The Annals of Probability
1984, Vol. 12, No. 1, 154–172

ON THE INFLUENCE OF EXTREMES ON THE RATE OF CONVERGENCE IN THE CENTRAL LIMIT THEOREM

BY PETER HALL

Australian National University

Let $\bar{X}$ be the mean of a random sample from a distribution which is symmetric about its unknown mean μ and has known variance σ^2. The classical method of constructing a hypothesis test or confidence interval for μ is to use the normal approximation to $n^{1/2}(\bar{X} - \mu)/\sigma$. In order to make this procedure more robust, we might lightly trim the mean by removing extremes from the sample. It is shown that this procedure can greatly improve the rate of convergence in the central limit theorem, but only if the new mean is rescaled in a rather complicated way. From a practical point of view, the removal of extreme values does *not* make the test or confidence interval more robust.

1. Introduction Let $X_1, X_2, \cdots, X_n$ be independent observations from a symmetric distribution with unknown centre μ and known variance σ^2. If we wish to test a hypothesis about μ, or construct a confidence interval for μ, then classical statistical theory suggests that we base our procedure on the normal approximation to the sample mean, $\bar{X} = n^{-1} \sum_1^n X_j$. However, to allow for the possibility that the underlying distribution has rather large tails, we might modify the sample mean to make it more robust against departures from a normal model. One very simple modification is to delete k extremes from the sample, and replace $\bar{X}$ by the mean of the remaining $n - k$ observations. Provided $k/n \to 0$ as $n \to \infty$, the new mean, $\bar{X}(k)$ say, will have the same asymptotic variance as $\bar{X}$.

Huber (1981, page 5) has described a robust procedure as one whose performance is "close to the nominal value calculated at the model". Thus, if the significance level of a robust test is equal to α under a normal model, the level should be close to α in the case of a longer-tailed distribution. Suppose our test can be made more robust by deleting extremes from the sample. Then the removal of outliers must improve the rate of convergence to the normal limit.

Our aim in the present paper is to investigate the effects of extremes on rates of convergence. Hatori, Maejima and Mori (1979) considered a related problem in the case of the law of large numbers. They showed there that the rate of convergence can be improved by deleting a number of extremes, even if the number is fixed as $n \to \infty$. See also Mori (1976, 1977).

Our main conclusions may be summarized by two points.

(i) From a theoretical viewpoint, the removal of extremes can substantially improve the rate of convergence in the central limit theorem. Provided the trimmed mean $\bar{X}(k)$ is scaled in the right way, the rate of convergence can be made arbitrarily close to $O(n^{-1})$ for a suitably chosen sequence $k(n)$ satisfying

Received October 1982; revised May 1983.

AMS 1980 *subject classifications.* Primary, 60F05, 62G30; secondary, 60G50.

Key words and phrases. Central limit theorem, extremes, order statistics, rate of convergence, sums of independent random variables.

$k(n)/n \to 0$ as $n \to \infty$, even if the underlying distribution does not have any finite moments higher than the second.

(ii) The improvements described in part (i) *cannot* be achieved using a simple scale factor. Indeed, if $\mu = 0$, if

$$\Delta_n^{(1)} = \sup_{-\infty<z<\infty} | P\{(n-k)^{1/2}\bar{X}(k) \le \sigma z\} - \Phi(z) |$$

denotes the rate of convergence after k extremes have been removed from the sample, and if

$$\Delta_n^{(2)} = \sup_{-\infty<z<\infty} | P(n^{1/2}\bar{X} \le \sigma z) - \Phi(z) |$$

denotes the rate based on the entire sample, then the ratio

$$(\Delta_n^{(1)} + n^{-1})/(\Delta_n^{(2)} + n^{-1})$$

does not converge to zero as $n \to \infty$. Therefore the rate of convergence of

$$(n-k)^{1/2}\bar{X}(k)/\sigma$$

to the standard normal law is at least as slow as that of $n^{1/2}\bar{X}/\sigma$, up to terms of order n^{-1}; see Theorem 4. (In the definitions of $\Delta_n^{(1)}$ and $\Delta_n^{(2)}$, Φ denotes the standard normal distribution function.)

The scale factor needed to achieve the improved rate of convergence described in (i) may be chosen as a function of the last-removed extreme. Sometimes it may be taken equal to the expectation of this function; see Proposition 3. However, the optimal scale factor is of a rather involved nature, and is most unlikely to be known in practice. Since the simple, commonly used scale factor of $(n-k)^{1/2}\sigma^{-1}$ does not improve the rate of convergence, the overall conclusion of our study is a rather pessimistic one. There appears to be no practical way of improving the rate of convergence of $n^{1/2}\bar{X}/\sigma$ to normality, and so of constructing a robust test or confidence interval, simply by trimming the mean. It may be possible to improve the rate of convergence by studentizing $\bar{X}(k)$ in a suitable manner, but there are considerable technical problems in describing the fine asymptotic behaviour of such a scheme.

We discuss now our definition of an "extreme". In practice, an extreme is often determined according to a mixture of several criteria: its apparent distance from the "centre" of the distribution, a test for extremality, a plot of the data, and perhaps some knowledge of the circumstances under which the data were recorded. The following model seems to be a reasonable approximation to the rather ad-hoc procedure which is carried out in reality. We assume that the underlying distribution is symmetric about $\mu = 0$, and delete those k observations which are *largest in absolute value*. Denote the sum of the remaining $n - k$ observations by ${}^{(k)}S_n$. Then the trimmed mean is given by $\bar{X}(k) = (n-k)^{-1}\,{}^{(k)}S_n$. We do not assume that the outliers come from a contaminant distribution, but rather that they represent extremes from the underlying distribution.

We conclude this section by describing the notation used in the remainder of the paper. Let $X, X_1, X_2, \cdots$ be independent and identically distributed observations with zero mean, unit variance and distribution function F, and let $X_{n1}, X_{n2}, \cdots, X_{nn}$ denote the sample values $X_1, X_2, \cdots, X_n$ arranged in *decreasing* order of magnitude: $|X_{n1}| \geq |X_{n2}| \geq \cdots \geq |X_{nn}|$. Then ${}^{(k)}S_n = S_n - \sum_{j=1}^{k} X_{nj}$, where $S_n = \sum_{j=1}^{n} X_j$. Set $G(x) = P(|X| \leq x)$, let $k = k(n)$ denote a sequence of positive integers satisfying $k/n \to 0$ as $n \to \infty$, and let G_{nk} denote the distribution function of $|X_{nk}|$. To avoid trivialities we assume that $G(x) < 1$ for all x.

2. Rates of convergence. We begin with a limit theorem for the case of a general summand distribution. Define

$$\mu(x) = \{G(x)\}^{-1} \int_{|u| \leq x} u \, dF(u)$$

whenever $G(x) > 0$, and set $\mu(x) = 0$ if $G(x) = 0$. Let Y have the distribution of X given that $|X| \leq x$, and define

$$\mu_n(x) = E\{(Y - \mu) I(|Y - \mu| \leq n^{1/2})\}$$

and

$$\sigma_n^2(x) = E\{(Y - \mu)^2 I(|Y - \mu| \leq n^{1/2})\},$$

where $\mu = \mu(x)$ and $I(E)$ denotes the indicator function of an event E. Both $\mu_n(x)$ and $\sigma_n(x)$ are defined to be zero if $G(x) = 0$. Let $\mathscr{F}_{nk}$ denote the σ-field generated by $|X_{nk}|$, and define the random variables A_n and B_n by $A_n = m^{1/2}\sigma_n(|X_{nk}|)$ and $B_n = m\{\mu(|X_{nk}|) + \mu_n(|X_{nk}|)\}$, where $m = n - k$. An extension of the argument used to prove Theorem 3 below may be used to show that in a certain sense, the norming functions A_n and B_n are optimal in the class of $\mathscr{F}_{nk}$-measurable functions. The rate of convergence using these functions is described by

$$\Delta_n = \sup_{-\infty < z < \infty} |P({}^{(k)}S_n \leq A_n z + B_n) - \Phi(z)|,$$

and is closely approximated by the more tractable sequence,

$$\delta_n = n \int_0^\infty E\left[\min\left\{1, \left(\frac{Y - \mu}{n^{1/2}}\right)^4\right\}\right] dG_{nk}(x)$$
$$+ n^{-1/2} \left| \int_0^\infty \{\sigma_n(x)\}^{-3} \, dG_{nk}(x) \int_{|u| \leq n^{1/2}} u^2 \, dP(Y - \mu \leq u) \right|.$$

THEOREM 1. *Suppose $E(X) = 0$ and $E(X^2) = 1$, and that $k/n \to 0$ as $n \to \infty$. Then the ratio $(\Delta_n + n^{-1/2})/(\delta_n + n^{-1/2})$ is bounded away from zero and infinity as $n \to \infty$.*

The interpretation of this result is greatly simplified if we suppose that the underlying distribution, F, is symmetric. This assumption entails $\mu(x) = \mu_n(x) \equiv 0$, $B_n \equiv 0$,

$$\sigma_n^2(x) = \{G(x)\}^{-1} \int_{u \leq \min(x, n)^{1/2}} u^2 \, dG(u)$$

and

$$\{1 + o(1)\}\delta_n = n^{-1} \int_{0 \le u \le n^{1/2}} u^4 P(|X_{nk}| \ge u)\, dG(u) \tag{2.1}$$
$$+ n \int_{u > n^{1/2}} P(|X_{nk}| \ge u)\, dG(u) + O(n^{-\lambda})$$

for all $\lambda > 0$. (See below for a proof.) The hypothesis of symmetry is imposed throughout almost all statistical work on trimming and robust estimation; see for example Andrews *et al* (1972) and Huber (1981). It ensures that the first order term in a Chebyshev-Edgeworth-Cramér expansion of the distribution of ${}^{(k)}S_n$ vanishes. Thus, if the summand distribution is sufficiently smooth and has sufficiently many finite moments, the order of approximation in the central limit theorem is $O(n^{-1})$ rather than $O(n^{-1/2})$.

To prove (2.1), note that in the case of symmetry and for each $\lambda > 0$,

$$\delta_n = n \int_0^\infty \{G(x)\}^{-1}\, dG_{nk}(x) \int_0^x \min\left\{1, \left(\frac{u}{n^{1/2}}\right)^4\right\} dG(u)$$
$$= \{1 + o(1)\}J + O(n^{-\lambda}),$$

where

$$J = n \int_0^\infty dG_{nk}(x) \int_0^x \min\left\{1, \left(\frac{u}{n^{1/2}}\right)^4\right\} dG(u) = J_1 + J_2,$$

$$J_1 = n^{-1} \int_0^{n^{1/2}} dG_{nk}(x) \int_0^x u^4\, dG(u) + n^{-1} \int_{n^{1/2}}^\infty dG_{nk}(x) \int_0^{n^{1/2}} u^4\, dG(u)$$
$$= n^{-1} \int_0^{n^{1/2}} u^4 P(|X_{nk}| \ge u)\, dG(u),$$

$$J_2 = n \int_{n^{1/2}}^\infty dG_{nk}(x) \int_{n^{1/2}}^x dG(u) = n \int_{n^{1/2}}^\infty P(|X_{nk}| \ge u)\, dG(u).$$

As our smoothness constraint we shall impose Cramér's continuity condition,

$$\limsup_{t\to\infty} |E(e^{itX})| < 1. \tag{C}$$

Our next theorem permits us to improve the boundary of the approximations in Theorem 1 from $n^{-1/2}$ to n^{-1}.

THOEREM 2. *Suppose $E(X) = 0$, $E(X^2) = 1$, $k/n \to 0$ as $n \to \infty$, and the distribution of X satisfies condition* (C). *Then the ratio $(\Delta_n + n^{-1})/(\delta_n + n^{-1})$ is bounded away from zero and infinity as $n \to \infty$.*

There is obviously a very wide latitude of choice for norming constants which depend on $|X_{nk}|$. Our next result shows that the choice of A_n is optimal in a certain sense. Here and in the remainder of this section we assume the underlying distribution is symmetric, which makes the estimates considerably easier to interpret.

THEOREM 3. *Assume the conditions of Theorem 2, and that the underlying distribution is symmetric. Let D_n, $n \geq 1$, be random variables such that D_n is $\mathscr{F}_{nk}$-measurable, $D_n > 0$ almost surely and*

$$P(D_n \geq n^{1/2}\varepsilon \mid |X_{nk}| > x_0) = 1$$

for some $\varepsilon > 0$, $x_0 > 0$ and all large n. Then

$$\liminf_{n\to\infty}\{\sup_{-\infty<z<\infty}|P({}^{(k)}S_n \leq D_n z) - \Phi(z)| + n^{-1}\}/(\delta_n + n^{-1}) > 0.$$

Between them, Theorems 2 and 3 provide a benchmark for assessing rates of convergence using either constant or random norming sequences. We shall interpret both theorems by referring to a special "test distribution" G_α, given by $G_\alpha(x) = 1 - x^{-\alpha}$ for $x \geq 1$ and $2 < \alpha < 4$. If $G \equiv G_\alpha$ then the underlying distribution F has variance $\alpha/(\alpha - 2)$, and is easily standardised.

PROPOSITION 1. *Suppose $G(x) = G_\alpha(x)$ for large x. Then*

$$\delta_n = \{1 + o(1)\}c(k)(4/\alpha - 1)^{-1}n^{-1}(k/n)^{1-4/\alpha}$$

as $n \to \infty$ and $k/n \to 0$, where $c(k) \equiv k^{4/\alpha-1}\Gamma(k - 4/\alpha + 1)/\Gamma(k)$ denotes a sequence of positive constants depending only on k and satisfying $c(k) \to 1$ as $k \to \infty$. In particular,

$$\delta_n \sim \text{const.}\ n^{4/\alpha-2} \tag{2.2}$$

if k is fixed.

Results similar to Proposition 1 may be obtained for distributions G satisfying $G(x) = 1 - x^{-\alpha}L(x)$, where L is a slowly varying function. Two consequences of Proposition 1 should be noted.

(i) Suppose X is symmetric and $P(|X| > x) = G_\alpha(x)$ for large x. Then the fastest rate of convergence in the central limit theorem for S_n, using optimal norming constants, is given by

$$\inf_{c>0,d}\{\sup_{-\infty<z<\infty}|P(S_n \leq cz + d) - \Phi(z)|\} \geq \text{const.}\ n^{1-\alpha/2}$$

for large n, whenever $2 < \alpha < 4$; see for example Theorem 1 of Hall (1980). Since $4/\alpha - 2 < 1 - \alpha/2$ whenever $2 < \alpha < 4$, then even by taking k equal to 1 we may improve the rate of convergence.

(ii) The rate of convergence for fixed $k \geq 1$ does not depend on the value of k. However, the constant in (2.2) is a strictly decreasing function of k. If we desire Δ_n to converge at a rate of $O(n^{-\gamma})$, where $2 - 4/\alpha < \gamma < 1$, we must choose k to be at least of order $n^{1-(1-\gamma)/(4/\alpha-1)}$. If $\alpha < 4$, there does not exist a sequence $k(n)$ diverging to infinity more slowly than n, and such that $\delta_n = O(n^{-1})$.

The test distribution G_α provides an informative guide to the behaviour of other distributions when extremes are removed from the sample. However, it excludes the interesting case where the variance is only just finite—i.e. where $P(|X| > x) = x^{-2}L(x)$, for a slowly varying function L. In this situation, the fastest rate of convergence in the central limit theorem for S_n is achieved with nonstandard norming constants; see Theorems 1 and 2 of Hall (1980). Our next result shows that when the variance is only just finite, choosing k larger than 1 (but still bounded) can improve on the rate of convergence. We shall confine our attention to the distribution given by $H_\beta(x) = 1 - x^{-2}(\log x)^{-\beta}$ for large x, where $\beta > 1$. Similar results may be obtained when $1 - G(x) = x^{-2}L(x)$ for a slowly varying function L, but they involve intricate conditions on the function L.

PROPOSITION 2. *Suppose $G(x) = H_\beta(x)$ for large x. Then*

$$2^{-2\beta}\delta_n \sim \begin{cases} \beta(\log n)^{-2\beta}\log\log n & \text{if } k = 1 \\ (k-1)^{-1}(\log n - \log k)^{-2\beta} & \text{if } k \geq 2 \end{cases}$$

as $n \to \infty$ and $k/n \to 0$.

We shall consider two consequences of Proposition 2.

(i) Suppose X is symmetric and $P(|X| > x) = H_\beta(x)$ for large x. Then (using Theorem 1 of Hall (1980)),

$$\inf_{c>0,d}\{\sup_{-\infty<z<\infty} |P(S_n \leq cz + d) - \Phi(z)|\} \geq \text{const. } (\log n)^{-\beta}$$

for large n, whenever $\beta > 1$. Therefore even by choosing k as small as 1 we may improve on the rate of convergence.

(ii) By taking $k \geq 2$ we achieve a significant improvement on the rate when $k = 1$. However, for fixed $k \geq 2$ the rate does not depend on k. If we wish Δ_n to converge to zero at a rate of $O(n^{-\gamma})$, where $0 < \gamma < 1$, we must choose k to be at least of order $n^\gamma(\log n)^{-2\beta}$.

The fast rates of convergence described in Theorem 2 can sometimes be achieved using a constant scale factor, rather than the random scale factor A_n. We illustrate this property by referring to the test distribution G_α. Set $a_n = \{E(A_n^2)\}^{1/2}$ and let

$$\Delta_n^{(0)} = \sup_{-\infty<z<\infty} |P({}^{(k)}S_n \leq a_n z) - \Phi(z)|.$$

PROPOSITION 3. *Suppose X is symmetric about $\mu = 0$, and $P(|X| \leq x) = G_\alpha(x)$, where $2 < \alpha < 4$. Then the ratio $(\Delta_n^{(0)} + n^{-1})/(\delta_n + n^{-1})$ is bounded away from zero and infinity as $n \to \infty$.*

Next we treat the case of a simple, known norming sequence. Let

$$\Delta_n^{(1)} = \sup_{-\infty<z<\infty} |P({}^{(k)}S_n/(n-k)^{1/2} \leq z) - \Phi(z)|$$

and

$$\Delta_n^{(2)} = \sup_{-\infty<z<\infty} |P(S_n/n^{1/2} \leq z) - \Phi(z)|.$$

The quantity $\Delta_n^{(1)}$ corresponds to a central limit approximation after extremes have been deleted from the sample, while $\Delta_n^{(2)}$ is a central limit approximation for the entire sample. Our next theorem shows that $\Delta_n^{(1)}$ converges to zero no more quickly than $\Delta_n^{(2)}$, up to terms of order n^{-1}. Therefore the operation of deleting extremes from the sample has no appreciable effect on the rate of convergence.

THEOREM 4. *Assume the conditions of Theorem 2. Then the ratio*

$$(\Delta_{n1} + n^{-1})/(\Delta_{n2} + n^{-1})$$

is bounded away from zero as $n \to \infty$.

Theorem 4 suggests that a hypothesis test or confidence interval based on the sample mean and population variance, cannot be made more robust by trimming out the extremes. This is despite the fact that trimming can greatly improve the rate of convergence in the central limit theorem, when the trimmed mean is scaled optimally.

3. Proofs. In the proofs we define $m = n - k$, and let the symbols $C, C_1, C_2, \cdots$ denote positive constants.

PROOF OF THEOREM 1. Let $Y, Y_1, Y_2, \cdots$ be independent and identically distributed random variables with the distribution of X truncated at $\pm x$, where $x > 0$ is fixed. Thus, $P(Y \leq y) = \{F(y) - F(x-)\}/G(x)$ for $|y| \leq x$. The characteristic function of $Y - \mu(x)$ is given by $\alpha(t \mid x) = E[\exp\{it(Y - \mu(x))\}]$. Now,

$$\begin{aligned} &|1 - \alpha(t \mid x)| \\ &= \left| \int_{|u| \leq x} [\exp\{it\mu(x)\} - e^{itu}]\, dF(u) \right| \Big/ G(x) \\ &\leq \left| \int_{|u| \leq x} (1 + itu - e^{itu})\, dF(u) \right| \Big/ G(x) + |\exp\{it\mu(x)\} - 1 - it\mu(x)|/G(x) \\ &\leq C_1 t^2 \exp\{|t\mu(x)|\} \end{aligned}$$

for all $x \geq x_1$ (say) and all t, where C_1 depends on neither x nor t. (It is assumed $G(x)$ is bounded away from zero for $x \geq x_1$.) The function μ is bounded uniformly in x, and so there exists $C_2 > 0$ and $\varepsilon_1 \in (0, 1]$ such that

$$|1 - \alpha(t \mid x)|^2 \leq C_2 t^4 \leq t^2/100 \tag{3.1}$$

and all $x \geq x_1$ and $|t| \leq \varepsilon_1$. In this case,

$$\log\{\alpha(t \mid x)\} = \alpha(t \mid x) - 1 + r_1(t \mid x),$$

where $|r_1(t \mid x) \leq |\alpha(t \mid x) - 1|^2$. Therefore

$$\{\alpha(t/m^{1/2} \mid x)\}^m = \exp[m\{\alpha(t/m^{1/2} \mid x) - 1\}]\{1 + r_{2n}(t \mid x)\}, \tag{3.2}$$

where by (3.1),

$$
\begin{aligned}
|r_{2n}(t\,|\,x)| &\le m\,|r_1(t/m^{1/2}\,|\,x)|\exp\{m\,|r_1(t/m^{1/2}\,|\,x)|\} \\
&\le C_2 m^{-1}t^4\exp(t^2/100)
\end{aligned} \tag{3.3}
$$

uniformly in $x \ge x_1$ and $|t| \le \varepsilon_1 m^{1/2}$.

Let $r_{3n}(t\,|\,x) = |\alpha(t\,|\,x) - 1 - it\mu_n(x) + \tfrac{1}{2}t^2\sigma_n^2(x)|$, and set $\mu = \mu(x)$ and $W = Y - \mu$. Then

$$
\begin{aligned}
r_{3n}(t\,|\,x) &\le |E\{(e^{itW} - 1 - itW + \tfrac{1}{2}t^2W^2)I(|W| \le n^{1/2})\}| \\
&\quad + |E\{(e^{itW} - 1)I(|W| > n^{1/2})\}| \\
&\le t^2E\{W^2\min(1, |tW|)I(|W| \le n^{1/2})\} \\
&\quad + 2E\{\min(1, |tW|)I(|W| > n^{1/2})\}.
\end{aligned} \tag{3.4}
$$

Hence there exists $x_2 \ge x_1$, $\nu_1 \ge 1$ and $\varepsilon_2 \in (0, \varepsilon_1]$ such that

$$
r_{4n}(t\,|\,x) \equiv m r_{3n}(t/m^{1/2}\,|\,x) \le t^2/100 + 1 \tag{3.5}
$$

whenever $x \ge x_2$, $n \ge \nu_1$ and $0 \le t \le \varepsilon_2 n^{1/2}$. Using (3.4),

$$
\begin{aligned}
|r_{3n}(t\,|\,x)|^2 &\le 2t^6[E\{|W|^3I(|W| \le n^{1/2})\}]^2 \\
&\quad + 8t^2[E\{|W|I(|W| > n^{1/2})\}]^2.
\end{aligned} \tag{3.6}
$$

Furthermore,

$$
[E\{|W|I(|W| > n^{1/2})\}]^2 \le E\{W^2I(|W| > n^{1/2})\}P(|W| > n^{1/2}).
$$

Let A be an increasing function on $[0, \infty)$ with $A(0) \ge 1$, $A(\infty) = \infty$ and $E\{X^2A^2(2|X|)\} < \infty$. Then for any $y > 0$,

$$
\begin{aligned}
&[E\{|W|^3I(|W| \le n^{1/2})\}]^2 \\
&\qquad \le [y^3 + E\{|W|A(|W|)W^2I(|W| \le n^{1/2})\}/A(y)]^2 \\
&\qquad \le 2[y^6 + \{A(y)\}^{-2}E\{W^2A^2(|W|)\}E\{W^4I(|W| \le n^{1/2})\}].
\end{aligned}
$$

Note that $E\{W^2A^2(|W|)\}$ is bounded as $x \to \infty$. Combining the results from (3.6) down, we see that for any $\eta > 0$ we may choose $x_3 \ge x_2$, $\nu_2 \ge \nu_1$ and $\varepsilon_3 \in (0, \varepsilon_2]$, and a large value of y, such that

$$
\begin{aligned}
r_{5n}(t\,|\,x) &\equiv m^2|r_{3n}(t/m^{1/2}\,|\,x)|^2 \\
&\le \eta t^2(1 + t^4)[n^{-1}E\{W^4I(|W| \le n^{1/2})\} + nP(|W| > n^{1/2})] \\
&\quad + C_3t^2(1 + t^4)n^{-1},
\end{aligned} \tag{3.7}
$$

where C_3 does not depend on t, x or m. Consequently,

$$
\begin{aligned}
&\exp[m\{\alpha(t/m^{1/2}\,|\,x) - 1 - i(t/m^{1/2})\mu_n(x) + \tfrac{1}{2}(t/m^{1/2})^2\sigma_n^2(x)\}] \\
&\quad = 1 + m\{\alpha(t/m^{1/2}\,|\,x) - 1 - i(t/m^{1/2})\mu_n(x) + \tfrac{1}{2}(t/m^{1/2})^2\sigma_n^2(x)\} + r_{6n}(t\,|\,x),
\end{aligned}
$$

where by (3.5) and (3.7),

$$
\begin{aligned}
|r_{6n}(t\,|\,x)| &\le r_{5n}(t\,|\,x)\exp\{r_{4n}(t\,|\,x)\} \\
\text{(3.8)}\qquad &\le e\eta t^2(1+t^4)e^{t^2/100}[n^{-1}E\{(Y-\mu)^4 I(|Y-\mu|\le n^{1/2})\} \\
&\qquad + nP(|Y-\mu|>n^{1/2})] + C_4t^2(1+t^4)e^{t^2/100}n^{-1}.
\end{aligned}
$$

Choose $x_4 \ge x_3$ and $\nu_3 \ge \nu_2$ so large that $\sigma_n^2(x) \ge 27/50$ whenever $x \ge x_4$ and $n \ge \nu_3$. For such values of x and n, and for $0 \le t \le \varepsilon_3 n^{1/2}$, we see from (3.2), (3.3), (3.5), (3.7) and (3.8) that

$$
\begin{aligned}
&\{\alpha(t/m^{1/2}\,|\,x)\}^m \exp\{-itm^{1/2}\mu_n(x)\} \\
\text{(3.9)}\qquad &\quad = [1 + m\{\alpha(t/m^{1/2}\,|\,x) - 1 - i(t/m^{1/2})\mu_n(x) + \tfrac{1}{2}(t/m^{1/2})^2\sigma_n^2(x)\}] \\
&\qquad \cdot \exp\{-t^2\sigma_n^2(x)/2\} + r_{7n}(t\,|\,x),
\end{aligned}
$$

where

$$|r_{7n}(t\,|\,x)| \le (C_5\eta t^2(1+t^8)nE[\min\{1, ((Y-\mu)/n^{1/2})^4\}] + C_6t^2(1+t^8)n^{-1})e^{-t^2/4},$$

in which C_5 is an *absolute* constant. Thus,

$$
\begin{aligned}
&\int_0^{\varepsilon_3(27/50)^{1/2}n^{1/2}} t^{-1}\left|r_{7n}\left(\frac{t}{\sigma_n(x)}\,\middle|\,x\right)\right| dt \\
\text{(3.10)}\qquad &\qquad \le \int_0^{\varepsilon_3 n^{1/2}} t^{-1}|r_{7n}(t\,|\,x)|\,dt \\
&\qquad \le C_7\eta nE\left[\min\left\{1, \left(\frac{Y-\mu}{n^{1/2}}\right)^4\right\}\right] + C_8n^{-1},
\end{aligned}
$$

where C_7 is an absolute constant.

Let $L_n(\cdot\,|\,x)$ denote the function whose Fourier-Stieltjes transform, as a function of t, equals

$$\text{(3.11)}\qquad \psi_n(t\,|\,x) \equiv m\{\alpha(t/m^{1/2}\sigma_n\,|\,x) - 1 - i(t/m^{1/2}\sigma_n)\mu_n + \tfrac{1}{2}(t/m^{1/2}\sigma_n)^2\sigma_n^2\}e^{-t^2/2},$$

where we have written σ_n for $\sigma_n(x)$ and μ_n for $\mu_n(x)$. Thus,

$$
\begin{aligned}
L_n(z\,|\,x) = m\int_{|u|\le n^{1/2}} &\left\{\Phi\left(z - \frac{u}{m^{1/2}\sigma_n}\right) - \Phi(z) + \frac{u}{m^{1/2}\sigma_n}\phi(z)\right. \\
&\left. - \frac{1}{2}\left(\frac{u}{m^{1/2}\sigma_n}\right)^2\phi'(z)\right\} dP(Y-\mu\le u) \\
&+ \int_{|u|>n^{1/2}}\left\{\Phi\left(z-\frac{u}{m^{1/2}\sigma_n}\right) - \Phi(z)\right\} dP(Y-\mu\le u).
\end{aligned}
$$

We shall apply the smoothing inequality (see Theorem 2, page 109 of Petrov, 1975), in which we take $T = \varepsilon_3(27/50)^{1/2}n^{1/2}$,

"$F(z)$" $= P\{\sum_1^m (Y_i - \mu) \le m^{1/2}\sigma_n z + m\mu_n\}$ and "$G(z)$" $= \phi(z) + L_n(z \mid x)$.

In view of (3.9), (3.10) and the smoothing inequality, we have

$$\sup_{-\infty<z<\infty} |P\{\textstyle\sum_1^m (Y_i - \mu) \le m^{1/2}\sigma_n z + m\mu_n\} - \Phi(z) - L_n(z \mid x)| \le C_7\eta n E[\min\{1, ((Y-\mu)/n^{1/2})^4\}] + O(n^{-1/2}) \tag{3.12}$$

uniformly in $x \ge x_4$, as $n \to \infty$.

During applications of this result we shall take $x = |X_{nk}|$. Writing $p = 1 - F(x_4)$ and using the normal approximation to the binomial via an Edgeworth expansion, we see that

$$P(|X_{nk}| \le x_4) = \textstyle\sum_{j=0}^{k-1} \binom{n}{j} p^j(1-p)^{n-j} = O(n^{-\lambda}) \tag{3.13}$$

for all $\lambda > 0$, since $k/n \to 0$. The estimates (3.12) and (3.13) form the heart of our proof.

We may deduce from (3.12) and (3.13) that with $A_n = m^{1/2}\sigma_n(|X_{nk}|)$ and $B_n = m\{\mu(|X_{nk}|) + \mu_n(|X_{nk}|)\}$, we have

$$\sup_{-\infty<z<\infty} |P(^{(k)}S_n \le A_n z + B_n) - \Phi(z) - E\{L_n(z \mid |X_{nk}|)\}| \le Cn\int_0^\infty E\left[\min\left\{1, \left(\frac{Y-\mu}{n^{1/2}}\right)^4\right\}\right] dG_{nk}(x) + O(n^{-1/2}) \tag{3.14}$$

as $n \to \infty$. Since

$$|\Phi(z+v) - \Phi(z) - v\phi(z) - \tfrac{1}{2}v^2\phi'(z) - \tfrac{1}{6}v^3\phi''(z)| \le (1/4!)v^4 \sup_w |\phi'''(w)|$$

uniformly in z, then

$$\sup_{-\infty<z<\infty}\left| L_n(z \mid x) + \frac{1}{6}\phi''(z)m^{-1/2}\sigma_m^{-3}\int_{|u|\le m^{1/2}} u^3\, dP(Y-\mu \le u)\right| \le C_8 nE\left[\min\left\{1, \left(\frac{Y-\mu}{n^{1/2}}\right)^4\right\}\right] \tag{3.15}$$

uniformly in $x \ge x_5$, where C_8 is an absolute constant. Combining (3.13), (3.14) and (3.15), we see that

$$\begin{aligned} \sup_{-\infty<z<\infty} &|P(^{(k)}S_n \le A_n z + B_n) - \Phi(z)| \\ &\le C\left(n\int_0^\infty E\left[\min\left\{1, \left(\frac{Y-\mu}{n^{1/2}}\right)^4\right\}\right] dG_{nk}(x)\right. \\ &\quad \left. + n^{-1/2}\left|\int_0^\infty \{\sigma_n(x)\}^{-3}\, dG_{nk}(x) \int_{|u|\le n^{1/2}} u^3\, dP(Y-\mu \le u)\right|\right) \\ &\quad + O(n^{-1/2}). \end{aligned} \tag{3.16}$$

Next we shall derive a lower bound to

$$\xi_n \equiv \sup_{-\infty<z<\infty} |E\{L_n(z\,|\,|X_{nk}|)\}|.$$

Let $w > 0$ and define $b_w(t) = (w - t)te^{t^2/2}$ if $0 \le t \le w$, $b_w(t) = 0$ otherwise. Set

$$\hat{b}_w(z) = \int_0^w e^{itz} b_w(t)\, dt.$$

It is readily proved that $\hat{b}_w$ is absolutely integrable, and so the Fourier transform $\hat{b}_w$ may be inverted to yield

$$b_w(t) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-itz} \hat{b}_w(z)\, dz = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-itz} \bar{\hat{b}}_w(z)\, dz. \tag{3.17}$$

Note too that

$$\psi_n(t\,|\,x) = \int_{-\infty}^{\infty} e^{itz}\, dL_n(z\,|\,x) = -it \int_{-\infty}^{\infty} e^{itz} L_n(z\,|\,x)\, dz,$$

where $\psi_n(t\,|\,x)$ is defined by (3.11). Thus,

$$\int_{-\infty}^{\infty} e^{itz} E\{L_n(z\,|\,|X_{nk}|)\}\, dz = -(it)^{-1} \int_0^{\infty} \psi_n(t\,|\,x)\, dG_{nk}(x). \tag{3.18}$$

Applying Parseval's formula to the pair of Fourier transforms (3.17) and (3.18), we see that

$$-\int_{-\infty}^{\infty} (it)^{-1} b_w(t)\, dt \int_0^{\infty} \psi_n(t\,|\,x)\, dG_{nk}(x) = \int_{-\infty}^{\infty} E\{L_n(z\,|\,|X_{nk}|)\} \hat{b}_w(z)\, dz.$$

Therefore

$$\left| \int_0^{\infty} dG_{nk}(x) \int_0^1 (1 - t)\psi_n(tw\,|\,x) \exp\left(\frac{t^2 w^2}{2}\right) dt \right| \le C_9(w)\xi_n,$$

where $C_9(w)$ does not depend on n. Taking real parts in this expression, we find that

$$m \left| \int_0^{\infty} dG_{nk}(x) \int_0^1 (1 - t) E\left[\cos\left\{\frac{tw(Y - \mu)}{m^{1/2}\sigma_n}\right\} - 1 + \frac{1}{2}\left\{\frac{tw(Y - \mu)}{m^{1/2}\sigma_n}\right\}^2 \right] dt \right|$$
$$\le C_9(w)\,\xi_n.$$

Subtracting two versions of this inequality, one with $w = 1$ and the other with

$w = 2$, we may deduce that

$$(3.19)\qquad n\int_0^\infty dG_{nk}(x)\int_0^1 (1-t)E\left[1-\cos\left\{\frac{t(Y-\mu)}{m^{1/2}\sigma_n}\right\}\right]^2 dt \le C_{10}\xi_n.$$

Since

$$\int_0^1 (1-t)(1-\cos\theta t)^2\,dt \ge C\min\left\{1, \left(\frac{m^{1/2}\sigma_n\theta}{n^{1/2}}\right)^4\right\}$$

for all large x and large n, then in view of (3.13) and (3.19), there exists a constant C_{11} such that

$$(3.20)\qquad n\int_0^\infty E\left[\min\left\{1, \left(\frac{Y-\mu}{n^{1/2}}\right)^4\right\}\right] dG_{nk}(x) \le C_{11}(\xi_n + n^{-1})$$

for all large n.

Returning to the estimate (3.12), we choose the constant η to equal $1/2C_7C_{11}$. (Recall that C_7 is an absolute constant.) It then follows from (3.12) and (3.20) that

$$\sup_{-\infty<z<\infty}|P({}^{(k)}S_n \le A_nz + B_n) - \Phi(z)| + O(n^{-1/2})$$
$$\ge \left(\frac{1}{2C_{11}}\right)n\int_0^\infty E\left[\min\left\{1, \left(\frac{Y-\mu}{n^{1/2}}\right)^4\right\}\right] dG_{nk}(x).$$

When this estimate is combined with (3.15), we see that

$$\sup_{-\infty<z<\infty}|P({}^{(k)}S_n \le A_nz + B_n) - \Phi(z)| + O(n^{-1/2})$$
$$(3.21)\qquad \ge C_{12}\left(n\int_0^\infty E\left[\min\left\{1, \left(\frac{Y-\mu}{n^{1/2}}\right)^4\right\}\right] dG_{nk}(x) + n^{-1/2}\left|\int_0^\infty \{\sigma_n(x)\}^{-3}\,dG_{nk}(x)\int_{|u|\le n^{1/2}} u^3\,dP(Y-\mu\le u)\right|\right).$$

Theorem 1 follows from (3.16) and (3.21).

PROOF OF THEOREM 2. Theorem 2 can be proved in the same manner as Theorem 1, provided we show that the term $O(n^{-1/2})$ in (3.12) may be sharpened to $O(n^{-1})$. Now, (3.12) was derived via the smoothing inequality, in which we took $T = \varepsilon_3(27/50)^{1/2}n^{1/2} \equiv \varepsilon n^{1/2}$, say. If instead we take $T = n$, then the desired version of (3.12) will follow from (3.9) and (3.10) if we show that

$$(3.22)\qquad \int_{\varepsilon n^{1/2}}^n t^{-1}\left|r_{7n}\left(\frac{t}{\sigma_n(x)}\,\middle|\,x\right)\right| dt = O(n^{-1})$$

uniformly in $x \ge x_5$, for some $x_5 \ge x_4$.

The left side of (3.22) is dominated by

$$\int_{\varepsilon n^{1/2}}^{(50/27)^{1/2}n} t^{-1}|r_{7n}(t\,|\,x)|\,dt$$

(3.23)
$$\leq \int_{\varepsilon n^{1/2}}^{\infty}\left[\left|\alpha\left(\frac{t}{m^{1/2}}\,\Big|\,x\right)\right|^{m} + \{1 + Cm(1+t^2)\}e^{-27t^2/100}\right]dt$$

for some $C > 0$. (Use the definition (3.9) of $r_{7n}(t\,|\,x)$.) Furthermore,

$$\sup_{t\geq\varepsilon n^{1/2}}|\alpha(t/m^{1/2}|\,x)| \leq \sup_{t\geq\varepsilon}|\alpha(t\,|\,x)|$$
$$\leq \{\sup_{t\geq\varepsilon}|E(e^{itX})|\}/G(x) + \{1 - G(x)\}/G(x).$$

We may deduce from Cramér's condition (C) that by choosing x_5 sufficiently large, this quantity is dominated by a number $\rho < 1$ and for all $x \geq x_5$. The result (3.22) now follows from (3.23).

Theorem 3 may be proved by adapting the proof of Theorem 3.1, page 87 of Hall (1982).

Proof of Theorem 4. By a slight modification of the argument leading to Theorem 2 (see in particular the inequalities (3.12) and (3.13)) it may be provded that the ratio $(\Delta_n' + n^{-1})/(\delta_n + n^{-1})$ is bounded away from zero and infinity as $n \to \infty$, where

$$\Delta_n' = \sup_{-\infty<z<\infty}|P(^{(k)}S_n \leq m^{1/2}z) - E[\Phi\{z/\sigma_n(|X_{nk}|)\}]|.$$

Let $\Psi(z) = \Phi(z^{-1/2})$ for $z > 0$. Then for any $x > 0$ and $z > 0$,

$$\Phi\{z/\sigma_n(x)\} - \Phi(z) = z^{-2}\{\sigma_n^2(x) - 1\}\Psi'(z^{-2}) + \tfrac{1}{2}z^{-4}\{\sigma_n^2(x) - 1\}^2\Psi''(z^{-2}\theta_n(x)),$$

where $\theta_n(x)$ lies between 1 and $\sigma_n^2(x)$. Replacing x by $|X_{nk}|$ and taking expectations, and noting that $E\{\sigma_n^2(|X_{nk}|)\} \to 1$ and

$$E\{\sigma_n^2(|X_{nk}|) - 1\}^2 \geq \{E\sigma_n^2(|X_{nk}|) - 1\}^2,$$

we see that

$$|E[\Phi\{z/\sigma_n(|X_{nk}|)\}] - \Phi(z)| + O(n^{-1}) \geq \tfrac{1}{2}z^{-2}|E\{\sigma_n^2(|X_{nk}|)\} - 1|\Psi'(z^{-2}).$$

Therefore

$$\begin{aligned}\eta_n &\equiv |E\{\sigma_n^2(|X_{nk}|)\} - 1|\\ &\leq C_1[\sup_{-\infty<z<\infty}|P(^{(k)}S_n \leq m^{1/2}z) - E[\Phi\{z/\sigma_n(|X_{nk}|)\}]|\\ &\qquad + \sup_{-\infty<z<\infty}|P(^{(k)}S_n \leq m^{1/2}z) - \Phi(z)| + n^{-1}]\\ &\leq C_2(\delta_n + \Delta_n^{(1)} + n^{-1}) \leq C_3(\Delta_n^{(1)} + n^{-1}),\end{aligned}$$

using Theorem 3. Consequently $\delta_n + \eta_n \leq C_4(\Delta_n^{(1)} + n^{-1})$. It is also true that $(\Delta_n^{(2)} + n^{-1})/(\delta_n^{(2)} + n^{-1})$ is bounded as $n \to \infty$, where

$$\delta_n^{(2)} = E\{X^2I(|X| > n^{1/2})\} + n^{-1}E\{X^4I(|X| \leq n^{1/2})\};$$

see Corollary 4.6.2, page 184 of Hall (1982). Therefore if we prove that $2(\delta_n + \eta_n) + O(n^{-1}) \geq \delta_n^{(2)}$, it will follow that $(\Delta_n^{(1)} + n^{-1})/(\Delta_n^{(2)} + n^{-1})$ is bounded away from zero as $n \to \infty$.

To this end, observe that

$$1 - E\{\sigma_n^2(|X_{nk}|)\} = \left[1 - \int_{x\geq 0} dG_{nk}(x) \int_{0\leq u\leq \min(x,n^{1/2})} u^2\, dG(u)\right]$$
$$- \left[\int_{x\geq 0} dG_{nk}(x) \left(\frac{1}{G(x)} - 1\right) \int_{0\leq u\leq \min(x,n^{1/2})} u^2\, dG(u)\right]$$
$$= \eta_{n1} - \eta_{n2},$$

say. Now,

$$\eta_{n1} = \int_{x\geq 0} dG_{nk}(x) \int_{u>\min(x,n^{1/2})} u^2\, dG(u)$$
$$= \int_{x\geq 0} E\{X^2 I(|X| > x)\}\, dG_{nk}(x) + \int_{x>n^{1/2}} dG_{nk}(x) \int_{n^{1/2}<u\leq x} u^2\, dG(u).$$

Since $P(|X_{nk}| \leq x_o) = O(n^{-\lambda})$ for all x_o and all $\lambda > 0$, then

$$\eta_{n2} = \{1 + o(1)\} \int_{x\geq 0} dG_{nk}(x)\{1 - G(x)\} \int_{0\leq u\leq \min(x,n^{1/2})} u^2\, dG(u) + O(n^{-1})$$
$$= \{1 + o(1)\} \int_{x\geq 0} P(|X| > x)\, dG_{nk}(x) + O(n^{-1}).$$

But $P(|X| > x)/E\{X^2 I(|X| > x)\} \to 0$ as $x \to \infty$, and so

$$1 - E\{\sigma_n^2(|X_{nk}|)\} + O(n^{-1})$$
$$= \{1 + o(1)\} \int_{x\geq 0} E\{X^2 I(|X| > x)\}\, dG_{nk}(x)$$
$$+ \int_{x>n^{1/2}} dG_{nk}(x) \int_{n^{1/2}<u\leq x} u^2\, dG(u)$$
$$\sim \eta_{n1} = \int_{0\leq x\leq n^{1/2}} dG_{nk}(x) \int_{x<u\leq n^{1/2}} u^2\, dG(u)$$
$$+ \int_{0\leq x\leq n^{1/2}} dG_{nk}(x) \int_{u>n^{1/2}} u^2\, dG(u)$$
$$+ \int_{x>n^{1/2}} dG_{nk}(x) \int_{u>n^{1/2}} u^2\, dG(u)$$
$$= \int_{0<u\leq n^{1/2}} u^2 P(|X_{nk}| < u)\, dG(u) + E\{X^2 I(|X| > n^{1/2})\}$$
$$\geq n^{-1} \int_{0<u\leq n^{1/2}} u^4 P(|X_{nk}| < u)\, dG(u) + E\{X^2 I(|X| > n^{1/2})\}.$$

Since

$$\{1 + o(1)\}\delta_n + O(n^{-1}) \geq n^{-1} \int_{0 \leq u \leq n^{1/2}} u^4 P(|X_{nk}| \geq u)\, dG(u)$$

(see (2.1)) then

$$2[\delta_n + 1 - E\{\sigma_n^2(|X_{nk}|)\}] + O(n^{-1}) \geq n^{-1}E\{X^4 I(|X| \leq n^{1/2})\} + E\{X^2 I(|X| > n^{1/2})\}.$$

The term within square brackets is dominated by $\delta_n + \eta_n$, and so the proof is complete.

PROOF OF PROPOSITION 1. Let U_{nk} be the kth largest value of an n-sample from the uniform distribution on (0, 1). Then using (2.1),

$$\delta_n \sim n^{-1} \int_0^{1-n^{-\alpha/2}} (1-u)^{-4/\alpha} P(U_{nk} > u)\, du + n \int_{1-n^{-\alpha/2}}^1 P(U_{nk} > u)\, du$$

$$= n^{-1}\left(\frac{4}{\alpha} - 1\right)^{-1} \left\{ \int_0^1 (1-u)^{1-4/\alpha}\, dP(U_{nk} \leq u) - \int_{1-n^{-\alpha/2}}^1 (1-u)^{1-4/\alpha}\, dP(U_{nk} \leq u) \right\}$$

$$+ \left(1 - \frac{\alpha}{4}\right)^{-1} n^{1-\alpha/2} P(U_{nk} > 1 - n^{-\alpha/2}) - \left(\frac{4}{\alpha} - 1\right)^{-1} n^{-1}.$$

Let $d(k) = \Gamma(k - 4/\alpha + 1)/\Gamma(k) \sim k^{1-4/\alpha}$ as $k \to \infty$. Then

$$\int_0^1 (1-u)^{1-4/\alpha}\, dP(U_{nk} \leq u) = d(k)\Gamma(n+1) \Big/ \Gamma\left(n - \frac{4}{\alpha} + 2\right) \sim d(k) n^{4/\alpha - 1}$$

as $n \to \infty$. Furthermore,

$$n^{1-\alpha/2} P(U_{nk} > 1 - n^{-\alpha/2}) \leq n^{-1} \int_{1-n^{-\alpha/2}}^1 (1-u)^{1-4/\alpha}\, dP(U_{nk} \leq u)$$

$$= n^{-1} k \binom{n}{k} \int_{1-n^{-\alpha/2}}^1 u^{n-k}(1-u)^{k-4/\alpha}\, du$$

$$\leq C\left\{\frac{n^{k-1}}{\Gamma(k)}\right\} (n^{-\alpha/2})^{k-4/\alpha+1}$$

$$= C d(k) n^{4/\alpha - 1} n^{(k+1-4/\alpha)(1-\alpha/2)-1} \Big/ \Gamma\left(k - \frac{4}{\alpha} + 1\right)$$

$$= o\{d(k) n^{4/\alpha - 1}\}.$$

Therefore $\delta_n \sim (4/\alpha - 1)^{-1} d(k) n^{4/\alpha - 2}$, from which follows Proposition 1.

PROOF OF PROPOSITION 2. In this case,

$$\delta_n \sim n^{-1} \int_0^{n^{1/2}} dG_{nk}(x) \int_0^x u^4\, dG(u) \tag{3.24}$$
$$+ n \int_{n^{1/2}}^{\infty} dG_{nk}(x) \left\{ n^{-2} \int_0^{n^{1/2}} u^4\, dG(u) + \int_{n^{1/2}}^x dG(u) \right\}.$$

We shall prove that the second term in this expression is negligible in comparison with the first. Now,

$$n\delta_{n1} \equiv \int_0^{n^{1/2}} dG_{nk}(x) \int_0^x u^4\, dG(u) \sim \int_0^{n^{1/2}} x^2 |\log x|^{-\beta}\, dG_{nk}(x)$$

$$\sim 2^{2\beta} \int_0^{G(n^{1/2})} (1-u)^{-1}\{-\log(1-u)\}^{-2\beta}\, dP(U_{nk} \le u)$$

$$= 2^{2\beta} k \binom{n}{k} \int_0^{1-n^{-1}(\log n)^{-\beta}2^{\beta}} u^{n-k}(1-u)^{k-2}\{-\log(1-u)\}^{-2\beta}\, du.$$

Suppose first that $k \ge 2$, and observe that

$$\int_{1-n^{-1}(\log n)^{-\beta}2^{\beta}}^{1} u^{n-k}(1-u)^{k-2}\{-\log(1-u)\}^{-2\beta}\, du$$

$$\le C_1(\log n)^{-2\beta} \int_{1-n^{-1}(\log n)^{-\beta}2^{\beta}}^{1} (1-u)^{k-2}\, du \le C_2(2^{\beta}n^{-1})^{k-1}(\log n)^{-\beta(k+1)}k^{-1},$$

while

$$\int_0^1 u^{n-k}(1-u)^{k-2}\{-\log(1-u)\}^{-2\beta}\, du$$

$$= \left\{(k-1)\binom{n-1}{k-1}\right\}^{-1} E\{-\log(1-U_{n-1,k-1})\}^{-2\beta}$$

$$\sim \left\{(k-1)\binom{n-1}{k-1}\right\}^{-1} (\log n - \log k)^{-2\beta}.$$

Therefore

$$\delta_{n1} \sim 2^{2\beta}(k-1)^{-1}(\log n - \log k)^{-2\beta}. \tag{3.25}$$

When $k = 1$,

$$(2\beta - 1)n^{-1}k\binom{n}{k}\int_0^{1-n^{-1}(\log n)^{-\beta}2^{\beta}} u^{n-k}(1-u)^{k-2}\{-\log(1-u)\}^{-2\beta}\,du$$

$$= \{1 - n^{-1}(\log n)^{-\beta}2^{\beta}\}^{n-1}(\log n + \beta\log\log n - \beta\log 2)^{1-2\beta}$$

$$+ (n-1)\int_0^{1-n^{-1}(\log n)^{-\beta}2^{\beta}} u^{n-2}\{-\log(1-u)\}^{1-2\beta}\,du;$$

$$(n-1)\int_0^1 u^{n-2}\{-\log(1-u)\}^{1-2\beta}\,du = E\{-\log(1-U_{n-1,1})\}^{1-2\beta}$$

$$= (\log n)^{1-2\beta}\left\{1 + O\left(\frac{1}{\log n}\right)\right\};$$

and

$$(n-1)\int_{1-n^{-1}(\log n)^{-\beta}2^{\beta}}^1 u^{n-2}\{-\log(1-u)\}^{1-2\beta}\,du \le C_1 n\int_0^{n^{-1}(\log n)^{-\beta}2^{\beta}} (-\log u)^{1-2\beta}\,du$$

$$\le C_2(\log n)^{1-3\beta}.$$

Therefore

$$(2\beta - 1)2^{-2\beta}\delta_{n1}\{1 + o(1)\}$$

$$= -(\log n)^{1-2\beta}\{1 + \beta(1-2\beta)(\log n)^{-1}\log\log n$$

$$+ O(1/\log n)\} + (\log n)^{1-2\beta}\{1 + O(1/\log n)\} \tag{3.26}$$

$$\sim \beta(2\beta - 1)(\log n)^{-2\beta}\log\log n.$$

Since

$$n^{-1}\int_0^{n^{1/2}} u^4\,dG(u) + n\int_{n^{1/2}}^{\infty} dG(u) \le C(\log n)^{-\beta},$$

the second term in the expansion (3.24) is dominated by

$$C_1(\log n)^{-\beta}P(|X_{nk}| > n^{1/2}) \le C_1(\log n)^{-\beta}\binom{n}{k}P(|X_1|, \cdots, |X_k| > n^{1/2})$$

$$\le C_2(\log n)^{-(k+1)\beta}2^{\beta k}/k! = o(\delta_{n1})$$

in either the case $k \ge 2$ or $k = 1$. Proposition 2 follows on combining (3.24) – (3.26).

Proof of Proposition 3. The proof is similar in many respects to those of Propositions 1 and 2, and so we provide only an outline. The following results are readily proved:

$$\left(1-\frac{2}{\alpha}\right)E\{\sigma_n^2(|X_{nk}|)\}$$
$$= k\binom{n}{k}\int_0^1 u^{n-k-1}(1-u)^{k-1}[1-\{1-\min(u, 1-n^{-\alpha/2})\}^{1-2/\alpha}]\,du,$$

$$\left(1-\frac{2}{\alpha}\right)^2 E\{\sigma_n^4(|X_{nk}|)\}$$
$$= k\binom{n}{k}\int_0^1 u^{n-k-2}(1-u)^{k-1}[1-\{1-\min(u, 1-n^{-\alpha/2})\}^{1-2/\alpha}]^2\,du.$$

If we could replace the term $\min(u, 1-n^{-\alpha/2})$ by u in each of these expressions, the quantity $(1-2/\alpha)^2\operatorname{var}\{\sigma_n^2(|X_{nk}|)\}$ would equal

$$k\binom{n}{k}\left\{B(n-k-1, k) - 2B\left(n-k-1, k-\frac{2}{\alpha}+1\right) + B\left(n-k-1, k-\frac{4}{\alpha}+2\right)\right\}$$
$$- \left[k\binom{n}{k}\left\{B(n-k, k) - B\left(n-k, k-\frac{2}{\alpha}+1\right)\right\}\right]^2$$
$$= (n-k)^{-1}(n-k-1)^{-1}$$
$$\cdot\left[n(n-1) - 2n^{1+2/\alpha}\{1+O(n^{-1})\}\frac{\Gamma(k-2/\alpha+1)}{\Gamma(k)} + n^{4/\alpha}\{1+O(n^{-1})\}\frac{\Gamma(k-4/\alpha+2)}{\Gamma(k)}\right]$$
$$-(n-k)^{-2}\left[n - n^{2/\alpha}\{1+O(n^{-1})\}\frac{\Gamma(k-2/\alpha+1)}{\Gamma(k)}\right]^2$$
$$= (n-k)^{-2}n^{4/\alpha}\{1+O(n^{-1})\}\left[\frac{\Gamma(k-4/\alpha+2)}{\Gamma(k)} - \left\{\frac{\Gamma(k-2/\alpha+1)}{\Gamma(k)}\right\}^2\right]$$
$$= O\left\{n^{-1}\left(\frac{k}{n}\right)^{1-4/\alpha}\right\}$$

as $n \to \infty$. The error caused by replacing $\min(u, 1-n^{-\alpha/2})$ by u may be shown to be of this order of magnitude, and so

$$\operatorname{var}\{\sigma_n^2(|X_{nk}|)\} = O\{n^{-1}(k/n)^{1-4/\alpha}\} \tag{3.27}$$

as $n \to \infty$.

Let $V_n = \sigma_n^2(|X_{nk}|)$, $v_n = E(V_n)$ and $\Psi(z) = \Phi(z^{-1/2})$. Arguing as in the early part of the proof of Theorem 4, we see that with

$$\Delta_n'' = \sup_{-\infty<z<\infty}|P({}^{(k)}S_n \le a_n z) - E\{\Phi(z v_n^{1/2}/V_n^{1/2})\}|,$$

the ratio $(\Delta_n'' + n^{-1})/(\delta_n + n^{-1})$ is bounded away from zero and infinity as $n \to \infty$. Therefore in view of Theorem 3, Proposition 3 will follow if we prove that with

$$\eta_n \equiv \sup_{-\infty<z<\infty} |E\{\Phi(zv_n^{1/2}/V_n^{1/2})\} - \Phi(z)|,$$

we have $\eta_n \le C(\delta_n + n^{-1})$. From a short Taylor expansion of $\Phi(zv_n^{1/2}/V_n^{1/2}) = \Psi(z^{-2}V_n/v_n)$, and by considering the cases $|X_{nk}| \ge x_0$ and $|X_{nk}| < x_0$ for a large x_0, we may deduce that

$$\eta_n = O\{\operatorname{var}(V_n) + n^{-1}\}$$

as $n \to \infty$. The desired result now follows from (3.27) and Proposition 1.

Acknowledgment. I am grateful to the referee for his detailed comments.

REFERENCES

ANDREWS, D. F. *et al* (1972). *Robust Estimates of Location: Survey and Advances.* Princeton University Press, Princeton, N.J.

HALL, P. (1980). Characterizing the rate of convergence in the central limit theorem. *Ann. Probab.* **8** 1037–1048.

HALL, P. (1982). *Rates of Convergence in the Central Limit Theorem.* Pitman, London.

HATORI, H., MAEJIMA, M. and MORI, T. (1979). Convergence rates in the law of large numbers when extreme terms are excluded. *Z. Wahrsch. verw. Gebiete* **47** 1–12.

HUBER, P. J. (1981). *Robust Statistics.* Wiley, New York.

MORI, T. (1976). The strong law of large numbers when extreme terms are excluded from sums. *Z. Wahrsch. verw. Gebiete* **36** 189–194.

MORI, T. (1977). Stability for sums of i.i.d. random variables when extreme terms are excluded. *Z. Wahrsch. verw. Gebiete* **40** 159–167.

PETROV, V. V. (1975). *Sums of Independent Random Variables.* Springer, Berlin.

DEPARTMENT OF STATISTICS
THE FACULTIES
THE AUSTRALIAN NATIONAL UNIVERSITY
G.P.O. BOX 4
CANBERRA ACT 2601
AUSTRALIA