Shrinkage and Bayesian Shrinkage Estimation of the Expected Length of a $M/M/1$ Queue System

A. Kiapour*† and M. Naghizadeh Qomi‡

† Babol branch, Islamic Azad University
‡ University of Mazandaran

Received: 2/2/2019   Approved: 11/19/2019

Abstract. In this paper, shrinkage and Bayesian shrinkage estimation of the expected length ($l$) in a $M/M/1$ queue system is considered. A shrinkage estimator of $l$ is considered when a priori about $l$ as $l_0$ is available. The bias and the risk of shrinkage estimators are derived under a scale-invariant squared error loss (SISEL) function. A class of Bayes shrinkage estimators for $l$ is proposed which is a generalization of Bayes shrinkage estimator and a relative performance of proposed estimators and the maximum likelihood estimator (MLE) is performed. A simulated data is given to illustrate the proposed results. Finally, we conclude with a summary of our contributions.

Keywords. Bayes shrinkage estimator, expected length, $M/M/1$ queue, scale-invariant squared error loss function.

MSC 2010: 62F10; 62F15; 60K25.

1 Introduction

Queueing systems are studied in many interesting probabilistic, operational, and statistical problems. For designing of queues and comparison of the
performance of these systems, it is necessary to estimate the parameters of the queuing system such as the arrival rate, service time and traffic intensity. Much of estimation procedures in the literature has focused on their development for classical and Bayesian approaches. However, no work has focused strictly on the development of semiclassical estimation based on a point guess of the parameters of queue.

Clarke (1957) derived the MLE of the arrival rate and service time in a $M/M/1$ queue system. The Bayes estimator of the traffic intensity in this type of systems is obtained by Armero and Bayarri (1994). Sharma and Kumar (1999) and Mukherjee and Chowdhury (2005) obtained The Bayes estimator of the traffic intensity under squared error loss (SEL) and linear-exponential (LINEX) loss functions, respectively. Dey (2008) obtained Bayes estimators of traffic intensity and various queue characteristics under the assumptions of different priors under the SEL function. Srinivas et al. (2011) considered Maximum likelihood and uniform minimum variance unbiased estimators for some measures of $M/M/1$ queue. Ren and Wang (2012) obtained the Bayes estimator of the traffic intensity under a precautionary loss function.

Consider a $M/M/1$ queuing system with the mean arrival rate $\lambda$ and mean service time $\frac{1}{\mu}$. Let the random variable $X$ representing the number of customers in the system under steady state has the distribution specified by the Geometric probability mass function (p.m.f)

$$P(x|\rho) = (1 - \rho)\rho^x, \quad x = 0, 1, 2, ...$$ (1)

where $\rho = \frac{\lambda}{\mu}$, $0 < \rho < 1$, represent the traffic intensity for the given queuing system $M/M/1$. One of important characteristic of a queue system is expected length, $E(X) = \frac{\rho}{1 - \rho} = l$. Let $X^n = (X_1, ..., X_n)$ be a sample of random variables has geometric distribution (1). Then the likelihood function corresponding to the p.m.f (1) on the basis of currently observed data $x^n = (x_1, ..., x_n)$ from $X^n$ is given by

$$L(\rho) = (1 - \rho)^n \rho^t, \quad x = 0, 1, 2, ...$$ (2)

where $t$ is an observation of $T = \sum_{i=1}^{n} X_i$. Then, the MLE of $\rho$ is $\hat{\rho} = \frac{T}{n+t}$. Using the invariant property, the MLE of $l$ is $\hat{l}_{ML} = \frac{T}{n} = \bar{X}$. This estimator is called as sample information.

In some practical situations, experimenter has nonsample information about unknown parameter $\rho$ (or equivalently $l$) as a prior point guess value.
A. Kiapour and M. Naghizadeh Qomi

Therefore, the researcher can improve the MLE or any other natural estimator by shrinking it towards the guess value and construct a linear shrinkage estimator in the hope that it will perform better than the natural estimator. This type of estimation is in the area of semiclassic estimation. In Bayesian perspective, we can combine experimental data with prior subjective knowledge on parameters to obtain an improved estimation.

Our aim of this paper is shrinkage and Bayesian shrinkage estimation of \( l \) in the presence of nonsample information under a SISEL function of the form

\[
L(l, \hat{l}) = \left( \frac{\hat{l}}{l} - 1 \right)^2, \tag{3}
\]

where \( \hat{l} \) is an estimator of \( l \). It is mentioned here that one can consider the pretest estimators which is studied by several authors, see Arabi Belaghi et al. (2014, 2015a,b), Naghizadeh Qomi and Barmoodeh (2015) and Kiapour and Naghizadeh Qomi (2016). But, we focus on Bayesian shrinkage estimation by constructing a class of Bayesian shrinkage estimators and compare its performance with the MLE.

To do this, we first introduce a linear shrinkage estimator of \( l \) in Section 2. The bias and risk of MLE and the shrinkage estimator are derived and performance of the shrinkage estimator with respect to the MLE is discussed analytically and numerically. A Bayes shrinkage estimator for expected queue length under the SISEL function is introduced in Section 3. Moreover, we construct a class of Bayes shrinkage estimators of \( l \) under the SISEL function and study its performance with respect to the MLE. A simulated data is used for illustrating the results in Section 4. We end the paper with some concluding remarks in Section 5.

2 A Linear Shrinkage Estimator of \( l \)

Following Thompson (1968) consider a linear shrinkage estimator of \( l \) of the form

\[
\hat{l}_s = k\hat{l}_{ML} + (1 - k)l_0, \quad 0 \leq k \leq 1 \tag{4}
\]

where \( \hat{l}_{ML} \) is the MLE of \( l \) and \( k \) is the shrinkage factor specified by the experimenter according to his belief in the guess value \( l_0 \). The estimator (4) is a combination of sample information (\( \hat{l}_{ML} \)) and nonsample information...
(l_0), see Saleh (2006) as a comprehensive reference in this area.

Selection a value for the shrinkage parameter \( l \) such that it may be regarded as an optimal estimator is a key question. A choice of shrinkage factor \( k \) is it to chose the parameter \( k \) in a data-driven fashion by explicitly minimizing the risk of shrinkage estimator \( \hat{l}_s \), i.e.,

\[
k_{\text{opt}} = \arg \min_{k \in [0,1]} R(l, \hat{l}_s).
\] (5)

### 2.1 Bias of Shrinkage Estimator

In this subsection, we investigate the bias of shrinkage estimator (4). we have the following definition for risk-unbiasedness due to Lehmann (1951).

**Definition 1.** An estimator \( \hat{l} \) of \( l \) is said to be risk-unbiased under the loss function \( L(l, \hat{l}) \) if it satisfies

\[
E[L(l, \hat{l})] \leq E[L(l', \hat{l})], \quad \forall l' \neq l.
\] (6)

Under SISEL function, the above definition reduces to the following lemma.

**Lemma 1.** Under SISEL function, an estimator \( \hat{l} \) of \( l \) is risk-unbiased if \( E(\hat{l}^2) = lE(\hat{l}) \).

**Proof.** Using the Definition, we have

\[
E[L(l, \hat{l})] - E[L(l', \hat{l})] = E[\hat{l}^2] - 1^2 - E[\hat{l}'^2] - 1^2 = (\frac{1}{l^2} - \frac{1}{l'^2})E(\hat{l}^2) + 2(\frac{1}{l} - \frac{1}{l'})E(\hat{l}).
\]

If we set \( E(\hat{l}^2) = lE(\hat{l}) \), we get

\[
E[L(l, \hat{l})] - E[L(l', \hat{l})] = -(\frac{l - l'}{ll'})^2 E(\hat{l}^2) < 0.
\]

Therefore, an estimator \( \hat{l} \) of \( l \) is risk-unbiased under the SISEL function, if it satisfies in condition \( E(\hat{l}^2) = lE(\hat{l}) \).
Then the bias function is defined as

\[ \text{Bias}_l(\hat{l}) = E(\hat{l}^2) - lE(\hat{l}). \]

For computing the bias of shrinkage estimator (4), we get

\[ E(\hat{l}_s) = kl + (1 - k)l_0, \]

and

\[ E(\hat{l}_s^2) = E[k\bar{X} + (1 - k)l_0]^2 = \frac{k^2l(1 + l)}{n} + (kl + (1 - k)l_0)^2. \]

Therefore, the bias of \( \hat{l}_s \) is given by

\[ \text{Bias}_l(\hat{l}_s) = \frac{k^2l(1 + l)}{n} - (kl + (1 - k)l_0)(1 - k)(l - l_0). \]

Figure 1 shows the bias of shrinkage estimator \( \hat{l}_s \) for selected sample sizes \( n = 5, 15, 30, 50 \) and shrinkage factors \( k = 0.1, 0.3, 0.5, 0.8 \) with respect to the guess value \( l_0 \) when \( \rho = 0.75 \) (or equivalently \( l = \frac{\rho}{1 - \rho} = 3 \)). From Figure 1, we observe that the bias may be negative, zero or positive, then we can state that the shrinkage estimator may be negatively risk-biased, risk-unbiased or positively risk-biased.

### 2.2 Comparison between Shrinkage Estimator and the MLE

In a loss function or decision theoretic analysis, the quality of an estimator is qualified in its risk function; that is, for an estimator \( \delta(x) \) of \( l \), the risk function is \( R(l, \delta) = E(L(l, \delta)) \). The risk function of \( \hat{l}_{ML} \) under the loss function (3) is given by

\[ R(l, \bar{X}) = E[\frac{\bar{X}}{l} - 1]^2 = \frac{l + 1}{nl}. \]
Also, the risk of $\hat{l}_s$ under the loss function (3) is given by

$$R(l, \hat{l}_s) = E[\hat{l}_s - 1]^2 = E[k\hat{l}_{ML} + (1 - k)l_0 - 1]^2$$

$$= \frac{k^2(1 + l)}{n l} + [(1 - k)(1 - \Delta)]^2,$$

where $\Delta = l_0/l$. Note that an SISEL-optimal shrinkage factor $k$ given in (5) is obtained as

$$k_{opt} = \frac{nl(1 - \Delta)^2}{1 + l + nl(1 - \Delta)^2}$$

which is a function of $\Delta$ and is unknown, but it can be estimated using the sample data.

The relative efficiency of shrinkage estimator with respect to the MLE is computed as
Figure 2. Plots for RE between shrinkage estimators and the MLE for selected values of $n = 5, 15, 30, 50$ and $k = 0.2, 0.4, 0.6, 0.8$ with respect to $\Delta$ when $l = 3$.

From (7), if we set $\Delta = 1$, i.e., $l = l_0$, then we have $RE = 1/k^2$, which indicates that $\hat{l}_s$ performs better than $\hat{l}_{ML}$. Figure 2 shows the RE between shrinkage estimator $\hat{l}_s$ and the MLE for selected sample sizes $n = 5, 15, 30, 50$ and shrinkage factors $k = 0.2, 0.4, 0.6, 0.8$ with respect to the guess value $l_0$ when $\rho = 0.75$ (or equivalently $l = \frac{\rho}{1-\rho} = 3$). From Figure 2, it is observed...
Table 1. The range of $\Delta$ that shrinkage estimator dominates the MLE.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$k$</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>0.2</td>
<td>[0.6348,1.3651]</td>
<td>[0.5445,1.4554]</td>
<td>[0.4037,1.5963]</td>
<td>[0.1055,1.8943]</td>
</tr>
<tr>
<td>30</td>
<td>0.2</td>
<td>[0.7417,1.2581]</td>
<td>[0.6779,1.3220]</td>
<td>[0.5783,1.4216]</td>
<td>[0.3675,1.6324]</td>
</tr>
<tr>
<td>50</td>
<td>0.2</td>
<td>[0.7999,1.9999]</td>
<td>[0.7505,1.2494]</td>
<td>[0.6734,1.3265]</td>
<td>[0.5101,1.4899]</td>
</tr>
<tr>
<td>70</td>
<td>0.2</td>
<td>[0.8309,1.1690]</td>
<td>[0.7891,1.2108]</td>
<td>[0.7239,1.2760]</td>
<td>[0.5859,1.4140]</td>
</tr>
</tbody>
</table>

that shrinkage estimator performs well with respect to the MLE when $\Delta$ is in the vicinity of 1. Moreover, when the guess value is near to the true value of $l$ and sample size $n$ is fixed, shrinkage estimators corresponding with small shrinkage factors have good performance, than other shrinkage estimators.

Looking at shrinkage estimators in Figure 2 shows that the proposed shrinkage estimates have RE close to zero in the broad range of parameter space. For this reason, we should specify some rang that can help the reader to choose that interval when wants to apply estimates. We can find some parameters for the shrinkage estimates to be uniformly better than the MLE. The range of $\Delta$ that shrinkage estimator dominates the MLE (the so-called effective interval) for selected values of $n = 15, 30, 50, 70$ and $k = 0.2, 0.4, 0.6, 0.8$ and $l = 3$ are summarized in Table 1.

3 A Class of Bayes Shrinkage Estimators

In this section, we first derive the Bayes estimator of $l$ under the loss function (3). Then, we compute a Bayes shrinkage estimator using prior information and a point guess value $l_0$. This reach us to construct a class of Bayes shrinkage estimators. We perform a relative performance for illustrating the proposed class of Bayes shrinkage estimators.

Let $X^n = (X_1, ..., X_n)$ be a sample of random variables has geometric distribution (1). Consider a beta prior for $\rho$, $Beta(\alpha, \beta)$, with p.d.f.

$$
\pi_{\alpha,\beta}(\theta) = \frac{1}{B(\alpha, \beta)} \rho^{\alpha-1}(1 - \rho)^{\beta-1}, \quad \alpha > 0, \beta > 0, \theta > 0.
$$

where $B(.,.)$ is the Beta function, i.e., $Beta(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1 - x)^{\beta-1}dx$. Combining the prior information (8) with the likelihood function (2), the
posterior p.d.f. of $\rho$ given $X^n$ is

$$
\pi_{\alpha, \beta}(\theta|X^n) = \frac{1}{Be(T + \alpha, n + \beta)} \rho^{T + \alpha - 1}(1 - \rho)^{n + \beta - 1}, \quad \alpha > 0, \beta > 0, \theta > 0.
$$

(9)

which is $Beta(T + \alpha, n + \beta)$ where $T = \sum_{i=1}^{n} X_i$. Therefore, the Bayes estimator of $l$ under the loss function (3) is given by

$$
\hat{l}_B = \hat{l}_B(X^n) = \frac{E[\frac{1}{\hat{T}}|X^n]}{E[\frac{1}{\hat{T}}|X^n]} = \frac{Be(\alpha - 1, \beta + 1)}{Be(\alpha - 2, \beta + 2)} = \frac{T + \alpha - 2}{n + \beta + 1}.
$$

(10)

Following Naghizadeh Qomi (2017), if we set $E[\hat{l}_B(X^n)] = l_0$, then we have

$$
\frac{E[T] + \alpha - 2}{n + \beta + 1} = l_0.
$$

(11)

Using the fact $E[T] = nl_0$ by a little algebra we reach to

$$
\alpha = l_0(\beta + 1) + 2.
$$

(12)

Substituting the value of $\alpha$ from (12) in (10), the Bayes estimator is given by

$$
\hat{l}_{BS} = \frac{T + l_0(\beta + 1)}{n + \beta + 1} = \frac{n}{n + \beta + 1} \bar{X} + \frac{\beta + 1}{n + \beta + 1} l_0 = \lambda \hat{l}_{ML} + (1 - \lambda) l_0,
$$

(13)

where $\lambda = \frac{n}{n + \beta + 1}$. The Bayes estimator (13) is expressed in the form of a weighted sum of the sample data and the prior information, and has a form of shrinkage estimator given in (4), which is called as a Bayes shrinkage estimator.

We can construct a class of Bayes shrinkage estimator of the form

$$
\hat{l}_{BS}^\lambda = \frac{n + (\beta + 1) \lambda^*}{n + \beta + 1} \hat{l}_{ML} + \frac{(\beta + 1)(1 - \lambda^*)}{n + \beta + 1} l_0,
$$

$$
= A \hat{l}_{ML} + (1 - A) \lambda_0,
$$

(14)

where $A = (n + (\beta + 1) \lambda^*)/(n + \beta + 1)$ with $\lambda^* \in [0, 1]$. The estimator (14) is
a generalization of Bayes shrinkage estimator proposed in (13). If we apply (14) with $\lambda^* = 0$, we get Bayes shrinkage estimator (13).

### 3.1 Consideration of Bias

For computing the bias of $\hat{l}_{BS}^{\lambda^*}$ under the SISEL function, we have

$$E(\hat{l}_{BS}^{\lambda^*}) = E[A\bar{X} + (1 - A)l_0] = Al + (1 - A)l_0$$

and

$$E(\hat{l}_{BS}^{\lambda^*})^2 = E[A\bar{X} + (1 - A)l_0]^2 = \frac{A^2l(1 + l)}{n} + (Al + (1 - A)l_0)^2$$

Therefore, the bias of $\hat{l}_{BS}^{\lambda^*}$ under the SISEL function is given by

$$\text{Bias}(\hat{l}_{BS}^{\lambda^*}) = E(\hat{l}_{BS}^{\lambda^*})^2 - lE(\hat{l}_{BS}^{\lambda^*}) = \frac{A^2l(1 + l)}{n} - (Al + (1 - A)l_0)(1 - A)(l - l_0).$$

The bias of Bayes shrinkage estimator $\hat{l}_{BS}^{\lambda^*}$ is plotted in Figure 3 for selected sample sizes $n = 5, 20, 40, 80$, prior parameters $\beta = 1, 3, 5, 7$ and $\lambda^* = 0$ with respect to $l_0$ when $\rho = 0.75$ (or equivalently $l = \frac{\rho}{1-\rho} = 3$). From Figure 3, we observe that the bias may be negative, zero or positive, then we can state that the shrinkage estimator may be negatively risk biased, risk unbiased or positively risk biased.

### 3.2 Relative Performance

The risk of generalized Bayes shrinkage estimator $\hat{l}_{BS}^{\lambda^*}$ under the loss function (3) is given by

$$R(l, \hat{l}_{BS}^{\lambda^*}) = E[\frac{\hat{l}_{BS}^{\lambda^*}}{l} - 1]^2$$

$$= E[\frac{Al_{ML} + (1 - A)l_0}{l} - 1]^2$$

$$= \frac{A^2(l + 1)}{nl} + (1 - A)^2(\Delta - 1)^2.$$
The estimated value of $\hat{\lambda}^*$ has been obtained by substituting the estimator of $l$ as $\hat{l}_{ML} = \bar{X}$.

The relative efficiency of Bayes shrinkage estimator with respect to the MLE is computed as

$$RE(\hat{l}_{BS}, \hat{l}_{ML}) = \frac{R(l, \bar{X})}{R(l, \hat{l}_{BS}^{*})} = \frac{l + 1}{A^2(l + 1) + nl(1 - A)^2(\Delta - 1)^2}$$

Figure 4 shows the RE between shrinkage estimator $\hat{l}_{BS}^{*}$ and the MLE for
selected sample sizes $n = 5, 20, 40, 80$, prior parameters $\beta = 1, 3, 5, 7$ and $\lambda^* = 0$ with respect to $\Delta$ when $l = 3$. It is observed from Figure 4 that Bayes shrinkage estimators performs well with respect to the MLE when $\Delta$ is in the vicinity of 1. Moreover, for fixed sample size $n$, Bayes shrinkage estimators corresponding with larger values of $\beta$ have good performance than other Bayes shrinkage estimators when the guess value is near to the true value of $l$. For investigating values of $\lambda^*$ on shrinkage estimators, we plotted the RE in Figure 5 for selected sample sizes $n = 5, 20, 40, 80$, prior parameters $\beta = 1, 3, 5, 7$ and $\Delta = 1$ with respect to $\lambda^*$ when $l = 3$. It is clear from Figure 5 that the RE is a decreasing function of $\lambda^*$. Moreover, as mentioned, Bayes shrinkage estimators with larger values of
Figure 5. Plots of RE between Bayes shrinkage estimator and the MLE for selected values of $n = 5, 20, 40, 80$ and $\beta = 1, 3, 5, 7$ with respect to $\lambda^*$ when $l = 3$ and $\Delta = 1$.

$\beta$ have higher efficiency than other Bayes shrinkage estimators when $\Delta = 1$.

4 A Simulated Example

A sample of size $n = 20$ is simulated from a geometric distribution with $\rho = 0.4$. The data are as follows:

$1, 0, 1, 0, 1, 1, 1, 2, 2, 0, 0, 1, 2, 0, 2, 2, 1, 1, 1, 0$.

The MLE of $l$ based on data is $\hat{l}_{ML} = \bar{x} = 0.95$. We consider selected underestimated and overestimated values of $l_0 = MLE$ as $l_0 = 0.2, 0.8, 1.5, 4, 8$.

\[ J. \text{Statist. Res. Iran} \ 15 \ (2018): \ 301-316 \]
and summarize the Bayesian shrinkage estimates of $l$ with $\lambda^* = 0$ and $\beta = 0, 1, 5, 10, 15, 20$ in Table 2. We observe from Table 2 that the Bayesian shrinkage estimates increases when the value of $\beta$ increases for $l_0 > MLE = 0.95$.

## 5 Concluding Remarks

Shrinkage and Bayesian shrinkage estimators for the expected length of a $M/M/1$ queue system are obtained under SISEL function when the queuing analyst has a priori information. A class of Bayes shrinkage estimators is proposed and its Bias performance and relative performance with respect to the MLE is studied numerically and graphically. Our findings show that shrinkage estimators have higher efficiency than the MLE when the guess of expected length is close to the true value, that is $\Delta = 1$. Moreover, shrinkage estimators with small value of shrinkage factors perform better than other shrinkage estimators for fixed sample size and $\Delta = 1$. Employing a prior distribution for traffic intensity and combining it with a point guess, investigation of a class of Bayes shrinkage estimators show that Bayes shrinkage estimators with larger values of hyperparameters $\alpha$ and $\beta$ perform better than other estimators when $\Delta = 1$.

### Acknowledgment

The authors are grateful to the Editor and reviewers for making helpful comments and suggestions on an earlier version of this paper.
References


Azadeh Kiapour
Department of Statistics,
Babol Branch, Islamic Azad University,
Babol, Iran.
email: Kiapour@baboliau.ac.ir

Mehran Naghizadeh Qomi
Department of Statistics,
University of Mazandaran,
Babolsar, Iran.
email: m.naghizadeh@umz.ac.ir