Computers Math. Applic. Vol. 24, No. 1/2, pp. 7–15, 1992
Printed in Great Britain.
0097-4943/92 $5.00 + 0.00

EVALUATING MEAN SOJOURN TIME ESTIMATES FOR THE $M/M/1$ QUEUE

HIROSHI TOYOIZUMI
NTT Telecommunication Networks Laboratories
Musashino-shi, Midori-cho, 3–9–11, Japan

Abstract—Mean sojourn time is one of the most important performance measures for queueing systems. It is difficult to obtain the real sojourn time of a customer directly, so it is also difficult to estimate the mean sojourn time. In this paper, we propose a new and relatively simple estimate of the mean sojourn time in a single server queue, using the number of arrivals and the number of departures. This method can be used for evaluating the quality and the performance of call processing in communication switching systems, for example. We evaluate the accuracy of this estimate for an $M/M/1$ queue, using some results obtained by Jenkins. This estimate is compared with two other standard estimates of the mean sojourn time obtained from the sequence of actual arrival and departure times.

1. INTRODUCTION

Mean sojourn time (waiting time plus service time) in a queueing system is used to analyze the performance of the system and the quality of service to the customers. For example, communication switching systems are required to maintain a certain quality and performance in call processing. Engineers and administrators have to check the mean sojourn time of calls in the switching systems. However, we do not have a good method of estimating the mean sojourn time so far. In most cases, measuring the sojourn time of each customer is ineffective, since collecting actual time instants of arrival and departure is cumbersome work and we cannot expect good accuracy in the estimate because of the positive correlation between the sojourn times of customers.

Here, we present a new and relatively simple estimate of the mean sojourn time, which should be easily applied to real systems. The object of this paper is to evaluate the relative efficiency of our estimate and two standard methods for estimating the mean sojourn time.

First, we briefly look at our estimate. We consider observing, over a certain period of time, the number of arrivals and the number of departures only at regular intervals, obtaining a series of numbers of customers in the system and their mean. The estimates of the average number of customers and arrival rate lead to an estimate of the mean sojourn time by applying Little's formula. We call this the Customer Count Estimate (CCE) of the mean sojourn time.

We investigate two other estimates of the mean sojourn time for comparison with the CCE. One is the Direct Estimate (DE), which is to collect the time instants of arrival and departure of a customer and use them to obtain the sojourn time of the customer and the mean sojourn time. The other is the Test-Customer Estimate (TCE), which is to generate test-customers at regular intervals, and put them into the system to measure their sojourn times.

When we evaluate the efficiency of an estimate of some performance measure of a queueing system, we should deal with the correlation of queueing processes such as queue length, waiting time and so on. (See Reynolds [1] and Gafarian and Ancker [2].) There have been some studies on the estimates of the mean sojourn time and waiting time. Jenkins [3] evaluated the efficiency of the DE for an $M/M/1$ queue, using results of the cross-covariance of the arrival process and the queue length process. Blomqvist [4] studied the covariance of successive waiting times of an actual customer and evaluated the effect of the positive correlation on the DE in an $M/G/1$ queue. Also, the DE for a many-server queue was considered by Olsson [5]. On the other hand,

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TeX

TCE was studied by Matsuda [6] following the results on the covariance of the virtual waiting time process obtained by Ott [7]. To investigate the efficiency of the CCE, we extend Jenkins' results, which are for the continuous time measurement of the number of customers in the system, to discrete time measurement.

Throughout this paper, we consider the equilibrium state of an $M/M/1$ queue with arrival rate λ, service rate μ and traffic intensity $\rho = \lambda/\mu < 1$ under the first-come- first-served discipline.

2. THE CUSTOMER COUNT ESTIMATE (CCE) AND ITS UNBIASEDNESS

First, we state the CCE more precisely. To estimate the mean sojourn time W, which includes the service time, we have to observe $A(t)$, the number of arrivals during $(0,t]$, and $D(t)$, the number of departures during $(0,t]$. $A(t)$ and $D(t)$ are observed at the regular intervals $(t_i)_{i=1,2,\ldots,n}$, where Δt is the scanning interval and n is the number of observations during the measurement period. Hence, we obtain the sequences $(A(t_i))_{i=1,2,\ldots,n}$ and $(D(t_i))_{i=1,2,\ldots,n}$. Let $L(t)$ be the number of customers in the system at time t. Since $L(t) = A(t) - D(t)$, we can estimate the observed average number of customers in the system as

$$\bar{L} \equiv \frac{1}{n} \sum_{i=1}^{n} L(t_i) = \frac{1}{n} \left(\sum_{i=1}^{n} A(t_i) - \sum_{i=1}^{n} D(t_i) \right).$$

We can estimate the observed arrival rate as

$$\bar{\lambda} \equiv \frac{1}{t_n} A(t_n).$$

Finally, after applying the well-known Little's formula, we can deduce an estimate of the mean sojourn time in the system as

$$W_c \equiv \frac{\bar{L}}{\bar{\lambda}} = \frac{\sum_{i=1}^{n} (A(t_i) - D(t_i))\, \Delta t}{A(t_n)}. \tag{1}$$

REMARK 1. In order to apply the CCE to real systems, we have to prepare four different counters in the system. Two counters, $A(t)$ and $D(t)$, indicate the number of arrivals and the number of departures. The other two, $\sum A(t_i)$ and $\sum D(t_i)$, are the respective cumulations of $A(t)$ and $D(t)$. These two counters read $A(t)$ and $D(t)$ at the regular interval Δt, and add the values of $A(t)$ and $D(t)$ to themselves.

REMARK 2. Instead of W_c, we can use another estimate of W, such as

$$W_c' \equiv \frac{\sum_{i=1}^{n} (A(t_i) - D(t_i))\, \Delta t}{D(t_n)}.$$

In fact, due to the nature of Little's formula, it is more appropriate to use W_c' rather than W_c, since the denominator of estimate (1) is the throughput of the system. However, assuming that we observe the system over a long period, i.e., a sufficiently large t_n, $D(t_n)$ will be well approximated by $A(t_n)$. Hence, it is sufficient to consider W_c instead of W_c'.

Now, we proceed to show that our CCE is unbiased. An approximation is obtained by expanding $W_c = I/A$ into a two-dimensional Taylor Series about the point $(a_0, i_0) = (E(A), E(I))$, where $I \equiv \sum_{i=1}^{n} L(t_i)\, \Delta t$ and $A \equiv A(t)$. Taking the expectation and retaining only the terms up to second order, we have an approximation of $E(W_c)$:

$$E(W_c) \cong \frac{E(I)}{E(A)} + \frac{E(I)}{E^3(A)} \operatorname{Var}(A) - \frac{1}{E^2(A)} \operatorname{Cov}(I, A). \tag{2}$$

(See Kendall-Stuart [8].) Thus, we have to calculate Cov (I, A) to show the unbiasedness of our estimate.

LEMMA 1. *For the $M/M/1$ queue,*

$$\lim_{n\to 0} \frac{\text{Cov}\,(I,A)}{t_n} = \frac{\rho}{(1-\rho)^2}, \tag{3}$$

for any $\Delta t > 0$ in I.

PROOF. From the definitions of I and A, we have

$$\text{Cov}\,(I,A) = \sum_{i=1}^{n} \gamma_{AL}(t_i)\,\Delta t,$$

where $\gamma_{AL}(t)$ is the covariance function of $A(t)$ and $L(t)$, i.e.,

$$\begin{aligned}\gamma_{AL}(t) &\equiv \text{Cov}\,(A(t), L(t)) \\ &= E(\{A(t) - E(A(t))\}\,\{L(t) - E(L(t))\}).\end{aligned}$$

By the use of the Cauchy-Schwartz inequality, we obtain for any non-negative t

$$|\gamma_{AL}(t)| \leq \{\text{Var}\,(L(t))\,\text{Var}\,(A(t))\}^{1/2} = Ct^{1/2}, \tag{4}$$

where $C = (\lambda \text{Var}\,(L))^{1/2}$ is a positive constant that is independent of t.

For the $M/M/1$ queue, Jenkins [3] proved that

$$\lim_{t\to\infty} \gamma_{AL}(t) = \frac{\rho}{(1-\rho)^2}.$$

Hence, for each $\epsilon > 0$, there exists $T_0 = i_0 \Delta t > 0$ with $i_0 \in \mathrm{N}$ such that

$$\left|\gamma_{AL}(t) - \frac{\rho}{(1-\rho)^2}\right| < \epsilon \qquad \text{for all } t > T_0. \tag{5}$$

Applying T_0 instead of t to (4), we have the upper bound of $\gamma_{AL}(t)$ for $t \leq T_0$ as

$$|\gamma_{AL}(t)| \leq CT_0^{1/2}. \tag{6}$$

Using (5) and (6), we have

$$\begin{aligned}\left|\frac{\sum_{i=0}^{n} \gamma_{AL}(t_i)\,\Delta t}{t_n} - \frac{\rho}{(1-\rho)^{1/2}}\right| \\ &\leq \left|\frac{\sum_{i=0}^{i_0} \gamma_{AL}(t_i)}{t_n}\right| + \left|\frac{\sum_{i=i_0+1}^{n} \gamma_{AL}(t_i)}{t_n} - \frac{\rho}{(1-\rho)^2}\right| \\ &\leq \frac{i_0 \Delta t\, CT_0^{1/2}}{t_n} + \sum_{i=i_0+1}^{n} \frac{|\gamma_{AL}(t_i) - \rho/(1-\rho)^2|\,\Delta t}{t_n} + \frac{i_0 \rho}{n(1-\rho)^2} \\ &\leq \frac{CT_0^{2/3}}{t_n} + \frac{(n-i_0)\,\epsilon \Delta t}{t_n} + \frac{i_0 \rho}{n(1-\rho)^2}.\end{aligned}$$

The right-most of the above inequalities tends to ϵ as $n \to \infty$. Since ϵ can be taken arbitrarily, the proof is completed. ∎

PROPOSITION 1. *Let W_c be the CCE of W in the $M/M/1$ queue. Then, W_c is asymptotically unbiased, i.e.,*

$$\lim_{n\to\infty} E(W_c) = W.$$

PROOF. From the fact that $E(L(t_i)) = \rho/(1-\rho)$ for the $M/M/1$ queue and from Lemma 1, we have

$$\begin{aligned} E(A) &= \lambda t_n, \\ \text{Var}\,(A) &= \lambda t_n, \\ E(I) &= \sum_{i=1}^{n} E(L(t_i))\,\Delta t = \frac{\rho\, t_n}{(1-\rho)}, \end{aligned}$$

and

$$\lim_{n\to\infty} \frac{\text{Cov}\,(I,A)}{t_n} = \frac{\rho}{(1-\rho)^2}.$$

Substituting the above equalities into (2) gives

$$\lim_{n\to\infty} E(W_c) = \lim_{n\to\infty} \left\{ \frac{1/\mu}{1-\rho} + \frac{\rho}{\lambda^2 t_n(1-\rho)} - \frac{\text{Cov}\,(I,A)}{(\lambda t_n)^2} \right\} = \frac{1/\mu}{1-\rho} = W. \qquad \blacksquare$$

REMARK 3. In this proof, we use the covariance of the arrival and queue length processes, which was calculated through the Laplace transform of their joint distribution by Jenkins. However, it might be impossible to obtain the same results for an $M/G/1$ queue by the same method.

3. ACCURACY OF THE CUSTOMER COUNT ESTIMATE (CCE)

In much the same way as in (2), we can calculate the variance of W_c through the approximation

$$\text{Var}\,(W_c) \cong \left\{ \frac{E(I)}{E(A)} \right\}^2 \left\{ \frac{\text{Var}\,(I)}{E^2(I)} + \frac{\text{Var}\,(A)}{E^2(A)} - \frac{2\text{Cov}\,(I,A)}{E(I)\,E(A)} \right\}. \tag{7}$$

(See [8].) In order to estimate the variance, we begin with the following lemma.

LEMMA 2. *For the $M/M/1$ queue,*

$$\lim_{n\to\infty} \frac{\text{Var}\,(I)}{t_n} \cong \frac{\rho}{(1-\rho)^2} \frac{\Delta t(1+e^{-\alpha\Delta t})}{1-e^{-\alpha\Delta t}},$$

for any $\Delta t > 0$ in I, where $\alpha = \mu(1-\rho)^2/(1+\rho)$.

PROOF. The covariance function of $L(t)$ is approximated by an exponential function such as

$$\gamma_{LL}(t) \equiv \text{Cov}\,(L(s), L(s+t)) \cong \text{Var}\,(L)\, e^{-\alpha t}, \tag{8}$$

for some positive constants α and t, $s \geq 0$, where $L \equiv \lim_{t\to\infty} L(t)$. We have a number of ways to determine the constant α, but it will be best to match its integral value, since we have to deal with $\sum_{i=1}^{n} \gamma_{LL}(t_i)\,\Delta t$, which will be similar to $\int_0^{t_n} \gamma_{LL}(\tau)\,d\tau$. We will use the result obtained by Jenkins [3] to estimate the integral value of $\gamma_{LL}(t)$. Jenkins showed that

$$\int_0^{\infty} \gamma_{LL}(t)\,dt = \frac{\rho^2(1+\rho)^2}{\lambda(1-\rho)^4}$$

by using the Laplace transform of $\gamma_{LL}(t)$. Hence, we can take

$$\alpha = \frac{\mu(1-\rho)^2}{1+\rho},$$

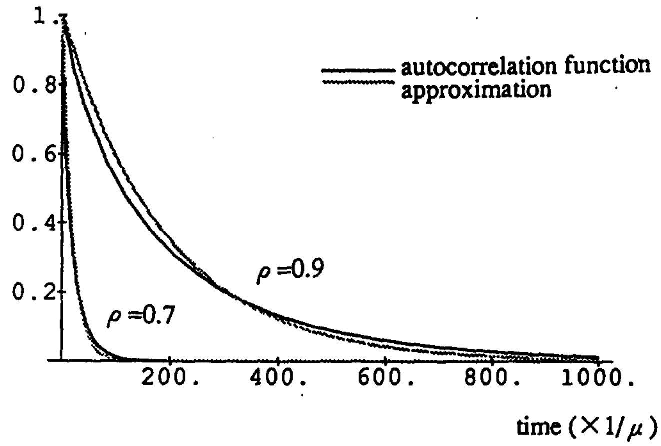

Figure 1. Autocorrelation function of queue length process and its approximation.

since Var $(L) = \rho/(1-\rho)^2$. We can see how our approximation matches the actual $\gamma_{LL}(t)$ in Figure 1.

On the other hand, easy computation gives

$$\text{Var }(I) = \left(\sum_{i=1}^{n} L(t_i)\,\Delta t\right) = (\Delta t)^2 \sum_{\substack{i=1\\j=1}}^{n} \gamma_{LL}(t_j - t_i)$$

$$= (\Delta t)^2 \left\{ n \left(\text{Var }(L) + 2\sum_{k=1}^{n-1} \gamma_{LL}(t_k) \right) - 2\sum_{k=1}^{n-1} k\gamma_{LL}(t_k) \right\}. \tag{9}$$

Using the approximation (8), the second term of (9) is evaluated as

$$2\sum_{k=1}^{n-1} \gamma_{LL}(t_k) \cong 2\text{Var }(L) \left\{ \sum_{k=0}^{n-1} (e^{-\alpha\Delta t})^k - 1 \right\}. \tag{10}$$

Let $n \to \infty$ in (10) to give

$$2\sum_{k=1}^{\infty} \gamma_{LL}(t) \cong \text{Var }(L) \frac{2e^{-\alpha\Delta t}}{1 - e^{-\alpha\Delta t}}. \tag{11}$$

Next, we deal with the third term of (9). Using the approximation (8), we have

$$-2(\Delta t)^2 \sum_{k=1}^{n-1} k\gamma_{LL}(t_k) \cong -2(\Delta t)^2 \,\text{Var }(L) \sum_{k=0}^{n-1} k(e^{-\alpha\Delta t})^k. \tag{12}$$

Similarly, let $n \to \infty$ in (12), to give

$$-2\sum_{k=1}^{\infty} k\gamma_{LL}(t) \cong \text{Var }(L)\, \frac{2(\Delta t)^2\, e^{-\alpha\Delta t}}{(1 - e^{-\alpha\Delta t})^2}. \tag{13}$$

In order to get an *a priori* estimate of the function on the right hand side of (13), we define

$$f(t) = \frac{t^2 e^{-\alpha t}}{(1 - e^{-\alpha t})^2}.$$

It is not difficult to show that $f(t)$ has the following properties:

$$f(t) \text{ decreases monotonically,}$$

$$f(0) = \frac{1}{\alpha^2}$$

and

$$\lim_{n\to\infty} f(t) = 0,$$

from which we obtain

$$0 \le f(t) \le \frac{1}{\alpha^2} \qquad \text{for all } t \ge 0.$$

Consequently, we can evaluate (13) as

$$\left| -2\sum_{k=1}^{\infty} k\gamma_{LL}(t) \right| \le \frac{2\text{Var }(L)}{\alpha^2} < \infty,$$

which means that the third term of (9) can be ignored.

Gathering up the results obtained here, we have

$$\lim_{n\to\infty} \frac{\text{Var }(I)}{t_n} \cong \frac{\Delta t\, \text{Var }(L)(1+2e^{-\alpha\Delta t})}{1-e^{-\alpha\Delta t}} = \text{Var }(L)\,\frac{\Delta t(1+e^{-\alpha\Delta t})}{1-e^{-\alpha\Delta t}}. \qquad \blacksquare$$

REMARK 4. Although the exact transient solution of $L(t)$ will be used for the evaluation of $\gamma_{LL}(t)$, the approximation we use here has great computational advantage.

We are now in a position to treat the asymptotic behaviour of the variance of the CCE, which is our main result.

PROPOSITION 2. *Let W_c be the CCE of W in the $M/M/1$ queue with the service rate μ and the traffic intensity ρ. Then, we have*

$$\lim_{n\to\infty} t_n \text{ Var }(W_c) \cong \frac{1}{\mu^2(1-\rho)^2}\left\{\frac{\Delta t(1+e^{-\alpha\Delta t})}{\rho(1-e^{-\alpha\Delta t})} - \frac{1+\rho}{\lambda(1-\rho)}\right\},$$

where $\alpha = \mu(1-\rho)^2/(1+\rho)$ and $\Delta t > 0$.

PROOF. Letting n tend to infinity in (7) gives

$$\begin{aligned}\lim_{n\to\infty} t_n \text{Var }(W_c) &\cong \lim_{n\to\infty} \left(\frac{E(L)}{\lambda}\right)^2 \left\{\left(\frac{1-\rho}{\rho}\right)^2 \frac{\text{Var }(I)}{t_n} + \frac{1}{\lambda} - \frac{2(1-\rho)}{\lambda\rho}\frac{\text{Cov }(I,A)}{t_n}\right\} \\ &= \frac{1}{\mu^2(1-\rho)^2}\left\{\frac{\Delta t(1+e^{-\alpha\Delta t})}{\rho(1-e^{-\alpha\Delta t})} - \frac{1+\rho}{\lambda(1-\rho)}\right\} \qquad \blacksquare\end{aligned}$$

4. THE TEST-CUSTOMER ESTIMATE (TCE) AND THE DIRECT ESTIMATE (DE)

We can easily show that the TCE and DE are unbiased estimates, so we proceed to evaluate the variance of the estimates.

First, we study the TCE. Since Ott [7] estimated the covariance function of the virtual waiting time process in the $M/G/1$ queue, we use his result to obtain the accuracy of the TCE. For the TCE, the sum of the virtual waiting time and the service time is measured by the test-customers, which are generated at regular interval Δt. Here, the test-customers are assumed to have no effect on other customers. This assumption might be valid when the intervals between test-customer are considerably long. If the intervals are short, the TCE no longer gives an unbiased estimate of the mean waiting time in the system, since we cannot ignore the effect of the test-customers, which increases the waiting time of real customers.

Let $V(t)$ be the virtual waiting time at time t and S_i be the service time of the i^{th} test-customer. Then, the estimate of the mean sojourn time obtained by the test-customer is

$$W_t = \frac{1}{n}\sum_{i=1}^{n}(V(t_i) + S_i).$$

Proposition 3. *For the $M/M/1$ queue, we have*

$$\lim_{n\to\infty} t_n \, Var\,(W_t) \cong \frac{\rho^3(2-\rho)}{\lambda^2(1-\rho)^2}\,\frac{\Delta t(1+e^{-\beta\Delta t})}{\rho(1-e^{-\beta\Delta t})} + \frac{\Delta t}{\mu^2},$$

where $\beta = \lambda(2-\rho)(1-\rho)^2/\rho(3-\rho)$ and $\Delta t > 0$.

Proof. We use the approximation used by Matsuda [6], who showed that the covariance function of the virtual waiting time is well approximated as

$$\begin{aligned}\gamma_{VV}(t) &\equiv \text{Cov}\,(V(t), V(0)) \\ &\cong \text{Var}\,(V)\, e^{-\beta t},\end{aligned} \tag{14}$$

where $V \equiv \lim_{t\to\infty} V(t)$. The constant β is determined by Ott's result about the covariance function of the virtual waiting time. Applying Ott's result to the $M/M/1$ case, we have

$$\int_0^\infty \gamma_{VV}(t)\,dt = \frac{\rho^4(3-\rho)^2}{\lambda^3(1-\rho)^4}.$$

Hence, we take

$$\beta = \frac{\lambda(2-\rho)(1-\rho)^2}{\rho(3-\rho)},$$

by matching the integral value of $\gamma_{VV}(t)$ as in the proof of Lemma 2. In much the same way as in Section 3, we have

$$\begin{aligned}\lim_{n\to\infty} t_n \text{Var}\,(W_t) &= \lim_{n\to\infty}\left[(\Delta t)^2\left\{n\left(\text{Var}\,(V) + 2\sum_{k=1}^{n-1}\gamma_{VV}(t_k)\right) - 2\sum_{k=1}^{n-1} k\gamma_{VV}(t_k)\right\}\right. \\ &\qquad \left. + \frac{t_n \text{Var}\,(S_i)}{n}\right] \\ &\cong \text{Var}\,(V)\frac{\Delta t(1+e^{-\beta\Delta t})}{1-e^{-\beta\Delta t}} + \frac{\Delta t}{\mu^2}\end{aligned}$$

∎

The DE is discussed and analysed by Jenkins [3] in relation to the Maximum Likelihood Estimate. Here we state his result without the proof.

Proposition 4. *Let W_d be the DE of W in the $M/M/1$ queue, i.e.,*

$$W_d = \frac{1}{A(t)}\sum_{i=1}^{A(t)} W_i,$$

where W_i is the sojourn time of the i^{th} real customer. Then, we have the asymptotic variance of W_d as

$$\lim_{t\to\infty} t\,\text{Var}\,(W_d) = \frac{\rho^2(1+\rho)^2}{\lambda^3(1-\rho)^4}.$$

We can easily find the relation between the CCE and DE by letting $\Delta t \to 0$ in Proposition 2.

Corollary.

$$\lim_{\Delta t\to 0}\left(\lim_{n\to\infty} t_n \,\text{Var}\,(W_c)\right) = \lim_{t\to\infty} t\,\text{Var}\,(W_d).$$

5. RELATIVE EFFICIENCY OF THE ESTIMATES BY NUMERICAL RESULTS

In this section, we study the relative efficiency of three estimates, the CCE, TCE and DE. We have to consider:

(1) how easily we can measure the original data (the number of arrivals, the sojourn time of customers and so on),
(2) how easily we can process the data, and
(3) how much accuracy we can expect.

Now, we proceed to a more detailed look at (1), (2) and (3) for these estimates.

(1) Measurement
For large ρ, we have to deal with an immense volume of data for the DE, whereas the volume of data is reasonable for the CCE and TCE with the scanning interval Δt selected so as to produce the same accuracy as the DE. (See (3) below.)

(2) Data processing
We have to prepare clocks and time recorders in systems for the TCE and DE. They are more cumbersome than the counters of the CCE.

(3) Accuracy
We already know that these three estimates are unbiased, so we should compare their coefficients of variance, calculated from the results obtained in Sections 3 and 4. We set the asymptotic coefficient of variance as

$$AC(W) = \lim_{t\to\infty} \frac{\{\mathrm{Var}\ (W)\,t\}^{1/2}}{E(W)}.$$

Hence, the coefficient of variance is well approximated by $AC(W)t^{-1/2}$ for sufficiently large t. If we have an equality $AC(W_c) = \alpha AC(W_d)$, W_c should be observed for a period α^2 times as long as the measurement period of W_d. Thus, we can consider that $AC(W_c)^2/AC(W_d)^2$ expresses the length of the measurement period, i.e., the efficiency of the estimate W_c. In Figure 2, we compare $AC(W_c)^2/AC(W_d)^2$ with $AC(W_t)^2/AC(W_d)^2$. Figure 2 shows that there are no significant differences between these estimates for large ρ. It also shows that we do not have to make the scanning interval Δt so small, since the positive correlation of the process decreases the accuracy of the estimates.

On the basis of (1), (2) and (3) above, we can conclude that the CCE is a convenient method for estimating the sojourn time.

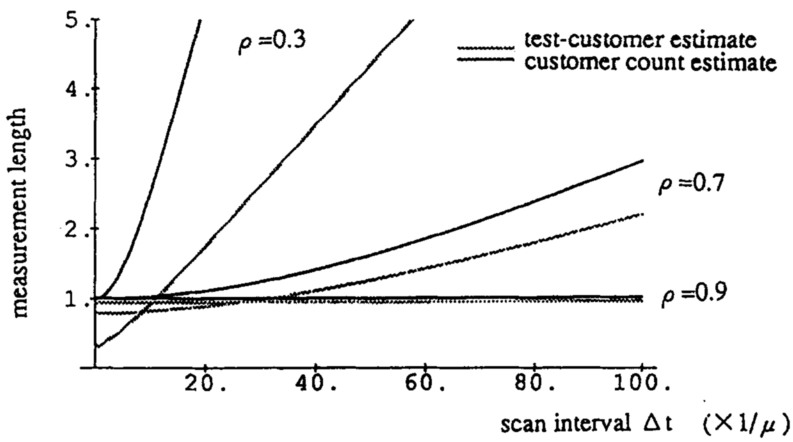

Figure 2. Relative efficiency of the estimates (measurement length having the same accuracy as the direct estimate).

REFERENCES

1. J.F. Reynolds, The covariance structure of queues and related process—A survey of recent work, *Adv. Appl. Prob.* **7**, 383–415 (1975).
2. A.V. Gafarian and C.J. Ancker, Mean value estimation from digital computer simulation, *Operations Res.* **14**, 25–44 (1966).

3. J.H. Jenkins, The relative efficiency of direct and maximum likelihood estimates of the mean waiting time in the simple $M/M/1$ queue, *J. Appl. Prob.* **9**, 396–403 (1972).
4. N. Blomqvist, The covariance function of the $M/G/1$ queueing system, *Skand. Aktuarie. Tidskr.* **50**, 1036–1040 (1967).
5. K.M. Olsson, Accuracy of observed mean values of waiting times, queue lengths and proportion of delayed calls in Markovian many-server systems, In *Proc. 8th ITC*, 254/1–7 (1976).
6. J. Matsuda, An approximation for the auto-covariance function of the virtual waiting time in an $M/G/1$ and its applications (in Japanese), *IECE Japan* **J67-B**, 121–126 (1984).
7. T.J. Ott, The covariance function of the virtual waiting time process in an $M/G/1$ queue, *Adv. Appl. Prob.* **9**, 158–168 (1977).
8. M. Kendall and A. Stuart, *The Advanced Theory of Statistics* **1**, 4th edn., p. 246, Charles Griffin, London, (1977).