ASYMPTOTIC THEORY FOR NONPARAMETRIC CONFIDENCE INTERVALS.
ASYMPTOTIC THEORY FOR NONPARAMETRIC CONFIDENCE INTERVALS

by

Peter W. Glynn

TECHNICAL REPORT NO. 63

July 1982

Prepared under Contract N00014-76-C-0578 (NR 042-343)*

for the

Office of Naval Research

Approved for public release: distribution unlimited.

Reproduction in Whole or in Part is Permitted for any
Purpose of the United States Government

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

*This research was also partially supported under
National Science Foundation Grant MCS79-09139.
1. Introduction

The problem of assigning nonparametric confidence intervals has recently been the focus of renewed attention. One impetus has been the development of the "bootstrap" method by EFRON (1979) as a general nonparametric statistical tool. BICKEL and FREEDMAN (1981), as well as SINGH (1981), have shown that the bootstrap's distributional approximation is asymptotically valid in a wide variety of circumstances, while EFRON (1981) has studied, in particular, the bootstrap's viability for setting confidence intervals. The recognition that computing power is increasingly available has allowed statisticians to consider confidence interval methods, such as the bootstrap, that are computationally more complex but statistically better behaved than previous techniques. The pivotal transformation of JOHNSON (1978) is another such procedure.

Nonparametric confidence interval methodology has also attracted considerable study in the Monte Carlo simulation literature; see CRANE and LEMOINE (1977), FISHMAN (1978), and LAW and KELTON (1982), for example. The idea is to assign confidence intervals to point estimators obtained from a simulation output sequence, in order to give the simulator an assessment of the estimates' variability.

The simulation applications mentioned above dictate that we analyze the confidence interval problem for ratio estimators. To be
precise, we shall consider the problem of estimating \( r = E(Y_n)/E(r_n) \)
from a sequence of independent and identically distributed (i.i.d.)
random vectors (r.v.'s) \( \{(Y_n, r_n); n \geq 1\} \), where \( E(|Y_n| + |r_n|) < \infty \)
and \( E(r_n) \neq 0 \). Of course, the classical nonparametric situation is
captured as a special case, by setting \( r_n = 1 \).

The organization of this chapter is as follows. In Section 2, we
show that ratio estimators arise naturally in the context of the simu-
lation and/or statistical analysis of ergodic quantities associated
with regenerative stochastic processes. Section 3 discusses the basic
central limit theorem (CLT) on which all the confidence interval
methods to be considered in this chapter will be based. Asymptotic
error analysis of these techniques requires certain tools from the
theory of Edgeworth expansions. In Section 4, results of
BHATTACHARAYA and GHOSH (1978) and GOTZE and HIPP (1978) are
extended to accommodate the generalizations required by the ratio estimator
problem.

In Section 5, we obtain a rigorous Edgeworth expansion for the
ratio estimator pivot statistic. This extends the work of CHUNG
(1946) from the classical case to the ratio problem (the formulas
there contain some errors, however; see WALLACE (1958), p. 642). This
enables us, in Section 6, to analyze the error asymptotics of the
ratio pivot confidence interval, as well as two related intervals. In
particular, we are able to precisely identify the effect of the
Student t-correction (i.e., using Student t-quantiles rather than
normal quantiles in the limit approximation) on coverage.
In Section 7, we extend Johnson's pivotal transformation to ratio pivots, and show that it corrects for the asymmetry effects of order $n^{-1/2}$ ($n$ is the sample size) that occur in the standard pivot. Section 8 presents a second-order pivotal transformation which corrects coverage error in the standard pivot to order $n^{-1}$. It turns out that this second-order pivot is the nonparametric analogue of a transformation suggested by HOTELLING and FRANKEL (1938) to "normalize" Student $t$-variates. Section 9 discusses computational issues and displays results of Monte Carlo sampling experiments in which the coverage characteristics of the pivotal transformations were compared with those of the untransformed pivot.

2. Some Applications of Ratio Estimator Confidence Intervals

The possibility of extending confidence interval methodology from the classical framework to the ratio estimator context has been previously studied in the statistical literature. For example, ROY and POTTHOFF (1958) discuss this problem in the case where $(Y_n, Z_n)$ has a bivariate normal distribution. Their motivation stemmed from applications in which a comparison of $EV_n$ and $EY_n$, in terms of their ratio, is desired. For instance, in evaluating the effect of a treatment, the ratio of the mean of the treated population to the mean of the untreated population is of interest.

More recently, this problem has attracted considerable attention in the simulation community. Consider a measurable regenerative stochastic process $\{X_t; t \geq 0\}$ (see SMITH (1955) for a complete
discussion). Then, there exist random times $T_1 < T_2 < \cdots$ with $T_n + \epsilon$ such that the vectors $((Y_k(f), \tau_k); k \geq 1)$ are i.i.d., where

$$Y_k(f) = \int_{T_k}^{T_{k+1}} f(X_s)ds$$

$$\tau_k = T_{k+1} - T_k$$

for any suitably measurable real-valued function $f$. It can be shown (see [13]) that if $E(Y_n(|f| + \tau_n)) < \infty$, then

$$\int_0^t f(X_s)ds/t + r(f) \equiv EY_n(f)/E\tau_n \quad a.s.$$

Hence, development of confidence intervals for ergodic quantities such as $r(f)$, in the context of regenerative processes, leads naturally to the study of ratio estimators. For a complete discussion of the simulation issues related to regenerative ratio estimates, we refer the reader to IGLEHART (1978), and Chapter 6 of RUBINSTEIN (1981).

Of course, it is clear that the regenerative approach is equally applicable to the nonparametric analysis of statistical data modelled as a regenerative stochastic process (for example, finite state Markov chains).
3. Confidence Intervals for Ratio Estimators

For the remainder of this chapter, we assume that $E(|Y_n| + |\tau_n|) < \infty$, $\tau_n \neq 0$, and $0 < \sigma^2(Z_1) < \infty$, where $Z_k = Y_k - \tau_k$. Also, without loss of generality, we assume that $E\tau_n > 0$ (otherwise, we pass to $(\bar{Y}_n, \bar{\tau}_n)$). For a generic sequence $(\eta_i; i \geq 1)$ of i.i.d. r.v's, we shall use the notation $\bar{\eta}_k = \sum_{i=1}^{k} \eta_i / k$, and $\eta_k = k^{1/2}(\bar{\eta}_k - \mathbb{E}\eta_1)$. The r.v.'s

$$r_n = 1 - \frac{\bar{Y}_n / \bar{\tau}_n}{\tau_n \neq 0}$$

$$v_n = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \tau_i)^2$$

$$t_n = 1 - \frac{n^{1/2}(\bar{r}_n - r)}{v_n^{1/2}}$$

(we interpret a product involving an indicator to be zero if the indicator is zero) play an important role in ratio estimator confidence intervals. To be precise, it is not difficult to show that $r_n + r$ a.s. and that

$$P_n(x) = P(t_n \leq x) + \int_{-\infty}^{x} \phi(u) du = \Phi(x)$$
where \( \phi(u) = (2\pi)^{-1/2} \exp(-u^2/2) \). The CLT (3.1) proves that 

\[ [L_n(p), R_n(p)] \quad (0 \leq p \leq a) \]

is an approximate 100 \((1-a)\%\) confidence interval for \( r \), where

\[
L_n(p) = r_n - v_n(p+1-a) I_{\{t_n > 0\}} - v_n(p) I_{\{t_n < 0\}}
\]

\[
R_n(p) = r_n - v_n(p) I_{\{t_n > 0\}} - v_n(p+1-a) I_{\{t_n < 0\}}
\]

\[
v_n(p) = z(p) v_n^{1/2} (a^{1/2} t_n)
\]

and \( z(p) = \Phi^{-1}(p) \). In order to study the error asymptotics of the above intervals, we introduce the error descriptors

\[
\varepsilon_n^L(p) = \mathbb{P}(r < L_n(p)) - (a-p)
\]

\[
\varepsilon_n^R(p) = \mathbb{P}(r \geq R_n(p)) - p
\]

\[
\varepsilon_n(p) = \mathbb{P}(L_n(p) \leq r < R_n(p)) - (1-a)
\]

The term \( \varepsilon_n(p) \) measures the coverage probability error in the interval \([L_n(p), R_n(p)]\), whereas the terms \( \varepsilon_n^L(p), \varepsilon_n^R(p) \) provide the one-sided coverage errors. The one-sided errors will assist us in evaluating the degree to which the above nonparametric confidence interval captures the asymmetry which is present in parametric confidence intervals (see [9] for a discussion of this point).
In analogy with the classical nonparametric case, two other intervals are natural to consider. Let \( t_n^* \) be the pivot obtained from \( t_n \) by replacing \( v_n \) with \( v_n^* \), where \( v_n^* = ((n-1)/n) v_n \). Intervals \([L_n^*(p), R_n^*(p)]\), with errors \( e_n^{L}(p)*, e_n^{R}(p)*, \) and \( e_n(p)* \), are defined analogously to (3.2) and (3.3).

A second alternative is to use the Student t-distribution with \( k \) degrees of freedom. Let \( z_k(p) \) be the p'th quantile of such a distribution. Then, \( z_k(p) \to z(p) \) as \( k \to \infty \) (PEISER (1943)), and thus, in light of the parametric theory for the bivariate normal case, it is of interest to consider the intervals, \([L'_n(p), R'_n(p)]\), with errors \( e_n^{L}(p)', e_n^{R}(p)', \) and \( e_n(p)' \), constructed by substituting \( z_{n-1}(p) \) for \( z(p) \) in \( v_n(p) \).

Before concluding this section, we observe that if \( E|\gamma_n|^{s+2} < \infty \) for \( s \geq 0 \), Chebyshev's inequality implies that

\[
P(\gamma_n \leq 0) \leq \frac{E(\gamma_n^s)}{(n^{1/2} E\gamma_n^s)^{s/2}}
\]

where \( s \) is an even integer lying in the interval \((s, s+2]\). It is easily verified algebraically that \( E(\gamma_n^s)^{s/2} \) remains bounded, so

\[
P(\gamma_n \leq 0) = o(n^{-s/2}).
\]

Then,

\[
\begin{align*}
e_n^{L}(p) &= -G_n(s(p+1-s)) + o(n^{-s/2}) \\
\gamma_n^{R}(p) &= G_n(s(p)) + o(n^{-s/2}) \\
\gamma_n(p) &= G_n(s(p+1-s)) - G_n(s(p)) + o(n^{-s/2})
\end{align*}
\]
where \( G_n(x) = F_n(x) - \Phi(x) \), provided \( \mathbb{E}|\tau_n|^{s+2} < \infty \). Analogous expressions hold for the errors for the other confidence intervals. Thus, the discussion of confidence interval error leads to study of asymptotic expansions for \( G_n(x) \).

4. **Edgeworth Expansions for Smooth Statistics**

BHATTACHARAYA and GHOSH (1978) have recently shown that the "delta method" for deriving Edgeworth expansions is rigorously correct for a wide class of statistics. To be precise, suppose that \( \{V_n; n \geq 0\} \) is a sequence of i.i.d. \( m \)-dimensional r.v.'s, and let \( f_1, \ldots, f_k \) be real-valued Borel measurable functions on \( \mathbb{R}^m \). Put

\[
U_i = (f_i(V_i), \ldots, f_k(V_i))
\]

\[
\mu = \mathbb{E}U_i
\]

and let \( H_0, H_1, \ldots, H_s \) be real-valued functions on \( \mathbb{R}^k \) such that \( H_i \) is continuously differentiable of order \( s+2 \) on a neighborhood of \( \mu \). The objective is to establish an asymptotic expansion for the distribution of

\[
(4.1) \quad \Lambda_n = a_1^{1/2} \left( \sum_{k=0}^{s} n^{-k/2} (H_k(U_n) - H_k(\mu))c_n \right) + a_n
\]

where
\[a_n = \sum_{k=1}^{s} n^{-k/2} a_{k,s} + o(n^{-s/2})\]

and

\[c_n = 1 + \sum_{k=2}^{s+1} n^{-k/2} c_{k,s} + O(n^{-(s+1)/2})\]

are sequences of deterministic constants. The form of \(n^{-1/2} A_n\) outside a neighborhood of \(\mu\) can be taken as an arbitrary real-valued measurable function of \(\bar{U}_n\) (the constants \(a_n, c_n\) add a flexibility which will be necessary later; see (7.2) and (8.3), for example).

We shall henceforth assume, in our study of \(A_n\), that the covariance matrix \(\Sigma\) associated with \(U\) is non-singular. This can be done, without loss of generality, by replacing \(1, f_1, \ldots, f_k\) by a maximal (in terms of number of elements) collection \(1, f_{i_1}, \ldots, f_{i_p}\) of functions linearly independent as elements of the \(L^2\) space of r.v.'s (see [1], p. 442, for details).

The "delta method" begins by expanding \(H_s(u)\) to \((s+1-i)\) terms, as a Taylor series about \(u = \mu\). This yields a polynomial \(H_s(u)\) of degree \((s+1-i)\) and gives rise to a differential approximation \(A_{s,n}\) of \(A_n\) as follows:

\[A_{s,n} = \sum_{k=0}^{s} n^{-k/2} [H_{s,k}(\bar{U}_n) - H_{s,k}(u)] c_n + a_n\]

This can be re-written in the form

\[A_{s,n}(\hat{U}_n) = \sum_{k=0}^{s} n^{-k/2} f_{s,k}(\hat{U}_n) + o(n^{-s/2})\]
where $P_{s,k}$ is a polynomial of degree $k+1$ and $o(n^{-s/2})$ is independent of $\overline{U_n}$. Assuming that $U_n$ has sufficiently many finite moments, the $j$'th cumulant $\kappa_{j,n}$ of $F_{s,n}$ has the form

\begin{equation}
\kappa_{j,n} = \sum_{k=0}^{s} n^{-k/2} b_{j,k} + o(n^{-s/2})
\end{equation}

Note that $\tilde{\kappa}_{j,n}$ depends exclusively on moments of order less than $s+1$. Let $\sigma^2$ be the variance of $P_{s,0}(\overline{U_n})$, which we henceforth assume to be positive and finite. A reasonable approximation to the characteristic function of $F_{s,n}$ (and hence $A_n$) is therefore

\begin{equation}
\exp(-\sigma^2 t^2/s) \exp(1 + \sum_{j=1}^{s+2} \frac{(it)^j}{j!} (\tilde{\kappa}_{j,n} - b_{j,0}))
\end{equation}

Expanding the second exponential yields the expression

\begin{equation}
\exp(-\sigma^2 t^2/2) \left[ 1 + \sum_{r=1}^{s} n^{-r/2} \kappa_r(it) \right] + o(n^{-s/2})
\end{equation}

where $\kappa_1, \ldots, \kappa_s$ are polynomials not depending on $n$. The Fourier transform $\hat{F}_{s,n}(it)$ corresponds to the signed measure.
\[ (4.7) \quad \phi_{s,n}(v) dv = \left[ 1 + \sum_{r=1}^{s} n^{-r/2} \pi_r (-d/dv) \right] \phi(v/a) \, dv/a \]

which is the formal Edgeworth expansion of the distribution of \( \Lambda_n \).

(4.8) **THEOREM 1** Suppose that \( E|U_n|^3 < \infty \). Then,

\[ (4.9) \quad P(\Lambda_n \leq x) = \int_{-\infty}^{x} \phi(v/a) \, dv/a + o(n^{-1/2}) \]

where \( o(n^{-1/2}) \) is uniform in \( x \).

ii) Suppose that \( E|U_i|^{s+2} < \infty \), and that \( U_1 + \cdots + U_i \) has a non-zero Lebesgue density component (in \( R^k \)) for some \( i \). Then,

\[ (4.10) \quad P(\Lambda_n \in B) = \int_{B} \phi_{s,n}(v) dv + o(n^{-s/2}) \]

where \( o(n^{-s/2}) \) is uniform over all Borel sets \( B \). The function \( \phi_{s,n} \) can be calculated via the "delta method" (4.2) through (4.7).

Although the proof given in [1] restricts its attention to the case where \( a_n = 0, c_n = 1 \), and \( H_k = 0 \) for \( k > 1 \), the argument readily extends to the more general situation considered here, the only complication being additional notational complexity.

In some related work, CHIBISOV (1972) proved Theorem 4.8 (ii) in the case where \( T_n \) was of the polynomial form (4.3) (no identification of \( \phi_{s,n} \) as the expansion obtained via the "delta method" was made, however). An extension to the general non-polynomial case was effected via the following "perturbation" theorem (see [4], p. 629).
(4.11) THEOREM. Suppose that \( \Lambda_n^2 = \Lambda_n + n^{-1} \chi_n \) where \( \Lambda_n \) satisfies the assumptions of Theorem 4.8(ii) and \( \mathbb{P}(|\chi_n| > n^{1/2} \rho_n) = o(n^{-s/2}) \) for some sequence \( \rho_n \to 0 \). Then,

\[
\mathbb{P}(\Lambda_n^2 \leq x) = \int_{-\infty}^{x} \phi_{s,n}(v) dv + o(n^{-s/2})
\]

where \( o(n^{-s/2}) \) is uniform in \( x \).

It is clear that the \( \phi_{s,n} \) of Theorem 4.11 must be that obtained via the "delta method".

We remark that the density assumption on the \( U_i \)'s in Theorem 4.8(ii) follows if \( V_i \) has a Lebesgue density component which is positive on an open set where \( 1, f_1, \ldots, f_k \) are linearly independent as continuous functions (see Lemma 2.2, [11]). Also, we note that the moment conditions in Theorem 4.8 are norm independent, due to the fact that all norms on finite-dimensional spaces are equivalent.

Because of the potentially large number of derivatives required in calculating the expansions \( E_{s,k}(u) \), it is convenient to consider a modification of the "delta method". Towards this end, let \( \theta(j;u) \) be polynomials of degree \( p(s+1) \), and set

\[
\theta_n = \sum_{j=0}^{m} n^{-j/2} \theta(j; \hat{U}_n) .
\]
(4.12) PROPOSITION. Suppose that \( a_n = \lambda_{s,a}^P - \theta_0 \) satisfies \( n^{s/2} R_n \to 0 \) in probability (i.e., \( R_n = o_p(n^{-s/2}) \)). Then, if \( R_n a_n (s+1)p < \), we have \( R_n^P a_n = R_0^P + o(n^{-s/2}) \).

Proof. The remainder \( R_n \) can be written in the form

\[
R_n = \sum_{j=0}^{a} n^{-j/2} R(j; \hat{u}_n)
\]

where \( R(j;u) \) are polynomials in \( u \) of degree \( p(s+1) \). We now show that \( R(j;u) \) vanishes for \( j \leq s \). Starting with \( j = 0 \), observe that

\[
R(0; \hat{u}_n) \to R(0; N)
\]

(\( \to \) denotes weak convergence), where \( N \) is a multivariate normal r.v. with non-singular covariance matrix \( \Sigma \). Evidently, since \( n^{s/2} R_n = o_p(1) \) and \( R_n \to R(0; \hat{u}_n) = o_p(1) \), it must be that \( R(0; N) \) is degenerate at \( 0 \). On the other hand, if \( R(0;u) \) depends non-trivially on \( u_k \) (say), then the Jacobian of the transformation \( u \to (u_1, ..., u_{k-1}, R(0;u), u_{k+1}, ..., u_k) \) is non-singular, and thus it follows that \( R(0; N) \) has a Lebesgue density. This contradiction forces \( R(0; u) \) to vanish identically. Repeating the argument \( s \) more times proves that

\[
R_n = \sum_{j=s+1}^{a} n^{-j/2} R(j; \hat{u}_n)
\]
Under the moment conditions given here, $E_R(j; \hat{U}_n) + E_R(j; U)$ and consequently $E_R = o(n^{-1})$, proving our result.

Our final goal in this section is to show that the Edgeworth expansion (4.5) remains valid, in a certain sense, when the density assumption on the distribution of $U$ is dropped. Let $C_b^r(R)$ be the class of all bounded infinitely differentiable functions and take $C_b^r(R) = \{ f : D^n f \in C_b^r(R), \text{for all } n \}$ ($D = d/dx$). The class $C_b^r(R)$ includes the trigonometric functions $\sin(tx), \cos(tx)$, as well as the Schwartz class $\mathcal{S}$ (see BHATTACHARAYA and RAO (1976), p. 257). We first need the following proposition.

(4.13) **Proposition.** (i) Let $f \in C_b^r(R)$, and suppose $\alpha$ is a multiindex (i.e., a non-negative integral vector) with $|\alpha| = \sum_1^n \alpha_i \leq m+2$. Then, if $EU_{\alpha}^{m+2} < \infty$, there exists a multivariate Edgeworth expansion $\xi_{n,\alpha}$ of the distribution $U$ such that

(4.14) $E(\hat{U}_u^2 f(p; \hat{U}_n)) = \int u^\alpha f(p-u) \xi_{n,\alpha}(u) du + o(n^{-1/2})$

holds, for any vector $p$.

(ii) If $EU_{\alpha}^{m+2} < \infty$, then the change-of-variables formula

(4.15) $\int f(A_{n,\alpha}(u)) \xi_{n,\alpha}(u) du = \int f(y) \psi_{n,\alpha}(y) dy + o(n^{-1/2})$

holds for all bounded measurable $f$. 
Proof. (i) We use Theorem 3.6 of [14], and observe that boundedness of the derivatives of $f$ implies that

$$D^\beta(u^z f) = O(nu^{m+2})$$

where $D^\beta = D_1^{\beta_1} \cdots D_k^{\beta_k}$, $D_j = \partial/\partial x_j$. For all multiindices $\beta$ with $|\beta| \leq m$. This is sufficient for (i), in the presence of $\int |u|^m du < \infty$.

(ii) Lemma 2.1 of [1] proves that (4.15) holds uniformly over all indicator functions $f$. For an arbitrary bounded $f$, approximate $f$ by a finite linear combination of indicators $f_j,n$ such that

$$|f(x) - \sum d_{j,n} f_j,n(x)| < 2^{-n} \quad \text{for all } x.$$

Then, letting $\Phi(f)$ be the difference between the two integrals in (4.15), and using the Hahn decomposition on the signed measure $\Phi(\cdot)$, shows that

$$|\Phi(f)| \leq 2^{-n} (\int |u| \delta_{m,n}(u) \, du + \int |y| \delta_{m,n}(y) \, dy) + |\sum d_{j,n} \Phi(f_j,n)|$$

$$\leq O(2^{-n}) + \sup |f(x)| \cdot o(a^{-s/2}),$$

where the uniformity over indicators is used in the final step.
Our next theorem shows that Theorem 4.8 continues to hold, in expectation, when the density assumption on \( U \) is deleted. We impose rather strong assumptions on \( U \), and the class of test functions \( f \) allowed, in order to simplify the exposition.

(4.16) **Theorem.** Suppose that \( U \) has finite moments of all orders. Then, for any \( s \),

\[
E\left(\Lambda_n - \frac{1}{2} \int f(y) \psi_{s,n}(y) \, dy + o(n^{-s/2})\right)
\]

for all \( f \in C^s_0(\mathbb{R}) \). The function \( \psi_{s,n} \) can be identified through the "delta method."

**Proof.** First, observe that for any \( \epsilon > 0 \), there exists \( K > 0 \) such that

\[
P\{\left| U_n - \mu \right| > \epsilon\} \leq P\{\left| U_n \right| > K \} + o(n^{-s/2})
\]

The probability on the right-hand side of (4.18) is \( o(n^{-s/2}) \) (see Corollary 17.12 of [2]); and hence

\[
E\left(\Lambda_n - \frac{1}{2} \int f(y) \psi_{s,n}(y) \, dy + o(n^{-s/2})\right)
\]

Choose \( \epsilon \) sufficiently small that \( D^{(s+2-k)} H_k(u) \) is continuous on an \( \epsilon \)-neighborhood of \( \mu \) for all \( k \). Expanding \( \Lambda_n \) on \( \{\left| U_n - \mu \right| < \epsilon\} \) yields
(4.19) \[ \Lambda_n = \Lambda_{s,n} + n^{-(s+1)/2} \sum_{k=0}^{s} c_n (\hat{U}_n - v)^{s+2-k} H_k(\eta_{k,n})/(s+2-k)! \]

where \( V = (D_1, \ldots, D_k) \) and \( |\eta_{k,n} - \mu| < \varepsilon \). Note that \( H_k(\eta_{k,n}) \) are bounded r.v.'s in (4.19).

Thus, we can write \( f(T_n) \) on \( \{ |\hat{U}_n - \mu| < \varepsilon \} \) as

\[
f(\Lambda_n) = \sum_{k=0}^{s} (\Lambda_n - P_{o,n}(\hat{U}_n))^k (D^k f) (P_{o,n}(\hat{U}_n))/k! + (\Lambda_n - P_{o,n}(\hat{U}_n))^{s+1} (D^{s+1} f)(\eta_n)/(s+1)!
\]

which evidently can be re-written as

\[
f(\Lambda_n) = \sum_{j=0}^{s} n^{-j/2} \Gamma_{j,n}(\hat{U}_n; \eta_{k,n}; \eta_n).
\]

For \( j \leq s \), \( \Gamma_{j,n} \) has the form \( \hat{U}_n^g g(p^* \hat{U}_n) \) \( g \in C_0^\infty(\mathbb{R}) \) whereas for \( j > s \), \( \Gamma_{j,n} \) is the product of functions of this form with bounded functions. Thus, for \( j \leq s \),

\[
E(\Gamma_{j,n}; \{ |\hat{U}_n - \mu| < \varepsilon \}) = E(\Gamma_{j,n} + o(n^{-s/2})) = \int \Gamma_{j,n}(u) \xi_{s,n}(u) du + o(n^{-s/2}),
\]

the first equality by uniform integrability of \( \{ \Gamma_{j,n} \} \), the second by Proposition 4.13(1). A similar argument for \( j > s \) shows that
and hence

$$E\left( \sum_{j=s+1}^{g} n^{-j/2} \Gamma_{j,n}(u) \xi_{s,n}(u) du \right) + o(n^{-1/2})$$

Applying Proposition 4.13(ii) completes the proof of (4.17). The identification of $\phi_{s,n}(y)$ as that derived from the "delta method" follows from a proof identical to that found on pages 445-6 of [1].

We remark that an immediate consequence of Theorem 4.16 is that, under the assumptions stated, the characteristic function of $\Lambda_n$ can be expanded as

$$E \exp(it\Lambda_n) = \phi_{s,n}(it) + o(n^{-1/2})$$

5. **Edgeworth Expansions for Ratio Estimator Pivots**

As Section 4 illustrates, the key to obtaining Edgeworth expansions is the calculation of cumulants (see (4.4)) of the differential approximation $\Lambda_{s,n}$. The required moments will be derived from Proposition 4.12. In this section, we will calculate Edgeworth expansions for $t_n$ and $t_n^*$ to order $n^{-1}$. This represents a
different approach from that of GEARY (1947) and GAYEN (1949), who formally expanded the distribution of the pivot $t_n$ (for the classical case where $\tau_n = 1$) in Charlier-type series.

We start by observing that the pivotal quantities $t_n$ and $t^*_n$ are invariant to the transformation $(Y, \tau^*_i) \rightarrow (\alpha Y, \alpha \tau^*_1)$ for $\alpha \neq 0$. In particular, by taking $\alpha = 1/\sigma(Z)$, we can assume throughout our calculations, via a passage to $(Y_1, \tau^*_1)$, that $\sigma(Z) = 1$. However, in stating our final conclusions, dependence on $\sigma(Z)$ will be made explicit.

Our first order of business is to expand $t_n$ and $t^*_n$ in differential-type approximations $t_{2,n}$ and $t^*_{2,n}$. Set $V_k = (Y_k, \tau_k)$ and let $f_1(v) = v_1$, $f_2(v) = v_2$, $f_3(v) = v_3^2$, $f_4(v) = v_1 v_2$, $f_5(v) = v_2^2$. Observe that $\tau_n - r = Z_n/\tau_n$ and that $I(v_n > 0, \tau_n = 0)$ is identically 1 on a neighborhood of $\mu$. Thus,

\[
(5.1) \quad t^*_n = \frac{Z_n}{\tau_n} \left( 1 - \frac{1}{2} (v^*_n - 1) + \frac{3}{4} (v^*_n - 1)^2 + o(n^{-1}) \right)
\]

where

\[
(5.2) \quad v^*_n - 1 = \frac{Z_n}{\tau_n} - 2 \frac{Y_1 \tau_1}{\tau_n} \sum_{i=1}^{n} \frac{\tau_i}{\tau_n} + \frac{Z_n}{\tau_n} (\tau_n + r) \sum_{i=1}^{n} \frac{\tau_i}{\tau_n} \frac{\tau_i^2}{\tau_n}
\]

and $W_i = Z_i^2 - 1$. Expanding $1/\tau_n$ in a Taylor series about $1/\Sigma Y$, we obtain

19
where we use $O_p(n^{-k/2})$ to represent a sequence of r.v.'s $X_n$ such that $n^{k/2} X_n$ remains bounded in probability (see p. 8 of Serfling (1980)). Also,

\begin{align*}
(5.4) \quad 2 \sum_{i=1}^{n} Y_i Y_i/n - (\bar{Y}_n)^2/n = 2 \sum_{i=1}^{n} Z_i Z_i/n - (\bar{Z}_n)^2/n.
\end{align*}

Relations (5.1) through (5.4) imply that

\begin{align*}
(5.5) \quad t^*_n &= 2n \hat{\omega}_n/\hat{\omega}_n^2 + \hat{Z}_n/n^{1/2} \\
&- Z_n Q/n + 3(\hat{\omega}_n - 2\alpha \hat{Z}_n)^2/n) + o_p(n^{-1})
\end{align*}

where

\begin{align*}
Q_n &= 2n \tau_1/\hat{\omega}_n - 2Z_1 \tau_1/(n \hat{\omega}_n) + Z_1 \delta
\end{align*}

and $\alpha = E\tau/(\sigma(Z)\hat{\omega}_n)$ and $\delta = E\tau^2/(E\hat{\omega}_n)^2$. In order to evaluate the cumulants required, some moment identities are needed.
(5.6) THEOREM (1) Let \( \{A_n; n \geq 0\} \) be a sequence of i.i.d. r.v.'s with \( E A_n^k < \infty \) for all \( k \). Then,

(a) \( E A_n^3 = n^{-1/2} E A_1^3 \)

(b) \( E A_n^4 = 3(E A_1^2)^2 + n^{-1/2} E A_1^4 - 2n^{-1/2}(E A_1^2)^2 \)

(c) \( E A_n^5 = 10 n^{-1/2}(E A_1^3)(E A_1^2) + O(n^{-1}) \)

(d) \( E A_n^6 = 15(E A_1^3)^2 + O(n^{-1/2}) \)

(e) \( E A_n^{2k+1} = O(n^{-1/2}) \).

(ii) Suppose that \( \{(A_n, B_n); n \geq 1\} \) is a sequence of i.i.d. r.v.'s with \( E A_n < \infty, E B_n < \infty \) for all \( k \). Then,

(f) \( E A_n^{2k} = n^{-1/2} E A_1^{2k} \)

(g) \( E A_n^{2k-1} = 3(E A_1^{2k})(E A_1^2) + O(n^{-1/2}) \)

(h) \( E A_n^{2k+1} = 4 n^{-1/2}(E A_1^3) E A_1^{2k} + 6n^{-1/2}(E A_1^{2k})(E A_1^2) E A_1^2 + O(n^{-1}) \)
\[ \begin{align*}
(i) \quad E_{n}^{k+j} &= 3(E_{1}^{2})(E_{1}^{2})^{2} + 12(E_{1}^{2}E_{1}^{2})^{2} E_{1}^{2} + O(n^{-1/2}) \\
(j) \quad E_{n}^{k+j} &= 15(E_{1}^{2}E_{1}^{2}) (E_{1}^{2})^{2} + O(n^{-1/2}) \\
(k) \quad E_{n}^{k+j} &= O(n^{-1/2}) \quad \text{if} \ k+j \ \text{is odd.}
\end{align*} \]

Proof. For (i), observe that exchangeability of the sequence 
\{A_n: n \geq 0\} provides a recursion in \( k \), namely

\[
E\left( \sum_{i=1}^{n} A_{1}^{i} \right)^{k+1} = E\left( \sum_{i=1}^{n} A_{1}^{i} \right) \left( \sum_{i=1}^{n} A_{1}^{i} \right)^{k}
\]

\[
= nE(A_{1}^{i} \left( \sum_{i=1}^{n} A_{1}^{i} \right)^{k})
\]

\[
= n \sum_{i=1}^{k} \left( \begin{array}{c} k \\ i \end{array} \right) E(A_{1}^{i}) E\left( \sum_{j=1}^{n} A_{1}^{j} \right)^{k-1}
\]

where \( A_{1}^{i} = A_{1}^{i} - EA_{1}^{i} \). Solving the recursion with initial condition 
\( EA_{1}^{i} = 0 \) proves (i). For (ii), apply (i) to \( A_{n}(s,t) = sA_{n} + tB_{n} \). Both sides of equations (a) through (e) are then polynomials in 
\( (s,t) \). Identifying coefficients yields results (f) through (k).
Now, let \( \beta = \frac{E[Z_1^3]}{\sigma^3(Z_1)}, \lambda = (E[Z_1^4]/\sigma^4(Z_1)) - 3, \) and \( \gamma = \frac{E[Z_1^2]}{\sigma^2(Z_1)E[z_1]}. \) Then,

\[
\begin{align*}
E[Z_1^2 \hat{W}_1] &= \beta, \\
E[Z_1^2 \hat{W}_1^2] &= \lambda + 2
\end{align*}
\]

(5.7)

\[
\begin{align*}
E[Z_1^2 \hat{Q}_1] &= \delta + 2\sigma^2 - 2\gamma, \\
E[W_1^2] &= \lambda + 2
\end{align*}
\]

In view of Proposition 4.12, (5.5) through (5.7) provide the asymptotic expansion

\[
E[t^*_{z,n}] = (\alpha - \beta/2)n^{-1/2} + o(n^{-1})
\]

Similar reasoning on the higher moments proves that

\[
E(t^*_z)^2 = 1 + (6\alpha^2 + 6\gamma + 2\beta^2 - 3\delta - 10\alpha\beta)n^{-1} + o(n^{-1})
\]

\[
E(t^*_z)^3 = (9\alpha - 7\beta/2)n^{-1/2} + o(n^{-1})
\]

\[
E(t^*_z)^4 = 3 + (120 \alpha^2 + 60 \gamma + 28 \beta^2 - 30 \delta - 140 \alpha \beta - 2\lambda - 6)n^{-1}
\]

+ o(n^{-1}).

The relevant expansions for \( t_n \) may be easily obtained by using the relation \( t_n = t^*_n(1 - 1/2n + o(n^{-2})). \) Consequently,
\[ E_{t_2, n} = E(t_2^*) \]
\[ E_{t_2, n} = E(t_2^*)^2 - n^{-1} \]
\[ E_{t_2, n} = E(t_2^*)^3 \]
and
\[ E_{t_2, n} = E(t_2^*)^4 - 6n^{-1} \]

up to terms of order \( o(n^{-1}) \).

Note that in the classical case where \( \tau_n = 1 \), we have \( \alpha = 0 \)
and \( \gamma = \delta = 1 \). The moment formulas for \( t_n \), when appropriately
simplified, are then in agreement with those found in [11] and [12].
It should also be noted that in the classical case, the approximate
skewness \( E(t_2, n)^3 \) (of \( t_n(t^*_n) \)) is \( -7\beta/2 \). This verifies
the empirical observation that positive skewness in the distribution
of \( Z_t \) leads to negative skewness in the pivots \( t_n \) and \( t^*_n \) (see
SOPHISTER (1928), and NEYMAN and PEARSON (1928)).

The appropriate cumulants \( \kappa_{j,n} \) of \( t_n \) are given by

(5.8) \[ \kappa_{j,n} = \sum_{k=0}^{2} n^{-k/2} b_{j,k} + o(n^{-1}) \]

where

\[ b_{1,1} = \alpha - \beta/2 \]
\[ b_{2,0} = 1 \]
\[ b_{2,2} = 6\gamma + 5\alpha^2 + 7\beta^2/4 - 3\delta - 9\alpha\beta - 1 \]
\[ b_{3,1} = 6\alpha - 2\beta \]
\[ b_{4,2} = 24\gamma + 60\alpha^2 + 12\beta^2 - 12\delta - 60\alpha\beta - 2\lambda - 6 \]
and all other $b_{j,k}$ are zero. For $t_n^*$, the $b_{j,k}^*$ of the corresponding cumulants $\kappa_{j,n}^*$ satisfy $b_{j,k}^* = b_{j,k}$, excepting that $b_{2,2}^* = b_{2,2} + 1$.

The following distribution function $y_{2,n}(x)$ is obtained from $\kappa_{j,n}$ in the same way as passage was made from (4.4) to (4.7):

\[(5.9) \quad y_{2,n}(x) = \phi(x) - b_{1,1} \phi(x)/n^{1/2} - (b_{2,2} + b_{1,1}^2)x \phi(x)/2n + b_{3,1}(1-x^2) \phi(x)/6n^{1/2} - (b_{4,2} + 4b_{1,1} b_{3,1}) (x^3 - 3x) \phi(x)/24n - b_{3,1}^2 (x^5 - 10x^3 + 15x) \phi(x)/72n\]

Also, let $y_{2,n}^*(x)$ be the function obtained from (5.9) by substituting $b_{j,k}^*$ in place of $b_{j,k}$.

\[(5.10) \text{THEOREM (1) If } E(r_n^6 + \tau_n^6) < \infty, \text{ then} \]

$$F(t_n \leq x) = \Phi(x) + O(n^{-1/2})$$

where $O(n^{-1/2})$ is uniform in $x$. 

25
(ii) Suppose that \( E(|u|^k + |v|^k) < \infty \) for all \( k \). Then,

\[
E(f(t_n)) = \int f(y) Y_{2,n}(dy) + o(n^{-1})
\]

for all \( f \in C_0^\infty (\mathbb{R}) \).

(iii) Suppose \( E(Y_n^3 + v_n^3) < \infty \), and that \((Y_1, v_1)\) has a distribution with a Lebesgue density component which is positive on some open set in the plane. Then,

\[
P(t_n \in B) = \int_B Y_{2,n}(dy) + o(n^{-1})
\]

where \( o(n^{-1}) \) is uniform over all Borel sets \( B \).

(iv) Suppose \( \tau_n \equiv 1 \). Then, if \( EY_n^3 < \infty \), and if the distribution of \( Y_n \) has a Lebesgue density component which is positive on some interval, the analogue of (iii) above holds. The function \( Y_{2,n} \) is obtained from (5.9) by formal substitution.

(v) Results (i) to (iv) are valid for \( t_n^* \) under the same assumptions as for \( t_n \), provided that \( Y_{2,n}^* \) is substituted in place of \( Y_{2,n} \).

Proof. The functions \( f_1, \ldots, f_5 \), being distinct polynomials, are linearly independent so Theorems 4.8 and 4.16 can be applied,
yielding (i) to (iii). Part (iv) is handled as a special case, by
setting \( V_1 = Y_1 \), and \( f_1(v) = v, \ f_2(v) = v^2 \), and applying the same
argument as for (iii).

As previously mentioned, a particularly important application of
ratio estimation lies in the domain of ergodic analysis of
regenerative stochastic processes. It frequently occurs that the
regenerative sequence \( (Y_i, \tau_i); \ i \geq 1 \) constructed is such that \( Y_1 \)
has a Lebesgue density component, whereas \( \tau_1 \) is a lattice r.v.
For example, this is the case that arises when \( \{X_t; \ t \geq 0\} \) is a
continuous time process constructed from a discrete time regenerative
process \( \{X_n\} \) via the formula

\[
X_t = \sum_{n=1}^{\infty} X_n I_{\{n \leq s < n+1\}}(t) .
\]

Our next result addresses this class of processes.

(5.11) THEOREM. Suppose that \( \mathbb{E}(Y_n^6 + \tau_n^6) < \infty \) and that \( Z_1 \) has a
distribution with a Lebesgue density component which is positive on an
interval. Then,

\[
P(t_n \leq x) = Y_{1,n}(x) + o(n^{-1/2})
\]

\[
P(t_n^6 \leq x) = Y_{1,n}(x) + o(n^{-1/2})
\]
uniformly in \( x \), where \( \hat{v}_{1,n}(x) \) is obtained from \( \hat{v}_{2,n}(x) \) by deleting terms with coefficient \( n^{-1} \).

**Proof.** The pivot \( t_n \) can be expanded as

\[
(5.12) \quad t_n = I_{\{v_n > 0, \tau_n \neq 0\}} \left( \hat{v}_n - \frac{\hat{r}_n \hat{z}_n}{2n^{1/2}} + \frac{\hat{a}_n^2}{n^{1/2}} \right) (1 + O(\frac{1}{n^2})) + \chi_n / 6n
\]

where \( O(1/n) \) is deterministic and

\[
\chi_n = I_{\{v_n > 0, \tau_n \neq 0\}} (\hat{u}_n \cdot \hat{y})^3 H(\xi_n)
\]

and \( H \) corresponds to \( t_n \) via (4.1). Now, observe that Theorem 4.8(ii), with \( s = 1 \), is applicable to the first term in (5.12).

Select \( \epsilon \) small enough so that \( (D^\tau H)(u) \) is bounded for

\( \|u - \mu\| \leq \epsilon \) for all multiindices \( \tau \) with \( |\tau| = 3 \). Let

\[
\rho_n = K(\ln n/n)^{1/2} \quad \text{for } K \text{ to be chosen later. Then,}
\]

\[
P(\|X_n\| > \rho_n n^{1/2}) \leq P(\|X_n\| > \rho_n n^{1/2}; \|\tilde{U}_n - \mu\| \leq \epsilon) + P(\|\tilde{U}_n - \mu\| > \epsilon)
\]

\[
\leq P(\|X_n\| > \rho_n n^{1/2}; \|\tilde{U}_n - \mu\| \leq \epsilon) + K \ln n.
\]

Choose \( K \) sufficiently small so that (4.18) applies. Thus,
\[ P(|X_n| > r_n n^{1/2}) = o(n^{-1/2}) , \]

and hence Theorem 4.11 implies the result for \( t_n \). Precisely the same argument works for \( t_n' \).

6. Applications to Ratio Estimator Confidence Interval Estimation

In this section, we apply the Edgeworth expansions of Section 5 to analysis of nonparametric ratio estimator confidence intervals.

(6.1) THEOREM (i) Suppose \( E(Y_n^6 + \zeta_n^6) < \infty \). Then, \( \epsilon_n^f(p) \), \( \epsilon_n^r(p) \), and \( \epsilon_n(p) \) are all \( O(n^{-1/2}) \), uniformly in \( p \).

(ii) Under the assumptions of Theorem 5.11,

\[ \epsilon_n^f(p) = p + 1 - s - \Psi_{1,n}(z(p+1-s)) + o(n^{-1/2}) \]
\[ \epsilon_n^r(p) = \Psi_{1,n}(z(p)) - p + o(n^{-1/2}) \]
\[ \epsilon_n(p) = \Psi_{1,n}(z(p+1-s)) - \Psi_{1,n}(z(p)) - (1-s) + o(n^{-1/2}) \]

uniformly in \( p \).

(iii) Under either assumptions (iii) or (iv) of Theorem 5.10,

\[ \epsilon_n^f(p) = p + 1 - s - \Psi_{2,n}(z(p+1-s)) + o(n^{-1}) \]
\[ e_n^F(p) = \nu_{2,n}(z(p)) - p + o(n^{-1}) \]
\[ e_n(p) = \nu_{2,n}(z(p+1-c)) - \nu_{2,n}(z(p)) - (1-a) + o(n^{-1}) \]

uniformly in \( p \).

(iv) Results (i) to (iii) are valid for \( e_n(p)^*, e_n^F(p)^* \), and 
\( e_n(p)^* \) under the same assumptions as for the \( e_n \) errors, provided 
\( \nu_{2,n}^* \) is substituted in place of \( \nu_{2,n} \).

Proof. The results follow immediately from (3.5), and Theorems 5.10 
and 5.11. 1

These expressions show that under reasonable assumptions the 
coverage errors \( e_n(p), e_n(p)^* \) are \( O(n^{-1/2}) \) for \( p \neq a/2 \), whereas 
for \( p = a/2 \), the coverage errors are \( O(n^{-1}) \). Thus, using confidence 
intervals based on \( p = a/2 \) leads to intervals that are asymptotically 
optimal in the sense of having shortest possible length and most 
accurate coverage rate. However, it is important to realise that the 
one-sided coverage errors are \( O(n^{-1/2}) \) for all \( p \), including 
\( p = a/2 \). Hence it must be that \( [L_n(a/2), R_n(a/2)] \) (similarly for 
\( [L_n^*(a/2), R_n^*(a/2)] \)) achieves coverage error of \( O(n^{-1}) \) via 
cancellation of one-sided errors of order \( O(n^{-1/2}) \). This suggests 
that a "corrected interval", in the sense of one-sided error, can be 
obtained by shifting the interval slightly. This is in agreement with 
parametric confidence interval theory, where intervals tend to be
asymmetric about the point estimate. We shall examine this question further in Section 7.

The coverage errors $\varepsilon_n(a/2)$ and $\varepsilon_n(a/2)^*$ are given by

$$\varepsilon_n(a/2) = -(b_{2,2} + b_{1,1}^2)x_{\alpha} \phi(x_{\alpha})/n$$

$$-(b_{4,2} + 4b_{1,1} b_{3,1}) (x_{\alpha}^3 - 3x_{\alpha}) \phi(x_{\alpha})/12n$$

$$-b_{3,1}^2 (x_{\alpha}^5 - 10x_{\alpha}^3 + 15x_{\alpha}) \phi(x_{\alpha})/36n$$

$$\varepsilon_n(a/2)^* = \varepsilon_n(a/2) - x_{\alpha} \phi(x_{\alpha})/n$$

where $x_{\alpha} = z(1 - \alpha/2)$. Recalling the definition of the $b_{j,k}$'s (see (5.7)) we see that $\varepsilon_n(a/2)$ and $\varepsilon_n(a/2)^*$ have a tendency to be negative, particularly if the $Z_i$'s are highly skewed (i.e., $\beta^2$ is large). This tendency for nonparametric confidence intervals to undercover has been exhibited empirically; see IGLEHART (1975), for example. The procedure of Section 8 will attempt to deal with this coverage rate problem.

Note that the $t_n$ coverage error is always biased upwards from that of $t_n^*$ by an amount $z(1-\alpha/2) \phi(z(1-\alpha/2))/n$. This is an attractive property of $t_n^*$ in comparison to $t_n$, in view of the undercoverage mentioned above. The cost associated with using $t_n$, rather than $t_n^*$, is that the $t_n$ interval is longer, asymptotically, by an amount $o(1) z(1-\alpha/2)/(\log n)^{3/2}$.

A similar analysis can be performed for the intervals $[L_n^*(a/2)$, $R_n^*(a/2)]$. PEISER (1943) showed that
Thus, using the uniformity in \( x \) of the expansion \( \Psi_{2,n}(x) \), we get

\[
\varepsilon_\alpha(n^{1/2}) = \Psi_{2,n}(z_{n-1}(1-\alpha/2)) - \Psi_{2,n}(z_{n-1}(\alpha/2)) - (1-\alpha) + o(n^{-1})
\]

\[
= \zeta_\alpha(n/2) + (z_\alpha^3 + x_n) \phi(x_n)/2n + o(n^{-1})
\]

where \( \zeta_\alpha = z(1-\alpha/2) \). Thus, the coverage rate for the interval

\[ [L_n'(\alpha/2), R_n'(\alpha/2)] \]

tends to be larger than that of \( t_n' \), by an amount

\[
(x_\alpha^3 + x_n) \phi(x_n)/2n.
\]

For highly skewed populations, this gives intervals based on Student t-quantiles an advantage over those based on normal quantiles. The use of Student t-quantiles comes at the cost of an interval which is longer by an amount \( o(2)(x_\alpha^3 + x_n)/(\mu) \) \( n^{3/2} \), however.

Note that for samples from populations with normal \( Y_1 \) and \( \tau_1 = 1 \), \( \varepsilon_\alpha(n^{1/2}) = o(n^{-1}) \), as expected.

7. Johnson's Pivotal Transformation

In Section 6, it was shown that under reasonably general assumptions, the one-sided coverage errors for \( t_n \) and \( t_n^* \) are of order \( O(n^{-1/2}) \). These errors arise due to asymmetry effects related to skewness and ratio estimator bias. In a recent paper JOHNSON (1978) considered, in the case where \( \tau_n = 1 \), a transformation of the pivot \( t_n \) derived on the basis of Cornish-Fisher expansions (see CORNISH and FISHER (1937) for a discussion of these expansions).

Empirical evidence collected by Johnson indicated that the
transformation led to intervals that reflected the "correct" degree of asymmetry. We now investigate the pivotal transformation of Johnson, using the machinery developed in Section 4.

Consider the sequence

\[ T_n = c_n + \theta_n n^{-1/2} + \rho_n(t_n)^2 n^{-1/2} \]

where \( \theta_n = \theta(\bar{u}_n, v_n) \), \( \rho_n = \rho(\bar{u}_n, v_n) \) and \( \theta(\cdot), \rho(\cdot) \) are functions analytic on a neighborhood of \( (\mu, \sigma(z)) \). Let \( \theta = \theta(\mu, \sigma^2(z)) \), \( \rho = \rho(\mu, \sigma^2(z)) \), and observe that

\[ T_n' = \hat{z}_n + (a^2_n - \hat{w}_n z_n^2 + \theta + \rho z_n^2) n^{-1/2} + a^{-1} x_n \]

where \( x_n \) is \( O_p(1) \).

We now use Proposition 4.12, Theorem 5.6, and relation (5.7) to obtain the cumulant expressions

\[ \kappa_1(T_{1,n}) = (-\beta/2 + \alpha + \theta + \rho) n^{-1/2} + o(n^{-1}) \]
\[ \kappa_2(T_{1,n}) = 1 + o(n^{-1}) \]
\[ \kappa_3(T_{1,n}) = (-2\beta + 6\alpha + 6\rho) n^{-1/2} + o(n^{-1}) \]

Observe that by setting \( \theta = \beta/6, \rho = \beta/3 - \alpha \), all three cumulants above are reduced to \( O(n^{-1}) \). This suggests setting \( \theta_n = \beta_n/6, \rho_n = \beta_n/3 - \alpha_n \) where
Let $T_n^*$ be defined through (7.1) and (7.3), substituting $t_n^*$ and $v_n^*$ for $t_n$ and $v_n$.

(7.4) **Theorem (i)** Suppose that $E(|\tau_n|^k + |\varphi_n|^k) < \infty$ for all $k$.

Then,

$$E(\gamma) = \int f(y) \phi(y) dy + o(n^{-1/2})$$

for all $f \in C^\infty_0(\mathbb{R})$.

(ii) If $E(|\tau_n|^9 + |\varphi_n|^9) < \infty$, and if the density assumption of Theorem 5.11 holds, then

$$P(T_n \leq x) = \Phi(x) + o(n^{-1/2})$$

uniformly in $x$.

(iii) Results (i) and (ii), under the assumptions stated, are valid for $T_n^*$.

**Proof.** First, observe that $\gamma_n$ (see (7.2)) is the remainder term from Taylor's theorem for $T_n$. On the set $\{v_n > 0, \tau_n \neq 0\}$, $\gamma_n$ has the form
\[ x_n = (\hat{U}_n \cdot \hat{v})^3 H_n(x_n)/6 \]

when \( H_n(x_n) \) is bounded on \([\mu - \varepsilon, \mu + \varepsilon] \). Now, apply Theorems 4.8(ii) and 4.16, as in the proof of Theorem 5.11, to obtain (ii).

For (i), write

\[ Ef(T_n) = Ef(T'_n) + Ef(Df(\eta_n)/n + o(n^{-1/2})) \]

and argue as in the proof of Theorem 4.16. The proofs for \( T^n \) can be handled similarly.

We remark that the moment assumptions in Theorem 7.4 follow from the fact that \( U_1 \) must be expanded to include \( Y_{1,1} \) with \( k+j = 3 \), due to the presence of \( \beta_n \) in \( T_n \).

For the classical case where \( \tau_n = 1 \), the transformed pivots \( T_n \) and \( T_n^* \) are precisely the statistics suggested by Johnson, up to a term which is \( O_p(n^{-1}) \). Note that Theorem 7.4 gives rigorous substance to the statement that \( T_n(T_n^*) \) "normalizes" \( t_n(t_n^*) \) in the sense of creating a r.v. which is closer to a normal. This is not surprising, in light of the fact that Johnson's calculations were based on Cornish-Fisher expansions, which are "normalization" series (see [30], p. 643).

Theorem 7.4 can be easily applied to coverage error asymptotics to yield the following result: If \( (Y_n, T_n) \) satisfies the assumptions of Theorem 7.4(iii), then all the coverage errors (one-sided as well as two-sided) for intervals based on \( T_n \) or \( T_n^* \)
are $o(n^{-1/2})$ uniformly in the parameter $p$. Thus, the Johnson pivotal transformation corrects for asymmetry effects.

6. A Second-Order Pivotal Transformation

As discussed in Section 6, nonparametric confidence intervals have a tendency to undercover at small sample sizes. However, the analysis of the symmetric intervals $[L_n(x/2), R_n(x/2)]$ showed that the coverage error is basically determined by the term in $n^{-1}$ of the asymptotic expansion of $P(t_n \leq x)$. This suggests that any attempt to correct the coverage rate of the symmetric intervals $[L_n(x/2), R_n(x/2)]$ must deal with higher order error terms than those considered by the Johnson pivotal transformation.

Consider the statistic

$$T_n = T_n + v_n t_n n^{-1} + \omega_n t_n^3 n^{-1}$$

where $v_n = v(U_n) + o_p(n^{-3/2})$, $\omega_n = \omega(U_n) + o_p(n^{-3/2})$, and $v(+), \omega(\cdot)$ are functions analytic on a neighborhood of $\mu$. Before proceeding with an expansion of $T_n$, we state the following approximations:

$$\beta_n = \bar{M}_n - 3\bar{X}_n \gamma/\sigma(Z) + o_p(n^{-3/2})$$

$$\alpha_n = \bar{M}_n - \bar{X}_n \delta/\sigma(Z) + o_p(n^{-3/2})$$
where \( M_1 = Z_1^2/\sigma^3(Z_1), N_1 = Z_1 \tau_1/(\sigma(Z) \text{ E}_t) \). Substituting (8.2) into (8.1) yields the following expansion (for the purposes of calculation, we take \( \sigma(Z) = 1 \)):

\[
\begin{align*}
\hat{T}_n &= \left\{ Z_n \left( 1 - \frac{1}{2} (W_n - 2aZ_n^2 + Z_n \tilde{Q}_n n^{-1/2}) n^{-1/2} \right) \\
&+ \frac{3}{4} (W_n - 2aZ_n^2) n^{-1} \right\} (1 - n^{-1}) \\
&+ \lambda n^{-1/2} + \rho Z_n^2 (1 - (W_n - 2aZ_n^2) n^{-1/2}) n^{-1/2} + M_n / 6a - Z_n \gamma / 2n \\
&+ (\lambda_n / 3 - Z_n \gamma - \hat{Z}_n \delta) Z_n^2 n^{-1} \\
&+ \nu Z_n n^{-1} + \omega Z_n^3 n^{-1} + O(n^{-3/2}) \\
&= \hat{T}_n + a^{-3/2} \chi_n .
\end{align*}
\]

Proposition 5.12 can be used to calculate the cumulants of \( \hat{T}_n^1 \):

\[
\begin{align*}
\kappa_1(\hat{T}_n^1) &= O(n^{-3/2}) \\
\kappa_2(\hat{T}_n^1) &= 1 + (36 - 7\gamma - 9a^2 + 24\beta^2 + 36 + 7\lambda / 3 + 6 + 2v + 6\omega)n^{-1} + O(n^{-3/2}) \\
\kappa_3(\hat{T}_n^1) &= O(n^{-3/2}) \\
\kappa_4(\hat{T}_n^1) &= (6\lambda + 4\beta^2 / 3 + 18 + 4a\beta + 12\gamma - 3a^2 + 24\omega)n^{-1} + O(n^{-3/2}) .
\end{align*}
\]

Once again, a judicious choice of \( v \) and \( \omega \) can reduce the order of the above cumulants to \( O(n^{-3/2}) \). In fact, choosing
\[ v_n = \gamma_n / 2 + 13\beta_n^2 / 72 - 3/4 - 5\lambda_n / 12 \]

\[ \omega_n = \gamma_n + 3\alpha_n^2 / 12 - \delta_n / 2 + \alpha_n \beta_n / 2 - \beta_n^2 / 18 - 3/4 - \lambda_n / 4 \]

where

\[
\gamma_n = I_{\{v_n > 0, \tau_n \neq 0\}} \sum_{i=1}^{n} \left( Y_i \tau_n \gamma_i \right)^2 / (\tau_n v_n)
\]

\[
\lambda_n = I_{\{v_n > 0, \tau_n \neq 0\}} \sum_{i=1}^{n} \left( Y_i \tau_n \gamma_i \right)^4 / (\tau_n v_n)^2 - 3
\]

\[
\delta_n = I_{\{\tau_n \neq 0\}} \sum_{i=1}^{n} \tau_i^2 / \tau_n^2
\]

yields a pair \( v = v(\mu), \omega = \omega(\mu) \) which reduces the above cumulant expressions (8.4) to \( O(n^{-3/2}) \).

(8.5) THEOREM (i) Suppose that \( E(\gamma_n^k + |\tau_n|^k) < 0 \) for all \( k \).
Then,

\[
Rf(\hat{\gamma}_n) = \int f(y) \phi(y) dy + o(n^{-1}) , \quad \text{for all} \ f \in C_2^\infty(R).
\]

(ii) If \( E(Y_n^16 + \tau_n^16) < 0 \), and if \( (Y_n, \tau_n) \) satisfies the density assumptions (5.10) (iii) or (iv), then

38
\[ P(T_n \leq x) = \phi(x) + o(n^{-1}) \]

uniformly in \( x \).

Proof. We apply Theorems 4.8(ii) and 4.16 to \( T_n \) (note that now fourth moments of \( (Y_n, \tau_n) \) must be included in \( U_n \)). The perturbation \( X_n \) of (8.3) is a Taylor series remainder similar to that found in (7.2). One then argues as in the proof of Theorem 7.4 for \( T_n \).

It is interesting to examine the situation when sampling from \( (Y_1, \tau_1) \), where \( \tau_1 \equiv 1 \) and \( Y_1 \) is normally distributed. In this case,

\[ T_n = \tau_n - (t_n + t_n^3) / 4n + o_p(n^{-3/2}) \]

This, up a term of order \( o_p(n^{-3/2}) \), is the Hotelling-Frankel transformation, which was derived in [15]. This transformation was designed as a device to transform a Student \( t \)-variate with \( n-1 \) degrees of freedoms into a r.v. with a "more" normal distribution. Theorem 8.5 thus shows that \( \hat{T}_n \) is the nonparametric analogue of the Hotelling-Frankel transformation.

Theorem 8.5 has important consequences for confidence interval estimation. In particular, under assumption (8.5) (ii), the result proves that all the coverage errors (one-sided and two-sided) for intervals based on \( \hat{T}_n \) are \( o(n^{-1}) \) uniformly in \( p \). Thus, the transformation (8.1) improves coverage rates, as well as corrects for asymmetry effects.
9. Numerical Results

In this section, we report the results of a Monte Carlo study of the coverage characteristics of "normal quantile" confidence intervals based on the pivots $t_n$, $T_n$, and $\hat{T}_n$.

(9.1) EXAMPLE. Choose $\tau_n \equiv 1$ and let $Y_n$ have an exponential distribution centered at 0 (i.e., $P(Y_n > y) = \exp(-(y+1))$ for $y > -1$). This example was studied in [9].

(9.2) EXAMPLE. Let $\tau_n \equiv 1$, and suppose $Y_n$ has a chi-square distribution with 10 degrees of freedom. This example, as well as (9.1), was considered in [19].

(9.3) EXAMPLE. Let $(W_n; n \geq 1)$ be the sequence of consecutive customer waiting times in an M/M/1 queue with arrival intensity $\lambda = 5$ and service intensity $\mu = 10$. The process $(W_n)$ is then a Markov chain which takes on the value 0 infinitely often. Returns to 0 constitute regeneration times for $(W_n)$ and thus a sequence $((Y_i, \tau_i))$ of appropriate regenerative pairs can be constructed, with a goal of estimating $E W$, the stationary waiting time. See Iglehart (1971) for more details on this process.

(9.4) EXAMPLE. Let $(B_t; t \geq 0)$ be the busy-time process obtained from the M/M/1 queue of Example 9.3; i.e., $B_t$ is 1 or 0 depending on whether or not the server is busy at time $t$. This process regenerates itself at those instants at which a customer arrives to
find a free server. Based on this sequence of regeneration times, a
certainty interval for the long-run proportion of time that the
server is busy can be derived, yielding a sequence \((Y_i, \tau_i)\) of
regenerative pairs.

For Examples 9.1 and 9.2, 2500 replications of the sampling
experiment were created; for Examples 9.3 and 9.4, 1000 replications.
Pseudo-random numbers were obtained from the Learmonth-Levis random
number generator (see LEARMONTH and LEWIS (1973) for a description).
The goal was to estimate \(P(\phi_n \leq z(0.05))\), \(P(\phi_n \geq z(0.95))\) and
\(P(z(0.05) \leq \phi_n \leq z(0.95))\) for \(\phi_n = \tau_n, \hat{T}_n,\) and \(\hat{T}_n\).

Note that both \(T_n\) and \(\hat{T}_n\) are non-linear in the parameter \(r\).
Thus in order to determine \(100(1-\alpha)\%\) confidence interval boundaries
based on these statistics, the zeros of some non-linear equations
must be found. Specifically, in the case of \(\hat{T}_n\), one first considers
the cubic polynomial

\[
(9.5) \quad f_n(x) = \theta_n x^{-1/2} + x(1 + \nu_n n^{-1}) + x^2 \rho_n n^{-1/2} + x^3 \omega_n n^{-1}.
\]

Given some fixed \(\varepsilon > 0\), one then finds solutions \(x_n(i)\) satisfying

\(f_n(x_n(i)) = z_i\) such that \(|z_i - x_n(i)| < \varepsilon\), for \(z_1 = z(p+1-\alpha)\),

\(z_2 = z(p)\). Let \(E_n\) be the event that such solutions exist uniquely,
with \(x_n(1) > x_n(2)\). On \(\nu_n > 0, \hat{T}_n \neq 0\), set

\[
\hat{T}_n(p) = r_n - \nu_n^{1/2} (I_{\hat{T}_n} x_n(1) + (1-I_{\hat{T}_n}) z(p+1-\alpha))/(n^{1/2} \hat{T}_n).
\]

\[
(9.6) \quad \hat{E}_n(p) = r_n - \nu_n^{1/2} (I_{\hat{T}_n} x_n(2) + (1-I_{\hat{T}_n}) z(p))/(n^{1/2} \hat{T}_n).
\]
(9.7) PROPOSITION. (i) If \( E(T_n^4 + \tau_n^4) < \infty \), then \( 1 - P(E_n) \to \frac{k}{2} \).

(ii) Under assumption (ii) of Theorem 8.5, the error asymptotics of \( [\hat{L}_n(p), \hat{E}_n(p)] \) are \( o(n^{-1}) \), uniformly in \( p \).

(iii) If \( E(T_n^4 + \tau_n^4) < \infty \), then \( \hat{E}_n(p) - \hat{L}_n(p) = E_n(p) - L_n(p) = o(n^{-1}) \) a.s.

**Proof.** Because of the continuity of \( \lambda_n, v_n, \rho_n, \) and \( \omega_n \) in \( \tilde{U}_n \), there exists \( \delta \) such that \( |\tilde{U}_n - \mu| < \delta \) implies all four estimators are within \( \eta \) of their limits. Hence, \( |\tilde{U}_n - \mu| < \delta \) implies that

\[
|f_n(x) - x| \leq K \eta n^{-1/2}
\]

(9.8)

\[
|\langle Df_n \rangle(x) - 1| \leq K \eta n^{-1/2}
\]

for some \( K \), uniformly in \( x \) on \([z_i-\varepsilon, z_i+\varepsilon]\). Thus, for \( n \) sufficiently large, \( f \) is monotone on \([z_i-\varepsilon, z_i+\varepsilon]\), with

\[
f_n(z_i+\varepsilon) - x_i > \varepsilon/2, \quad f_n(z_i-\varepsilon) - x_i < -\varepsilon/2,
\]

provided \( |\tilde{U}_n - \mu| < \delta \).

So, \( E_n \geq (|\tilde{U}_n - \mu| < \delta) \) for \( n \) sufficiently and hence \( 1 - P(E_n) \leq P(|\tilde{U}_n - \mu| > \delta) \leq O(n^{-1}) \) (see (4.18)), provided \( E(T_n^4 + \tau_n^4) < \infty \).

Relation (9.8) also proves (iii), with the assistance of the strong law of large numbers. For (ii), we use the fact that Theorem 4.8 allows the \( E_i \)'s to be defined arbitrarily outside a neighborhood of \( \mu \).
For the pivots $T_n$ and $T^*_n$, a different approach is more attractive. Observe that the exponential pivots $T^e_n$ and $T^{*e}_n$ defined by

$$T^e_n = I_{\{\rho_n \neq 0\}} \left( \frac{n}{2\rho_n} \left( \exp(\frac{-n\rho_n}{n^{1/2}}) - 1 \right) + \theta_n n^{-1/2} \right) + I_{\{\rho_n = 0\}} (t_n + \theta_n n^{-1/2})$$

are attractive. Observe that the exponential pivots $T^e_n$ and $T^{*e}_n$ defined similarly) satisfy $T^e_n = T_n + O(n^{-1})$, $T^{*e}_n = T^*_n + O(n^{-1})$.

The pivots $T^*_n$ and $T^{*e}_n$ are monotone in the parameter $r$, thereby avoiding some of the complications inherent in using $T_n$ or $T^*_n$. An argument based on Theorem 4.16 proves that confidence intervals based on $T^*_n$ or $T^{*e}_n$ enjoy the same error asymptotics as those for $T_n$ or $T^*_n$, up to order $O(n^{-1/2})$.

It should be noted, however, that the coverage estimates in Examples 9.1 through 9.4 were computed using the estimators $T_n$, $T^*_n$, and $T_n$ explicitly. In other words, because the value of $r$ was known for each of the examples, the three pivots were explicitly calculated to determine which of the intervals $I_1 = (-\infty, \alpha(0.05))$, $I_2 = [\alpha(0.05), \alpha(0.95)]$, and $I_3 = [\alpha(0.95), +\infty)$ covered the pivots. In practice, of course, one would have to explicitly calculate the confidence interval boundaries. For the pivot $T^*_n$, this would require finding roots of the cubic $f_n(x)$, and substituting into (9.6).

Given the fact that $f_n(x) = x + O_p(n^{-1/2})$, Newton's method for root-solving should be quite well-behaved, and hence the numerical difficulties involved in solving (9.5) should not be too significant.
Table 1 displays the results for the exponential and chi-square examples. Table 2 illustrates the behavior of the pivots for the M/M/1 queueing process examples. It should be pointed out that for the \( W_n \) process, \( \{ (Y_n, \tau_n) \} \) does not satisfy assumptions (iii) or (iv) of Theorem (5.10), since \( \tau_n \) is a lattice r.v. in this case. However, it can be shown that the other three examples do satisfy the conditions of Theorem 5.10.

Note that Examples 9.1 through 9.3 appear to confirm the error asymptotics of Sections 6, 7, and 8. The pivot \( T_n \) tends to "balance" the one-sided coverage probabilities, moving them towards their correct values of 0.05. This confirms the asymmetry correction induced by the Johnson pivotal transformation. The pivot \( T_n \) goes one step further: It seems to deal reasonably well with the overall confidence interval coverage rate. In Example 9.4, all three methods do well. Such an outcome is not surprising, in light of the fact that the sample skewness \( \beta \), for a sample size of 1000, was only 0.01.
<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Pivot</th>
<th>Exponential Coverage Rates</th>
<th>Chi-square Coverage Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I₁</td>
<td>I₂</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tₙ</td>
<td>0.210</td>
<td>0.758</td>
<td>0.032</td>
</tr>
<tr>
<td>Tₙ</td>
<td>0.173</td>
<td>0.793</td>
<td>0.034</td>
</tr>
<tr>
<td>T'ₙ</td>
<td>0.166</td>
<td>0.784</td>
<td>0.050</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tₙ</td>
<td>0.154</td>
<td>0.820</td>
<td>0.026</td>
</tr>
<tr>
<td>Tₙ</td>
<td>0.106</td>
<td>0.851</td>
<td>0.043</td>
</tr>
<tr>
<td>T'ₙ</td>
<td>0.082</td>
<td>0.874</td>
<td>0.044</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tₙ</td>
<td>0.136</td>
<td>0.838</td>
<td>0.026</td>
</tr>
<tr>
<td>Tₙ</td>
<td>0.096</td>
<td>0.860</td>
<td>0.044</td>
</tr>
<tr>
<td>T'ₙ</td>
<td>0.061</td>
<td>0.897</td>
<td>0.042</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tₙ</td>
<td>0.121</td>
<td>0.855</td>
<td>0.024</td>
</tr>
<tr>
<td>Tₙ</td>
<td>0.088</td>
<td>0.870</td>
<td>0.042</td>
</tr>
<tr>
<td>T'ₙ</td>
<td>0.053</td>
<td>0.908</td>
<td>0.039</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tₙ</td>
<td>0.113</td>
<td>0.864</td>
<td>0.023</td>
</tr>
<tr>
<td>Tₙ</td>
<td>0.075</td>
<td>0.881</td>
<td>0.044</td>
</tr>
<tr>
<td>T'ₙ</td>
<td>0.040</td>
<td>0.925</td>
<td>0.035</td>
</tr>
<tr>
<td>Sample Size</td>
<td>Pivot</td>
<td>Waiting Times $W_n$ Coverage Rates</td>
<td>Busy Time $B_\tau$ Coverage Rates</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>-----------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$I_1$</td>
<td>$I_2$</td>
</tr>
<tr>
<td>1000 replications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>$t_n$</td>
<td>0.333</td>
<td>0.649</td>
</tr>
<tr>
<td></td>
<td>$T_n$</td>
<td>0.201</td>
<td>0.756</td>
</tr>
<tr>
<td></td>
<td>$\hat{T}_n$</td>
<td>0.065</td>
<td>0.802</td>
</tr>
<tr>
<td>80</td>
<td>$t_n$</td>
<td>0.256</td>
<td>0.726</td>
</tr>
<tr>
<td></td>
<td>$T_n$</td>
<td>0.176</td>
<td>0.787</td>
</tr>
<tr>
<td></td>
<td>$\hat{T}_n$</td>
<td>0.025</td>
<td>0.891</td>
</tr>
<tr>
<td>120</td>
<td>$t_n$</td>
<td>0.242</td>
<td>0.741</td>
</tr>
<tr>
<td></td>
<td>$T_n$</td>
<td>0.148</td>
<td>0.809</td>
</tr>
<tr>
<td></td>
<td>$\hat{T}_n$</td>
<td>0.020</td>
<td>0.910</td>
</tr>
<tr>
<td>160</td>
<td>$t_n$</td>
<td>0.219</td>
<td>0.767</td>
</tr>
<tr>
<td></td>
<td>$T_n$</td>
<td>0.131</td>
<td>0.833</td>
</tr>
<tr>
<td></td>
<td>$\hat{T}_n$</td>
<td>0.018</td>
<td>0.937</td>
</tr>
<tr>
<td>200</td>
<td>$t_n$</td>
<td>0.194</td>
<td>0.792</td>
</tr>
<tr>
<td></td>
<td>$T_n$</td>
<td>0.117</td>
<td>0.846</td>
</tr>
<tr>
<td></td>
<td>$\hat{T}_n$</td>
<td>0.016</td>
<td>0.948</td>
</tr>
<tr>
<td>Sample Size</td>
<td>Pivot</td>
<td>Waiting Times $W_n$ 1000 replications</td>
<td>Busy Times $B_n$ 1000 replications</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>---------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td></td>
<td>Coverage Rates</td>
<td>Coverage Rates</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$I_1$</td>
<td>$I_2$</td>
<td>$I_3$</td>
</tr>
<tr>
<td>$t_n$</td>
<td>0.164</td>
<td>0.818</td>
<td>0.018</td>
</tr>
<tr>
<td>400</td>
<td>$T_n$</td>
<td>0.095</td>
<td>0.858</td>
</tr>
<tr>
<td></td>
<td>$\hat{T}_n$</td>
<td>0.019</td>
<td>0.934</td>
</tr>
<tr>
<td>$t_n$</td>
<td>0.124</td>
<td>0.852</td>
<td>0.024</td>
</tr>
<tr>
<td>800</td>
<td>$T_n$</td>
<td>0.075</td>
<td>0.873</td>
</tr>
<tr>
<td></td>
<td>$\hat{T}_n$</td>
<td>0.027</td>
<td>0.945</td>
</tr>
<tr>
<td>$t_n$</td>
<td>0.121</td>
<td>0.844</td>
<td>0.035</td>
</tr>
<tr>
<td>1200</td>
<td>$T_n$</td>
<td>0.091</td>
<td>0.850</td>
</tr>
<tr>
<td></td>
<td>$\hat{T}_n$</td>
<td>0.038</td>
<td>0.926</td>
</tr>
<tr>
<td>$t_n$</td>
<td>0.108</td>
<td>0.857</td>
<td>0.035</td>
</tr>
<tr>
<td>1600</td>
<td>$T_n$</td>
<td>0.076</td>
<td>0.861</td>
</tr>
<tr>
<td></td>
<td>$\hat{T}_n$</td>
<td>0.037</td>
<td>0.924</td>
</tr>
<tr>
<td>$t_n$</td>
<td>0.103</td>
<td>0.859</td>
<td>0.038</td>
</tr>
<tr>
<td>2000</td>
<td>$T_n$</td>
<td>0.073</td>
<td>0.858</td>
</tr>
<tr>
<td></td>
<td>$\hat{T}_n$</td>
<td>0.044</td>
<td>0.911</td>
</tr>
</tbody>
</table>

47
REFERENCES


Asymptotic Theory for Nonparametric Confidence Intervals

Peter W. Glynn

Department of Operations Research
Stanford University
Stanford, CA 94305

Office of Naval Research - Dept. of the Navy
Arlington, VA 22217

This document has been approved for public release and sale; its distribution is unlimited.

In practice, nonparametric confidence intervals often have undesirable small sample asymmetry and convergence characteristics. By applying Edgeworth expansion theory, we obtain asymptotic expansions for the errors associated with nonparametric confidence intervals. This analysis isolates the various elements that contribute to error. We then proceed to develop first and second order corrections to the standard nonparametric interval, which deal with asymmetry problems and coverage difficulties, respectively.