MANAGEMENT SCIENCE
Vol. 36, No 7, July 1990
Printed in U S A

CONTROL VARIATES FOR QUANTILE ESTIMATION*

JASON C. HSU AND BARRY L. NELSON
Department of Statistics, Ohio State University, Columbus, Ohio 43210-1271
Department of Industrial and Systems Engineering, Ohio State University, Columbus, Ohio 43210

New point and interval estimators for quantiles that employ a control variate are introduced. The properties of these estimators do not depend on the usual assumption of joint normality between the random variable of interest and the control. Illustrative examples for queueing and stochastic activity network models are given. In those examples, the new estimators are superior to the standard estimator in terms of the mean squared error of the point estimator and the length of the confidence interval.
(CONTROL VARIATES; QUANTILES; SIMULATION; VARIANCE REDUCTION; MAXIMUM LIKELIHOOD; UNBIASED TESTS)

1. Introduction

Let Y be a random variable with an unknown distribution, F_Y, but for which realizations can be obtained. This paper considers estimating the value y_q such that $\Pr\{Y \leq y_q\} = q$ for prespecified q $(0 < q < 1)$. The value y_q is called the qth quantile of Y.

Much of the literature on simulation output analysis concentrates on estimating $E[Y]$, the expected value of Y. Quantiles provide additional information about the distribution of Y. In fact, in some problems the quantiles of Y are the parameters of primary interest.

For example, Y could be a proposed test statistic whose distribution under the null hypothesis is difficult to evaluate numerically. One might then be interested in estimating the critical values $y_{0.90}$, $y_{0.95}$, and $y_{0.99}$ by simulating Y under the null hypothesis. As a second example, Y might be the delay in queue experienced by a customer arriving to a service system. Then 50% of the customers experience delays less than $y_{0.50}$, but 5% of the customers experience delays longer than $y_{0.95}$.

Straightforward estimation of y_q is based on the order statistics of Y (see §2 below). However, sometimes one can observe a *control* random variable X that is statistically dependent on Y and whose qth quantile, x_q, is known. §3 presents improved estimators based on pairs (X, Y) and x_q. §4 introduces a new confidence interval procedure. §5 presents some simple numerical examples. Some conclusions and recommendations are offered in §6.

2. The Standard Method

Let $Y_1, Y_2, \ldots, Y_n$ be an independent and identically distributed (i.i.d.) sample from a distribution F_Y that is absolutely continuous. Let $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$ be the sample Y values ordered from smallest to largest; these are the order statistics of the sample. If $k = [nq] + 1$, where $[\cdot]$ is the largest integer function, then $Y_{(k)}$ is the standard estimator of y_q (see David 1981 and Juritz, Juritz and Stephens 1983 for properties of this estimator).

Since $E[Y_{(k)}] \neq y_q$ and $\Pr\{Y_{(k)} \leq y_q\} \neq 1/2$, in general, $Y_{(k)}$ is neither mean nor median unbiased (an estimator is median unbiased if the true parameter is the median of the estimator). Thus, one may want to interpolate between order statistics. In this study we utilize the quantile function of the S statistical package (Becker and Chambers 1984), in which $Y_{(i)}$ is taken to be the $(i - 0.5)/n$th sample quantile, $i = 1, 2, \ldots, n$,

* Accepted by James R. Wilson; received January 6, 1988. This paper has been with the authors 4½ months for 2 revisions.

0025-1909/90/3607/835$01.25

and linear interpolation is employed for q between these values. When $q < 0.5/n$, the estimator is $Y_{(1)}$, and when $q > (n - 0.5)/n$, the estimator is $Y_{(n)}$. We call this interpolated estimator the "no control variate" (No CV) estimator. Other interpolation schemes have been proposed by Harrell and Davis (1982), Kaigh and Lachenbruch (1982), and Yang (1985). Kappenman (1987) integrates and inverts a kernel estimator of the density of Y.

The standard estimator is intuitively appealing. In the remainder of this section we show that the standard estimator arises from either of two estimation methods: inverting a test of hypothesis and (nonparametric) maximum likelihood estimation. In the sections that follow we derive new quantile estimators by applying these same two methods in conjunction with a control variate.

To understand how an estimator is derived by inverting a test of hypothesis, one must first understand the connection between tests and confidence sets (see also Lehmann 1986, p. 90). The explanation given below is more general than quantile estimation.

Suppose a random variable Y has a distribution that depends on $\theta \in \Theta$, the parameter space. For each $\theta^* \in \Theta$, let A_{θ^*} be the acceptance region of a size-α test for $\mathbf{H}$: $\theta = \theta^*$; i.e., $\Pr_{\theta^*}\{A_{\theta^*}\} = 1 - \alpha$. Then the set $C(Y) = \{\theta\colon Y \in A_\theta\}$ is an exact $(1 - \alpha)100\%$ confidence set for θ, since $\Pr_\theta\{\theta \in C(Y)\} = \Pr_\theta\{Y \in A_\theta\} = 1 - \alpha$ for every $\theta \in \Theta$.

However, the most familiar confidence sets (normal and t confidence intervals, for example) are derived by the pivotal method, and do not use the full power of this connection. Suppose Y has the same dimension as θ, and $Y - \theta$ has a known probability distribution P_0 that does not depend on θ. Choose the set A_0 such that $P_0\{A_0\} = 1 - \alpha$; then clearly the set $Y - A_0 = \{\theta\colon \theta = Y - a, a \in A_0\}$ is an exact $(1 - \alpha)100\%$ confidence set for θ. This is referred to as the pivotal method because $Y - A_0$ is obtained by pivoting $Y - \theta = A_0$. Notice that the pivotal method amounts to setting $A_\theta = A_0 + \theta$ for all θ in the testing framework above.

As an illustration of deriving an estimator by inverting a test, consider one-sample Hodges-Lehmann estimation: Suppose $Y_1, Y_2, \ldots, Y_n$ are i.i.d. with symmetric distribution $F(y - \theta)$. For testing $\mathbf{H}$: $\theta = \theta^*$, let R_i denote the rank of $|Y_i - \theta^*|$ in the joint ranking from least to greatest of $|Y_1 - \theta^*|, \ldots, |Y_n - \theta^*|$. Let $I_i = 1$ if $Y_i - \theta^* \geq 0$, and 0 otherwise. Then the signed rank test rejects $\mathbf{H}$ against the alternative $\mathbf{K}$: $\theta \leq \theta^*$ or $\mathbf{K}$: $\theta \geq \theta^*$ if

$$T = \sum_{i=1}^{n} I_i R_i \geq t \qquad \text{or} \qquad T = \sum_{i=1}^{n} I_i R_i \leq \frac{n(n+1)}{2} - t$$

respectively, where t is the smallest integer such that the test is of level α. By the correspondence between tests and confidence sets, a $1 - \alpha$ lower confidence bound is $W^{(a)}$, where $W^{(1)} \leq \cdots \leq W^{(n(n+1)/2)}$ are the ordered $(Y_i + Y_j)/2$ values with $i \leq j$ and $a = n(n+1)/2 - t + 1$ (Hollander and Wolfe 1973, p. 35). Similarly, a $1 - \alpha$ upper confidence bound is $W^{(b)}$, with $b = t$. A point estimator is derived by letting α approach $1/2$, which drives both $W^{(a)}$ and $W^{(b)}$ toward the median of $W^{(1)} \leq \cdots \leq W^{(n(n+1)/2)}$, which is the usual Hodges-Lehmann point estimator of θ (Hollander and Wolfe 1973, p. 33).

Returning to quantile estimation, a uniformly best estimator of y_q among median unbiased estimators based on Y that assumes no knowledge of F_Y can be obtained by inverting one-sided sign tests (Lehmann 1986, pp. 94–95 and pp. 120–121). When n is large, this best median unbiased estimator is typically an estimator that randomizes between $Y_{(k)}$ and $Y_{(k-1)}$ or $Y_{(k+1)}$. However, by the Rao-Blackwell Theorem, a nonrandomized version with smaller risk relative to any convex loss function (such as mean square error) can be obtained by taking the conditional expectation with respect to some sufficient statistic, the set of order statistics in this case. The resulting nonrandomized estimator is then a linear combination of $Y_{(k)}$ and $Y_{(k-1)}$ or $Y_{(k+1)}$. This nonrandomized estimator, which is no longer exactly median unbiased, is typically different from No CV, but not by much. Thus, the estimator No CV can be thought of as approximately the best median unbiased estimator based on Y when n is large.

Another method for deriving estimators is maximum likelihood estimation. Since we consider the case where F_Y is unknown, it is not possible to apply the maximum likelihood method directly. However, by taking a nonparametric approach as in the sign test—reducing the data to the number of observations less than or equal to y_q—it is possible to apply the principle of maximum likelihood in the following way: Let $M = \#\{Y_i \leq y_q\}$. The distribution of the count M is binomial with parameters n and q; that is

$$\Pr\{M = m \mid y_q\} = \binom{n}{m} q^m (1-q)^{n-m} \tag{1}$$

for $m = 0, 1, \ldots, n$. We take (1) as our likelihood function. Of course, M cannot actually be observed because y_q is not known. However, given a sample $y_1, \ldots, y_n$ and treating y_q as a variable, the value of y_q that maximizes (1) is any one that makes the sample value of M equal to m^*, where $(n+1)q - 1 < m^* \leq (n+1)q$ (Johnson and Kotz 1969, p. 53); this occurs if we set $y_q = y_{(m^*)}$. Notice the difference between this approach and standard maximum likelihood estimation: Given a sample, the sample value of M that we observe *depends on our candidate value for the parameter* y_q. Although our likelihood function is not a function of the sample directly, the value of M does depend on the sample through the relationship $M = \#\{Y_i \leq y_q\}$. The value m^* is nearly the same as k, so that No CV is also approximately a nonparametric maximum likelihood estimator of y_q, in the sense we use the term here.

The discussion above shows that we may derive the estimator No CV (at least for large n) by inverting unbiased tests or by nonparametric maximum likelihood estimation. Below we use these two methods, in conjunction with a control variate, to derive new estimators.

3. Control-Variate Estimators

Control variates (CVs) is a well-known variance reduction technique that estimates some characteristic of Y by exploiting knowledge about a random variable X that can be observed simultaneously with Y, and that is statistically dependent on Y. See Bratley, Fox, and Schrage (1987) for an introduction to CVs.

We now assume that there exists an X such that (X, Y) has joint distribution $F_{X,Y}$, which is absolutely continuous, and the qth quantile x_q of the marginal distribution of X is known. Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be an i.i.d. sample of (X, Y), and let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote the order statistics of X. In this section we develop estimators of y_q based on simulated pairs (X, Y) and x_q.

3.1. *A Regression-Based Estimator*

If we assume that $\mathrm{Cov}[X_{(k)}, Y_{(k)}] \neq 0$, then we might consider the classical control-variate estimator $Y_{(k)} - \beta(X_{(k)} - x_q)$ (e.g., Hammersley and Handscomb 1964, Chapter 5; Bratley, Fox, and Schrage 1987, Chapter 2). We refer to this estimator as the "regression estimator" (Reg). Unfortunately, $\mathrm{Cov}[X, Y] \neq 0$ does not guarantee that $\mathrm{Cov}[X_{(k)}, Y_{(k)}] \neq 0$. However, under the assumption of regression dependence (Tong 1980), we can show that $\mathrm{Cov}[X_{(k)}, Y_{(k)}] \neq 0$, so Reg might be expected to do better than the standard estimator $Y_{(k)}$.

The performance of Reg depends on β. The value of β that minimizes the variance of Reg is $\beta^* = \mathrm{Cov}[X_{(k)}, Y_{(k)}]/\mathrm{Var}[X_{(k)}]$. Typically, β^* is not known and must be estimated. In the context of quantile estimation, estimating β^* requires partitioning the size n sample into subsamples and calculating estimates of y_q from the smaller samples. Unfortunately, when estimating extreme quantiles single-sample estimators are usually less biased (Juritz, Juritz and Stephens 1983). If we fix the value of β arbitrarily, then Reg may be more variable than No CV. As an aside, we mention that fixing β guarantees that Reg is unbiased when estimating mean values, but it is not sufficient in the case of quantile estimation since $X_{(k)}$ is not an unbiased estimator of x_q.

Our new control-variate estimators can be computed without partitioning the sample. Thus, they are not directly comparable to Reg and we do not consider Reg further.

3.2. *New Quantile Estimators*

To motivate the new estimators derived below, consider Figure 1, which is a plot of a random sample of 100 pairs (X, Y); X and Y are strongly dependent. The vertical solid line is $X = x_{0.95}$, the known 0.95th quantile of X. The horizontal solid line represents a candidate for $y_{0.95}$, the unknown 0.95th quantile of Y. To estimate $y_{0.95}$ based on Y alone we would put the estimate somewhere between $Y_{(95)}$ and $Y_{(96)}$. Observe, however, that while the expected number of X's $> x_{0.95}$ is 5, in this sample there are 8 X's $> x_{0.95}$. Because X and Y are strongly dependent, one would guess that the number of Y's $> y_{0.95}$ in this sample is also 8, which would put $y_{0.95}$ somewhere between $Y_{(92)}$ and $Y_{(93)}$, as indicated by the dashed horizontal line. More generally, a large difference between the number of X's $> x_q$ and the number of Y's $> c$ is evidence against the candidate value c for y_q.

We can visualize the observed data in the (X, Y) plane as follows: Each hypothesized value c of y_q corresponds to a horizontal line $Y = c$ which, together with the known vertical line $X = x_q$, divides the (X, Y) plane into four quadrants or cells (see Figure 1). For notation, let

$N_{00}(c)$ = number of (X, Y) with $X \le x_q$ and $Y \le c$,
$N_{01}(c)$ = number of (X, Y) with $X \le x_q$ and $Y > c$,
$N_{10}(c)$ = number of (X, Y) with $X > x_q$ and $Y \le c$,
$N_{11}(c)$ = number of (X, Y) with $X > x_q$ and $Y > c$.

The $N_{ij}(c)$, $i, j = 1, 2$, are random variables. Let $n_{00}(c)$, $n_{01}(c)$, $n_{10}(c)$, and $n_{11}(c)$ be their realized values. Intuitively, if no knowledge is assumed concerning the joint distri-

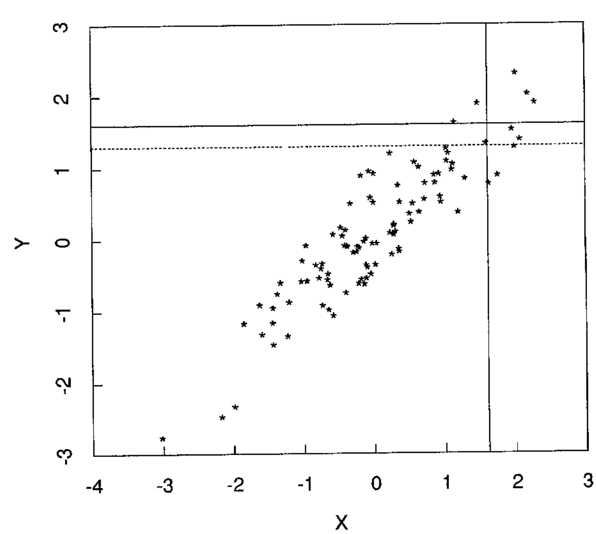

FIGURE 1. Sample Scatter Plot of 100 (x, y) Pairs, Known qth Quantile of X, and Candidates for qth Quantile of Y.

bution $F_{X,Y}$, then the essential information concerning y_q is contained in the four numbers $N_{00}(c)$, $N_{01}(c)$, $N_{10}(c)$, and $N_{11}(c)$. For a hypothesized value c of y_q, let

$p_{00}(c) = \Pr\{X \le x_q,\ Y \le c\}$,
$p_{01}(c) = \Pr\{X \le x_q,\ Y > c\}$,
$p_{10}(c) = \Pr\{X > x_q,\ Y \le c\}$,
$p_{11}(c) = \Pr\{X > x_q,\ Y > c\}$.

For any fixed c, the cell counts have a multinomial distribution; that is

$$\Pr\{N_{00}(c) = n_{00},\ N_{01}(c) = n_{01},\ N_{10}(c) = n_{10},\ N_{11}(c) = n_{11}\}$$
$$= \frac{n!}{n_{00}!n_{01}!n_{10}!n_{11}!}(q - p_{01}(c))^{n_{00}}p_{01}(c)^{n_{01}}p_{10}(c)^{n_{10}}(1 - q - p_{10}(c))^{n_{11}} \quad (2)$$

where $n_{00} + n_{01} + n_{10} + n_{11} = n$. Equation (2) will be used repeatedly to derive new estimators.

Two general methods of estimating an unknown parameter are inverting tests of hypotheses and the maximum likelihood method. In §3.2.1 and §3.2.2 we derive estimators of y_q by inverting tests for hypothesized values of y_q based on observed pairs (X, Y) and x_q. In §3.2.3 it is shown that, even with no knowledge of the joint distribution $F_{X,Y}$ beyond x_q, it is still possible to apply the maximum likelihood method to estimate y_q by reducing the data to $n_{00}(c)$, $n_{01}(c)$, $n_{10}(c)$, and $n_{11}(c)$.

3.2.1. *An Estimator Based on Inverting an Unbiased Test.* The hypothesis $\mathbf{H}$: $y_q = c$ is the same as $\mathbf{H}$: $p_{00}(c) + p_{01}(c) = p_{00}(c) + p_{10}(c)$ $(=q)$ or, equivalently, $\mathbf{H}$: $p_{01}(c) = p_{10}(c)$. Thus, estimators of y_q can be obtained from tests of the hypothesis $\mathbf{H}$: $p_{01}(c) = p_{10}(c)$. An approximately median unbiased estimator of y_q is derived in the Appendix by inverting uniformly most powerful unbiased tests for $\mathbf{H}$: $p_{01}(c) = p_{10}(c)$, namely McNemar's test. A different estimator of y_q is derived in the next section by inverting the likelihood ratio test for $\mathbf{H}$: $p_{01}(c) = p_{10}(c)$.

The resulting point estimator is a linear interpolation between $(X_{(m)}, Y_{(m)})$ and $(X_{(m+1)}, Y_{(m+1)})$ at x_q, where m is the number of X's less than or equal to x_q. (Also, when $m = n$, we take the estimate of y_q to be $Y_{(n)}$. When $m = 0$, we take the estimate of y_q to be $Y_{(1)}$.) We refer to this approximately median unbiased interpolated estimator as "Med Unb."

The performance of Med Unb relative to No CV depends on the unknown distribution $F_{X,Y}$. However, even without knowledge of $F_{X,Y}$, we can compare the two estimators in one respect.

If $c = y_q$, then $p_{01}(c) = p_{10}(c) = p$; notice that $0 \le p \le \min\{q, 1 - q\}$. Let $K = N_{00}(y_q) + N_{10}(y_q)$, the number of Y's less than or equal to y_q. Ideally, an estimator of y_q lies somewhere between $Y_{(K)}$ and $Y_{(K+1)}$, since $Y_{(K)} \le y_q$ and $Y_{(K+1)} > y_q$ by definition. The No CV estimator predicts that K is $k = [nq] + 1$. Med Unb predicts that K is $M = N_{00}(y_q) + N_{01}(y_q)$. One way to compare the estimators is to compare the differences $\Delta = k - K$ and $\Delta' = M - K = N_{01}(y_q) - N_{10}(y_q)$; that is, the difference in the number of order statistics between No CV or Med Unb and $Y_{(K)}$. The first two moments are easily calculated: $\mathrm{E}[\Delta] = k - nq$ versus $\mathrm{E}[\Delta'] = 0$, and $\mathrm{Var}[\Delta] = nq(1 - q)$ versus $\mathrm{Var}[\Delta'] = 2np$. Thus, the expected difference for Med Unb is 0, while the expected difference for No CV only converges to 0. Also, $\mathrm{Var}[\Delta'] \le \mathrm{Var}[\Delta]$ if and only if $p \le q(1 - q)/2$, and $\mathrm{Var}[\Delta'] \to 0$ as $p \to 0$. Thus, when $p \le q(1 - q)/2$ Med Unb tends to be closer to $Y_{(K)}$ than No CV.

3.2.2. *An Estimator Based on Inverting the Likelihood Ratio Test.* For fixed c, the likelihood function, as a function of $(p_{00}(c), p_{01}(c), p_{10}(c), p_{11}(c))$ and $(n_{00}(c), n_{01}(c), n_{10}(c), n_{11}(c))$, is proportional to

$$L = (q - p_{01}(c))^{n_{00}(c)}p_{01}(c)^{n_{01}(c)}p_{10}(c)^{n_{10}(c)}(1 - q - p_{10}(c))^{n_{11}(c)}.$$

The maximum likelihood estimators of $p_{01}(c)$ and $p_{10}(c)$ are (see the Appendix)

$$\hat{p}_{01} = \frac{q n_{01}(c)}{n_{00}(c) + n_{01}(c)}, \qquad \hat{p}_{10} = \frac{(1-q) n_{10}(c)}{n_{10}(c) + n_{11}(c)},$$

with asymptotic variance-covariance matrix

$$n^{-1} \begin{pmatrix} \dfrac{p_{01}(c)(q - p_{01}(c))}{q} & 0 \\ 0 & \dfrac{p_{10}(c)(1 - q - p_{10}(c))}{(1-q)} \end{pmatrix}.$$

The asymptotic likelihood ratio test for **H**: $p_{01}(c) = p_{10}(c)$ is based on the statistic

$$T = \frac{\sqrt{n}(\hat{p}_{01} - \hat{p}_{10})}{\sqrt{\dfrac{\hat{p}_{01}(q - \hat{p}_{01})}{q} + \dfrac{\hat{p}_{10}(1 - q - \hat{p}_{10})}{1-q}}}.$$

(Bickel and Doksum 1977, p. 212). In the Appendix we show that the point estimator obtained by inverting this test is the value $\hat{c}$ such that $\hat{p}_{01} = \hat{p}_{10}$. Since such a $\hat{c}$ may not exist, we take $\hat{c} = \sup\{c \mid \hat{p}_{01} - \hat{p}_{10} \geq 0\}$. We refer to this estimator as ILRT, for inverted likelihood ratio test.

The distribution of ILRT depends on $F_{X,Y}$. However, in the special case when $p_{01}(y_q) = p_{10}(y_q) = 0$, ILRT, like Med Unb, sets $\hat{c} = Y_{(M)}$, and thus $\hat{p}_{01} = \hat{p}_{10} = 0$ exactly.

3.2.3. *A Maximum Likelihood Estimator.* Recall that if $c = y_q$, then $p_{01}(c) = p_{10}(c) = p$. Therefore, the distribution of the cell counts $N_{00}(c)$, $N_{01}(c)$, $N_{10}(c)$, and $N_{11}(c)$ when $c = y_q$ is

$$\frac{n!}{n_{00}(c)!n_{01}(c)!n_{10}(c)!n_{11}(c)!}(q-p)^{n_{00}(c)} p^{n_{01}(c)+n_{10}(c)} (1-q-p)^{n_{11}(c)} = k(n;c)g(n;c,p) \quad (3)$$

where $k(n;c)$ is the multinomial term and $g(n;c,p)$ is the product of probabilities. We take (3) as our likelihood function, which depends on the observed sample indirectly through the cell counts.

Given a sample and treating p and $c = y_q$ as variables, we seek the values p^* and c^* that maximize (3). Notice that p is a nuisance parameter and c is the parameter of interest. We only need to consider values of c equal to the order statistics of Y, $\{Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}\}$, since the cell counts $n_{00}(c)$, $n_{01}(c)$, $n_{10}(c)$, $n_{11}(c)$, and thus the value of (3), change at those values.

No closed-form expression for (p^*, c^*) has been found. However, for fixed c, the value of p, denoted $p^*(c)$, that maximizes (3) is (see the Appendix)

$$\frac{q(n - n_{00}(c)) + (1-q)(n - n_{11}(c))}{2n} - \frac{\sqrt{(q(n - n_{00}(c)) - (1-q)(n - n_{11}(c)))^2 + 4q(1-q)n_{11}(c)n_{00}(c)}}{2n}.$$

An efficient recursion exists for calculating $k(n;c)$ as c increases (see the Appendix), which leads to an algorithm that steps through the possible values of c, determines $p^*(c)$ and the corresponding value of (3) for each c, and sets c^* equal to the value that maximizes (3). This estimator, obtained by maximizing the nonparametric likelihood function (3) with respect to y_q, will be referred to as the nonparametric maximum likelihood estimator (NPMLE).

4. Confidence Intervals

Although seldom reported in Monte Carlo studies, an interval estimate of y_q lends some insight into the precision of the point estimate. The standard $(1-\alpha)100\%$ confidence interval based on Y alone is $(Y_{(l_s)}, Y_{(u_s)})$, where $0 \le l_s < u_s \le n$ are integers such that

$$\sum_{k=l_s}^{u_s-1} \binom{n}{k} q^k (1-q)^{n-k} \approx 1-\alpha.$$

The confidence interval is nonparametric since the constants l_s and u_s depend only on n and q. Determining l_s and u_s is computationally expensive when n is large, but large n allows a normal approximation to the binomial distribution:

$$\sum_{k=l_s}^{u_s-1} \binom{n}{k} q^k (1-q)^{n-k} \approx \Phi\left(\frac{u_s - 1 - nq + 1/2}{\sqrt{nq(1-q)}}\right) - \Phi\left(\frac{l_s - nq - 1/2}{\sqrt{nq(1-q)}}\right)$$

where Φ is the distribution function of the standard normal distribution. The right-hand side equals $1-\alpha$ if $(l_s - nq - 1/2)/\sqrt{nq(1-q)} = -z_{1-\alpha/2}$ and $(u_s - 1 - nq + 1/2)/\sqrt{nq(1-q)} = z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution. Since integers l_s and u_s that satisfy these equations may not exist, we set $l_s = [nq - z_{1-\alpha/2}\sqrt{nq(1-q)} + 1/2]$ and $u_s = [nq + z_{1-\alpha/2}\sqrt{nq(1-q)} + 1/2] + 1$, which guarantees that $\Pr\{Y_{(l_s)} \le y_q < Y_{(u_s)}\} \ge 1-\alpha$ when the normal approximation to the binomial is adequate.

We seek an improved interval estimator by exploiting knowledge of the control variate X. In principle, we can invert McNemar's test (§3.2.1) or the likelihood ratio test (§3.2.2) to obtain confidence intervals. However, unless $\alpha = 1/2$ the endpoints of these intervals are difficult to compute. We consider instead a direct extension of the standard interval.

Suppose $c = y_q$. Then, suppressing the dependence of N_{ij} and p_{ij} on c for notational convenience,

$$\begin{aligned}\Pr\{N_{10} = r, N_{00} = s \mid N_{10} + N_{11} = n - m\} &= \binom{n-m}{r}\left(\frac{p}{1-q}\right)\left(\frac{1-q-p}{1-q}\right)^{n-m-r}\binom{m}{s}\left(\frac{q-p}{q}\right)\left(\frac{p}{q}\right)^{m-s} \\ &= B(r; n-m, p/(1-q))B(s; m, (q-p)/q) \qquad (4)\end{aligned}$$

for $r = 1, 2, \ldots, n-m$, $s = 1, 2, \ldots, m$, where $B(\cdot\,; n, p)$ is the binomial mass function with parameters n and p, and $p = p_{01} = p_{10}$ (Lehmann 1986, p. 158). The quantity $N_{10} + N_{11}$ is the number of observations with X value greater than x_q, which does not depend on knowledge of y_q. The random variables N_{10} and N_{00} are, together, the number of observations with Y value less than or equal to y_q. Equation (4) shows that, conditional on $N_{10} + N_{11}$, N_{10} and N_{00} are independent binomial random variables. Thus, for integers $0 \le l < u \le n$,

$$\begin{aligned}\Pr\{Y_{(l)} \le y_q < Y_{(u)} \mid N_{10} + N_{11} = n - m\} &= \sum_{i=l}^{u-1} \Pr\{N_{10} + N_{00} = i \mid N_{10} + N_{11} = n - m\} \\ &= \sum_{i=l}^{u-1}\sum_{j=0}^{i} B(j; n-m, p/(1-q))B(i-j; m, (q-p)/q). \qquad (5)\end{aligned}$$

Appropriate l and u yield a confidence interval that covers y_q with probability approximately $1-\alpha$, conditional on knowledge of the control variate. The interval estimator is exact, up to discreteness, and is nonparametric. There are, however, two difficulties: First, p is not known; to compute the interval we will substitute an estimate for p. Second, determining l and u such that (5) is close to $1 - \alpha$ is computationally expensive; below we construct a normal approximation. In the derivation that follows p is assumed known.

Conditional on $N_{10} + N_{11} = n - m$, the mean and variance of $N_{10} + N_{00}$ are

$$\mu(n-m) = \frac{(n-m)p}{1-q} + \frac{m(q-p)}{q} \quad \text{and}$$

$$\sigma^2(n-m) = \frac{(n-m)p(1-q-p)}{(1-q)^2} + \frac{m(q-p)p}{q^2}$$

respectively. We approximate the conditional distribution of $N_{10} + N_{00}$ by a normal distribution with mean $\mu(n-m)$ and variance $\sigma^2(n-m)$. An argument completely analogous to the standard interval leads to the conditional interval $(Y_{(l_c)}, Y_{(u_c)})$, where $l_c = [\mu(n-m) - z_{1-\alpha/2}\sigma(n-m) + 1/2]$ and $u_c = [\mu(n-m) + z_{1-\alpha/2}\sigma(n-m) + 1/2] + 1$. The approximation should be good when n is large or p is not too extreme.

We can give a rough argument for why the conditional interval should be superior to the standard interval. Let $M = n - (N_{10} + N_{11})$. Comparing the normal approximations, we notice that (l_s, u_s) are constants, but (l_c, u_c) are random variables since $n - M$ is a random variable. However, if p is known, then

$$\mathrm{E}[\mu(n-M)] = nq \quad \text{and} \quad \mathrm{E}[\sigma^2(n-M)] = \frac{np(2q(1-q)-p)}{q(1-q)}.$$

Then since

$$nq(1-q) - \frac{np(2q(1-q)-p)}{q(1-q)} = \frac{n}{q(1-q)}(q(1-q)-p)^2 \geq 0$$

we expect $l_s \leq l_c < u_c \leq u_s$; that is, the conditional interval should be shorter on average.

Both ILRT and NPMLE yield estimates of p as byproducts. The effect of substituting an estimate of p is illustrated below.

5. Examples

This section presents examples that illustrate the new quantile estimators and give some idea of their potential effectiveness. In all cases y_q is actually known, allowing point estimator bias and interval estimator coverage to be evaluated. Thus, these examples are not realistic in the sense that experimentation would never be used to estimate y_q when it can be easily calculated. However, they are illustrative of the type of problem for which the new quantile estimators can be used, and the control variates are like the control variates that would be available in more complex problems.

Experiments consisted of 100 samples of $n = 100$ and $n = 400$ (X, Y) pairs. In all cases $q = 0.95$. Experiments were performed on a Pyramid $90x$ super mini-computer.

Measures of point-estimator performance are mean square error (MSE), variance (Var), and bias (Bias). In addition, boxplots of the estimators are presented. The box in a boxplot contains the middle half of the data (i.e., from the 0.25th sample quantile to the 0.75th sample quantile); a horizontal line is drawn through the box at the median of the data. The whiskers extending from the box reach to the most extreme data points, or ± 1.5 times the interquartile range above and below the median, whichever is least. Points beyond the limits are plotted individually by "$*$". Thus, boxplots summarize the sampling distribution of the estimators.

Measures of interval-estimator performance are the mean, variance and coefficient of variation of the interval halfwidth, and the probability the interval contains y_q. These measures are displayed via halfwidth versus midpoint plots (Kang and Schmeiser 1989). Points inside the 45 degree angle with its base at y_q indicate intervals that contained y_q; points to the right indicate intervals whose lower endpoint was greater than y_q, and points to the left indicate intervals whose upper endpoint was less than y_q.

5.1. *M/M/1 Queue*

The $M/M/1$ queue is a single server, first-come-first-served service system in which customers arrive according to a Poisson process and service times are i.i.d. negative exponential random variables. Let Y be the delay in queue (not including service) experienced by the lth ($l > 0$) customer to arrive to an $M/M/1$ queue that had $h \geq 0$ customers present at time 0. The control variate X is the difference between the sum of the service times of the first $l + h - 1$ customers and the sum of the interarrival times of the first l customers. The distribution of X is the difference of independent Erlang random variables and the distribution of Y is a mixture of Erlangs (Kelton and Law 1985).

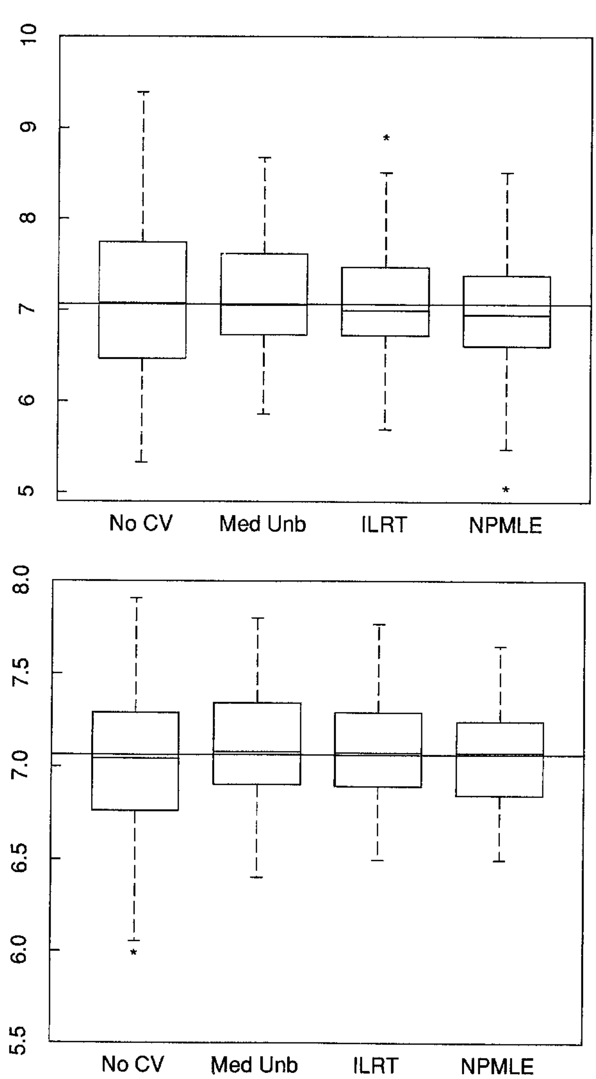

FIGURE 2. Boxplots of Estimates of $y_{0.95}$ for the $M/M/1$ Example with (a) $n = 100$, (b) $n = 400$.

Observations (X, Y) were generated by a FORTRAN simulation using IMSL subroutine rnexp to generate interarrival and service times. The value of $x_{0.95}$ was obtained as follows: Since $X = E_1 - E_2$, where E_1 and E_2 are independent Erlang random variables, the cdf of X was expressed as a single-variable integral by conditioning on E_2. Numerical integration via IMSL routine qdagi was used to determine $F_X(x)$ for any x. A bisection search was employed to find the value $x_{0.95}$ such that $F_X(x_{0.95}) \approx 0.95$, with relative error 0.001 for the numerical integration and absolute error 0.0001 for the bisection search. The cdf of Y was evaluated using the algorithm of Kelton and Law (1985), and the same bisection method was used to determine $y_{0.95}$.

The example below is an $M/M/1$ queue with arrival rate 0.9 customers/unit time, service rate 1 customer/unit time, and $h = 0$ customers present at time 0. We consider the delay in queue in the 10th arriving customer. Thus, with probability 0.95 the 10th customer to arrive for service waits no more than $y_{0.95}$ time units for service to begin. The sample correlation between X and Y, based on 40,000 pairs, was 0.76 which seems to indicate strong dependence.

Figures 2(a) and (b) show boxplots of the 100 values of each estimator for $n = 100$ and $n = 400$, respectively. Table 1 gives the numerical values of MSE, variance, and bias for these experiments. The MSE reductions for the best control variate estimator in each case is approximately 50%. At the larger sample size ILRT and NPMLE seem to perform better than Med Unb.

Figures 3(a) and (b) show midpoint by halfwidth plots for the standard and conditional confidence intervals when $1 - \alpha = 0.95$ for $n = 100$ and $n = 400$, respectively. Plusses represent the standard interval and circles represent the conditional interval. Points that are lower (shorter halfwidth) and centered within the 45 degree angle (y_q in the middle of the interval) are preferred. The probability of coverage is apparent from the number of points outside the 45 degree angle, since there are 100 intervals constructed via each method. Table 2 gives the numerical values of the mean, variance and coefficient of variation of the halfwidth, and the estimated probability of coverage.

In experiments not reported, estimators for p from both ILRT and NPMLE were evaluated, and the estimator from NPMLE appeared to be slightly better when n and/or p is small; they performed equally well when n is large and/or p is not too extreme. However, ILRT does have the advantage that, when $p = 0$, the estimate from ILRT is always 0. Nevertheless, the estimator from NPMLE was used to construct the conditional interval in all experiments reported here.

When $p = 0$ the conditional interval becomes two adjacent order statistics. Even though $p > 0$ in this example, when the sample size is small the estimator of p is quite variable

TABLE 1

Point Estimator Performance for $M/M/1$ and SAN Examples

Data Set		No CV	Med Unb	ILRT	NPMLE
$M/M/1$	MSE	0.766	0.384	0.375	0.484
$n = 100$	Var	0.756	0.376	0.375	0.477
	Bias	0.096	0.088	0.001	−0.084
$n = 400$	MSE	0.164	0.103	0.075	0.076
	Var	0.162	0.100	0.075	0.076
	Bias	−0.042	0.057	0.015	−0.001
SAN	MSE	0.357	0.162	0.191	0.202
$n = 100$	Var	0.357	0.157	0.191	0.182
	Bias	0.013	0.068	−0.011	−0.140
$n = 400$	MSE	0.082	0.043	0.035	0.042
	Var	0.081	0.043	0.035	0.039
	Bias	−0.034	0.010	−0.013	−0.050

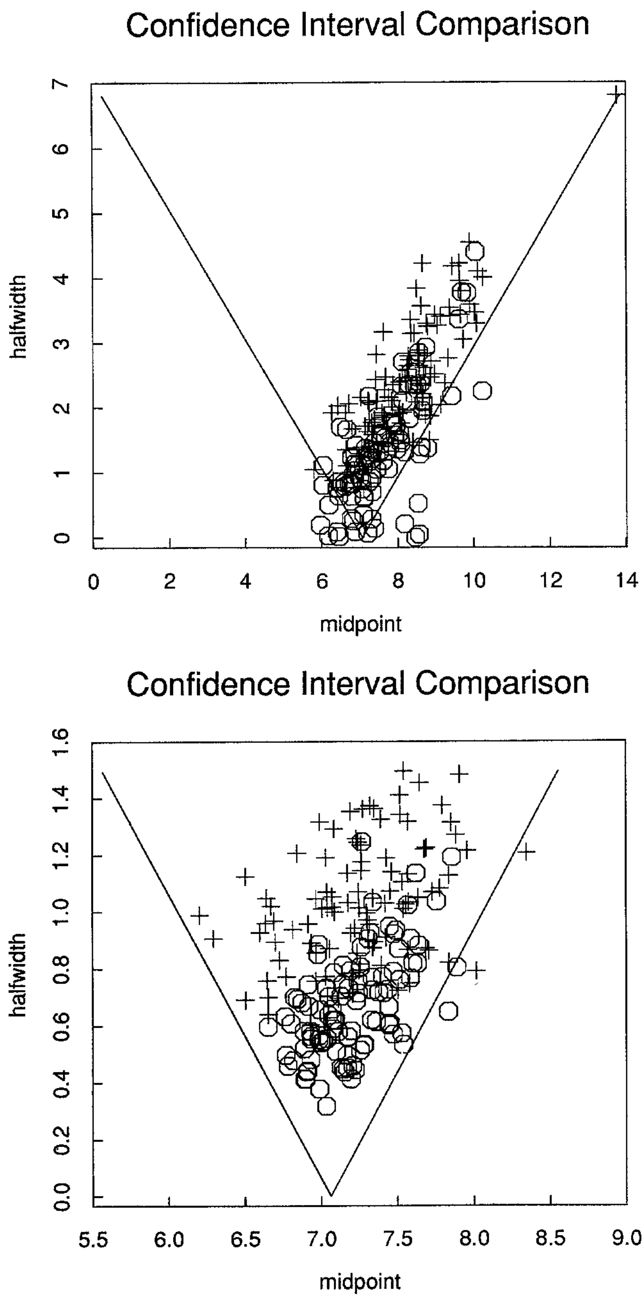

FIGURE 3. Halfwidth Versus Midpoint Plots for Standard (Plus) and Conditional (Circle) Intervals for the $M/M/1$ Example with $\alpha = 0.05$, $q = 0.95$ and (a) $n = 100$, (b) $n = 400$.

so that the estimate of p may be zero. This explains the short intervals that do not cover $y_{0.95}$ in Figure 3(a). Figure 3(b) shows that the problem disappears when $n = 400$ and the estimate of p is more stable. Certainly when the sample size is large, the conditional interval produces shorter intervals on average while maintaining the desired coverage probability. The results of experiments not reported showed little or no undercoverage problem when q is less extreme, in which case the estimate of p would tend to be greater than 0.

TABLE 2

Confidence Interval Performance for M/M/1 and SAN Examples (Mean, Variance and Coefficient of Variation of the Halfwidth, and Coverage)

Data Set		Standard	Conditional
$M/M/1$	Mean	2.52	1.33
$n = 100$	Variance	0.85	0.74
	Coef. Variation	0.37	0.65
	Coverage	0.97	0.81
$n = 400$	Mean	1.03	0.67
	Variance	0.05	0.03
	Coef. Variation	0.21	0.27
	Coverage	0.98	0.98
SAN	Mean	1.74	0.82
$n = 100$	Variance	0.57	0.39
	Coef. Variation	0.43	0.76
	Coverage	0.94	0.74
$n = 400$	Mean	0.67	0.40
	Variance	0.02	0.02
	Coef. Variation	0.21	0.37
	Coverage	0.99	0.92

5.2. *Stochastic Activity Network*

Stochastic activity networks (SANs) are used to model and manage the progress of large projects. A SAN is composed of a collection of nodes and directed arcs, with one node representing the beginning of the project and another representing the completion of the project. The arcs represent activities that are part of the project and must be completed in sequence. The time required to complete each activity is often modeled as a random variable. One quantity of interest is the time required to complete the project.

Consider the SAN in Figure 4. Let T_i be the time required to complete activity i, $i = 1, 2, \ldots, 5$, and let $Y = \max\{T_1 + T_2, T_1 + T_3 + T_5, T_4 + T_5\}$, the time to complete the project. Thus, with probability 0.95 the project will be completed by or before $y_{0.95}$. In this example we assume that the activity times are i.i.d. negative exponentially distributed random variables with mean 1. As the control variate we use $X = T_1 + T_3 + T_5$, which has an Erlang distribution. The sample correlation between X and Y was 0.87.

Observations (X, Y) were generated by a FORTRAN simulation using IMSL subroutine rnexp to generate activity times. The value of $x_{0.95}$ was determined by the S function qgamma, which numerically inverts the gamma (and thus Erlang) distribution. The cdf of Y, which was derived via a sequence of conditioning arguments, is

$$F_Y(y) = 1 - e^{-3y} + (\tfrac{1}{2}y^2 - 3y - 3)e^{-2y} + (-\tfrac{1}{2}y^2 - 3y + 3)e^{-y}$$

for $y \geq 0$. A bisection search was used to find $y_{0.95}$.

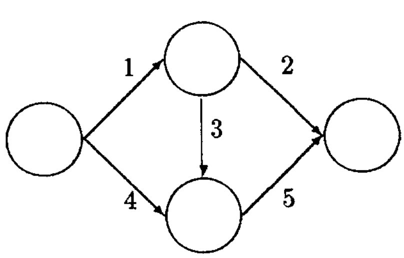

FIGURE 4. Stochastic Activity Network (SAN) Example.

Figures 5(a) and (b) show boxplots of the 100 values of each estimator for $n = 100$ and $n = 400$, respectively. Table 1 gives the numerical values of MSE, variance, and bias for these experiments. The MSE reductions for the best control variate estimator in each case is more than 50%. Med Unb and ILRT seem to be the least biased.

Figures 6(a) and (b) show midpoint by halfwidth plots for the standard and conditional confidence intervals for $n = 100$ and $n = 400$, respectively. Table 2 gives the numerical values of the mean, variance and coefficient of variation of the halfwidth, and the estimated probability of coverage. Again, undercoverage is apparent when $n = 100$, but coverage is indistinguishable from 0.95 when $n = 400$.

6. Conclusions

Variance reduction research has concentrated on efficiently estimating population means and variances, which are just two of the characteristics of the population (see Nelson 1987a for a survey of variance reduction). Quantiles provide additional infor-

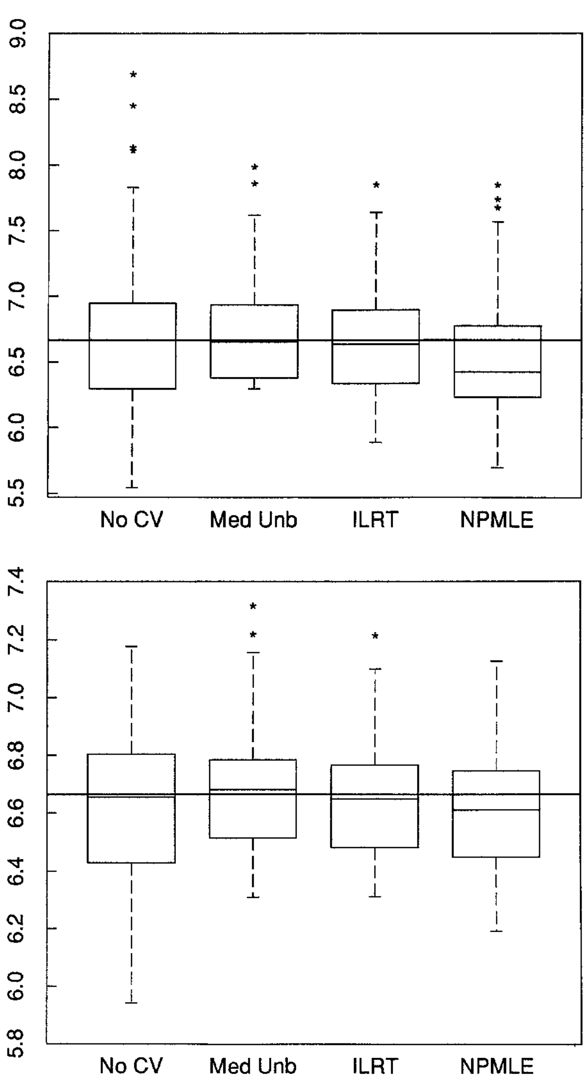

FIGURE 5. Boxplots of Estimates of $y_{0.95}$ for the SAN Example with (a) $n = 100$, (b) $n = 400$.

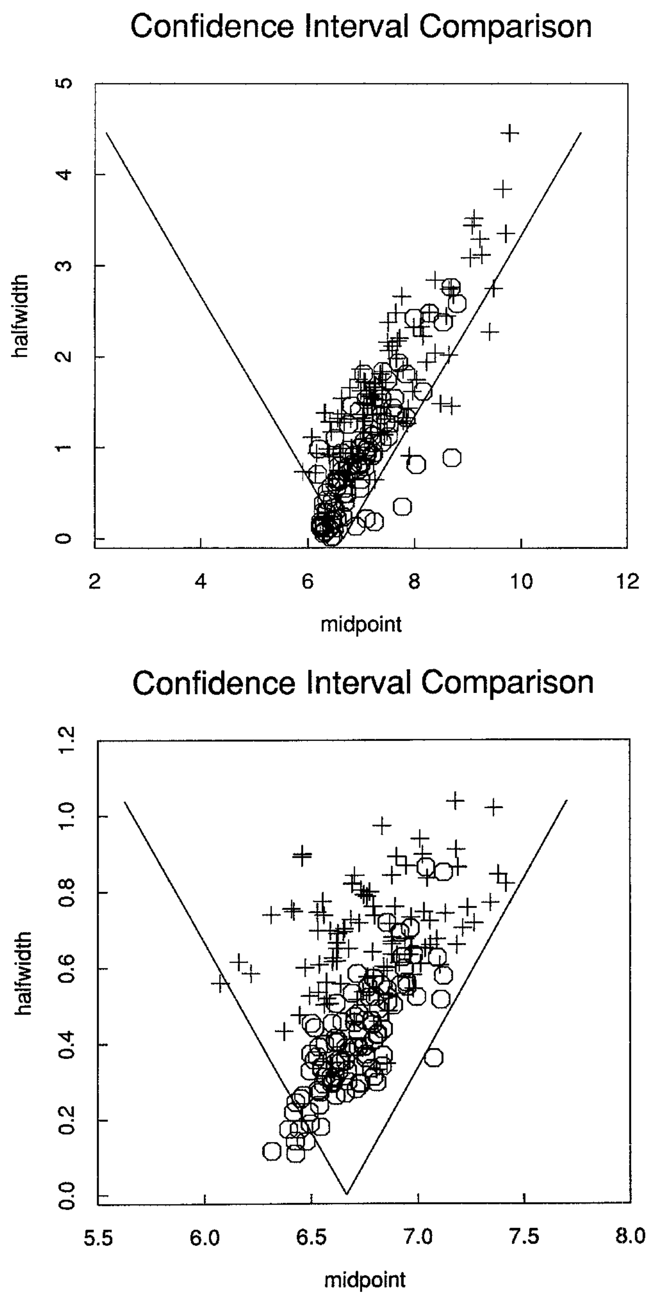

FIGURE 6. Halfwidth Versus Midpoint Plots for Standard (Plus) and Conditional (Circle) Intervals for the SAN Example with $\alpha = 0.05$, $q = 0.95$ and (a) $n = 100$, (b) $n = 400$.

mation about the population, and can in fact be the parameters of primary interest in certain problems. Thus, it is important to develop good techniques for estimating quantiles.

Techniques based on regression have been the primary focus of control variate research (Glynn and Whitt 1989, Nelson 1987b, and Rothery 1982 are some exceptions). We propose quantile estimators that are based on estimating the joint probabilistic behavior of the variable of interest and the control variate.

The empirical evaluation presented here shows the three new control variate estimators to be promising. Based on these results and many others not presented, we recommend

ILRT, the estimator based on inverting the likelihood ratio test, as the best general-purpose quantile estimator. This recommendation is based on the observation that ILRT performed well, if not always the best, on all cases considered. An additional benefit of ILRT over Med Unb is that it yields an estimate of p as a byproduct which can be used to form the conditional confidence interval. We recommend the conditional interval when n is large or q is not extreme.

An important next step is to consider estimating quantiles of the limiting distribution of a stationary stochastic process. This problem is of interest to simulators, but is complicated by dependence within the data. Heidelberger and Lewis (1984), Iglehart (1976), and Seila (1982) propose estimators, but do not consider variance reduction.[1]

[1] Jason Hsu's research is made possible in part by Grant No. 1R01 CA41168, awarded by the National Cancer Institute. Barry Nelson's research was partially supported by National Science Foundation Grant No. ECS-8707634. Computing resources were provided by the Mathematical Sciences Computing Laboratory of The Ohio State University. The authors acknowledge the helpful comments of two referees and the Departmental Editor.

Appendix

Derivation of the Median Unbiased Estimator

We define $X_{(0)} = X_{(1)}$, $Y_{(0)} = Y_{(1)}$, $X_{(n+1)} = X_{(n)}$, and $Y_{(n+1)} = Y_{(n)}$, for convenience. The derivation of Med Unb depends on the following lemma:

LEMMA. *A uniformly most powerful unbiased (UMPU) size-α test, ϕ_c^-, for* $\mathbf{H}$: $y_q = c$ *versus* $\mathbf{K}$: $y_q > c$ *based on $N_{00}(c)$, $N_{01}(c)$, $N_{10}(c)$, and $N_{11}(c)$ exists. It is McNemar's test, which is a conditional test that rejects for small values of $N_{10}(c)$, conditional on $N(c) = N_{01}(c) + N_{10}(c)$.*

Let $n(c)$ be the realized value of $N(c)$, and let b be the positive integer such that

$$\alpha_{b-1} = \sum_{j=0}^{b-1} \binom{n(c)}{j}\left(\frac{1}{2}\right)^{n(c)} < \alpha \leq \sum_{j=0}^{b} \binom{n(c)}{j}\left(\frac{1}{2}\right)^{n(c)} = \alpha_b .$$

Let

$$\gamma = \frac{(\alpha - \alpha_{b-1})}{\binom{n(c)}{b}\left(\frac{1}{2}\right)^{n(c)}}$$

so that $(1-\gamma)\alpha_{b-1} + \gamma\alpha_b = \alpha$. Then the UMPU test rejects if $n_{10}(c) < b$, or with probability γ when $n_{10}(c) = b$.

PROOF. The hypothesis and alternative $\mathbf{H}$: $y_q = c$ versus $\mathbf{K}$: $y_q > c$ is the same as $\mathbf{H}$: $p_{01}(c) = p_{10}(c)$ versus $\mathbf{K}$: $p_{01}(c) > p_{10}(c)$, for which the one-sided McNemar's test is UMPU (Lehmann 1986, §4.9).

Similarly, the UMPU test, ϕ_c^+, for $\mathbf{H}$: $y_q = c$ versus $\mathbf{K}$: $y_q < c$ based on $N_{00}(c)$, $N_{01}(c)$, $N_{10}(c)$, and $N_{11}(c)$ is the one-sided McNemar's test which rejects if $n_{01}(c) < b$, or with probability γ when $n_{01}(c) = b$.

Thus, by the usual correspondence between tests and confidence sets (Lehmann 1986, Theorem 3.4), $y_q^- = \inf\{c \mid \mathbf{H}\colon y_q = c$ is accepted by $\phi_c^-\}$ is a level $1 - \alpha$ lower confidence bound for y_q. Likewise, $y_q^+ = \sup\{c \mid \mathbf{H}\colon y_q = c$ is accepted by $\phi_c^+\}$ is a level $1 - \alpha$ upper confidence bound for y_q. One way to derive a median unbiased estimator is to look for a common value of y_q^- and y_q^+ when $\alpha = 1/2$ (Lehmann 1986, pp. 94–95). We next show that the resulting point estimator is a randomized estimator that selects $Y_{(m)}$ or $Y_{(m+1)}$, each with probability $1/2$, where m is the number of X's less than or equal to x_q.

Let $m = n_{00}(c) + n_{01}(c)$, which does not depend on the hypothesized value c of y_q. When $\alpha = 1/2$, the critical value is $b = n(c)/2 = (n_{01}(c) + n_{10}(c))/2$. Thus, since each hypothesis $\mathbf{H}$: $y_q = c$ with $c < Y_{(m)}$ gives $n_{01}(c) > n_{10}(c)$, it is rejected by ϕ_c^-. And, since each $\mathbf{H}$: $y_q = c$ with $c \geq Y_{(m+1)}$ gives $n_{01}(c) < n_{10}(c)$, it is rejected by ϕ_c^+. Thus, $c \in [Y_{(m)}, Y_{(m+1)})$ are the only candidate estimates.

Every $\mathbf{H}$: $y_q = c$ with $c \in [Y_{(m)}, Y_{(m+1)})$ gives $n_{01}(c) = n_{10}(c)$. Thus, for $c \in [Y_{(m)}, Y_{(m+1)})$, $\mathbf{H}$: $y_q = c$ is rejected by ϕ_c^- with probability γ. The same $\mathbf{H}$: $y_q = c$ is also rejected by ϕ_c^+ with probability γ.

When $\alpha = 1/2$, $\gamma = 1/2$ due to the symmetry of the binomial distribution with parameter $1/2$. Let U be a random variable that is uniformly distributed on the interval $(0, 1)$, and that is independent of X and Y. Suppose ϕ_c^- rejects when $U < 1/2$. Then $y_q^- = Y_{(m+1)}$ when $U < 1/2$, and $Y_{(m)}$ otherwise. Suppose ϕ_c^+ rejects when $U \geq 1/2$. Then $y_q^+ = Y_{(m+1)}$ when $U < 1/2$, and $Y_{(m)}$ otherwise. Thus, if the same auxiliary random variable U is employed for both the tests, then $y_q^- = y_q^+$.

This randomized estimator can be expected to have good properties, since it is derived from UMPU tests. However, randomization is not very appealing in practice. Further, according to the Rao-Blackwell Theorem (Lehmann 1983, pp. 50–51), a nonrandomized version with smaller risk (expected loss) relative to any strictly convex loss function (e.g., mean square error) can be obtained by taking the conditional expectation with respect to a sufficient statistic (the order statistics of X and Y in this case). The Rao-Blackwellized estimator is $(Y_{(m)} + Y_{(m+1)})/2$. We chose to linearly interpolate between $(X_{(m)}, Y_{(m)})$ and $(X_{(m+1)}, Y_{(m+1)})$ at x_q as described in §3.2.1.

Derivation of ILRT

The maximum likelihood estimators of $p_{01}(c)$ and $p_{10}(c)$ are obtained by solving the system of equations

$$\frac{\partial \log L}{\partial p_{01}} = \frac{-n_{00}(c)}{q - p_{01}(c)} + \frac{n_{01}(c)}{p_{01}(c)} = 0, \qquad \frac{\partial \log L}{\partial p_{01}} = \frac{n_{10}(c)}{p_{10}(c)} + \frac{-n_{11}(c)}{1 - q - p_{10}(c)} = 0.$$

Let $\hat{p}_{01}$ and $\hat{p}_{10}$ be the solution; we drop the argument c for convenience. The inverse of the asymptotic variance-covariance matrix of $(\hat{p}_{01}, \hat{p}_{10})'$ is (Kendall and Stuart 1979, p. 59)

$$-\begin{pmatrix} \mathrm{E}\dfrac{\partial^2 \log L}{\partial p_{01}^2} & \mathrm{E}\dfrac{\partial^2 \log L}{\partial p_{01}\partial p_{10}} \\ \mathrm{E}\dfrac{\partial^2 \log L}{\partial p_{10}\partial p_{01}} & \mathrm{E}\dfrac{\partial^2 \log L}{\partial p_{10}^2} \end{pmatrix} = \begin{pmatrix} \dfrac{nq}{p_{01}(q - p_{01})} & 0 \\ 0 & \dfrac{n(1-q)}{p_{10}(1 - q - p_{10})} \end{pmatrix}. \tag{A1}$$

The asymptotic variance-covariance matrix is obtained by inverting (A1).

The hypothesis $\mathbf{H}$: $y_q = c$ is equivalent to $\mathbf{H}$: $p_{01}(c) = p_{10}(c)$. Let

$$T = \frac{\sqrt{n}(\hat{p}_{01} - \hat{p}_{10})}{\sqrt{\dfrac{\hat{p}_{01}(q - \hat{p}_{01})}{q} + \dfrac{\hat{p}_{10}(1 - q - \hat{p}_{10})}{1 - q}}}$$

which converges in distribution to the standard normal distribution if $\mathbf{H}$ is correct. Thus, an asymptotic level-α test for $\mathbf{H}$: $p_{01}(c) = p_{10}(c)$ versus $\mathbf{K}^-$: $p_{01}(c) > p_{10}(c)$ rejects if $T > z_{1-\alpha}$. Similarly, $\mathbf{H}$ is rejected at level α in favor of $\mathbf{K}^+$: $p_{01}(c) < p_{10}(c)$ if $T < -z_{1-\alpha}$.

When $\alpha = 1/2$, $\mathbf{H}$ is rejected in favor of $\mathbf{K}^-$ if $T > 0$, and in favor of $\mathbf{K}^+$ if $T < 0$. Since $\hat{p}_{01}(c) - \hat{p}_{10}(c)$ is a nonincreasing function of c, then

$$\hat{c} = \inf\{c \mid \mathbf{H} \text{ is accepted versus } \mathbf{K}^-\} = \sup\{c \mid \mathbf{H} \text{ is accepted versus } \mathbf{K}^+\}$$

must satisfy $\hat{p}_{01}(\hat{c}) = \hat{p}_{10}(\hat{c})$.

Derivation of NPMLE

In this section we show that NPMLE is the nonparametric maximum likelihood estimator of y_q. When c is fixed, $k(n; c)$ is a constant; thus the likelihood function is proportional to $g(n; c, p)$. Let $\bar{m} = n_{10}(c) + n_{11}(c)$, the number of Xs greater than x_q, which does not depend on c. Then

$$g(n; c, p) = (q - p)^{r(c) - n_{10}(c)} p^{n - \bar{m} - r(c) + 2n_{10}(c)} (1 - q - p)^{\bar{m} - n_{10}(c)}$$

where $r(c) = n_{00}(c) + n_{10}(c)$, the number of Ys not exceeding c. Notice that $g(n; c, 0) = 0$, unless $n_{01} = n_{10} = 0$; $g(n; c, q) = 0$, unless $n_{00} = 0$; $g(n; c, 1 - q) = 0$, unless $n_{11} = 0$; $g(n; c, p) > 0$ for $0 < p < \min\{q, 1 - q\}$; and g is bounded. Thus, g has a maximum at some $p \in (0, \min\{q, 1 - q\})$, except for the special cases mentioned (we deal with these cases later). Since g is continuous and differentiable in this interval, the maximum must occur where the $\partial g/\partial p = 0$. We temporarily drop the argument c, since it is fixed.

Let $a = -q(n - r + n_{10}) - (1 - q)(n - \bar{m} + n_{10})$ and $d = q(1 - q)(n - \bar{m} - r + 2n_{10})$. Considerable algebra shows that

$$\frac{\partial g}{\partial p} = p^{n - \bar{m} - r + 2n_{10} - 1}(q - p)^{r - n_{10} - 1}(1 - q - p)^{\bar{m} - n_{10} - 1}(np^2 + ap + d). \tag{A2}$$

Clearly, (A2) is zero at $p = 0$, at $p = q$ and at $p = 1 - q$. We seek the maximum between 0 and min $\{q, 1 - q\}$. From the last term on the right-hand side of (A2), it is also zero at

$$p = \frac{-a \pm \sqrt{a^2 - 4nd}}{2n}.$$

Tedious algebra shows that $a^2 \geq 4nd$. Computing $\partial^2 g/\partial p^2$ shows that a local maximum occurs at

$$p^*(c) = \frac{-a - \sqrt{a^2 - 4nd}}{2n}. \tag{A3}$$

Finally, we can show that $0 \le p^* \le \min\{q, 1-q\}$, which proves that the absolute maximum in the interval occurs at p^*. Noting that $r(c) - n_{10}(c) = n_{00}(c)$, and $\bar{m} - n_{10}(c) = n_{11}(c)$, (A3) can be algebraically reduced to the result in §3.2.3.

The special cases are handled as follows: If $n_{01}(c) = n_{10}(c) = 0$ then $p^*(c) = 0$, which can be shown to maximize the likelihood function. If $n_{00}(c) = n_{11}(c) = 0$ then the likelihood function is maximized at $p = \min\{q, 1-q\}$. If only $n_{00}(c) = 0$ then the likelihood function is maximized at $p = \min\{1-q, q(n_{01}(c) + n_{10}(c))/n\}$. If only $n_{11}(c) = 0$ then the likelihood function is maximized at $p = \min\{q, (1-q)(n_{01}(c) + n_{10}(c))/n\}$.

Calculation of $k(n; c)$ as c increases is facilitated by the following recursion: The function $n_{10}(c)$ is a nondecreasing function that increases in unit jumps only when c equals an order statistic of Y. Suppose $c \in [Y_{(i)}, Y_{(i+1)})$, for some $i \in \{1, 2, \ldots, n-1\}$. If $c' \in [Y_{(i+1)}, Y_{(i+2)})$, then $k(n; c') = k(n; c)h(c, c')$, where

$$h(c, c') = \begin{cases} \dfrac{n_{11}(c)}{n_{10}(c)+1} & \text{if} \quad n_{10}(c') > n_{10}(c), \\[2ex] \dfrac{n_{01}(c)}{n_{00}(c)+1} & \text{otherwise.} \end{cases}$$

References

BECKER, R. A. AND J. M. CHAMBERS, *S—An Interactive Environment for Data Analysis and Graphics*, Wadsworth, Belmont, CA, 1984.

BICKEL, P. J. AND K. A. DOKSUM, *Mathematical Statistics: Basic Ideas and Selected Topics*, Holden-Day, San Francisco, 1977.

BRATLEY, P., B. L. FOX AND L. E. SCHRAGE, *A Guide to Simulation*, 2nd Ed., Springer-Verlag, New York, 1987.

DAVID, H. A., *Order Statistics*, 2nd Ed., Wiley, New York, 1981.

GLYNN, P. W. AND W. WHITT, "Indirect Estimation via $L = \lambda W$," *Oper. Res.*, 37 (1989), 82–103.

HAMMERSLEY, J. M. AND D. C. HANDSCOMB, *Monte Carlo Methods*, Chapman and Hall, London, 1964.

HARRELL, F. E. AND C. E. DAVIS, "A New Distribution-free Quantile Estimator," *Biometrika*, 69 (1982), 635–640.

HEIDELBERGER, P. AND P. A. W. LEWIS, "Quantile Estimation in Dependent Sequences," *Oper. Res.*, 32 (1984), 185–209.

HOLLANDER, M. AND D. WOLFE, *Nonparametric Statistical Methods*, Wiley, New York, 1973.

IGLEHART, D. L., "Simulating Stable Stochastic Systems; VI. Quantile Estimation," *J. Assoc. Comput. Mach.*, 23 (1976), 347–360.

JOHNSON, N. L. AND S. KOTZ, *Discrete Distributions*, Houghton-Mifflin, Boston, 1969.

JURITZ, J. M., J. W. F. JURITZ AND M. A. STEPHENS, "On the Accuracy of Simulated Percentage Points," *J. Amer. Statist. Assoc.*, 78 (1983), 441–444.

KAIGH, W. D. AND P. A. LACHENBRUCH, "A Generalized Quantile Estimator," *Comm. Statist.*, A11 (1982), 2217–2238.

KANG, K. AND B. W. SCHMEISER, "Methods for Evaluating and Comparing Confidence-interval Procedures," *Oper. Res.*, forthcoming.

KAPPENMAN, R. F., "Improved Distribution Quantile Estimation," *Comm. Statist.*, B16 (1987), 307–320.

KELTON, W. D. AND A. M. LAW, "The Transient Behavior of the $M/M/s$ Queue, with Implications for Steady-state Simulation," *Oper. Res.*, 33 (1985), 378–396.

KENDALL, M. AND A. STUART, *The Advanced Theory of Statistics, Vol. 2. Inference and Relationship*, Griffin and Company, London, 1979.

LEHMANN, E. L., *Theory of Point Estimation*, Wiley, New York, 1983.

———, *Testing Statistical Hypotheses*, 2nd Ed., Wiley, New York, 1986.

NELSON, B. L., "A Perspective on Variance Reduction in Dynamic Simulation Experiments," *Comm. Statist.*, B16 (1987a), 385–426.

———, "On Control Variate Estimators," *Computers & Oper. Res.*, 14 (1987b), 219–225.

ROTHERY, P., "The Use of Control Variates in Monte Carlo Estimation of Power," *Appl. Statist.*, 31 (1982), 125–129.

SEILA, A. F., "A Batching Approach to Quantile Estimation in Regenerative Simulations," *Management Sci.*, 28 (1982), 573–581.

TONG, Y. L., *Probability Inequalities in Multivariate Distributions*, Academic Press, New York, 1980.

YANG, S., "A Smooth Nonparametric Estimator of a Quantile Function," *J. Amer. Statist. Assoc.*, 80 (1985), 1004–1011.