韓國 軍事運營分析 學會誌
第12卷 第2號, 1986.12.

Jacknife and Bootstrap Estimation of the Mean Number of Customers in Service for an M/G/∞

Park, Dong Keun*

Abstract

This thesis studies the estimation from interarrival and service time data of the mean number of customers in service at time t for an M/G/∞ queue. The assumption is that the parametric form of the service time distribution is unknown and the empirical distribution of twe service time is used in the estimate the mean number of customers in service. In the case in which the customer arrival rate is known the distribution of the estimate is derived and an approximate normal confidence interval procedure is suggested. The use of the nonparametric methods, which are the jackknife and the bootstrap, to estimate variability and construct confidence intervals for the estimate is also studied both analytically and by simulation.

I. Instruction

The concern of this thesis is inference problems for a particularly simple queueing model, the M/G/∞ queue. In this model, customers arrive according to a poisson process with rate λ and there are an unlimited number of independent servers. Service times for each server are independent, identically distributed with distribution function F. Let X(t) be the number of customers being arrived at time t. It is well known that if there are no customers being served at time 0,

$$P\{X(t)=n\}=\frac{[M(t)]^K}{K!}\exp[-M(t)] \tag{1-1}$$

then where $M(t)=\lambda\int_0^t \bar{F}(s)\,ds$ with $\bar{F}(t)=1-F(t)$ [Ref. 2]. Thus the distribution of X(t) is poisson and is characterized by its mean M(t).

In this thesis we will assume that the service time distribution F(t) is unknown and must be estimated from service time data and that the arrival process is known to be poisson, except possibly for its rate λ. We will study the estimation of the mean number of customers being served at time t, M(t).

* 해군본부

We generally divide the estimation method into two cases which we shall call "parametric estimation" and "nonparametric estimation". In the parametric estimation case, a particular probabilistic model is specified for the service time distribution and the parameters of the distribution are estimated. In the nonparametric estimation method, the empirical survivor function is used in the estimate of the expected number of customers. The jackknife procedure and bootstrap procedure are considered and compared by simulation.

In most cases, parametric assumptions concerning the service time distribution are difficult to justify. Hence nonparametric estimation procedure may well be preferred to parametric estimation when actual data is used. However, the nonparametric estimates can be expected to be less efficient than the parametric ones.

II. Jackknife estimation method

In this chapter, we will study the jackknife procedure for obtaining a confidence interval for M(t). The jackknife was first introduced by Quenouille (1949) for the purpose of reducing the estimate bias, and the procedure was later utilized by Tukey (1958), to develop a general method for obtaining approximate confidence intervals [Refs. 7,8].

The basic idea of the jackknife estimation method is to assess the effect of each of the groups into which the data have been devided, not by the results for that group alone, but rather through the effect upon the body of data that results from omitting that group. The two bases of the jackknife are that we make the desired calculation for all the data, and then, after dividing the data into groups, we make the calculations for each of the slightly reduced bodies of data obtained by leaving out just one of the groups. A special case of jackknife estimation is called the "complete jackknife estimation", where the number of subgroups is n (the size of sample); the i^{th} subgroup is obtained by deleting the i^{th} observation; thus the size of each subgroup is n-1 [Ref. 9]. Attention will be restricted to complete jackknife estimation in this study.

Let $\hat{M}^{i}_{n-1}(t)$ be the estimated mean number of customers being served at time t on the portion of the sample that omits the i^{th} sample. Let $\hat{M}_{all}(t)$ be the corresponding estimator for the entire sample and define the i^{th} pseudo-value by

$$\hat{M}_{i}(t) = n\,\hat{M}_{all}(t) - (n-1)\,\hat{M}^{i}_{n-1}(t) \tag{2-1}$$

The jackknife estimate $\hat{M}_J(t)$ and an estimate $\hat{S}^2_J$ of its variance are given by

$$\hat{M}_J(t) = \frac{1}{n}\sum_{i=1}^{n}\hat{M}_i(t) \tag{2-2}$$

$$\hat{S}^2_J = \frac{1}{n\,n-1}\sum_{i=1}^{n}[\hat{M}_i(t) - \hat{M}_J(t)]^2 \tag{2-3}$$

Tukey (1958) proposes that the n estimated pseudo values be treated as approximately independent and identically distriboted random variables [Ref. 9]. Hence, the statistic

$$\frac{\sqrt{n}\,(\hat{M}_J(t)-\hat{M}_{all}(t))}{[\frac{1}{n-1}\sum_{i=1}^{n}(\hat{M}_i(t)-\hat{M}_J(t))^2]^{1/2}} \qquad (2\text{-}4)$$

has an approximate t-distribution with n-1 degrees of freedom, which leads to the approximate 100(1-α)% confidence interval

$$\hat{M}_J(t) \pm t_{1-\frac{\alpha}{2}}\hat{S}_J \qquad (2\text{-}5)$$

where $t_{1-\frac{\alpha}{2}}$ is the upper $1-\frac{\alpha}{2}$ critical point of the t-distribution with n-1 degrees of freedom. The confidence interval given by equation 2.5 is a function of the estimated variance. In the remainder of this section, we will describe several methods of implementing the confidence interval procedure. We will also obtain an analytic expression for the jackknife estimate and its variance estimate for the case in which the arrival rate λ is known.

1. Jackknife Estimate with Known Arrival Rate

The jackknife estimate is based on sequentially deleting point S and recomputing the estimator. Removing point S_i from data set gives a different empirical probability distribution $\hat{F}_{n-1}$ with mass $\frac{1}{n-1}$ at $S_1, S_2, \ldots, S_{i-1}, S_{i+1}, \ldots, S_n$ and a corresponding recomputed value of the estimate. In the jackknife process, the i^{th} pseudo value is

$$\hat{M}_i(t) = \begin{cases} \lambda t & \text{if} \quad i > k \\ \lambda s & \text{if} \quad i < k \end{cases} \qquad (2\text{-}6)$$

for the fixed time t and where λ is the constant value. Accordingly the pseudo values $\hat{M}_i(t)$ have just K+1 different values. The jackknife estimate is

$$\hat{M}_J(t) = \lambda[\frac{1}{n}\sum_{i=1}^{k} S_i - \frac{n-k}{n}t] \qquad (2\text{-}7)$$

This result is exactly the same as the original estimate. This is because the estimate $\hat{M}_J(t)$ is unbiased. We will now describe two selected procedures to obtain confidence intervals for the jackknife estimate. Tukey suggested that the statistic in equation 2.4 has an approximate t-distribution with n-1 degrees of freedom, which leads to the approximate two-sided 100(1-α)% confidence interval

$$\hat{M}_J(t) \pm t_{1-\frac{\alpha}{2}}\sqrt{\mathrm{Var}[M_J(t)]} \qquad (2\text{-}8)$$

for $M_N(t)$, where $t_{1-\frac{\alpha}{2}}$ is the upper $1-\frac{\alpha}{2}$ critical point of the t-distribution with n-1 degrees of freedom. However, the n estimate pseudo values have just K+1 different values. Hence, another possible procedure is to adjust the degrees of freedom of the t-distribution, that is, subtract one from the number of different pseudo values (K+1), and use the result as the degrees of freedom. The length of confidence interval generated using the adjusted degrees of freedom (K) is slightly wider than that generated using the usual degrees of freedom (n-1) and the coverage rate should be increased.

2. Jackknife Estimate with Unknown Arrival Rate

In this subsection, it will be assumed that the rate of Poisson arrival process is unknown and must also be estimated. The maximum likelihood estimate is $\hat{\lambda} = n/\sum_{i=1}^{n} Y_i$, where Y_i is the interarrival time between i^{th} and $(i-1)^{th}$ customers. A nonparametric estimate of mean number of customers being served at time t is given by

$$\hat{M}_N(t) = \hat{\lambda}\left\{\frac{1}{n}\sum_{i=1}^{K} S_i + \frac{n-k}{n}t\right\} \qquad (2\text{-}9)$$

where S_i's are the order statistics. It is assumed that the S_i's and Y_i's are independent. The variable K is the number of S_i's which are less than t. The data consist of two independent random samples,

$$S_1, S_2, \ldots\ldots, S_n \sim F \text{ and } Y_1, Y_2, \ldots\ldots, Y_n \sim Q$$

F and Q being two possibly different distribution on the real line with Q, the exponential distribution with mean $\frac{1}{\lambda}$. From equation 2.9, the estimate $\hat{M}_N(t)$ is the product of two estimates. One is the function of Y_i, $\hat{\lambda} = n / \sum_{i=1}^{n} Y$, and the other is the function of S_i, $H(s) = \frac{1}{n}\sum_{i=1}^{k} S_i + \frac{n-k}{n}t$. There are many possible ways to perform a two-sample jackknife procedure. We will call one method the "paired sample jackknife" procedure. Since the size of both samples is the same, we make the one set of observations by pairing respective observations, that is, $(s_1, y_1), (s_2, y_2), \ldots, (s_n, y_n)$. As with the one-sample jackknife, we estimate the $\hat{M}(t)$ for all the data, and then, we estimate $\hat{M}_{n-1}^{i}(t)$ based on the remaining data obtained by leaving out just the i^{th} pair. Thus the i^{th} pseudo value $\hat{M}_i(t)$ is

$$\hat{M}_i(t) = n\hat{M}_{all}(t) - (n-1)\hat{M}_{n-1}^{i}(t) \qquad (2\text{-}10)$$

and the jackknife estimate $\hat{M}_J(t)$ and variance estimate are given by

$$\hat{M}_J(t) = \frac{1}{n}\sum_{i=1}^{n} \hat{M}_i(t)$$

$$\hat{S}_J^2 = \frac{1}{n(n-1)}\sum_{i=1}^{n} [\hat{M}_i(t) - \hat{M}_J(t)]^2$$

Based on these statistics, an approximate two-sided 100(1-α)% confidence interval is given by

$$\hat{M}_J(t) \pm t_{1-\frac{\alpha}{2}} \hat{S}_J \qquad (2\text{-}11)$$

where $t_{1-\frac{\alpha}{2}}$ is the upper $1-\frac{\alpha}{2}$ point of t-distribution with n-1 degrees of freedom. A second method is called the "separated sample jackknife" procedure. Since we assumed that the X_i's and Y_i's are independent, we can perform the jackknife procedure separately for each sample, and then, estimates which combine jackknife estimates and the jackknife variance estimate can be computed.

Let $\hat{M}_{J1}(t)$ be the jackknife estimate of $\hat{\lambda}$ and $\hat{V}_Y$ be the jackknife variance estimate for $\hat{\lambda}$. Let $\hat{M}_{J2}(t)$ be the jackknife estimate of $\int_0^t F(s)\, ds$ and $\hat{V}_s$ be its jackknife variance estimate. Then the combined jackknife estimate of $\hat{M}_{JC}(t)$ is

$$\hat{M}_{JC}(t) = \hat{M}_{J1}(t)\, \hat{M}_{J2}(t) \qquad (2\text{-}12)$$

and the combined jackknife variance estimate is

$$\hat{S}^2_{JC} = \hat{V}_Y \cdot \hat{V}_S + \hat{V}_Y[\hat{M}_{J2}(t)]^2 + \hat{V}_S[\hat{M}_{J1}(t)]^2 \qquad (2\text{-}13)$$

The approximate two-sided 100(1-α)% confidence interval is given by

$$\hat{M}_{JC}(t) \pm t_{1-\frac{\alpha}{2}} \hat{S}_{JC}/\sqrt{n} \qquad (2\text{-}14)$$

where $t_{1-\frac{\alpha}{2}}$ is the upper $1-\frac{\alpha}{2}$ point of t-distribution with n-1 degree of freedom.

III. Bootstrap Estimation Method

Efron (1979) introduced the bootstrap method for estimating the distribution of a statistic computed from observations [Ref. 10]. The bootstrap estimate is obtained by replacing the unknown distribution by the empirical distribution of the data in the definition of the statistical function. In practice, the distribution of the statistic is approximated by Monte Carlo methods.

For convenience, the arrival rate is assumed to be known and equal to 1, then the nonparametric estimate $\hat{M}_N(t)$ is just a function of service times. This is a one-sample problem. The bootstrap procedure is as follows:

1) Suppose that the data points $x_1, x_2, ..., x_n$ are independent observations from the unknown distribution F. Then the true estimate is

$$M_N(t) = \int_0^t \bar{F}(s)\, ds \qquad (3\text{-}1)$$

2) We can estimate the distribution F by the empirical probability distribution F_n.

F_n : mass $\frac{1}{n}$ on each observed data point x_i

$i = 1, 2, ..., n$

3) The bootstrap estimate of $M_N(t)$ is

$$\hat{M}_B(t) = \int_0^t \bar{F}_n(s)\ ds \tag{3-2}$$

To obtain an estimate of variability for $\hat{M}_B(t)$, we procede as follows;

(a) Construct F_n, the empirical distribution function, as just described

(b) Draw a bootstrap sample x_1^*, x_2^*, ..., x_n^* by independent random sampling from F_n.

Notice that we are not getting a permutation distribution since the values of X_i^* are selected with replacement from the set $(x_1, x_2, ..., x_n)$. As a point comparison, the ordinary jackknife can be thought of as drawing samples of size n-1 without replacement.

(c) Compute an estimate of $M_N(t)$ for each bootstrap replication, $M_N^*(t)$, that is, the value of statistic evaluated for the bootstrap sample.

$$M_n^*(t) = \frac{1}{n}\sum_{X_i^* \le t} X_i^* + \frac{1}{n}[n - \sum_{i=1}^{n} I(X_i^* \le t)]t \tag{3-3}$$

where $I(x \leq t) = \begin{cases} 1 & \text{if } x \leq t \\ 0 & \text{otherwise} \end{cases}$

(d) Do step (b) some large number "B" times obtaining independent bootstrap replications $M_N^*(t)$, $M_N^*(t)$, ..., $M_N^{*B}(t)$.

Based on the bootstrap replications, the approximate estimate of $M_N(t)$ and its variance are obtained by

$$\hat{M}_B(t) = \frac{1}{B}\sum_{i=1}^{B} M_N^{*i}(t) \tag{3-4}$$

$$\hat{V}ar[\hat{M}_B(t)] = \frac{1}{B-1}\sum_{i=1}^{B}[M_N^{*i}(t) - \hat{M}_B(t)]^2 \tag{3-5}$$

So far we have considered the problem, where the arrival rate is known. The bootstrap methodology also applies if the arrival rate is unknown and is estimated from interarrival data. Suppose the data consist of a random sample $X = (X_1, X_2, ..., X_n)$ from unknown service time distribution F and an independent sample $Y = (Y_1, Y_2, ..., Y_n)$ from the exponential interarrival time distribution G with unknown parameter λ. One bootstrap procedure to estimate the expected number of customers being served at time t is to construct F_n and G_n, the empirical probability distribution corresponding to F and G. Bootstrap samples $X_i^* \sim F_n$, $i=1,2, ..., n$, $Y_j^* \sim G_n$, $j=1,2, ..., n$, are independently drawn, an estimate of $M_N(t)$

$$M_N^*(t) = \frac{n}{\sum_{i=1}^{n} Y_i^*} \left\{ \frac{1}{n} \sum_{X_i^* \leq t} X_i^* + \frac{1}{n} \left[n - \sum_{i=1}^{n} I(X_i^* \leq t) \right] t \right\} \tag{3-6}$$

is calculated. As before there are a large number B of bootstrap replications. For this case, the bootstrap estimate of $M_N(t)$ and its variance are still given by equations 3.4 and 3.5. There appears to be no closed form of the analytical variance of $\hat{M}_B(t)$ in this case. Now we will describe methods to obtain approximate confidence intervals for the bootstrap estimate $\hat{M}_B(t)$.

1. The Percentile Method

A simple method for assigning approximate confidence intervals to the nonparametric estimate $M_N(t)$ is as follows:
Let

$$\hat{C}(t) = \frac{\#\text{ of } \{M_N^*(t) \leq t\}}{B} \tag{3-7}$$

be the cumulative distribution function of the bootstrap distribution of $M_N(t)$; B is the number of bootstrap replications. For a given $0 < \alpha < 0.5$, define

$$\hat{L}(\alpha) = \hat{C}^{-1}(\alpha),\ \hat{U}(\alpha) = \hat{C}^{-1}(1\text{-}\alpha) \tag{3-8}$$

Usually denoted simply by $\hat{L}$ and $\hat{U}$. This definition runs into complications when we actually try to compute quantiles $\hat{L}$ and $\hat{U}$ from a set of bootstrap replications. To overcome these difficulties, we order the bootstrap replications from smallest to largest, obtaining the sorted data M_N^{*i} (t), for i=1 to B. Letting represent any fraction between 0 and 1; take $Q(\alpha)$ to be M_N^{*i} (t) whenever Q is one of the functions $\alpha = \frac{i-0.5}{B}$, for i=1 to B. Thus $\hat{L}(\alpha)$ turns out to be the $(B * \alpha + 0.5)^{th}$. M_N^{*i} (t) and $\hat{U}(\alpha)$ to be the $(B * (1\text{-}\alpha) + 0.5)^{th}$ $M_N^{*i}(t)$. The percentile method consists of taking

$$[\hat{L}(\alpha),\ \hat{U}(\alpha)] \tag{3-9}$$

as an approximate $1\text{-}2\alpha$ confidence interval for $\hat{M}_B$ (t) since $\alpha = C(\hat{L})$, $1\text{-}\alpha = C(\hat{U})$, the percentile method interval consists of the central $1\text{-}2\alpha$ proportion of the bootstrap distribution.

2. The Bias-corrected Percentile Method

Efron (1980) suggests the following bias correction for the percentile confidence interval procedure [Ref. 11]. He argues that is $\hat{M}_B(t)$ is not the median of the bootstrap replication distribution, then a bias correction to the percentile method is called for. To be specific, define

$$Z_0 = \Phi^{-1} \left\{ \hat{C}(\hat{M}_B(t)) \right\} \tag{3-10}$$

where $\hat{C}(t) = \frac{\# \text{ of } \{M_V^*(t) \leq t\}}{B}$ as in equation 3.7, and Φ is the cumulative distribution function function for a standard normal variate. The bias corrected percentile method consists of taking

$$\left[\hat{C}^{-1}\{\Phi(2Z_0 - Z_\alpha)\},\ \hat{C}^{-1}\{\Phi(2Z_0 + Z_\alpha)\} \right] \tag{3-11}$$

as an approximate $1\text{-}2\alpha$ central confidence interval for $\hat{M}_B(t)$. Here z_α is the upper point for a standard normal $\Phi(z_\alpha) = 1\text{-}\alpha$.

IV. Simulation Results

The purpose of the simulation in this chapter is toassess the performance of the nonparametric estimation methods, the jackknife and the bootstrap. Since the estimate of M(t), the mean number of customers being served at time t, is a function of the customer arrival rate and the integral of the servivor function of the service time distribution, two simulation cases are done. The first simulation case was performed to estimate $M_N(t)$, the nonparametric estimate of M(t), as a function of the service times with the arrival rate assumed to be known and set equal to 1. The second case assumed that the customer arrival rate is also unknown and must be estimated using interarrival times.

In each replication of the simulation for case 1, 50 independent service times from a specified service time distribution were generated. For the bootstrap procedure, 500 bootstrap replications were performed. The simulation was replicated 300 times. For the purpose of comparison, we considered four types of service time distributions, which were the exponential, the mixed exponential, the gamma, the lognormal distribution. The arrival process is known to be poisson process with known rate $\lambda = 1$. The same generated service times were used for each estimation procedure in a replication. This reduces the variability of the differences in performance between the procedures.

To illustrate the efficiency of the nonparametric estimation methods, we simulated two possible ways to construct the approximate confidence interval for $M_N(t)$ for the bootstrap and the jackknife methods. For the jackknife estimation method, one procedure was to construct the confidence interval with the regular degrees of freedom, n-1, and the other used the reduced degrees of freedom, which is the number of different pseudo values. For the bootstrap estimation method, one way used the percentile method by the Monte Carlo process, and the other used the bias-corrected percentile method; there were 500 bootstrap replications. Nominal 68%, 80% and 90% confidence intervals were constructed for each replication using each method. It was noted whether the confidence interval formed by a given method covered the true value M(t). The entire process was independently replicated with R = 300 times. From these R replications we computed, for each method, the proportion $\hat{p}$ of the R confidence intervals which contained M(t), as well as the average length of the confidence intervals. If a method was performing adequately, $\hat{p}$ should be near $1\text{-}\alpha$, and a small mean length is desirable.

Table I. Coverage and Length of 100(1-α)% C.I for Exponential with μ=2, λ = 1 at T = 1

		68%		80%		90%	
		Length (s.d)	C.R.	Length (s.d)	C.R.	Length (s.d)	C.R.
Normal C.I. Procedure		0.0888 (0.0089)	19.33 69.33 11.33	0.1147 (0.0101)	12.00 83.67 4.33	0.1477 (0.0141)	4.67 92.00 3.30
Jackknife	Reduced d. f	0.0911 (0.0089)	17.67 71.33 11.00	0.1187 (0.0102)	11.00 85.33 3.67	0.1548 (0.0142)	3.33 94.00 2.67
	Regular d. f	0.0897 (0.0090)	18.38 70.67 11.00	0.1163 (0.0103)	11.33 84.33 4.33	0.1505 (0.0144)	4.00 93.00 3.00
Bootstrap	Percentile method	0.0884 (0.0098)	18.67 69.00 12.33	0.1132 (0.0112)	12.00 82.00 6.00	0.1462 (0.0154)	4.33 92.00 4.67
	Biascorrect method	0.0887 (0.0098)	16.67 71.00 12.33	0.1137 (0.0112)	10.67 83.00 6.33	0.1467 (0.0154)	2.67 92.67 4.67

In order to compare the performance of these procedures to the normal confidence interval procedure, simulations were conducted, and nominal 68%, 80% and 90% confidence limits were constructed for time t = 1 for each replication. By the central limit thorem, the distribution of $M_N(t)$ is asymtotically normal distributed as the number of data points $n \rightarrow \infty$ Thus, the 100(1-α)% normal confidence interval is given by

$$M_N(t) \pm Z_{1-\frac{\alpha}{2}} \sqrt{Var[M_N(t)]} \qquad (4\text{-}1)$$

where $Z_{1-\frac{\alpha}{2}}$ is the upper $1-\frac{\alpha}{2}$ point of the standard normal distribution.

Each cell in the tables contain the average and standard deviation of confidence interval length; and the proportion of intervals that are too high, (e.g. $M(t) < L$), where L is the lower bound of interval; the proportion of intervals covering the true value M(t), $\hat{p}$; the proportion of interval are too low, (e.g. $M(t) < U$), where U is the upper bound of interval.

The overall examination of the tabulations of confidence limit coverage and also the average and standard deviation of confidence interval length suggest that the bootstrap procedure is slightly better than the jackknife procedure; however, the difference is negligible.

Table II. Coverage and Length of 100(1-α)% C.I. for Mixed Exponential with μ_1 = 2, μ_2 = .75, P_1 = .2, λ = 1 at T = 1.

		68%		80%		90%	
		Length (s.d)	C.R.	Length (s.d)	C.R.	Length (s.d)	C.R.
Normal C.I. Procedure		0.0914 (0.0084)	14.67 71.67 13.67	0.1169 (0.0102)	10.00 83.00 7.00	0.1509 (0.0143)	5.00 90.67 4.33
Jackknife	Reduced d. f	0.0936 (0.0085)	14.00 73.00 13.00	0.1208 (0.0103)	9.67 83.33 7.00	0.1581 (0.0143)	4.33 91.67 4.00
	Regular d. f	0.0923 (0.0085)	14.67 72.00 13.33	0.1185 (0.0103)	10.00 83.00 7.00	0.1538 (0.0145)	4.33 91.33 4.33
Bootstrap	Percentile method	0.0911 (0.0094)	14.33 71.67 14.00	0.1156 (0.0112)	11.33 81.33 7.33	0.1490 (0.0153)	4.33 89.67 4.33
	Biascorrect method	0.0913 (0.0094)	14.00 71.00 15.00	0.1161 (0.0112)	9.00 83.00 8.00	0.1495 (0.0153)	4.00 89.67 6.33

Table III. Coverage and Length of 100(1-α)% C.I. for Gamma with β=1, K=2, λ=1 at T = 1

		68%		80%		90%	
		Length (s.d)	C.R.	Length (s.d)	C.R.	Length (s.d)	C.R.
Normal C.I. Procedure		0.0595 (0.0100)	22.00 68.33 9.67	0.0774 (0.0120)	12.33 80.33 7.33	0.0989 (0.0174)	10.67 86.33 3.00
Jackkneife	Reduced d. f	0.0617 (0.0102)	21.00 70.00 9.00	0.0813 (0.0122)	11.33 82.67 6.00	0.1062 (0.0178)	8.33 89.00 2.67
	Regular d. f	0.0601 (0.0102)	21.67 69.33 9.00	0.0785 (0.0121)	12.33 80.67 7.00	0.1008 (0.0177)	10.33 86.67 4.00
Bootstrap	Percentile method	0.0589 (0.0091)	21.33 69.97 9.00	0.0766 (0.0122)	12.00 80.33 7.67	0.0981 (0.0179)	9.33 86.67 3.00
	Biascorrect method	0.0595 (0.0100)	19.33 69.33 11.33	0.0777 (0.0121)	10.67 81.00 8.33	0.0990 (0.0180)	8.67 87.00 4.33

Table IV. Coverage and Length of 100(1-α)% C.I. for Lognormal with ξ = .193, δ^2=1, λ=1 at T = 1.

		68%		80%		90%	
		Length (s.d)	C.R.	Length (s.d)	C.R.	Length (s.d)	C.R.
Normal C.I. Procedure		0.0769 (0.0075)	16.67 71.00 12.33	0.1007 (0.0096)	10.00 78.00 12.00	0.1272 (0.0126)	6.67 89.33 4.00
Jackknife	Reduced d. f	0.0787 (0.0076)	16.33 71.33 12.33	0.1208 (0.0097)	9.67 79.67 10.67	0.1330 (0.0127)	6.33 89.67 4.00
	Regular d. f	0.0763 (0.0076)	16.67 70.33 13.00	0.1021 (0.0097)	10.00 79.00 11.00	0.1297 (0.0128)	6.67 89.33 4.00
Bootstrap	Percentile method	0.0763 (0.0084)	16.67 69.33 14.00	0.0998 (0.0105)	10.33 77.00 12.67	0.1258 (0.0133)	7.33 88.67 5.00
	Biascorrect method	0.0768 (0.0084)	16.00 68.33 15.67	0.1004 (0.0106)	8.67 78.33 13.00	0.1263 (0.0133)	6.33 88.67 5.00

Although the method of the reduced degrees of freedom used in the jackknife and the bias-correct percentile method applied in the bootstrap improved the coverage rate, the variance was inflated. Furthermore, the amount of improvement was small and not significant. Hence, the original procedures for constructing the confidence interval for the jackknife and bootstrap are preferred in this case. Note that the coverage rates are skewed left slightly but almost balanced. It is a reason that the normal confidence interval procedure performs well.

Results will now be reported for the simulation of the case in which the arrival rate of the Poisson process is also unknown and must be estimated from interarrival time data. More computations are required for this case; however, the procedure is same. Each replication of the simulation generated 50 independent service times and 50 independent exponential interarrival times having mean 1. Confidence intervals were computed using both separated and paired jackknife procedures and the percentile method for the bootstrap. The number of bootstrap replications was 1000. Nominal 80% confidence limits were computed for each replication. The simulation was replicated 300 times.

Table V. Coverage and Length of 80% C.I. with Unknown Arrival Rate (N=50, R=300, B=1000).

	Jackknife				Bootstrap	
	Separated		Paired		Percentile	
	Length (s.d)	C.R.	Length (s.d)	C.R.	Length (s.d)	C.R.
Exponential (μ = 2)	0.3447 (0.1329)	5.33 79.67 15.00	0.3282 (0.0641)	8.00 82.00 10.00	0.3162 (0.0585)	12.67 83.00 4.33
Mixed exponen (μ_1=2, μ_2=.75 P_1=.2)	0.2558 (0.1088)	75.67	0.2688 (0.0569)	5.67 82.33 12.00	0.2553 (0.0490)	12.67 83.00 4.33
Gamma (β=1, k=2)	0.3646 (0.1509)	5.00 78.67 19.33	0.3504 (0.0703)	6.67 81.67 11.67	0.3375 (0.0657)	16.33 75.67 8.00
Lognormal (§ = .193, δ^2=1)	0.3259 (0.1279)	4.67 77.67 16.67	0.3231 (0.0624)	6.33 79.33 14.33	0.3115 (0.0585)	14.67 75.67 10.00

The quantities in the left part of each cell are the average and standard deviation (with parenthesis) of coverage interval length. The right part of each cell contains three quantities; the top value is the proportion of intervals that are too high; the center value is the proportion of intervals that cover the true value, $\hat{p}$; and the bottom part is the proportion of intervals that are too low.

In Table V, the average length from the bootstrap shows outstanding performance with a small value of standard deviation. The paired jackknife procedure performs as well as the bootstrap procedure. This procedure reduced the standard deviation by more than half of that in the separated jackknife procedure, and also improved the coverage rate. From the results of coverage rate in the table, it can be recognized immediately that the jackknife estimate, regardless of the application method, is often too low, while the bootstrap estimate tends may a little too high but is almost balanced in the number of confidence intervals that are too high or too low. It is the reason that the bias-corrected percentile method was not required in this case.

In this simulation, all the coverage rates fall within $\pm 2\sqrt{\frac{\alpha(1-\alpha)}{300}}$ of 1-α (75.38, 84.61). Note that the average length of the confidence interval in the gamma service time case is the highest. When the arrival rate was known, the gamma service time case had the smallest average length. This indicates that the variability of the estimated arrival rate may be the dominate effect in the width of the confidence interval. Obviously the paired jackknife procedure performs very well. The bootstrap procedure still has the best performance; however the value of coverage rate fluctuates greatly for the different service time distributions. The paired jackknife improved the coverage rate tremendously. Although the average lengths in the paired jackknife procedure are slightly bigger than those in the bootstrap procedure, the overall performance is better than the bootstrap.

In general, all the confidence interval procedures performed very well for the exponential service time case, regardless of the level of the confidence interval. In most cases, the average length produced by the bootstrap procedure is the smallest, but the value of the coverage rate fluctuates for different service time distribution. The overall examination of the tabulations suggests that the paired jackknife procedure performs very well compared to the separated jackknife procedure and in some cases shows better performance than the bootstrap procedure.

V. Conclusion

In general, the nonparametric methods of the bootstrap and the jackknife performs very well, regardless the complexing of the estimation problem. Of course, if the parameter estimation methods can be applied, the results are clearly superior. However, the application of the parametric estimaton is a highly limited because the parametric assumption is ofen difficult to verify. When the estimate is simple enough, which is the nonparametric estimate when the arrival rate is known, and the asymptotic distribution of estimate can be obtained, the nonparametric normal confidence interval proprocedure performs well, and more complicate computations such as the jackknife and the bootstrap method are not required. However, the jackknife and the bootstrap method have a good performance for the more complicated problem in which the arrival rate is unknown. The bootstrap confidence intervals show the best performance but the paired jackknife procedure achieve the same level performance with less computation than the bootstrap in this problem.

References

1. Gross, D. and Harris, C.M., Fundamentals of Queueing Theory, John Wiley and Sons, 1974.
2. Brown, M. and Ross S.M. "Some Results for Infinite Server Queues", J. Appl. Prob., 6, pp. 604-611, 1969.
3. Takacs, L., Introduction to the Theory of Queues, Oxford University Press, 1962.
4. Cox, D.R., "Some Problems of Statistical Analysis Connected with Congestion, Proceedings of the Symposium on Congestion Theory, pp. 289-316, 1964.
5. Mirasol, N.M., "The Output of an M/G Queueing System is Poisson", Opns. Res., 11, pp. 282-284, 1963.
6. Lewis, P.A.W. and Uribe, L., "The New Naval Postgraduate School Random Number Package LLRANDOMII, NPS 55-81-005, Monterey, California, 1981.
7. Quenouille, M. H., "Notes an Bias in Estimation", Biometrika, 4, pp. 353-360, 1956.
8. Tukey, J.W., "Bias and Confidence in Not Quite Large Samples, Ann. Math. Statist., 29, p. 614, 1958.
9. Miller, R.G., "The Jackknife–a Review", Biometrika, 61, pp. 1-15, 1974.
10. Efron, B, "Bootstrap Methods: Another Look at the Jackknife", Ann. Statist., 7, pp. 1-26, 1979.
11. Efron, B., The Jackknife, the Bootstrap, and other Resampling Plans", Technical Report No. 63, Dept. of Statistics, Stanford University, 1980.
12. Feller, W. An Introduction to Probability Theory and its Applications, Vol. 2, New York, Wiley, 1966.
13. Efron, B., "Nonparametric Standard Errors and Confidence Intervals", Canadian Journal of Statistics, 9, pp. 139-172, 1981.