Bull. Korean Math. Soc. **46** (2009), No. 1, pp. 107–116

QUEUE LENGTH DISTRIBUTION IN A QUEUE WITH RELATIVE PRIORITIES

Jeongsim Kim

Abstract. We consider a single server multi-class queueing model with Poisson arrivals and relative priorities. For this queue, we derive a system of equations for the transform of the queue length distribution. Using this system of equations we find the moments of the queue length distribution as a solution of linear equations.

1. Introduction

We consider a multi-class queueing model with relative priorities. In the relative priority service discipline for a single server (processor) system with K classes of customers, if at some service completion there are n_j customers of class j, $j = 1, \ldots, K$, then the next customer to commence service is selected from class i customers with probability

$$\frac{n_i p_i}{\sum_{j=1}^{K} n_j p_j}, \quad i = 1, \ldots, K.$$

Once a customer has started service, it is served without interruption until completion.

Relative priority model is related to the well-known model of discriminatory processor sharing (DPS), see the recent survey [1]. An essential difference with DPS is that for DPS all customers in the system are served simultaneously by a single processor, whereas in relative priority model, the processor serves customers one at a time until their service has been completed.

A single server multi-class queueing model with relative priorities was first suggested in [2]. For the analysis of queueing model with relative priorities it seems that Haviv and van der Wal [3] is the only known result in open literatures. Haviv and van der Wal [3] obtained the mean waiting times for the M/G/1 queue with relative priorities.

Received May 5, 2008.
2000 *Mathematics Subject Classification.* 60K25.
Key words and phrases. M/G/1 queue, relative priority model, queue length distribution.
This work was supported by the research grant of the Chungbuk National University in 2007.

In this paper we consider a single server multi-class queueing model with Poisson arrivals and relative priorities. For this queue, we derive a system of equations for the transform of the queue length distribution. Using this system of equations we find the moments of the queue length distribution as a solution of linear equations. Besides, numerical examples are given.

2. Transform of the queue length distribution

We consider an M/G/1 queue with relative priorities and K classes of customers. Each class i customer has a positive priority parameter p_i, $i = 1, \ldots, K$. Customers of class i arrive in a Poisson stream with rate λ_i. The overall arrival rate is $\lambda = \sum_{i=1}^{K} \lambda_i$. The service times of class i customers, denoted by random variable X_i, have an identical distribution function $B_i(t)$ with Laplace-Stietjes transform $B_i^*(s) = \int_0^\infty e^{-st} dB_i(t)$. The traffic intensity for class i customer is $\rho_i = \lambda_i \mathbb{E}[X_i]$ and the total traffic intensity is $\rho = \sum_{j=1}^{K} \rho_j$.

Let $N_i(t)$, $i = 1, \ldots, K$, be the number of class i customers in the system at time t. Let τ_n be the nth departure epoch. Then $\{(N_1(\tau_n+), \ldots, N_K(\tau_n+)) : n = 1, 2, \ldots\}$ is a Markov chain, called an embedded Markov chain (EMC). We observe that

$$\begin{aligned}
&\mathbb{P}((N_1(\tau_{n+1}+), \ldots, N_K(\tau_{n+1}+)) \\
&\quad = (l_1, \ldots, l_K) \mid (N_1(\tau_n+), \ldots, N_K(\tau_n+)) = (n_1, \ldots, n_K)) \\
&= \begin{cases} \sum_{i=1}^{K} \frac{\lambda_i}{\lambda} b_i(l_1, \ldots, l_K) & \text{if } (n_1, \ldots, n_K) = (0, \ldots, 0), \\ \sum_{i=1}^{K} \frac{n_i p_i}{n_1 p_1 + \cdots + n_K p_K} b_i\left((l_1, \ldots, l_K) - (n_1, \ldots, n_K) + \mathbf{1}_i\right) & \text{if } (n_1, \ldots, n_K) \neq (0, \ldots, 0), \end{cases}
\end{aligned} \tag{1}$$

where

$$b_j(l_1, \ldots, l_K) = \begin{cases} 0 & \text{if } (l_1, \ldots, l_K) = (0, \ldots, 0), \\ \int_0^\infty e^{-\lambda t} \frac{\lambda_1^{l_1} \cdots \lambda_K^{l_K} t^{l_1 + \cdots + l_K}}{l_1! \cdots l_K!} dB_j(t) & \text{otherwise}, \end{cases}$$

which is the probability that the number of each class i customer which arrives during the service time of class j customer, is l_i, $i = 1, \ldots, K$. Further, $\mathbf{1}_i$ denotes a K-dimensional row vector whose ith component is 1 and all other components are 0.

It can be shown that the EMC is positive recurrent if and only if $\rho < 1$. We assume $\rho < 1$. Let $\pi(l_1, \ldots, l_K)$ be the stationary distribution of the EMC:

$$\pi(l_1, \ldots, l_K) = \lim_{n \to \infty} \mathbb{P}\left((N_1(\tau_n+), \ldots, N_K(\tau_n+)) = (l_1, \ldots, l_K)\right).$$

Let us denote by $\Pi(z_1, \ldots, z_K)$ the probability generating function of $\pi(l_1, \ldots, l_K)$:

$$\Pi(z_1, \ldots, z_K) \equiv \sum_{l_1=0}^{\infty} \cdots \sum_{l_K=0}^{\infty} \pi(l_1, \ldots, l_K) z_1^{l_1} \cdots z_K^{l_K}. \tag{2}$$

From (1), we have

$$
\begin{aligned}
&\Pi(z_1,\ldots,z_K)\\
&= \pi(0,\ldots,0)\sum_{i=1}^{K}\frac{\lambda_i}{\lambda}B_i^*(\lambda-\sum_{j=1}^{K}\lambda_j z_j)+\sum\cdots\sum_{(n_1,\ldots,n_K)\neq(0,\ldots,0)}\pi(n_1,\ldots,n_K)\\
&\quad\times\sum_{i=1}^{K}\frac{n_i p_i}{n_1p_1+\cdots+n_Kp_K}z_1^{n_1}\cdots z_i^{n_i-1}\cdots z_K^{n_K}B_i^*(\lambda-\sum_{j=1}^{K}\lambda_j z_j).
\end{aligned}\tag{3}
$$

Since equation (3) is somewhat cumbersome to work with directly, we define

$$
r(z_1,\ldots,z_K)\equiv\sum\cdots\sum_{(l_1,\ldots,l_K)\neq(0,\ldots,0)}\frac{\pi(l_1,\ldots,l_K)}{l_1p_1+\cdots+l_Kp_K}z_1^{l_1}\cdots z_K^{l_K}.\tag{4}
$$

Further, we note that

$$
\pi(0,\ldots,0)=1-\rho.\tag{5}
$$

In the following theorem we have a system of equations for the probability generating function of the queue length distribution.

Theorem 1. (a) $\Pi(z_1,\ldots,z_K)$ *and* $r(z_1,\ldots,z_K)$ *are related by*

$$
\Pi(z_1,\ldots,z_K)=1-\rho+\sum_{i=1}^{K}p_i z_i\frac{\partial}{\partial z_i}r(z_1,\ldots,z_K).\tag{6}
$$

(b) $r(z_1,\ldots,z_K)$ *satisfies*

$$
\begin{aligned}
&\sum_{i=1}^{K}p_i\left(z_i-B_i^*(\lambda-\sum_{j=1}^{K}\lambda_j z_j)\right)\frac{\partial}{\partial z_i}r(z_1,\ldots,z_K)\\
&=(\rho-1)\left(1-\sum_{i=1}^{K}\frac{\lambda_i}{\lambda}B_i^*(\lambda-\sum_{j=1}^{K}\lambda_j z_j)\right).
\end{aligned}\tag{7}
$$

Proof. Part (a) is immediate from (2), (4) and (5). Now, by using (4) and (5), we can rewrite (3) as

$$
\begin{aligned}
\Pi(z_1,\ldots,z_K)&=(1-\rho)\sum_{i=1}^{K}\frac{\lambda_i}{\lambda}B_i^*(\lambda-\sum_{j=1}^{K}\lambda_j z_j)\\
&\quad+\sum_{i=1}^{K}p_i\frac{\partial}{\partial z_i}r(z_1,\ldots,z_K)B_i^*(\lambda-\sum_{j=1}^{K}\lambda_j z_j).
\end{aligned}
$$

Part (b) follows immediately from substituting (6) into the above equation. □

Note that the equation (7) reduces to

$$
\Pi(z_1,\ldots,z_K)=\frac{1-\rho}{1-\rho z},
$$

which is the probability generating function of the queue length in the ordinary M/M/1 queue.

Theorem 1 helps us obtain the moments of the queue length for the M/G/1 queue with relative priorities, as we see in the following section.

3. Moments of the queue length distribution

In this section we show how to obtain the first and second moments of the queue length. To do this, we define the following moments:

$$\begin{aligned} L^j_{i_1\cdots i_j} &\equiv \left.\frac{\partial^j}{\partial z_{i_1}\cdots\partial z_{i_j}}\Pi(z_1,\ldots,z_K)\right|_{z_1=\cdots=z_K=1}, \\ R^j_{i_1\cdots i_j} &\equiv \left.\frac{\partial^j}{\partial z_{i_1}\cdots\partial z_{i_j}}r(z_1,\ldots,z_K)\right|_{z_1=\cdots=z_K=1}, \end{aligned}$$

where $j = 1, 2, \ldots,$ $1 \le i_l \le K$ and $1 \le l \le j$. We also define, for $1 \le l_k \le j$,

$$\begin{aligned} &\mathcal{R}^{j-k+1}_{i_1\cdots i_j i}(i_{l_1},\ldots,i_{l_k}) \\ \equiv &\left.\frac{\partial^{j-k+1}}{\partial z_{i_1}\cdots\partial z_{i_{l_1-1}}\partial z_{i_{l_1+1}}\cdots\partial z_{i_{l_k-1}}\partial z_{i_{l_k+1}}\cdots\partial z_{i_j}\partial z_i}r(z_1,\ldots,z_K)\right|_{z_1=\cdots=z_K=1}, \end{aligned}$$

which means that we take partial derivatives with respect to variables $z_{i_1},\ldots,$ z_{i_j}, z_i except $z_{i_{l_1}},\ldots,z_{i_{l_k}}$. For example,

$$\mathcal{R}^j_{i_1\cdots i_j i}(i_l) = \left.\frac{\partial^j}{\partial z_{i_1}\cdots\partial z_{i_{l-1}}\partial z_{i_{l+1}}\cdots\partial z_{i_j}\partial z_i}r(z_1,\ldots,z_K)\right|_{z_1=\cdots=z_K=1}.$$

More specifically, $\mathcal{R}^3_{1236}(2) = R^3_{136}$, $\mathcal{R}^2_{123}(2) = R^2_{13}$, $\mathcal{R}^1_{246}(2,4) = R^1_6$. We have the following relations on the moments of the queue length.

Theorem 2. *For $j = 1, 2, \ldots,$ $1 \le i_l \le K$ and $1 \le l \le j$,*

$$L^j_{i_1\cdots i_j} = \sum_{i=1}^{K} p_i R^{j+1}_{i_1\cdots i_j i} + \sum_{l=1}^{j} p_{i_l} R^j_{i_1\cdots i_j}, \tag{8}$$

and

$$\begin{aligned} &\sum_{l=1}^{j} p_{i_l} R^j_{i_1\cdots i_j} - \sum_{i=1}^{K}\sum_{k=1}^{j}\sum_{1\le l_1<\cdots<l_k\le j} p_i\phi_{i,i_{l_1}\cdots i_{l_k}}\mathcal{R}^{j-k+1}_{i_1\cdots i_j i}(i_{l_1},\ldots,i_{l_k}) \\ &= (1-\rho)\sum_{i=1}^{K}\frac{\lambda_i}{\lambda}\phi_{i,i_1\cdots i_j}, \end{aligned} \tag{9}$$

where $\phi_{i,i_1\cdots i_k}$ is given by

$$\begin{aligned}\phi_{i,i_1\cdots i_k} &\equiv \frac{\partial^k}{\partial z_{i_1}\cdots\partial z_{i_k}} B_i^*(\lambda - \sum_{j=1}^{K}\lambda_j z_j)\Big|_{z_1=\cdots=z_K=1} \\ &= \lambda_{i_1}\cdots\lambda_{i_k}\mathbb{E}(X_i^k).\end{aligned}$$

Proof. The equation (8) is derived by taking partial derivatives of (6) with respect to $z_{i_1}, z_{i_2}, \ldots, z_{i_j}$ and evaluating at $z_1 = \cdots = z_K = 1$. Now we derive (9). For notational convenience, we set in equation (7)

$$I_i(z_1, \ldots, z_K) = p_i\left(z_i - B_i^*(\lambda - \sum_{j=1}^{K}\lambda_j z_j)\right).$$

Note that $I_i(1, \ldots, 1) = 0$. By taking partial derivatives of (7) with respect to $z_{i_1}, z_{i_2}, \ldots, z_{i_j}$ and evaluating at $z_1 = \cdots = z_K = 1$, we have

(10)
$$\sum_{i=1}^{K}\sum_{k=1}^{j}\sum_{1\le l_1<\cdots<l_k\le j}\frac{\partial^k}{\partial z_{i_{l_1}}\cdots\partial z_{i_{l_k}}} I_i(z_1, \ldots, z_K)\Big|_{z_1=\cdots=z_K=1}\cdot\mathcal{R}^{j-k+1}_{i_1\cdots i_j i}(i_{l_1}, \ldots, i_{l_k})$$
$$= (1-\rho)\sum_{i=1}^{K}\frac{\lambda_i}{\lambda}\phi_{i,i_1\cdots i_j}.$$

Observe that for $k = 1$, the left hand side of (10) becomes

$$\sum_{l_1=1}^{j} p_{i_{l_1}} R^j_{i_1\cdots i_j} - \sum_{i=1}^{K}\sum_{l_1=1}^{j} p_i\phi_{i,i_{l_1}}\mathcal{R}^j_{i_1\cdots i_j i}(i_{l_1}),$$

and for $k \ge 2$,

$$\sum_{i=1}^{K}\sum_{k=2}^{j}\sum_{1\le l_1<\cdots<l_k\le j}\frac{\partial^k}{\partial z_{i_{l_1}}\cdots\partial z_{i_{l_k}}} I_i(z_1, \ldots, z_k)\Big|_{z_1=\cdots=z_K=1}\cdot\mathcal{R}^{j-k+1}_{i_1\cdots i_j i}(i_{l_1}, \ldots, i_{l_k})$$
$$= -\sum_{i=1}^{K}\sum_{k=2}^{j}\sum_{1\le l_1<\cdots<l_k\le j} p_i\phi_{i,i_{l_1}\cdots i_{l_k}}\mathcal{R}^{j-k+1}_{i_1\cdots i_j i}(i_{l_1}, \ldots, i_{l_k}).$$

Combining the above two equations, we can rewrite (10) as (9). This finishes the derivation of (9). □

Now we describe in detail how to calculate the first and second moments of the queue length. From (9) with $j = 1$, a system of linear equations for R^1_k, $1 \le k \le K$, is given as follows:

(11)
$$p_k R^1_k - \lambda_k\sum_{i=1}^{K} p_i\mathbb{E}(X_i)R^1_i = \frac{\lambda_k}{\lambda}(1-\rho)\rho,\ k = 1, \ldots, K.$$

From (9) with $j = 2$, a system of linear equations for R^2_{kl}, $1 \le k, l \le K$, is given as follows:
(12)

$$(p_k + p_l)R^2_{kl} - \lambda_k \sum_{i=1}^{K} p_i \mathbb{E}(X_i) R^2_{li} - \lambda_l \sum_{i=1}^{K} p_i \mathbb{E}(X_i) R^2_{ki} - \lambda_k \lambda_l \sum_{i=1}^{K} p_i \mathbb{E}(X_i^2) R^1_i$$
$$= (1-\rho)\frac{\lambda_k \lambda_l}{\lambda} \sum_{i=1}^{K} \lambda_i \mathbb{E}(X_i^2), \ 1 \le k \le l \le K.$$

From (9) with $j = 3$, a system of linear equations for R^3_{klm}, $1 \le k, l, m \le K$, is given as follows:
(13)

$$\begin{aligned}
&(p_k + p_l + p_m)R^3_{klm} - \lambda_k \sum_{i=1}^{K} p_i \mathbb{E}(X_i) R^3_{lmi} - \lambda_l \sum_{i=1}^{K} p_i \mathbb{E}(X_i) R^3_{kmi} \\
&- \lambda_m \sum_{i=1}^{K} p_i \mathbb{E}(X_i) R^3_{kli} - \lambda_k \lambda_l \sum_{i=1}^{K} p_i \mathbb{E}(X_i^2) R^2_{mi} - \lambda_l \lambda_m \sum_{i=1}^{K} p_i \mathbb{E}(X_i^2) R^2_{ki} \\
&- \lambda_m \lambda_k \sum_{i=1}^{K} p_i \mathbb{E}(X_i^2) R^2_{li} - \lambda_k \lambda_l \lambda_m \sum_{i=1}^{K} p_i \mathbb{E}(X_i^3) R^1_i \\
= &(1-\rho)\frac{\lambda_k \lambda_l \lambda_m}{\lambda} \sum_{i=1}^{K} \lambda_i \mathbb{E}(X_i^3), \qquad 1 \le k \le l \le m \le K.
\end{aligned}$$

Thus we can obtain the first and second moments of the queue length as follows. Using (11), we first obtain R^1_k, $k = 1, \dots, K$. Secondly, we obtain R^2_{kl}, $k, l = 1, \dots, K$, by solving the system of linear equations (12). Lastly, we obtain the mean queue length L^1_k, $k = 1, \dots, K$, from

$$L^1_k = \sum_{i=1}^{K} p_i R^2_{ki} + p_k R^1_k \tag{14}$$

by (8). Then, solving the system of linear equations (13) yields R^3_{klm}, $k, l, m = 1, \dots, K$. Similarly, we obtain the second moment L^2_{kl}, $k, l = 1, \dots, K$, from

$$L^2_{kl} = \sum_{i=1}^{K} p_i R^3_{kli} + p_k R^2_{kl} + p_l R^2_{kl} \tag{15}$$

by (8).

In summary, we have the following procedure for the first and second moments of the queue length.

Procedure for the first and second moments of the queue length

1. Calculate R^1_k, $k = 1, \dots, K$, by solving (11).
2. Calculate R^2_{kl}, $k, l = 1, \dots, K$, by solving (12).
3. Calculate the mean queue length L^1_k, $k = 1, \dots, K$, by (14).

4. Calculate R^3_{klm}, $k, l, m = 1, \ldots, K$, by solving (13).
5. Calculate the second moment of the queue length L^2_{kl}, $k, l = 1, \ldots, K$, by (15).

Remark. In a similar manner we can obtain the higher moments of the queue length by using the equations (8) and (9).

4. Numerical examples

In this section we present some numerical examples to compute the mean, variance and squared coefficient of variation of the queue length. We consider the case of $K = 2$ customer classes with priority parameters p_1 and p_2. We assume equal loads of $\rho_1 = \rho_2 = 0.35$, and hence $\rho = 0.7$.

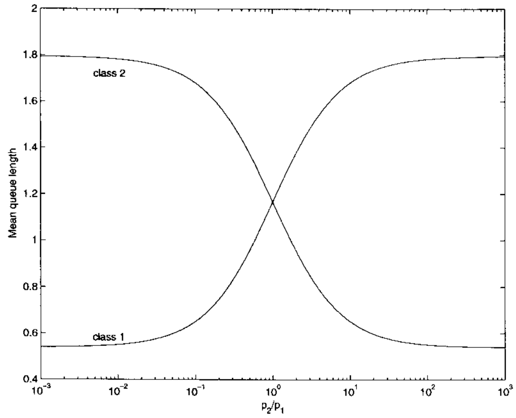

FIGURE 1. Mean queue length for Example 1.

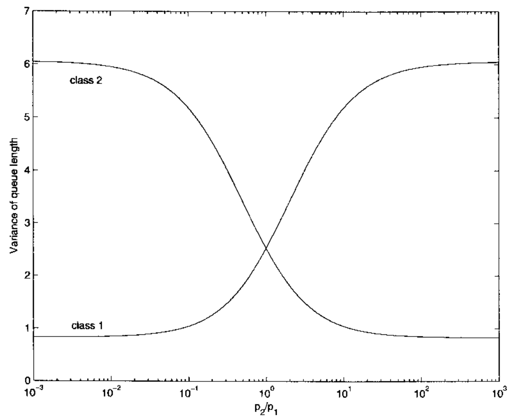

FIGURE 2. Variance of queue length for Example 1.

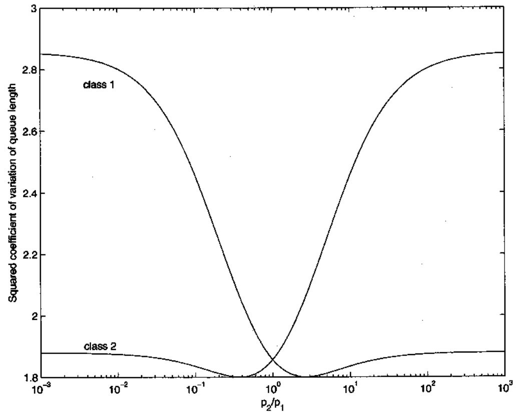

FIGURE 3. Squared coefficient of variation of queue length for Example 1.

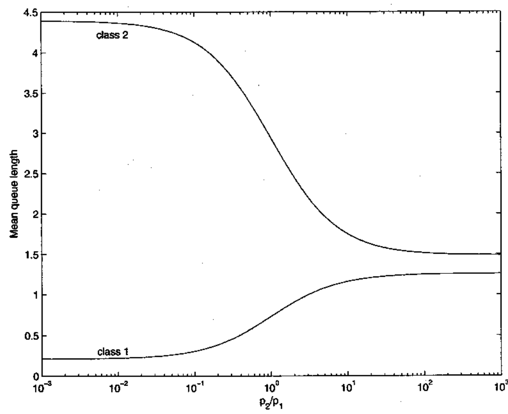

FIGURE 4. Mean queue length for Example 2.

Example 1. We assume that the service times of customers in each class have exponential distribution with density $b(t) = e^{-t}$. Hence $\lambda_1 = \lambda_2 = 0.35$.

Figs. 1-3 show the mean, variance and squared coefficient of variation for the queue length of each class, respectively, varying the priority parameter ratio p_2/p_1. We observe, as expected, that in the case of $p_2/p_1 = 1$, the mean, variance and squared coefficient of variation for the queue length of class 1 customer equal to those of class 2 customer. It is also observed that the curves in each figure are symmetric with respect to the vertical line passing through $p_2/p_1 = 1$, as expected. From Figs. 1 and 2, we see, as expected, that the mean and variance of the queue length of class 1 customer (resp. class 2 customer) increase (resp. decrease) as the priority parameter ratio p_2/p_1 increases.

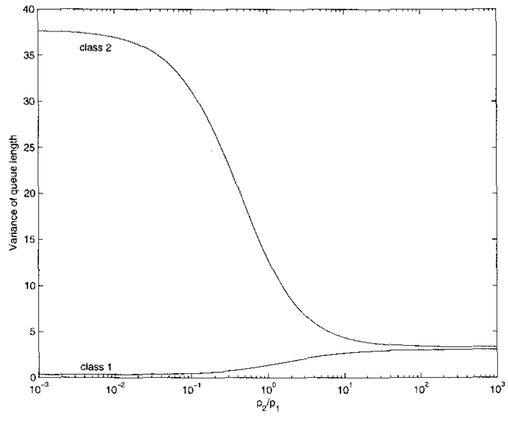

FIGURE 5. Variance of queue length for Example 2.

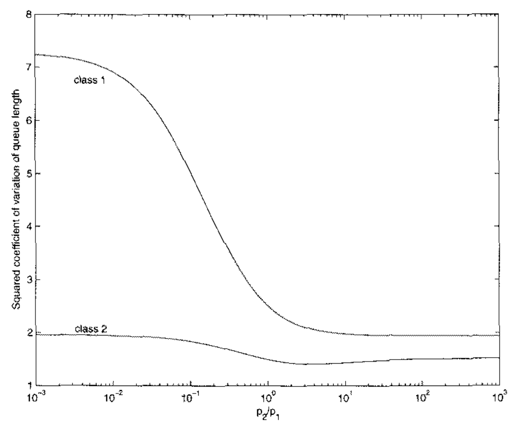

FIGURE 6. Squared coefficient of variation of queue length for Example 2.

Example 2. We assume that the service times of class 1 and class 2 customers have exponential distributions with densities $b(t) = e^{-t}$ and $b(t) = 4e^{-4t}$, respectively. Hence $\lambda_1 = 0.35$ and $\lambda_2 = 1.4$.

In Figs. 4-6, we plot the mean, variance and squared coefficient of variation for the queue length of each class, respectively, varying the priority parameter ratio p_2/p_1. As illustrated in Figs. 4 and 5, the mean and variance of the queue length of class 1 customer (resp. class 2 customer) increase (resp. decrease) as the priority parameter ratio p_2/p_1 increases, as we expect. Further, the mean and variance of the queue length of class 2 customer are larger than those of class 1 customer. Fig. 6 illustrates that the squared coefficient of variation of the queue length of class 2 customer is less affected by the priority parameter ratio p_2/p_1.

References

[1] E. Altman, A. Avrachenkov, and U. Ayesta, *A survey on discriminatory processor sharing*, Queueing Syst. **53** (2006), no. 1-2, 53–63.
[2] M. Haviv and J. van der Wal, *Equilibrium strategies for processor sharing and random queues with relative priorities*, Probab. Eng. Inform. Sci. **11** (1997), no. 4, 403–412.
[3] ———, *Waiting times in queues with relative priorities*, Oper. Res. Lett. **35** (2007), no. 5, 591–594.

Department of Mathematics Education
Chungbuk National University
Cheongju 361-763, Korea
E-mail address: jeongsimkim@chungbuk.ac.kr