The Annals of Statistics
1992, Vol. 20, No. 2, 1040–1061

AN ASYMPTOTIC THEORY FOR SLICED INVERSE REGRESSION

BY TAILEN HSING[1] AND RAYMOND J. CARROLL[2]

Texas A & M University

Sliced inverse regression [Li (1989), (1991) and Duan and Li (1991)] is a nonparametric method for achieving dimension reduction in regression problems. It is widely applicable, extremely easy to implement on a computer and requires no nonparametric smoothing devices such as kernel regression. If Y is the response and $X \in \mathbf{R}^p$ is the predictor, in order to implement sliced inverse regression, one requires an estimate of $\Lambda = E\{\mathrm{cov}(X|Y)\} = \mathrm{cov}(X) - \mathrm{cov}\{E(X|Y)\}$. The inverse regression of X on Y is clearly seen in Λ. One such estimate is Li's (1991) two-slice estimate, defined as follows: The data are sorted on Y, then grouped into sets of size 2, the covariance of X is estimated within each group and these estimates are averaged. In this paper, we consider the asymptotic properties of the two-slice method, obtaining simple conditions for $n^{1/2}$-convergence and asymptotic normality. A key step in the proof of asymptotic normality is a central limit theorem for sums of conditionally independent random variables. We also study the asymptotic distribution of Greenwood's statistics in nonuniform cases.

1. Introduction. Suppose we are given a sample (Y_i, X_i) for $i = 1, \dots, n$, where Y_i is the response and X_i is a vector of predictors of dimension p. Kernel regression [Eubank (1988)] is a popular method for estimating the mean function $m(x) = E(Y|X = x)$ and the variance function $v(x) = \mathrm{var}(Y|X = x)$. However, if p is large, it is well known that the kernel method is inefficient. This has led to the development of dimension reduction techniques, including projection pursuit regression [Friedman and Stuetzle (1981), Hall (1989)], average derivative estimation [Härdle and Stoker (1989)], generalized additive models [Hastie and Tibshirani (1986)] and so on.

All of the preceding methods involve some sort of nonparametric smoothing. Li (1991) took an entirely different and promising approach to dimension reduction, developing sliced inverse regression, or SIR.

The basic assumption of sliced inverse regression is that the distribution of Y given X depends only on K linear combinations of X, say $(\theta_1^T X, \dots, \theta_K^T X)$. This is a quite general formulation. In this paper, we provide relatively simple conditions under which Li's two-slice estimate of $\Theta = (\theta_1, \dots, \theta_K)$ converges at the rate $O_p(n^{-1/2})$.

Received March 1990; revised September 1991.
[1]Research supported by NSF Grant DMS-88-14006 and AFOSR Contract F49620 85C 0144.
[2]Research supported by grants from the National Institutes of Health.
AMS 1980 *subject classification*. Primary 62G05.
Key words and phrases. Dimension reduction, generalized linear models, Greenwood's statistic, projection pursuit, regression, sliced inverse regression.

The following is a short summary of sliced inverse regression. One should consult Duan and Li (1991), Li (1989, 1991), the references within and the discussion and rejoinder to Li (1991) for more details and background. The main idea of sliced inverse regression is to relate the inverse regression curve $E(X|Y)$ to Θ. Li (1991) makes the assumption that for any $b \in \mathbf{R}^p$, the conditional expectation $E(b^T X|\theta_1^T X, \dots, \theta_K^T X)$ is linear in $\theta_1^T X, \dots, \theta_K^T X$; in the rejoinder to his paper, Li (1991) discusses the wide applicability of this assumption. It thus follows (his Theorem 3.1) that the centered inverse regression curve $E(X|Y) - EX$ is contained in the linear subspace spanned by $\theta_k^T \Omega_X$ for $k = 1, \dots, K$, where Ω_X denotes the covariance matrix of X.

These facts enable us to identify and estimate Θ. Define

$$\Lambda = E\{\text{cov}(X|Y)\} = \text{cov}(X) - \text{cov}\{E(X|Y)\}. \tag{1.1}$$

If $(\eta_1, \dots, \eta_K)$ are the eigenvectors associated with the K largest nonzero eigenvalues of $I - \Omega_X^{-1/2}\Lambda\Omega_X^{-1/2}$, then $\theta_j = \Omega_X^{-1/2}\eta_j$ for $j = 1, \dots, K$. Since one can estimate Ω_X at rate $n^{1/2}$, the rate for estimating Θ is determined by how well one estimates Λ.

Li (1991) discusses two methods for estimating Θ. One method is based on categorizing the response Y into a new response Y_* with $H \geq K$ levels. As long as the distribution of Y_* given X can be minimally specified by the linear combinations of $\theta_k^T X$ for $k = 1, \dots, K$, estimation of (1.1) with Y replaced by Y_* is easy and yields $n^{1/2}$-consistent estimates of Θ. In this paper, we focus on the second method of estimating Λ, namely Li's two-slice estimate discussed in his Remark 5.3, a method which avoids categorization. Let $Y_{(1)} \leq \cdots \leq Y_{(n)}$ be the order statistics of the responses and define $X_{(i)*}$ to be the value of X associated with the ith order statistic in Y, that is, the concomitant value of the ith order statistic [Yang (1977)]. Then the two-slice estimate of Λ is

$$\hat{\Lambda}_n = n^{-1} \sum_{i=1}^{[n/2]} \{X_{(2i)*} - X_{(2i-1)*}\}\{X_{(2i)*} - X_{(2i-1)*}\}^T. \tag{1.2}$$

The main goal of this paper is to study the rate of convergence of $\hat{\Lambda}_n$ to Λ and the asymptotic distribution of $\hat{\Lambda}_n - \Lambda$. In Section 2, we show that simple conditions entail root n convergence and asymptotic normality.

Notice that an alternative estimator of Λ is

$$\tilde{\Lambda}_n = (2n)^{-1} \sum_{i=2}^{n} \{X_{(i)*} - X_{(i-1)*}\}\{X_{(i)*} - X_{(i-1)*}\}^T.$$

The proofs in Section 2 show that the asymptotic distributions of $\hat{\Lambda}_n$ and $\tilde{\Lambda}_n$ are essentially the same. If we specialize to the scalar case $p = 1$ and take $X = m(Y)$ for some monotone function m, then $\Lambda = E\{\text{cov}(X|Y)\} = 0$ and $\tilde{\Lambda}_n$ becomes $(2n)^{-1}$ times Greenwood's statistic:

$$T_n = \sum_{i=1}^{n-1} \{X_{(i+1)} - X_{(i)}\}^2. \tag{1.3}$$

Greenwood's statistic was introduced by Greenwood (1946) to test the hypothesis that the X's are uniformly distributed [cf. Read (1983, 1988) and Pyke (1965)]. Under uniformity, the distribution of T_n has been studied and its percentage points tabulated [cf. Stephens (1981)]. In Section 3, we investigate the asymptotic distribution of T_n for situations where the tail distribution of Y is regularly varying and exponential-like. Those results not only provide insight into the main theme of this paper, but also extend the traditional theory of Greenwood's statistics which considered mainly distributions with bounded support.

The Appendix includes a number of technical results, including a central limit theorem for sums of conditionally independent random variables which is of independent interest.

2. Asymptotics of $\hat{\Lambda}_n$. The outline of this section is as follows. In Theorem 2.1, we show that $\hat{\Lambda}_n$ is an $n^{1/2}$-consistent estimate of Λ, under simple and general conditions. Corollary 2.2 provides an asymptotic expansion of $\hat{\Lambda}_n$ about Λ in terms of the concomitants $\varepsilon_{(i)*}$. Theorem 2.3 uses this expansion to obtain an asymptotic normal limit distribution for $\hat{\Lambda}$. We make strong use here of Theorem A.4 in the Appendix, which provides limit distributions for sums of conditionally independent random variables.

Define

$$m(y) = E(X|Y=y),$$
$$\varepsilon = X - m(Y).$$

Also let $\varepsilon_{(i)*}$ be the concomitant of the ith order statistic of the Y's. These concomitants are conditionally independent with mean zero given the order statistics of the Y's [cf. Yang (1977)].

For convenience of notation we will use the absolute value sign to denote the norm of $\mathbf{R}^k$ for any finite k. For $B > 0$ and $n \geq 1$, let $\Pi_n(B)$ be the collection of all the n-point partitions $-B \leq y_{(1)} \leq \cdots \leq y_{(n)} \leq B$ of $[-B, B]$. For a function G defined on $\mathbf{R}$ and taking values in $\mathbf{R}^k$, define the smoothness condition:

$$\lim_{n\to\infty} \sup_{\mathbf{y}\in\Pi_n(B)} n^{-1/r} \sum_{i=2}^{n} \left|G(y_{(i)}) - G(y_{(i-1)})\right| = 0, \qquad \forall B > 0, \tag{2.1}$$

where r is some positive constant. Recall that a function G on $[-B, B]$ is bounded variation if its total variation norm

$$\lim_{n\to\infty} \sup_{y\in\Pi_n(B)} \sum_{i=2}^{n} \left|G(y_{(i)}) - G(y_{(i-1)})\right|$$

is finite. Thus (2.1) is weaker than bounded variation. Condition (2.1) clearly holds for any $r > 0$ if G has a bounded derivative on every finite interval. For a continuous G, (2.1) holds for $r \geq 1$.

THEOREM 2.1. *Suppose that* $E|X|^4 < \infty$, *that* (2.1) *holds for* $G = m$ *and* $r = 4$ *and that for* $B_0 > 0$, *there exists a nondecreasing (real-valued) function* $\tilde{m}$ *on* (B_0, ∞), *such that*

$$\begin{aligned} &\tilde{m}^4(y) P\{|Y| > y\} \to 0 \quad \text{as } y \to \infty, \\ &|m(x) - m(y)| \le |\tilde{m}(|x|) - \tilde{m}(|y|)| \quad \text{for } x, y \in (-\infty, B_0) \text{ or } (B_0, \infty). \end{aligned} \tag{2.2}$$

Then

$$n^{1/2}(\hat{\Lambda}_n - \Lambda) = O_p(1). \tag{2.3}$$

REMARKS. (i) The moment condition $E|X|^4 < \infty$ in the above result is essentially required without further restriction on m. To see this, consider the case where $X = m(Y)$ for some monotone function m (so that $\Lambda = 0$). If for some $\delta > 0$, $P\{X > x\} = O(x^{-4-\delta})$ as $x \to \infty$, then it follows from Theorem 3.1 below that $n^{1/2}(\hat{\Lambda}_n - \Lambda) \neq O_p(1)$, violating (2.3). See the remark after Theorem 3.1.

(ii) Whereas (2.1) is a local condition, (2.2) is a global condition. Note (2.2) is quite mild under $E|X|^4 < \infty$. To see this, we will argue heuristically by considering the situation where m does not behave erratically so that $\tilde{m}$ can in fact be taken to satisfy $\tilde{m}(|y|) = O(|m(y)| + |m(-y)|)$ as $y \to \infty$. Since

$$E|m(Y)|^4 = E\big(|E\{X|Y\}|^4\big) \le E\big(E\{|X|^4 \big| Y\}\big) = E|X|^4 < \infty,$$

it is then plausible to assume also that $E\tilde{m}^4(|Y|) < \infty$. Thus by monotonicity of $\tilde{m}$,

$$\begin{aligned} \lim_{y \to \infty} \tilde{m}^4(y) P\{|Y| > y\} &\le \lim_{y \to \infty} \tilde{m}^4(y) P\{\tilde{m}^4(|Y|) \ge \tilde{m}^4(y)\} \\ &= \lim_{u \to \infty} u P\{\tilde{m}^4(|Y|) > u\} = 0. \end{aligned}$$

(iii) Clearly if the components of m are monotone, then (2.3) holds under $E|X|^4 < \infty$.

PROOF OF THEOREM 2.1. It suffices to consider the case that X is scalar, the general case being similar. Since $X = m(Y) + \varepsilon$, we have that

$$\begin{aligned} n^{1/2}(\hat{\Lambda}_n - \Lambda) &= n^{-1/2} \sum_{i=1}^{[n/2]} \{m(Y_{(2i)}) - m(Y_{(2i-1)})\}^2 \\ &\quad + n^{-1/2} \sum_{i=1}^{[n/2]} \Big[\{\varepsilon_{(2i)*} - \varepsilon_{(2i-1)*}\}^2 - \Lambda\Big] \\ &\quad + 2n^{-1/2} \sum_{i=1}^{[n/2]} \{m(Y_{(2i)}) - m(Y_{(2i-1)})\}\{\varepsilon_{(2i)*} - \varepsilon_{(2i-1)*}\} \\ &= C_1 + C_2 + C_3. \end{aligned}$$

$C_1 = o_p(1)$ by Lemma A.3. Since the concomitants $\varepsilon_{(i)*}$ are independent with mean zero given the order statistics, it is easy to see that C_2 has mean zero and that its variance is $O(1)$ as long as ε has finite fourth moment and

$$W_n = n^{-1} \sum_{i=1}^{[n/2]} E\left(\varepsilon_{(2i)*}^2 \varepsilon_{(2i-1)*}^2\right) = O(1).$$

By Cauchy–Schwarz, it is easy to see that $Em^4(Y) \le E(E(X^4|Y)) = E(X^4) < \infty$ and hence $E\varepsilon^4 < \infty$ and $|W_n| \le n^{-1}\sum_{i=1}^n E(\varepsilon_{(i)*}^4) = E(\varepsilon^4) < \infty$. Finally, by Lemma A.1 and Lemma A.3,

$$\begin{aligned} C_3 &\le 2n^{-1/2} \sum_{i=1}^{[n/2]} \left| m(Y_{(2i)}) - m(Y_{(2i-1)}) \right| \left| \varepsilon_{(2i)*} - \varepsilon_{(2i-1)*} \right| \\ &\le 2n^{-1/2}(\varepsilon_{(n)} - \varepsilon_{(1)}) \sum_{i=1}^{[n/2]} \left| m(Y_{(2i)}) - m(Y_{(2i-1)}) \right| = o_p(1), \end{aligned}$$

where $\varepsilon_{(i)}$ is the ith order statistic (not the concomitant) of the ε's. This concludes the proof. □

We have also proved the following corollary.

COROLLARY 2.2.

$$\begin{aligned} & n^{1/2}\left(\hat{\Lambda}_n - \Lambda\right) \\ &\quad = n^{-1/2} \sum_{i=1}^{[n/2]} (\varepsilon_{(2i)*} - \varepsilon_{(2i-1)*})(\varepsilon_{(2i)*} - \varepsilon_{(2i-1)*})^T - n^{1/2}\Lambda + o_p(1) \\ &\quad = S_n + o_p(1). \end{aligned}$$

The study of the asymptotic distribution of S_n is facilitated by Corollary 2.2 and the following notation. For a symmetric $(p \times p)$ matrix $D = (d^{(jk)})$, let $\mathrm{vech}(D) = (d^{(11)}, \dots, d^{(p1)}, d^{(22)}, \dots, d^{(pp)})^T$ be the $((p(p+1))/2 \times 1)$ vector of the unique elements of D. Let $\mathrm{vec}(D) = (d^{(11)}, \dots, d^{(p1)}, d^{(21)}, d^{(22)}, \dots, d^{(pp)})^T$ be the $(p^2 \times 1)$ vector of all elements of D. There exist matrices Ψ and Γ such that $\mathrm{vech}(D) = \Psi\, \mathrm{vec}(D)$ and, for any vectors a, b with $a = (a^{(1)}, \dots, a^{(p)})^T$, $\mathrm{vec}(ab^T) = \Gamma\, \mathrm{vec}(ba^T)$. For any vector λ, define the matrix $C = (c^{(jk)})$ by

$$\lambda^T \mathrm{vech}(ab^T + ba^T) = \lambda^T \Psi(I + \Gamma)\mathrm{vec}(ab^T) = \sum_j \sum_k c^{(jk)} a^{(j)} b^{(k)}.$$

Define

$$V(y) = E\left(\varepsilon\varepsilon^T \middle| Y = y\right); \qquad R(y) = \mathrm{cov}\left\{\mathrm{vech}(\varepsilon\varepsilon^T) \middle| Y = y\right\};$$

$$\mathscr{G}(y, z) = \sum_{j,k,l,m} c^{(jk)} c^{(lm)} V^{(jl)}(y) V^{(km)}(z),$$

$$\sigma_\lambda^2 = \lambda^T \mathrm{cov}[\mathrm{vech}\{V(Y)\}]\lambda + \lambda^T E\left[\mathrm{cov}\left\{\mathrm{vech}(\varepsilon\varepsilon^T) \middle| Y\right\}\right]\lambda + \tfrac{1}{2} E\mathscr{G}(Y, Y).$$

THEOREM 2.3. *In addition to the assumptions of Theorem* 2.1, *suppose that for all* $1 \le i, j \le p$ *and* $1 \le k, l \le (p(p+1))/2$, (2.1) *holds for* $r = 1$, *and* $G = V^{(ij)}$ *and* $R^{(kl)}$. *Then* $n^{1/2}(\hat{\Lambda}_n - \Lambda) \to_d Z$, *where, for every* $(p(p+1))/2$-*dimensional vector* λ, $\lambda^T \operatorname{vech}(Z) \sim N(0, \sigma_\lambda^2)$.

PROOF. We will apply Theorem A.4 with $s_n = n^{1/2}$, $m_n = [n/2]$, $\mathscr{F}_n$ = σ-field generated by $(Y_1, \dots, Y_n)$ and

$$X_{ni} = \lambda^T \operatorname{vech}\Big\{(\varepsilon_{(2i)*} - \varepsilon_{(2i-1)*})(\varepsilon_{(2i)*} - \varepsilon_{(2i-1)*})^T - 2\Lambda\Big\}.$$

By Corollary 2.2, $n^{1/2}\lambda^T \operatorname{vech}(\hat{\Lambda}_n - \Lambda) = n^{-1/2}\sum_1^{[n/2]} X_{ni} + o_p(1)$. We will consider only the case that n is even. In this framework, (A.9) holds by definition. Also, (A.12) is easily handled, as

$$\begin{aligned}
&s_n^{-1} \sum_1^{m_n} E(X_{ni}|\mathscr{F}_n)\\
&= n^{-1/2} \sum_1^{[n/2]} \lambda^T \operatorname{vech}\Big\{E\Big(\varepsilon_{(2i)*}\varepsilon_{(2i)*}^T + \varepsilon_{(2i-1)*}\varepsilon_{(2i-1)}^T - 2\Lambda\Big|\mathscr{F}_n\Big)\Big\}\\
&= n^{-1/2} \sum_1^{n} \lambda^T \operatorname{vech}\{V(Y_i) - \Lambda\}\\
&\Rightarrow_{\mathscr{L}} N\big(0, \lambda^T \Sigma_1 \lambda\big),
\end{aligned}$$

where $\Sigma_1 = \operatorname{cov}[\operatorname{vech}\{V(Y)\}]$. It thus remains to verify (A.10) and (A.11). With $\tilde{X}_{nk} = X_{nk} - E(X_{nk}|\mathscr{F}_n)$, we see that $E(\tilde{X}_{nk}^2|\mathscr{F}_n) = \operatorname{Var}(X_{nk}|\mathscr{F}_n)$, so that by conditional independence of the concomitants given $\mathscr{F}_n$ and since $E(\varepsilon|Y) = 0$, we find that the l.h.s. of (A.10) is

$$n^{-1} \sum_1^{[n/2]} E\Big(\tilde{X}_{nk}^2\Big|\mathscr{F}_n\Big) = n^{-1} \sum_{i=1}^{n} \operatorname{Var}\big\{\lambda^T \operatorname{vech}\big(\varepsilon_i \varepsilon_i^T\big)|Y_i\big\} + n^{-1} \sum_1^{[n/2]} \mathscr{G}_i,$$

where

$$\begin{aligned}
\mathscr{G}_i &= \mathscr{G}\big(Y_{(2i)}, Y_{(2i-1)}\big)\\
&= \operatorname{Var}\Big\{\lambda^T \operatorname{vech}\Big(\varepsilon_{(2i)*}\varepsilon_{(2i-1)*}^T + \varepsilon_{(2i-1)*}\varepsilon_{(2i)*}^T\Big)\Big|\mathscr{F}_n\Big\}\\
&= \operatorname{Var}\left\{\sum_j \sum_k c^{(jk)} \varepsilon_{(2i)*}^{(j)} \varepsilon_{(2i-1)*}^{(k)} \middle| \mathscr{F}_n\right\}\\
&= \sum_j \sum_k \sum_l \sum_m c^{(jk)} c^{(lm)} V^{(jl)}\big(Y_{(2i)}\big) V^{(km)}\big(Y_{(2i-1)}\big).
\end{aligned}$$

Observe that $E|X|^4 < \infty$ implies

$$E|\varepsilon|^4 < \infty, \qquad E\big|V(Y)\big|^2 < \infty \quad \text{and} \quad E\big|R(Y)\big| < \infty. \tag{2.4}$$

Thus it follows from Lemma A.2 that

$$\left| n^{-1} \sum_1^{[n/2]} \mathscr{G}(Y_{(2i)}, Y_{(2i-1)}) - n^{-1} \sum_1^{[n/2]} \mathscr{G}(Y_{(2i)}, Y_{(2i)}) \right|$$
$$\leq n^{-1} \sum_j \sum_k \sum_l \sum_m c^{(jk)} c^{(lm)} V^{(jl)}(Y_{(2i)}) \left| V^{(km)}(Y_{(2i)}) - V^{(km)}(Y_{(2i-1)}) \right|$$
$$\to_p 0$$

by the smoothness and finite second moment condition on the components of V. Thus

$$(2.5) \qquad \begin{aligned} n^{-1} \sum_1^{[n/2]} \mathscr{G}(Y_{(2i)}, Y_{(2i-1)}) &= n^{-1} \sum_1^{[n/2]} \mathscr{G}(Y_{(2i)}, Y_{(2i)}) + o_p(1) \\ &= n^{-1} \sum_1^{[n/2]} \mathscr{G}(Y_{(2i-1)}, Y_{(2i-1)}) + o_p(1). \end{aligned}$$

Also by (2.4), $E|\mathscr{G}(Y, Y)| < \infty$, so that summing the last two terms in (2.5), we have $n^{-1}\sum_1^{[n/2]} \mathscr{G}_i \to_p (1/2) E\mathscr{G}(Y, Y)$. Hence, (A.10) holds with

$$\sigma^2 = E\left[\operatorname{Var}\{\lambda^T \operatorname{vech}(\varepsilon\varepsilon^T) \big| Y\}\right] + \tfrac{1}{2} E\mathscr{G}(Y, Y).$$

The proof is completed by showing (A.11), which is done in Lemma A.5 in the Appendix. □

3. Asymptotics of Greenwood's statistic. An important purpose of this section is to show that the assumption $E|X|^4 < \infty$ in Theorem 2.3 is essentially necessary. To illustrate the point we consider the naive setting that $p = 1$ and $X = m(Y)$ for some monotone function m. As pointed out in the Introduction, the asymptotic distributional behavior of $\hat{\Lambda}_n$ is then closely related to that of Greenwood's statistic T_n defined in (1.3). The distribution of T_n is generally quite difficult to handle [cf. Moran (1947) and Pyke (1965)]. In this section, we single out some special cases for which the asymptotic distribution of T_n can be precisely determined. As a result, for those cases, we can monitor the precise behavior of $\hat{\Lambda}_n$ and in turn gain understanding of the role of the finite fourth moment condition in Theorem 2.3.

Throughout let $X_1, \dots, X_n$ denote a sample of i.i.d. random variables and let T_n be as defined in (1.3). It is easy to observe that when the X's do not have bounded support, the extreme values of the sample will have a dominant effect on the asymptotic distribution of T_n. Therefore, to study the asymptotic distribution of T_n, the conditions appropriate are those that describe tail probabilities of X. In this paper we are obviously not interested in giving a complete general theory in that regard. With the main purpose described previously in mind, we consider the situations where tail probabilities of X are regularly varying and exponential-like.

In the following, let E_i^-, E_i^+, $i \geq 1$, be i.i.d. unit exponential random variables and define

$$\Gamma_j^- = \sum_1^j E_i^-, \qquad \Gamma_j^+ = \sum_1^j E_j^+.$$

A function f is said to be regularly varying at ∞ with index γ if

$$f(x) = x^\gamma L(x), \qquad x > 0,$$

for some slowly varying function L [cf. Feller (1971)]. We now investigate the asymptotic distribution of T_n assuming that both tails of X are regularly varying with the same index. Examples of distributions satisfying this are the nonnormal stable distributions.

THEOREM 3.1. *Suppose $1 - F(x)$ and $F(-x)$ are regularly varying with index $-\alpha < 0$. Suppose further that $F(-x)/(1 - F(x)) \to p/(1 - p)$ as $x \to \infty$, where $0 \leq p \leq 1$. Write $F_{|X|}(x) = P\{|X| \leq x\}$ and let its left-continuous inverse function be $F_{|X|}^{-1}(x)$. Then, as $n \to \infty$,*

$$\left\{F_{|X|}^{-1}\left(1 - \frac{1}{n}\right)\right\}^{-2} T_n \to_d p^{2/\alpha} \sum_{j=1}^{\infty} \left\{\left(\Gamma_j^-\right)^{-1/\alpha} - \left(\Gamma_{j+1}^-\right)^{-1/\alpha}\right\}^2$$

$$+ (1 - p)^{2/\alpha} \sum_{j=1}^{\infty} \left\{\left(\Gamma_j^+\right)^{-1/\alpha} - \left(\Gamma_{j+1}^+\right)^{-1/\alpha}\right\}^2.$$

PROOF. First write

$$\frac{1}{\{F_{|X|}^{-1}(1 - (1/n))\}^2} \sum_{i=1}^{n-1} (X_{(i+1)} - X_{(i)})^2$$

$$= \frac{1}{\{F_{|X|}^{-1}(1 - (1/n))\}^2} \left\{ \sum_{i=1}^{k-1} (X_{(i+1)} - X_{(i)})^2 + \sum_{i=k}^{n-k} (X_{(i+1)} - X_{(i)})^2 + \sum_{i=n-k+1}^{n-1} (X_{(i+1)} - X_{(i)})^2 \right\}$$

$$=: T_{n,1}(k) + T_{n,2}(k) + T_{n,3}(k).$$

By a well-known result in extreme value theory [cf. Proposition 1.11 and Corollary 4.19 of Resnick (1987)],

$$\left(-\frac{X_{(1)}}{F_{|X|}^{-1}(1 - (1/n))}, \dots, -\frac{X_{(k-1)}}{F_{|X|}^{-1}(1 - (1/n))}, \frac{X_{(n)}}{F_{|X|}^{-1}(1 - (1/n))}, \dots, \frac{X_{(n-k+1)}}{F_{|X|}^{-1}(1 - (1/n))} \right)$$

$$\to_d \left(\left\{\frac{\Gamma_1^-}{p}\right\}^{-1/\alpha}, \dots, \left\{\frac{\Gamma_{k-1}^-}{p}\right\}^{-1/\alpha}, \left\{\frac{\Gamma_1^+}{1-p}\right\}^{-1/\alpha}, \dots, \left\{\frac{\Gamma_{k-1}^+}{1-p}\right\}^{-1/\alpha} \right).$$

Applying the continuous mapping theorem [Billingsley (1968), Theorem 5.1], as $n \to \infty$,

$$T_{n,1}(k) + T_{n,3}(k) \to_d p^{2/\alpha} \sum_{j=1}^{k-1} \left\{ \left(\Gamma_j^-\right)^{-1/\alpha} - \left(\Gamma_{j+1}^-\right)^{-1/\alpha} \right\}^2$$

$$+ (1-p)^{2/\alpha} \sum_{j=1}^{k-1} \left\{ \left(\Gamma_j^+\right)^{-1/\alpha} - \left(\Gamma_{j+1}^+\right)^{-1/\alpha} \right\}^2,$$

and this limit converges almost surely to

$$p^{2/\alpha} \sum_{j=1}^{\infty} \left\{ \left(\Gamma_j^-\right)^{-1/\alpha} - \left(\Gamma_{j+1}^-\right)^{-1/\alpha} \right\}^2 + (1-p)^{2/\alpha} \sum_{j=1}^{\infty} \left\{ \left(\Gamma_j^+\right)^{-1/\alpha} - \left(\Gamma_{j+1}^+\right)^{-1/\alpha} \right\}^2$$

as $k \to \infty$ since

$$\sum_{k+1}^{\infty} \left\{ \left(\Gamma_j^\pm\right)^{-1/\alpha} - \left(\Gamma_{j+1}^\pm\right)^{-1/\alpha} \right\}^2 \le \left(\Gamma_{k+1}^\pm\right)^{-2/\alpha} \to_{\text{a.s}} 0 \quad \text{as } k \to \infty.$$

Also note that

$$T_{n,2}(k) \le \left\{ \frac{X_{(n-k+1)}}{F_{|X|}^{-1}(1-(1/n))} - \frac{X_{(k)}}{F_{|X|}^{-1}(1-(1/n))} \right\}^2$$

$$\to_d \left\{ (1-p)^{1/\alpha} \left(\Gamma_{k-1}^+\right)^{-1/\alpha} - p^{1/\alpha} \left(\Gamma_k^-{}^1\right)^{-1/\alpha} \right\}^2 \quad \text{as } n \to \infty$$

$$\to_{\text{a.s.}} 0 \quad \text{as } k \to \infty.$$

Thus the result follows from Billingsley [(1968), Theorem 4.2]. □

Thus for the situation that $p = 1$ and $X = m(Y)$, where m is some monotone function and X satisfies the assumption of Theorem 3.1, we have

$$n^{1/2}\hat{\Lambda}_n = \frac{\left\{F_{|X|}^{-1}(1-(1/n))\right\}^2}{n^{1/2}} \left\{ F_{|X|}^{-1}\left(1 - \frac{1}{n}\right) \right\}^{-2} T_n$$

$$= \frac{\left\{F_{|X|}^{-1}(1-(1/n))\right\}^2}{n^{1/2}} O_p(1).$$

Thus $n^{1/2}\hat{\Lambda}_n \to 0$ if $\alpha > 4$ and is stochastically unbounded if $\alpha < 4$.

Next we consider a different situation. First assume that F is unit exponential. By Renyi's representation of order statistics [cf. David (1981)], denoting by $\{E_i\}$ a sequence of i.i.d. unit exponential random variables,

$$\left(X_{(k)}, 1 \le k \le n\right) =_d \left(\sum_{1}^{k} E_j/(n-j+1), 1 \le k \le n \right).$$

Therefore,

$$\sum_{1}^{n-1} (X_{(i+1)} - X_{(i)})^2 =_d \sum_{1}^{n-1} \left(\frac{E_{i+1}}{n-i} \right) =_d \sum_{1}^{n-1} \left(\frac{E_i}{i} \right)^2.$$

Since

$$E \sum_{1}^{\infty} \left(\frac{E_i}{i} \right)^2 = 2 \sum_{1}^{\infty} \frac{1}{i^2} < \infty,$$

we obtain

$$\sum_{1}^{n-1} (X_{(i+1)} - X_{(i)})^2 \to_d \sum_{1}^{\infty} \left(\frac{E_i}{i} \right)^2. \tag{3.1}$$

It is interesting to note that the above elegant argument hides the fact that (3.1) depends on the exponential distribution assumption in a rather mild way. In the following we will show that (3.1) holds for a class of distributions whose tails are exponential-like. Instead of using Renyi's representation which only applies to the exponential distribution, we use the domain of attraction theory.

Consider a distribution G such that $G(x) < 1$ for all x and

$$\overline{G}(x) = 1 - G(x) = c(x)\exp\left(- \int_{-\infty}^{x} dt/\phi(t) \right), \tag{3.2}$$

where $c(x) \to$ some positive constant as $x \to \infty$, ϕ is positive and differentiable with $\phi'(x) \to 0$ as $x \to \infty$. (3.2) is sometimes called the von Mises condition. A distribution function G satisfying the condition can be shown to be in the domain of attraction of the Gumbel distribution; namely, there exist $a_n > 0$ and b_n such that as $n \to \infty$,

$$G^n(a_n x + b_n) \to \exp(-e^{-x}), \qquad -\infty < x < \infty$$

[see Resnick (1987), Proposition 1.1]. We will confine to a subclass $\mathscr{E}$ of such distributions G for which $\phi(x) \to \phi > 0$ as $x \to \infty$, c is differentiable from some point on and

$$\liminf_{x \to \infty} \left(\frac{1}{\phi(x)} - \frac{c'(x)}{c(x)} \right) > 0.$$

For any $G \in \mathscr{E}$, it is clear that for Δ large enough, there exists $\delta > 0$ such that

$$\frac{\overline{G}(x+v)}{\overline{G}(v)} = \exp\left\{ - \int_{v}^{x+v} \left(\frac{1}{\phi(t)} - \frac{c'(t)}{c(t)} \right) dt \right\} \le \exp(-\delta x), \tag{3.3}$$

$$x > 0, v > \Delta.$$

The class $\mathscr{E}$ contains many distributions of interest, including the exponential distribution, the Gumbel distribution, a class of subexponential distributions and so on. However, $\mathscr{E}$ does not contain, for example, the normal or the log normal distributions [cf. Bingham, Goldie and Teugels (1987)].

Suppose henceforth that both $F(x)$ and $1 - F(-x)$ are in $\mathscr{E}$ with pairs of functions (c_+, ϕ_+) and (c_-, ϕ_-), respectively, in the representation (3.2). Also let δ and Δ be such that the inequality in (3.3) holds for $G(x)$ equal to both $F(x)$ and $1 - F(x)$. We consider the asymptotic distribution of $T_n = \sum_{j=1}^{n-1}(X_{(i+1)} - X_{(i)})^2$.

LEMMA 3.2. *For $\varepsilon \in (0,1)$ with $|F^{-1}(\varepsilon)| \vee |F^{-1}(1-\varepsilon)| > \Delta$, we have*

$$\lim_{k\to\infty} \limsup_{n\to\infty} P\left\{ \sum_{i=[n(1-\varepsilon)]}^{n-k} \left(X_{(i+1)} - X_{(i)}\right)^2 > \eta \right\} = 0,$$
$$(3.4)\qquad \lim_{k\to\infty} \limsup_{n\to\infty} P\left\{ \sum_{i=k}^{[n\varepsilon]} \left(X_{(i+1)} - X_{(i)}\right)^2 > \eta \right\} = 0$$

for all $\eta > 0$.

PROOF. We shall only prove (3.4). Define $X^{(\Delta)} = XI(|X| \le \Delta)$ and $\overline{X}^{(\Delta)} = X - X^{(\Delta)}$. Clearly,

$$\sum_{i=[n(1-\varepsilon)]}^{n-k} \left(X_{(i+1)} - X_{(i)}\right)^2 \le 2 \sum_{i=[n(1-\varepsilon)]}^{n-k} \left(X^{(\Delta)}_{(i+1)} - X^{(\Delta)}_{(i)}\right)^2 + 2 \sum_{i=[n(1-\varepsilon)]}^{n-1} \left(\overline{X}^{(\Delta)}_{(i+1)} - \overline{X}^{(\Delta)}_{(i)}\right)^2.$$

It follows from (3.3) that $X_{[n(1-\varepsilon)]} \to_p F^{-1}(1-\varepsilon)$. Thus

$$P\left\{ \sum_{i=[n(1-\varepsilon)]}^{n-k} \left(X^{(\Delta)}_{(i+1)} - X^{(\Delta)}_{(i)}\right)^2 > \eta \right\} \le P\{X_{([n(1-\varepsilon)])} > \Delta\} \to 0,$$

showing that $\sum_{i=[n(1-\varepsilon)]}^{n-k}(X^{(\Delta)}_{(i+1)} - X^{(\Delta)}_{(i)})^2 \to_p 0$ as $n \to \infty$. Thus, to show (3.4), it suffices to show that

$$(3.5)\qquad \lim_{k\to\infty} \limsup_{n\to\infty} \sum_{i=[n(1-\varepsilon)]}^{n-k} E\left(\overline{X}^{(\Delta)}_{(i+1)} - \overline{X}^{(\Delta)}_{(i)}\right)^2 = 0.$$

Integrating by parts,

$$E\left(\overline{X}^{(\Delta)}_{(i+1)} - \overline{X}^{(\Delta)}_{(i)}\right)^2 = 2\int_0^\infty xP\left\{\overline{X}^{(\Delta)}_{(i+1)} - \overline{X}^{(\Delta)}_{(i)} > x\right\} dx.$$

But, by (3.3),

$$\begin{aligned} P\left\{\overline{X}^{(\Delta)}_{(i+1)} - \overline{X}^{(\Delta)}_{(i)} > x\right\} &= P\{X_{(i+1)} - X_{(i)} > x, X_{(i)} > \Delta\} \\ &= \int_{v=\Delta}^\infty P\left\{X_{(i+1)} - X_{(i)} > x \mid X_{(i)} = v\right\} dF_{X_{(i)}}(v) \\ &= \int_{v=\Delta}^\infty \left[\frac{\overline{F}(x+v)}{\overline{F}(v)}\right]^{n-i} dF_{X_{(i)}}(v) \le \exp(-(n-i)\delta x). \end{aligned}$$

Hence

$$\sum_{i=[n(1-\varepsilon)]}^{n-k} E\left(\overline{X}_{(i+1)}^{(\Delta)} - \overline{X}_{(i)}^{(\Delta)}\right)^2 \le 2 \sum_{i=[n(1-\varepsilon)]}^{n-k} \int_0^\infty x e^{-(n-i)\delta x}\, dx$$

$$= \frac{2}{\delta^2} \sum_{i=[n(1-\varepsilon)]}^{n-k} \frac{1}{(n-i)^2} \le \frac{2}{\delta^2} \sum_k^\infty \frac{1}{i^2},$$

from which (3.5) follows. □

LEMMA 3.3. *For every fixed $k \ge 2$,*

$$\sum_{i=1}^{k-1} \left(X_{(i+1)} - X_{(i)}\right)^2 + \sum_{i=n-k+1}^{n-1} \left(X_{(i+1)} - X_{(i)}\right)^2$$

$$\to_d \phi_- \sum_{j=1}^{k} \left(-\log \Gamma_j^- + \log \Gamma_{j+1}^-\right)^2 + \phi_+ \sum_{j=1}^{k} \left(-\log \Gamma_j^+ + \log \Gamma_{j+1}^+\right)^2.$$

PROOF. The proof follows from the fact [cf. Proposition 1.1 and Corollary 4.19 of Resnick (1987)]

$$\left(\frac{-X_{(1)} + F^{-1}(1/n)}{\phi_-}, \dots, \frac{-X_{(k)} + F^{-1}(1/n)}{\phi_-}, \frac{X_{(n)} - F^{-1}(1-(1/n))}{\phi_+}, \right.$$

$$\left. \dots, \frac{X_{(n-k+1)} - F^{-1}(1-(1/n))}{\phi_+}\right)$$

$$\to_d \left(-\log \Gamma_1^-, \dots, -\log \Gamma_k^-, -\log \Gamma_1^+, \dots, -\log \Gamma_k^+\right) \quad \text{as } n \to \infty,$$

and the continuous mapping theorem. □

LEMMA 3.4. *Let E_i, $i \ge 1$, be i.i.d. unit exponentials and $\Gamma_j = \sum_i^j E_i$. Then $\sum_{j=1}^i (-\log \Gamma_j + \log \Gamma_{j+1})^2 \to_p \sum_{j=1}^\infty (-\log \Gamma_j + \log \Gamma_{j+1})^2$ as $i \to \infty$.*

PROOF.

$$E\left(-\log \Gamma_j + \log \Gamma_{j+1}\right)^2 = E\left[\log\left(1 + \frac{E_{j+1}}{\Gamma_j}\right)\right]^2 \le E\left(\frac{E_{j+1}^2}{\Gamma_j^2}\right) = E\left(E_{i+1}^2\right) E\left(\frac{1}{\Gamma_j^2}\right)$$

$$= 2\int_0^\infty \frac{1}{t^2} \frac{t^{j-1}}{\Gamma(j)} e^{-t}\, dt = 2\frac{\Gamma(j-2)}{\Gamma(j)} \int_0^\infty \frac{t^{j-3}}{\Gamma(j-2)} e^{-t}\, dt$$

$$= \frac{2}{(j-1)(j-2)}, \qquad j \ge 3.$$

Thus

$$E \sum_{j=k}^{\infty} (-\log \Gamma_j + \log \Gamma_{j+1})^2 \le 2 \sum_{j=k}^{\infty} \frac{1}{(j-1)(j-2)} \to 0 \quad \text{as } k \to \infty,$$

showing the lemma. □

THEOREM 3.5. *Suppose both $F(x)$ and $1 - F(-x)$ are in $\mathscr{E}$ and F^{-1} is continuous. Then*

$$T_n \to_d \phi_+ \sum_{j=1}^{\infty} (-\log \Gamma_j^+ + \log \Gamma_{j+1}^+) + \phi_- \sum_{j=1}^{\infty} (-\log \Gamma_j^- + \log \Gamma_{j+1}^-)^2.$$

PROOF. Let $\varepsilon > 0$ be as in Lemma 3.2. Now write

$$\begin{aligned} T_n &= \sum_{i=1}^{n-1} (X_{(i+1)} - X_{(i)})^2 \\ &= \sum_{i=1}^{k-1} (X_{(i+1)} - X_{(i)})^2 + \sum_{i=k}^{[n\varepsilon]} (X_{(i+1)} - X_{(i)})^2 + \sum_{i=[n\varepsilon]+1}^{[n(1-\varepsilon)]-1} (X_{(i+1)} - X_{(i)})^2 \\ &\quad + \sum_{i=[n(1-\varepsilon)]}^{n-k} (X_{(i+1)} - X_{(i)})^2 + \sum_{i=n-k+1}^{n-1} (X_{(i+1)} - X_{(i)})^2 \\ &=: \sum_{m=1}^{5} T_{n,m}(k). \end{aligned}$$

Since, by Lemmas 3.3 and 3.4,

$$\begin{aligned} T_{n,1}(k) + T_{n,5}(k) &\to_d \phi_- \sum_{j=1}^{k-1} (-\log \Gamma_j^- + \log \Gamma_{j+1}^-)^2 \\ &\qquad + \phi_+ \sum_{j=1}^{k-1} \left(-\log \Gamma_j^2 + \log \Gamma_{j+1}^+\right)^2 \quad \text{as } n \to \infty \\ &\to_d \phi_- \sum_{j=1}^{\infty} (-\log \Gamma_j^- + \log \Gamma_{j+1}^-)^2 \\ &\qquad + \phi_+ \sum_{j=1}^{\infty} (-\log \Gamma_j^+ + \log \Gamma_{j+1}^+)^2 \quad \text{as } k \to \infty, \end{aligned}$$

it suffices to show [cf. Billingsley (1968), Theorem 4.2] that for any $\eta > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} P\{T_{n,2}(k) > \eta\} = 0,$$

$$\lim_{n\to\infty} P\{T_{n,3}(k) > \eta\} = 0,$$

$$\lim_{k\to\infty} \limsup_{n\to\infty} P\{T_{n,4}(k) > \eta\} = 0,$$

which follows readily from Lemma 3.2 and an application of the following result. □

LEMMA 3.6. *Suppose F is boundedly supported. Then*

$$\sum_{i=1}^{n-1} (X_{(i+1)} - X_{(i)})^2 \to_p 0$$

if and only if F^{-1}, the left-continuous inverse of F, is continuous.

PROOF. Notice that F^{-1} is continuous if and only if S, the support of F, is a compact interval. The only if part if trivial. To prove the if part, let k be any positive integer and $Q_{k,0}, \dots, Q_{k,k}$ be defined by

$$Q_{k,i} = \frac{k-i}{k} \inf(S) + \frac{i}{k} \sup(S), \qquad 0 \le i \le k.$$

Also let $p_{k,i} = F(Q_{k,i})$. Under these assumptions,

$$X_{([np_{k,i}])} \to_p Q_{k,i} \quad \text{as } n \to \infty, 0 \le i \le k.$$

Thus

$$\sum_{i=1}^{n-1} (X_{(i+1)} - X_{(i)})^2 \le \sum_{j=0}^{k-1} (X_{([np_{k,j+1}])} - X_{([np_{k,j}])})^2$$

$$\to_p \sum_{j=0}^{k-1} (Q_{k,j+1} - Q_{k,j})^2 = \frac{\{\sup(S) - \inf(S)\}^2}{k}.$$

Since the left-hand side is independent of k, the conclusion of the if part follows upon letting $k \to \infty$. □

Finally we remark that in our proofs, exponent 2 in T_n can be replaced by any finite number no less than 1 and all the results will remain correct with minor modification.

APPENDIX

Technical results. We include in the Appendix the technical results as well as some technical details required in Section 2.

LEMMA A.1. *Suppose that $Z_1, \dots, Z_n$ are an i.i.d. sample and r is a positive constant. Let $Z_{(i)}$ be the ith order statistic. Then*

$$n^{-1/r}(|Z_{(n)}| + |Z_{(1)}|) = o_p(1)$$

if and only if $x^r P\{|Z| > x\} \to 0$ as $x \to \infty$.

PROOF. We deal only with the largest order statistic. It follows that for any $v > 0$,

$$P\{Z_{(n)} \le vn^{1/r}\} = P^n\{Z \le vn^{1/r}\},$$

which tends to one if and only if $x^r P\{Z > x\} \to 0$ as $n \to \infty$. On the other hand, it is obvious that $P\{Z_{(n)} > -vn^{1/r}\} \to 1$. This ends the proof. □

LEMMA A.2. *Let H and G be real-valued functions on $(-\infty, \infty)$ such that H is bounded on any compact set and G satisfies* (2.1) *with $r = 1$. Suppose that p and q are constants in $(1, \infty)$ with $1/p + 1/q = 1$ and for which $E|G(Y)|^p < \infty$ and $E|H(Y)|^q < \infty$. Then*

$$\text{(A.1)} \qquad n^{-1} \sum_{i=2}^{n} \left(\left|H(Y_{(i)})\right| + \left|H(Y_{(i-1)})\right|\right)\left|G(Y_{(i)}) - G(Y_{(i-1)})\right| \to_p 0.$$

(A.1) *also holds under* (2.1) *with $r = 1$, $E|G(Y)| < \infty$ and $H \equiv 1$.*

PROOF. For any $0 < \delta < 1/2$, there exists a compact set such that the probability that both $Y_{[n\delta]}$ and $Y_{[n(1-\delta)]}$ belong to the set tends to 1. By this and the assumption that H is bounded on any compact set, the proof is based on the following two facts:

$$\text{(A.2)} \qquad n^{-1} \sum_{[n\delta]}^{[n(1-\delta)]} \left|G(Y_{(i)}) - G(Y_{(i-1)})\right| \to_p 0, \qquad 0 < \delta < \tfrac{1}{2},$$

and (by Hölder's inequality)

$$\begin{aligned}
&\lim_{\delta \to 0} \limsup_{n \to \infty} P\left\{n^{-1} \sum_{1}^{[n\delta]} \left|G(Y_{(i)})\right|^p > c\right\} \\
&\quad + \lim_{\delta \to 0} \limsup_{n \to \infty} P\left\{n^{-1} \sum_{1}^{[n\delta]} \left|H(Y_{(i)})\right|^q > c\right\} \\
\text{(A.3)} \qquad &\quad + \lim_{\delta \to 0} \limsup_{n \to \infty} P\left\{n^{-1} \sum_{[n(1-\delta)]}^{n} \left|G(Y_{(i)})\right|^p > c\right\} \\
&\quad + \lim_{\delta \to 0} \limsup_{n \to \infty} P\left\{n^{-1} \sum_{[n(1-\delta)]}^{n} \left|H(Y_{(i)})\right|^q > c\right\} = 0 \qquad c > 0.
\end{aligned}$$

We now show (A.2). Fix $\delta \in (0, 1/2)$. Let F be the distribution function of Y and F^{-1} the left-continuous inverse of F. Define $A_n = I\{Y_{([n\delta])} > F^{-1}(\beta)\}$ and $B_n = I\{Y_{([n(1-\delta)])} < F^{-1}(1-\beta)\}$ for $0 < \beta < \delta$. For some $\beta > 0$, $E(A_n) \to 1$ and $E(B_n) \to 1$. Thus (A.2) follows from

$$n^{-1} \sum_{[n\delta]}^{[n(1-\delta)]} \left|G(Y_{(i)}) - G(Y_{(i-1)})\right| A_n B_n \to_p 0,$$

which, in turn, follows from (2.1) with $r = 1$. Next we show (A.3). The proofs for the four terms tending to zero are identical. To demonstrate we consider only the first term. Choose $\beta > 0$ such that

$$\text{(A.4)} \qquad E\left|G(Y)\right|^p I\{Y \le F^{-1}(\beta)\} < c.$$

For this β there exists $0 < \delta < 1/2$ such that

$$P\{Y_{([n\delta])} > F^{-1}(\beta)\} \to 0 \quad \text{as } n \to \infty. \tag{A.5}$$

Then

$$P\left\{n^{-1}\sum_{1}^{[n\delta]} \left|G(Y_{(i)})\right|^{p} > c\right\}$$

$$\le P\{Y_{([n\delta])} > F^{-1}(\beta)\} + P\left\{Y_{([n\delta])} \le F^{-1}(\beta), n^{-1}\sum_{1}^{[n\delta]} \left|G(Y_{(i)})\right|^{p} > c\right\}$$

$$\le P\{Y_{([n\delta])} > F^{-1}(\beta)\} + P\left\{n^{-1}\sum_{1}^{n} \left|G(Y_{i})\right|^{r} I\{Y_i \le F^{-1}(\beta)\} > c\right\}.$$

These terms converge in probability to zero by (A.5) and (A.4), respectively. This shows (A.3) and concludes the proof. □

The following result uses slightly different ideas.

Lemma A.3. *Let G be a real-valued function on $(-\infty, \infty)$ and let r be a positive constant. Suppose that for all $B > 0$, (2.1) holds and for some $B_0 > 0$, there exists a nondecreasing function $\tilde{G}$ on (B_0, ∞) such that*

$$\tilde{G}^{r}(y)P\{|Y| > y\} \to 0 \quad \text{as } y \to \infty$$

and

$$|G(x) - G(y)| \le \left|\tilde{G}(|x|) - \tilde{G}(|y|)\right| \quad \text{for } x, y \in (-\infty, -B_0) \text{ or } (B_0, \infty). \tag{A.6}$$

Then for each $s \ge 1$,

$$n^{-s/r}\sum_{i=2}^{n} \left|G(Y_{(i)}) - G(Y_{(i-1)})\right|^{s} \to_p 0.$$

Proof. If Y is boundedly supported, then the result follows easily from the condition (2.1). Suppose that the support of Y is unbounded. It suffices to show that

$$n^{-s/r}\sum_{[n\delta]}^{[n(1-\delta)]} \left|G(Y_{(i)}) - G(Y_{(i-1)})\right|^{s} \to_p 0, \qquad 0 < \delta < \tfrac{1}{2} \tag{A.7}$$

and, for $c > 0$,

$$\lim_{\delta\to 0}\limsup_{n\to\infty} P\left\{n^{-s/r}\sum_{1}^{[n\delta]} \left|G(Y_{(i)}) - G(Y_{(i-1)})\right|^{s} > c\right\} + \lim_{\delta\to 0}\limsup_{n\to\infty} P\left\{n^{-s/r}\sum_{[n(1-\delta)]}^{n} \left|G(Y_{(i)}) - G(Y_{(i-1)})\right|^{s} > c\right\} = 0. \tag{A.8}$$

(A.7) is proved in much the same way as (A.2), hence we only prove (A.8). Choose $\delta > 0$ small enough so that $EA_n \to 1$, where $A_n = I\{Y_{([n\delta])} < -B_0\}$. By (A.6) and the monotonicity of $\tilde{G}$, we have that

$$n^{-s/r} \sum_{i=1}^{[n\delta]} \left|G(Y_{(i)}) - G(Y_{(i-1)})\right|^s A_n \le n^{-s/r}\left|\tilde{G}(Y_{(1)}) - \tilde{G}(Y_{([n\delta])})\right|^s$$

which tends to zero by Lemma A.1. The other tail can be handled the same way. This completes the proof. □

THEOREM A.4. *Let $\{s_n\}$ be a sequence of positive constants, $\{X_{nk}\}$ a triangular array of random variables for $k = 1, \dots, m_n$ and $n = 1, 2, 3, \dots$, and $\mathscr{F}_n$ a sequence of σ-fields. Define $\tilde{X}_{nk} = X_{nk} - E(X_{nk}|\mathscr{F}_n)$. Finally, assume that*

(A.9) $\quad X_{n1}, \dots, X_{n,m_n}$ *are conditionally independent given* $\mathscr{F}_n$;

(A.10) $$s_n^{-2} \sum_1^{m_n} E\left(\tilde{X}_{nk}^2 \middle| \mathscr{F}_n\right) \to_p \sigma^2;$$

(A.11) *for every* $c > 0$, $$s_n^{-2} \sum_1^{m_n} E\left\{\tilde{X}_{nk}^2 I\left(|\tilde{X}_{nk}| > cs_n\right)\middle|\mathscr{F}_n\right\} \to_p 0;$$

(A.12) $$s_n^{-1} \sum_1^{m_n} E(X_{nk}|\mathscr{F}_n) \Rightarrow_{\mathscr{L}} \text{some distribution } G.$$

Then the limiting distribution of $s_n^{-1}\sum_1^{m_n} X_{nk}$ is the convolution of G and $N(0, \sigma^2)$.

REMARK. (A.11) is a conditional version of the classical Lindeberg condition. In this paper, the major difficulty in applying Theorem A.4 is the verification of (A.11); see Theorem 2.3 and Lemma A.5. It is interesting to compare Theorem A.4 with other existing conditional central limit theorems, in particular, central limit theorems in the contexts of martingales and exchangeability. For example, a major difference between this result and a well-known martingale central limit theorem, Theorem 3.2 and Corollary 3.1 of Hall and Heyde (1980), is that there the roles of (A.10) and (A.12) are in some sense reversed, making the limit distribution a mixture instead of a convolution; see Hall and Heyde (1980) and Taylor, Daffer and Patterson (1985).

PROOF OF THEOREM A.4. Let $\tilde{s}_n^2 = \sum_1^{m_n} E(\tilde{X}_{nk}^2|\mathscr{F}_n)/\sigma^2$. By (A.10), it suffices to prove the theorem for $T_n = \tilde{s}_n^{-1}\sum_1^{m_n} X_{nk}$. Write

$$Z_n = \exp\left\{(it/\tilde{s}_n)\sum_1^{m_n} E\left(X_{nk}\middle|\mathscr{F}_n\right)\right\}.$$

The characteristic function of T_n is

$$E\{\exp(itT_n)\} = E\left\{Z_n \exp\left(\sum_1^{m_n} \frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right)\right\} = A_n + B_n,$$

where

$$A_n = E\left(Z_n \exp\left[\sum_1^{m_n} E\left\{\exp\left(\frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right) - 1 \middle| \mathscr{F}_n\right\}\right]\right);$$

$$B_n = EZ_n D_n = EZ_n\left(\prod_1^{m_n} E\left\{\exp\left(\frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right)\middle|\mathscr{F}_n\right\}\right.$$
$$\left. - \exp\left[\sum_1^{m_n} E\left\{\exp\left(\frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right) - 1 \middle| \mathscr{F}_n\right\}\right]\right).$$

Define

$$C_n = \exp\left[\sum_1^{m_n} E\left\{\exp\left(\frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right) - 1 - \frac{it\tilde{X}_{nk}}{\tilde{s}_n} + \frac{t^2\tilde{X}_{nk}^2}{2\tilde{s}_n^2}\middle| \mathscr{F}_n\right\}\right].$$

Since $E(\tilde{X}_{nk}|\mathscr{F}_n) = 0$, we see that

$$A_n = E\left[Z_n C_n \exp\left\{-\frac{t^2}{2\tilde{s}_n^2}\sum_1^{m_n} E\left(\tilde{X}_{nk}^2|\mathscr{F}_n\right)\right\}\right] = E\left(e^{-\sigma^2 t^2/2} Z_n C_n\right),$$

so that

$$\left|A_n - e^{-\sigma^2t^2/2}E(Z_n)\right| \le e^{-\sigma^2t^2/2}E(|Z_n|\,|C_n - 1|) = e^{-\sigma^2t^2/2}E|C_n - 1|.$$

Thus, in view of (A.12), the theorem follows if we prove that $C_n \to_{L_1} 1$ and $D_n \to_{L_1} 0$. Now observe that with probability 1, both C_n and D_n are bounded for each fixed t. C_n is bounded since

$$\left|\sum_1^{m_n} E\left\{\exp\left(\frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right) - 1 - \frac{it\tilde{X}_{nk}}{\tilde{s}_n} + \frac{t^2\tilde{X}_{nk}^2}{2\tilde{s}_n^2}\middle|\mathscr{F}_n\right\}\right| \le \sum_1^{m_n} E\left\{\frac{t^2\tilde{X}_{nk}^2}{\tilde{s}_n^2}\middle|\mathscr{F}_n\right\} = t^2\sigma^2.$$

The second term of D_n can be considered in exactly the same way, showing that $|D_n| \le 1 + \exp(t^2\sigma^2)$. Thus L_1 convergence of C_n and D_n is equivalent to convergence in probability. We therefore first show that $\log(C_n) \to_p 0$. Now, for $c > 0$,

$$\left|\exp\left(\frac{ixt}{\tilde{s}_n}\right) - 1 - \frac{ixt}{\tilde{s}_n} + \frac{x^2t^2}{2\tilde{s}_n^2}\right| \le \begin{cases} \dfrac{cx^2|t|^3}{6\tilde{s}_n^2}, & \text{if } |x| \le c\tilde{s}_n; \\[2ex] \dfrac{x^2t^2}{\tilde{s}_n^2}, & \text{if } |x| > c\tilde{s}_n. \end{cases}$$

Thus

$$|\log(C_n)| \le \frac{1}{6}\sum_1^{m_n} E\left\{ I\big(|\tilde{X}_{nk}| \le c\tilde{s}_n\big) \frac{c\tilde{X}_{nk}^2|t|^3}{\tilde{s}_n^2} \,\middle|\, \mathscr{F}_n \right\}$$

$$+ \sum_1^{m_n} E\left\{ I\big(|\tilde{X}_{nk}| > c\tilde{s}_n\big) \frac{\tilde{X}_{nk}^2 t^2}{\tilde{s}_n^2} \,\middle|\, \mathscr{F}_n \right\}$$

$$\le \frac{c\sigma^2|t|^3}{6} + \left(\frac{t^2}{\tilde{s}_n^2}\right) \sum_1^{m_n} E\Big\{\tilde{X}_{nk}^2 I\big(|\tilde{X}_{nk}| > c\tilde{s}_n\big)\Big|\mathscr{F}_n\Big\}.$$

By (A.10) and (A.11), the second term converges to zero in probability. Letting $c \to 0$ shows that $\log(C_n) \to_p 0$. Next we show that $D_n \to_p 0$. For each $\varepsilon > 0$, there exists $0 < \delta < 1$ such that if $|x| \le \delta$, then $|\exp(x) - 1 - x| \le \varepsilon|x|$. Fix such a (ε, δ) pair. Let $G_n = \max_{1 \le k \le m_n} |E\{\exp(it\tilde{X}_{nk}/\tilde{s}_n) - 1|\mathscr{F}_n\}|$. Using the inequality $|\prod_1^m x_i - \prod_1^m y_i| \le \sum_1^m |x_i - y_i|$ for $|x_i|, |y_i| \le 1$ and the fact that $E(\tilde{X}_{nk}|\mathscr{F}_n) = 0$, we have that

$$|D_n| I(G_n \le \delta)$$

$$\le \sum_1^{m_n} \left| \exp\left[E\left\{ \exp\left(\frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right) - 1 \,\middle|\, \mathscr{F}_n \right\}\right] - E\left\{ \exp\left(\frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right) \,\middle|\, \mathscr{F}_n \right\} \right| I(G_n \le \delta)$$

$$= \sum_1^{m_n} \left| \exp\left[E\left\{ \exp\left(\frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right) - 1 \,\middle|\, \mathscr{F}_n \right\}\right] - 1 \right.$$

$$\left. - E\left\{ \exp\left(\frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right) - 1 \,\middle|\, \mathscr{F}_n \right\} \right| I(G_n \le \delta)$$

$$\le \varepsilon \sum_1^{m_n} \left| E\left\{ \exp\left(\frac{it\tilde{X}_{nk}}{\tilde{s}_n}\right) - 1 \,\middle|\, \mathscr{F}_n \right\} \right| \le \frac{\varepsilon t^2}{2\tilde{s}_n^2} \sum_1^{m_n} E\big(\tilde{X}_{nk}^2 \big| \mathscr{F}_n\big) = \frac{\varepsilon\sigma^2 t^2}{2}.$$

Again using the fact that $|D_n| \le 1 + e^{t^2\sigma^2}$, this means that $|D_n| \le \varepsilon\sigma^2 t^2/2 + (1 + e^{t^2\sigma^2}) I(G_n > \delta)$. Since ε was arbitrary, it suffices to show that $G_n \to_p 0$. But for any fixed $c > 0$,

$$G_n \le \frac{2|t|}{\tilde{s}_n} \max_{1 \le k \le m_n} E\big(|\tilde{X}_{nk}| \,\big|\mathscr{F}_n\big)$$

$$\le \frac{2s_n c|t|}{\tilde{s}_n} + \frac{2|t|}{\tilde{s}_n} \max_{1 \le k \le m_n} E\Big(|\tilde{X}_{nk}| I\big(|\tilde{X}_{nk}| > cs_n\big)\Big|\mathscr{F}_n\Big)$$

$$\le \frac{2s_n c|t|}{\tilde{s}_n} + \frac{2|t|}{cs_n\tilde{s}_n} \sum_1^{m_n} E\Big(\tilde{X}_{nk}^2 I\big(|\tilde{X}_{nk}| > cs_n\big)\Big|\mathscr{F}_n\Big).$$

By (A.10) and (A.11), the second term converges in probability to zero for any fixed c; letting c then go to zero concludes that $G_n \to_p 0$ and hence $D_n \to_p 0$. This completes the proof. □

LEMMA A.5. *Assume that the conditions of Theorem 2.3 hold and let* $s_n, m_n, \mathscr{F}_n, X_{ni}$ *be as defined in the proof of Theorem 2.3. Then* (A.11) *holds.*

PROOF. Note that

$$\tilde{X}_{ni} = U_{i1} + U_{i2} - U_{i3},$$

where

$$\begin{aligned} U_{i1} &= \lambda^T \operatorname{vech}\{\varepsilon_{(2i)*}\varepsilon_{(2i)*}^T - V(Y_{(2i)})\}; \\ U_{i2} &= \lambda^T \operatorname{vech}\{\varepsilon_{(2i-1)*}\varepsilon_{(2i-1)*}^T - V(Y_{(2i-1)})\}; \\ U_{i3} &= \lambda^T \operatorname{vech}\{\varepsilon_{(2i)*}\varepsilon_{(2i-1)*}^T + \varepsilon_{(2i-1)*}\varepsilon_{(2i)*}^T\} \\ &= \lambda^T \Psi(I + \Gamma)\operatorname{vec}\{\varepsilon_{(2i)*}\varepsilon_{(2i-1)*}^T\}. \end{aligned}$$

To prove (A.11), it suffices to show that for $1 \le j, k \le 3$, and for any $\eta > 0$,

$$K_{jk} = n^{-1} \sum_{1}^{[n/2]} E\{U_{ij}^2 I(|U_{ik}| > \eta n^{1/2}) | \mathscr{F}_n\} \to_p 0.$$

For $j = k = 1$ or $j = k = 2$, we have

$$\begin{aligned} K_{jk} \le 2\|\lambda\|^2 n^{-1} \sum_{i=1}^{n} E\Big[\big(|\varepsilon_i|^4 + |V(Y_i)|^2\big) \\ \times\big\{I\big(|\varepsilon_i|^2 > \eta n^{1/2}/2\big) + I\big(|V(Y_i)| > \eta n^{1/2}/2\big)\big\}\Big|Y_i\Big], \end{aligned}$$

which converges in probability to zero by dominated convergence since $E|\varepsilon|^4 < \infty$ and $E|V(Y)|^2 < \infty$.

If $j = 1$, $k = 2$ (and similarly for $j = 2$, $k = 1$), since U_{i1} and U_{i2} are independent given $\mathscr{F}_n$, we get

$$\begin{aligned} K_{12} &= n^{-1} \sum_{1}^{[n/2]} E(U_{i1}^2 | \mathscr{F}_n) E\{I(|U_{i2}| > \eta n^{1/2}) | \mathscr{F}_n\} \\ &= n^{-1} \sum_{1}^{[n/2]} \lambda^T R(Y_{(2i)})\lambda E\{I(|U_{i2}| > \eta n^{1/2}) | \mathscr{F}_n\}. \end{aligned}$$

Again applying Lemma A.2, we have $K_{12} = K_{12*} + o_p(1)$, where

$$K_{12*} = n^{-1} \sum_{1}^{[n/2]} \lambda^T R(Y_{(2i-1)})\lambda E\{I(|U_{i2}| > \eta n^{1/2}) | \mathscr{F}_n\};$$

$$|K_{12*}| \le n^{-1} \sum_{1}^{n} |\lambda^T R(Y_i)\lambda| E\{I(|\lambda \operatorname{vech}(\varepsilon\varepsilon^T - V(Y_i))| > \eta n^{1/2})\},$$

which converges in probability to zero since $E|R(Y)| < \infty$ and by an application of dominated convergence.

For $j = 3$, $k = 1$ (and similarly if $j = 3$, $k = 2$), first recall that $U_{i3} = \sum_l \sum_m c^{(lm)} \varepsilon^{(l)}_{(2i)*} \varepsilon^{(m)}_{(2i-1)*}$. Since $\varepsilon_{(2i)*}$ and $\varepsilon_{(2i-1)*}$ are independent given $\mathscr{F}_n$, we have

$$\begin{aligned} K_{31} = n^{-1} \sum_{1}^{[n/2]} \sum_l \sum_m \sum_r \sum_s c^{(lm)} c^{(rs)} V^{(ms)}\big(Y_{(2i-1)}\big) \\ \times E\Big\{\varepsilon^{(l)}_{(2i)*} \varepsilon^{(r)}_{(2i)*} I\big(|U_{i1}| > \eta n^{1/2}\big) \Big| \mathscr{F}_n \Big\}. \end{aligned} \tag{A.13}$$

By an application of Lemma A.2, in (A.13) we may replace $Y_{(2i-1)}$ by $Y_{(2i)}$, in which case (A.13) is bounded by

$$\begin{aligned} n^{-1} \sum_{1}^{n} \sum_l \sum_m \sum_r \sum_s |c^{(lm)} c^{(rs)}| E\Big\{ \big| V^{(ms)}(Y_i) \varepsilon_i^{(l)} \varepsilon_i^{(r)} \big| \\ \times I\Big(\big| \lambda^T \{ \varepsilon_i \varepsilon_i^T - v(Y_i) \} \big| > \eta n^{1/2} \Big) \Big| Y_i \Big\}, \end{aligned}$$

which is easily seen to converge in probability to zero.

For $j = 1$, $k = 3$ (and similarly for $j = 2$, $k = 3$), it suffices to show that for any $\eta > 0$,

$$\begin{aligned} n^{-1} \sum_{1}^{[n/2]} E\Big\{ \big[\lambda^T \operatorname{vech}\big\{ \varepsilon_{(2i)*} \varepsilon^T_{(2i)*} - V\big(Y_{(2i)}\big) \big\} \big]^2 \\ \times I\Big(\big| \lambda^T \operatorname{vech}\big(\varepsilon_{(2i)*} \varepsilon_{(2i-1)*} \big) \big| > \eta n^{1/2} \Big) \Big| \mathscr{F}_n \Big\} \to_p 0. \end{aligned}$$

Using the fact that $I(|X_1 X_2| > c) \le I(|X_1| > c^{1/2}) + I(|X_2| > c^{1/2})$, by considering individual elements it suffices to show that $T_{n1} \to_p 0$, $T_{n2} \to_p 0$, where for any (l, m, r),

$$T_{n1} = n^{-1} \sum_{1}^{[n/2]} E\Big\{ \big\{ \varepsilon^{(l)}_{(2i)*} \varepsilon^{(m)}_{(2i)*} - V^{(lm)}\big(Y_{(2i)}\big) \big\}^2 I\Big(|\varepsilon^{(r)}_{(2i)*}| > \eta^{1/2} n^{1/4} \Big) \Big| \mathscr{F}_n \Big\}.$$

$$T_{n2} = n^{-1} \sum_{1}^{[n/2]} E\Big\{ \big\{ \varepsilon^{(l)}_{(2i)*} \varepsilon^{(m)}_{(2i)*} - V^{(lm)}\big(Y_{(2i)}\big) \big\}^2 I\Big(|\varepsilon^{(r)}_{(2i-1)*}| > \eta^{1/2} n^{1/4} \Big) \Big| \mathscr{F}_n \Big\}.$$

Since

$$T_{n1} \le n^{-1} \sum_{1}^{n} E\Big\{ \big\{ \varepsilon_i^{(l)} \varepsilon_i^{(m)} - V^{(lm)}(Y_i) \big\}^2 I\Big(|\varepsilon_i^{(r)}| > \eta^{1/2} n^{1/4} \Big) \Big| Y_i \Big\},$$

we clearly have $T_{n1} \to_p 0$ by the conditions $E|\varepsilon|^4 < \infty$, $E|V(Y)|^2 < \infty$ and dominated convergence. Again, applying Lemma A.2 we have

$$\begin{aligned} T_{n2} = n^{-1} \sum_{1}^{[n/2]} E\Big\{ \big\{ \varepsilon^{(l)}_{(2i-1)*} \varepsilon^{(m)}_{(2i-1)*} - V^{(lm)}\big(Y_{(2i-1)}\big) \big\}^2 \\ \times I\Big(|\varepsilon^{(r)}_{(2i-1)*}| > \eta^{1/2} n^{1/4} \Big) \Big| \mathscr{F}_n \Big\} + o_p(1), \end{aligned}$$

where the first term has expectation converging to zero by dominated convergence since $E|R(Y)| < \infty$.

Finally, for $j = k = 3$, it suffices to show that for any (l, m, r, s, u)

$$n^{-1} \sum_{1}^{[n/2]} E\Big\{\varepsilon_{(2i)*}^{(l)} \varepsilon_{(2i)*}^{(m)} \varepsilon_{(2i-1)*}^{(r)} \varepsilon_{(2i-1)*}^{(s)} I\Big(|\varepsilon_{(2i)*}^{(u)}| > \eta^{1/2} n^{1/4}\Big)\Big|\mathscr{F}_n\Big\} \to_p 0.$$

This is a consequence of the Schwarz inequality and dominated convergence since $E|\varepsilon|^4 < \infty$. □

Acknowledgments. The authors wish to thank an Associate Editor and the referees for their valuable comments.

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variations*. Cambridge Univ. Press.

DAVID, H. A. (1981). *Order Statistics*. Wiley, New York.

DUAN, N. and LI, K. C. (1991). Slicing regression: A link-free regression method. *Ann. Statist.* **19** 505–530.

EUBANK, R. L. (1988). *Spline Smoothing and Nonparametric Regression*. Dekker, New York.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. Wiley, New York.

FRIEDMAN, J. and STUETZLE, W. (1981). Projection pursuit regression. *J. Amer. Statist. Assoc.* **76** 817–823.

GREENWOOD, M. (1946). The statistical study of infectious deseases. *J. Roy. Statist. Soc. Ser. A* **109** 85–109.

HALL, P. (1989). On projection pursuit regression. *Ann. Statist.* **17** 573–588.

HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Applications*. Academic, New York.

HÄRDLE, W. and STOKER, T. (1989). Investigating smooth multiple regression by the method of average derivatives. *J. Amer. Statist. Assoc.* **84** 986–995.

HASTIE, T. and TIBSHIRANI, R. (1986). Generalized additive models. *Statist. Sci.* **1** 297–318.

LI, K. C. (1989). Sightseeing with SIR: A transformation based projection pursuit method. Technical report, Dept. Mathematics, UCLA.

LI, K. C. (1991). Sliced inverse regression for dimension reduction (with discussion). *J. Amer. Statist. Assoc.* **86** 316–327.

MORAN, P. A. P. (1947). The random division of an interval. *J. Roy. Statist. Soc. Ser. B* **9** 92–98.

PYKE, R. (1965). Spacings. *J. Roy. Statist. Soc. Ser. B* **27** 395–449.

READ, C. B. (1983). Greenwood's statistic. In *Encyclopedia of Statistical Sciences* **3** 522–523. Wiley, New York.

READ, C. B. (1988). Spacings. In *Encyclopedia of Statistical Sciences* **8** 566–569. Wiley, New York.

RESNICK, S. (1987). *Extreme Values, Regular Variation, and Point Process*. Springer, New York.

STEPHENS, M. A. (1981). Further percentage points for Greenwood's statistic. *J. Roy. Statist. Soc. Ser. A* **144** 364–366.

TAYLOR, R. L., DAFFER, P. Z. and PATTERSON, R. F. (1985). *Limit Theorems For Sums of Exchangeable Random Variables*. Rowman and Allanheld, Totowa.

YANG, S. S. (1977). General distribution theory of the concomitants of order statistics. *Ann. Statist.* **5** 996–1002.

DEPARTMENT OF STATISTICS
TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77843-3143