Edgeworth Expansion in Regression Models

MAHER B. QUMSIYEH*

Indiana University and Bethlehem University

Communicated by the Editors

Asymptotic expansions for the standardized as well as the studentized least squares estimate in multiple linear regression models are obtained without assuming normal errors and under simple assumptions that are easy to check. © 1990 Academic Press, Inc.

1. INTRODUCTION

Asymptotic expansions in regression models have been widely studied and many papers exist on the subject. For non-linear models most of the papers deal with the case of normal errors, see, for example, Kunitomo [9] and Skovgaard [10]. The papers that omit the normality assumption substitute it with hard to analyze conditions that are difficult to check in real situations such as condition $B_k^{(1)}$ in Ivanov and Zwanzig [7] and the conditions $C, D_x, E_x,$ and $F_x$ in Ivanov and Zwanzig [8]. In these two papers Ivanov and Zwanzig obtain an asymptotic expansion for the standardized least squares estimate in non-linear regression models. An asymptotic expansion for the studentized least squares estimate was not obtained. See also Zwanzig [11].

In this paper we obtain asymptotic expansions for the studentized least squares estimate in multiple linear regression models without assuming normal errors and under simple assumptions that are easy to check. We also show, under suitable easily verifiable conditions and without assuming normal errors, that the standardized least squares estimate has a density and that an Edgeworth expansion for the density is valid. Using this expansion for the density, we can obtain an Edgeworth expansion for the distribution of the standardized least squares estimate.

Received August 12, 1988; revised November 8, 1989.

AMS 1980 subject classifications: primary 62E20, 62J05.

Key Words and Phrases: linear models, least squares, Edgeworth expansions.

* Part of this work is part of the author’s Ph.D. dissertation written at Indiana University under the direction of Professor R. N. Bhattacharya.
Consider the linear regression model

\[ Y = X\beta + \varepsilon. \] (1.1)

In this equation, \( Y \) is an \( n \times 1 \) observable vector of random variables, \( X \) is an \( n \times k \) matrix of known fixed numbers, \( \beta \) is a \( k \times 1 \) vector of unknown parameters to be estimated from the data and \( \varepsilon \) is an \( n \times 1 \) unobservable vector of random variables. The dependence on \( n \) is always omitted in this paper unless confusion seems likely. Attention is restricted to the conventional least squares estimate \( \hat{\beta} \) of \( \beta \) given by

\[ \hat{\beta} = (X'X)^{-1} X'Y, \] (1.2)

where \( X' \) is the transpose of \( X \) and where we assume \( X \) to be of full rank \( k \).

Assume the components \( \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \) of \( \varepsilon \) are i.i.d. with mean 0 and finite positive variance \( \sigma^2 \). Let \( \hat{\sigma}^2 \) be the least squares estimate of \( \sigma^2 \), \( \hat{\sigma}^2 = \frac{Y'Y - Y'X(X'X)^{-1} X'Y}{n} \) and consider the statistics \( T_n \) and \( W_n \) given by

\[ T_n = (X'X)^{1/2} (\hat{\beta} - \beta)/\sigma, \] (1.3)
\[ W_n = (X'X)^{1/2} (\hat{\beta} - \beta)/\hat{\sigma}. \] (1.4)

Note that since \( X \) is of full rank, \( X'X \) is positive definite and has a positive definite square root.

Under quite general conditions on the matrix \( X \), it can be shown that \( T_n \) and \( W_n \) are asymptotically normal with mean zero and covariance matrix \( I_k \), the \( k \times k \) identity matrix. For details see for example Anderson [1] or Eicker [6]. In Section 3 we show, under suitable assumptions on the matrix \( X \) and on the distribution of the \( \varepsilon_i \)'s, that \( T_n \) has a bounded density and that an Edgeworth expansion for this density is valid. In Section 4 we derive a two-term Edgeworth expansion for the distribution of \( W_n \) (Theorem 4.5) and we show how one can use that to obtain a better approximation for the distribution of \( W_n \) than the normal approximation for estimation and testing purposes (Remark 4.6). An Edgeworth expansion for the correlation model in which \( X \) is considered random is also valid (Remark 4.7).

2. Notations and Some Lemmas

In the linear model (1.1) considered before, write \( X \) as \( (X_1, \ldots, X_n)' \), where the \( X_i \)'s are \( k \times 1 \) vectors for \( i = 1, \ldots, n \), \( \varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)' \). Let \( A = X'X = \sum_{i=1}^n X_iX_i' \), \( \lambda_n = \) the smallest eigenvalue of \( A \), \( M_n = \max\{\|X_i\|: i = 1, \ldots, n\} \), where \( \| \cdot \| \) is the Euclidean norm.

We want to make the following assumptions:
A1. \( \varepsilon_i \)'s are i.i.d. with mean zero and finite variance \( \sigma^2 \).
A2. \( X \) is of full rank \( k \).
A3. The characteristic function \( \gamma \) of \( \varepsilon_i \) is integrable.
A4. \( \varepsilon_i \)'s have finite \( s \)th absolute moment, for some integer \( s \geq 3 \) and \( \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \|X_i\|^s < \infty \).
A5. \( \lim \frac{\lambda_n}{n} > 0 \), \( M_n = O(n^\delta) \) for some \( \delta \in (0, \frac{1}{2}) \).

Note that A4 implies that
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \|X_i\|^r < \infty \quad \text{for} \quad 1 \leq r \leq s.
\]

In Section 4 assumption A3 will be weakened to \( \overline{A}_3 \) and assumption A4 will be strengthened to \( \overline{A}_4 \), where

\( \overline{A}_3 \). \( \varepsilon_i \)'s have a non-zero absolutely continuous component which has a positive density on an open subset of \( \mathbb{R} \).
\( \overline{A}_4 \). \( \varepsilon_i \)'s have finite 2s\( \text{th absolute moment for some integer } s \geq 3 \) and \( \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \|X_i\|^s < \infty \).

Now under assumption A5, there exist a positive integer \( N \) and a constant \( K \) such that
\[
n/\lambda_n \leq K \quad \text{and} \quad M_n \leq Kn^\delta \quad \text{for all } n \geq N. \tag{2.2}
\]

Let \( \mu_r \) be the \( r \)th moment of \( \varepsilon_i \), \( \rho_r \) the \( r \)th absolute moment of \( \varepsilon_i \) and \( \chi_r \) the \( r \)th cumulant of \( \varepsilon_i \). Write
\[
\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^2. \tag{2.3}
\]

The following notations will be used in the following sections:

- \( \nu \) a nonnegative integral vector \( (\nu \in (\mathbb{Z}^+)^k) \).
- \( |\nu| = \sum_{i=1}^{k} \nu_i \), where \( \nu = (\nu_1, \ldots, \nu_k)^t \).
- \( \nu! = \nu_1! \cdots \nu_k! \).
- \( x^\nu = x_1^{\nu_1}x_2^{\nu_2} \cdots x_k^{\nu_k} \) [\( x = (x_1, \ldots, x_k)^t \) \( \in \mathbb{R}^k \) or \( \mathbb{C}^k \)].
- \( D^\nu = D_1^{\nu_1} \cdots D_k^{\nu_k} \), where \( D_j = \partial/\partial x_j, 1 \leq j \leq k \).
- \( \chi_r(X) \) \( r \)th cumulant of the random vector \( X \).
- \( \Phi_k \) standard normal distribution on \( \mathbb{R}^k \).
- \( \varphi_k \) density of \( \Phi_k \).
$B^c$ complement of a set $B$ in $\mathbb{R}^k$.

$A$ symmetric difference between two sets.

$\|A\|$ norm of a matrix $A$.

Bhattacharya and Ranga Rao [5] will be referred to as BR because of its frequent usage.

The following lemmas will be used in the proofs of the main results.

**Lemma 2.1.** Let $\xi \in \mathbb{R}^k$, $B(\xi, c) = B = \{i: 1 \leq i \leq n, |\xi' A^{-1/2} X_i| > c \|\xi\|/n^{1/2}\}$, and $\Gamma_n(\xi, c)$ be the number of elements in $B$, $\#B$, say. Then under assumption A2 and for $0 < c < 1$ one has

$$\Gamma_n(\xi, c)/n \geq (1 - c^2)/[(nM_n^2/\lambda_n) - c^2].$$

**Proof.** Let $B' = \{i: 1 \leq i \leq n, |\xi' A^{-1/2} X_i| \leq c \|\xi\|/n^{1/2}\}$ and note that

$$\sum_{i=1}^n (\xi' A^{-1/2} X_i)^2 = \sum_{i=1}^n (\xi' A^{-1/2} X_i A^{-1/2} \xi) = \|\xi\|^2.$$ Hence,

$$\|\xi\|^2 = \sum_{i=1}^n (\xi' A^{-1/2} X_i)^2 + \sum_{i=1}^n (\xi' A^{-1/2} X_i A^{-1/2} \xi) \leq \Gamma_n(\xi, c) \|\xi\|^2 M_n/\lambda_n + (n - \Gamma_n(\xi, c))n^{1/2} \|\xi\|^2/n.$$

I.e., $(1 - c^2) \|\xi\|^2 \leq [(\Gamma_n(\xi, c)/n)](nM_n^2/\lambda_n) - c^2 \|\xi\|^2$.

Now the result follows from this inequality, since $\lambda_n = \min\|y\|=1 y' A y = \min\|y\|=1 y' \sum_{i=1}^n X_i X_i' y \leq nM_n^2$, and so $nM_n^2/\lambda_n \geq 1 > c$. 

**Lemma 2.2.** Assume that the $n \times k$ matrix $X$ is of full rank and that the smallest eigen value $\lambda$ of $X'X$ goes to $\infty$ as $n \to \infty$. Let $s$ be an arbitrary positive integer. Then for large enough $n$ one can find, among the rows of $X$, $s$ disjoint sets with $k$ independent rows in each.

**Proof (By induction on $s$).** $X$ is of full rank implies the lemma is true for $s = 1$. Now assume that for fixed $n = n_0$ we can find among the rows of $X$, $r - 1$ disjoint sets with $k$ independent rows in each ($r \geq 2$). We need to show that, choosing $n$ large enough, one can find among the $n_0 + 1$, $n_0 + 2$, ..., $n$ rows another set of $k$ independent vectors.

Write $X$ as $[B_n B_{n-n_0}]$, where $B_n$ consist of the first $n_0$ rows and $B_{n-n_0}$ are the remaining $n - n_0$ rows. Then, $X'X = B_n' B_n + B_{n-n_0} B_{n-n_0}$.

Let $y \in \mathbb{R}^k$, then

$$\min_{\|y\|=1} \{y' B_{n-n_0} B_{n-n_0} y\} = \min_{\|y\|=1} \{y' X' X y - y' B_n B_n y\} \geq \lambda_n - A,$$

where $A$ is the largest eigen value of $B_n' B_n$. The left-hand side of (2.4) is the smallest eigen value of $B_{n-n_0} B_{n-n_0}$ and the right-hand side goes to $\infty$ as $n \to \infty$. Therefore, for some $n > n_0$, $B_{n-n_0} B_{n-n_0}$ is positive definite, which implies that $B_{n-n_0}$ is of full rank; i.e., it contains $k$ independent rows.
Lemma 2.3. Let $\hat{\sigma}^2$ be the least squares estimate of $\sigma^2$ and $\hat{\sigma}^2$ be as in (2.2). Then for $a > 0$ and under assumption A1, one has $P\{|\hat{\sigma}^2 - \sigma^2| > a\} \leq 2\mu_4 M_n^4/(a^2\lambda_n^2$).

Proof. The lemma is obviously true if $\mu_4 = \infty$, so assume $\mu_4 < \infty$ and note that
\[
\hat{\sigma}^2 = \frac{Y'Y - X'X^{-1}X'Y}{n} \quad \text{and} \quad \sigma^2 = \frac{[\epsilon'X + \epsilon']X^{-1}X'X' + \epsilon' + \epsilon]}{n} = \frac{\epsilon' \epsilon - \epsilon'X'X^{-1}X' \epsilon}{n}.
\]

Therefore, by A1 and using Chebyshev's inequality, one has
\[
P\{|\hat{\sigma}^2 - \sigma^2| > a\} \leq E\left[\left|\sum_{i,j} X_i'X^{-1}X_j \epsilon_i \epsilon_j\right| n^2\right]/a^2n^2 = (1/a^2n^2) \left[ 2 \sum_{i \neq j} (X_i'X^{-1}X_j)^2 E(\epsilon_i^2 \epsilon_j^2) + \sum (X_i'X^{-1}X_j)^2 E(\epsilon_i^2) \right] \leq (1/a^2n^2)\lambda_n^2[2n(n-1)M_n^4 \sigma^4 + n M_n^4 \mu_4] \leq (2M_n^4 \mu_4)/(a^2\lambda_n^2$.

3. Asymptotic Expansion for the Distribution of the Standardized Least Squares Estimate

Using the same notation as before and assuming, without loss of generality, that $\sigma = 1$ we have
\[
T_n = A^{1/2}(\hat{\beta} - \beta)/\sigma = A^{1/2}(A^{-1}X'Y - \beta) = A^{-1/2}(X'Y - X'X \hat{\beta}) = A^{-1/2}X' \epsilon = A^{-1/2} \sum_{i=1}^n X_i \epsilon_i = n^{1/2}Z,
\]
where $Z = (1/n) \sum_{i=1}^n Z_i$ and $Z_i = n^{1/2}X_i^{-1/2}X_i \epsilon_i$. Clearly under assumptions A1 through A4 the $Z_i$'s are independent with mean 0 and finite $s$th absolute moment. Let, $V_j = \text{Cov}(Z_j)$, $V = (1/n) \sum_{j=1}^n V_j = \text{Cov}(A^{-1/2}X' \epsilon) = L_k$, and $\bar{\rho}_r = (1/n) \sum_{j=1}^n E\|Z_i\|^r$. Noted that by A4 and A5, there exist an integer $N'(r)$ such that $\bar{\rho}_r \leq c(r) \bar{\rho}_r$ for $n \geq N'(r)$, where $c(r)$ is a constant depending on $r$ only, $r = 1, 2, \ldots, s$. 

Let $\chi_v(Z_j)$ be the $v$th cumulant of $Z_j$, and write $\tilde{\chi}_v = (1/n) \sum_{i=1}^n \chi_v(Z_i)$. It is easy to see that $\tilde{\chi}_v = (1/n) \sum_{i=1}^n (n^{1/2}A^{-1/2}X_i)^v \chi_{|v|}$, where as before $\chi_{|v|}$ is the $|v|$th cumulant of $\varepsilon_i$. Now let $z \in \mathbb{C}^k$, and define, for each positive integer $s$, $\chi_s(z) = s! \sum_{|v|=s} \tilde{\chi}_v z^v / |v|!$ and

$$P_s(z, \{\tilde{\chi}_v\}) = \sum_{m=1}^{s!} \frac{1}{m!} \sum_{j_1, \ldots, j_m} \chi_{j_1 + 2(z)} \chi_{j_2 + 2(z)} \cdots \chi_{j_m + 2(z)} \cdot$$

where $\sum^*$ is the sum over all $m$-tuples of positive integers $j_1, \ldots, j_m$ satisfying $\sum_{i=1}^m j_i = s$. Using these notations and following the lines of the proof of Theorem 9.9 of BR one can show that under assumptions A1 through A5 there exist two positive constants $c_1(s, k)$ and $c_2(s, k)$ depending only on their arguments such that, for all $\xi \in \mathbb{R}^k$ satisfying $|\xi| \leq c_1(s, k) n^{(1/2) - \delta}$, one has, for all nonnegative integral vectors $v$, $0 \leq |v| \leq s$, and for large enough $n$,

$$D_v \left\{ f_n(\xi) - e^{-\|\xi\|^2/2} \left[ 1 + \sum_{r=1}^{s-3} P_r(i\xi, \{\tilde{\chi}_v\}) \right] \right\} \leq c_2(s, k) n^{-s/2} \|\xi\|^s + \|\xi\|^{3(s-2) + |v|} e^{-\|\xi\|^2/4},$$

(3.3)

where $f_n(\xi)$ is the characteristic function of $T_n$.

Also, under the same conditions and as in Lemma 14.3 of BR, there exist two positive constants $c_3(s, k, \rho_s)$ and $c_4(v, k)$ such that, for $|v| \leq s$ and for $\xi \in \mathbb{R}^k$ satisfying $|\xi| \leq c_3(s, k, \rho_s) n^{1/2}$, one has

$$|D_v f_n(\xi)| \leq c_4(v, k) [1 + |\xi|^{|v|}] e^{-\|\xi\|^2/3},$$

(3.4)

The next theorem is the main result of this section.

**Theorem 3.1.** Under assumptions A1 through A5, the distribution $Q_n$ of $T_n$ has a bounded density $q_n$ and we have the following asymptotic expansion for $q_n$:

$$\sup_{x \in \mathbb{R}^k} \left( 1 + \|x\|^{s'} \right) \left| q_n(x) - \left[ 1 + \sum_{r=1}^s n^{-r/2} P_r(-D, \{\tilde{\chi}_v\}) \right] \varphi_k(x) \right| = O(n^{-(s-2)/2}).$$

**Proof.** The characteristic function $f_n(\xi)$ of $T_n$ is given by

$$f_n(\xi) = E[\exp(i\xi' T_n)] = E \left[ \exp \left( i\xi' \sum_{i=1}^n A^{-1/2}X_i \varepsilon_i \right) \right] = \prod_{i=1}^n \gamma(\xi' A^{-1/2}X_i).$$
Since $X = (X_1, ..., X_r)'$ is of full rank, there exist $k$ independent rows among the rows $X_1', ..., X_r'$ of $X$, say $X_1', ..., X_k'$, where $1 \leq r_i \leq n$ for $i = 1, ..., k$, which implies $\{A^{-1/2}X_r\}_{i=1}^{k}$ are $k$ independent vectors. Therefore the linear transformation $u_j = X_j'A^{-1/2}\xi$ ($j = 1, ..., k$) is one to one with nonzero Jacobian $J$. Let $u = (u_1, ..., u_k)'$; then we have

$$
\int_{\mathbb{R}^k} |f_n(\xi)| \, d\xi = \int_{\mathbb{R}^k} \left| \prod_{i=1}^{k} \gamma(X_i'A^{-1/2}\xi) \right| \, d\xi \\
\leq \int_{\mathbb{R}^k} \left| \prod_{i=1}^{k} \gamma(X_i'A^{-1/2}\xi) \right| \, d\xi \\
= \int_{\mathbb{R}^k} \left| \prod_{i=1}^{k} \gamma(u_i) \right| \cdot |J| \, du \\
= |J| \left[ \int_{\mathbb{R}^k} (\gamma(t)) \, dt \right]^k < \infty \quad \text{(by A3)}
$$

i.e., $f_n(\xi)$ is integrable, which implies that $Q_n$ has a bounded density $q_n$ given by $q_n(x) = (1/2\pi)^k \int_{\mathbb{R}^k} \exp(-i\xi'x) \cdot f_n(\xi) \, d\xi$.

Now by the Fourier inversion formula we obtain for $|\nu| \leq s$,

$$
\left| \chi^v \left\{ q_n(x) - \left[ 1 + \sum_{r=1}^{s-3} n^{-r/2}P_r(-D, \{\bar{x}_r\}) \right] \phi_k(x) \right\} \right| \\
= (1/2\pi)^k \left| \int_{\mathbb{R}^k} \exp(-i\xi'x) \, D^v \\
\times \left\{ f_n(\xi) \left[ 1 + \sum_{r=1}^{s-3} n^{-r/2}P_r(i\xi, \{\bar{x}_r\}) \right] e^{-i\xi'\tilde{x}/2} \right\} \, d\xi \right| \\
\leq (1/2\pi)^k \left[ \int_{B} |D^v \left\{ f_n(\xi) \left[ 1 + \sum_{r=1}^{s-3} n^{-r/2}P_r(i\xi, \{\bar{x}_r\}) \right] e^{-i\xi'\tilde{x}/2} \right\} | \, d\xi \right] \\
+ \int_{B'} |D^v f_n(\xi)| \, d\xi \\
+ \int_{B'} |D^v \left\{ 1 + \sum_{r=1}^{s-3} n^{r/2}P_r(i\xi, \{\bar{x}_r\}) \right\} e^{-i\xi'\tilde{x}/2} \} | \, d\xi \\
= (1/2\pi)^k \left[ I_1 + I_2 + I_3 \right], \quad \text{(3.5)}
$$

where $B = \{ \xi \in \mathbb{R}^k : \|\xi\| \leq c_4(s, k) n^{(1/2) - \delta} \}$ and $c_4(s, k)$ is as before.

The theorem now follows if we can show that the right-hand side of (3.5) is $O(n^{-(s-2)/2})$ taking $\nu = (0, s, 0, ..., 0)$, $(0, s, 0, ..., 0)$, By (3.3) it is easily seen that $I_1 = O(n^{-(s-2)/2})$. Also, over $B'$ the integrand of $I_3$ decays exponentially fast and so $I_3 = O(n^{-(s-2)/2})$, and so, to show that the right-
hand side of (3.5) is $O(n^{-(s-2)/2})$, we need only to show that $I_2$

 is $O(n^{-(s-2)/2})$. To show this, write $B' = B_1 \cup B_2$, where, $B_1 = \{ \xi \in \mathbb{R}^k : c_1(s, k) < \| \xi \| < c_3(s, k, \rho_n) n^{1/2} \}$ and $B_2 = B' - B$. Then, $I_2 \leq \int_{B_1} |D^y f_n(\xi)| \, d\xi + \int_{B_2} |D^y f_n(\xi)| \, d\xi$. Using (3.4) one easily sees (since $\delta < \frac{1}{2}$) that $\int_{B_1} |D^y f_n(\xi)| \, d\xi = O(n^{-(s-2)/2})$, and so we need only to show that $\int_{B_2} |D^y f_n(\xi)| \, d\xi$ is $O(n^{-(s-2)/2})$. To do this let $v = (v_1, \ldots, v_k)'$ be a nonnegative integral vector, $|v| \leq s$. Then

$$|D^y \gamma(X_i A^{-1/2} \xi)| = \left| D^y \int \exp(iX_j A^{-1/2} \xi \xi) \, dP_{ij} \right| \leq (M_n^{1/2})^{(|v|)} \rho_{|v|}$$

(3.6)

where $P_{ij}$ is the probability distribution of $\varepsilon_j$.

Now using Leibniz formula for differentiation of a product of functions, we see that $D^y f_n(\xi) = D^y \prod_{i=1}^n \gamma(X_i A^{-1/2} \xi)$ is the sum of $n^{|v|}$ terms each of the form

$$g(\xi) = \prod_{i \notin \{i_1, \ldots, i_j \}} \gamma(X_{i} A^{-1/2} \xi) \cdot (D^n[\gamma(X_{i} A^{-1/2} \xi)])^{i_1} \times \cdots \times (D^n[\gamma(X_{i} A^{-1/2} \xi)])^{i_j}$$

(3.7)

where $i_1, \ldots, i_j$ are distinct indices in $\{1, \ldots, n\}$, $r_1, \ldots, r_j$ are positive integers and $\alpha_1, \ldots, \alpha_j$ are nonnegative integral vectors satisfying $\sum_{i=1}^j r_i \alpha_i = v$ and $j \leq |v|$. Also, using Lemma 2.2, take $n$ large enough so that $\chi$ has at least $s + 1$ disjoint sets with $k$-independent rows in each, and since $j \leq |v| \leq s$, one can find, among $\{X_i A^{-1/2} \}_{i \notin \{i_1, \ldots, i_j \}}$, $k$ independent rows, say, $\{X_{l_i} A^{-1/2} \}_{i=1, \ldots, k}$, $l_i \in \{1, \ldots, n\}$ for $i = 1, \ldots, k$, $l_i \notin \{i_1, \ldots, i_j \}$; and so, using (3.6)-(3.7), we have

$$\int_{B_2} |g(\xi)| \, d\xi \leq [M_n^{1/2}]^{(|v|)} \rho_{|v|} \int_{B'_i} \prod_{i \notin \{i_1, \ldots, i_j \}} |\gamma(X_{i} A^{-1/2} \xi)| \, d\xi$$

$$= [M_n^{1/2}]^{(|v|)} \rho_{|v|} \int_{B'_i} \prod_{i \notin \{i_1, \ldots, i_j \}} |\gamma(X_{i} A^{-1/2} \xi)|$$

$$\times \prod_{i=1}^k \int_{B'_i} \prod_{i \notin \{i_1, \ldots, i_j \}} |\gamma(X_{i} A^{-1/2} \xi)| \, d\xi.$$  

(3.8)

Now fix $0 < c < 1$. Using Lemma 2.1 and choosing $n$ large enough so that (2.2) is satisfied, we get

$$\Gamma_n(\xi, c) \geq [n(1 - c^2)]/[(nM_n^{2/3}/\lambda_n) - c^2]$$

$$= q \cdot n^{1 - 2\delta}, \quad \text{where} \quad q = (1 - c^2)/K^3,$$
\( n \geq N \). Also, \( \# \{ i \in \{ 1, \ldots, n \} : i \notin \{ i_1, \ldots, i_k \} \cup \{ l_1, \ldots, l_k \} \} \) and \( |X_i^tA^{-1/2}\xi| > c \| \xi \|/n^{1/2} \geq F_n(\xi, c) - |v| - k \).

Now by A3 we have \( \sup_{|v| < cc_3(s, k, \rho_\delta)} |\gamma(t)| = d < 1 \). Furthermore, for \( \xi \in B_2 \) we have \( c \| \xi \|/n^{1/2} > cc_3(s, k, \rho_\delta) \).

The last four relations imply that if \( \xi \in B_2 \) and \( n \geq N \) then
\[
\prod_{i \notin \{ i_1, \ldots, i_k \} \cup \{ l_1, \ldots, l_k \}} |\gamma(X_i^tA^{-1/2}\xi)| \leq d^{\left[ q_\delta(n^{1/2}) - |v| - k \right]}.
\]

Substituting (3.9) in (3.8) one gets
\[
\int_{B_2} |g(\xi)| \, d\xi \leq \left[ \frac{M}{\lambda^2} \right]^{|v|} \rho_{|v|} \, d^{\left[ q_\delta(n^{1/2}) - |v| - k \right]} \times \int_{B_2} \prod_{i = 1}^k |\gamma(X_i^tA^{-1/2}\xi)| \, d\xi.
\]

Now consider the linear transformation \( u_i = X_i^tA^{-1/2}\xi \) for \( i = 1, \ldots, k \) and let \( u = (u_1, \ldots, u_k)' \), then \( \xi = A^{1/2}[X_1, \ldots, X_k]^{-1} u \) \((X_1, \ldots, X_k)\) are independent and so \( [X_1, \ldots, X_k] \) is invertible. The Jacobian of the above transformation is given by \( J = \det[A^{1/2}[X_1, \ldots, X_k]^{-1}] \) and hence using A5 one gets
\[
|J|^2 \leq (nM^2)^k/\det([X_1, \ldots, X_k]' [X_1, \ldots, X_k]) = O(n^{k+2\delta k}).
\]

Using this transformation in (3.10) one gets
\[
\int_{B_2} |g(\xi)| \, d\xi \leq \left[ \frac{M}{\lambda^2} \right]^{|v|} \rho_{|v|} \, d^{\left[ q_\delta(n^{1/2}) - |v| - k \right]} O(n^{k+2\delta k}) \left\{ \int_R |\gamma(t)| \, dt \right\}^k
\]
and so
\[
\int_{B_2} |D^s f_n(\xi)| \, d\xi \leq n^{\delta k} \cdot \left[ \frac{M}{\lambda^2} \right]^{|v|} \rho_{|v|} \cdot d^{\left[ q_\delta(n^{1/2}) - |v| - k \right]} \times O(n^{k+2\delta k}) \left\{ \int_R |\gamma(t)| \, dt \right\}^k = O(n^{-(s-2)/2}).
\]

The next corollary gives an Edgeworth expansion for the distribution \( Q_n \) of \( T_n \); its proof follows immediately from Theorem 3.1.

**Corollary 3.2.** Under the assumptions of Theorem 3.1 and if the integer \( s \) in A4 (never smaller than 3) is also larger than \( k \) then
\[
\left\| Q_n - \int \left[ 1 + \sum_{r=1}^{s-3} n^{-r/2} P_r(-D, \{ \bar{\alpha}, \}) \right] \varphi_k(x) \, dx \right\| = O(n^{-(s-2)/2}).
\]
4. ASYMPTOTIC EXPANSION FOR THE DISTRIBUTION OF
THE STUDENTIZED LEAST SQUARES ESTIMATES

In this section we are going to derive a two-term Edgeworth expansion for the studentized statistic $W_n$ given in (1.4).

Define the $(k + 1) \times 1$ random vector $R_n$ by

\[
R_n = \begin{pmatrix} \sum_{i=1}^{n} A^{-1/2} X_i e_i \\ n^{1/2} \sum_{i=1}^{n} (e_i^2 - \sigma^2)/n \end{pmatrix},
\]

where $\sigma^2$ is as in (2.3). Clearly $R_n = n^{1/2} \bar{Y}$, where

\[
\bar{Y} = (1/n) \sum_{i=1}^{n} Y_i \quad \text{and} \quad Y_i = \left( n^{1/2} A^{-1/2} X_i e_i \right) / (\bar{e}_i^2 - \sigma^2).
\]

Let $F_i$ be the distribution of $Y_i$ and $\hat{F_i}$ be the Fourier transform of $F_i$. Then, under assumption A1, the $Y_i$'s are independent with mean 0, and if we assume $A4$, $A5$, and $n \geq N$ to satisfy (2.2), we obtain $E \| Y_i \|'' \leq 2' (E \| n^{1/2} A^{-1/2} X_i e_i \|'' + E |e_i^2 - \sigma^2|') \leq 2' [K'' \| X_i \|'' + c_5] \leq s$. Here $c_5 = c_5(r, \rho_2r)$ is a constant. Therefore (see (2.1))

\[
\lim (1/n) \sum_{i=1}^{n} E \| Y_i \|'' < \infty. \tag{4.1}
\]

Now let $S_i = \text{Cov}(Y_i)$, $S = (1/n) \sum_{i=1}^{n} S_i$, $\chi_s(Y_i)$ the $v$th cumulant of $Y_i$, $\bar{\chi}_s = (1/n) \sum_{i=1}^{n} \chi_s(Y_i)$. We claim that under assumption $A3$, $S$ is positive definite. To show this, first note that

\[
S = \text{Cov} \begin{pmatrix} \sum_{i=1}^{n} A^{-1/2} X_i e_i \\ n^{1/2} \sum_{i=1}^{n} (e_i^2 - \sigma^2)/n \end{pmatrix} = \begin{pmatrix} \sigma^2 I_k & (\mu_3/n^{1/2}) \sum_{i=1}^{n} A^{-1/2} X_i \\ (\mu_3/n^{1/2}) \sum_{i=1}^{n} A^{-1/2} X_i & E(e_i^2 - \sigma^2)^2 \end{pmatrix}.
\]

Now, let $\mathbf{1}$ be the $n \times 1$ vector all of whose elements are ones. Then
where \( n \) is the largest eigenvalue of \( X'AX \) which is 1, since the matrix is idempotent. Using this one obtains

\[
\det(S) = \det(\sigma^2 I_k) \cdot \{E(\varepsilon_i^2 - \sigma^2)^2 - (\mu_2^2/\mu_4) \sum_j (X'_j A^{-1} X_j) \}
\]

\[
= (\mu_2)^k \cdot \left[ \mu_4 - \mu_2^2 - (\mu_2^2/\mu_4) \cdot X'AX^{-1}X' \right] \geq (\mu_2)^{k-1} \cdot \left[ \mu_2 \mu_4 - \mu_2^2 - \mu_3^2 \right] = c_6,
\]

where \( c_6 = c_6(k, \mu_2, \mu_3, \mu_4) \). Also, by \( A3, 1, \varepsilon_1, \varepsilon_2^i \) are linearly independent and so \( \text{cov}[(1, \varepsilon_1, \varepsilon_2^i)^T] \) is positive definite, i.e., \( 0 < \det(\text{cov}[(1, \varepsilon_1, \varepsilon_2^i)^T]) = \mu_2 \mu_4 - \mu_2^2 - \mu_3^2 \). This and the assumption that \( \mu_2 \) is positive show that \( c_6 \) is positive and so \( \det(S) \) is positive which implies \( S \) is positive definite.

Now \( S \) is positive definite implies that the smallest eigenvalue \( \lambda_n \) of \( S \) is positive. Write \( S^{-1} = B^2 \); then by (4.2) we have

\[
\|B\| = 1/\lambda_n^{1/2} \leq c_7(k, \mu_2, \mu_3, \mu_4) < \infty
\]

and

\[
\|B^{-1}\| \leq c_8 = c_8(k, \mu_2, \mu_4).
\]

Now let \( \tilde{\rho}_r = (1/n) \sum_{i=1}^n E \|BY_i\|^r \). Then (4.1) and (4.3) imply that \( \tilde{\rho}_r \) is bounded for \( 0 \leq r \leq s \); i.e., there exist \( c_9 = c_9(k, \rho_2) \) such that

\[
\tilde{\rho}_r \leq c_9 < \infty.
\]

Using these notations we have the following theorem.

**Theorem 4.1.** Under assumptions A1, A2, A3, A4, and A5, there exist positive constants \( c_{10} = c_{10}(s, k, \rho_2) \) and \( c_{11} = c_{11}(s, k, \rho_2) \) such that, for every \( t \in \mathbb{R}^{k+1} \) satisfying

\[
\|t\| \leq c_{10} n^{(1/2) - \delta},
\]

one has, for all nonnegative integral vectors \( v \in \mathbb{Z}^+ \) \( k+1 \), \( 0 \leq |v| \leq s \),

\[
\left| D^v \prod_{i=1}^n \hat{F}_i(Bt/n^{1/2}) - e^{-\|t\|^2/2} \left[ 1 + \sum_{r=1}^{s-3} n^{-r/2} P_r(iBt, \{\tilde{\xi}_r\}) \right] \right| \leq c_{11} n^{-(s-3)/2} \left[ \|t\|^{\gamma - |v|} + \|t\|^{\gamma(s-3) + |v|} \right] \cdot e^{-\|t\|^2/2},
\]

where \( P_r(z, \{\tilde{\xi}_r\}) \) is as defined in the previous section using dimension \( k+1 \) instead of \( k \).
Proof. This theorem follows directly from Theorem 9.9 of BR. We only need to show that (4.6) implies condition (9.37) of that theorem which is

$$
\|t\| \leq d_n, \quad \|s\| \leq c(s, k) n^{1/2} \cdot \bar{\rho}^{-1/2},
$$

(4.7)

where $d_n = \sup \{ a > 0 : t' St \leq a^2 \}$. We only need to show that (4.6) implies $|\hat{F}_i(Bt/n^{1/2}) - 1| < h$. To show this, let $\hat{G}$ be the characteristic function of $[\varepsilon_i^2 - \sigma^2]$ then, since $\hat{G}$ is continuous at 0, there exists $\delta' > 0$ such that $|\hat{G}(z) - 1| < \frac{1}{2}$ for $\|z\| < \delta'$.

Now take $n > N$ to satisfy Lemma 2.2 and let

$$
e_{10} = \min \{ c(s, k) \cdot c_0^{-1/\bar{s} - 2}, \delta'/\left[ c_7 (1 + K^{3/2}) \right] \}. \quad (4.8)
$$

We need to show that if $t \in \mathbb{R}^{k+1}$ satisfies (4.6) then it satisfies (4.7). To show this, write $Bt = [t_1, t_2]'$, where $t_1 \in \mathbb{R}^k$, $t_2 \in \mathbb{R}$. Then $\hat{F}_i(Bt/n^{1/2}) = \hat{G}([t_1' A^{-1/2} X_i t_2/n^{1/2}']) = \hat{G}(b)$, say, and using (4.3), (4.6), (4.8), one obtains

$$
\|b\| \leq |t_2|/n^{1/2} + \|t_1' A^{-1/2} X_i\| \leq \|Bt\|/n^{1/2} + M_n \|Bt\|^2/n^{1/2}
\leq c_7 c_{10} n^{(1/2) - \delta} \left[ 1 + K^{3/2} \right] n^{1/2} \leq \delta'
$$

and so $|\hat{F}_i(Bt/n^{1/2}) - 1| < \frac{1}{2}$ for all $1 \leq i \leq n$. Therefore the first condition of (4.7) is satisfied. The second part of (4.7) follows easily using (4.6), (4.8), and (4.5).

Before giving an Edgeworth expansion for the distribution of $R_n$ we need the following lemma.

**Lemma 4.2.** Let the assumptions be as in Theorem 4.1 and let $c_{12}$ be a positive constant. Then there exists a constant $d, 0 < d < 1$, such that for $t_1 \in \mathbb{R}^k$, $t_2 \in \mathbb{R}, t = [t_1, t_2]' \in \mathbb{R}^{k+1}$ satisfying $\|t\| > c_{12}$ and for $i \in B(t_1, c)$ for some fixed $c, 0 < c < 1$, one has $|\hat{F}_i(t)| \leq d$, where $B(\xi, c)$ is as in Lemma 2.1.

**Proof.** Let, as before, $G$ be the distribution of $[\varepsilon_i^2 - \sigma^2]'$ and $\hat{G}$ be its characteristic function. Then, by Lemma 1.4 of Bhattacharya [2], we have $\lim_{\|t\| \to \infty} |\hat{G}(t)| < 1$. Therefore, given $0 < c < 1$ and $c_{12} > 0$ we have

$$
\sup_{\|t\| > cc_{12}/2} |\hat{G}(t)| = d < 1
$$

(4.9)

Next assume $t = [t_1, t_2]' \in \mathbb{R}^{k+1}$ satisfies $\|t\| > c_{12}$. Then $\hat{F}_i(t) = \hat{G}([t_1' n^{1/2} A^{-1/2} X_i t_2'])$ and either $\|t_1\| > c_{12}/2$ or $\|t_2\| > c_{12}/2$. If $\|t_1\| > c_{12}/2$ and $i \in B(t_1, c)$, then $\|t_1' n^{1/2} A^{-1/2} X_i\| > c \|t_1\| > cc_{12}/2$, and if $\|t_2\| > c_{12}/2$ then $\|t_2\| > cc_{12}/2$ (since $0 < c < 1$). These imply that $\|t_1' n^{1/2} A^{-1/2} X_i t_2\| > cc_{12}/2$. Therefore by (4.9) $|\hat{G}([t_1' n^{1/2} A^{-1/2} X_i t_2'])| < d$ and so $|\hat{F}_i(t)| < d$. \[\square\]
The next theorem gives an Edgeworth expansion for the distribution of $R_n$; its proof follows the lines of the proof of theorem 20.1 of BR by using Lemma 4.2, Theorem 4.1, and the relations we have just before Theorem 4.1.

**Theorem 4.3.** Assume the hypothesis of Theorem 4.1. Also assume that $s$ in $\tilde{A}$ is greater than $k + 1$. Then, using the same notations as before, we have

$$
\sup_{E \in \mathcal{B}} \left| P(BR_n \in E) - \int_E \left[ 1 + \sum_{r=1}^{r-3} n^{-r/2} P_r(-D, \{\tilde{x}_r\}) \right] \phi_k(x) \, dx \right|
$$

$$
= O\left( (n^{-s + 1/2} + 1) \right),
$$

where $\mathcal{B}$ is any class of Borel subsets of $\mathbb{R}^{k+1}$ satisfying

$$
\sup_{E \in \mathcal{B}} \Phi_k(\partial E) = O(1), \quad (\partial E)^c \text{ is the set of points in } \mathbb{R}^{k+1} \text{ within } \eta \text{ from the boundary of } E, \quad \text{and } \tilde{x}_r = (1/n) \sum_{i=1}^{n} x_i(BY_i).
$$

To get an Edgeworth expansion for the distribution of the studentized least squares $W_n$ we first give an expansion for $\tilde{W}_n$ defined by $\tilde{W}_n = A^{1/2}(\tilde{\beta} - \beta)/\tilde{\sigma}$, all notations being the same as before.

Define $H: \mathbb{R}^k x (\sigma^2, \infty) \to \mathbb{R}^k$ by $H([x_1, \ldots, x_{k+1}]) = [x_1, \ldots, x_k]'/(x_{k+1} + \sigma^2)^{1/2}$. Then clearly $\tilde{W}_n = n^{1/2} [H(\tilde{Y}) - H(0)]$.

Since $H$ is continuously differentiable up to order $s - 1$, one has the Taylor expansion of $H$ around 0,

$$
g_n(x) = n^{1/2} \left[ H(0 + x/n^{1/2}) - H(0) \right]
$$

$$
= \text{grad } H(0) x' + (n^{-1/2} 2!) \sum_{|a| = s - 1} D^a H(0) x^a
$$

$$
+ \cdots + \left[ n^{-(s-2)/2} / (s - 1)! \right] \sum_{|a| = s - 1} \left[ D^a H(0) + R_{x,n}(x) \right] x^a
$$

$$
= h_n(x) + \left[ n^{-(s-2)/2} / (s - 1)! \right] \sum_{|a| = s - 1} \left[ D^a H(0) + R_{x,n}(x) \right] x^a.
$$

Clearly $\tilde{W}_n = g_n(n^{1/2} \tilde{Y})$.

Let $\tilde{W}_n = h_n(n^{1/2} \tilde{Y})$ and $\kappa_v$ be the $v$th cumulant of $\tilde{W}_n$ after deleting all terms of $O(n^{-(s - 2)/2})$. The formal Edgeworth expansion of the characteristic function of $\tilde{W}_n$ is given by

$$
\hat{\tilde{W}}_{n,v}(t) = \left[ 1 + \sum_{r=1}^{s-3} n^{-r/2} \tilde{P}_r(z, \{\kappa_v\}) \right] \exp(-t \tilde{S}t/2),
$$

where $\tilde{S} = (\text{grad } H(0)) S(\text{grad } H(0))'$ and $\tilde{P}_r(z, \{\kappa_v\})$ is the coefficient of $n^{-r/2}$ in the expansion of $\exp\{\sum_{1 \leq |v| \leq s} (\kappa_v z^v/|v|!) + \|t\|^2/2\}$. Clearly, $\tilde{S} = I_k$ in our set up.
The next theorem follows from Theorem 2 of Bhattacharya and Ghosh [4] or Theorem 1 of Bhattacharya [3] and the proof is therefore omitted.

**Theorem 4.4.** Under the hypothesis of Theorem 4.3, one has

$$\sup_{E \in \mathcal{B}_0} \left| P(\bar{W}_n \in E) - \int_E \left[ 1 + \sum_{r=1}^{s-3} n^{-r/2} \bar{P}_r \left( -D, \{\kappa_r\} \right) \right] \varphi_k(x) \, dx \right| = o(n^{-(s-3)/2}),$$

where $\mathcal{B}_0$ is a class of Borel subsets of $\mathbb{R}^k$ satisfying $\sup_{E \in \mathcal{B}_0} \Phi_k(\partial E)^n = O(\eta)$.

Finally we are going to show that the two term expansion for the distribution of $\bar{W}_n$ is also valid for the studentized statistic $\bar{W}_n$. To show this, let $E \in \mathcal{B}_0$, and assume $s = \max \{4, k + 2\}$ in $A4$. Then by Theorem 4.4 we have

$$\left| P(W_n \in E) - \int_E \left[ 1 + \frac{1}{n^{1/2}} \bar{P}_1 \left( -D, \{\kappa_r\} \right) \right] \varphi_k(x) \, dx \right| = \left| P(W_n \in E) - P(\bar{W}_n \in E) \right| + o(n^{1/2}) \quad (4.11)$$

and

$$\left| P(W_n \in E) - P(\bar{W}_n \in E) \right| \leq P\left\{ (W_n \in E) \Delta (\bar{W}_n \in E) \right\} \leq P\{ \| W_n - \bar{W}_n \| \geq 1/(n^{1/2} \log n) \}$$

$$+ P\{ (W_n \in E) \Delta (\bar{W}_n \in E) \cap \{ \| W_n - \bar{W}_n \| < 1/(n^{1/2} \log n) \} \}$$

$$= P\{ \| W_n - \bar{W}_n \| \geq 1/(n^{1/2} \log n) \} + P\{ \bar{W}_n \in (\partial E)^{1/(n^{1/2} \log n)} \}$$

$$= P\{ \| W_n - \bar{W}_n \| \geq 1/(n^{1/2} \log n) \} + o(n^{-1/2}). \quad (4.12)$$

Also

$$P\{ \| W_n - \bar{W}_n \| \geq 1/(n^{1/2} \log n) \}$$

$$= P\{ \| A^{1/2}(\hat{\beta} - \beta) \| \cdot (1/\hat{\delta}) - (1/\hat{\delta}) \| > 1/(n^{1/2} \log n) \}$$

$$= P\{ \| A^{1/2}(\bar{\beta} - \beta)/\hat{\delta} \| \cdot 2n^{1/2}(\log n)^2 \cdot (|\hat{\delta} - \hat{\delta}/\hat{\delta}|/\hat{\delta}) > 2 \log n \}$$

$$\leq P\{ \| \bar{W}_n \| > (2 \log n)^{1/2} \} + P\{ (|\hat{\delta} - \hat{\delta}|/\hat{\delta}) > n^{-1/2}(\log n)^{-3/2} \}$$

$$= P\{ |\hat{\delta}^2 - \hat{\delta}^2| > (1/2) \hat{\delta}(\hat{\delta} + \hat{\delta}) n^{-1/2}(\log n)^{-3/2} \} + o(n^{-1/2}). \quad (4.13)$$
Now if one takes $\delta$ in A5 to be less than $\frac{1}{8}$, and $n$ large enough so that $n^{-1/4} < \sigma^2/2$ and (2.2) holds, one gets

$$P\{|\hat{\sigma}^2 - \sigma^2| > \frac{1}{2} \delta (\hat{\sigma} + \sigma)\ n^{-1/2} (\log n)^{-3/2}\}$$

$$\leq P\{|\hat{\sigma}^2 - \sigma^2| > \frac{1}{4} \sigma^2 n^{-1/2} (\log n)^{-3/2}\} + P\{|\hat{\sigma}^2 - \sigma^2| > n^{-1/4}\}$$

$$\leq 32 \mu_4 M_n^4 \cdot n (\log n)^3 (\hat{\lambda}_n^2 \sigma^4)$$

$$+ P\{|\hat{\sigma}^2 - \sigma^2| > n^{-1/4}\} \quad \text{(by Lemma 2.3)}$$

$$= P\{|\hat{\sigma}^2 - \sigma^2| > n^{-1/4}\} + o(n^{-1/2}). \quad (4.14)$$

Furthermore, using Lemma 2.3, one has

$$P\{|\hat{\sigma}^2 - \sigma^2| > n^{-1/4}\} \leq P\{|\hat{\sigma}^2 - \sigma^2| > \frac{1}{2} n^{-1/4}\} + P\{|\hat{\sigma}^2 - \sigma^2| > \frac{1}{4} n^{-1/4}\}$$

$$\leq 8 \mu_4 M_n^4 \cdot n^{1/2} \hat{\lambda}_n^2 + P\{|\hat{\sigma}^2 - \sigma^2| > \frac{1}{2} n^{-1/4}\}$$

$$- o(n^{-1/2}). \quad (4.15)$$

The last equality follows from A5 and the fact that $n^{1/2}(\hat{\sigma}^2 - \sigma^2)$ has a two term Edgeworth expansion if $E |\varepsilon_i|^6 < \infty$ and A1 holds.

Relations (4.11) through (4.15) imply that

$$E[W_n \in E] = \int_E [1 + (1/n^{1/2}) \bar{P}_1(-D, \{\kappa_i\})] \varphi_\lambda(x) \, dx = o(n^{-1/2}).$$

Thus we have arrived at the following theorem which gives a two term Edgeworth expansion for the studentized least squares estimate $W_n$.

**Theorem 4.5.** Assume the hypothesis of Theorem 4.4 with $\delta$ in A5 less than $\frac{1}{8}$ and $s = \max\{4, k+2\}$ in A4. Then $\sup_{E \in \mathcal{G}_0} |P(W_n \in E) - \int_E [1 + (1/n^{1/2}) \bar{P}_1(-D, \{\kappa_i\})] \varphi_\lambda(x) \, dx| = o(n^{-1/2})$, where $\mathcal{G}_0$ is as in Theorem 4.4.

**Remark 4.6.** In Theorem 4.5 if one uses for the population moments that appear in $\bar{P}_1(-D, \{\kappa_i\})$ other estimates which can be calculated from the data and which converge a.s. to the population moments, one obtains an approximate value for the distribution of the studentized least squares $A^{1/2}(\hat{\beta} - \beta)/\hat{\sigma}$ which is better than the normal approximation (for purposes of estimation and testing).

**Remark 4.7.** An Edgeworth expansion for the distribution of $\hat{\beta}$ given in (1.2) in the correlation model, where $X$ is considered random, is also valid under suitable assumptions such as those given in Theorem 2 of Bhattacharya and Ghosh [4] on the $(k + 1) \times 1$ vectors $[X_i, Y_i]$, where the notations are the same as before, again the two term expansion gives a better estimate than the normal one, as in the previous remark.
EDGEWORTH EXPANSIONS IN REGRESSION

ACKNOWLEDGMENT

I thank Professor R. N. Bhattacharya for his support and constant guidance throughout the preparation of this paper.

REFERENCES