

Small-sample theory for steady state confidence intervals

Chia-Hon Chien
BNR, Inc.
P.O. Box 7277
Mountain View, CA 94039

ABSTRACT

The purpose of this paper is to develop a nonparametric method for obtaining a confidence interval for the mean of a stationary sequence. As indicated in literature, nonparametric confidence intervals in practice often have undesirable small-sample asymmetry and coverage characteristics. These phenomena are partially due to the fact that the third and fourth cumulants of the point estimator for the stationary mean, unlike those of the standard normal random variable, are not zero. We will apply Edgeworth and Cornish-Fisher expansions theory to obtain asymptotic expansions for the errors associated with confidence intervals. The analysis isolates various elements that contribute to errors and makes it possible for us to estimate each element and hopefully correct the errors to a smaller order. We will use Glynn's method to develop first and second order correction terms for the confidence intervals. These procedures, in the meantime, also improve the asymptotic order of confidence interval accuracy.

1. INTRODUCTION

The purpose of this paper is to develop a nonparametric method for obtaining a confidence interval for the mean of a stationary process with certain regularity conditions. Our starting point is the batch means method. We hope that by using the idea of Johnson-Glynn pivotal transformations we can obtain better confidence intervals for the quantities of interest. We make no assumption that the observed data are sampled from either iid, regenerative, or ARMA processes. We believe this assumption is more general case and robust for the output analysis of the real world simulation experiments. The procedures we purpose do not require the selection of any critical constants that can not be reasonably preset. This fact and the less restrictive nature of the assumptions will also allow us to implement these procedures as a software package for the output analysis of many real world simulation studies.

The basic approach of the procedures we purpose follow from Johnson (1978), Glynn (1982), and Titus (1985). As indicated in Glynn (1982), nonparametric confidence intervals in practice often have undesirable small-sample asymmetry and coverage characteristics. We will apply Edgeworth expansion theory to obtain asymptotic expansions for the errors associated with confidence intervals. The analysis isolates various elements that contribute to errors. We then use Glynn's method to develop first and second order correction terms to the confidence intervals. These procedures, in the meantime, also improve the asymptotic order of confidence interval accuracy in the sense that the actual coverage rate of the corrected confidence interval is closer to the nominal coverage than previous methods.

Johnson (1978) is perhaps the first author to use these procedures. He derives a first order correction for t -statistics for independent and identically distributed samples. Glynn (1982) extends this idea to a second order correction for ratio estima-

tors of regenerative processes. Titus (1985) applies the same idea to asymptotically stationary autoregressive processes of finite orders.

The organization of this paper is as follows. In Section 2, we develop the necessary background for cumulants and Edgeworth expansions, while Cornish-Fisher expansions and some new uniqueness properties are discussed in Section 3. Section 4 describes how the cumulants for a stationary process, with some regularity conditions, can be calculated.

In Section 5, we derive first and second order Johnson-Glynn pivots for the corrections of the errors in confidence interval coverage. Finally, some numerical examples are displayed in Section 6.

2. EDGEWORTH EXPANSION

Consider a random variable X with distribution function F_X . The moments $\mu'_r = EX^r$, $r \geq 0$, and central moments $\mu_r = E(X - EX)^r$, $r \geq 0$, are some very useful constants for measuring the properties of X and, in some cases, uniquely characterize the distribution function F_X (Chung 1974 pp. 98-99). For some statistical analyses, however, there is another series of constants, the cumulants, which are closely related to the moments but would be more convenient from a theoretical point of view.

Let $\Phi_X(t)$ be the characteristic function of X , then subject to existence the cumulants κ_r 's of X are formally defined by (Kendall and Stuart 1977 p. 69)

$$\begin{aligned}\Phi_X(t) &= \int_{-\infty}^{\infty} e^{itx} dF_X \\ &= \sum_{r=0}^{\infty} \mu'_r (it)^r / r! \\ &= \exp \left\{ \sum_{r=1}^{\infty} \kappa_r (it)^r / r! \right\}\end{aligned}$$

Thus κ_r is the coefficient of $(it)^r / r!$ in $\log \Phi_X(t)$ when the expansion in power series exists; $\log \Phi_X(t)$ may be called as *cumulant generating function* (c.g.f.) or the *second characteristic function*. However, the first terminology is somehow misleading in the sense that $\log \Phi_X(t)$ exists (for some neighborhood of 0) even if the moments and cumulants do not exist.

The cumulants that involve two or more random variables are also of interest. Consider random variables X_1, \dots, X_r , such that $E|X_j|^r < \infty$, $j = 1, \dots, r$. The r th order joint cumulant of X_1, \dots, X_r is given by $\text{cum}(X_1, \dots, X_r) = \sum (-1)^{p-1} (p-1)! (E \prod_{j \in \nu_1} X_j) \cdots (E \prod_{j \in \nu_p} X_j)$ where the summation extends over all partitions ν_1, \dots, ν_p , $p = 1, \dots, r$, of the set $\{1, \dots, r\}$ (Billinger 1981 p.19). For convenience we will define

$$\kappa_{ij}(X, Y) = \text{cum}(\underbrace{X, \dots, X}_{i \text{ terms}}, \underbrace{Y, \dots, Y}_{j \text{ terms}}).$$

Suppose F_X , $\Phi_X(t)$, and $\{\kappa_r(X), r \geq 1\}$, and F_Y , $\Phi_Y(t)$, and $\{\kappa_r(Y), r \geq 1\}$ are distribution function, characteristic function, and cumulants of random variables X and Y , respectively. Since $\Phi_X(t) = \exp \left\{ \sum_{r=1}^{\infty} \kappa_r(X) (it)^r / r! \right\}$ and $\Phi_Y(t) = \exp \left\{ \sum_{r=1}^{\infty} \kappa_r(Y) (it)^r / r! \right\}$, we have

$$\Phi_X(t) = \exp \left\{ \sum_{r=1}^{\infty} [\kappa_r(X) - \kappa_r(Y)] (it)^r / r! \right\} \Phi_Y(t).$$

Now if the characteristic function $\Phi_Y(t)$ is absolutely integrable over $(-\infty, \infty)$, then F_Y would be absolutely continuous and

$$f_Y(y) \equiv F'_Y(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\{-ity\} \Phi_Y(t) dt.$$

The density f_Y would be bounded and continuous (Chung 1974 p. 155). It follows that, subject to existence, the characteristic function of $D^r F_Y$ will be $(-it)^r \Phi_Y(t)$, where D denotes the differential operator. The uniqueness of the Fourier transform will now give

$$F_X(x) = \exp \left\{ \sum_{r=1}^{\infty} [\kappa_r(X) - \kappa_r(Y)] \frac{(-D)^r}{r!} \right\} F_Y(x),$$

and similarly

$$f_X(x) = \exp \left\{ \sum_{r=1}^{\infty} [\kappa_r(X) - \kappa_r(Y)] \frac{(-D)^r}{r!} \right\} f_Y(x).$$

The special case that Y is a normal random variable is most important. Let $\alpha(y) = (1/\sqrt{2\pi}) \exp\{-t^2/2\}$. If Y has density function $f_Y(y) = \beta(y) = \alpha((y-m)/\sigma)$ then $\Phi_Y(t) = \exp\{imt - \sigma^2 t^2/2\}$, we have $\kappa_1(Y) = m$, $\kappa_2(Y) = \sigma^2$, and $\kappa_r(Y) = 0$, for each $r \geq 2$. The above equation becomes

$$f_X(x) = \exp \left\{ -\frac{\kappa_1(X) - m}{1!} D + \frac{\kappa_2(X) - \sigma^2}{2!} D^2 - \frac{\kappa_3(X)}{3!} D^3 + \frac{\kappa_4(X)}{4!} D^4 - \dots \right\} \beta(x).$$

This is the Edgeworth expansion of Type A (Kendall and Stuart 1977 p. 170).

Up to this point an underlying assumption is that the functions f_X and F_X possess convergent Type A series. That this is not always the case. For a discussion of the convergence properties see Kendall and Stuart (1977) pp. 173-174 and Cramér (1946) p. 223.

For practical applications, however, it is usually of little value to know the convergence properties of the expansions. What we are really interested in is whether a small number of terms would suffice to give good approximations of the functions f_X and F_X . If this is the case, we would not be too concerned about the convergence properties. On the other hand, if the series actually converges but a satisfactory approximation can only be obtained after a large number of terms have been calculated, then this Type A series would be of very little use.

3. CORNISH-FISHER EXPANSIONS

Suppose that we have a sequence of random variables $\{Y_n\}$

which are asymptotically normal in the sense that there exist sequences $\{\mu_n\}$ and $\{\sigma_n\}$ so that, as n increases, $(Y_n - \mu_n)/\sigma_n \Rightarrow N(0, 1)$. For large n , it is often possible to approximate the distribution of $(Y_n - \mu_n)/\sigma_n$ by a normal, but for small to moderate n this may not be a good approximation. Under some circumstance we are able to use a polynomial transformation such as

$$\xi_n = b_{n,0} + b_{n,1} \left(\frac{Y_n - \mu_n}{\sigma_n} \right) + b_{n,2} \left(\frac{Y_n - \mu_n}{\sigma_n} \right)^2 + \dots,$$

where $b_{n,i}$'s are of order $n^{1/2}$ or smaller. By choosing $b_{n,i}$'s appropriately, we may make the distribution of ξ_n much closer to normality than that of $(Y_n - \mu_n)/\sigma_n$.

It turns out that this problem is embedded in the following larger and more natural question. Suppose Z_n has distribution function F_{Z_n} , and ξ is a $N(0, 1)$ variate. If x and ξ are corresponding quantiles of F_{Z_n} and Φ respectively, namely, $F_{Z_n}(x) = \Phi(\xi)$, solving ξ in terms of x and x in terms of ξ will be of interest.

Specifically, suppose we are given random variables $S_n, n = 1, 2, \dots$, with $ES_n = 0$, and for each $i \geq 2$, the i th cumulant of S_n , $\kappa_i(S_n) = O(n)$ as $n \rightarrow \infty$. We are interested in the standardized random variable $Z_n \equiv S_n/\kappa_2^{1/2}(S_n)$ as $n \rightarrow \infty$.

We use this notation: $\kappa_i \equiv \kappa_i(S_n)$, $l_i \equiv \kappa_i(Z_n) = \kappa_i/\kappa_2^{i/2}$. It is easy to see that $l_1 = 0$, $l_2 = 1$, and for each $i \geq 2$, $l_i = O(n^{1-i/2})$. By taking $m = 0$ and $\sigma^2 = 1$, the Edgeworth expansion of Type A becomes

$$f_{Z_n}(x) = \exp \left\{ -\frac{l_3}{6} D^3 + \frac{l_4}{24} D^4 - \frac{l_5}{120} D^5 + \dots \right\} \alpha(x).$$

If we apply a similar argument to F_X , expand the operator, and retain only the leading terms, we will have (Cramér 1946)

$$F_{Z_n}(x) \approx \Phi(x) - \frac{l_3}{6} H_2(x) \alpha(x) - \left[\frac{l_4}{24} H_3(x) + \frac{l_5^2}{72} H_5(x) \right] \alpha(x) - \left[\frac{l_3^2}{1296} H_8(x) + \frac{l_3 l_4}{144} H_6(x) + \frac{l_5}{120} H_4(x) \right] \alpha(x).$$

We can see that if we only include the first term on the right hand side this is the usual central limit theorem while all the remaining terms, with coefficients of small order of n , represent the error terms in the approximation of central limit theorem.

From Kendall and Stuart (1977) the problems of the relationship between ξ and x have the following formal solution.

$$\xi \approx x - \left[\frac{l_3}{6} (x^2 - 1) \right] + \left[\frac{l_3^2}{36} (4x^3 - 7x) - \frac{l_4}{24} (x^3 - 3x) \right].$$

$$x \approx \xi + \frac{l_3}{6} (\xi^2 - 1) + \left[\frac{l_4}{24} (\xi^3 - 3\xi) - \frac{l_3^2}{36} (2\xi^3 - 5\xi) \right].$$

Notice that for each of these equations the three terms on the right hand side are of orders $O(1)$, $O(n^{-1/2})$, and $O(n^{-1})$, respectively. These equations are the Cornish-Fisher expansions.

We state the following new interpretation of Cornish-Fisher expansions. These properties also uniquely characterize the Cornish-Fisher expansions. We note that similar interpretations of Cornish-Fisher expansions with one more term can also be derived; see Chien (1988) for details.

Proposition 1.

Suppose ξ is a $N(0,1)$ random variable, $g(X_n) = X_n + a_0 + a_1X_n + a_2X_n^2$, where X_n is a random variable such that $\kappa_1(X_n) = l_1 = O(n^{-1/2})$, $\kappa_2(X_n) = 1 + l_2 = 1 + O(n^{-1})$, $\kappa_3(X_n) = l_3 = O(n^{-1/2})$. If $a_i = O(n^{-1/2})$, $i = 0, 1, 2$, then $g(X_n) = X_n - l_1 - 1/6l_3(X_n^2 - 1)$ is the unique (up to the order of $O(n^{-1/2})$) second order polynomial of X_n such that $|\kappa_i(\xi) - \kappa_i(g(X_n))| = O(n^{-1})$ for each i , $1 \leq i \leq 3$.

Proposition 2.

Suppose ξ is a $N(0,1)$ random variable, $h(\xi) = \xi + a_0 + a_1\xi + a_2\xi^2$, and X_n is a random variable such that $\kappa_1(X_n) = l_1 = O(n^{-1/2})$, $\kappa_2(X_n) = 1 + l_2 = 1 + O(n^{-1})$, $\kappa_3(X_n) = l_3 = O(n^{-1/2})$. If $a_i = O(n^{-1/2})$, $i = 0, 1, 2$, then $h(\xi) = \xi + l_1 + 1/6l_3(\xi^2 - 1)$ is the unique (up to the order of $O(n^{-1/2})$) second order polynomial of ξ such that $|\kappa_i(X_n) - \kappa_i(h(\xi))| = O(n^{-1})$ for each i , $1 \leq i \leq 3$.

4. CUMULANTS FOR STATIONARY PROCESSES

Suppose $\{X_n\}$ is a discrete time stationary process with mixing constants $\{\alpha_n\}$ in the sense that $|P(A \cap B) - P(A)P(B)| \leq \alpha_n$ for $A \in \sigma(X_1, \dots, X_k)$, which is the σ -field generated by random variables X_1, \dots, X_k , $B \in \sigma(X_{k+n}, X_{k+n+1}, \dots)$, which is the σ -field generated by random variables $X_{k+n}, X_{k+n+1}, \dots$, for each k and n , $k \geq 1$, $n \geq 1$. If $\alpha_n \rightarrow 0$, then X_k and X_{k+n} are approximately independent for large n .

We now cite a theorem from Titus (1985) p. 16.

Theorem (Titus).

For a stationary sequence $\{X_n\}$ such that (1) for some positive integer j , each mixed moments of the form $EX_{n_1} \dots X_{n_l}$ is bounded for all l , $l \leq 4(j-1)$. (2) the sequence is mixing with $\alpha_n = O(n^{-2(j-1+\epsilon)})$ for some $\epsilon > 0$. Then $\kappa_j(S_n)$, where $S_n \equiv \sum_{i=1}^n X_i$, is of $O(n)$ as $n \rightarrow \infty$.

5. JOHNSON-GLYNN PIVOTS

Suppose $\{X_i, 1 \leq i \leq m\}$ is a discrete time stationary process which satisfies some regularity conditions; e.g., such as those in Section 4, so that $\kappa_i(S_m) = O(m)$ as $m \rightarrow \infty$. We first review the batch means method for obtaining a point estimate and confidence interval for the true mean value EX , where X is a generic X_i .

Let $\bar{X} = (1/m) \sum_{i=1}^m X_i$ be the sample mean, which we shall use as the point estimate. Assume $m = nb$. Here b is the batch size and n is the number of batches. For each i , $1 \leq i \leq n$, let $Y_i = (1/b) \sum_{j=1}^b X_{(i-1)b+j}$, the i th batch mean, and $V_n = (1/n) \sum_{i=1}^n (Y_i - \bar{X})^2$. Then by central limit theorem for stationary processes (Brillinger 1981) we have

$$t_{n,b} = \frac{(\bar{X} - EX)}{(V_n/n)^{1/2}} \Rightarrow N(0,1),$$

so that we can construct a confidence interval for EX .

Define $\sigma^2 = m \cdot \text{var}(\bar{X})$ then $V_n \approx \text{var}(Y_i) \approx \sigma^2/b$ so that $(V_n/(\sigma^2/b) - 1)$ is small and we have

$$\begin{aligned} t_{n,b} &= n^{1/2} \frac{\bar{X} - EX}{V_n^{1/2}} \\ &\approx \frac{m^{1/2}}{\sigma} (\bar{X} - EX) \left\{ 1 - \frac{1}{2} \left(\frac{V_n}{\sigma^2/b} - 1 \right) \right. \\ &\quad \left. + \frac{3}{8} \left(\frac{V_n}{\sigma^2/b} - 1 \right)^2 - \frac{5}{16} \left(\frac{V_n}{\sigma^2/b} - 1 \right)^3 + \dots \right\}. \end{aligned}$$

It is clear that if we could calculate various mixed cumulants of $(\bar{X} - EX)$, $(\bar{X} - EX)(V_n/(\sigma^2/b) - 1)$, $(\bar{X} - EX)(V_n/(\sigma^2/b) - 1)^2$, ..., then we would be able to obtain various cumulants of $t_{n,b}$. The results are summarized as follows.

Proposition 3.

Leading terms of the first four cumulants of $t_{n,b}$ are as follows:

$$\begin{aligned} \kappa_1(t_{n,b}) &\approx -(1/2)(\sqrt{mb}/\sigma^3) \kappa_{1,1}(\bar{Y}, (1/n) \sum Y_i^2) \\ \kappa_2(t_{n,b}) &\approx 1 + (b/\sigma^2)^2 \kappa_{2,1}((1/n) \sum Y_i^2) \\ &\quad - (mb/\sigma^2) \kappa_{2,1}(\bar{Y}, (1/n) \sum Y_i^2) \\ &\quad - E[(1/n) \sum Y_i^2 / (\sigma^2/b) - 1] + 3/n \\ \kappa_3(t_{n,b}) &\approx (\sqrt{m}/\sigma^3) \{ \kappa_3(\bar{Y}) \\ &\quad - 3(\sigma^2/m) \kappa_{1,1}(\bar{Y}, (1/n) \sum Y_i^2 / (\sigma^2/b)) \} \\ \kappa_4(t_{n,b}) &\approx -6(mb/\sigma^4) \kappa_{2,1}(\bar{Y}, (1/n) \sum Y_i^2) \\ &\quad + 3(b/\sigma^2)^2 \kappa_2((1/n) \sum Y_i^2) + 12/n \end{aligned}$$

It can also be shown that we have $\kappa_1(t_{n,b}) = O(m^{-1/2})$, $\kappa_2(t_{n,b}) = 1 + O(n^{-1}) + O(b^{-1})$, $\kappa_3(t_{n,b}) = O(m^{-1/2})$, and $\kappa_4(t_{n,b}) = O(n^{-1})$. We also note that the $O(b^{-1})$ term in $\kappa_2(t_{n,b})$ would disappear should $\{Y_i\}$ be independent.

The first four cumulants of a standard normal random variable are 0, 1, 0, and 0. If we want to use a normal approximation to generate a confidence interval, we would like to have the cumulants of $t_{n,b}$ to be as close to those of the standard normal random variable as possible. This argument leads to two naive, but natural, approaches: (1) Minimizing $\max_{1 \leq i \leq 4} |\kappa_i(t_{n,b}) - \kappa_i(\xi)|$. (2) Minimizing $\sum_{i=1}^4 |\kappa_i(t_{n,b}) - \kappa_i(\xi)|^2$. In both cases we can see that the optimal b should be chosen such that b and n are of the same order. This relationship of the batch size and the number of batches gives us a $t_{n,b}$ which is closer to a standard normal random variable. Hence we feel this is a preferred choice of batch size for the traditional batch means method.

Based on the choice that $b \sim O(m^{1/2})$ we can derive the formal Edgeworth expansion

$$F_i(x) \approx \Phi(x) + \phi(x) \left[(\kappa_3/6 - \kappa_1) + (\kappa_4/8 - (\kappa_2 - 1)/2)x - (\kappa_3/6)x^2 - (\kappa_4/24)x^3 \right]$$

and the formal Cornish-Fisher expansions

$$\xi \sim g(x) \approx x + [(\kappa_3/6 - \kappa_1) + (\kappa_4/8 - (\kappa_2 - 1)/2)x - (\kappa_3/6)x^2 - (\kappa_4/24)x^3]$$

$$x \sim h(\xi) \approx \xi + [(\kappa_1 - \kappa_3/6) + ((\kappa_2 - 1)/2 - (\kappa_4/8))\xi + (\kappa_3/6)\xi^2 + (\kappa_4/24)\xi^3]$$

Again we can have the following new uniqueness properties of the Cornish-Fisher expansions for the batch means method. For the Cornish-Fisher expansions with one more term see Chien (1988) for details.

Proposition 4.

Suppose ξ is a $N(0,1)$ random variable, $g(X_n) = X_n + a_0 + a_1X_n + a_2X_n^2$, where X_n is a random variable such that

$\kappa_1(X_n) = l_1 = O(n^{-1/2})$, $\kappa_2(X_n) = 1 + l_2 = 1 + O(n^{-1/2})$, $\kappa_3(X_n) = l_3 = O(n^{-1/2})$. If $a_i = O(n^{-1/2})$, $i = 0, 1, 2$, then $g(X_n) = X_n - l_1 - 1/2l_2X_n - 1/6l_3(X_n^2 - 1)$ is the unique (up to the order of $O(n^{-1/2})$) second order polynomial of X_n such that $|\kappa_i(\xi) - \kappa_i(g(X_n))| = O(n^{-1})$ for each i , $1 \leq i \leq 3$.

Proposition 5.

Suppose ξ is a $N(0, 1)$ random variable, $h(\xi) = \xi + a_0 + a_1\xi + a_2\xi^2$, where X_n is a random variable such that $\kappa_1(X_n) = l_1 = O(n^{-1/2})$, $\kappa_2(X_n) = 1 + l_2 = 1 + O(n^{-1/2})$, $\kappa_3(X_n) = l_3 = O(n^{-1/2})$. If $a_i = O(n^{-1/2})$, $i = 0, 1, 2$, then $h(\xi) = \xi + l_1 + 1/2l_2\xi + 1/6l_3(\xi^2 - 1)$ is the unique (up to the order of $O(n^{-1/2})$) second order polynomial of ξ such that $|\kappa_i(X_n) - \kappa_i(h(\xi))| = O(n^{-1})$ for each i , $1 \leq i \leq 3$.

We now discuss how to use the Cornish-Fisher expansions to produce the corrected confidence intervals. Following the terminology in Glynn (1982) we define $t = (\bar{X} - EX)/(V_n/n)^{1/2}$ as the 0th order correction (no correction); $T = t + [(\kappa_3/6 - \kappa_1) - (\kappa_2 - 1)/2t - (\kappa_3/6)t^2]$ as the 1st order correction; and $T^* = t + [(\kappa_3/6 - \kappa_1) + (\kappa_4/8 - (\kappa_2 - 1)/2)t - (\kappa_3/6)t^2 - (\kappa_4/24)t^3]$ as the 2nd order correction. Notice that t is the traditional t -statistic associated with the traditional batch means method; T is the unique second order polynomial of t in the sense of Proposition 4; and T^* is the unique third order polynomial of t .

To construct confidence intervals it is more convenient to use the inverted Cornish-Fisher expansions (Hall 1983). For δ -quantile point z_δ of the standard normal distribution function, the confidence intervals for the three corrections are

$$[\bar{X} - h'(z_\delta)(V_n/n)^{1/2}, \bar{X} - h'(-z_\delta)(V_n/n)^{1/2}],$$

where $h'(z) = z$ for t , $h'(z) = z + [(\kappa_1 - \kappa_3/6) + (\kappa_2 - 1)/2z + (\kappa_3/6)z^2]$ for T , and $h'(z) = z + [(\kappa_1 - \kappa_3/6) + ((\kappa_2 - 1)/2 - (\kappa_4/8))z + (\kappa_3/6)z^2 + (\kappa_4/24)z^3]$ for T^* .

6. NUMERICAL EXAMPLES

In this section, we report the results of two Monte Carlo studies of the coverage statistics of "normal quantile" confidence intervals based on the Johnson-Glynn pivots $t_{n,b}$, $T_{n,b}$, and $T^*_{n,b}$.

Example 1.

Let Y_i have an exponential distribution centered at 0. (i.e. $P(Y_i > y) = e^{-(y+1)}$, $y > -1$; 0, otherwise). This example was studied in Efron (1981) and Glynn (1982). See Table 1 for the empirical results.

Example 2.

Let $\{W_i : i \geq 1\}$ be the sequence of consecutive customer waiting times in an M/M/1 queue with arrival rate $\lambda = 0.5$ and service rate $\mu = 1$. This example was studied in Glynn (1982) and Titus (1985). See Tables 2 and 3 for the empirical results.

Let us define I_1, I_2, I_3 as follows: I_1 represents the percentage of replications for which the exact value lies to the left of the confidence interval; I_2 represents the percentage that the exact value lies in the confidence interval; and I_3 represents the percentage that the exact value lies to the right of the confidence interval. Thus I_2 is the usual coverage percentage, and

I_1 and I_3 are the tails.

For each of these examples, we will repeat 2500 independent replications and report empirical coverage percentages I_1, I_2, I_3 , sample mean of the length of the confidence interval (SM), sample standard deviation (SSD), and sample coefficient of variation (SCV), which is the ratio of sample standard deviation over sample mean (SSD/SM) of the 90% confidence interval.

Table 1: Centered-Exp(1) Random Variables

Centered exp(1) r.v. 2500 replications							
Sample Size	Pivot	Coverage Rate			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
m=5	normal	4.08	71.40	24.52	1.12	0.65	0.58
	t	1.88	78.64	19.48	1.45	0.84	0.58
	1st	1.52	82.20	16.28	1.61	1.01	0.63
	2nd	1.60	81.88	16.52	1.58	0.98	0.62
m=10	normal	3.08	79.04	17.88	0.91	0.39	0.42
	t	1.88	82.52	15.60	1.01	0.43	0.42
	1st	2.16	86.60	11.24	1.17	0.58	0.49
	2nd	2.16	86.52	11.32	1.15	0.56	0.48
m=15	normal	2.76	82.76	14.48	0.78	0.27	0.34
	t	1.76	84.88	13.36	0.84	0.29	0.34
	1st	2.64	88.84	8.52	0.96	0.39	0.41
	2nd	2.64	88.80	8.56	0.94	0.38	0.40
m=20	normal	2.40	85.48	12.12	0.69	0.21	0.30
	t	1.92	86.68	11.40	0.73	0.22	0.30
	1st	2.44	89.92	7.64	0.82	0.29	0.36
	2nd	2.64	89.52	7.84	0.81	0.28	0.35
m=25	normal	2.28	86.32	11.44	0.63	0.17	0.27
	t	1.84	87.44	10.72	0.65	0.18	0.27
	1st	2.72	89.84	7.44	0.72	0.24	0.32
	2nd	2.84	89.68	7.48	0.72	0.23	0.32

Table 2: M/M/1 Waiting Times

M/M/1 Waiting Times $\rho = 0.5$ 2500 replications							
Sample Size		m=1000					
Batch Size	Pivot	Coverage Rate			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
b=20	normal	4.56	81.04	14.40	0.45	0.15	0.33
	t	4.12	82.00	13.88	0.46	0.15	0.33
	1st	6.08	84.08	9.84	0.52	0.21	0.40
	2nd	6.44	83.68	9.88	0.51	0.20	0.39
b=25	normal	3.92	82.08	14.00	0.47	0.16	0.35
	t	3.76	82.68	13.56	0.48	0.17	0.35
	1st	5.20	85.68	9.12	0.54	0.28	0.43
	2nd	5.40	85.24	9.34	0.53	0.22	0.42
b=40	normal	3.56	82.72	13.72	0.48	0.19	0.39
	t	2.92	84.00	13.08	0.50	0.19	0.39
	1st	4.00	86.68	9.32	0.57	0.29	0.50
	2nd	4.16	86.48	9.36	0.56	0.28	0.49
b=50	normal	3.72	82.12	14.16	0.48	0.19	0.40
	t	2.92	84.08	13.00	0.51	0.20	0.40
	1st	4.20	86.08	9.72	0.58	0.29	0.50
	2nd	4.40	85.84	9.76	0.57	0.28	0.49
b=100	normal	3.88	81.24	14.88	0.48	0.21	0.44
	t	2.60	84.92	12.48	0.53	0.23	0.44
	1st	2.84	87.64	9.52	0.60	0.32	0.53
	2nd	2.96	87.44	9.60	0.59	0.31	0.52
b=125	normal	4.84	79.80	15.36	0.47	0.22	0.47
	t	3.00	84.32	12.68	0.54	0.25	0.47
	1st	3.12	86.72	10.16	0.61	0.34	0.56
	2nd	3.32	86.36	10.32	0.60	0.33	0.55

Table 3: M/M/1 Waiting Times

M/M/1 Waiting Times $\rho = 0.5$ 2500 replications							
Sample Size		m=2000					
Batch Size	Pivot	Coverage Rate			Length of C.I.		
		I_1	I_2	I_3	SM	SSD	SCV
b=20	normal	4.76	82.24	13.00	0.32	0.082	0.25
	t	4.64	82.56	12.80	0.33	0.083	0.25
	1st	6.56	83.72	9.72	0.36	0.10	0.29
	2nd	6.80	83.48	9.72	0.35	0.10	0.29
b=25	normal	4.00	83.36	12.64	0.34	0.089	0.27
	t	3.92	83.48	12.60	0.34	0.090	0.27
	1st	5.68	85.12	9.20	0.37	0.12	0.32
	2nd	5.80	85.00	9.20	0.37	0.11	0.31
b=40	normal	3.36	84.36	12.28	0.35	0.10	0.30
	t	3.08	84.88	12.04	0.36	0.11	0.30
	1st	4.92	86.52	8.56	0.39	0.15	0.38
	2nd	5.04	86.24	8.72	0.39	0.14	0.37
b=50	normal	3.00	84.84	12.16	0.35	0.11	0.31
	t	2.92	85.16	11.92	0.38	0.11	0.31
	1st	4.48	86.92	8.60	0.40	0.16	0.39
	2nd	4.68	86.60	8.72	0.40	0.15	0.38
b=100	normal	3.24	84.80	11.96	0.36	0.12	0.34
	t	2.52	86.20	11.28	0.38	0.13	0.34
	1st	3.40	87.88	8.72	0.42	0.18	0.42
	2nd	3.80	87.48	8.72	0.41	0.17	0.41
b=125	normal	3.76	83.28	12.96	0.35	0.13	0.36
	t	2.88	85.16	11.96	0.38	0.14	0.36
	1st	3.56	87.20	9.24	0.42	0.19	0.46
	2nd	3.60	87.12	9.28	0.42	0.19	0.44

Notice that the empirical coverage percentages are essentially the sample means of iid binomial random variables with suitable parameter p . For 2500 replications and the fixed 90% confidence interval, $I_1/100$ and $I_3/100$ are the sample means of 2500 iid binomial random variables with p approximately equals 0.05 (5%) which has standard deviation to be about 0.00436 (0.44%). Similarly, for $I_2/100$ the p value of the corresponding iid binomial random variables is approximately 0.90 (90%), thus the standard deviation of the sample mean is 0.006 (0.6%). These are, we feel, acceptable levels of accuracy for such experiment.

From these examples, the first and second order corrections, as to be expected, tend to balance the one-sided coverage probabilities, moving them towards their required values of 0.05 (5%). This confirms the asymptotic corrections induced by Johnson-Glynn pivotal transformations. Note that essentially in every case, there is some improvement in actual coverage percentage from 0th order pivot (the traditional way) to the 1st and 2nd order pivots.

7. DISCUSSION

Johnson-Glynn pivotal transformations have provided a new way of generating confidence intervals. In applying this approach to the batch means method, they appear to behave well empirically and seem to be a robust procedure. However, there are three possible defects from a mathematical viewpoint: more computation time needed, longer confidence intervals on average, and more variable intervals. On the other hand, as shown in Chien (1988), the increase of computing time is relatively small, and the increase of length in confidence interval is asymptotically negligible as n increases. Moreover, due to the fact that many confidence intervals do have undercoverage problem, this increase of length seems to be a necessity rather than a nuisance.

ACKNOWLEDGEMENTS

The author is greatly indebted to Professors Donald L. Iglehart, Peter W. Glynn, and Gerald S. Shedler of Stanford University for their helps and encouragements. He also wishes to thank National Science Foundation (NSF Grant DCR-85-09668) and Army Research Office (ARO Contract DAAG29-84-K-0030) for financial supports of this research.

REFERENCES

- Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. John Wiley & Sons, New York.
- Bhattacharya, R. N., and Ghosh, J. K. (1978). On the validity of formal Edgeworth expansion. *The Annals of Statistics* 6, 434-451.
- Bhattacharya, R. N., and Ranga Rao, R. (1976). *Formal Approximation and Asymptotic Expansions*. Wiley, New York.
- Billingsley, P. (1979). *Probability and Measure*. John Wiley & Sons, New York.
- Brillinger, D. R. (1981). *Time Series Analysis: Data Analysis and Theory*. Holt, Rinehart & Winston, New York.
- Chien, C. (1988). Small-sample theory for steady state confidence intervals. Forthcoming Technical Report, Department of Operations Research, Stanford University, Stanford, California.
- Chung, K. L. (1974). *A Course in Probability Theory*, Second Edition. Academic Press, New York.
- Cornish, E. A., and Fisher, R. A. (1937). Moments and cumulants in the specification of distributions. *Rev. Int. Statist. Inst.* 5, 307-320.
- Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton, New Jersey.
- Glynn, P. W. (1982). Asymptotic theory for nonparametric confidence intervals. Technical Report No. 19, Department of Operations Research, Stanford University, Stanford, California.
- Hall, P. (1983). Inverting an Edgeworth expansion. *The Annals of Statistics* 39, 1264-1273.
- Johnson, N. J. (1978). Modified t -tests and confidence intervals for asymmetrical populations. *Journal of American Statistical Association* 73, 536-544.
- Kendall, M. G., and Stuart, A. (1977). *The Advanced Theory of Statistics*, Fourth Edition, vol. 1. Macmillan, New York.
- Law, A. (1983). Statistical analysis of simulation output data. *Operations Research* 31, 983-1029.
- Law, A., and Kelton, D. (1984). Confidence intervals for steady state simulations: I. A survey of fixed sample size procedures. *Operations Research* 32, 1221-1239.
- Lukacs, E. (1970). *Characteristic Functions*, Second Edition. Hafner, New York.
- Schmeiser, B. (1982). Batch size effects in the analysis of simulation output. *Operations Research* 30, 556-568.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

Titus, B. D. (1985). Modified confidence intervals for the mean of an autoregressive process. Technical Report No. 34, Department of Operations Research, Stanford University, Stanford, California.

Wallace, D. L. (1958) Asymptotic approximations of distribution functions. *The Annals of Mathematical Statistics* 29, 635-654.

Withers, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution. *The Annals of Statistics* 11, 577-587.

AUTHOR'S BIOGRAPHY

Chia-Hon Chien received the B.S. degree in electrical engineering from National Taiwan University in 1982, M.S. degrees in operations research and in statistics from Stanford University in 1986, and the Ph.D. degree in operations research from Stanford University in 1988. He is now a member of scientific staff at Bell-Northern Research at Mountain View, CA. His research interests include simulation, queueing networks, and performance evaluation.

Chia-Hon Chien
BNR, Inc.
P.O. Box 7277
Mountain View, CA 94039