibration | | | |
| | $n=5$ | $n=10$ | $n=20$ | $n=30$ | $n=5$ | $n=10$ | $n=20$ | $n=30$ |
| CAN | 0.699 | 0.792 | 0.825 | 0.852 | 0.810 | 0.870 | 0.889 | 0.894 |
| SB | 0.695 | 0.799 | 0.823 | 0.853 | 0.812 | 0.873 | 0.888 | 0.893 |
| BB | 0.647 | 0.765 | 0.808 | 0.846 | 0.763 | 0.858 | 0.885 | 0.888 |
| PB | 0.822 | 0.860 | 0.879 | 0.880 | 0.889 | 0.897 | 0.899 | 0.912 |
| Average Area for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
| CAN | 1.587 | 1.768 | 1.837 | 1.843 | 1.652 | 1.733 | 1.817 | 1.858 |
| SB | 0.325 | 0.179 | 0.092 | 0.062 | 0.342 | 0.175 | 0.091 | 0.062 |
| BB | 0.250 | 0.157 | 0.087 | 0.059 | 0.261 | 0.154 | 0.085 | 0.059 |
| PB | 0.432 | 0.323 | 0.228 | 0.187 | 0.443 | 0.312 | 0.228 | 0.188 |
| Relative Coverage Area for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
| CAN | 0.441 | 0.448 | 0.449 | 0.462 | 0.490 | 0.502 | 0.489 | 0.481 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| SB | 2.141 | 4.462 | 8.931 | 13.869 | 2.375 | 4.989 | 9.721 | 14.398 |
| BB | **2.589** | **4.859** | **9.328** | **14.382** | **2.919** | **5.567** | **10.376** | **14.942** |
| PB | 1.904 | 2.666 | 3.848 | 4.718 | 2.009 | 2.873 | 3.945 | 4.858 |

**Table 4. Simulation results for confidence regions of** $E_4/H_4^{Po}/1$ **to** $H_4^{Po}/H_4^{Pe}/1$

| Coverage Area Percentages for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Estimation Approaches | Before Calibration | | | | After Calibration | | | |
| | $n=5$ | $n=10$ | $n=20$ | $n=30$ | $n=5$ | $n=10$ | $n=20$ | $n=30$ |
| CAN | 0.752 | 0.838 | 0.866 | 0.863 | 0.842 | 0.88 | 0.901 | 0.895 |
| SB | 0.751 | 0.842 | 0.866 | 0.861 | 0.848 | 0.885 | 0.907 | 0.898 |
| BB | 0.685 | 0.811 | 0.855 | 0.855 | 0.805 | 0.875 | 0.896 | 0.892 |
| PB | 0.821 | 0.869 | 0.886 | 0.872 | 0.899 | 0.898 | 0.904 | 0.914 |
| Average Area for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
| CAN | 1.151 | 1.188 | 1.2249 | 1.2446 | 1.1431 | 1.2046 | 1.2189 | 1.2296 |
| SB | 0.2505 | 0.1227 | 0.0622 | 0.0418 | 0.2489 | 0.1244 | 0.0618 | 0.0414 |
| BB | 0.1872 | 0.107 | 0.0579 | 0.0399 | 0.1861 | 0.1084 | 0.0576 | 0.0394 |
| PB | 0.4312 | 0.3164 | 0.2256 | 0.1852 | 0.4356 | 0.3146 | 0.2236 | 0.1836 |
| Relative Coverage Area for $(\rho_1^F = 0.8, \rho_2^F = 0.2)$ | | | | | | | | |
| CAN | 0.653 | 0.705 | 0.707 | 0.693 | 0.737 | 0.731 | 0.739 | 0.728 |
| SB | 2.998 | 6.861 | 13.932 | 20.596 | 3.408 | 7.113 | 14.687 | 21.707 |
| BB | **3.659** | **7.577** | **14.776** | **21.45** | **4.325** | **8.072** | **15.558** | **22.632** |
| PB | 1.904 | 2.746 | 3.928 | 4.708 | 2.064 | 2.855 | 4.042 | 4.977 |

**Table 5. Maximum percentage increase in coverage percentage area due to calibration technique for**
$(\rho_1^F = 0.8, \rho_2^F = 0.2)$

| Queueing Network Model | $n=5$ | $n=10$ | $n=20$ | $n=30$ |
|---|---|---|---|---|
| $M/E_4/1$ to $E_4/H_4^{Pe}/1$ | 5.4 | 2.1 | 3.0 | 0.7 |
| $M/H_4^{Pe}/1$ to $H_4^{Pe}/E_4/1$ | 10.2 | 7.0 | 2.0 | 5.1 |
| $E_4/H_4^{Pe}/1$ to $H_4^{Pe}/M/1$ | 11.7 | 9.3 | 7.7 | 4.2 |
| $E_4/H_4^{Po}/1$ to $H_4^{Po}/H_4^{Pe}/1$ | 12.0 | 6.4 | 4.1 | 4.2 |

## REFERENCES


[1] Trivedi K.S. (2000), Probability and Statistics with Reliability, Queueing and Computer Science Applications, Prentice Hall, India, ISBN 81-203-0508-6.

[2] Jackson J.R. (1957), Networks of Waiting Lines, Operations Research 5, 518- 521.

[3] Jackson J.R.(1963), Jobshop-Like Queueing Systems Management Science, Vol. 10, pp. 131-142.

[4] Disney R. L. (1975) , Random flow in queueing networks: a review and a critique. Trans. A.I.E.E., 7, pp. 268-288.

[5] Kleinrock L.(1976), Queueing Systems, Vol. II (1976), Computer Applications, John Wiley & Sons, New York.

[6] Thiruvaiyaru D., Basawa I.V. and Bhat U.N. (1991), Estimation for a class of simple queueing network. Queueing Systems 9, 301-312.

[7] Thiruvaiyaru D, and Basava I.V.(1996) , Maximum likelihood estimation for queueing networks, In Stochastic Processes and Statistical Inference (Eds., B.L.S. Prakasa Rao and B. R. Bhat), pp.132-149, New Age International Publications, New Delhi.

[8] Gross D. and Harris C.M.(2002),Fundamentals of Queueing Theory, Third Edition,John Wiley and Sons,INC.

[9] Efron B. (1979), Bootstrap methods: another look at the jackknife. Annals of Statistics 7, 1-26.

[10] Efron B. (1982), The jackknife, the bootstrap, and other resampling plans. CBMS-NSF regional conference series in applied mathematics monograph 38. Philadelphia, PA: SIAM.

[11] Efron B. (1987) Better bootstrap confidence intervals. J Am Stat Assoc; 82: 171-200.

[12] Efron B. and Gong G.(1983), A leisurely look at the boot strap, the jackknife and cross validation, The American Statistician, Vol.37, pp.36-48.

[13] Efron B. and Tibshirani R.J. (1986), Bootstrap Method for standard errors, confidence intervals and other measures of statistical accuracy, Statistical Science, Vol. 1, pp. 54-77.

[14] Rubin D.B. (1981), The Bayesian bootstrap, The Annals of Statistics, Vol.9, pp. 130-134.

[15] Ke J. C. and Chu Y. K. (2009), Comparison on five estimation approaches of intensity for a queueing system with short run, Computational Statistics, Vol. 24 No.- 4,pp. 567-582.

[16] Gedam V.K. and Pathare S.B. (2013), Comparison of different confidence intervals of intensities for an open queueing network with feedback. Am J Oper Res, Vol.3 No.2, 307–327.

[17] Gedam V.K. and Pathare S.B.(2013), Calibrated confidence intervals for intensities of a two stage open queueing network with feedback. J Stat Math, Vol. 4, No.1 , 151–161.

[18] Gedam V.K. and Pathare S.B. (2014),Calibrated confidence intervals for intensities of a two stage open queueing network. J Stat. Appl. Pro., Vol. 3, No.1, 33–44.

[19] Pathare S.B. and Gedam V.K.(2014), Some estimation approaches of intensities for a two stage open queueing network. Stat Optim Inform Comput, Vol. 2, No.1, 33–46.

[20] Gedam V.K. and Pathare S.B. (2015), Estimation approaches of mean response time for a two stage open queueing network model. Stat Optim Inform Comput, Vol.3, No.3, 249-258.

[21] Gedam V.K. and Pathare S.B. (2015), Use of the calibration approach in confidence intervals for mean response times of an open queueing network with feedback. SIMULATION: Transactions of The Society for Modeling and Simulation International , Vol. 91, No 6 , pp. 553-565.

[22] Efron B. and Tibshirani R.J. (1993), An Introduction to the bootstrap, Chapman and Hall, New York.

[23] Loh W.Y. (1987) Calibrating confidence coefficient. J Am Stat Assoc; 82: 155-162.

[24] Beran R.(1987) Prepivoting to reduce level error of confidence sets. Biometrika; 74: 457-468.

[25] Hall P. (1986), On the bootstrap and confidence intervals. Ann Stat. 14: 1431-1452.

[26] Hall P. and Martin M.A.(1988), On bootstrap resampling and iteration. Biometrika; 75: 661-671.

[27] Loh W.Y. (1991), Bootstrap calibration for confidence interval construction and selection. Stat sinica, 1: 477-491.

[28] Kale B.K.(1999), A First Course on Parametric Inference, Narosa Publishing House, London.

[29] Anderson T.W.(1984), An Introduction to Multivariate Statistical Analysis, Second edition, Wiley series in probability and mathematical statistics, New York.