

Accuracy of the Bootstrap Approximation

by

Arup Bose  
Purdue University  
West Lafayette, IN 47907

and

G.J. Babu  
Pennsylvania State University  
University Park, PA 16802

Technical Report # 89-10

Department of Statistics  
Purdue University

April 1989

First Revision December 1989  
Second Revision September 1990

# ACCURACY OF THE BOOTSTRAP APPROXIMATION

by

Arup Bose  
Purdue University

and

G.J. Babu\*  
Pennsylvania State University

## ABSTRACT

The sampling distribution of several commonly occurring statistics are known to be closer to the corresponding bootstrap distribution than the normal distribution, under some conditions on the moments and the smoothness of the population distribution. These conditional approximations are suggestive of the unconditional ones considered in this paper, though one cannot be derived from the other by elementary methods. In this paper, probabilistic bounds are provided for the deviation of the sampling distribution from the bootstrap distribution. The rate of convergence to one, of the probability that the bootstrap approximation outperforms the normal approximation, is obtained. These rates can be applied to obtain the  $L^p$  bounds of Bhattacharya and Qumsiyeh (1989) under weaker conditions. The results apply to Studentized versions of functions of multivariate means and thus cover a wide class of common statistics. As a consequence we also obtain approximations to percentiles of Studentized means and their appropriate modifications. The results indicate the accuracy of the bootstrap confidence intervals both in terms of the actual coverage probability achieved and also the limits of the confidence interval.

AMS 1985 subject classifications. Primary 60F05, 60F25, Secondary 62E20, 62G99

Key words and phrases. Bootstrap distribution, Edgeworth expansion, Cramer's condition, strongly non lattice variables, percentiles.

---

\*Research supported in part by NSA Grant MDA 904-90-H-1001.

## 1. INTRODUCTION

The asymptotic accuracy of the bootstrap procedure has been studied by several authors. Under some regularity conditions (smoothness and moments), for several types of statistics, the difference between the bootstrap distribution (which is random) and the actual distribution is  $o(n^{-1/2})$  a.s. These results hold, for example, for the sample mean,  $t$ -statistic, quantiles and certain smooth functions of multivariate sample means. See Bickel and Freedman (1981), Singh (1981), and Babu and Singh (1984). However, all these results are a.s. results and there is no way of knowing if this improvement over the normal approximation (the latter being typically  $O(n^{-1/2})$ ), holds with a probability approaching one at a fast rate. This issue is specially relevant in confidence interval problems, where both the eventual coverage probability and the accuracy of the critical point are important. See Hall (1986) for a discussion of such problems.

In this paper we study the probabilistic aspects of the bootstrap approximation. Let  $X, X_1, \dots, X_n$  be i.i.d. random vectors in  $\mathbb{R}^m$  with distribution function  $F$ . Let  $K_n = K_n(X_1, \dots, X_n)$  be a statistic and  $H_n$  be the distribution function of  $K_n - k_n(F)$  where  $\{k_n(F)\}$  is a sequence of constants. Appropriate bootstrap version of  $H_n$  is denoted by  $H_n^*$ . Our main result states that under certain conditions on  $K_n, k_n$  the moments and smoothness of  $F$ ,

$$(1.1) \quad P \left( \sup_x |H_n(x) - H_n^*(x)| \geq \varepsilon_n n^{-1/2} \right) \leq C n^{-(1+\gamma)},$$

where  $\varepsilon_n \rightarrow 0$  is an appropriate sequence and  $\gamma > 0$  depends on the moment assumptions. Using Lemma 2.1 of Babu and Bose (1988) and (1.1), one can derive confidence intervals

and estimates of associated errors like

$$\sup_{0 \leq \alpha \leq 1} |P(K_n - k_n(F) \leq H_n^{*-1}(\alpha)) - \alpha| \leq \varepsilon_n n^{-\frac{1}{2}} + Cn^{-(1+\gamma)}.$$

We provide an estimate of  $\varepsilon_n$ . This leads to a better understanding of confidence intervals and coverage probability. The inequality (1.1) is obtained under strong non lattice (s.n.l.) assumption on the distribution  $F$ . The estimates are improved under additional moment conditions and Cramer's condition. The class of statistics to which our results are applicable include sample means and their smooth functions. Result (1.1) is not true when  $X_i$ 's are lattice variables. This is clear from statement (1.7) of Singh (1981).

Under stronger moment conditions than those assumed in deriving (1.1), Babu and Bose (1988) proved that

$$(1.2) \quad P \left( \sup_x |H_n(x) - H_n^*(x)| \geq Cn^{-1}(\log n)^\delta \right) \leq Cn^{-1}.$$

This in turn shows that the coverage probability of the bootstrap critical point is accurate with error  $O(n^{-1}(\log n)^\delta)$ . Clearly (1.1) and (1.2) do not follow from each other but there is some similarity between the results.

Abramovitch and Singh (1985) have shown that a statistic  $K_n$  can often be modified to obtain  $\widehat{K}_n$  which satisfies  $\sup_x |P(\widehat{K}_n \leq x) - \Phi(x)| = O(n^{-1})$ . Further if  $\widehat{K}_n^*$  is an appropriate bootstrap version of  $\widehat{K}_n$  then  $\sup_x |P(\widehat{K}_n \leq x) - P^*(\widehat{K}_n^* \leq x)| = o(n^{-1})$  a.s. We also extend our result (1.1) and provide probabilistic bounds for such modified statistics. For instance we show that if  $H_n$  denotes the c.d.f. of a modified  $t$ -statistic and  $H_n^*$  is the corresponding bootstrap version, then for  $0 < t < 1$ ,  $H_n^{*-1}(t) - H_n^{-1}(t) = o(n^{-1})$  w.p.  $1 - O(n^{-(1+\delta)})$ .

Our results thus complement the a.s. results available in the literature and show that the chances of the bootstrap distribution outperforming the normal approximation are very high. These results can be used to improve estimated confidence intervals.

The original version of this paper was written in 1985. Subsequently we discovered that  $L^p$  estimates for the difference between the bootstrap cumulative distribution function (cdf) and the original cdf were obtained by Bhattacharya and Qumsiyeh (1989). Some of their results can be derived from the present work and be proved under weaker conditions. See Remarks 2.7 and 2.8. In one case requirement of moment of order  $s^2$  is reduced to  $s(s-1)$  and in another, requirement of Cramer's condition is weakened to the distribution being merely strongly non lattice, thereby allowing some discrete distributions.

The technical details of the proofs are given in the appendix. We make repeated use of the techniques of Bhattacharya and Ranga Rao (1978) in obtaining bounds for the difference between distribution functions and their Edgeworth expansions. The a.s. results mentioned earlier follow from our results but with additional assumptions. On the other hand, our probabilistic bounds cannot be obtained from the a.s. results.

## 2. THE MAIN RESULTS

In this section we will state the main results and sketch some of the proofs. Most of the notations used in the statements of the results are explained below. Additional notations used in the proofs are explained in the appendix.

Let  $X_k, k \geq 1$  be independent observations from a common cdf  $G$  on  $\mathbf{R}^m$  and let  $F_n$  be the empirical distribution (cdf) which puts mass  $1/n$  at each  $X_k, k = 1, \dots, n$ . Let

$G_n(x) = F_n(x + \bar{X}_n)$  where  $\bar{X}_n = n^{-1} \sum_{k=1}^n X_k$ . Let  $X_{ni}^*, i = 1, \dots, n$  (or  $X_i^*, i = 1, \dots, n$  in short) be i.i.d. from  $F_n$ . Let  $D$  denote the dispersion matrix of  $G$  and let  $V_n$  denote the sample dispersion matrix. Whenever  $V_n$  is positive definite,  $T_n$  denotes a matrix such that  $T_n V_n^{-1} T_n' = I$ . The distribution function of  $n^{-1/2} \sum_{k=1}^n (X_k - EX_1)$  is denoted by  $Q_n$  and  $Q_n^*$  denotes the (conditional) distribution function of  $n^{-1/2} \sum_{k=1}^n (X_k^* - \bar{X}_n)$  given  $X_1, \dots, X_n$ . In general the presence of (\*) indicates that we are dealing with a bootstrap quantity. The  $s$ th absolute moment of  $X$ ,  $E\|X_1\|^s$  for  $s \geq 0$  is denoted by  $\rho_s$ ,  $\chi_\nu = \chi_\nu(G)$  denotes the  $\nu$ th cumulant of  $G$  and  $\psi_{n,s-2} = \sum_{r=0}^{s-2} n^{-r/2} P_r(-\Phi_{0,D}; \{\chi_\nu\})$  is the usual signed measure associated with Edgeworth expansions (see Bhattacharya and Ranga Rao (1978), page 54, for a detailed discussion). For a real valued measurable function  $f$  on  $\mathbf{R}^m$ , let  $M_s(f) = \sup_x |f(x)|(1 + \|x\|^2)^{-1}$ ,  $w(f, \varepsilon, x) = \sup\{|f(z) - f(y)| : \|x - y\| < \varepsilon \text{ and } \|z - x\| < \varepsilon\}$  and  $\omega_f(\mathbf{R}^m) = \sup\{|f(x)| : x \in \mathbf{R}^m\}$ . To state our results, we define for  $s \geq 3$ ,

$$B_{n,s} = \int f d \left( Q_n - \sum_{r=0}^{s-2} n^{-r/2} P_r(-\Phi_{0,D} : \{\chi_\nu\}) \right)$$

and its bootstrap version,

$$B_{n,s}^* = \int f d \left( Q_n^* - \sum_{r=0}^{s-2} n^{-r/2} P_r(-\Phi_{0,V_n} : \{\chi_\nu^*\}) \right).$$

Finally

$$\Delta_{n,s}(G, \delta) = E_G \left( \|X_1\|^s I \left( \|X_1\| \geq \delta n^{1/2} \right) \right).$$

The two standard assumptions needed in deriving Edgeworth expansions are strong non-lattice (s.n.l.) of  $G$  or that  $G$  satisfies Cramer's condition,  $\lim_{\|t\| \rightarrow \infty} \sup |\widehat{G}(t)| < \theta < 1$ , where  $\widehat{G}$  denotes the characteristic function of  $G$ . We of course get much stronger results in the latter case.

**Theorem 2.1.** *Let  $G$  be s.n.l.,  $D$  be positive definite and  $\rho_{3+\gamma} < \infty$  for some  $\gamma > 3$ . Then*

$$P\left(\sup_x |Q_n(D^{1/2}x) - Q_n^*(T_n x)| \geq n^{-1/2}\varepsilon_n\right) = O\left(n^{-(1+\gamma)/4}\varepsilon_n^{-(2+\gamma)}\right),$$

*provided  $n^2\varepsilon_n^2 - C(3 + \gamma + 1)\log n \rightarrow \infty$ .*

*In particular it follows that for some  $\varepsilon_n \rightarrow 0$  and for some  $\delta > 0$ ,*

$$\sum_{n=1}^{\infty} P\left(\sup_x |Q_n(D^{1/2}x) - Q_n^*(T_n x)| \geq n^{-1/2}\varepsilon_n\right) \leq C \sum_{n=1}^{\infty} n^{-(1+\delta)} < \infty. \quad \blacktriangle$$

Babu and Singh (1984) have shown that  $\sup_x |Q_n(D^{1/2}x) - Q_n^*(T_n x)| = o(n^{-1/2})$  a.s.

Theorem 2.1 provides an estimate of the rate of this convergence. Also note that the above theorem yields much more than the complete convergence of  $\sup_x |Q_n(D^{1/2}x) - Q_n(T_n x)|$ .

In fact it gives a  $r$ -quick limit result and may be useful in sequential confidence band estimation of the distribution function. See Babu and Singh (1982). See Theorem 2.6 for a similar result for studentized statistics.

Theorem 2.1 can be derived from a more technical Theorem 2.2 given below, using moderate deviations results of Michel (1976).

**Theorem 2.2.** *Suppose  $G$  is s.n.l.,  $D$  is positive definite and  $\rho_{3+\gamma} < \infty$  for some  $\gamma \geq 0$ .*

*Suppose  $f$  is a real valued bounded measurable function and  $(\varepsilon_n)$  and  $(\delta_n)$  are two sequences*

*between 0 and 1 such that  $\max\left(n^2\delta_n^2, n\varepsilon_n^{2/3}\right) - C(3 + \gamma + 1)\log n \rightarrow \infty$ . Then*

$$\begin{aligned} & P\left(|B_{n,3}^*| \geq n^{-1/2}\varepsilon_n \left(1 + \omega_f(\mathbf{R}^m) + \int \omega\left(f \circ T_n^{-1}, \eta n^{-1/2}, y\right) \exp(-y^2/100) dy\right)\right) \\ & \leq C \left[ n^{-(3+\gamma-1)/3} \varepsilon_n^{-(3+\gamma)/3} + (1 - P(C_1 \leq \|T_n\| \leq C_2)) \right. \\ & \quad + P(\rho_3(G_n) \geq C_3) + P(\rho_3(G_n) \geq C_4\varepsilon_n/\delta_n) \\ & \quad \left. + n^{-(3+\gamma-2)/4} \delta_n^{-(3+2\gamma)/2} \varepsilon_n^{-1/2} \Delta_{n,3+\gamma}(G, \delta_n) \right]. \quad \blacktriangle \end{aligned}$$

The next theorem is an improvement on Theorem 2.2 under additional assumptions.

**Theorem 2.3.** *Suppose  $G$  satisfies Cramer's condition,  $D$  is positive definite and  $M_{s'}(f)$*

*$< \infty$  for some  $s' \leq s$ .  $(\varepsilon_n)$  and  $(\delta_n)$  are sequences between 0 and 1 such that*

*$\max \left( n^2 \delta_n^2, n \varepsilon_n^{2/s} \right) - C(s + \gamma + 1) \log n \rightarrow \infty$  and  $\rho_{s+\gamma} < \infty$  for some  $s \geq 3$  and  $\gamma \geq 0$ .*

*Then*

$$\begin{aligned} & P \left( |B_{n,s}^*| \geq n^{-(s-2)/2} \varepsilon_n \left( 1 + M_{s'} + \int \omega(f \circ T_n^{-1}, \exp(-dn), y) \exp(-y^2/100) dy \right) \right) \\ & \leq C \left[ n^{-(s+\gamma-1)} \varepsilon_n^{-(s+\gamma)/s} + (1 - P(C_1 \leq \|T_n\| \leq C_2)) \right. \\ & \quad + P(\rho_3(G_n) \geq C_3) + P \left( \rho_s(G_n) \geq C_4 \frac{\varepsilon_n}{\delta_n} \right) \\ & \quad \left. + n^{-(s+\gamma-2)/4} \delta_n^{-(s+2\gamma)/2} \varepsilon_n^{-1/2} \Delta_{n,s+\gamma}(G, \delta_n) \right]. \end{aligned}$$

*The constants  $C_i$ 's  $i = 1, \dots, 4$  in Theorems 2.1, 2.2 and 2.3 can be chosen, and  $C$  depends on  $C_i$ 's,  $\gamma, \eta$  and  $G$ , and is independent of  $f$ . ▲*

*Remark 2.4.* Note that in the above theorems, the random errors of empirical Edgeworth expansion have been approximated by non random errors, except the modulus of continuity term. ▲

*Remark 2.5.* The bounds in Theorems 2.2 and 2.3 can be significantly simplified under additional conditions. In the following discussion,  $s$  has a value 3 for Theorem 2.2. If  $s + \gamma > 4$  then the second term in both the bounds can be dropped. Further if  $s + \gamma > 6$  then the first and the third terms can also be dropped. Finally if  $\gamma > s$ , then the entire bound can be replaced by

$$o \left( n^{-\gamma/s} (\varepsilon_n / \delta_n)^{-(s+\gamma)/s} \right) + O \left( n^{-(s+\gamma-2)/4} \delta_n^{-(s+2\gamma)/2} \varepsilon_n^{-1/2} \Delta_{n,s+\gamma}(G, \delta_n) \right),$$

provided we also have  $\delta_n / \varepsilon_n = O(1)$ . The above simplifications can be achieved by using



moderate deviation result (Lemma A3) on  $\|X_i\|, \|X_i\|^2, \|X_i\|^3$  and  $\|X_i\|^g$ . We omit the details. ▲

## STUDENTIZED STATISTICS

A version of Theorem 2.1 is true for studentized statistics. However, we will not state the result for studentized statistics in its most general form to preserve clarity of exposition. ▲

Let  $H$  be a function from  $\mathbf{R}^m$  to  $\mathbf{R}$  which is thrice continuously differentiable in a neighborhood of  $\mu = E(X_i)$  and let  $\ell(z)$  denote the vector of first order partial derivatives of  $H$  evaluated at  $z$ . Suppose  $\lambda(\cdot) = (\lambda_1(\cdot), \dots, \lambda_g(\cdot))$  is a continuous function from  $\mathbf{R}^m$  to  $\mathbf{R}^g$  and  $\nu(\cdot)$  is twice continuously differentiable real valued function on  $\mathbf{R}^g$ . Define the studentized statistic

$$t(X, F) = n^{1/2}(H(\bar{X}_n) - H(\mu)) / \nu \left( n^{-1} \sum_{k=1}^n \lambda(X_k) \right),$$

and let  $t^*(X^*, F_n)$  be its bootstrap version,

$$t^*(X^*, F_n) = n^{1/2} \left( H(\bar{X}_n^*) - H(\bar{X}_n) \right) / \nu \left( n^{-1} \sum_{k=1}^n \lambda(X_k^*) \right).$$

Further assume that

$$\nu(E_F \lambda(X_1)) = [\ell'(\mu) D\ell(\mu)]^{1/2} = \sigma,$$

and

$$\nu(E_{F_n} \lambda(X_1^*)) = [\ell'(\bar{X}_n) V_n \ell(\bar{X}_n)]^{1/2} = \sigma_n.$$

Let  $L(X_i)$  be a linearly independent subcollection of  $(X_i, \lambda(X_i))$  with the property that all the elements of  $(X_i, \lambda(X_i))$  can be expressed as linear combinations of  $L(X_i)$ .

**Theorem 2.6.** *Let  $L(X_1)$  be strong non-lattice and  $E\|L(X_1)\|^{3+\gamma} < \infty$  for some  $\gamma > 3$ . Then  $P\left(\sup_x |P_F(t(X, F) \leq x) - P^*(t^*(X, F_n) \leq x)| \geq n^{-1/2}\varepsilon_n\right) = o(n^{-(1+\delta)})$  for some  $\varepsilon_n \rightarrow 0$  and  $\delta > 0$ . ▲*

The proof of this theorem uses Taylor expansion arguments given in the proof of Theorem 4 of Babu and Singh (1984), followed by an application of Theorem 2.1. We omit the details. In fact, one could also provide bounds for  $\varepsilon_n$  and  $\delta$  and extend the result to an  $s$ -term expansion along the lines of Theorem 2.3 under stronger conditions.

*Remark 2.7.* We will now derive Theorem 2.1 of Bhattacharya and Qumsiyeh (1989) (henceforth referred as BQ) from Theorem 2.3. We will prove the result only for sample means whereas BQ work with studentized means. This case can also be dealt with, by using the Taylor expansion and obtaining three term Edgeworth expansions for  $t(X, F)$  and  $t^*(X, F_n)$  along the lines of proof of Theorem 2.3.

Using  $s = 4$  in Theorem 2.3, we get

$$Q_n^*(T_n x) - \sum_{r=0}^2 n^{-r/2} \int_{-\infty}^x P_r^*(y) \phi(y) dy = o_p(n^{-1})$$

for some polynomials  $P_r^*$ 's.

Further,

$$Q_n(D^{1/2} x) - \sum_{r=0}^2 n^{-r/2} \int_{-\infty}^x P_r(y) \phi(y) dy = o(n^{-1}).$$

Thus it follows that

$$\begin{aligned} n(Q_n(D^{1/2} x) - Q_n^*(T_n x)) &= n^{1/2} \int_{-\infty}^x (P_1(y) - P_1^*(y)) \phi(y) dy \\ &\quad + \int_{-\infty}^x (P_2(y) - P_2^*(y)) \phi(y) dy + o_p(n^{-1}). \end{aligned}$$

The coefficients involved in the polynomials are smooth functions of the moments. By CLT,  $n^{1/2}(P_1(x) - P_1^*(x)) \rightarrow N(0, \sigma_0^2(x))$  where  $\sigma_0^2(x)/\phi^2(x)$  is a polynomial in  $x$ . By SLLN, the second term in the right side of the above equality  $\rightarrow 0$  a.s. This shows that  $n(Q_n(D^{1/2}x) - Q_n(T_n x)) \xrightarrow{\mathcal{D}} N(0, \sigma_0^2(x))$ . Note that we need  $E\|X_1\|^6 < \infty$  to use the CLT since  $P_1$  involves the third moment. For studentized statistics we need  $E\|X_1\|^8$  since in that case  $P_1$  involves the fourth moment. ▲

*Remark 2.8.* Next we show how Proposition 3.1 of BQ follows from our results.

Note that under Cramer's condition,

$$(2.1) \quad \begin{aligned} P\left(\sup_x |Q_n^*(x) - \sum_{r=0}^{s-2} n^{-r/2} P_r^*(x)| \geq n^{-(s-2)/2} \varepsilon_n\right) \\ = o\left(n^{-(s-2)/2}\right) \text{ for some } \varepsilon_n \rightarrow 0 \text{ provided } \rho_{s(s-1)} < \infty. \end{aligned}$$

This follows by using moderate deviation bounds on the right side of Theorem 2.3. It is clear that the inequality (2.1) is the main fact needed in the proof of Proposition 3.1 of BQ. The other auxiliary results needed in the proof are straightforward consequences of moderate deviation results. It may be pointed out that BQ need  $\rho_{s^2} < \infty$ . Our proof improves the result by proving it under the weaker assumption of  $\rho_{s(s-1)} < \infty$ . Also note that BQ need the Cramer's condition. For the case  $s = 3$ , we do not need Cramer's condition but only s.n.l. Thus we can allow many discrete distributions. ▲

## MODIFIED STATISTICS

The order of errors can be improved by modifying the statistics as in Abramovitch and Singh (1985) (henceforth referred as AS). In principle this modification can be carried out for statistics admitting suitable Edgeworth expansions. Moreover as AS have pointed

out, this modification can be carried out to as many steps as we please with corresponding reduction in error. To keep the discussion simple, we will deal with only one-step modified  $t$  statistic. In this discussion  $X_i$ 's are assumed to be one-dimensional. Let

$$t_n = n^{1/2} (\bar{X}_n - E(X_1)) / s_n,$$

where

$$s_n^2 = n^{-1} \sum_{k=1}^n (X_k - \bar{X}_n)^2.$$

Let  $t_{1n}$  be the one-step modification of  $t_n$  (see AS for a detailed discussion) given by

$$t_{1n} = t_n + n^{-1} \sum (X_i - \bar{X}_n)^3 (2t_n^2 + 1) \left(6n^{1/2} s_n^3\right)^{-1}.$$

Under sufficient conditions on  $G$ ,  $t_{1n}$  has an expansion (uniformly in  $x$ ) of the form (we omit the proofs),

$$(2.2) \quad P(t_{1n} \leq x) = \Phi(x) + n^{-1} p_1(x) + n^{-3/2} p_2(x) + o\left(n^{-3/2}\right).$$

A similar expansion holds for the bootstrap version  $t_{1n}^*$  (by an appropriate strengthening of Theorem 2.2) and

$$(2.3) \quad \begin{aligned} P\{\sup_x |P^*(t_{1n}^* \leq x) - \Phi(x) - n^{-1} p_1^*(x) - n^{-3/2} p_2^*(x)| \geq n^{-3/2} \varepsilon_n\} \\ = o\left(n^{-(1+\delta)}\right) \text{ for some } \delta > 0 \text{ and } \varepsilon_n \rightarrow 0 \end{aligned}$$

From the above two equations it easily follows (as in Remark 2.7) that,

$$(2.4) \quad n^{3/2} (P(t_{1n} \leq x) - P^*(t_{1n}^* \leq x)) \xrightarrow{\mathcal{D}} N(0, \sigma^2(x)),$$

and

$$(2.5) \quad \sum_{n=1}^{\infty} P\{\sup_x |P(t_{1n} \leq x) - P^*(t_{1n}^* \leq x)| \geq n^{-1} \varepsilon_n\} \leq \sum_{n=1}^{\infty} n^{-(1+\delta)} < \infty,$$

for some  $\delta > 0$  and  $\varepsilon_n \rightarrow 0$ .

Using formal Cornish-Fisher expansion, Johnson (1978) modified the  $t$ -statistic to eliminate the effect of population skewness. AS procedure, though different from Johnson's approach, gives essentially the same result when applied to the first step correction of Student's  $t$ -statistic considered above. The procedure developed by AS immediately suggests higher order corrections and makes the idea of these modifications clear mathematically. Our results (2.4) and (2.5) strengthen Theorem 5 of AS. The modifications of  $t$ -statistic and the results of this section are useful in practice as the bootstrap of the modified  $t$  automatically corrects for skewness and other factors and gives a better approximation (2.5) than the bootstrap of  $t$ -statistic.

## APPLICATION TO PERCENTILES AND CONFIDENCE INTERVALS

Our results on direct Edgeworth expansions can be converted into results for percentiles. Some of these results are given below. Again, to keep the discussion simple, we do not give exact estimates for the  $o(\cdot)$  terms. To state the results, define

$$\begin{aligned} \bar{Q}_n(x) &= Q_n(D^{1/2}x), & \bar{Q}_n^*(x) &= Q_n(T_n x), \\ H_n(x) &= P(t_n \leq x), & H_n^*(x) &= P^*(t_n^* \leq x), \\ H_{1n}(x) &= P(t_{1n} \leq x), & \text{and } H_{1n}^*(x) &= P^*(t_{1n}^* \leq x). \end{aligned}$$

**Theorem 2.9.** *Fix  $t \in (0, 1)$ . Let  $\gamma \geq 0$ .*

*a) If  $\rho_{6+\gamma} < \infty$  and  $X_1$  is s.n.l. then*

$$(2.6) \quad \bar{Q}_n^{*-1}(t) - \bar{Q}_n^{-1}(t) = o\left(n^{-1/2}\right) \text{ w.p. } 1 - O\left(n^{-(1+\delta)}\right).$$

*b) If  $\rho_{12+\gamma} < \infty$  and  $(X_1, X_1^2)$  is s.n.l. then*

$$(2.7) \quad \bar{H}_n^{*-1}(t) - \bar{H}_n^{-1}(t) = o\left(n^{-1/2}\right) \text{ w.p. } 1 - O\left(n^{-(1+\delta)}\right).$$

c) If  $\rho_{24+\gamma} < \infty$  and  $(X_1, X_1^2)$  satisfies Cramer's condition then

$$(2.8) \quad \overline{H}_{1n}^{*-1}(t) - \overline{H}_{1n}^{-1}(t) = o(n^{-1}) \text{ w.p. } 1 - O(n^{-(1+\delta)}).$$

In (2.6), (2.7), (2.8),  $\delta \geq 0$  depends on  $\gamma$ . Further  $\delta > 0$  if  $\gamma > 0$  and  $o(\cdot)$  terms are all nonrandom quantities. ▲

*Remark 2.10.* Note that for the  $t$ -statistic little more than the existence of 12 moments are assumed in Theorem 2.9b). With some additional effort and messy arguments, it may be possible to reduce this requirement to existence of 6 moments. See Babu and Bai (1990) for results on Edgeworth expansions under minimal moment conditions. Similar remarks apply to parts a) and c) of the theorem. Under stronger moment conditions, the above theorem gives mathematically stronger forms of results given in Babu and Singh (1984) and Abramovitch and Singh (1985). With these extra conditions their results follow from ours. However, our results cannot be derived from theirs since those results are all a.s. results with no corresponding probability statements. ▲

*Remark 2.11.* We sketch below how our results may be applied to confidence interval problems. For a general discussion on bootstrap confidence intervals, see Hall (1988). Suppose  $X_i$ 's are real-valued with unknown mean  $\mu = EX_1$ , for which a  $100(1 - \alpha)\%$  confidence interval is desired. To keep the discussion simple, we will limit ourselves to a derivation of one sided confidence interval for a fixed  $\alpha, 0 < \alpha < 1$ . We will assume that conditions of Theorem 2.6, or more generally conditions for (2.5) and (2.8) to hold, are

true. Recall that

$$t_n = \frac{n^{\frac{1}{2}}(\bar{X}_n - \mu)}{s_n} = \frac{\hat{\mu} - \mu}{\nu_n}, \text{ say}$$

$$t_{1n} = t_n + n^{-\frac{1}{2}}\mu_n(2t_n^2 + 1), \text{ where}$$

$$\mu_n = \Sigma(X_i - \bar{X}_n)^3 / (6ns_n^3).$$

Using Theorem 2.6 and Lemma 2.1 of Babu and Bose (1989) and noting that the maximum jump of  $H_n$  is  $o(n^{-\frac{1}{2}})$ ,

$$|P(t_n < H_n^{*-1}(\alpha)) - \alpha| = o(n^{-\frac{1}{2}}),$$

yielding

$$P(\mu > \hat{\mu} - \nu_n H_n^{*-1}(\alpha)) = \alpha + o(n^{-\frac{1}{2}}).$$

It is interesting to note that AS derive such confidence intervals, correct up to  $o(n^{-\frac{1}{2}})$ , but use the modified statistics  $t_{1n}$  to do so. We may use  $t_{1n}$  and improve as follows. Define

$$t_{1\alpha}^* = H_{1n}^{*-1}(\alpha),$$

$$b_n = t_{1\alpha}^* - n^{-\frac{1}{2}}\mu_n(1 + 2t_{1\alpha}^{*2}) + \frac{4}{n}\mu_n^2 t_{1\alpha}^* (1 + 2t_{1\alpha}^{*2}).$$

Consider the function

$$\eta(x) = x + n^{-\frac{1}{2}}\mu_n(1 + 2x^2).$$

It is easily seen that on the set  $|x| \leq \varepsilon n^{\frac{1}{2}}, \varepsilon > 0$ , (small)  $\eta$  is a strictly increasing function w.p.  $1 - o(n^{-(1+\delta)})$ . Define  $A_n = \{|t_{1\alpha}^*| \leq \log n, |\mu_n| \leq c, |t_n| \leq \varepsilon n^{\frac{1}{2}}\}$  where  $C$  is sufficiently large. Note that  $P(A_n^c) = o(n^{-(1+\delta)})$ . On  $A_n, \eta(b_n) = t_{1\alpha}^* + a_n$ , where  $a_n = o(n^{-\frac{3}{2}}(\log n)^6)$ .

Using this, it can be easily seen after some algebra that

$$|P(t_n < b_n) - \alpha| = o(n^{-(1+\delta)}) + |P(H_{1n}^*(t_{1n} - a_n) < \alpha) - \alpha|.$$

The second term above equals

$$|P(H_{1n}(t_{1n}) < \alpha + H_{1n}(t_{1n}) - H_{1n}(t_{1n} - a_n) - (H_{1n}^*(t_{1n} - a_n) - H_{1n}(t_{1n} - a_n)))|$$

Noting that  $\sup_x |H_{1n}(x) - H_{1n}(x + \delta)| \leq C\delta + o(n^{-1})$ , and using Theorem 2.9, it follows that,

$$|P(t_n < b_n) - \alpha| = o(n^{-1}).$$

Hence

$$P(\mu > \hat{\mu} - \nu_n b_n) = \alpha + o(n^{-1}).$$

As is clear from the work of Hall (1986), the  $o(n^{-1})$  term, derived above cannot be uniform in  $\alpha, 0 < \alpha < 1$ . To derive bounds which hold uniformly in  $\alpha$ , we proceed as follows.

Define

$$\begin{aligned} \tilde{t}_{1\alpha}^* &= \max(-\log n, \min(t_{1\alpha}^*, \log n)) \\ \tilde{b}_n &= \tilde{t}_{1\alpha}^* - n^{-\frac{1}{2}}\mu_n(1 + 2\tilde{t}_{1\alpha}^{*2}) + \frac{4}{n}\mu_n^2(1 + 2\tilde{t}_{1\alpha}^{*2}). \end{aligned}$$

Proceeding as before, but using

$$\tilde{A}_n = \{|\tilde{t}_{1\alpha}^*| < \log n, |\mu_n| \leq c, |t_n| < 2\log n\}$$

it can be shown that for some  $a > 0$ ,

$$\sup_{0 < \alpha < 1} |P(t_{1n} < \tilde{t}_{1\alpha}^*) - \alpha| = o(n^{-1}(\log n)^a).$$

We omit the details, which involve use of the moderate deviation result in Lemma A3, Theorem 2.9 and Lemma 2.1 of Babu and Bose (1986). This shows that

$$\sup_{0 < \alpha < 1} |P(\mu > \hat{\mu} - \nu_n \tilde{b}_n) - \alpha| = o(n^{-1}(\log n)^a). \quad \blacktriangle$$



*Proofs of the Theorems:* Theorem 2.9 is a consequence of Theorems 2.1, 2.3 and equations (2.2) and (2.4). Theorem 2.1 follows by using the bound in Remark 2.5 and choosing  $(\varepsilon_n)$  and  $(\delta_n)$  suitably. We omit the details. Theorem 2.6 is essentially a Taylor series argument together with Theorems 2.1 and 2.2. Thus it is enough to prove Theorems 2.2 and 2.3. We will prove Theorem 2.3 below. Proof of Theorem 2.2 is similar except that we have to use the second parts of Lemma A1 and A5 since the Cramer's condition is replaced by s.n.l. We omit the details.

To prove Theorem 2.3, without loss of generality assume  $EX_i = 0$ . Denote  $\Delta_n^*(\delta) = \Delta_{n,s}(G_{T_n}, \delta)$ ,  $A^*(\delta) = A_s(G_{T_n}, \delta)$ ,  $\theta_n^* = \theta_n(G_{T_n})$ . (See appendix for the notations.) By Lemma A1 of appendix, on the set

$$(2.9) \quad n^{-(s-2)/2} \Delta_n^*(1) \leq 1/(8m), 0 < C_1 \leq \lambda_{T_n} \leq \Lambda_{T_n} \leq C_2 < \infty,$$

$$(2.10) \quad \begin{aligned} |B_{n,s}^*| &\leq C n^{-(s-2)/2} A^*(\delta) \left[ 1 + M_{s'} + \int \omega(f \circ T_n^{-1}, \exp(-dn), y) \exp(-y^2/100) dy \right] \\ &+ C(1 + M_{s'}) \left[ \exp(-dn) + n^{-(s-2)/2} \Delta_n^*(1) \right] \\ &+ C \sup_{0 \leq \alpha \leq s+m+1} n^{s+m} \exp(dn) (\theta_n^*)^{n-\alpha}, \end{aligned}$$

for some  $d > 0$  depending only on  $\theta_n^*$  given in (A.1).

To prove the theorem, we need probabilistic bounds for  $A^*(\delta)$ ,  $\Delta_n^*(1)$  and  $\theta_n^*$ . Recall that  $C_1 \leq \lambda_{T_n} \leq \Lambda_{T_n} \leq C_2$ . Note that

$$(2.11) \quad \begin{aligned} \theta_n^* &= P_{G_{T_n}} \left( \|X\| \geq n^{1/2} \right) + \sup \left\{ |\widehat{G}_{T_n}(t)| : (16\rho_3(G_{T_n}))^{-1} \leq \|t\| \leq \exp(dn) \right\} \\ &= \theta_{1n} + \theta_{2n}, \text{ say.} \end{aligned}$$

Fixing  $\alpha$ ,  $0 < \alpha < 1/2$ , we have

$$P(\theta_{1n} \geq \alpha) = P \left( n^{-1} \sum_{k=1}^n I \left( \|T_n(X_k - \bar{X}_n)\| \geq n^{1/2} \right) \geq \alpha \right)$$

$$\begin{aligned}
&\leq P\left(n^{-1} \sum_{k=1}^n I\left(\|X_k - \bar{X}_n\| \geq Cn^{1/2}\right) > \alpha\right) \\
&\leq \alpha^{-1} P\left(\|X_k\| + \|\bar{X}_n\| \geq Cn^{1/2}\right) \\
(2.12) \quad &\leq \alpha^{-1} \left[ P\left(\|X_k\| \geq Cn^{1/2}\right) + P\left(\|\bar{X}_n\| \geq Cn^{1/2}\right) \right].
\end{aligned}$$

By Markov's inequality and Lemma A3 with  $2 + \delta = s + \gamma$ , the right side of (2.12) is bounded by

$$(2.13) \quad C\alpha^{-1} \left[ n^{-(s+\gamma)/2} \Delta_{n,s+\gamma}(C, G) + o\left(n^{-(3s+3\gamma-2)/2}\right) \right].$$

Since  $X_1$  satisfies Cramer's condition, by Lemma A5,

$$(2.14) \quad P(\theta_{2n} \geq \alpha) \leq P(\rho_3(G_n) \geq C_3) + o\left(n^{-(s+\gamma-1)}\right).$$

On the other hand,

$$\begin{aligned}
A^*(\delta) &\leq \Delta_{n,s}(G_n, C\delta) + C\delta n^{-1} \sum_{k=1}^n \|X_k - \bar{X}_n\|^s I(\|X_k - \bar{X}_n\| \leq C\delta n^{1/2}) \\
&= A_{1\delta}^* + A_{2\delta}^* \text{ say.}
\end{aligned}$$

Thus

$$\begin{aligned}
(2.15) \quad P(A_{2\delta}^* \geq \varepsilon_n) &\leq P\left(n^{-1} \sum_{k=1}^n \|X_k\|^s \geq C\varepsilon_n/\delta_n\right) \\
&\quad + P\left(\|\bar{X}_n\| \geq C(\varepsilon_n/\delta_n)^{1/s}\right).
\end{aligned}$$

By the given condition, using Lemma A3, the second term is

$$(2.16) \quad o\left(n^{-(s+\gamma-1)}(\varepsilon_n/\delta_n)^{-(s+\gamma)/s}\right).$$

By Lemma A4 of appendix

$$\begin{aligned}
(2.17) \quad P(A_{1\delta}^* \geq \varepsilon_n) &\leq 2n^{-(s+\gamma-2)/4} \delta_n^{-(s+2\gamma)/2} \varepsilon_n^{-1/2} \Delta_{n,s+\gamma}(G, \delta_n) \\
&\quad + o\left(n^{-(s+\gamma-1)} \varepsilon_n^{-(s+\gamma)/s}\right).
\end{aligned}$$

Note that when  $\limsup_n \theta_n < 1$  the last term in (2.10) is exponentially small. Hence the theorem follows from (2.9), (2.10), (2.13), (2.15), (2.16) and (2.17).  $\blacktriangle$

## APPENDIX

We first list all the notations. Then we state all the Lemmas we used to prove the main results. The proofs in most cases are either omitted or merely sketched.

Suppose  $G$  is any distribution on  $\mathbf{R}^m$ ,  $m \geq 1$ , ( $G$  may be random) and  $X_i(G) = X_i = (X_{i1}, \dots, X_{im})$  are i.i.d. observations on  $G$ . Let  $D(G) = D$  be the dispersion matrix of  $X_1$  which is assumed to be positive definite. Let  $\alpha = (\alpha_1, \dots, \alpha_m)$  denote a multi-index with  $\alpha_i \geq 0$  as integers and let  $\mu_\alpha(G) = \mu_\alpha = E(X^\alpha) = E(X_{11}^{\alpha_1} \dots X_{1m}^{\alpha_m})$ . Let  $\chi_\alpha(G)$  be the  $\alpha^{th}$  cumulant of  $G$ . Suppose  $\rho_s(G) = E_G \|X_1\|^s$ . Let  $P_r(-\Phi_{0,V} : \{\chi_\nu\})$  denote a signed measure whose density is a polynomial multiple of the  $N(0, V)$  density and the coefficients of the polynomial are specific functions of the cumulants  $\{\chi_\nu\}$ . See Bhattacharya and Ranga Rao (1978), page 54 for details. Let

$$\psi_{n,k} = \sum_{r=0}^k n^{-r/2} P_r(-\Phi_{0,V} : \{\chi_\nu\}),$$

$$\Delta_{n,s}(G, \delta) = \Delta_n(G, \delta) = \Delta_n(\delta) = E_G \left( \|X_1\|^s I(\|X_1\| \geq \delta n^{1/2}) \right),$$

and 
$$A_s(G, \delta) = A(\delta) = \delta E_G \left( \|X_1\|^s I(\|X_1\| \leq \delta n^{1/2}) \right) + \Delta_n(\delta).$$

Suppose  $T$  is a positive definite matrix such that  $TD^{-1}T' = I$ .

$$\text{Let } S_n = \sum_{k=1}^n (X_k - E(X_1)), \quad \widehat{G}(t) = E_G \exp(it' X_1),$$

$$(A.1) \theta_n(G, d) = 2P_G \left( \|X_1\| \geq n^{1/2} \right) + \sup \left\{ |\widehat{G}(t)| : (16\rho_3(G))^{-1} \leq \|t\| \leq \exp(dn) \right\},$$

$$\alpha_n(G, \eta) = 2P_G \left( \|X_1\| \geq n^{1/2} \right) + \sup \{ |\widehat{G}(t)| : (16\rho_3(G))^{-1} \leq \|t\| \leq \eta^{-1}n \},$$

$$F_n(x) = n^{-1} \sum_{k=1}^n I(X_k \leq x), \text{ and } G_n(x) = F_n(x + \bar{X}_n),$$

where  $\bar{X}_n = n^{-1} \sum_{k=1}^n X_k$ . Let

$$G_T(x) = P_G(TX_1 \leq x), \quad Q_n(x) = P_G\left(n^{-1/2}S_n \leq x\right).$$

For any square symmetric matrix  $A$ ,  $\lambda_A$  (respectively  $\Lambda_A$ ) = minimum (resp. max.) eigen value of  $A$ . Recall that for  $f : \mathbf{R}^m \rightarrow \mathbf{R}$ ,

$$M_r(f) = \sup_x (1 + \|x\|^{-r})|f(x)| \text{ and}$$

$$w(f, \varepsilon, x) = \sup\{|f(z) - f(y)| : z, y \in B(x, \varepsilon)\},$$

where

$$B(x, \varepsilon) = \{y : \|y - x\| < \varepsilon\}.$$

Let  $C_i$ , denote generic constants. Their dependence on other quantities will be clear from the context.

Our first Lemma is the key Lemma used to prove the main results. The first part of the Lemma is a modified version of Theorem 20.1 of Bhattacharya and Ranga Rao (1976) (see page 208) and the second part is a modification of their Theorem 20.8 (see page 218).

**Lemma A1.** *Let  $X_1, \dots, X_n$  be i.i.d. with  $EX_1 = 0$ ,  $\rho_{s+\gamma} < \infty$  for some  $s \geq 3, \gamma \geq 0$ .*

*Whenever  $n^{-(s-2)/2}\Delta_n(1) \leq 1/(8m)$ , and  $0 < C_1 \leq \lambda_T \leq \Lambda_T < C_2 < \infty$ , then we have*

$$\begin{aligned} & \left| \int f d(Q_n - \sum_{r=0}^{s-2} n^{-r/2} P_r(-\Phi_{0,D} : \{\chi_\nu\})) \right| \\ & \leq C n^{-(s-2)/2} A(G_T, \delta) \left[ 1 + M_s(f) + \int \omega(f \circ T^{-1}, \exp(-dn), y) \exp(-y^2/100) dy \right] \\ & \quad + (1 + M_s) \left[ \exp(-dn) + n^{-(s-2)/2} \Delta_n(1) \right] \\ & \quad + C \sup_{0 \leq \alpha \leq s+m+1} n^{s+m} \exp(dn) (\theta_n(G_T))^{n-\alpha}. \end{aligned}$$

The constant  $C$  depends on  $C_1, C_2, s, m$  and  $d$  and is independent of  $G$ .

If  $f$  is a bounded function, then

$$\begin{aligned} \left| \int f d(Q_n - \psi_{n,1}) \right| &\leq C n^{-1/2} A_3(G_T, \delta) \left[ 1 + O(n^{-1}) + \omega_f(\mathbf{R}^m) + \int \omega(f, \eta n^{-1/2}, y) \phi(y) dy \right] \\ &\quad + C \sup_{0 \leq \alpha \leq m+1} n^{m+1} (\alpha_n(G_T, \eta))^{n-\alpha} \\ &\quad + C n^{-1/2} \Delta_{n,3}(1) \end{aligned}$$

The constant  $C$  depends on  $C_1, C_2, m$  and  $\eta$ . ▲

*Proof.* The result follows by carefully keeping track of all the bounds appearing in the proof of Theorem 20.1 and 20.8 of BR. We also need to use a lemma of Sweeting given below for bounded  $f$ , and its modification given in Babu and Singh (1984) for  $f$  with  $M_r(f) < \infty$ . These Lemmas replace the convolution Lemma used by BR to prove their theorems. ▲

Let  $K$  be a kernel on  $\mathbf{R}^m$  such that  $\alpha = K(\|x\| \leq 1) > 3/4$ . Let  $f$  be a measurable function on  $\mathbf{R}^m$  bounded by 1. Let  $K_\varepsilon(x) = K(\varepsilon x)$ .

**Lemma A2.** (*T.J. Sweeting*) We have

$$\begin{aligned} \left| \int f d(P - Q) \right| &\leq 2 \|K_{n^{1/2}\eta^{-1}} * (P - Q)\| + 2/n + 2K \left( \|x\| \geq n^{1/4}\eta^{-1} \right) \\ &\quad + \sup_{\|x\| \leq 2(\log n)n^{-1/4}} \int \omega(f, x - y, 2\eta n^{-1/2}) dQ(y). \end{aligned}$$

While dealing with the bootstrap expansions, the terms appearing on the right side of Lemma A1 are random. The random moment terms are controlled by the following Lemma due to Michel (1976). ▲

**Lemma A3.** Let  $(Z_i), i \geq 1$  be i.i.d. mean zero random variables with  $E|Z_i|^{2+\beta} < \infty$  for some  $\beta \geq 0$ . For any sequence  $t_n \geq 0$  such that  $t_n^2 - (\beta + 1)\log n \rightarrow \infty$ , we have

$$P(n^{-1/2} \left| \sum_{i=1}^n Z_i \right| \geq t_n) = o(n^{-\beta/2} t_n^{-(2+\beta)}). \quad \blacktriangle$$

The truncated moment  $\Delta_n$  appearing in Lemma A1 can be controlled by the following Lemma.

**Lemma A4.** *Let  $X_i, i \geq 1$  be i.i.d. with  $EX_i = 0$  and  $\rho_{s+\gamma} < \infty$  for some  $\gamma \geq 0$ . For all positive sequences  $(\varepsilon_n)$  and  $(\delta_n)$  between 0 and 1 such that  $n\varepsilon_n^{2/s} - C(s+\gamma+1)\log n \rightarrow \infty$  and  $n\delta_n^2 - C(s+\gamma+1)\log n \rightarrow \infty$ , we have*

$$(A.2) \quad P(\Delta_{n,s}(F_n, \delta_n) \geq \varepsilon_n) \leq 2n^{-(s+\gamma-2)/4} \delta_n^{-(s+2\gamma)/2} \varepsilon_n^{-1/2} \Delta_{\eta, s+\gamma}(G, \delta_n),$$

and

$$(A.3) \quad P(\Delta_{n,s}(G_n, \delta_n) \geq \varepsilon_n) \leq \text{the bound above} + o\left(n^{-(s+\gamma-1)} \varepsilon_n^{-(s+\gamma)/s}\right). \quad \blacktriangle$$

*Proof.* Define  $p_n = P(\|X_1\| \geq \delta n^{1/2})$

$$\begin{aligned} P(\Delta_{\eta,s}(F_n, \delta) \geq \varepsilon_n) &= P\left(\sum_{k=1}^n \|X_k\|^s I(\|X_k\| \geq \delta n^{1/2}) \geq n\varepsilon_n\right) \\ &= \sum_{k=1}^n \binom{n}{k} (1-p_n)^{n-k} P\left(\sum_{j=1}^k \|X_j\|^s I(\|X_j\| \geq \delta n^{1/2}) \geq n\varepsilon_n, \|X_j\| \geq \delta n^{1/2}, j = 1, \dots, k\right) \\ (A.4) \quad &\leq \sum_{k=1}^n \binom{n}{k} (1-p_n)^{n-k} p_n^{k/2} P^{1/2}\left(\sum_{j=1}^k \|X_j\|^s I(\|X_j\| \geq \delta n^{1/2}) \geq n\varepsilon_n\right). \end{aligned}$$

Note that  $P\left(\sum_{j=1}^k \|X_j\|^s I(\|X_j\| \geq \delta n^{1/2}) \geq n\varepsilon_n\right)$

$$\begin{aligned} &\leq (n\varepsilon_n)^{-1} k \Delta_{n,s}(G, \delta) \\ (A.5) \quad &\leq (n\varepsilon_n)^{-1} k \left(\delta n^{1/2}\right)^{-(s+\gamma)} \Delta_{n, s+\gamma}(G, \delta). \end{aligned}$$

Further,

$$(A.6) \quad \sum_{k=1}^m \binom{n}{k} k p_n^{k/2} (1-p_n)^{n-k} \leq 2n p_n^{1/2}$$

and

$$p_n \leq (\delta n^{1/2})^{-(s+\gamma)} \Delta_{n,s+\gamma}(G, \delta)$$

(A.2) now follows from (A.4), (A.5) and (A.6).

To prove (A.3), note that

$$(A.7) \quad \Delta_n(G_n, 2\delta) \leq C \|\bar{X}_n\|^s + C \Delta_n(F_n, \delta) + C n^{-1} \sum_{k=1}^n \|\bar{X}_n\|^s I\left(\|\bar{X}_n\| \geq \delta n^{1/2}\right).$$

By Lemma A3 with  $\beta = s + \gamma - 2$ ,  $\|X_i\| = Z_i$ , we have

$$(A.8) \quad P(\|\bar{X}_n\|^s \geq C \varepsilon_n) = o\left(n^{-(s+\gamma-1)} \varepsilon_n^{-(s+\gamma)/s}\right),$$

and

$$(A.9) \quad P\left(\|\bar{X}_n\| \geq \delta_n n^{1/2}\right) = o\left(n^{-(3s+3\gamma-2)/2} \delta_n^{-(s+\gamma)}\right).$$

However, by the conditions on  $\delta_n$ , the bound in (A.9) is dominated by the bound in (A.8). Thus the second part of the Lemma follows by using (A.7), (A.8) and (A.2).  $\blacktriangle$

The characteristic function terms appearing in Lemma A1 are controlled by the following Lemma.

**Lemma A5.** (i) If  $0 < \alpha < 1$ ,  $\delta = \alpha^2/300m$  and  $G$  satisfies Cramer's condition and  $\rho_{s+\gamma} < \infty$  for some  $\gamma \geq 0$  then

$$P\left(\sup_{0 < C \leq \|t\| \leq \exp(n\delta)} |\hat{F}_n(t)| \geq \alpha\right) = o\left(n^{-(s+\gamma-1)}\right).$$

(ii) If  $G$  is non lattice,  $\rho_{s+\gamma} < \infty$  for some  $\gamma \geq 0$  then for any  $\eta > 0$

$$P\left(\sup_{0 < C \leq \|t\| \leq \eta} |\hat{F}_n(t)| \geq \alpha\right) = o\left(n^{-(s+\gamma-1)}\right). \quad \blacktriangle$$

The proofs follow by modifying arguments given in the proof of Lemma 2 of Babu and Singh (1984).

### ACKNOWLEDGEMENTS

We are grateful to the referee, the Associate Editor and the Editor for their very helpful comments. Thanks are also due to Professor J.K. Ghosh of the Indian Statistical Institute, Calcutta, for his suggestions.

### REFERENCES

- Abramovitch, L. and Singh, K. (1985). Edgeworth corrected pivotal statistics and the bootstrap. *Ann. Statist.*, **13**, 116–132.
- Babu, G.J. and Bai, Z.D. (1990). Edgeworth expansions under minimal moment conditions. Tech. Report No. 90–17, Center for multivariate analysis, Department of Statistics, Pennsylvania State University, University Park, PA, 16802.
- Babu, G.J. and Bose, Arup (1988). Bootstrap confidence intervals. *Statistics and Probability Letters*, **7**, 151–160.
- Babu, G.J. and Singh, K. (1982). On  $r$ -quick limit sets for empirical and related processes based on mixing random variables. *J. Multivariate Analysis*, **12**, 508–525.
- Babu, G.J. and Singh, K. (1984). On one term Edgeworth correction by Efron's bootstrap. *Sankhyā, Series A*, **46**, 219–232.
- Bhattacharya, R.N. and Ghosh, J.K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.*, **6**, 434–451.
- Bhattacharya, R.N. and Qumsiyeh, M. (1989). Second order and  $L^p$  comparisons between the bootstrap and empirical Edgeworth expansion methodologies. *Ann. Statist.*, **17**, No. 1, 160–169.
- Bhattacharya, R.N. and Ranga Rao, R. (1976). *Normal approximation and asymptotic expansion*. Wiley, New York.
- Bickel, P.J. and Freedman, D. (1981). Some asymptotics on the bootstrap. *Ann. Statist.*, **9**, 1196–1217.
- Hall, P. (1986). On the bootstrap and confidence intervals. *Ann. Statist.*, **14**, 1431–1452.
- Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals (with discussion). *Ann. Statist.*, **16**, No. 1, 927–985.



- Johnson, N.J. (1978). Modified  $t$ -tests and confidence intervals for asymmetrical population. *J. Amer. Statist. Assoc.*, **73**, 536–544.
- Michel, R. (1976). Non uniform central limit bounds with applications to probabilities of deviation. *Ann. Prob.*, **4**, 102–106.
- Singh, K. (1981). On asymptotic accuracy of Efron's bootstrap. *Ann. Statist.*, **9**, 1187–1195.
- Sweeting, T. J. (1977). Speeds of convergence for the multidimensional central limit theorem. *Ann. Prob.*, **5**, 28–41.

ARUP BOSE  
DEPARTMENT OF STATISTICS  
PURDUE UNIVERSITY  
WEST LAFAYETTE, IN 47907  
email: ha0@l.cc.purdue.edu

G. J. BABU  
DEPARTMENT OF STATISTICS  
219 POND LABORATORY  
PENNSYLVANIA STATE UNIVERSITY  
UNIVERSITY PARK, PA 16802  
email: GJB6@PSUVM.BITNET