View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

JOURNAL OF MULTIVARIATE ANALYSIS 42, 226–244 (1992)

# Edgeworth Expansions for Errors-in-Variables Models

GUTTI JOGESH BABU*

*The Pennsylvania State University*

AND

Z. D. BAI[†]

*Temple University*

*Communicated by the Editors*

Edgeworth expansions for sums of independent but not identically distributed multivariate random vectors are established. The results are applied to get valid Edgeworth expansions for estimates of regression parameters in linear errors-in-variable models. The expansions for studentized versions are also developed. Further, Edgeworth expansions for the corresponding bootstrapped statistics are obtained. Using these expansions, the bootstrap distribution is shown to approximate the sampling distribution of the studentized estimators, better than the classical normal approximation. © 1992 Academic Press, Inc.

## 1. INTRODUCTION

Singh [17] has shown that the bootstrap approximation of the sampling distribution of the student's $t$-statistic is asymptotically better than the standard normal approximation. Using Edgeworth expansions, Babu and Singh [1, 2], showed for the first time that for a wide class of studentized statistics, bootstrap automatically corrects for skewness and hence gives a better approximation than the normal approximation. In the same spirit, Bose [8] showed that bootstrapping leads to a better approximation in the case of autoregressive processes. In all these results the maximum difference

Received September 11, 1990; revised January 10, 1992.

AMS 1980 subject classifications: 60F05, 62E20, 62J05.

Key words and phrases: bootstrap, Edgeworth expansions, errors-in-variable models, skewness.

* Research supported in part by NSA Grant MDA904-90-H-1001 and by NSF Grant DMS-9007717.

[†] Research supported by the Air Force Office of Scientific Research under Grant AFOSR-89-0279.



0047-259X/92 $5.00

Copyright © 1992 by Academic Press, Inc.
All rights of reproduction in any form reserved.

between the sampling distribution and the corresponding bootstrap distribution evaluated at the same point when multiplied by $\sqrt{n}$, tends to zero as $n \to 0$. In the case of i.i.d. random variables, the best one can do with normal approximation is

$$\sqrt{n}(P(\sqrt{n}(\bar{X}-\mu) \leqslant x s_n) - \Phi(x)) \to \mu_3(2x^2+1)\,\varphi(x)/6\sigma^3 \neq 0,$$

unless the third moment $\mu_3$ of $X$ is 0, where $\mu$ and $\sigma$ denote mean and variance of $X$, $s_n$ denotes the sample standard deviation, and $\Phi$ and $\varphi$ denote normal distribution and normal density. Due to this, bootstrap gives far superior result than the normal approximation.

In this paper we consider errors-in-variables regression models with homogeneous residuals and obtain Edgeworth expansions for the estimates of slope. To formulate the problem, consider the simple linear errors-in-variable (EIV) model $(X_i, Y_i)$,

$$X_i = U_{in} + \delta_i, \qquad Y_i = \omega + \beta U_{in} + \varepsilon_i, \tag{1.1}$$

where $(\delta_i, \varepsilon_i)$ are independent with $E(\delta_i) = E(\varepsilon_i) = 0$, and $U_i$ are unknown nuisance parameters. The EIV models have been studied extensively in the literature; see, among others, Kendall and Stuart [14], Gleser [11, 12], Fuller [10], Birch [7], York [18], Jones [13], and Madansky [16]. Initially, we concentrate on the case when $(\delta_i, \varepsilon_i)$ are independent copies of $(\delta, \varepsilon)$, $\lambda = \sigma_\varepsilon^2/\sigma_\delta^2$ is known, and $\delta$ and $\varepsilon$ are independent, where $\sigma_\varepsilon^2 = \mathrm{var}(\varepsilon)$ and $\sigma_\delta^2 = \mathrm{var}(\delta)$. It is well known that the least squares estimators of $\beta$ and $\omega$ are given by

$$\hat{\beta}_1 = \hat{h} + \mathrm{sign}(S_{XY})(\lambda + \hat{h}^2)^{1/2} \qquad \text{and} \qquad \hat{\omega}_1 = \bar{Y} - \hat{\beta}_1 \bar{X}, \tag{1.2}$$

where $\bar{X}$ and $\bar{Y}$ denote the sample means of $X_1, ..., X_n$ and of $Y_1, ..., Y_n$,

$$S_{XY} = \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}), \tag{1.3}$$

$$S_{XX} = \sum_{i=1}^{n} (X_i - \bar{X})^2, \qquad S_{YY} = \sum_{i=1}^{n} (Y_i - \bar{Y})^2, \tag{1.4}$$

and

$$\hat{h} = (S_{YY} - \lambda S_{XX})/2S_{XY}. \tag{1.5}$$

The least squares method gives the same estimates as in (1.2), when both $\sigma_\varepsilon$ and $\sigma_\delta$ are known. Instead, if $\sigma_\delta$ alone is known, $S_{XX} > n\sigma_\delta^2$ and $S_{YY} > S_{XY}^2/(S_{XX} - n\sigma_\delta^2)$, then the least squares estimators of $\beta$ and $\omega$ are given by

$$\hat{\beta}_2 = S_{XY}/(S_{XX} - n\sigma_\delta^2) \qquad \text{and} \qquad \hat{\omega}_2 = \bar{Y} - \hat{\beta}_2 \bar{X}. \tag{1.6}$$

On the other hand, if $\sigma_\varepsilon$ alone is known, $S_{XX} > S_{XY}^2/(S_{YY} - n\sigma_\varepsilon^2)$, and $S_{YY} > n\sigma_\varepsilon^2$, then the least squares estimators of $\beta$ and $\omega$ are given by

$$\hat{\beta}_3 = (S_{YY} - n\sigma_\varepsilon^2)/S_{XY} \qquad \text{and} \qquad \hat{\omega}_3 = \bar{Y} - \hat{\beta}_3 \bar{X}. \tag{1.7}$$

See Fuller [10] and Jones [13].

Even though the residuals in the EIV model considered here are assumed to be i.i.d. random variables, the statistics of interest turn out to be functions of means of independent but not identically distributed random vectors. This is mainly due to the large number of nuisance parameters. A result on Edgeworth expansions for independent but non-identically distributed random vectors is established in Section 2. Using these, two-term Edgeworth expansions for $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ are derived in that section. In Section 3, the bootstrap estimator of the sampling distributions of $\hat{\beta}_i$ are shown to correct for the skewness. Though the final results on Studentized versions of the bootstrap approximation are not entirely surprising in view of the results of Singh [17] and Babu and Singh [1, 2], some effort is required to deduce the results in the nonstatinary case.

The methods developed here can be extended to the case of known $\lambda = \text{var}(\varepsilon_i)/\text{var}(\delta_i)$ for all $i$, when $(\delta_i, \varepsilon_i)$ are independent, but not identically distributed.

## 2. The Main Results

### 2.1. *Notation and Main Assumptions*

Let $(\delta, \varepsilon)$ be a random vector and let $F_\delta$ and $F_\varepsilon$ denote, respectively, the conditional distributions of $\varepsilon$ given $\delta$ and of $\delta$ given $\varepsilon$.

*Assumption* 1. $E(\delta) = E(\varepsilon) = 0$, $E(\delta^6 + \varepsilon^6) < \infty$,

$$P\{\delta > 0 \colon F_\delta \text{ is not purely discrete}\} > 0,$$

$$P\{\delta < 0 \colon F_\delta \text{ is not purely discrete}\} > 0,$$

$$P\{\varepsilon > 0 \colon F_\varepsilon \text{ is not purely discrete}\} > 0,$$

and

$$P\{\varepsilon < 0 \colon F_\varepsilon \text{ is not purely discrete}\} > 0.$$

*Remark.* Assumption 1 holds in particular if $\varepsilon$ and $\delta$ are independent and continuous random variables. Assumption 1 is essentially used to establish the strongly non-lattice structure of the distribution of $\xi_*$, defined below. The following example shows that Assumption 1 can not be weakened substantially, even when $\varepsilon$ and $\delta$ are independent.

EXAMPLE. Let $\delta$ and $\varepsilon$ be i.i.d. and $\varepsilon$ takes values 0, 1, and $\sqrt{2}$ with probability $\frac{1}{3}$. Then $\varepsilon$ is non-lattice. If

$$a = c = -1/\sqrt{2}, \qquad d = f = 1 + 1/\sqrt{2}, \qquad \text{and} \quad b = 0,$$

then

$$\zeta = a\varepsilon^2 + b\varepsilon\delta + c\delta^2 + d\varepsilon + f\delta$$

is lattice, i.e.,

$$P(\zeta = 0) = \tfrac{1}{9}, \qquad P(\zeta = 1) = P(\zeta = 2) = \tfrac{4}{9}.$$

The characteristic function of $(\varepsilon^2, \delta^2, \varepsilon\delta, \varepsilon, \delta, u\varepsilon, u\delta\}$ is 1 in absolute value at $\mathbf{t}' = (a, a, 0, d, d, 0, 0)$ for all $u$. This violates (2.4) below, which is needed to estimate the error term in the expansion. As is well known, the formal Edgeworth expansions are not valid for the means of independent copies of the vectors $(\varepsilon^2, \delta^2, \varepsilon, \delta)$, since the distribution of this vector is lattice.

*Assumption* 2. For each $n$, the sequence $u_{1n}, ..., u_{nn}$ of constants satisfies

$$\text{(i)} \quad \sum u_{jn} = 0,$$

$$\text{(ii)} \quad s_n^2 = \frac{1}{n} \sum u_{jn}^2 \to \mu > 0,$$

and

$$\text{(iii)} \quad \sup_n \frac{1}{n} \sum |u_{jn}^3| \leqq M < \infty.$$

These conditions on $u_{jn}$ are needed to show that the dispersion of $\sqrt{n}\, \xi_n$ below converges and that its third order moments are not too large.

*Throughout this paper, let* $\{(\delta_j, \varepsilon_j),\ j = 1, 2, ..., n\}$ *denote a sequence of* i.i.d. *samples drawn from* $(\delta, \varepsilon)$, *and we use the notation* $g_n \ll h_n$ *to denote* $g_n = O(h_n)$.

For ease of notation, *we drop the subscript n from* $u_{in}$ *and from* $\xi_{jn}$ defined below. Let

$$\xi_j = \xi_{jn} = (\varepsilon_j^2 - E\varepsilon_j^2, \delta_j^2 - E\delta_j^2, \varepsilon_j\delta_j, \varepsilon_j, \delta_j, u_j\varepsilon_j, u_j\delta_j)'.$$

and

$$\xi_* = (\varepsilon^2 - E\varepsilon^2, \delta^2 - E\delta^2, \varepsilon\delta, \varepsilon, \delta)',$$

for $j = 1, \ldots, n$. Let $A$ denote the dispersion matrix of $\xi_*$ and $\Sigma$ denote the dispersion matrix of $(\delta, \varepsilon)$. Then the dispersion matrix $B_n$ of

$$\sqrt{n}\,\xi_n = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\xi_j\right)$$

is given by the partition matrix

$$B_n = \begin{pmatrix} A & 0 \\ 0 & s_n^2\Sigma \end{pmatrix} \to B = \begin{pmatrix} A & 0 \\ 0 & \mu\Sigma \end{pmatrix},$$

as $n \to \infty$. Let $\lambda_1$ and $\lambda_2$ denote the smallest eigenvalues of $A$ and $\Sigma$, respectively. Then for all large $n$, the smallest eigenvalue of $B_n$ is not less than $b = \frac{1}{2}\min(\lambda_1, \frac{1}{2}\mu\lambda_2) > 0$. Consequently, $B_n - bI_7$ is a positive definite matrix for all large $n$.

### 2.2. *Edgeworth expansions*

Let $G_n$ and $Q_n$ denote respectively the distribution and the formal two-term Edgeworth expansion of $\sqrt{n}\,\xi_n$. We now state the main theorem.

THEOREM 1. *Let $f$ be any measurable function bounded by* 1. *Under the Assumptions* 1 *and* 2,

$$\int f d(G_n - Q_n) \leqslant \theta_n + \int (\sup\{|f(y) - f(x)| : |x - y| \leqq \theta_n\})\, \varphi_{B_n}(x)\, dx,$$

*where $\theta_n = o(n^{-1/2})$ and $\varphi_{B_n}$ denotes the density function of a normal random vector with mean zero and dispersion $B_n$.*

To prove the theorem, we need the following lemmas. Lemma 1 guarantees existence of a large number of "good" $u_j$'s. Using only these and dropping the rest, a suitable upper bound for the characteristic function of $n\xi_n$ is obtained in Lemma 2. This bound is similar to the one required in the standard proofs for the i.i.d. random vectors (see Bhattacharya and Ghosh [4]). These two lemmas form part of the main contributions of this paper. Lemma 3 gives a bound for the differences of probabilities in terms of the derivatives of the characteristic functions. Finally, Lemma 4 estimates the differences of the derivatives of the characteristic functions.

Let $N_{d,D}^{\pm} = \#\{j \leqq n : d < \pm u_j \leqslant D\}$ and $N_{d,D} = N_{d,D}^{+} + N_{d,D}^{-}$ for any $0 < d < D < \infty$.

LEMMA 1. *Under Assumption* 2, *there exist constants $0 < d < D < \infty$ and $\eta > 0$, such that*

$$N_{d,D}^{\pm} \geqq \eta n.$$

*Proof.* We need only to prove the lemma for $N^+_{d,D}$. Proof for $N^-_{d,D}$ is similar. For fixed $\tau > 0$, when $n$ is large enough, we have by (ii) that

$$D_1^2 N_{d_1,D_1} + nd_1^2 + \sum_{j=1}^{n} I_{[|u_j| > D_1]} u_j^2 \geqq n(\mu - \tau).$$

Thus, by (iii),

$$N_{d_1,D_1} \geqq D_1^{-2} n(\mu - \tau - d_1^2 - M/D_1). \tag{2.1}$$

By (i)–(iii) we also have

$$d_1 N_{d_1,D_1} \leqq 2 \sum u_j I[u_j > 0] \leqq 2(d + M/D^2)n + 2DN^+_{d,D}. \tag{2.2}$$

From (2.1) and (2.2), it follows that

$$N^+_{d,D} \geqq \tfrac{1}{2} nD^{-1}(d_1 D_1^{-2}(\mu - \tau - d_1^2 - M/D_1) - 2(d + M/D^2)). \tag{2.3}$$

Lemma 1 is proved by choosing $D_1$ large, $\tau$ and $d_1$ small enough, and then choosing $D$ large and $d$ small (for example, by letting $\tau = d_1^2 = M/D_1 = \mu/4$ and $d = M/D^2 = \mu d_1/(20D_1^2)$, then we can choose $\eta = \mu d_1/(40\mathrm{DD}_1^2)$). This completes the proof.

LEMMA 2. *Under Assumptions* 1 *and* 2, *for any* $0 < 2a < b < \infty$, *there exists a constant* $\rho = \rho(a, b) < 1$ *such that*

$$\sup_{2a \leqq |\mathbf{t}| \leqq b} \left| \prod_{j=1}^{n} E(\exp\{i\mathbf{t}'\xi_j\}) \right| \leqq \rho^n. \tag{2.4}$$

*Proof.* Let $\mathbf{t} = (t_1, ..., t_7)'$, $\mathbf{t}_s = (t_1, t_3, t_4, t_6)'$, $\mathbf{t}_r = (t_2, t_3, t_5, t_7)'$, $\xi_j^{(s)} = (\varepsilon_j^2, \varepsilon_j\delta_j, \varepsilon_j, u_j\varepsilon_j)'$, and $\xi_j^{(r)} = (\delta_j^2, \varepsilon_j\delta_j, \delta_j, u_j\delta_j)'$. We shall prove (2.4) for the cases $a \leqq |\mathbf{t}_s| \leqq b$ and $a \leqq |\mathbf{t}_r| \leqq b$, respectively.

We first consider the case where $a \leqq |\mathbf{t}_s| \leqq b$. Without loss of generality, we can assume that $t_4 \geqq 0$. Write $\mathscr{E} = \{j \leqq n\colon t_6 u_j \geqq 0,\ d \leqq |u_j| \leqq D\}$, and $N = \#\{\mathscr{E}\}$, where $d$ and $D$ are as in Lemma 1. Then we have $N \geqq \eta n$.

If $j \in \mathscr{E}$, $t_3\delta_j \geqq 0$, $\delta_j \neq 0$, then $t_1$ and $t_3\delta_j + t_4 + t_6 u_j$ cannot vanish simultaneously. In this case, if

$$P(\mathbf{t}'\xi_j^{(s)} = v + wk,\ k = 1, 2, ..., \text{for some } v \text{ and } w) = 1,$$

then it follows that $F_{\delta_j}$ concentrates on the countable set

$$\{x\colon t_1 x^2 + (t_3\delta_j + t_4 + t_6 u_j)x = v + wk,\ k = 1, 2, ..., \text{for some } v \text{ and } w\}.$$

This implies that, in addition, if $F_{\delta_j}$ is not purely discrete, then

$$|E(\exp\{i\mathbf{t}'_s\xi_j^{(s)}\} \mid \delta_j)| < 1.$$

Note that

$$P\{t_3\delta_j \geqq 0, \delta_j \neq 0 \text{ and } F_{\delta_j} \text{ is not purely discrete}\} > 0.$$

Hence we get for $j \in \mathscr{E}$,

$$|E(\exp\{i\mathbf{t}'\xi_j\})| \leqq E\,|E(\exp\{i\mathbf{t}'_s\xi_j^{(s)}\}\mid \delta_j)| < 1.$$

Similarly, if $a \leqq |\mathbf{t}_r| \leqq b$, we can assume that $t_5 \geqq 0$ and define $\mathscr{E}^* = \{j \leqq n : t_7 u_j \geqq 0,\ d \leqq |u_j| \leqq D\}$. Hence when $j \in \mathscr{E}^*$ we have

$$|E(\exp\{i\mathbf{t}'\xi_j\})| \leqq E\,|E(\exp\{i\mathbf{t}'_r\xi_j^{(r)}\}\mid \varepsilon_j)| < 1.$$

Since $E(\exp(i\mathbf{t}'\xi_j))$ is a continuous function of $\mathbf{t}$ and $u_j \in \{x : d < |x| < D\}$, there exists a constant $\rho_1 \in (0, 1)$ such that

$$|E(\exp\{i\mathbf{t}'\xi_j\})| \leqq \rho_1, \qquad \text{for all} \quad j \in \mathscr{E} \text{ and } a \leqq |\mathbf{t}| \leqq b.$$

This establishes the lemma with $\rho = \rho_1^\eta$.

LEMMA 3 (Lemma 1 of Babu and Singh[3]). *Let $P$ be a probability on $R^k$ and $Q$ denote a signed measure with density $[1 + n^{-1/2}p]\varphi_V$, where $p$ is a polynomial and $V$ is a positive definite matrix. Let $\gamma_1$, $\gamma_2^{-1}$, and the coefficients of $p$ be bounded by $N > 0$, where $\gamma_1$ and $\gamma_2$ denote the maximum and the minimum eigenvalues of $V$. Then for any bounded real valued measurable function $f$ and $\theta$ positive,*

$$\left|\int f d(P-Q)\right| \leqq c(k) \max_{|\alpha| \leqq k+1} \int_{\|\mathbf{t}\| \leqq c\theta^{-1}\sqrt{n}} |D^\alpha(\hat{P} - \hat{Q})(\mathbf{t})|\, d\mathbf{t}$$

$$+ c(N) \int (\sup\{|f(y) - f(x)| : |y - x| \leqq \theta n^{-1/2}\})$$

$$\times \varphi_V(\mathbf{x})\, d\mathbf{x} + o(n^{-1/2}).$$

*Here $\hat{P}$ and $\hat{Q}$ stand for the characteristic functions of $P$ and $Q$, and $c(k)$, $c$, and $c(N)$ are positive constants. These constants and $o(n^{-1/2})$ term depend upon $f$ only through its bound.*

Let $V_n$ be a symmetric positive definite matrix satisfying $V_n^2 = B_n^{-1}$. Let $G_n^*$ and $Q_n^*$ denote, respectively, the distribution and the two-term formal Edgeworth expansion of $\sqrt{n}(\overline{\xi_n^*} - E(\overline{\xi_n^*}))$, where

$$\overline{\xi_n^*} = \frac{1}{n} \sum_{i=1}^{n} \xi_j I[|\xi_j| \leqq \sqrt{n}].$$

Some of the arguments in the proofs of Lemma 4 and Corollary 1 below depend on results and arguments of Bhattacharya and Rao [5]. To avoid repetition and making the proofs unnecessarily long, the reader is referred to the appropriate results of Bhattacharya and Rao [5] at various places.

LEMMA 4. *Suppose Assumptions* 1 *and* 2 *hold. Then for any* $\alpha = (\alpha_1, ..., \alpha_7)$, *with* $\alpha_i \geqq 0$ *and* $|\alpha| = \sum_{i=1}^{7} \alpha_i \leqq 8$, *and for any* $|\mathbf{t}| \leqq n^{1/12}$, *we have*

$$|D^\alpha(\hat{G}_n^*(V_n\mathbf{t}) - \hat{Q}_n^*(V_n\mathbf{t}))| \ll \beta_n n^{-1/2}(|\mathbf{t}|^{8-|\alpha|} + |\mathbf{t}|^{18+|\alpha|}) \exp\{-\tfrac{1}{4}|\mathbf{t}|^2\}, \tag{2.5}$$

*where* $\hat{G}_n^*$ *and* $\hat{Q}_n^*$ *denote the characteristic functions of* $G_n^*$ *and* $Q_n^*$, *and*

$$\beta_n = n^{-7/2} \sum_{j=1}^{n} E(|V_n\xi_j|^8 I[|\xi_j| \leqq \sqrt{n}]) \to 0. \tag{2.6}$$

*Further, for all* $|\mathbf{t}| \leqq c\sqrt{n}$, *for some* $c > 0$,

$$|D^\alpha(\hat{G}_n^*(V_n\mathbf{t}))| \ll (1 + |\mathbf{t}|^{|\alpha|}) \exp(-(\tfrac{5}{24})|\mathbf{t}|^2). \tag{2.7}$$

*Proof.* Note that for some $c > 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} E|\xi_j|^3 \leqq c\{E\varepsilon^6 + E\delta^6 + ME|\varepsilon|^3 + ME|\delta|^3\} < \infty, \tag{2.8}$$

where $M$ is the constant defined in Assumption 2. The inequality (2.7) now follows from Corollary 14.4 of Bhattacharya and Rao [5].

To prove (2.5), first note that for all $j$ and $n$, Assumption 2 implies

$$|u_j| \leqq (Mn)^{1/3}. \tag{2.9}$$

Hence for $|\mathbf{s}| < n^{-5/12}$,

$$\begin{aligned}|\hat{g}_n^*(V_n\mathbf{s}) - 1| &\leqq \tfrac{1}{2}E(s'V_n(\xi_j^* - E\xi_j^*))^2 \\ &\ll |s|^2 E|\xi_j|^2 \ll |s|^2 u_j^2 \ll n^{-1/6} \to 0,\end{aligned}$$

where $\hat{g}_n^*(\mathbf{s}) = E(\exp\{i\mathbf{s}'(\xi_j^* - E\xi_j^*)\})$. Thus, for all large $n$ and $|\mathbf{t}| \leqq n^{1/12}$, we have

$$\sup_{1 \leqq j \leqq n} |\hat{g}_j^*(V_n\mathbf{t}n^{-1/2}) - 1| \leqq \tfrac{1}{2}. \tag{2.10}$$

The inequality (2.5) now follows from Theorem 9.9 of Bhattacharya and Rao [5].

To prove (2.6), we have by (2.8) and (2.9) that

$$\begin{aligned}\beta_n &\ll n^{-7/2}\sum_{j=1}^{n} E\left(|\xi_j|^8 I[|\xi_j| \leqq \sqrt{n}]\right)\\ &\ll n^{-7/2}\sum_{j=1}^{n} E\left(|\xi_j|^8 I[|\xi_j| \leqq n^{5/12}]\right)\\ &\quad + n^{-1}\sum_{j=1}^{n} E\left(|\xi_j|^3 I[|\xi_j| \geqq n^{5/12}]\right)\\ &\ll n^{-17/12}\sum_{j=1}^{n} E|\xi_j|^3\\ &\quad + n^{-1}\sum_{j=1}^{n} E\left(|\xi_j|^3 I[\varepsilon_j^2+\delta_j^2 \geqq cn^{1/6}]\right) \to 0,\end{aligned}$$

for some $c>0$. This completes the proof.

*Proof of Theorem* 1. We have for any (measurable real) function $f$ bounded by 1,

$$\begin{aligned}&|E[f(\sqrt{n}(\xi_n))] - E[f(\sqrt{n}(\xi_n^*))]|\\ &\quad\ll 2\sum_{j=1}^{n} P(|\xi_j| \geqq \sqrt{n}/2)\\ &\quad\ll n^{-3/2}\sum_{j=1}^{n} E\left(|\xi_j|^3 I(|\xi_j| > \sqrt{n}/2)\right) = o(n^{-1/2}). \qquad (2.11)\end{aligned}$$

The last equality follows from the proof of (2.6).

By using the frist three moments of $\xi_n$ and $\xi_n^*$ we get

$$\int fd(Q_n - Q_n^*) = o(n^{-1/2}) \quad \text{and} \quad \int f dQ_n^* - \int f_{a_n} dQ_n^* = o(n^{-1/2}), \qquad (2.12)$$

where $a_n = -\sqrt{n}\, E\xi_n^* = o(n^{-1/2})$ and $f_a(x) = f(x+a)$.

The theorem now follows from (2.11), (2.12), Lemmas 2, 3, and 4, as the eigenvalues of $V_n$ and $V_n^{-1}$ are bounded.

COROLLARY 1. *Suppose $\varepsilon$ and $\delta$ are independent and continuous random variables satisfying $E(\varepsilon) = E(\delta) = 0$ and $E(\varepsilon^6 + \delta^6) < \infty$. Then under Assumption* 2, *$\sqrt{n}(\hat{h} - h)$ and $\sqrt{n}(\hat{\beta}_1 - \beta)$ both have valid two-term Edgeworth expansions, where*

$$h = (2\beta)^{-1}(\beta^2 - \lambda) \quad \textit{and} \quad \lambda = \sigma_\varepsilon^2/\sigma_\delta^2.$$

*Proof.* It is not difficult to show that

$$2(\hat{h}-h)=s_n^{-2}g_1(\xi_n)-s_n^{-4}g_2(\xi_n)+r_n,$$

where for any vector $\mathbf{z}=(z_1, ..., z_7)'$,

$$g_1(\mathbf{z})=\beta^{-1}(z_1-\lambda z_2-2hz_3+2\gamma z_6-2\gamma\beta z_7-z_4^2+\lambda z_5^2+2hz_4z_5),$$
$$g_2(\mathbf{z})=\beta^{-2}(z_1-\lambda z_2-2hz_3+2\gamma z_6-2\gamma\beta z_7)(z_3+z_6+\beta z_7),$$
$$\gamma=(2\beta)^{-1}(\beta^2+\lambda)=\beta-h,$$

and

$$\sqrt{n}\,|r_n|\ll\sqrt{n}\,|\xi_n|^3\leqq n^{-2/3},\qquad \text{when}\quad |\xi_n|\leqq n^{-7/18}. \tag{2.13}$$

By Bikelis' [6] inequality we have

$$P\{|\xi_n|\geqq n^{-7/18}\}\leqq 2\Phi(-n^{1/9})+cn^{-3/2}\sum_{j=1}^{n}E\left(|\xi_j|^3/(1+n^{1/9})^3\right)$$
$$=o(n^{-1/2}). \tag{2.14}$$

Since $g_1$ and $g_2$ are polynomials of degree two,

$$\hat{\beta}_1=\hat{h}+\text{sign}(S_{XY})(\lambda+\hat{h}^2)^{1/2},$$

and $P[\text{sign}(S_{XY})\neq\text{sign}(\beta)]=O(n^{-2})$, the corollary follows from (2.13) and (2.14) as in Lemma 2.1 of Bhattacharya and Rao [5].

COROLLARY 2. *Suppose $\varepsilon$ and $\delta$ are independent and continuous random variables satisfying $E(\varepsilon)=E(\delta)=0$ and $E(\varepsilon^6+\delta^6)<\infty$. Then under Assumption 2, $\sqrt{n}\,(\hat{\beta}_2-\beta)$ and $\sqrt{n}(\hat{\beta}_3-\beta)$ both have valid two-term Edgeworth expansions.*

*Proof.* Note that

$$\hat{\beta}_2-\beta=s_n^{-2}g_3(\xi_n)-s_n^{-4}g_4(\xi_n)+r_{n2},$$
$$\hat{\beta}_3-\beta=s_n^{-2}g_5(\xi_n)-s_n^{-4}g_6(\xi_n)+r_{n3}, \tag{2.15}$$

where for any vector $\mathbf{z}=(z_1, z_2, ..., z_7)'$,

$$g_3(\mathbf{z})=z_3-\beta z_2-z_4z_5+\beta z_5^2+z_6-\beta z_7,$$
$$g_4(\mathbf{z})=(z_2+2z_7)(z_3-\beta z_2+z_6-\beta z_7),$$
$$g_5(\mathbf{z})=\beta^{-1}(z_1-\beta z_3+\beta z_6-\beta^2 z_7-z_4^2+\beta z_4z_5),$$
$$g_6(\mathbf{z})=\beta^{-2}(z_1-\beta z_3+\beta z_6-\beta^2 z_7)(z_3+z_6+\beta z_7),$$

and

$$\sqrt{n}\,|r_{ni}| \ll \sqrt{n}\,|\xi_n|^3 \qquad \text{when} \quad |\xi_n| \leqq n^{-7/18},$$

for $i=2$ and 3. The rest of the proof is similar to that of Corollary 1.

2.3. *The Expressions for the Edgeworth Expansions of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$*

Let

$$\Gamma_n = \frac{1}{n}\sum_{i=1}^{n} u_i^3,$$

$$S_1 = \varepsilon^2 - \lambda\delta^2 - 2h\varepsilon\delta, \qquad S_{1j} = \varepsilon_j^2 - \lambda\delta_j^2 - 2h\varepsilon_j\delta_j,$$

$$S_2 = \varepsilon\delta - \beta(\delta^2 - \sigma_\delta^2), \qquad S_{2j} = \varepsilon_j\delta_j - \beta(\delta_j^2 - \sigma_\delta^2),$$

$$S_3 = \varepsilon^2 - \sigma_\varepsilon^2 - \beta\varepsilon\delta, \qquad S_{3j} = \varepsilon_j^2 - \sigma_\varepsilon^2 - \beta\varepsilon_j\delta_j,$$

$$S_4 = \varepsilon - \beta\delta, \quad \text{and} \quad S_{4j} = \varepsilon_j - \beta\delta_j.$$

The formal two-term Edgeworth expansions $Q_{ni}$ of $\sqrt{n}\,(\hat{\beta}_i - \beta)$, $i = 1, 2, 3$, which are shown to be valid by Corollaries 1 and 2, are given by

$$Q_{ni}(x) = \Phi(x/\sigma_i) + n^{-1/2}\{\alpha_{i1} - \sigma_i^{-2}\alpha_{i3}(x^2\sigma_i^{-2} - 1)\}\,\varphi(x/\sigma_i)\,\sigma_i^{-1}.$$

The expressions for $\alpha_{i\kappa}$ and $\sigma_i$, $i = 1, 2, 3$ and $\kappa = 1, 3$, can be computed using approximate cumulants and are given by

$$\sigma_1^2 = (2\gamma s_n^2)^{-2}\,\{E(S_1^2) + 8\beta\gamma^3 s_n^2\sigma_\delta^2\}, \tag{2.16}$$

$$\alpha_{11} = 2s_n^{-4}h\sigma_\delta^2(\lambda\sigma_\delta^2(\beta^2+\lambda)^{-1} + s_n^2) + \lambda(2\gamma)^{-1}\,\sigma_1^2,$$

$$\alpha_{13} = (2\gamma s_n^2)^{-3}\,[E(S_1^3) + 8\gamma^3\Gamma_n E(S_4^3) + 12 s_n^2\gamma^2 E(S_1 S_4^2)],$$

$$\sigma_2^2 = s_n^{-4}\{E(S_2^2) + s_n^2(\sigma_\varepsilon^2 + \beta^2\sigma_\delta^2)\}, \tag{2.17}$$

$$\alpha_{21} = \beta\sigma_\delta^2 s_n^{-2} + \beta s_n^{-4}(E(\delta^2 - \sigma_\delta^2)^2 + 2s_n^2\sigma_\delta^2)\},$$

$$\alpha_{23} = s_n^{-6}\{E(S_2^3) + \Gamma_n E(S_4^3) + 3s_n^2 E(S_2 S_4^2)\},$$

$$\sigma_3^2 = \beta^{-2}s_n^{-4}\{E(S_3^3) + \beta^2 s_n^2(\sigma_\varepsilon^2 + \beta^2\sigma_\delta^2)\}, \tag{2.18}$$

$$\alpha_{31} = \beta^{-1}s_n^{-4}(\sigma_\varepsilon^2\sigma_\delta^2 + \beta^2 s_n^2\sigma_\delta^2 - 2s_n^2\sigma_\varepsilon^2),$$

and

$$\alpha_{33} = \beta^{-3}s_n^{-6}\{E(S_3^3) + \beta^3\Gamma_n E(S_4^3) + 3\beta^2 s_n^2 E(S_3 S_4^2)\}.$$

2.4. *Estimators of the Asymptotic Variances of $\sqrt{n}\,\hat{\beta}_1$, $\sqrt{n}\,\hat{\beta}_2$, and $\sqrt{n}\,\hat{\beta}_3$*

The asymptotic variances $\sigma_j^2$ of $\sqrt{n}\,\hat{\beta}_j$, $j = 1, 2, 3$, are derived in (2.16)–(2.18). They involve unknown parameters. In this subsection we shall propose consistent estimators for these variances and derive their asymptotic properties.

For simplicity of writing, we shall use the notation $R_n \sim T_n$, whenever

$$\sqrt{n}P\{|R_n - T_n| \geq \varepsilon/\sqrt{n}\} \to 0, \qquad \text{for any} \quad \varepsilon > 0, \text{ as } n \to \infty,$$

and the notation $R_n \ll T_n$ to indicate the existence of an $N > 0$ such that

$$P\{R_n \geq NT_n\} = o(n^{-1/2}), \qquad \text{as} \quad n \to \infty.$$

The jackknife-type arguments lead to the estimators

$$\hat{\sigma}_1^2 = n\hat{\beta}_1^2(4S_{XY}^2(\hat{h}^2 + \lambda))^{-1} \times \sum_{i=1}^{n} [(Y_i - \bar{Y})^2 - \lambda(X_i - \bar{X})^2 - 2\hat{h}(X_i - \bar{X})(Y_i - \bar{Y})]^2, \tag{2.19}$$

$$\hat{\sigma}^2 = n(S_{XX} - n\sigma_\delta^2)^{-2} \sum_{i=1}^{n} [(X_i - \bar{X})(Y_i - \bar{Y} - \hat{\beta}_2(X_i - \bar{X})) + \hat{\beta}_2\sigma_\delta^2]^2, \tag{2.20}$$

and

$$\hat{\sigma}_3^2 = nS_{XY}^{-2} \sum_{i=1}^{n} [(Y_i - \bar{Y})(Y_i - \bar{Y} - \hat{\beta}_3(X_i - \bar{X})) - \sigma_\varepsilon^2]^2, \tag{2.21}$$

of $\sigma_1^2$, $\sigma_2^2$, and $\sigma_3^2$, respectively. To obtain their asymptotic properties let

$$\hat{R}_1 = \frac{1}{n} \sum_{i=1}^{n} \{[(Y_i - \bar{Y})^2 - \lambda(X_i - \bar{X})^2 - 2h(X_i - \bar{X})(Y_i - \bar{Y})]^2 - 4\sigma_1^2 s_n^4 \gamma^2\}, \tag{2.22}$$

and

$$\hat{R}_2 = 8h(\hat{h} - h)[\lambda\sigma_\delta^4 + 2s_n^2\sigma_\delta^2\beta\gamma)]. \tag{2.23}$$

By Corollary 1 we have $n^{1/3}|\hat{h} - h|^2 \sim 0$ and $n^{-1/4}|\hat{h} - h| \sim 0$. Since $|u_i| \leq (Mn)^{1/3}$, we have

$$n^{-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2 (X_i - \bar{X})^2 \ll \frac{1}{n} \sum_{i=1}^{n} [(1 + \beta^2) u_i^4 + (\varepsilon_i - \bar{\varepsilon})^4 + (\delta_i - \delta)^4] \ll n^{1/3}, \tag{2.24}$$

and

$$\left| n^{-1} \sum_{i=1}^{n} \{[(Y_i - \bar{Y})^2 - \lambda(X_i - \bar{X})^2 - 2h(X_i - \bar{X})(Y_i - \bar{Y})](X_i - \bar{X})(Y_i - \bar{Y}) + 2h\sigma_\delta^2[\lambda\sigma_\delta^2 + 2\beta\gamma s_n^2]\} \right| \ll n^{-1/4}. \tag{2.25}$$

Consequently,

$$n^{-1}\sum_{i=1}^{n}[(Y_i-\bar{Y})^2-\lambda(X_i-\bar{X})^2-2\hat{h}(X_i-\bar{X})(Y_i-\bar{Y})]^2$$
$$\sim 4\sigma_1^2 s_n^4\gamma^2+\hat{R}_1+\hat{R}_2. \tag{2.26}$$

We also have

$$\hat{R}_1=\frac{1}{n}\sum_{i=1}\{[(\varepsilon_i-\bar{\varepsilon})^2-\lambda(\delta_i-\delta)^2-2h(\delta_i-\delta)(\varepsilon_i-\bar{\varepsilon})$$
$$+2\gamma u_i(\varepsilon_i-\varepsilon)-2\gamma\beta u_i(\delta_i-\delta)]^2-4\sigma_1^2 s_n^4\gamma^2\}$$
$$\sim\frac{1}{n}\sum_{i=1}^{n}\{(S_{1i}+2\gamma u_i S_{4i})^2-4\sigma_1^2 s_n^4\gamma^2\}$$
$$-4[(v_\varepsilon+\lambda h v_\delta)\bar{\varepsilon}-(hv_\varepsilon-\lambda^2 v_\delta)\bar{\delta}],$$

where $v_\varepsilon=E\varepsilon^3$ and $v_\delta=E\delta^3$. For the denominator of $\hat{\sigma}_1^2$ we have

$$4n^{-2}S_{XY}^2(\hat{h}^2+\lambda)=n^{-2}(4\lambda S_{XY}^2+(S_{YY}-\lambda S_{XX})^2)$$
$$\sim 4s_n^2\beta^2\gamma^2+4\beta s_n^2[2\beta\gamma\overline{u\varepsilon}+2\lambda\gamma\overline{u\delta}+h(\overline{\varepsilon^2}-\lambda\overline{\delta^2})+2\lambda\overline{\varepsilon\delta}].$$

Therefore,

$$\hat{\sigma}_1^{-1}\sim\sigma_1^{-1}(1+(2\beta\gamma s_n)^{-2}[4\beta^2\gamma\overline{u\varepsilon}+4\lambda\beta\gamma\overline{u\delta}+2\beta h(\overline{\varepsilon^2}-\lambda\overline{\delta^2})+4\lambda\beta\overline{\varepsilon\delta}]$$
$$-\frac{1}{8}\sigma_1^{-2}s_n^{-4}\gamma^2\left\{\frac{1}{n}\sum_{i=1}^{n}\{(S_{1i}+2\gamma u_i S_{4i})^2-4\sigma_1^2 s_n^4\gamma^2\}\right.$$
$$-4[(v_\varepsilon+\lambda h v_\delta)\bar{\varepsilon}-(hv_\varepsilon-\lambda^2 v_\delta)\bar{\delta}]$$
$$\left.+8h(\hat{h}-h)[\lambda\sigma_\delta^4+s_n^2\sigma_\delta^2 2\beta\gamma]\right\}-\beta^{-1}(\hat{\beta}-\beta)\Big). \tag{2.27}$$

Similarly, we have,

$$\hat{\sigma}_2^{-1}\sim\sigma_2^{-1}(1+s_n^{-1}(2\overline{u\delta}+\overline{\delta^2-\sigma_\delta^2})$$
$$-\frac{1}{2}\sigma_2^{-2}s_n^{-4}\left\{n^{-1}\sum_{i=1}^{n}[(S_{2i}-u_i S_{4i})^2-s_n^4\sigma_2^2]+2\beta v_\delta(\bar{\varepsilon}-\delta)\right\}\Big), \tag{2.28}$$

and

$$\hat{\sigma}_3^{-1}\sim\sigma_3^{-1}(1+\beta^{-1}s_n^{-2}(\overline{\varepsilon\delta}+\overline{u\varepsilon}+\beta\overline{u\delta})$$
$$-\frac{1}{2}(\beta\sigma_3 s_n^2)^{-2}\left\{\frac{1}{n}\sum_{i=1}^{n}[(S_{3i}+\beta u_i S_{4i})^2-\beta^2\sigma_3^2 s_n^4]+2v_\varepsilon(2\bar{\varepsilon}-\beta\delta)\right\}\Big), \tag{2.29}$$

### 2.5. *Two-Term Edgeworth Expansions for Studentized $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$*

In this subsection, we give results on the Edgeworth expansions for Studentized versions of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ without the detailed proofs.

THEOREM 2. *Suppose $\varepsilon$ and $\delta$ are independent and continuous random variables satisfying $E(\varepsilon) = E(\delta) = 0$ and $E(\varepsilon^{2\theta} + \delta^{2\theta}) < \infty$, for some $3 < \theta < 4$. In addition to Assumption 2, suppose*

$$\sup_n \frac{1}{n} \sum |u_i|^\theta < \infty. \tag{2.30}$$

*Then the formal two-term Edgeworth expansions for $\sqrt{n}(\hat{\beta}_r - \beta)/\hat{\sigma}_r$, $r = 1, 2, 3$, are valid.*

*Proof.* The proofs for the three cases are almost the same, so as an illustration, we only sketch the proof for the case $r = 2$. Define

$$\boldsymbol{\eta}_j' = (\xi_j', S_{2j}^2 - ES_2^2, S_{2j}u_jS_{4j} - \beta^2 u_j v_\delta, u_j^2(S_{4j}^2 - ES_4^2))$$

and $\bar{\boldsymbol{\eta}}_n$ to be the sample mean of $\boldsymbol{\eta}_j$. From (2.15) and (2.28), as in the proof of Corollary 1, one sees that $\sqrt{n}(\hat{\beta}_2 - \beta)/\hat{\sigma}_2$ is a smooth function of $\bar{\boldsymbol{\eta}}_n$ and

$$\hat{\sigma}_2^{-1} = \sigma_2^{-1} A_k + R_n^{(k)},$$

where $k$ is an integer not less than $\theta/2(\theta - 3)$, $A_k$ is a polynomial in $\bar{\boldsymbol{\eta}}_n$ of degree $k$, and

$$R_n^{(k)} \ll |\bar{\boldsymbol{\eta}}_n|^{k+1}.$$

This implies, for any $\Delta > 0$, that

$$P\{|R_n^{(k)}| \geqslant \Delta/\sqrt{n}\} = o(1/\sqrt{n}). \tag{2.31}$$

Following the proof of Lemma 2, we can establish that for any $0 < a < b < \infty$, there exists a $\rho = \rho(a, b) < 1$, such that

$$\sup_{2a \leqslant |\mathbf{t}| \leqslant b} \left| \prod_{j=1}^{n} E(\exp\{i\mathbf{t}'\boldsymbol{\eta}_j\}) \right| \leqslant \rho^n. \tag{2.32}$$

Using (2.32) and the arguments similar to the proof of Part B of the Theorem of Chibisov [9], one can show that $\sqrt{n}(s_n^{-2} g_3(\xi_n) - s_n^{-4} g_4(\xi_n))A_k$ has a valid two-term Edgeworth expansion. Thus, the validity of the two-term Edgeworth expansion for $\sqrt{n}(\hat{\beta}_2 - \beta)/\hat{\sigma}_2$ follows from (2.31) and the estimation of $r_{n2}$ in Corollary 2. The sketch of the proof is complete.

Let

$$Q_{nj}^{(t)}(x)=\Phi(x)+n^{-1/2}\{\alpha_{j1}^{(t)}-\alpha_{j3}^{(t)}(x^2-1)\}\ \varphi(x) \tag{2.33}$$

denote the two term Edgeworth expansion for $\sqrt{n}(\hat{\beta}_j-\beta)/\hat{\sigma}_j$, $j=1, 2, 3$, where the expressions for the coefficients $\alpha_{j\kappa}^{(t)}$ can be computed using approximate cumulants. They are given by

$$\begin{aligned}
\alpha_{11}^{(t)}=&\ hs_n^{-4}\sigma_\delta^2(2\lambda\sigma_\delta^2(\beta^2+\lambda)^{-1}+s_n^2)\,\sigma_1^{-1}+\lambda\beta^{-1}\sigma_1(\beta^2+\lambda)^{-1}\\
&+s_n^{-4}\beta(\beta^2-\lambda)(\beta^2+\lambda)^{-3}\,\sigma_1^{-1}[E(\varepsilon^2-\lambda\delta^2)^2-4\lambda^2\sigma_\delta^4]\\
&-\sigma_1^{-3}s_n^2\gamma(E(S_1^3)+8\gamma^3\Gamma_n E(S_4^3)+12\gamma^2s_n^2E(S_1S_4^2))\\
&+2\sigma_1^{-3}s_n^2\gamma(v_\varepsilon^2+2h\lambda v_\varepsilon v_\delta-\lambda^3v_\delta^2)\\
&-8h\sigma_1^{-1}s_n^2\sigma_\delta^2\gamma^2[\lambda\sigma_\delta^2-s_n^2(\beta^2+\lambda)]-\beta^{-1}\sigma_1,\\
\alpha_{13}^{(t)}=&\ \beta^3s_n^{-6}(\beta^2+\lambda)^{-3}\,\sigma_1^{-3}[E(S_1^3)+8\gamma^3\Gamma_nE(S_4^3)+12s_n^2\gamma^2E(S_1S_4^2)],\\
\alpha_{21}^{(t)}=&\ \sigma_2^{-1}\{\beta\sigma_\delta^2s_n^{-2}+\beta s_n^{-4}(E(\delta^2-\sigma_\delta^2)^2+2s_n^2\sigma_\delta^2)\\
&-\beta s_n^{-4}[E(\delta^2-\sigma_\delta^2)^2+s_n^2\sigma_n^2]\\
&-\tfrac{1}{2}\sigma_2^{-2}s_n^{-6}[E(S_2^3)+3\beta s_n^2(E(\delta^2-\sigma_\delta^2)^2+\sigma_\varepsilon^2\sigma_\delta^2)-\Gamma_n(v_\varepsilon-\beta^3v_\delta)]\},\\
\alpha_{23}^{(t)}=&\ \sigma_2^{-3}s_n^{-6}\{E(S_2^3)+\Gamma_nE(S_4^3)+3s_n^2E(S_2S_4^2)\},\\
\alpha_{31}^{(t)}=&\ \sigma_3^{-1}\{\beta^{-1}s_n^{-4}(\sigma_\varepsilon^2\sigma_\delta^2+\beta^2s_n^2\sigma_\delta^2-2s_n^2\sigma_\varepsilon^2)-\beta^{-1}s_n^{-4}[\sigma_\varepsilon^2\sigma_\delta^2-s_n^2\sigma_\varepsilon^2+s_n^2\sigma_\delta^2]\\
&-\tfrac{1}{2}\beta^{-3}\sigma_3^{-2}s_n^{-6}[E(S_3^3)+3\beta^2s_n^2(E(\varepsilon^2-\sigma_\varepsilon^2)^2\\
&+2\beta^2\sigma_\varepsilon^2\sigma_\delta^2)+\beta^3\Gamma_n(v_\varepsilon-\beta^3v_\delta)-8v_\varepsilon^2],
\end{aligned}$$

and

$$\alpha_{33}^{(t)}=\beta^{-3}s_n^{-6}\{E(S_3^3)+\beta^3\Gamma_nE(S_4^3)+3\beta^2s_n^2E(S_3S_4^2)\}.$$

## 3. Bootstrapping

Babu and Singh [2] have shown, using Edgeworth expansions, that for a wide class of statistics, the bootstrap approximation of the sampling distribution is superior to the classical approximation. Further in this case, the bootstrap automatically corrects for skewness. In this section, we shall show that the same holds for the studentized $\hat{\beta}_i$, $i=1, 2$, and 3.

Let $(X_i^*, Y_i^*)$, $i=1, 2, ..., n$, be a simple random sample with replacement from $(X_i, Y_i)$, $i=1, 2, ..., n$. Let $h^*$, $\beta_j^*$, $\sigma_j^{*2}$, $j=1, 2, 3$, denote the bootstrapped versions of $\hat{h}$, $\hat{\beta}_j$, $\hat{\sigma}_j^2$, respectively, where $(X_i, Y_i)$ are replaced by $(X_i^*, Y_i^*)$.

If $X_j^* = X_i = u_i + \delta_i$, then we define $v_j^* = u_i$, $\delta_j^* = \delta_i$ and $\varepsilon_j^* = \varepsilon_i$. Let

$$u_j^* = v_j^* - \overline{v^*}, \qquad \text{with} \quad \overline{v^*} = n^{-1} \sum v_j^*,$$

$$s_n^{*2} = \frac{1}{n} \sum_{j=1}^{n} u_j^{*2}, \tag{3.1}$$

$$\boldsymbol{\eta}_j^0 = \boldsymbol{\eta}_{jn}^0 = (\varepsilon_j^{*2}, \delta_j^{*2}, \varepsilon_j^* \delta_j^*, \varepsilon_j^*, \delta_j^*, u_j^* \varepsilon_j^*, u_j^* \delta_j^*)',$$

and

$$\boldsymbol{\eta}_j^* = \boldsymbol{\eta}_{jn}^* = \boldsymbol{\eta}_j^0 - E_B \boldsymbol{\eta}_j^0,$$

where $E_B$ denotes the expectation under bootstrap measure $P_B$. Clearly, $\boldsymbol{\eta}_j^*$'s are i.i.d. random vectors and for each $j$,

$$E_B(\boldsymbol{\eta}_j^0) = \left( \overline{\varepsilon^2}, \overline{\delta^2}, \overline{\varepsilon\delta}, \bar{\varepsilon}, \bar{\delta}, \frac{n-1}{n} \overline{u\varepsilon}, \frac{n-1}{n} \overline{u\delta} \right)'.$$

Theorem 3. *In addition to the Assumptions of Theorem* 1, *suppose* (2.30) *holds for some* $\theta > 3$. *Then we have, with probability one, that for any measurable function* $f$ *bounded by one,*

$$\int f d(G_n^B - Q_n^B) \leqslant \theta_n + \int (\sup\{|f(\mathbf{y}) - f(\mathbf{x})| : |\mathbf{x} - \mathbf{y}| \leqslant \theta_n\}) \, \varphi_{B_n^*}(\mathbf{x}) \, d\mathbf{x},$$

*where* $\theta_n = o(n^{-1/2})$,

$$B_n^* = \frac{1}{n} \sum_{j=1}^{n} (\boldsymbol{\xi}_j - \boldsymbol{\xi}_n)(\boldsymbol{\xi}_j - \boldsymbol{\xi}_n)',$$

*and* $G_n^B$ *and* $Q_n^B$ *denote, respectively, the distribution and formal two-term Edgeworth expansion of* $\sqrt{n}\, \bar{\boldsymbol{\eta}}_n^*$.

*Proof.* The proof is similar to that of Theorem 1. We only need to verify that for any fixed $0 < a < b < \infty$, with probability one,

$$\limsup_{n \to \infty} \sup_{a \leqslant |\mathbf{t}| \leqslant b} \left| \frac{1}{n} \sum_{j=1}^{n} \exp\{i\mathbf{t}'\boldsymbol{\xi}_j\} \right| < 1, \tag{3.2}$$

$$s_n^{*2} - s_n^2 \to 0, \tag{3.3}$$

and

$$\frac{1}{n} \sum_{j=1}^{n} |u_j^*|^3 \leqslant 1. \tag{3.4}$$

To prove (3.2), we get from the proof of Lemma 2 a constant $\rho = \rho(a, b) \in (0, 1)$ such that

$$\sup_{a \leqslant |\mathbf{t}| \leqslant b} \frac{1}{n} \sum_{j=1}^{n} |E(\exp\{i\mathbf{t}'\xi_j\})| \leqslant \rho. \tag{3.5}$$

It is a routine matter to prove that there exist positive constants $c$ and $c_1$ independent of $\mathbf{t}$ such that for any fixed $\mathbf{t}$,

$$P\left\{\left|\frac{1}{n} \sum_{j=1}^{n} \{\exp\{i\mathbf{t}'\xi_j\} - E(\exp\{i\mathbf{t}'\xi_j\})\right| \geqslant (1-\rho)/4\right\} \leqslant c_1 e^{-cn}.$$

Evidently, in the ball $\{\mathbf{t} : |\mathbf{t}| \leqslant b\}$ we can choose $\mathbf{t}_1, ..., \mathbf{t}_K$, such that $K = K_n \leqslant n^8$ and for any $\mathbf{t}$ belonging to the ball, $|\mathbf{t} - \mathbf{t}_k| < n^{-1}$ for some $k \leqslant K$. Thus

$$P\left\{\left|\frac{1}{n} \sum_{j=1}^{n} \{\exp\{i\mathbf{t}_k'\xi_j\} - E(\exp\{i\mathbf{t}_k'\xi_j\})\right| \geqslant (1-\rho)/4 : \text{for some } k \leqslant K\right\}$$
$$\leqslant c_1 n^8 e^{-cn}. \tag{3.6}$$

By the law of large numbers, we have

$$n^{-2} \sum_{j=1}^{n} |\xi_j| \to 0, \text{ a.s.} \tag{3.7}$$

Equation (3.2) follows from (3.5)–(3.7) on noticing that $|\mathbf{t} - \mathbf{t}_k| \leqslant n^{-1}$ implies

$$\left|\frac{1}{n} \sum_{j=1}^{n} \{\exp\{i\mathbf{t}_k'\xi_j\} - \exp\{i\mathbf{t}'\xi_j\}\right| \leqslant n^{-2} \sum_{j=1}^{n} |\xi_j|.$$

By the condition $|u_j| \leqslant n^{1/\theta}$ and applying Bennett's inequality, one can prove that

$$\bar{v}^* \to 0, \frac{1}{n} \sum_{j=1}^{n} v_j^{*2} - s_n^2 \to 0,$$

and

$$\frac{1}{n} \sum_{j=1}^{n} |v_j^{*3}| \ll 1, \text{ a.s.,} \tag{3.8}$$

which imply (3.3) and (3.4). As an illustration, note that (3.8) follows from the Bennett's inequality

$$P_B\left(\left|\sum_{j=1}^{n}|v_j^{*3}|-\sum_{i=1}^{n}|u_i^3|\right|\geqslant n\tau\right)\leqslant 2\exp\left\{-n\tau^2\Big/\left[2\left(\frac{1}{n}\sum_{j=1}^{n}u_j^6+n^{3/\theta}\tau c_1\right)\right]\right\}$$
$$\leqslant 2\exp\{-cn^{1-3/\theta}\},$$

for some $c$, $c_1>0$.

Using the arguments similar to the proofs of Theorems 2 and 3, we can establish the following result.

THEOREM 4. *Under the conditions of Theorem 2 we have, with probability one, that*

$$P_B(\sqrt{n}(\beta_j^*-\hat{\beta}_j)/\sigma_j^*\leqslant x)=\Phi(x)+n^{-1/2}(\alpha_{j1}^*-\alpha_{j3}^*(x^2-1))\,\varphi(x)+o(n^{-1/2}).$$

Here for $j=1,2,3$, $\alpha_{j\kappa}^*$ are defined in a way similar to $\alpha_{j\kappa}^{(t)}$, in which the distribution nof $(\delta,\varepsilon)$ is replaced by its empirical distribution and $s_n^2$ is replaced by $s_n^{*2}$. Further it is not difficult to prove $\alpha_{j\kappa}^*-\alpha_{j\kappa}^{(t)}\to 0$, a.s.

The following corollary is immediate from Theorems 2 and 4.

COROLLARY 3. *Under the conditions of Theorem 4, for $j=1,2,3$,*

$$\sqrt{n}\sup_x|P(\sqrt{n}(\hat{\beta}_j-\beta)/\hat{\sigma}_j\leqslant x)-P_B(\sqrt{n}(\beta_j^*-\hat{\beta}_j)/\sigma_j^*\leqslant x)|\to 0,$$

*with probability one.*

An alternative way to implement the bootstrap is to resample from the residuals. In such a case, one needs to obtain preliminary estimates of the nuisance parameters. Geometrical considerations lead to the necessary modifications of the estimates of the residuals, to match the original structure. For the details of this method see Linder and Babu [15]. These results are not entirely satisfactory.

## REFERENCES

[1] BABU, G. J., AND SINGH, K. (1983). Inference on means using the bootstrap. *Ann. Statist.* **11** 999–1003.

[2] BABU, G. J., AND SINGH, K. (1984). On one term Edgeworth correction by Efron's bootstrap. *Sankhya Ser. A* **46** 219–232.

[3] BABU, G. J., AND SINGH, K. (1985). Edgeworth Expansions for sampling without replacement from finite populations. *J. Multivariate Anal.* **17** 261–278.

[4] BHATTACHARYA, R. N., AND GHOSH, J. K. (1978). On the validity of formal Edgeworth expansions. *Ann. Statist.* **6** 434–471.

[5] BHATTACHARYA, R. N., AND RAO, R. (1976). *Normal Approximation and Asymptotic Expansions.* Wiley, New York.

[6] BIKELIS, A. (1966). Estimates of the remainder term in the central limit theorem. *Litovsk. Mat. Sb.* **6** 323–346.

[7] BIRCH, M. W. (1964). A note on the maximum likelihood estimation of a linear structural relationship. *J. Amer. Statist. Assoc.* **59** 1175–1178.

[8] BOSE, A. (1988). Edgeworth correction by bootstrap in autoregressions. *Ann. Statist.* **16** 1709–1722.

[9] CHIBISOV, D. M. (1980). An asymptotic expansion for the distribution of a statistic admitting a stochastic expansion. I. *Theory Probab. Appl.* **25** 732–744.

[10] FULLER, W. A. (1987). *Measurement Error Models.* Wiley, New York.

[11] GLESER, L. J. (1983). Functional, Structural and ultrastructural errors-in-variables models. *Proceedings of Business and Economic Statistics Section*, pp. 57–66. American Statistical Association, Washington, DC.

[12] GLESER, L. J. (1985). A note on G. R. Dolby's unreplicated ultrastructural model. *Biometrika* **72** 117–124.

[13] JONES, T. A. (1979). Fitting straight lines when both variables are subject to error. *Math. Geol.* **11** 1–25.

[14] KENDALL, M. G., AND STUART, A. (1979). *The Advanced Theory of Statistics*, Vol. 2, 4th ed. Macmillan, New York.

[15] LINDER, E., AND BABU, G. J. (1990). Bootstrapping the linear functional relationship with known error variance ratio. Preprint.

[16] MADANSKY, A. (1959). A fitting of straight lines when both variables are subject to error. *J. Amer. Statist. Assoc.* **54** 179–205.

[17] SINGH, K. (1981). On asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187–1195.

[18] YORK, D. (1966). Least squares fitting of a straight line. *Canad. J. Phys.* **44** 1079–1086.