The Annals of Statistics
1991, Vol. 19, No. 4, 2209–2225

NORMALIZING TRANSFORMATIONS AND BOOTSTRAP CONFIDENCE INTERVALS[1]

BY SADANORI KONISHI

Institute of Statistical Mathematics and Pennsylvania State University

This paper considers the problem of constructing approximate confidence intervals for functional parameters in the nonparametric case. The approach based on transformation theory is applied to improve standard confidence intervals. The accelerated bias-corrected percentile interval introduced by Efron relies on the existence of a normalizing transformation with bias and skewness corrections, although calculation does not require explicit knowledge of its functional form. We formally construct such a transformation and estimate bias and skewness correction factors for nonparametric situations. The resulting interval is shown to be second-order accurate. To this end Edgeworth expansions for the distributions of transformed statistics are derived, using the von Mises expansion.

1. Introduction. In recent years intensive investigations have been made concerning the problem of constructing approximate confidence intervals. Work has been done both for parametric and nonparametric situations. Standard confidence intervals may be constructed based on consistent estimators with asymptotic normality. The common weakness of these intervals appears to lie in accuracy, since the standardized quantities are often quite skewed and biased, especially for small sample sizes. Several authors have improved the standard confidence intervals, using Edgeworth or Cornish–Fisher expansions [Abramovitch and Singh (1985), Bartlett (1953), Beran (1984), Johnson (1978), Hall (1983), Hinkley and Wei (1984), Peers and Iqbal (1985) and Withers (1983)].

The bootstrap method provides an alternative procedure for constructing nonparametric confidence intervals. In practical applications the percentile interval [Efron (1979)] has been widely used, and the advantages and disadvantages have been pointed out both in theoretical and practical aspects. A comprehensive survey of work in this area was given by DiCiccio and Romano (1988). Efron (1987) improved the percentile interval by taking bias and skewness corrections into account, and introduced the accelerated bias-corrected percentile interval called the BC_a interval. The BC_a interval is

Received February 1989; revised May 1990.

[1]Part of this research was done while the author was visiting the Center for Multivariate Analysis, Pennsylvania State University. Research supported by the Department of Statistics, Pennsylvania State University, and the Ministry of Education Grant 63530015.

AMS 1980 *subject classifications*. 62E20, 62G15.

Key words and phrases. Bootstrap confidence intervals, Edgeworth expansions, statistical functional, transformations, von Mises expansion.

constructed based on the existence of a normalizing transformation. The advantage of this method is that the calculation does not require the exact form of a transformation, but only its existence. Hall (1988) showed without knowledge of a transformation that the BC_a interval is second-order correct in the context of the smooth function model for which estimators can be expressed as a function of multivariate vector means.

In the present paper we formally construct a normalizing transformation with bias and skewness corrections required in Efron (1987), using an Edgeworth expansion based on a functional Taylor series expansion. The result in Section 3 shows that the BC_a interval is second-order correct for the nonparametric situations. It also provides a unified approach to a normalizing transformation theory and can be used to construct confidence intervals with second-order accuracy for the parametric situations. In Section 4 we propose an approximate confidence interval for a functional parameter in a nonparametric model, which can be constructed without bootstrap sampling. Some numerical results are presented at the end of Section 4.

2. Edgeworth expansion for the distributions of transformed statistics. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables with unknown distribution function F. Let $\hat{\theta}_n = \hat{\theta}(X_1, X_2, \ldots, X_n)$ be an estimator of θ, a parameter of interest which depends on F. We assume that there exists a suitably regular functional $T(\cdot)$ on the space of distribution functions on R^p such that $\theta = T(F)$ and $\hat{\theta}_n = T(F_n)$, where F_n is the empirical distribution function of $X_1, X_2, \ldots, X_n$.

Suppose that $\hat{\theta}_n$ admits the functional Taylor series expansion

$$\hat{\theta}_n = \theta + n^{-1}\sum_{i=1}^{n} T_1(X_i; F) + n^{-2}\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} T_2(X_i, X_j; F) + o_p(n^{-1}), \tag{2.1}$$

where $T_1(X_i; F)$ and $T_2(X_i, X_j; F)$ are defined as symmetric functions such that for an arbitrary distribution G on R^p,

$$\frac{d^i}{d\varepsilon^i} T((1-\varepsilon)F + \varepsilon G) = \int \cdots \int T_i(x_1, \ldots, x_i; F) \prod_{j=1}^{i} d\{G(x_j) - F(x_j)\}$$

at $\varepsilon = 0$ and

$$\int T_i(x_1, \ldots, x_i; F)\, dF(x_j) = 0 \quad \text{for } 1 \le j \le i.$$

Then it may be seen that $n^{1/2}(\hat{\theta}_n - \theta)$ is asymptotically normally distributed with mean 0 and variance

$$\sigma^2(F) = \int T_1^2(x; F)\, dF(x).$$

The standard confidence interval is constructed based on the normal approximation to the distribution of $n^{1/2}(\hat{\theta}_n - \theta)/\hat{\sigma}$, where $\hat{\sigma}$ is the estimated standard deviation of $\hat{\theta}_n$. In practice this approximation is not adequate, since the pivotal quantities are often not unbiased and quite skewed. The normal approximation may be improved by adjusting bias and reducing skewness. To this end we use an Edgeworth expansion for the distribution of a transformed statistic. For theoretical work on the functional Taylor series expansion, we refer to von Mises (1947), Reeds (1976) and Withers (1983).

Suppose that the variance of $\sigma^2(F)$ is estimated by

$$\hat{\sigma}^2(F_n) = n^{-1} \sum_{i=1}^{n} T_1^2(X_i; F_n). \tag{2.2}$$

We consider the bias and skewness of each of the following quantities:

$$T_\sigma = n^{1/2} \frac{\hat{\theta}_n - \theta}{\sigma(F)} \quad \text{and} \quad T_s = n^{1/2} \frac{\hat{\theta}_n - \theta}{\hat{\sigma}(F_n)}. \tag{2.3}$$

Substituting (2.1) in T_σ and calculating the moments and from them the cumulants of T_σ, we formally expand the first three cumulants in the form

$$\begin{aligned} \kappa_1(T_\sigma) &= n^{-1/2} b_\sigma(F) + O(n^{-3/2}), \\ \kappa_2(T_\sigma) &= 1 + O(n^{-1}), \\ \kappa_3(T_\sigma) &= n^{-1/2} k_\sigma(F) + O(n^{-3/2}), \end{aligned}$$

where

$$\begin{aligned} b_\sigma(F) &= \{2\sigma(F)\}^{-1} \int T_2(x, x; F)\, dF(x), \\ k_\sigma(F) &= \sigma^{-3}(F) \Big\{ \int T_1^3(x; F)\, dF(x) \\ &\quad + 3 \int\int T_1(x; F) T_1(y; F) T_2(x, y; F)\, dF(x)\, dF(y) \Big\}. \end{aligned} \tag{2.4}$$

For the Studentized quantity T_s, expanding $\hat{\sigma}(F_n)$ in a functional Taylor series and combining the resultant with (2.1), we have

$$\begin{aligned} T_s &= n^{-1/2} \sigma^{-1}(F) \sum_{i=1}^{n} T_1(X_i; F) \\ &\quad + n^{-3/2} \frac{1}{2} \Bigg[\sigma^{-1}(F) \sum_{i=1}^{n} \sum_{j=1}^{n} T_2(X_i, X_j; F) \\ &\quad - \sigma^{-3}(F) \sum_{i=1}^{n} \sum_{j=1}^{n} T_1(X_i : F) \Big\{ T_1^2(X_j; F) - \sigma^2(F) \\ &\quad + 2 \int T_1(z; F) T_2(X_j, z; F)\, dF(z) \Big\} \Bigg] + o_p(n^{-1/2}). \end{aligned} \tag{2.5}$$

Then the first three cumulants of T_s can be formally expanded as

$$\kappa_1(T_s) = n^{-1/2}b_s(F) + O(n^{-3/2}),$$
$$\kappa_2(T_s) = 1 + O(n^{-1}),$$
$$\kappa_3(T_s) = n^{-1/2}k_s(F) + O(n^{-3/2}),$$

where

$$\begin{aligned} b_s(F) = \frac{1}{2}\bigg[&\sigma^{-1}(F)\int T_2(x,x;F)\,dF(x) \\ &-\sigma^{-3}(F)\bigg\{\int T_1^3(x;F)\,dF(x) \\ &+2\int\int T_1(x;F)T_1(y;F)T_2(x,y;F)\,dF(x)\,dF(y)\bigg\}\bigg], \end{aligned} \tag{2.6}$$

$$\begin{aligned} k_s(F) = -\sigma^{-3}(F)\bigg\{&2\int T_1^3(x;F)\,dF(x) \\ &+3\int\int T_1(x;F)T_1(y;F)T_2(x,y;F)\,dF(x)\,dF(y)\bigg\}. \end{aligned}$$

Let $g(\hat{\theta}_n)$ be a one-to-one and twice continuously differentiable function in a neighborhood of $\hat{\theta}_n = \theta$. It is known that the limiting distribution of $n^{1/2}\{g(\hat{\theta}_n) - g(\theta)\}$ is normal with mean 0 and variance $\{\sigma(F)g'(\theta)\}^2$. Corresponding to each of T_σ and T_s in (2.3), we write

$$n^{1/2}\frac{g(\hat{\theta}_n) - g(\theta)}{vg'(\theta)}, \tag{2.7}$$

where $v = \sigma(F)$ or $v = \hat{\sigma}(F_n)$.

Under suitable regularity conditions a bias-corrected Edgeworth expansion for the distribution of the transformed variate (2.7) is

$$\begin{aligned} &P\left[n^{1/2}\frac{g(\hat{\theta}_n) - g(\theta)}{vg'(\theta)} - n^{-1/2}\left\{b_v + \frac{1}{2}\sigma(F)g''(\theta)g'(\theta)^{-1}\right\} < x\right] \\ &\quad = \Phi(x) - n^{-1/2}\left\{\frac{1}{6}k_v + \frac{1}{2}\sigma(F)g''(\theta)g'(\theta)^{-1}\right\}(x^2-1)\phi(x) \\ &\quad\quad + O(n^{-1}), \end{aligned} \tag{2.8}$$

where $\Phi(x)$ and $\phi(x)$ are the distribution function and the density of the standard normal distribution, respectively, and the bias b_v and the skewness k_v are given by

$$(b_v, k_v) = \begin{cases} (b_\sigma(F), k_\sigma(F)) \text{ in } (2.4), & \text{if } v = \sigma(F), \\ (b_s(F), k_s(F)) \text{ in } (2.6), & \text{if } v = \hat{\sigma}(F_n). \end{cases}$$

It may be seen from (2.8) that the asymptotic bias and skewness of the transformed variate are, respectively,

$$b_v + \tfrac{1}{2}\sigma(F)g''(\theta)g'(\theta)^{-1} \quad \text{and} \quad k_v + 3\sigma(F)g''(\theta)g'(\theta)^{-1}.$$

We formally expanded $\hat{\theta}_n$ in a functional Taylor series and obtained the Edgeworth expansion by calculating cumulants directly. Suitable conditions are required for the validity of these expansions. For the theory of Edgeworth expansions, we refer to Beran (1984), Bhattacharya and Ghosh (1978), Pfanzagl (1985), Takahashi (1988) and Withers (1983).

In the next section we discuss the BC_a interval from the point of view of transformation theory based on the Edgeworth expansion (2.8) with $v = \sigma(F)$. In Section 4 the Edgeworth expansion with $v = \hat{\sigma}(F_n)$ is used to construct an approximate confidence interval for θ.

3. Bootstrap confidence intervals. Nonparametric bootstrap methods provide a useful procedure for constructing confidence intervals for a parameter θ. Let $\hat{G}(x) = P_{F_n}\{\hat{\theta}_n^* < x\}$ be the cumulative distribution function of the bootstrap distribution of $\hat{\theta}_n^*$, where $\hat{\theta}_n^*$ is the estimator of θ based on bootstrap sample from F_n. The $\hat{G}(x)$ is the conditional distribution function of $\hat{\theta}_n^*$ and is, in practice, approximated by Monte Carlo sampling. Then, for a given α, the percentile interval is given by $[\hat{G}^{-1}(\alpha), \hat{G}^{-1}(1-\alpha)]$.

Efron (1987) improved the percentile interval by taking bias and skewness corrections into account and gave the accelerated bias-corrected percentile interval

$$\left[\hat{G}^{-1}(\Phi(z[\alpha])), \hat{G}^{-1}(\Phi(z[1-\alpha]))\right], \tag{3.1}$$

where $z[\alpha] = Z_0 + (Z_0 + z_\alpha)/\{1 - a(Z_0 + z_\alpha)\}$, z_α is the 100α percentile point of a standard normal variate and Z_0 and a are to be considered the bias and skewness corrections, respectively. The BC_a interval is constructed based on the existence of a transformation. It is assumed that there exists a monotone-increasing function g such that

$$\frac{g(\hat{\theta}_n) - g(\theta)}{1 + ag(\theta)} + Z_0 \tag{3.2}$$

is normal with mean 0 and variance 1.

The BC_a interval requires us to calculate the bootstrap distribution and two correction factors Z_0 and a included in (3.1) or (3.2). The advantage of this method is that the calculation does not require the exact form of a function g, but only its existence. In practice, it is necessary to show the existence of g and to estimate Z_0 and a, which satisfy the above condition asymptotically. We formally construct a transformation and estimate two correction factors such that the normal approximation to the distribution of (3.2) is valid with a remainder of order $o(n^{-1/2})$. In other words it will be shown that the BC_a interval is second-order correct in a nonparametric model.

The approach used here is based on an Edgeworth expansion for the distribution of a composite function of $\hat{\theta}_n$. The use of a composite function is motivated by the discussion in Section 10 of Efron (1987) and in DiCiccio and Tibshirani (1987) for a one-parameter model. In this section further investigation will be given for nonparametric situations.

3.1. *Normalizing transformation.* Let f be a monotone-increasing and differentiable function in a neighborhood of $\hat{\theta}_n = \theta$. Since the asymptotic variance of a transformed variate $f(\hat{\theta}_n)$ is $\{\sigma(F) f'(\theta)\}^2/n$, we first search for a function which satisfies the condition

$$\sigma(F) f'(\theta) = n^{1/2}. \tag{3.3}$$

In general, the asymptotic variance $\sigma^2(F)$ is a function of the parameter θ and also the moments of a population distribution F. Hence the variance stabilizing transformation f may depend on unknown parameters, and we write $f(\hat{\theta}_n) = f(\hat{\theta}_n, \eta)$, where $\eta = (\eta_1, \dots, \eta_q)$ is an unknown vector of parameters. Then we consider the quantity

$$T_f = n^{-1/2}\Big\{ f\big(\hat{\theta}_n, \hat{\eta}\big) - f(\theta, \hat{\eta})\Big\}, \tag{3.4}$$

where $\hat{\eta} = (\hat{\eta}_1, \dots, \hat{\eta}_q)$ is an estimator of η having an estimation error $O_p(n^{-1/2})$.

Suppose that f is a twice continuously differentiable function in a neighborhood of $(\hat{\theta}_n, \hat{\eta}) = (\theta, \eta)$, and that $\hat{\eta}$ admits a functional Taylor series expansion. Expanding $f(\hat{\theta}_n, \hat{\eta})$ and $f(\theta, \hat{\eta})$ in a Taylor's series around $(\hat{\theta}_n, \hat{\eta}) = (\theta, \eta)$ and $(\theta, \hat{\eta}) = (\theta, \eta)$, respectively, and substituting (2.1) and the corresponding expansion of $\hat{\eta}$ in the resultant yield

$$\begin{aligned} n^{1/2} T_f = {} & n^{-1} \sum_{i=1}^{n} T_1(X_i; F) f_\theta \\ & + n^{-2}\Bigg[\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} T_2(X_i, X_j; F) f_\theta + \frac{1}{2}\Bigg\{ \sum_{i=1}^{n} T_1(X_i; F)\Bigg\}^2 f_{\theta\theta} \\ & \qquad + \sum_{\alpha=1}^{q} \sum_{i=1}^{n} \sum_{j=1}^{n} T_1(X_i; F) U_1^{(\alpha)}(X_j; F) f_{\theta\eta_\alpha} \Bigg] \\ & + o_p(n^{-1/2}), \end{aligned}$$

where f_θ, $f_{\theta\theta}$ and $f_{\theta\eta_\alpha}$ are the partial derivatives of $f(\hat{\theta}_n, \hat{\eta})$ at $(\hat{\theta}_n, \hat{\eta}) = (\theta, \eta)$, and $U_1^{(\alpha)}(X_j; F)$ is the influence function of $\hat{\eta}_\alpha$. Then the first three cumulants of $n^{1/2} T_f$ can be expanded as

$$\begin{aligned} &\kappa_1\big(n^{1/2} T_f\big) = n^{-1/2}\{b_\sigma(F) + \tfrac{1}{2} c(F)\} + O(n^{-3/2}), \\ &\kappa_2\big(n^{1/2} T_f\big) = 1 + O(n^{-1}), \\ &\kappa_2\big(n^{1/2} T_f\big) = n^{-1/2}\{k_\sigma(F) + 3c(F)\} + O(n^{-3/2}), \end{aligned} \tag{3.5}$$

where $b_\sigma(F)$ and $k_\sigma(F)$ are given by (2.4) and

$$c(F) = n^{-1/2}\left\{\sigma^2(F) f_{\theta\theta} + 2\sum_{\alpha=1}^{q} f_{\theta\eta_\alpha}\int T_1(x;F)U_1^{(\alpha)}(x;F)\,dF(x)\right\}. \tag{3.6}$$

It should be noted that $n^{-1/2}\sigma^2(F) f_{\theta\theta}$ and $n^{-1/2} f_{\theta\eta_\alpha}$ included in $c(F)$ can be replaced by $-\partial\sigma(F)/\partial\theta$ and $-(\partial\sigma(F)/\partial\eta_\alpha)\sigma^{-2}(F)$, respectively, which are obtained by differentiating (3.3) with respect to θ and η_α.

Let $h(T_f)$ be a monotone-increasing and twice continuously differentiable function in a neighborhood of $T_f = 0$. It follows from (2.8) and (3.5) that an Edgeworth expansion for the distribution of $h(T_f)$ is given by

$$\begin{aligned} P\Bigg[& n^{1/2}\frac{h(T_f) - h(0)}{h'(0)} \\ & -n^{-1/2}\left\{b_\sigma(F) + \frac{1}{2}c(F) + \frac{1}{2}h''(0)h'(0)^{-1}\right\} < x\Bigg] \\ & = \Phi(x) - n^{-1/2}\left[\frac{1}{6}\{k_\sigma(F) + 3c(F)\} + \frac{1}{2}h''(0)h'(0)^{-1}\right] \\ & \quad\times (x^2-1)\phi(x) + O(n^{-1}). \end{aligned} \tag{3.7}$$

Comparing the standardized quantity in (3.7) with (3.2), we see that h is given as a solution of

$$n^{-1/2}h'(0) = 1 + ah(0). \tag{3.8}$$

A particular solution of this differential equation is

$$h(T_f) = a^{-1}\left\{\exp\left(n^{1/2}aT_f\right) - 1\right\}.$$

Taking this $h(T_f)$ in (3.7) yields

$$\begin{aligned} & P\left[a^{-1}\left\{\exp\left(n^{1/2}aT_f\right) - 1\right\} - n^{-1/2}\{b_\sigma(F) + \tfrac{1}{2}c(F) + \tfrac{1}{2}n^{1/2}a\} < x\right] \\ & \quad = \Phi(x) - n^{-1/2}\left[\tfrac{1}{6}\{k_\sigma(F) + 3c(F)\} + \tfrac{1}{2}n^{1/2}a\right](x^2-1)\phi(x) \\ & \qquad + O(n^{-1}). \end{aligned} \tag{3.9}$$

The approach to normality may be accelerated by choosing a to make the term of $O(n^{-1/2})$ in (3.9) to vanish, so the error involved is of order $O(n^{-1})$. This can be realized by choosing a to be

$$a = -n^{-1/2}\{\tfrac{1}{3}k_\sigma(F) + c(F)\}. \tag{3.10}$$

Then it can be seen that the Edgeworth expansion (3.9) is further reduced to

$$P\left[a^{-1}\left\{\exp\left(n^{1/2}aT_f\right) - 1\right\} + Z_0 < x\right] = \Phi(x) + O(n^{-1}), \tag{3.11}$$

where

$$Z_0 = -n^{-1/2}\{b_\sigma(F) - \tfrac{1}{6}k_\sigma(F)\}. \tag{3.12}$$

The transformation $h(T_f)$ with T_f defined by (3.4) can be rewritten as

$$h(T_f) = a^{-1}\frac{\left\{\exp\left(af\left(\hat{\theta}_n, \hat{\eta}\right)\right) - 1\right\} - \{\exp(af(\theta, \hat{\eta})) - 1\}}{1 + aa^{-1}\{\exp(af(\theta, \hat{\eta})) - 1\}}.$$

Comparing this with (3.2), we can show that a transformation required to construct the BC_a interval is

$$g(\hat{\theta}) = a^{-1}\left\{\exp\left(af\left(\hat{\theta}_n, \hat{\eta}\right)\right) - 1\right\}. \tag{3.13}$$

REMARK 1. If $a = 0$ in (3.10), it follows from (3.8) that the differential equation is reduced to $h'(0) = n^{1/2}$, so $h(T_f) = n^{1/2}T_f$. Hence, in the case where $a = 0$, it is not necessary to consider a composite function. Examples of this sort will be discussed in the next section through the case of Fisher's z-transformation for a correlation coefficient in a bivariate normal sample.

In a one-parameter model DiCiccio and Tibshirani (1987) also discussed the problem of constructing a function g which satisfies the condition (3.2) and gave an alternative confidence interval (BC_a^0 interval) by finding a function g in closed form. The functional form given by DiCiccio and Tibshirani (1987) coincides with $g(\hat{\theta}_n)$ in (3.13), but the skewness correction is different from the one in (3.10). As shown in Example 1 below, for a maximum likelihood estimator $\hat{\theta}_n$ in a one-parameter model, the bias and skewness corrections given by (3.10) and (3.12) are reduced to those suggested by Efron (1987).

EXAMPLE 1. Suppose that $l(\theta)$ is the log likelihood function of a random sample of size n drawn from a distribution depending upon an unknown parameter θ. Let $\hat{\theta}_n$ be the maximum likelihood estimator of θ. Take $T_\sigma = \kappa_2^{1/2}(\hat{\theta}_n - \theta)$, where κ_2 is the variance of a score function. Then it is known that

$$\begin{aligned} b_\sigma(F) &= n^{1/2}\kappa_2^{-3/2}(\kappa_{001} - 2\kappa_3)/6, \\ k_\sigma(F) &= n^{1/2}\kappa_2^{-3/2}(\kappa_{001} - \kappa_3), \\ c(F) &= n^{1/2}\kappa_2^{-3/2}(\kappa_3 - 2\kappa_{001})/6\ (= -\sigma'(F)), \end{aligned}$$

where $\kappa_{001} = E[\partial^3 l(\theta)/\partial\theta^3]$ and $\kappa_3 = E[\{\partial l(\theta)/\partial\theta\}^3]$. Substituting these in (3.10) and (3.12), we have $Z_0 = a = \kappa_3/(6\kappa_2^{3/2}) = \{\text{skewness of } \partial l(\theta)/\partial\theta\}/6$. Note that $f_{\theta\eta_\alpha}$ in (3.6) is 0. This agrees with the result in Theorem 2 of Efron (1987) and in DiCiccio and Tibshirani (1987).

3.2. *Second-order accuracy.* The bias and skewness correction factors Z_0 and a included in (3.11) have to be estimated from a sample. We replace, for example, $\sigma(F)$, $b_\sigma(F)$, $k_\sigma(F)$ and $C(F)$ by $\hat{\sigma} = \sigma(F_n)$, $\hat{b}_\sigma = b_\sigma(F_n)$, $\hat{k}_\sigma = k_\sigma(F_n)$ and $\hat{c} = c(F_n)$, respectively, which have estimation errors $O_p(n^{-1/2})$. Let

$$\hat{h} = \hat{a}^{-1}\left[\exp\left\{\hat{a}\left(f\left(\hat{\theta}_n, \hat{\eta}\right) - f(\theta, \hat{\eta})\right)\right\} - 1\right] + \hat{Z}_0,$$

where $\hat{a} = -n^{-1/2}(\hat{k}_\sigma/3 + \hat{c})$ and $\hat{Z}_0 = -n^{-1/2}(\hat{b}_\sigma - \hat{k}_\sigma/6)$. By using (2.1), $\hat{h}$ can be expanded in the form

$$\hat{h} = n^{-1/2}\sigma^{-1}(F)\sum_{i=1}^{n} T_1(X_i; F) + n^{-3/2}\Bigg[\tfrac{1}{2}\sigma^{-1}(F)\sum_{i=1}^{n}\sum_{j=1}^{n} T_2(X_i, X_j; F) + \left\{\sum_{i=1}^{n} T_1(X_i; F)\right\}^2 \left\{-\tfrac{1}{6}k_\sigma(F)\sigma^{-2}(F) + \tfrac{1}{2}n^{-1/2}f_{\theta\theta} - \tfrac{1}{2}\sigma^{-2}(F)c(F)\right\} + n^{-1/2}\sum_{\alpha=1}^{q}\sum_{i=1}^{n}\sum_{j=1}^{n} T_1(X_i; F)U_1^{(\alpha)}(X_j; F)f_{\theta\eta_\alpha}\Bigg] + Z_0 + o_p(n^{-1/2}).$$

Then the first three cumulants of $\hat{h}$ are of the form

$$\kappa_1(\hat{h}) = o(n^{-1/2}), \qquad \kappa_2(\hat{h}) = 1 + O(n^{-1}) \quad \text{and} \quad \kappa_3(\hat{h}) = o(n^{-1/2}),$$

which lead to

$$P(\hat{h} < x) = \Phi(x) + o(n^{-1/2}).$$

Hence the results derived in Sections 3.1 and 3.2 are summarized in the following theorem.

THEOREM 1. *Suppose that $\hat{\theta}_n$ has the Edgeworth expansion*

$$P\left[n^{1/2}\left(\hat{\theta}_n - \theta\right)/\sigma(F) < x\right] = \Phi(x) - n^{-1/2}\left\{b_\sigma(F) + \tfrac{1}{6}k_\sigma(F)(x^2 - 1)\right\}\phi(x) + O(n^{-1}),$$

where $b_\sigma(F)$ and $k_\sigma(F)$ are given by (2.4). *If $a = -n^{-1/2}\{k_\sigma(F)/3 + c(F)\} \neq 0$, then the transformation g and two correction factors $\hat{Z}_0$ and $\hat{a}$ ($\neq 0$) satisfying the condition*

$$P\left[\frac{g(\hat{\theta}_n) - g(\theta)}{1 + \hat{a}g(\theta)} + \hat{Z}_0 < x\right] = \Phi(x) + o(n^{-1/2}) \tag{3.14}$$

are given by

$$g(\hat{\theta}_n) = \hat{a}^{-1}\left[\exp\left\{\hat{a}f\left(\hat{\theta}_n, \hat{\eta}\right)\right\} - 1\right]$$

and

$$\hat{a} = -n^{-1/2}\left(\hat{k}_\sigma/3 + \hat{c}\right), \qquad \hat{Z}_0 = -n^{-1/2}\left(\hat{b}_\sigma - \hat{k}_\sigma/6\right), \tag{3.15}$$

where f is a solution of the differential equation (3.3) *and $\hat{b}_\sigma$, $\hat{k}_\sigma$ and $\hat{c}$ are, respectively, estimators of $b_\sigma(F)$ and $k_\sigma(F)$ in* (2.4) *and $c(F)$ in* (3.6), *which have estimation errors $O_p(n^{-1/2})$. If $a = 0$, then the transformation g is given by $g(\hat{\theta}_n) = f(\hat{\theta}_n, \hat{\eta})$.*

In most statistical applications the remainder in (3.14) may be replaced by $O(n^{-1})$. We derived Theorem 1 to show that Efron's BC_a interval is second-order correct in a nonparametric model. This theorem also gives a general procedure for finding approximate confidence intervals which achieve second-order accuracy in parametric problems. Several examples are given in the following discussion.

EXAMPLE 2. Let $\hat{\theta}_n/\theta$ be distributed according to a χ^2 distribution with n degrees of freedom. Consider the standardized quantity $T_\sigma = n^{1/2}(\hat{\theta}_n/n - \theta)/(\sqrt{2}\,\theta)$. Then $\sigma(F) = \sqrt{2}\,\theta$, $b_\sigma(F) = 0$, $k_\sigma(F) = 2\sqrt{2}$ and $c(F) = -\sqrt{2}$. An asymptotic variance stabilizing transformation defined by (3.3) is $f(\theta) = (n/2)^{1/2}\log\theta$. Hence

$$Z_0 = a = \{2/(9n)\}^{1/2} \quad \text{and} \quad g(\hat{\theta}_n) = (9n/2)^{1/2}(\hat{\theta}_n^{1/3} - 1),$$

and finally we have

$$P\Big[(9n/2)^{1/2}\Big\{(\hat{\theta}_n/\theta)^{1/3} - 1\Big\} + \{2/(9n)\}^{1/2} < x\Big] = \Phi(x) + O(n^{-1}).$$

This is the Wilson–Hilferty approximation for the central χ^2 distribution [see Wilson and Hilferty (1931)]. It is well known that this approximation produces highly accurate values even for small n. Efron (1987) considered the same example in the case where $n = 19$, and gave $a = 0.1081$, which is equal to $a = \{2/(9\cdot 19)\}^{1/2}$.

EXAMPLE 3. Let S be a sample covariance matrix based on a sample of size n from a p-variate normal distribution with positive definite covariance matrix Σ. Let $l_1 > l_2 > \cdots > l_p > 0$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$ be the ordered eigenvalues of S and Σ, respectively. Consider the standardized quantity $T_\sigma = n^{1/2}(l_i - \lambda_i)/(\sqrt{2}\,\lambda_i)$. If λ_i is a simple root, then $\sigma(F) = \sqrt{2}\,\lambda_i$, $b_\sigma(F) = (\sqrt{2})^{-1}\sum_{\alpha\neq i}^{p}\lambda_\alpha/(\lambda_i - \lambda_\alpha)$, $k_\sigma(F) = 2\sqrt{2}$ and $c(F) = -\sqrt{2}$ [see, e.g., Konishi (1981)]. Solving the differential equation $\sqrt{2}\,\lambda_i f'(\lambda_i) = n^{1/2}$ yields $f(\lambda_i) = (n/2)^{1/2}\log\lambda_i$. The bias and skewness correction factors are

$$Z_0 = -(2n)^{-1/2}\left\{\sum_{\alpha\neq i}^{p}\frac{\lambda_\alpha}{\lambda_i - \lambda_\alpha} - \frac{2}{3}\right\}$$

and

$$a = n^{-1/2}\frac{\sqrt{2}}{3} = \frac{1}{6} \text{ (asymptotic skewness of } T_\sigma).$$

Hence it follows from Theorem 1 that

$$P\left[\left(\frac{9n}{2}\right)^{1/2}\left\{\left(\frac{l_i}{\lambda_i}\right)^{1/3} - 1\right\} - (2n)^{-1/2}\left\{\sum_{\alpha\neq i}^{p}\frac{l_\alpha}{l_i - l_\alpha} - \frac{2}{3}\right\} < x\right]$$
$$= \Phi(x) + O(n^{-1}),$$

where the population eigenvalues are estimated by the corresponding sample eigenvalues. Note that the remainder is shown to be of order $O(n^{-1})$.

EXAMPLE 4. Let r be a correlation coefficient based on a sample of size n from a bivariate normal distribution with population correlation coefficient ρ. It is known [Hotelling (1953), page 212] that for $T_\sigma = n^{1/2}(r - \rho)/(1 - \rho^2)$,

$$\sigma(F) = 1 - \rho^2, \qquad b_\sigma(F) = -\rho/2, \qquad k_\sigma(F) = -6\rho \quad \text{and} \quad c(F) = 2\rho.$$

Hence we have

$$f(\rho) = (n^{1/2}/2)\log\{(1+\rho)/(1-\rho)\} = z(\rho),$$

say, $Z_0 = -n^{-1/2}\rho/2$ and $a = -n^{-1/2}\{k_\sigma(F)/3 + c(F)\} = 0$. Taking $g(\hat{\theta}_n) = z(r)$ in (3.14) gives

$$P\left[z(r) - z(\rho) - n^{-1/2}(r/2) < x\right] = \Phi(x) + O(n^{-1}).$$

This implies that the asymptotic variance stabilizing transformation $z(r)$ also yields a normalizing transformation, and that the confidence interval for ρ based on Fisher's z-transformation achieves second-order accuracy. Another example is the z-transformation for a sample canonical correlation coefficient in a normal sample. It can be verified [see, e.g., Konishi (1981)] that the skewness correction factor a is 0, and consequently a confidence interval for a population canonical correlation based on the z-transformation achieves second-order accuracy.

One advantage of Theorem 1 is that in parametric cases it unifies the general problem of finding normalizing transformations for estimators and produces accurate confidence intervals. One disadvantage is that the procedure requires finding a variance stabilizing transformation explicitly. In nonparametric cases it is generally difficult to find such a transformation in closed form. In order to avoid this difficulty, we introduce in the next section an alternative approximate confidence interval, which does not require bootstrap sampling. Tibshirani (1988) provided an algorithm to find a variance stabilizing transformation under certain restrictions and introduced the bootstrap t-interval based on a transformed variate.

4. Approximate confidence intervals based on transformations.

4.1. *Construction of confidence intervals.* A nonparametric confidence interval for a parameter θ is constructed based on an Edgeworth expansion for the distribution of a transformed statistic. The Edgeworth expansion (2.8) with $v = \hat{\sigma}(F_n)\ (= \hat{\sigma})$ implies that the bias-corrected Studentized quantity

$$(4.1) \qquad n^{1/2}\frac{g(\hat{\theta}_n) - g(\theta)}{\hat{\sigma}g'(\theta)} - n^{-1/2}\left\{b_s(F) + \frac{1}{2}\sigma(F)g''(\theta)g'(\theta)^{-1}\right\}$$

has asymptotically the standard normal distribution and its rate of convergence to normality is of order $n^{-1/2}$.

The problem is what function should be chosen to improve the normal approximation. Following the arguments of Section 3.1, we search for a function g which reduces the term of order $n^{-1/2}$ in the Edgeworth expansion (2.8) with $v = \hat{\sigma}(F_n)$ to 0, so the error involved is of order n^{-1}. This can be realized by finding a function which satisfies the second-order differential equation

$$\tfrac{1}{6}k_s(F) + \tfrac{1}{2}\sigma(F)g''(\theta)g'(\theta)^{-1} = 0. \tag{4.2}$$

In the nonparametric case it is generally difficult to obtain a function g in closed form, since $\sigma(F)$ and $k_s(F)$ may depend on θ but their functional forms are unknown. To avoid this difficulty arising in nonparametric situations, we first estimate the unknown parameters. Suppose that $b_s(F)$ and $k_s(F)$ are estimated by $\hat{b}_s = b_s(F_n)$ and $\hat{k}_s = k_s(F_n)$. Then a particular solution of the differential equation (4.2) is

$$g(\hat{\theta}_n) = \left(-3\hat{\sigma}/\hat{k}_s\right)\exp\left\{-\hat{k}_s\hat{\theta}_n/(3\hat{\sigma})\right\}.$$

Taking this $g(\hat{\theta}_n)$ in (4.1) and replacing $b_s(F)$ and $\sigma(F)$ by their estimates, we have

$$m^{-1}\{\exp(mT_s) - 1\} - n^{-1/2}\left(\hat{b}_s - \tfrac{1}{6}\hat{k}_s\right), \tag{4.3}$$

where $m = -\hat{k}_s/(3n^{1/2})$ and $T_s = n^{1/2}(\hat{\theta}_n - \theta)/\hat{\sigma}$.

We will show that the normal approximation to the distribution of the pivotal quantity (4.3) is valid with a remainder of order $o(n^{-1/2})$. It can be readily seen that the pivotal quantity can be expanded as

$$T_s - n^{-1/2}\left\{\hat{b}_s + \tfrac{1}{6}\hat{k}_s(T_s^2 - 1)\right\} + O_p(n^{-1}). \tag{4.4}$$

Substituting (2.5), $\hat{b}_s = b_s + O_p(n^{-1/2})$ and $\hat{k}_s = k_s + O_p(n^{-1/2})$ in (4.4) and calculating the cumulants yield the results that the bias and skewness of (4.4) are of order $o(n^{-1/2})$. Then we have

$$P\left[T_s - n^{-1/2}\left\{\hat{b}_s + \tfrac{1}{6}\hat{k}_s(T_s^2 - 1)\right\} < x\right] = \Phi(x) + o(n^{-1/2}). \tag{4.5}$$

This result also follows from Theorem 1 in Abramovitch and Singh (1985). Hence the normal approximation to the distribution of (4.3) is valid with a remainder of order $o(n^{-1/2})$. A confidence interval for a parameter θ is constructed by inverting the pivotal quantity (4.3) directly. The results are summarized in the following theorem.

THEOREM 2. *Suppose that the distribution function of* $T_s = n^{1/2}(\hat{\theta}_n - \theta)/\hat{\sigma}$ *with* $\hat{\sigma}^2$ *given by* (2.2) *can be expanded as*

$$\begin{aligned} P(T_s < x) = \Phi(x) &- n^{-1/2}\left\{b_s(F) + \tfrac{1}{6}k_s(F)(x^2 - 1)\right\}\phi(x) \\ &+ O(n^{-1}), \end{aligned} \tag{4.6}$$

where $b_s(F)$ *and* $k_s(F)$ *are given by* (2.6). *Let* $\hat{b}_s$ *and* $\hat{k}_s$ *be estimators of* $b_s(F)$ *and* $k_s(F)$, *respectively, which have estimation errors* $O_p(n^{-1/2})$. *Then* (i) *the pivotal quantity defined by* (4.3) *has the Edgeworth expansion*

$$P\left[m^{-1}\{\exp(mT_s) - 1\} - n^{-1/2}\left(\hat{b}_s - \tfrac{1}{6}\hat{k}_s\right) < x\right] = \Phi(x) + o(n^{-1/2}), \tag{4.7}$$

where $m = -\hat{k}_s/(3n^{1/2})$, *and consequently* (ii) *a confidence interval for* θ *is given by*

$$(4.8)\quad \theta \in \left[\hat{\theta}_n - n^{-1/2}(\hat{\sigma}/m)\log\left[1 + m\left\{\pm z_\alpha + n^{-1/2}\left(\hat{b}_s - \tfrac{1}{6}\hat{k}_s\right)\right\}\right]\right].$$

Let $\hat{\theta}_T(\alpha)$ be the α-endpoint of the confidence interval (4.8). It is easily seen that $\hat{\theta}_T(\alpha)$ can be expanded as

$$\hat{\theta}_T(\alpha) = \hat{\theta}_n - n^{-1/2}\hat{\sigma}\left[-z_\alpha + n^{-1/2}\left\{\hat{b}_s + \tfrac{1}{6}\hat{k}_s(z_\alpha^2 - 1)\right\}\right] + O_p(n^{-3/2}).$$

This implies that the confidence interval (4.8) is asymptotically equivalent to that based on a Cornish–Fisher inverse expansion for the percentile of the Edgeworth expansion (4.6). Hence from the discussion in Hall (1988), we have the following theorem.

THEOREM 3. *The* α*-endpoint* $\hat{\theta}_T(\alpha)$ *of the confidence interval* (4.8) *is second-order correct in the sense that* $\hat{\theta}_T(\alpha)$ *agrees with the exact* α*-endpoint to order* n^{-1}.

It follows from (4.5) that the polynomial transformation of T_s is also a pivotal quantity. However, it might be noticed that the use of this quantity does not give directly a confidence interval, and that the polynomial transformation is not a monotone function over the whole domain of T_s. We also note that both methods (4.5) and (4.7) depend on the transformations used and are not invariant under reparametrization. These two approximations are illustrated through an example of the sample correlation coefficient in the following.

EXAMPLE 5. Let r be the sample correlation coefficient as discussed in Example 4. We first obtain $T_1(x; F)$ and $T_2(x, y; F)$ for r by expanding r in a functional Taylor series. Substituting the resultant in (2.6) and calculating the asymptotic bias and skewness under the assumption of normality yield $b_s(F) = 3\rho/2$ and $k_s(F) = 6\rho$. Then from (4.7) we have

$$(4.9)\quad \begin{aligned} &P\left[n^{1/2}(-2r)^{-1}\{\exp(-2r(r-\rho)/(1-r^2)) - 1\} - n^{-1/2}r/2 < x\right] \\ &\quad = \Phi(x) + O(n^{-1}). \end{aligned}$$

Further, let $R_s = n^{1/2}(r - \rho)/(1 - r^2)$. It follows from (4.5) that the polynomial transformation of T_s is

$$(4.10)\quad P\left[R_s - n^{-1/2}\left\{\tfrac{3}{2}r + r(R_s^2 - 1)\right\} < x\right] = \Phi(x) + O(n^{-1}).$$

In the next section we compare these approximations and check the accuracy, using Monte Carlo simulation.

In parametric problems Konishi (1981, 1987) constructed the concept of normalizing transformations based on the rate of convergence to normality. The use of transformations to parametrize models has also been discussed by

Hougaard (1982) and DiCiccio (1984). There is a close relationship among them in the sense that transformations are chosen in such a way that the skewness of distributions vanishes or becomes smaller. For related work, see Konishi (1987).

REMARK 2. We used the transformation theory to improve the standard confidence interval and derived an approximate confidence interval. For the α-endpoint $\hat{\theta}_T(\alpha)$ given by (4.8), we consider a value of the bootstrap cumulative distribution function $P_{F_n}(\hat{\theta}_n^* < \hat{\theta}_T(\alpha))$. This approximates

$$P\Big[\hat{\theta}_n < \theta + 3\sigma(F)k_s^{-1}(F) \times \log\big[1 - n^{-1/2}\tfrac{1}{3}k_s(F)\{-z_\alpha + n^{-1/2}(b_s(F) - \tfrac{1}{6}k_s(F))\}\big]\Big].$$

Expanding the quantile in a Taylor series and applying an Edgeworth expansion for the distribution of T_σ yield

$$\begin{aligned} P\big[n^{1/2}(\hat{\theta}_n - \theta)/\sigma(F) &< z_\alpha - n^{-1/2}\{b_s(F) - \tfrac{1}{6}k_s(F) + \tfrac{1}{6}k_s(F)z_\alpha^2\}\big] \\ &= \Phi(z_\alpha) - n^{-1/2}\{b_s(F) - \tfrac{1}{6}k_s(F) + \tfrac{1}{6}k_s(F)z_\alpha^2 + b_\sigma(F) \\ &\qquad + \tfrac{1}{6}k_\sigma(F)(z_\alpha^2 - 1)\}\phi(z_\alpha) \\ &\quad + O(n^{-1}) \\ &= \Phi(z_\alpha) - n^{-1/2}\{b_s(F) - \tfrac{1}{6}k_s(F) + b_\sigma(F) - \tfrac{1}{6}k_\sigma(F) \\ &\qquad + \tfrac{1}{6}(k_s(F) + k_\sigma(F))z_\alpha^2\}\phi(z_\alpha) \\ &\quad + O(n^{-1}), \end{aligned} \tag{4.11}$$

where $(b_\sigma(F), k_\sigma(F))$ and $(b_s(F), k_s(F))$ are given by (2.4) and (2.6), respectively. It follows from (3.1) that

$$\Phi(z[\alpha]) = \Phi(z_\alpha) + (2Z_0 + az_\alpha^2)\phi(z_\alpha) + O(n^{-1}).$$

Comparing this with the last formula in (4.11) and noting that $b_s(F) - k_s(F)/6 = b_\sigma(F) - k_\sigma(F)/6$, we have

$$\begin{aligned} Z_0 &= -n^{-1/2}\{b_s(F) - \tfrac{1}{6}k_s(F)\} = -n^{-1/2}\{b_\sigma(F) - \tfrac{1}{6}k_\sigma(F)\} \\ &= -n^{-1/2}\tfrac{1}{6}\sigma^{-3}(F)\Big\{3\sigma^2(F)\int T_2(x,x;F)\,dF(x) - \int T_1^3(x;F)\,dF(x) \\ &\qquad -3\int\int T_1(x;F)T_1(y;F)T_2(x,y;F)\,dF(x)\,dF(y)\Big\}, \end{aligned} \tag{4.12}$$

$$a = -n^{-1/2}\tfrac{1}{6}\{k_s(F) + k_\sigma(F)\} = n^{-1/2}\tfrac{1}{6}\sigma^{-3}(F)\int T_1^3(x;F)\,dF(x).$$

If a is estimated by replacing F by the empirical distribution function F_n, then

$$\hat{a} = \frac{1}{6} \frac{\sum_{i=1}^{n} T_1^3(X_i; F_n)}{\{\sum_{i=1}^{n} T_1^2(X_i; F_n)\}^{3/2}},$$

which agrees with Efron's estimate [Efron (1987), page 178]. This implies that the approximate confidence interval given in Theorem 2 is asymptotically equivalent to the BC_a interval. The skewness correction a in (4.12) has a different form from the one given by (3.10), which is derived via transformation theory. We found, however, that two correction factors produce similar results in many problems. In a nonparametric setting we have been so far unable to make the relation clear. Hall (1988) showed that the BC_a interval is second-order correct in a smooth function model, and gave an interpretation of the skewness correction based on inverse Cornish–Fisher expansions of bootstrap critical points [see also Bickel (1988)].

4.2. *Numerical results.* A Monte Carlo study was performed to examine the normal approximations (4.9) and (4.10) to the distribution of the sample correlation coefficient in a normal sample. We first obtained the exact values of $P(R_s < r_0)$, where $R_s = n^{1/2}(r - \rho)/(1 - r^2)$, using Monte Carlo simulation; 1,000,000 repeated random samples were generated from a bivariate normal population for different combinations of ρ and n. Table 1 compares errors $(\times 10^4)$ in approximating the values of the probability $P(R_s < r_0)$ for a sample

TABLE 1
Errors in approximating the values of $P(R_s < r_0)$ *for* $n = 50$:
error = (*approximate value* − *exact value*) × 10^4

r_0	Exact	PT	KT	r_0	Exact	PT	KT
		$\rho = 0.1$				$\rho = 0.3$	
−2.0	0.0289	−90 (0.004)	−91 (0.004)	−2.2	0.0134	−54 (0.002)	−57 (0.002)
−1.9	0.0352	−96 (0.004)	−98 (0.005)	−2.0	0.0215	−68 (0.003)	−72 (0.003)
−1.8	0.0426	−102 (0.005)	−104 (0.005)	−1.8	0.0336	−81 (0.004)	−85 (0.004)
−1.7	0.0514	−108 (0.006)	−109 (0.006)	−1.6	0.0514	−92 (0.005)	−96 (0.005)
1.8	0.9487	106 (0.006)	108 (0.006)	2.0	0.9564	82 (0.006)	90 (0.005)
2.0	0.9635	95 (0.005)	97 (0.005)	2.2	0.9685	70 (0.005)	79 (0.004)
2.2	0.9746	81 (0.004)	83 (0.004)	2.4	0.9773	60 (0.004)	68 (0.003)
2.4	0.9824	67 (0.003)	68 (0.003)	2.6	0.9837	49 (0.004)	57 (0.003)
		$\rho = 0.7$				$\rho = 0.9$	
−1.7	0.0235	−24 (0.002)	−41 (0.002)	−1.6	0.0210	23 (0.0)	−2 (0.0)
−1.6	0.0315	−28 (0.002)	−46 (0.002)	−1.5	0.0297	24 (0.001)	−3 (0.001)
−1.5	0.0416	−32 (0.002)	−51 (0.002)	−1.4	0.0413	20 (0.001)	−8 (0.001)
−1.4	0.0539	−35 (0.003)	−55 (0.003)	−1.3	0.0555	17 (0.001)	−11 (0.001)
2.4	0.9660	−35 (0.005)	27 (0.003)	2.4	0.9595	−138 (0.003)	−10 (0.001)
2.6	0.9743	−43 (0.005)	21 (0.003)	2.6	0.9686	−153 (0.003)	−15 (0.001)
2.8	0.9805	−48 (0.005)	16 (0.003)	2.8	0.9758	−167 (0.003)	−20 (0.001)
3.0	0.9854	−53 (0.004)	11 (0.002)	3.0	0.9814	−179 (0.003)	−22 (0.001)

size $n = 50$ in the tail areas. Simulation entries in the table are estimated by averaging over 10,000 repeated Monte Carlo trials. The standard deviations for the means are given in parentheses. We use the notation KT and PT standing for the approximations (4.9) and (4.10), respectively.

It may be seen from Table 1 that the approximation PT is slightly superior to KT for low values of ρ, while KT is superior for high values of ρ in the upper tail areas. In the lower tails KT and PT have almost the same standard deviation, but in the upper tails KT has a somewhat smaller one. One important advantage of KT is that it can be readily inverted to construct a confidence interval for the parameter ρ. It should be pointed out that the approximation PT performs poorly in the neighborhood of the extreme value $(n^{1/2}/(2r))$ of a quadratic function, since the polynomial transformation for PT is not monotonically increasing or decreasing over the whole domain of R_s. This implies that the method based on the polynomial transformation should be applied with extreme care when a sample size is small and an asymptotic skewness is large.

Acknowledgments. I would like to thank an Associate Editor, two referees, Professor C. R. Rao, Professor G. J. Babu and Professor C. G. Khatri for helpful suggestions and comments. I am particularly indebted to one of the referees for constructive criticism and valuable comments which improved the quality of this paper considerably.

REFERENCES

ABRAMOVITCH, L. and SINGH, K. (1985). Edgeworth corrected pivotal statistics and the bootstrap. *Ann. Statist.* **13** 116–132.

BARTLETT, M. S. (1953). Approximate confidence intervals. *Biometrika* **40** 12–19.

BERAN, R. (1984). Jackknife approximations to bootstrap estimates. *Ann. Statist.* **12** 101–118.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

BICKEL, P. J. (1988). Discussion of "Theoretical comparison of bootstrap confidence intervals" by P. Hall. *Ann. Statist.* **16** 959–961.

DICICCIO, T. J. (1984). On parameter transformations and interval estimation. *Biometrika* **71** 477–485.

DICICCIO, T. J. and ROMANO, J. P. (1988). A review of bootstrap confidence intervals (with discussion). *J. Roy. Statist. Soc. Ser. B* **50** 338–354.

DICICCIO, T. J. and TIBSHIRANI, R. (1987). Bootstrap confidence intervals and bootstrap approximations. *J. Amer. Statist. Assoc.* **82** 163–170.

EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26.

EFRON, B. (1987). Better bootstrap confidence intervals (with discussion). *J. Amer. Statist. Assoc.* **82** 171–200.

HALL, P. (1983). Inverting an Edgeworth expansion. *Ann. Statist.* **11** 569–576.

HALL, P. (1988). Theoretical comparison of bootstrap confidence intervals (with discussion). *Ann. Statist.* **16** 927–985.

HINKLEY, D. and WEI, B.-C. (1984). Improvements of jackknife confidence limit methods. *Biometrika* **71** 331–339.

HOTELLING, H. (1953). New light on the correlation coefficient and its transforms (with discussion). *J. Roy. Statist. Soc. Ser. B* **15** 193–232.

HOUGAARD, P. (1982). Parametrizations of non-linear models. *J. Roy. Statist. Soc. Ser. B* **44** 244–252.

JOHNSON, N. J. (1978). Modified t tests and confidence intervals for asymmetrical populations. *J. Amer. Statist. Assoc.* **73** 536–544.

KONISHI, S. (1981). Normalizing transformations of some statistics in multivariate analysis. *Biometrika* **68** 647–651.

KONISHI, S. (1987). Transformations of statistics in multivariate analysis. In *Advances in Multivariate Statistical Analysis* (A. K. Gupta, ed.) 213–231. Reidel, Dordrecht.

PEERS, H. W. and IQBAL, M. (1985). Asymptotic expansions for confidence limits in the presence of nuisance parameters, with applications. *J. Roy. Statist. Soc. Ser. B* **47** 547–554.

PFANZAGL, J. (1985). *Asymptotic Expansions for General Statistical Models. Lecture Notes in Statist.* **31**. Springer, Berlin.

REEDS, J. (1976). On the definition of von Mises functionals. Ph.D. dissertation, Harvard Univ.

TAKAHASHI, H. (1988). A note on Edgeworth expansions for the von Mises functionals. *J. Multivariate Anal.* **24** 56–65.

TIBSHIRANI, R. (1988). Variance stabilization and the bootstrap. *Biometrika* **75** 433–444.

VON MISES, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* **18** 309–348.

WILSON, E. B. and HILFERTY, M. M. (1931). The distribution of chi-square. *Proc. Nat. Acad. Sci.* **17** 684–688.

WITHERS, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals. *Ann. Statist.* **11** 577–587.

INSTITUTE OF STATISTICAL MATHEMATICS
4-6-7 MINAMI-AZABU, MINATO-KU
TOKYO 106
JAPAN