J. Operations Research Soc. of Japan
Vol. 13, No. 2, October 1970

INEQUALITIES FOR MANY-SERVER QUEUE AND OTHER QUEUES

TAKEJI SUZUKI and YOSHIYUKI YOSHIDA

The Defense Academy

(Received August 7, 1970)

Abstract

An inequality for the mean of the equilibrium waiting time distribution in the queueing system with many servers is presented. The case with two servers, specially, is discussed in detail. Further, the two bulk queueing systems are considered in which the inequalities for the mean and variance of the equilibrium queue-length distribution are found.

1. Introduction

Several methods of getting upper and lower bounds for the moments of the equilibrium waiting time or queue-length distribution have been developed only for the single-server queueing system with first-come-first-served discipline by Kingman [5], Marshall [7] and the others.

In the many-server queueing-system $GI/G/s$, any method of finding upper and lower bounds for the mean $E(W)$ of the equilibrium waiting time distribution, supposedly, has not been presented before.*) Kingman

---

*) After the paper was recieved by the society, authors had a chance of reading Kingman's paper "Inequalities in the Theory of Queues, *J.R. Statist. Soc. B*, 32 (1970) 102–110" in which he dealt with inequalities of many-server queueing-system. In order to get a lower bound for $E(W)$ in the system $GI/G/s$ he considered a single-server queueing-system. This point of view is essentially same as stated in Kiefer and Wolfowitz [4]. This idea has been used in this paper too. Hewever the methods employed in this paper for getting upper bounds are different from that of Kingman and all results under the conditions imposed on theorems are better than his result.

[6], however, made the following conjecture:

> When the queueing system $GI/G/s$ is in a situation of heavy traffic, the equilibrium waiting time distribution is approximately negative exponential, with mean
>
> (1.1) $$\frac{V(T)+V(S/s)}{2\{E(T)-E(S/s)\}},$$
>
> where E and V denote the expectation and variance of a random variable respectively,

and these notations will be used throughout the paper.

In this paper we deal with the many-server queueing-system in the steady state and find some results for upper and lower bounds of $E(W)$. In the next section it is shown that (1.1) is generally true for $E(S)\leqq E(T)$ when $s\geqq 2$. In particular an upper bound for $E(W)$ in the queueing system $GI/M/s$ is presented for $E(S)<sE(T)$. Further, the queueing system $GI/G/2$ is discussed in detail and we obtain upper and lower bounds for $E(W)$. The method used for the case $GI/G/2$ is not seemed to be applicable for the general case $GI/G/s$.

Finally, we deal with two bulk queueing systems, M(*batch-arrival*)/G(*batch-service*)/1 and GI(*batch-arrival*)/M(*batch-service*)/1. For these queueing systems upper and lower bounds for the mean $E(Q)$ or variance of the equilibrium queue-length distribution are presented and it should be note that these upper bounds for $E(Q)$ are of similar form as the expression (1.1).

Now we shall consider the queueing systems in the equilibrium and assume the necessary moments exist. And the subscript n of a notation refers to the n-th customer. When it is not required to note the order of the customer the subscript will be dropped throughout the paper.

2. Many-server queue

The following notations are used throughout the section.

T_n=time between the n-th and the $(n+1)$-th arrival.

S_n=service time of the n-th customer.

s=number of servers.

$U_n=(S_n/s)-T_n$.

W_n=waiting time in the queue of the n-th customer.

$\rho=E(S)/\{sE(T)\}$.

2.1. *GI/G/s*

At first we will deal with the queueing system based on random walk introduced by Kiefer and Wolfowitz [4]. The queue discipline "first-come, first-served" is adopted.

Let $\tilde{W}_n$ be the following vector with s components

$$\tilde{W}_n=\{W_{n1},\ W_{n2},\cdots,\ W_{ns}\},$$

where $0\leqq W_{n1}\leqq W_{n2}\leqq\cdots\leqq W_{ns}$ and these notations were used by Kiefer and Wolfowitz. Then the basic relationship with respect to the random walk $\{\tilde{W}_n\}$ introduced by the above authors is

$$\tilde{W}_{n+1}=R\{(W_{n1}+S_n-T_n)^+,\ (W_{n2}-T_n)^+,\cdots,(W_{ns}-T_n)^+\},\tag{2.1}$$

where $R\{\tilde{X}\}$ is a vector formed by rearrenging the components of a vector $\tilde{X}$ in order of ascending magnitude and $x^+=\max(0,x)$. The waiting time of the n-th customer is of course just the first component W_{n1}, that is, $W_n=W_{n1}$. Kiefer and Wolfowitz proved that the sequence $\{\tilde{W}_n\}$ converges to a random vector in law regardless of initial vector value under the condition $E(U)<0$. We assume that $\tilde{W}_1$ obeys the limiting (or equilibrium) distribution.

Let
$$X_n=\sum_{i=2}^{s}(W_{ni}-W_n)$$
and
$$Z_n=W_n+S_n-T_n)^-+\sum_{i=2}^{s}(W_{ni}-T_n)^-,$$

where $x^-=-\min(0,x)$. From (2.1) we have a fundamental equation as follows,

$$(2.2)\qquad W_{n+1}+\frac{1}{s}X_{n+1}-\frac{1}{s}Z_n=W_n+\frac{1}{s}X_n+U_n.$$

For, using the equations $x=x^+-x^-$ and (2.1) we have

$$sW_{n+1}+X_{n+1}-(W_n+S_n-T_n)^- - \sum_{i=2}^{s}(W_{ni}-T_n)^-$$
$$=sW_n+X_n+S_n-sT_n.$$

Dividing both sides of the above equation by s, we get (2.2). If we take expectations in (2.2) under the condition $E(U)<0$, then

$$(2.3)\qquad E(Z)=sE(-U)=sE(T)-E(S),$$

since it is assumed that the system is in equilibrium.

Square both sides of (2.2) and note that $Z_nW_{n+1}=0$, being given

$$\left(W_{n+1}+\frac{1}{s}X_{n+1}\right)^2-\frac{2}{s^2}Z_nX_{n+1}+\frac{1}{s^2}Z_n^2$$
$$=\left(W_n+\frac{1}{s}X_n\right)^2+2W_nU_n+\frac{2}{s}U_nX_n+U_n^2.$$

Take expectations in the above equation, noting that U_n is independent of W_n as well as X_n, then

$$(2.4)\qquad 2E(W_n)E(-U_n)=E(U_n^2)+\frac{2}{s}E(U_n)E(X_n)$$
$$+\frac{2}{s^2}E(Z_nX_{n+1})-\frac{1}{s^2}E(Z_n^2).$$

Lemma 2.1. (see Gurland [2]) Let X be an arbitrary random variable whose values are taken on a subset S of the real numbers and let f and g be two functions defined on S. Then if f and g are both non-increasing or both non-decreasing, then

$$(2.5)\qquad E\{f(X)\}E\{g(X)\}\leqq E\{f(X)g(X)\}.$$

If f is non-decreasing and g is non-increasing or vice versa, then we have

(2.6) $E\{f(X)\}E\{g(X)\}\geqq E\{f(X)g(X)\}.$

In (2.5) and (2.6) we suppose one of f and g is continuous.

Lemma 2.2. In the queueing system $GI/G/s$,

$$E(Z_nX_{n+1})\leqq(s-1)E(T_n)E(X_{n+1})=(s-1)E(T)E(X).$$

Proof. Let $X_{ni}=W_{ni}-W_n$, $(i=2,3,\cdots,s)$. Then $X_n=\sum_{i=2}^{s}X_{ni}$. At first we will prove that

(2.7) $Z_nX_{n+1}\leqq(s-1)T_nX_{n+1}.$

From the definitions of Z_n and X_{n+1},

$$Z_n=0 \quad \text{if} \quad W_n+\min(S_n, X_{n2})\geqq T_n$$

and

$$X_{n+1}=0 \quad \text{if} \quad W_n+\max(S_n, X_{ns})\leqq T_n.$$

If either $W_n+S_n<T_n<W_n+X_{ns}$ or $W_n+X_{ns}<T_n<W_n+S_n$ holds, then at least one of terms of the sum $(W_n+S_n-T_n)^-+\sum_{i=2}^{s}(W_n+X_{ni}-T_n)^-$ is zero, that is,

$$Z_n\leqq(s-1)(-T_n)^-=(s-1)T_n.$$

Combining with the above results we have the required inequality (2.7).

Now we will take expectations of both sides of (2.7) and then

$$E(Z_nX_{n+1})\leqq(s-1)E(T_nX_{n+1}).$$

When $W_n, X_{n2},\cdots,X_{ns}$ and S_n are fixed, X_{n+1} is a non-increasing function of T_n. By lemma 2.1 it is deduced that

$$\begin{aligned}&E(T_nX_{n+1}|W_n, X_{n2},\cdots,X_{ns},S_n)\\ &\leqq E(T_n|W_n, X_{n2},\cdots,X_{ns},S_n)E(X_{n+1}|W_n, X_{n2},\cdots,X_{ns},S_n)\\ &=E(T)E(X_{n+1}|W_n, X_{n2},\cdots,X_{ns},S_n),\end{aligned}$$

since T_n is independent of $W_n, X_{n2},\cdots,X_{ns}$ and S_n. Taking expectations of both sides of the inequality on the σ-field $(\tilde{W}_1, S_1, T_1 \cdots, S_{n-1}, T_{n-1}, S_n)$ we have

$$E(T_n X_{n+1}) \leqq E(T)E(X).$$

This complets the proof.

Theorem 2.1. In the queuning system $GI/G/s$ with $s \geqq 2$,

$$E(W) \leqq \frac{V(T)+V(S/s)}{2\{E(T)-E(S/s)\}} \text{ for } \rho \leqq \frac{1}{s},$$

where the case $s=1$ with $\rho<1$ was treated by Marshall [7] and Kingman [5].

Proof. According to lemma 2.2, (2.3), (2.4) and $\rho \leqq 1/s$ it follows that

$$\begin{aligned} &2E(W)E(-U) \\ &\leqq E(U^2)+\frac{2}{s}E(U)E(X)+\frac{2}{s^2}(s-1)E(T)E(X) \\ &\quad -\frac{1}{s^2}E(Z^2) \\ &=E(U^2)-\frac{1}{s^2}E(Z^2)+\frac{2}{s^2}E(X)\{E(S)-E(T)\} \\ &\leqq E(U^2)-\frac{1}{s^2}E(Z^2) \leqq E(U^2)-\frac{1}{s^2}E^2(Z) \\ &=E(U^2)-E^2(U)=V(U)=V(T)+V(S/s). \end{aligned}$$

That is, the required result is derived.

Remark 2.1. Kiefer and Wolfowitz [4] considered the following single-server queueing-system $GI^*/G^*/1$ corresponding to a given $GI/G/s$. That is, the sequence $\{L_n;\ n=1, 2, \cdots\}$ of waiting times, where L_n is the waiting time of the n-th customer in the queueing system $GI^*/G^*/1$, satisfies the recurrence formula;

$$L_{n+1}=[L_n+S_n-sT_n]^+.$$

Then they proved that if $L_1 \leqq sW_1+X_1$,

$$L_n \leqq sW_n + X_n \qquad \text{for all } n.$$

From this fact, a lower bound for $E(W)$ may be obtained by using a lower bound for $E(L)$ and an upper bound for $E(X)$. From the point of view a lower bound for $E(W)$ in the queueing system $GI/G/2$ is obtained and presented in Theorem 2.4.

For the particular case $GI/M/s$ we have the following result.

Theorem 2.2. In the queueing system $GI/M/s$ with $\rho<1$,

$$E(W) \leqq \frac{V(T)+V(S/s)/\rho^2}{2\{E(T)-E(S/s)\}}.$$

Remark 2.2. In particular, Theorem 2.1 may be applied to the case $\rho \leqq 1/s$.

Proof. Kendall [3] show that when W is known to be positive, its conditional distribution is

$$\frac{1}{c}\exp\left\{-\frac{w}{c}\right\}dw \qquad (0<w<\infty),$$

where $c=E(S)/\{s(1-\lambda)\}$ and λ is unique root of the equation

$$\lambda=\int_0^\infty \exp\{-(1-\lambda)sx/E(S)\}\,dP(T\leqq x).$$

From his result we have

$$\begin{aligned}1-\frac{E(S)}{sc} &= \int_0^\infty \exp\left\{-\frac{x}{c}\right\}dP(T\leqq x)\\ &\leqq \int_0^\infty \left\{1-\frac{x}{c}+\frac{x^2}{2c^2}\right\}dP(T\leqq x)\\ &= 1-\frac{1}{c}E(T)+\frac{1}{2c^2}E(T^2),\end{aligned}$$

that is,

$$E(W)\leqq E(W|W>0)=c\leqq\frac{E(T^2)}{2\{E(T)-E(S/s)\}}.$$

Rewriting the right-hand side, our statement is true.

2.2. *GI/G/2*

In order to obtain upper and lower bounds for $E(W)$ of the queueing system $GI/G/2$ the following two lemmas are prepared.

Lemma 2.3. For $GI/G/2$,

$$E(Z_nX_{n+1}) \leqq \frac{E(S^2)}{2} - E(S)E(X). \tag{2.8}$$

Proof. We will show at first that under the condition $W_{n+1,2}>0$,

$$Z_n + X_{n+1} = |X_n - S_n|. \tag{2.9}$$

X_{n+1} may be written by W_n, S_n, T_n and X_n as follows.

$$\begin{aligned} X_{n+1} &= W_{n+1,2} - W_{n+1} \\ &= |(W_n+S_n-T_n)^+ - (W_n+X_n-T_n)^+|. \end{aligned}$$

In the case $X_n \geqq S_n$ we have that $W_n+X_n-T_n \geqq 0$ and $(W_n+X_n-T_n)^+ \geqq (W_n+S_n-T_n)^+$. Then,

$$\begin{aligned} Z_n+X_{n+1} &= (W_n+X_n-T_n)^- + (W_n+S_n-T_n)^- \\ &\quad + (W_n+X_n-T_n)^+ - (W_n+S_n-T_n)^+ \\ &= W_n+X_n-T_n - \{(W_n+S_n-T_n)^+ \\ &\qquad - (W_n+S_n-T_n)^-\} \\ &= W_n+X_n-T_n - (W_n+S_n-T_n) \\ &= X_n - S_n. \end{aligned}$$

In the similar manner we can derive (2.9) for the case $X_n < S_n$. Squaring and taking expectations of both sides of (2.9), then

$$\begin{aligned} &2E(Z_nX_{n+1} | W_{n+1,2}>0) \\ &= E\{(S_n-X_n)^2 | W_{n+1,2}>0\} \\ &\quad - E(X_{n+1}^2 | W_{n+1,2}>0) - E(Z_n^2 | W_{n+1,2}>0). \end{aligned} \tag{2.10}$$

Clearly we have the following inequality

$$E\{(S_n-X_n)^2 | W_{n+1,2}=0\} \geqq 0. \tag{2.11}$$

Multiplying both sides of (2.10) by $P(W_{n+1,2}>0)$ and both sides of (2.11) by $P(W_{n+1,2}=0)$, then

$$\begin{aligned}2E(Z_nX_{n+1})&=2E(Z_nX_{n+1}|W_{n+1,2}>0)P(W_{n+1,2}>0)\\&\leqq E\{(S_n-X_n)^2\}-E(X_{n+1}^2|W_{n+1,2}>0)P(W_{n+1,2}>0)\\&\qquad -E(Z_n^2|W_{n+1,2}>0)P(W_{n+1,2}>0)\\&\leqq E(S^2)-2E(S)E(X),\end{aligned}$$

since $X_{n+1}=0$ when $W_{n+1,2}=0$. Thus lemma 2.3 is true.

Lemma 2.4. For the queueing system $GI/G/2$,

$$E(S)-E(T)\leqq E(X)\leqq\frac{E(S^2)}{2E(S)}. \tag{2.12}$$

Proof. We will show that the inequality of the right-hand side is true. To do this we will present the following inequality

$$X_{n+1}\leqq|S_n-X_n|. \tag{2.13}$$

If $W_{n+1,2}>0$, the inequality is clear from (2.9). In the following we will prove the inequality being true generally.

When $S_n\leqq X_n$,

$$\begin{aligned}X_{n+1}&=(W_n+X_n-T_n)-(W_n+S_n-T_n)\\&=X_n-S_n\end{aligned}$$

for $W_n+S_n-T_n>0$ and

$$X_{n+1}=W_n+X_n-T_n\leqq X_n-S_n$$

for $W_n+X_n-T_n>0$ and $W_n+S_n-T_n\leqq 0$.

Thus (2.13) is true when $S_n\leqq X_n$. In the similar way we can also prove (2.13) to be true when $S_n>X_n$.

Now squaring both sides of (2.13) and taking these expectations, then we get

$$2E(X)E(S)\leqq E(S^2),$$

that is, the inequality of the right-hand side of (2.12) is proved.

In the next, we will show the inequality of the left-hand side of (2. 12) to be true. To do this we will given an inequality,

(2. 14) $$Z_n+X_{n+1}\geqq|S_n-X_n|.$$

For,

$$\begin{aligned} Z_n+X_{n+1}&=(W_n+X_n-T_n)^-+(W_n+S_n-T_n)^- \\ &\quad+|(W_n+X_n-T_n)^+-(W_n+S_n-T_n)^+| \\ &=|S_n-X_n|+2\{W_n+\max(S_n,X_n)-T_n\}^- \\ &\geqq|S_n-X_n|. \end{aligned}$$

From this inequality it follows

$$\begin{aligned} X_n+X_{n+1}&\geqq X_n+|S_n-X_n|-Z_n \\ &\geqq S_n-Z_n. \end{aligned}$$

Taking expectations in the above inequality, then the inequality of the left-hand side is derived immediately.

Theorem 2.3. For the queueing system $GI/G/2$

$$\frac{l}{2}-\frac{E(S^2)}{4E(S)}\leqq E(W)\leqq\frac{V(T)+V(S/2)+\dfrac{E(S^2)}{12}}{2\{E(T)-E(S/2)\}}-\frac{1}{3}E(X),$$

where l is a lower bound for the mean waiting time of the corresponding single-server system to the $GI/G/2$.

Proof. From lemma 2.3,

$$\begin{aligned} &\frac{1}{2}E(Z_nX_{n+1})+E(X)E(U) \\ &\leqq\frac{E(S^2)}{4}-E(S/2)E(X)+E(X)E(U) \\ &=\frac{E(S^2)}{4}-E(X)E(T) \\ &\leqq\frac{E(S^2)}{4}-E(Z_nX_{n+1}). \end{aligned}$$

Then,

$$\frac{3}{2}E(Z_nX_{n+1})+E(X)E(U)\leqq\frac{E(S^2)}{4}.$$

Dividing both sides of this inequality by 3 and adding $2E(X)E(U)/3$, we get

$$\frac{1}{2}E(Z_nX_{n+1})+E(X)E(U)\leqq\frac{E(S^2)}{12}+\frac{2}{3}E(X)E(U).$$

Together with (2.4) and the above inequality an upper bound for $E(W)$ is obtained.

To find a lower bound for $E(W)$ we will consider the corresponding single-server system stated in Remark of Theorem 2.1 to the given $GI/G/2$. The waiting time L_n of the n-th customer in the corresponding system satisfies the following inequality,

$$L_n\leqq 2W_n+X_n.$$

Taking expectations of both sides of the above inequality and using (2.12), we get a lower bound for $E(W)$.

We will close this section with an example. In the queueing system $M/M/s$, $E(W)$ is given by the following well-known formula;

$$E(W)=\frac{s^{s-1}\rho^s}{s!\,\mu(1-\rho)^2}P_0$$

and

$$P_0=\left[\sum_{n=0}^{s-1}\frac{a^n}{n!}+\frac{a^s}{(s-1)!\,(s-a)}\right]^{-1},$$

where $\mu=1/E(S)$, $\lambda=1/E(T)$, $a=\lambda/\mu$ and $\rho=a/s$.

Let $E^*(W)$ be the upper bound for $E(W)$ being given in Theorem 2.1 (when $\rho\leqq 1/s$) and Theorem 2.2. Then

$$E^*(W)=\begin{cases}\dfrac{1+\rho^2}{2(1-\rho)\lambda} & \text{if} \quad \rho\leqq 1/s,\\[2ex] \dfrac{1}{(1-\rho)\lambda} & \text{if} \quad 1/s<\rho<1.\end{cases}$$

In a particular case $s=2$, a lower bound for $E(W)$ can be obtained by Theorem 2.3. Using the expression (21) of Marshall's paper [7], we can obtain a lower bound $E_*(W)$ for $E(W)$.

$$E_*(W)=\frac{\lambda\{V(T)+V(S/2)\}}{2(1-\rho)}-\frac{1+\rho}{4\lambda}-\frac{E(S^2)}{4E(S)},$$
$$=\frac{1+\rho^2}{2(1-\rho)\lambda}-\frac{1+\rho}{4\lambda}-\rho.$$

For the simplicity we will put $\lambda=1$ and then get the following tables.

Table 2.1. $s=2$

ρ	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99
$E^*(W)$	0.56	0.65	0.77	0.96	1.25	2.50	3.33	5.00	10.00	25.00	100.00
$E(W)$	0.00	0.01	0.04	0.15	0.25	0.61	1.34	2.85	7.67	17.56	97.40
$E_*(W)$	0.00	0.00	0.00	0.00	0.10	0.30	0.93	2.40	7.20	17.10	97.00

Table 2.2. $s=3$

ρ	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99
$E^*(W)$	0.56	0.65	0.77	1.66	2.00	2.50	3.33	5.00	10.00	25.00	100.00
$E(W)$	0.00	0.00	0.03	0.09	0.20	0.53	1.15	2.60	7.35	17.25	97.15

In the equeueing system $M/M/2$, $E^*(W)$ and $E_*(W)$ are both good approximations for $E(W)$ as $\rho\to 1$. Similar argument will be also stated in the queueing system $M/M/3$.

3. Two bulk queues

3.1. *M(batch-arrival)/G(batch-service)/1*

Let t_n be the n-th departure time and Q_n be the queue-length just after t_n. Let X_n be the number of arrived customers during the service time in $(t_{n-1}+0,\ t_n]$ and Y_n be the capacity for service ending at time t_{n+1}. If $Q_n \leqq Y_n$, all customers in queue are served and if $Q_n > Y_n$, Y_n customers are served at t_n. Let C_n be the size of the n-th arrival batch and the sequence $\{C_n\}$ be mutually independent with a common distribution $P(C_n=j)=c_j$, $(j=1, 2, \cdots)$. We assume in this section that the arrival process of batches is Poissonian with parameter λ and service times are mutually independent with a common distribution function $G(t)$ $(0 \leqq t < \infty)$. Then the distribution of X_n is determined as follows.

$$P(X_n=j)=\int_0^\infty \sum_{k=0}^{j} \frac{(\lambda t)^k}{k!} c_j^{(k)} dG(t),$$

where the distribution $\{c_j^{(k)}, j \geqq 1\}$ is the k-fold convolution of the distribution $\{c_j, j \geqq 1\}$ with itself. Also it is assumed that the sequence $\{Y_n\}$ is mutually independent with a common distribution $P(Y_n=j)=b_j$, $(j=0, 1, \cdots)$ and independent of arrival batches as well as queue sizes. It was proved by Bhat [1] that if $E(X)<E(Y)$, the number of customers in queue after departure converges to a random variable regardless to initial queue size with probability one and if $E(X) \geqq E(Y)$ it diverges with probability one. We shall assume, of course, that $E(X)<E(Y)$ and Q_1 obeys the limiting (or equilibrium) distribution throughout the section. As we see Bhat's paper, the sequence $\{Q_n\}$ may be considered as a Markov chain satisfying the following recurrence formula.

(3.1) $$Q_{n+1}=[Q_n-Y_n]^+ + X_{n+1}.$$

Let $Z_n=[Y_n-Q_n]^+$. Hence equation (3.1) may be rewritten as follows,

(3.2) $$(Q_{n+1}-X_{n+1})-Z_n=Q_n-Y_n.$$

Taking expectations in (3.2), we have

(3.3) $$E(Z)=E(Y)-(X).$$

Squaring both sides of (3.2) and taking these expectations, then

(3.4)
$$-2E(Q_{n+1}X_{n+1})+E(X^2_{n+1})+E(Z_n^2)$$
$$=-2E(Q_n)E(Y_n)+E(Y_n^2),$$

where we use the relation $(Q_{n+1}-X_{n+1})\ Z_n=0$ and notice that Q_n is independent of Y_n.

Multiply both sides of (3.2) by X_{n+1} and take their expectations, then

(3.5)
$$E(Q_{n+1}X_{n+1})=E(X_{n+1})\{E(Q_n)-E(Y_n)\}$$
$$+E(X^2_{n+1})+E(X_{n+1})E(Z_n)$$

from (3.3).

Substitute (3.5) into (3.4),

$$2E(Q)\{E(Y)-E(X)\}$$
$$=E(Y^2)+2V(X)-E(X^2)-E(Z^2)$$
$$=V(Y)+V(X)+2E(X)\{E(Y)-E(X)\}-V(Z),$$

that is

(3.6)
$$E(Q)=E(X)+\frac{V(Y)+V(X)-V(Z)}{2E\{(Y)-E(X)\}}.$$

To obtain an expression of $V(Q)$ we will cube both sides of (3.2) and take their expectations. Then

(3.7)
$$-3E(Q^2_{n+1}X_{n+1})+3E(Q_{n+1}X^2_{n+1})-E(X^3_{n+1})-E(Z_n^3)$$
$$=-3E(Q_n^2)E(Y_n)+3E(Q_n)E(Y_n^2)-E(Y_n^3),$$

where we use the relation $(Q_{n+1}-X_{n+1})\ Z_n=0$ and note that Q_n is independent of Y_n. Also, squaring both sides of (3.2) and multiplying them by X_{n+1} and then taking expectations, we get

$$E(Q_{n+1}{}^2X_{n+1})=E(X_{n+1})E\{(Q_n-Y_n)^2\}+2E(Q_{n+1}X^2_{n+1})$$
$$-E(X^3_{n+1})-E(X_{n+1})E(Z_n^2),$$

where we note that X_{n+1} is independent of Q_n-Y_n. Further, multiply both sides of (3.2) by X^2_{n+1} and take expectations, then

$$E(Q_{n+1}X^2_{n+1})=E(X^2_{n+1})E(Q_n-Y_n+Z_n)+E(X^3_{n+1}).$$

Substituting these two equations into (3.7) and using (3.3) and (3.6), it follsws that

$$V(Q)=\frac{\{V(Y)+V(X)\}^2-\{V(Z)\}^2}{4\{E(Y)-E(X)\}^2} + \frac{1}{2}\{V(Y)+V(X)-V(Z)\}+E(X)E(Y) + \frac{E(X)\{V(Y)-V(X)\}}{E(Y)-E(X)}+\frac{E(X^3)-E(Y^3)+E(Z^3)}{3\{E(Y)-E(X)\}}. \tag{3.8}$$

Since $X_n \leqq Q_n$ from (3.1), we have

$$Z_n \leqq [Y_n - X_n]^+,$$

Then,

$$0 \leqq V(Z) \leqq V(Y)+V(X)$$

and

$$0 \leqq E(Z^3) \leqq E(Y^3).$$

Using these relations with (3.6) and (3.8), we have the following theorem.

Theorem 3.1. For the queueing system *M(batch-arrival)/G(batch-service)/*1,

$$E(X) \leqq E(Q) \leqq E(X)+\frac{V(Y)+V(X)}{2\{E(Y)-E(X)\}} \tag{3.9}$$

and

$$V(Q) \leqq \left[\frac{V(Y)+V(X)}{2\{E(Y)-E(X)\}}\right]^2+\frac{1}{2}\{V(Y)+V(X)\} + E(Y)E(X)+\frac{E(X)\{V(Y)-V(X)\}}{E(Y)-E(X)} + \frac{E(X^3)}{3\{E(Y)-E(X)\}}. \tag{3.10}$$

We will consider an example. $E(Q)$ and $V(Q)$ are given in the queueing system $M/M/1$ as follows.

$$E(Q)=\frac{\rho}{1-\rho}$$

and

$$V(Q)=\frac{\rho}{(1-\rho)^2}$$

Let $E^*(Q)$ be the upper bound and $E_*(Q)$ be the lower bound given in (3.9). Also let $V^*(Q)$ be the upper bound given in (3.10). Then,

$$E^*(Q)=\rho+J,\ E_*(Q)=\rho,$$

and $$V^*(Q)=J^2+(1-\rho)J+\frac{\rho(3\rho^2+4)}{3(1-\rho)},$$

where $J=\frac{\rho(1+\rho)}{2(1-\rho)}$. For several values of ρ, we have the following table.

Table 3.1.

ρ	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99
$E^*(Q)$	0.16	0.35	0.57	0.86	1.25	1.80	2.68	4.40	9.45	19.47	99.49
$E(Q)$	0.11	0.25	0.42	0.66	1.00	1.50	2.33	4.00	9.00	19.00	99.00
$E_*(Q)$	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99
$\sqrt{V^*(Q)}$	0.45	0.69	0.94	1.22	1.59	2.11	2.69	4.64	9.66	19.66	99.67
$\sqrt{V(Q)}$	0.35	0.56	0.87	1.05	1.41	1.94	2.78	4.47	9.49	19.49	99.49

We shall see from this table that the lower bound $E_*(Q)$ is no good approximation and both upper bounds are good approximation for all ρ.

3.2. *GI(batch-arrival)/M(batch-service)/1*

Let t_n be the n-th arrival time and X_n be the size of the batch arriving at t_{n-1}. The sequence $\{X_n\}$ is mutually independent with a common distribution $P(X_n=j)=b_j$, $(j=0, 1, \cdots)$. Customers are served in batches of variable capacity. Let C_n^* be the capacity of the n-th service and the sequence $\{C_n^*\}$ be mutually independent with a common distribution $P(C_n^*=j)=c_j^*$, $(j=0, 1, 2, \cdots)$. The service mechanism is such that when there is vacancy in the batch being served, the arriving customer will join the batch immediately untill its capacity is reached. The rest of the batch of arrivals will wait for the next service. Let Y_n be the total capacity of the batches that would be served during the period $[t_{n-1}, t_n)$ and Q_n be the number of customers including those who are being served before the n-th arrival.

Bhat [1] proved that if $E(X)<E(Y)$, Q_n converges to a random variable regardless of initial queue size with probability one and if $E(X)\geqq E(Y)$ diverges with probability one. We shall only consider the case $E(X)<E(Y)$ and assume that Q_1 obeys the limiting (or equilibrium) distribution.

The sequence $\{Q_n\}$ satisfies the following recurrence formula,

$$(3.11)\qquad Q_{n+1}=[Q_n+X_{n+1}-Y_{n+1}]^+.$$

Put $Z_n=[Q_n+X_{n+1}-Y_{n+1}]^-$, then (3.11) may be written by

$$(3.12)\qquad Q_{n+1}-Z_n=Q_n+X_{n+1}-Y_{n+1}.$$

Taking expectations of the above equation, we have

$$(3.13)\qquad E(Z)=E(Y)-E(X).$$

Squaring both sides of (3.12) and taking expectations, then

$$E(Z^2)=-2E(Q)E(Y-X)+E\{(X-Y)^2\},$$

where we note the relation $Q_{n+1}Z_n=0$. That is, by using (3.13)

$$(3.14)\qquad E(Q)=\frac{E\{(X-Y)^2\}-E(Z^2)}{2\{E(Y)-E(X)\}}=\frac{V(X)+V(Y)-V(Z)}{2\{E(Y)-E(X)\}}$$

To obtain the variance $V(Q)$ we will cube both sides of (3.12) and take their expectations, then we get

$$-E(Z^3)=3E(Q^2)E(X-Y)+3E(Q)E\{(X-Y)^2\}+E\{(X-Y)^3\}$$

where we note that $Q_{n+1}Z_n=0$. With this equation, (3.13) and (3.14) we have

$$(3.15)\qquad V(Q)=\frac{\{V(X)+V(Y)\}^2-\{V(Z)\}^2}{4\{E(Y)-E(X)\}}+\frac{1}{2}\{V(X)+V(Y)-V(Z)\}+E(X)E(Y)+\frac{E(X^3)-E(Y^3)+E(Z^3)}{3\{E(Y)-E(X)\}}$$

Also it is easily seen that $0\leqq E(Z^3)\leqq E(Y^3)$.

Theorem 3.2. For the queueing system

GI/(batch-arrival)/M(batch-service)/1

$$(3.16)\qquad E(Q)\leqq\frac{V(Y)+V(X)}{2\{E(Y)-E(X)\}}$$

and

$$(3.17)\qquad V(Q)\leqq\left[\frac{V(Y)+V(X)}{2\{E(Y)-E(X)\}}\right]^2+\frac{1}{2}\{V(Y)+V(X)\}+E(X)E(Y)+\frac{E(X^3)}{3\{E(Y)-E(X)\}}.$$

REFERENCES

[1] Bhat, U.N., Imbedded Markov chain analysis of single server bulk queues, *Jour. Aust. Math. Sci.*, 4 (1964), 244-262.

[2] Gurland, J., An inequality satisfied by the expectation of the reciprocal of a random variable, *Amer. Stat.*, 21, 3 (1967), 24-25.

[3] Kendall, D.G., Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain, *Ann. Math. Statist.*, 24 (1953), 338-354.

[4] Kiefer, J., and J. Wolfowitz, On the theory of queues with many servers, *Trans. Amer. Math. Soc.*, 78 (1955), 1-18.

[5] Kingman, J.F.C., On queues in heavy traffic, *J. Roy. Statist. Soc., Ser. B*, 24 (1962), 383-392.

[6] Kingman, J.F.C., The heavy traffic approximation in the theory of queues, Proceedings of the symposium on congestion theory, *The University of North Carolina Press* (1965).

[7] Marshall, K.T., Some inequalities in queuing, *Opns. Res.*, 16, 3 (1968), 651-665.

[8] Takács, L., Introduction to the theory of queues, *Oxford Univ. Press, N.Y.* (1962).