Ann. Inst. Statist. Math.
33 (1981), Part A, 449-462

LARGE SAMPLE PROPERTIES OF JAECKEL'S ADAPTIVE TRIMMED MEAN

PETER HALL

(Received June 11, 1980; revised July 6, 1981)

Summary

A critical examination of Jaeckel's (1971, *Ann. Math. Statist.*, 42, 1540-1552) study of his adaptive trimmed mean reveals that the theory is not applicable in many important cases, such as when the optimal trimming proportion is close to 0 or 1/2. This region includes the normal and double exponential distributions, among others, which have received considerable attention in the study of other adaptive location estimates. In this paper we obtain results which justify the use of Jaeckel's trimmed mean for a very large class of distributions. By restricting this class we obtain weak and strong rates of convergence which are much faster than those given by Jaeckel.

1. Introduction

Let $X_1, X_2, \cdots, X_n$ be independent random variables whose common distribution is symmetric about its median, θ. The α-trimmed mean of the sample is defined by

$$\bar{x}_\alpha=(n-2[\alpha n])^{-1}\sum_{i=[\alpha n]+1}^{n-[\alpha n]} X_{ni}\ , \qquad 0\leqq\alpha<1/2$$

where $X_{n1}\leqq X_{n2}\leqq\cdots\leqq X_{nn}$ denote the sample's order statistics and $[\alpha n]$ stands for the integer part of αn. The trimmed mean is an estimate of θ, and under mild regularity conditions, $n^{1/2}(\bar{x}_\alpha-\theta)$ is asymptotically normal $N(0, \sigma_\alpha^2)$. The asymptotically optimal choice of α, say α_0, is that one which minimizes σ_α^2. Jaeckel [7] proposed a method of estimating α_0. His approach was to construct an estimate of σ_α^2, say $\hat{\sigma}_\alpha^2$, and use the value of α which minimizes $\hat{\sigma}_\alpha^2$ as the trimming proportion in the construction of $\bar{x}_\alpha$. He proved a result which indicates that his adaptive procedure is "asymptotically as good as knowing and using the best α".

AMS (MOS) 1970 subject classification: Primary 62G05; Secondary 62E20.
Key words and phrases: Adaptive, estimate of location, rate of convergence, trimmed mean.

A critical examination of Jaeckel's work reveals a number of assumptions which seriously restrict the use of his estimate. Firstly, he stipulates that α_0 be not close to 0 or 1/2. This eliminates distributions like the normal (for which $\alpha_0=0$) and the double exponential ($\alpha_0=1/2$), both of which have received considerable attention in the construction of other adaptive estimates of location (see for example Prescott [9]). Jaeckel ([7], p. 1545) does discuss the possibility of a nonunique α_0, but in this case his procedure must be modified and requires the subjective choice of a sequence of constants $\{z_n\}$.

Other restrictive assumptions are that the estimate of α_0, say $\hat{\alpha}$, must be chosen from a predetermined range $[\alpha_1, \alpha_2]$ where $0<\alpha_1<\alpha_2<1/2$, and that the value of $\hat{\alpha}$ be unique. In this paper we relax all of these conditions, making Jaeckel's procedure more universally applicable. We also remove many of his restrictions on the smoothness of the underlying distribution—for the most part we do not even assume that there exists a density function.

Jaeckel's justification for his procedure was that his adaptive estimate satisfied the same central limit theorem as its optimal non-adaptive form, and so had the same asymptotic variance. The central limit theorem is only a "weak" measure of the rate of convergence of an estimate; a "strong" measure is provided by the law of the iterated logarithm. Under weaker conditions than Jaeckel imposed we show that his estimate satisfies the same law of the iterated logarithm as its optimal non-adaptive form.

Jaeckel [7] and Andrews *et al.* [1] report on Monte Carlo trials of the estimator. This work is mainly concerned with small sample sizes, but it is apparent that while the adaptive estimator is reasonably efficient for samples of about 20, the value of $\hat{\alpha}$ is still some distance from α_0. Jaeckel's result tells us only that $\bar{x}_{\hat{\alpha}}-\bar{x}_{\alpha_0}=o(n^{-1/2})$ in probability, and a faster rate of convergence is needed if the estimate is to be used for moderate sample sizes. To this end we prove that under slightly more restrictive conditions,

$$\bar{x}_{\hat{\alpha}}-\bar{x}_{\alpha_0}=\begin{cases}O(n^{-3/4}) & \text{in probability}\\ O((n^{-1}\log\log n)^{3/4}) & \text{a.s.}\end{cases}$$

(Here a.s. stands for "almost surely"; $\xrightarrow{p}$ will denote convergence in probability.)

2. Properties of the estimates

Let $X_1, X_2, \cdots$ be independent variables with a common continuous (but not necessarily absolutely continuous) distribution function $F(x-\theta)$

where θ is the unknown location parameter. Assume that the distribution is symmetric about θ, and set $F^{-1}(t)=\inf\{x\colon F(x)\geqq t\}$. If $0<\alpha<1/2$ then $n^{1/2}(\bar{x}_\alpha-\theta)$ has an asymptotic $N(0,\sigma_\alpha^2)$ distribution if and only if the $(1-\alpha)$th quantile is unique (Stigler [11]). In this case

$$\sigma_\alpha^2=2(1-2\alpha)^{-2}\left\{\int_0^{\xi_\alpha}x^2dF(x)+\alpha\xi_\alpha^2\right\}$$

where $\xi_\alpha=F^{-1}(1-\alpha)$. It makes little sense to compare the "spread" of normal and non-normal distributions on the basis of their variance, and so we are obliged to assume that ξ_α is uniquely defined for $0<\alpha\leqq 1/2$. That is, there are no gaps in the support of the distribution and F is both continuous and strictly increasing. In this case σ_α^2 is a continuous function of α.

Suppose $0<a<1/2$ and let $\mathscr{A}$ and $\mathscr{A}(a)$ be the sets of values in the ranges $0\leqq\alpha\leqq 1/2$ and $0\leqq\alpha\leqq a$, respectively, which minimize σ_α^2. If $1/2\in\mathscr{A}$ we assume that $\sigma_{1/2}^2\equiv\lim_{\alpha\to 1/2}\sigma_\alpha^2$ exists (finite), and therefore $(1-2\alpha)^{-1}\xi_\alpha$ is bounded as $\alpha\to 1/2$. Obviously $0\in\mathscr{A}$ or $\mathscr{A}(a)$ entails $\mathrm{E}(X_1^2)<\infty$. If $1/2\in\mathscr{A}$ then

$$\int_0^{\xi_\alpha}x^2dF(x)=O((1-2\alpha)^2)\int_0^{\xi_\alpha}dF(x)=o((1-2\alpha)^2)\ ,$$

and $\sigma_{1/2}^2=\lim_{\alpha\to 1/2}(1-2\alpha)^{-2}\xi_\alpha^2=\{2F'_+(0)\}^{-2}$. Consequently $1/2\in\mathscr{A}$ entails the existence of a non-zero right derivative $F'_+(0)$. (Possibly $F'_+(0)=\infty$.)

An estimate of σ_α^2 is given by

$$\hat{\sigma}_\alpha^2=(1-2\alpha)^{-2}\left\{n^{-1}\sum_{[\alpha n]+1}^{n-[\alpha n]}(X_{ni}-\bar{x}_\alpha)^2+\alpha(X_{n,[\alpha n]+1}-\bar{x}_\alpha)^2+\alpha(X_{n,n-[\alpha n]}-\bar{x}_\alpha)^2\right\}\ .$$

We choose α to minimize $\hat{\sigma}_\alpha^2$ but consider only values of α in the range $0\leqq\alpha\leqq 1/2-n^{-1/2}$. The reason for this restriction is that if $\alpha=\alpha(n)$ tends to 1/2 too quickly then $\hat{\sigma}_\alpha^2$ will not be a consistent estimator of $\sigma_{1/2}^2$. The variable $\hat{\sigma}_\alpha$ may be viewed as based on a naive estimator of the reciprocal of a density, similar to that proposed by Bloch and Gastwirth [4]. As is well known, such an estimator will not be consistent if the window size, $1/2-\alpha(n)$, converges to zero too quickly. In practical terms, the minimizing value of α should not be chosen "too close" to 1/2. (It will be clear from the proofs that the assumption $0\leqq\alpha(n)\leqq 1/2-Cn^{-1/2}$ for all n and any $C>0$, will suffice.)

Let $\mathscr{A}_n$ and $\mathscr{A}_n(a)$ be the sets of values in the ranges $0\leqq\alpha\leqq 1/2-n^{-1/2}$ and $0\leqq\alpha\leqq a$, respectively, which minimize $\hat{\sigma}_\alpha^2$. Our first concern is to show that the elements of $\mathscr{A}_n$ and $\mathscr{A}_n(a)$ converge to those of $\mathscr{A}$ and $\mathscr{A}(a)$. For each $\alpha\in\mathscr{A}_n$ or $\mathscr{A}_n(a)$, let $\beta=\beta(\alpha)$ be that element of $\mathscr{A}$ or $\mathscr{A}$(a), respectively, which is closest to α.

THEOREM 1. *If* $0<a<1/2$ *and* $\mathscr{A}(a)$ *is finite then*

$$\sup_{\alpha\in\mathscr{A}_n(a)}|\alpha-\beta(\alpha)|\to 0 \quad a.s. \tag{1}$$

If $\mathscr{A}$ *is finite, if* F *has a derivative in an interval* $(0, \varepsilon)$ *for some* $\varepsilon>0$, *and if* $F'(0+)$ *exists and is non-zero (possibly* $F'(0+)=\infty$*) then*

$$\sup_{\alpha\in\mathscr{A}_n}|\alpha-\beta(\alpha)| \xrightarrow{p} 0 . \tag{2}$$

The first part of Theorem 1 holds under weaker assumptions than Jaeckel's Lemma 3, and provides a strong version of his result.

THEOREM 2. *Suppose* $0<a<1/2$ *and* $\mathscr{A}(a)$ *is finite. If* $0\in\mathscr{A}(a)$, *assume that* $\mathrm{E}|X_1|^{2+\eta}<\infty$ *for some* $\eta>0$. *Then*

$$\sup_{\alpha\in\mathscr{A}_n(a)}|\bar{x}_\alpha-\bar{x}_{\beta(\alpha)}|=\begin{cases}o(n^{-1/2}) & \textit{in probability}\\ o((n^{-1}\log\log n)^{1/2}) & a.s.\end{cases}$$

If in addition F *has a derivative in an interval* $(0, \varepsilon)$, *and* $F'(0+)$ *exists and is non-zero, then* $\sup_{\alpha\in\mathscr{A}_n}|\bar{x}_\alpha-\bar{x}_{\beta(\alpha)}|=o(n^{-1/2})$ *in probability.*

From Theorem 2 and limit laws for the trimmed mean (see Shorack [10] and Wellner [12]) we may obtain weak and strong rates of convergence for the adaptive trimmed mean.

Our last result presents faster rates of convergence under more restrictive conditions. To simplify the proof we shall consider only the case of minimization in the range $0\leqq\alpha\leqq a$, where $a<1/2$, and assume that $\mathscr{A}(a)$ has only one element α_0, with $\alpha_0\neq 0$. Given a sequence of variables Δ_n, write $\Delta_n=\mathrm{O.B.}(\lambda)$ to denote that $\Delta_n=O(n^{-\lambda})$ in probability and $O((n^{-1}\log\log n)^{\lambda})$ a.s. Let $s(\alpha)=\sigma_\alpha^2$.

THEOREM 3. *Suppose* F *has a derivative in a neighbourhood of* ξ_{α_0}, *continuous and non-zero at* ξ_{α_0}, *and* s *has two derivatives in a neighbourhood of* α_0, *with* s'' *continuous and non-zero at* α_0. *Then*

$$\sup_{\alpha\in\mathscr{A}_n(a)}|\alpha-\alpha_0|=\mathrm{O.B.}(1/4) \quad \textit{and} \quad \sup_{\alpha\in\mathscr{A}_n(a)}|\bar{x}_\alpha-\bar{x}_{\alpha_0}|=\mathrm{O.B.}(3/4) .$$

3. The proofs

There is no loss of generality in assuming that $\theta=0$. In the work which follows some of the relations should strictly be qualified by "a.s.", but we shall omit this. During the proofs of Lemmas 1 and 2 we shall make use of the theorem of James [8], theorem A1 of Shorack [10] and theorem of 3 Wellner [12], which describe the asymptotic behaviour of the empiric process.

LEMMA 1. *Assume F has a derivative in an interval $(0, \varepsilon)$ for some $\varepsilon>0$, and $F'(0+)$ exists and is non-zero. Let m_n, $n\geqq 1$, be integers such that $m_n/n\to 0$ as $n\to\infty$ and $m_n\geqq n^{1/2}$. Then*

$$(3)\qquad \sup_{1\leqq m\leqq m_n}\left|X_{2n+1,n+1}-(2m+1)^{-1}\sum_{n-m+1}^{n+m+1}X_{2n+1,i}\right|=o(n^{-1/2})$$

and

$$(4)\qquad \sup_{n^{1/2}\leqq m\leqq m_n}|(n/m)(X_{2n+1,n+m}-X_{2n+1,n+1})-\{2F'(0+)\}^{-1}|=o(1)$$

in probability.

PROOF. We may assume that the distribution's support is confined to $(-\varepsilon, \varepsilon)$, and F' exists and is bounded away from zero on $(-\varepsilon, \varepsilon)$. For otherwise, choose ε so small that F' is bounded away from zero on $(0, \varepsilon)$, and replace F by any symmetric, absolutely continuous distribution G which agrees with F on $(-(1/2)\varepsilon, (1/2)\varepsilon)$, is bounded away from zero on $(-\varepsilon, \varepsilon)$ and satisfies $F(\varepsilon)=1$. Note that with probability tending to 1 all of the order statistics $X_{2n+1,i}$ with $|n-i|\leqq m_n+1$ lie in $(-(1/2)\varepsilon, (1/2)\varepsilon)$. For such an F, F^{-1} has a bounded derivative on $(0, 1)$, continuous at $x=1/2$. Let $G(x)=1-e^{-x}$, $H=F^{-1}\circ G$ and $c=H'(\log 2)=\{2F'(0+)\}^{-1}$. Then

$$(5)\qquad H(x)=H(y)+(x-y)c+(x-y)h(x, y)$$

where h is bounded on $(0, \infty)^2$ and $h(x, y)\to 0$ as $x, y\to\log 2$.

We may write $X_{ni}=H(Z_{ni})$, $1\leqq i\leqq n$, $n\geqq 1$, where the Z_i are independent exponential variables. In view of the expansion (5) with $x=Z_{ni}$ and $y=\mathrm{E}Z_{ni}$, (3) will follow if we prove that

$$(6)\qquad \sup_{1\leqq m\leqq m_n}\left|(Z_{2n+1,n+1}-\mathrm{E}Z_{2n+1,n+1})-(2m+1)^{-1}\sum_{n-m+1}^{n+m+1}(Z_{2n+1,i}-\mathrm{E}Z_{2n+1,i})\right|=o(n^{-1/2})$$

in probability, and

$$(7)\qquad \sup_{|n-i|\leqq m_n+1}\mathrm{E}|(Z_{2n+1,i}-\mathrm{E}Z_{2n+1,i})h(Z_{2n+1,i}, \mathrm{E}Z_{2n+1,i})|=o(n^{-1/2})\ .$$

The variables $Z_{2n+1,i}$, $1\leqq i\leqq 2n+1$, have the same distribution as

$$Z^*_{2n+1,i}=\sum_{1}^{i}Z_j/(2n+2-j)\ ,\qquad 1\leqq i\leqq 2n+1$$

(see David [5], p. 17), and the variables

$$L_n=\sup_{|n-i|\leqq m_n+1}|h(Z^*_{2n+1,i}, \mathrm{E}Z^*_{2n+1,i})|$$

are uniformly bounded and converge to zero in probability. If $|n-i|\leqq m_n+1$,

$$\mathrm{E}\,|\,(Z^*_{2n+1,i}-\mathrm{E}Z^*_{2n+1,i})h(Z^*_{2n+1,i},\,\mathrm{E}Z^*_{2n+1,i})\,|\leqq \mathrm{E}(L_nS_n)$$

where $S_n=\sup_{i\leqq n+m_n+1}|\sum_1^i(Z_j-1)/(2n+2-j)|$. Doob's inequality (Doob [6], p. 317) implies that $\mathrm{E}(S_n^2)=O(n^{-1})$, and so $\{n^{1/2}S_n,\ n\geqq 1\}$ is uniformly integrable. Consequently $\mathrm{E}(L_nS_n)=o(n^{-1/2})$, proving (7). To prove (6) we observe that

$$\begin{aligned}(Z^*_{2n+1,n+1}-\mathrm{E}Z^*_{2n+1,n+1})-(2m+1)^{-1}\sum_{n-m+1}^{n+m+1}(Z^*_{2n+1,i}-\mathrm{E}Z^*_{2n+1,i})\\ =(2m+1)^{-1}\Big\{\sum_{n-m+2}^{n+1}(Z_i-1)(m-n-1+i)(2n+2-i)^{-1}\\ -\sum_{n+2}^{n+m+1}(Z_i-1)(n+m+2-i)(2n+2-i)^{-1}\Big\}\ .\end{aligned}$$

We shall prove that

$$(8)\qquad \sup_{1\leqq m\leqq m_n}|(2m+1)^{-1}\sum_{i=1}^{m}(Z_i-1)i(m+n+1-i)^{-1}|=o(n^{-1/2})\ ;$$

condition (6) will follow from this and a similar result. Let $T_i=\sum_1^i(Z_j-1)$ and use Abel's method of summation to show that

$$\begin{aligned}\Delta_m&=\sum_{i=1}^{m}(Z_i-1)i(m+n+1-i)^{-1}\\ &=\sum_1^m(T_i-T_{i-1})i(m+n+1-i)^{-1}\\ &=-\sum_1^{m-1}T_i[(i+1)(m+n-i)^{-1}-i(m+n+1-i)^{-1}]+T_mm(n+1)^{-1}\\ &=-\sum_1^{m-1}T_i(m+n+1)(m+n+1-i)^{-1}(m+n-i)^{-1}+T_mm(n+1)^{-1}\ .\end{aligned}$$

Therefore

$$\sup_{1\leqq m\leqq m_n}m^{-1}|\Delta_m|\leqq\Big\{\sup_{1\leqq i\leqq m_n}|T_i|\Big\}\{(m_n+n+1)/n^2+n^{-1}\}\ .$$

Donsker's invariance principle (see Billingsley [2], p. 68) implies that $\sup_{i\leqq n}|T_i|=O(n^{1/2})$ in probability, and so (8) holds.

To prove (4), let $Y_{nm}=X_{2n+1,n+m}-X_{2n+1,n+1}$. In view of (5) and (7) we have

$$Y_{nm}-\mathrm{E}Y_{nm}=\{(Z_{2n+1,n+m}-\mathrm{E}Z_{2n+1,n+m})-(Z_{2n+1,n+1}-\mathrm{E}Z_{2n+1,n+1})\}c+o(n^{-1/2})$$

in probability, uniformly in $n^{1/2}\leqq m\leqq m_n$. Since

$$(Z^*_{2n+1,n+m}-EZ^*_{2n+1,n+m})-(Z^*_{2n+1,n+1}-\mathrm{E}Z^*_{2n+1,n+1})=\sum_{n+2}^{n+m}(Z_i-1)/(2n+2-i)$$

then it suffices to prove that

$$(9)\qquad \Delta_{n1}=\sup_{n^{1/2}\le m\le m_n}(n/m)\left|\sum_1^m(Z_i-1)/(n+1-i)\right|=o(1)$$

in probability, and

$$(10)\qquad \Delta_{n2}=\sup_{n^{1/2}\le m\le m_n}|(n/m)\mathrm{E}Y_{nm}-\{2F'(0+)\}^{-1}|\to 0\ .$$

Using Abel's method of summation again we find that

$$\left|\sum_1^m(Z_i-1)/(n+1-i)\right|\le\left\{\sup_{1\le i\le m}|T_i|\right\}\{m/(n-m)^2+1/(n-m)\}\ ,$$

and so for a constant C,

$$\Delta_{n1}\le Cn^{-1/2}\sup_{1\le i\le m_n}|T_i|=O((m_n/n)^{1/2})$$

in probability, proving (9). Finally, in view of (5) and (7),

$$\begin{aligned}\mathrm{E}(Y_{nm})&=H(\mathrm{E}Z_{2n+1,n+m})-H(\mathrm{E}Z_{2n+1,n+1})+o(n^{-1/2})\\&=\mathrm{E}(Z_{2n+1,n+m}-Z_{2n+1,n+1})(c+o(1))+o(n^{-1/2})\ ,\end{aligned}$$

and since $\mathrm{E}(Z_{2n+1,n+m}-Z_{2n+1,n+1})\sim m/n$, (10) is proven.

For $0\le\alpha<1/2$ define $\bar{y}_\alpha=n^{-1}\sum_{[\alpha n]+1}^{n-[\alpha n]}X_{ni}$.

LEMMA 2. *Suppose* $\varepsilon_n\downarrow 0$. *If* $0<\alpha<1/2$ *then*

$$\sup_{0<\varepsilon<\varepsilon_n}|\bar{y}_\alpha-\bar{y}_{\alpha+\varepsilon}|=\begin{cases}o(n^{-1/2}) & \text{in probability}\\ o((n^{-1}\log\log n)^{1/2}) & a.s.\end{cases}$$

If $\mathrm{E}|X_1|^{2+\eta}<\infty$ *for some* $\eta>0$ *then the result is also true for* $\alpha=0$.

PROOF. It suffices to prove that there exists a function $f(\delta)\to 0$ as $\delta\to 0$ and variables $Z_n(\delta)$ such that for each $\delta>0$, (i) $\sup_{0<\varepsilon<\varepsilon_n}|\bar{y}_\alpha-\bar{y}_{\alpha+\varepsilon}|\le f(\delta)Z_n(\delta)$, (ii) $n^{1/2}Z_n(\delta)$ has a limiting distribution not depending on δ, and (iii) $(n/\log\log n)^{1/2}Z_n(\delta)$, $n\ge 1$, has its a.s. limit points confined to a compact interval not depending on δ. Suppose $0<\alpha<1/2$ and $0<\alpha-\delta<\alpha<\alpha+\delta<1/2$, and define $Y_n^\delta=X_n$ if $-\xi_{\alpha-\delta}<X_n<-\xi_{\alpha+\delta}$; $-\xi_{\alpha-\delta}$ if $X_n\le-\xi_{\alpha-\delta}$; $-\xi_{\alpha+\delta}$ if $X_n\ge-\xi_{\alpha+\delta}$, and $Z_n^\delta=X_n$ if $\xi_{\alpha+\delta}<X_n<\xi_{\alpha-\delta}$; $\xi_{\alpha+\delta}$ if $X_n\le\xi_{\alpha+\delta}$; $\xi_{\alpha-\delta}$ if $X_n\ge\xi_{\alpha-\delta}$. Let G_δ and H_δ be the respective distribution functions of Y_1^δ and Z_1^δ, and Y_{ni}^δ and Z_{ni}^δ the order statistics. With probability tending to 1 as $m\to\infty$,

$$|\bar{y}_\alpha - \bar{y}_{\alpha+\varepsilon}| = \left| n^{-1} \sum_{[\alpha n]+1}^{[(\alpha+\varepsilon)n]} X_{ni} + n^{-1} \sum_{n-[(\alpha+\varepsilon)n]}^{n-[\alpha n]} X_{ni} \right| = \left| n^{-1} \sum_{[\alpha n]+1}^{[(\alpha+\varepsilon)n]} Y_{ni}^{\delta} + n^{-1} \sum_{n-[(\alpha+\varepsilon)n]+1}^{n-[\alpha n]} Z_{ni}^{\delta} \right| \tag{11}$$

for all $0<\varepsilon<\varepsilon_n$ and $n \geqq m$.

Define $a_n(t)=1$ if $[\alpha n]+1 \leqq nt \leqq [(\alpha+\varepsilon)n]$; 0 otherwise, and $a_n^*(t) = a_n(1+n^{-1}-t)$, $b_n(t)=\int_0^t a_n(u)du$ and $b_n^*(t)=\int_t^1 a_n^*(u)du = b_n(1-t)+O(n^{-1})$ uniformly in ε and t. Then

$$\int_0^1 b_n^*(t) dH_\delta^{-1}(t) = \int_0^1 b_n^*(1-t) dG_\delta^{-1}(t) = \int_0^1 b_n(t) dG_\delta^{-1}(t) + O(n^{-1})$$

uniformly in ε. (Here and below, $\int_a^b \cdot dB(t) = \int_{a<t\leqq b} \cdot dB(t)$ for any function B of bounded variation.) We may write $Y_{ni}^{\delta} = G_\delta^{-1}(U_{ni})$ and $Z_{ni}^{\delta} = H_\delta^{-1}(U_{ni})$ for $1 \leqq i \leqq n$ and $n \geqq 1$, where the $U_i = F(X_i)$ are uniform (0, 1) variables. Let G_n be the empiric d.f. of $U_1, \cdots, U_n$. Then

$$n^{-1} \sum_{[\alpha n]+1}^{[(\alpha+\varepsilon)n]} Y_{ni}^{\delta} = n^{-1} \sum_1^n G_\delta^{-1}(U_i) a_n(G_n(U_i)) = \int_0^1 G_\delta^{-1}(t) db_n(G_n(t)) = G_\delta^{-1}(1) b_n(1) - \int_0^1 b_n(G_n(t)) dG_\delta^{-1}(t)$$

and

$$n^{-1} \sum_{n-[(\alpha+\varepsilon)n]+1}^{n-[\alpha n]} Z_{ni}^{\delta} = H_\delta^{-1}(0+) b_n^*(0) + \int_0^1 b_n^*(G_n(t)) dH_\delta^{-1}(t) .$$

Therefore the right-hand side of (11) is dominated by

$$\Delta_n = \int_0^1 |b_n(G_n(t)) - b_n(t)| dG_\delta^{-1}(t) + \int_0^1 |b_n^*(G_n(t)) - b_n^*(t)| dH_\delta^{-1}(t) + O(n^{-1}) \quad .$$

Each term within modulus signs is dominated by $|G_n(t)-t|$, and so $\Delta_n \leqq f(\delta) Z_n(\delta)$ where $(\delta) = \int_0^1 d\{G_\delta^{-1}(t) + H_\delta^{-1}(t)\} \to 0$ as $\delta \to 0$, and $Z_n(\delta) = \sup|G_n(t) - t| + O(n^{-1})$ satisfies conditions (ii) and (iii).

It remains to consider the case $\alpha=0$. For $\delta>0$ define $Y_n^\delta = X_n$ if $X_n < -\xi_\delta$; $-\xi_\delta$ otherwise, and $Z_n^\delta = X_n$ if $X_n > \xi_\delta$; ξ_δ otherwise. Let G_δ and H_δ be the respective distribution functions of Y_1^δ and Z_1^δ. With probability tending to 1 as $m \to \infty$,

$$|\bar{y}_0 - \bar{y}_\varepsilon| = \left| n^{-1} \sum_1^{[\varepsilon n]} X_{ni} + n^{-1} \sum_{n-[\varepsilon n]+1}^{n} X_{ni} \right| = \left| n^{-1} \sum_1^{[\varepsilon n]} Y_{ni}^\delta + n^{-1} \sum_{n-[\varepsilon n]+1}^{n} Z_{ni}^\delta \right| \tag{12}$$

for all $0<\varepsilon<\varepsilon_n$ and $n\geqq m$. Define $c_n(t)=1$ if $nt\leqq[\varepsilon n]$; 0 otherwise, and $c_n^*(t)=c_n(1+n^{-1}-t)$, $d_n(t)=\int_0^t c_n(u)du$ and $d_n^*(t)=\int_t^1 c_n^*(u)du$. If $\mathrm{E}|X_1|^{2+\eta}<\infty$ and $(2+\eta)(1/2-\xi)>1$ then

$$\int_0^1 |d_n^*(t)-d_n(1-t)|\,dH_\delta^{-1}(t)\leqq\int_0^1 \min(1-t,\, n^{-1})dH_\delta^{-1}(t)$$
$$\leqq n^{-1/2-\xi}\int_0^1 (1-t)^{1/2-\xi}dH_\delta^{-1}(t)=O(n^{-1/2-\xi})\ .$$

Now, $Y_{ni}^\delta=G_\delta^{-1}(U_{ni})$ and $Z_{ni}^\delta=H_\delta^{-1}(U_{ni})$. Moreover,

$$n^{-1}\sum_1^{[\varepsilon n]} G^{-1}(U_{ni})=G_\delta^{-1}(1)d_n(1)-\int_0^1 d_n(G_n(t))dG_\delta^{-1}(t)$$

and

$$n^{-1}\sum_{n-[\varepsilon n]+1}^{n} H_\delta^{-1}(U_{ni})=H_\delta^{-1}(0+)d_n^*(0)+\int_0^1 d_n^*(G_n(t))dH_\delta^{-1}(t)\ .$$

Combining these estimates we see that the right side of (12) is dominated by

$$\varDelta_n=\int_0^1 |d_n(G_n(t))-d_n(t)|\,dG_\delta^{-1}(t)+\int_0^1 |d_n^*(G_n(t))-d_n^*(t)|\,dH_\delta^{-1}(t)$$
$$+O(n^{-1/2-\xi})$$

uniformly in ε. Therefore $\varDelta_n\leqq f(\delta)Z_n(\delta)$ where for $(2+\eta)(1/2-\xi)>0$,

$$f(\delta)=\int_0^1 [t(1-t)]^{1/2-\xi}d[G_\delta^{-1}(t)+H_\delta^{-1}(t)]\to 0 \quad \text{as} \quad \delta\to 0$$

and $Z_n(\delta)=\sup[t(1-t)]^{-1/2+\xi}|G_n(t)-t|+O(n^{-1/2-\xi})$. Here G_n is the empiric distribution function of $F(X_1),\cdots,F(X_n)$. By theorem A1 of Shorack [10] $n^{1/2}Z_n(\delta)$ has a limiting distribution not depending on δ [condition (ii) above], while by the theorem of James [8] or theorem 3 of Wellner [12], $(n/\log\log n)^{1/2}Z_n(\delta)$ has its a.s. limit points confined to a compact interval not depending on δ [condition (iii)].

PROOF of THEOREM 1. Let $\mathscr{A}_\infty(a)$ be the set of points such that $\mathrm{P}(\{\text{for some } \beta\in\mathscr{A}_n(a),\ |\alpha-\beta|<\varepsilon\} \text{ infinitely often})=1$ for all $\varepsilon>0$. (The tail σ-field of the sequence $X_1, X_2, \cdots$ is almost trivial, and so the probability on the left above takes only the values 0 and 1.) For $\alpha\in\mathscr{A}_\infty(a)$, let $\gamma=\gamma(\alpha)$ be the element of $\mathscr{A}_n(a)$ closest to α. If $\alpha\neq 0$ and $0<\xi\leqq a$,

$$\hat\sigma_\xi^2\geqq\hat\sigma_\gamma^2=(1-2\gamma)^{-2}\Big\{n^{-1}\sum_{[\gamma n]+1}^{n-[\gamma n]}X_{ni}^2-n^{-1}(n-2[\gamma n])\bar x_\gamma^2$$
$$+\gamma(X_{n,[\gamma n]+1}-\bar x_\gamma)^2+\gamma(X_{n,n-[\gamma n]}-\bar x_\gamma)^2\Big\}\ .$$

In view of the continuity of F and σ_α^2, the right side has σ_α^2 as an a.s. limit point, while the left side converges a.s. to σ_ξ^2. Therefore $\alpha \in \mathscr{A}(a)$. If $\alpha=0$ then $\liminf \gamma=0$ a.s., and if $0<\xi\leqq a$ and $\gamma<\eta\leqq a$ then

$$\begin{aligned}\hat{\sigma}_\xi^2 \geqq \hat{\sigma}_\gamma^2 &\geqq (1-2\gamma)^{-2} n^{-1} \sum_{[\gamma n]+1}^{n-[\gamma n]} (X_{ni}-\bar{x}_\gamma)^2 \\ &\geqq (1-2\gamma)^{-2} n^{-1} \sum_{[\eta n]+1}^{n-[\eta n]} (X_{ni}-\bar{x}_\gamma)^2 \geqq (1-2\gamma)^{-2} n^{-1} \sum_{[\eta n]+1}^{n-[\eta n]} (X_{ni}-\bar{x}_\eta)^2 .\end{aligned}$$

The right side has $2\int_0^{\xi_\eta} x^2 dF(x)$ as an a.s. limit point, while the left side converges a.s. to σ_ξ^2. (Note that $\liminf \gamma(n)=0$ a.s.) Letting $\eta\to 0$ we deduce that $0\in\mathscr{A}(a)$. Therefore $\mathscr{A}_\infty(a)\subseteqq\mathscr{A}(a)$. We shall prove that this implies condition (1).

Since $\mathscr{A}_n(a)$ is finite, there exists a point $\alpha_n\in\mathscr{A}_n(a)$ at which the supremum on the left side of (1) is achieved. In view of the Hewitt-Savage zero-one law (see Breiman [3], page 63) the quantities

$$\limsup_{n\to\infty} \{\alpha_n-\beta(\alpha_n)\} , \qquad \liminf_{n\to\infty} \{\alpha_n-\beta(\alpha_n)\}$$

are constant a.s. If we assume that (1) fails then either the lim usp is positive or the lim inf is negative. We may confine attention to the former case, and assume that for some $\delta>0$, $\limsup_{n\to\infty}\{\alpha_n-\beta(\alpha_n)\}=\delta$ a.s. Let A be the set of sample points ω at which this relation holds. Then $P(A)=1$, and for each $\omega\in A$ there exists an increasing sequence $n_k=n_k(\omega)$ defined by the smallest integer n such that $|\alpha_n(\omega)-\beta(\alpha_n(\omega))-\delta|<1/k$. Let

$$\lambda(\omega)=\limsup_{k\to\infty}\{\beta(\alpha_{n_k})\omega))+\delta\} ,$$

and choose an increasing subsequence $\{m_k\}$ of $\{n_k\}$ such that $\beta(\alpha_{m_k}(\omega))+\delta\to\lambda(\omega)$. Then $\alpha_{m_k}\to\lambda$ a.s., and for all $\varepsilon>0$,

$$P(\{\text{for some } \beta\in\mathscr{A}_n(a),\ |\lambda-\beta|<\varepsilon\} \ \text{infinitely often})=1 .$$

Since $\mathscr{A}(a)$ is finite then $\beta(\alpha_{m_k})+\delta$ can assume only a finite range of values, and so the random variable λ takes only a finite range of values. One of these values, say λ_0, must therefore satisfy both $P(\lambda=\lambda_0)>0$ and

$$P(\{\text{for some } \beta\in\mathscr{A}_n(a),\ |\lambda_0-\beta|<\varepsilon\} \ \text{infinitely often})=1$$

for all $\varepsilon>0$. This means that $\lambda_0\in\mathscr{A}_\infty(a)$, and consequently $\lambda_0\in\mathscr{A}(a)$. Since $\beta(\alpha)$ equals that value in $\mathscr{A}(a)$ which is closest to α, then on the set $\{\lambda=\lambda_0\}$ we have

$$|\alpha_{m_k}-\beta(\alpha_{m_k})|\leqq|\alpha_{m_k}-\lambda_0|=|\alpha_{m_k}-\lambda| .$$

But the left side converges a.s. to δ and the right side converges a.s. to zero as $k\to\infty$ — a contradiction. Therefore our original assumption—that (1) is false—must have been wrong.

A similar argument shows that (2) will follow if we prove that whenever

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty} P(\sup \mathscr{A}_n > 1/2-\varepsilon) > 0 ,$$

we have $1/2 \in \mathscr{A}$. We shall assume that the lim sup may be taken over odd n, and in order to simplify the notation, that

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty} P(\sup \mathscr{A}_{2n+1} > 1/2-\varepsilon) = 2\delta > 0 .$$

(The case of even n may be treated similarly.) With $m=2n+1$ and $\gamma=\sup\mathscr{A}_m$ we have for any $0<\xi<1/2$ and large n,

$$\hat\sigma_\xi^2 \geqq \hat\sigma_\gamma^2 \geqq \gamma(1-2\gamma)^{-2}\{(X_{m,[\gamma m]+1}-\bar x_\gamma)^2+(X_{m,m-[\gamma m]}-\bar x_\gamma)^2\} . \tag{13}$$

Lemma 1 implies that for any $\varepsilon>0$ and all large n,

$$\begin{aligned}&P(|(1-2\gamma)^{-1}(X_{m,m-[\gamma m]}-X_{m,n+1})-\{2F'(0+)\}^{-1}|<\varepsilon\\ &\quad\text{and}\quad n^{1/2}|X_{m,n+1}-\bar x_\gamma|<\varepsilon)>\delta .\end{aligned}$$

Since each $\alpha\in\mathscr{A}_n$ satisfies $\alpha\leqq 1/2-n^{-1/2}$ then $n^{1/2}(1-2\gamma)\geqq 2$, and so

$$P(|(1-2\gamma)^{-1}(X_{m,m-[\gamma m]}-\bar x_\gamma)-\{2F'(0+)\}^{-1}|<3\varepsilon/2)>\delta .$$

A similar argument can be applied to the first term on the right in (13) and so for all $\varepsilon>0$ and all large n we have with probability greater than δ that $\hat\sigma_\xi^2>\{2F'(0+)\}^{-2}-\varepsilon$. Since $\hat\sigma_\xi^2\xrightarrow{p}\sigma_\xi^2$ then $\sigma_\xi^2\geqq\{2F'(0+)\}^{-2}$, and so $1/2\in\mathscr{A}$.

PROOF OF THEOREM 2. An analogue of Lemma 2 in which the supremum is taken over $-\varepsilon_n<\varepsilon<0$, is easily proved. Theorem 2 follows easily from this, Lemmas 1 and 2 and Theorem 1. Note that

$$\begin{aligned}|\bar x_\alpha-\bar x_{\alpha+\varepsilon}|&\leqq n\{n-2[(\alpha+\varepsilon)n]\}^{-1}|\bar y_\alpha-\bar y_{\alpha+\varepsilon}|\\ &\quad+2n\{n-2[\alpha n]\}^{-1}\{n-2[(\alpha+\varepsilon)n]\}^{-1}|\{[(\alpha+\varepsilon)n]-[\alpha n]\}\bar y_\alpha| ,\end{aligned} \tag{14}$$

and $|\bar y_\alpha|$ is $O(n^{-1/2})$ in probability and $O((n^{-1}\log\log n)^{1/2})$ a.s.

PROOF OF THEOREM 3. Consider the expansions

$$\begin{aligned}(1-2\alpha)^2(\hat\sigma_\alpha^2-\sigma_\alpha^2)&=\left\{\int_{X_{n,[\alpha n]}}^{-\xi_\alpha}+\int_{\xi_\alpha}^{X_{n,n-[\alpha n]}}\right\}x^2dF_n(x)\\ &\quad+\int_{-\xi_\alpha}^{\xi_\alpha}x^2d\{F_n(x)-F(x)\}+\alpha\{(X_{n,[\alpha n]+1}-\bar x_\alpha)^2+(X_{n,n-[\alpha n]}-\bar x_\alpha)^2-2\xi_\alpha^2\}\end{aligned}$$

and

$$n^{-1}(n-2[\alpha n])\bar{x}_\alpha = \left\{\int_{X_{n,[\alpha n]}}^{-\xi_\alpha} + \int_{\xi_\alpha}^{X_{n,n-[\alpha n]}}\right\} x dF_n(x) + \int_{-\xi_\alpha}^{\xi_\alpha} x d\{F_n(x)-F(x)\} , \tag{15}$$

where F_n is the empiric distribution function and $\int_{a<x\leq b}$ has the obvious interpretation if $b<a$. Evaluate the integrals by parts; for example

$$\int_{X_{n,[\alpha n]}}^{-\xi_\alpha} x dF_n(x) = \xi_\alpha\{n^{-1}([\alpha n]+1) - F_n(-\xi_\alpha)\} - n^{-1}([\alpha n]+1)\{\xi_\alpha + X_{n,[\alpha n]}\} - \int_{X_{n,[\alpha n]}}^{-\xi_\alpha} F_n(x)dx$$

and

$$\int_{-\xi_\alpha}^{\xi_\alpha} x d\{F_n(x)-F(x)\} = \xi_\alpha\{F_n(\xi_\alpha)-F(\xi_\alpha)+F_n(-\xi_\alpha)-F(-\xi_\alpha)\} - \int_{-\xi_\alpha}^{\xi_\alpha}\{F_n(x)-F(x)\}dx .$$

Applying the probability transformation we deduce from the usual limit theorems for empiric and quantile processes that for small $\varepsilon>0$,

$$\sup_{|\alpha-\alpha_0|<\varepsilon}\{|\xi_\alpha + X_{n,[\alpha n]}| + |\xi_\alpha - X_{n,n-[\alpha n]}|\} + \sup_x |F_n(x)-F(x)| = \mathrm{O.B.}(1/2) .$$

Substituting the expansion of $\bar{x}_\alpha$ into that of $\hat{\sigma}_\alpha^2 - \sigma_\alpha^2$ we obtain

$$\sup_{|\alpha-\alpha_0|<\varepsilon} |\hat{\sigma}_\alpha^2 - \sigma_\alpha^2| = \mathrm{O.B.}(1/2) .$$

From this fact and since

$$0 \leq \sigma_\alpha^2 - \sigma_{\alpha_0}^2 = \sigma_\alpha^2 - \hat{\sigma}_\alpha^2 + \hat{\sigma}_\alpha^2 - \hat{\sigma}_{\alpha_0}^2 + \hat{\sigma}_{\alpha_0}^2 - \sigma_{\alpha_0}^2 \tag{16}$$

we see that $0 \leq \hat{\sigma}_\alpha^2 - \hat{\sigma}_{\alpha_0}^2 + \mathrm{O.B.}(1/2)$ uniformly in $|\alpha-\alpha_0|<\varepsilon$. With probability tending to 1 as $m\to\infty$, $|\alpha-\alpha_0|<\varepsilon$ for all $\alpha \in \mathscr{A}_n(a)$ and all $n \geq m$ (use Theorem 1). Since $\hat{\sigma}_\alpha^2 - \hat{\sigma}_{\alpha_0}^2 \leq 0$ for all $\alpha \in \mathscr{A}_n(a)$, then

$$\sup_{\alpha\in\mathscr{A}_n(a)} |\hat{\sigma}_\alpha^2 - \hat{\sigma}_{\alpha_0}^2| = \mathrm{O.B.}(1/2) .$$

Substituting this back into (16) we see that

$$\sup_{\alpha\in\mathscr{A}_n(a)} |\sigma_\alpha^2 - \sigma_{\alpha_0}^2| = \mathrm{O.B.}(1/2) .$$

But $\sigma_\alpha^2 - \sigma_{\alpha_0}^2 \sim (1/2)(\alpha-\alpha_0)^2 s''(\alpha_0)$ as $\alpha\to\alpha_0$, and so

$$\sup_{\alpha \in \mathscr{A}_n(a)} |\alpha - \alpha_0| = \text{O.B.}(1/4) , \tag{17}$$

proving the first part of the theorem. We shall prove only the "in probability" portion of the second part of Theorem 3; the "a.s." portion may be proved in the same way.

Let $\xi > 0$. In view of (17) we may choose C_1 so large that

$$\mathrm{P}\Big(\sup_{\alpha \in \mathscr{A}_n(a)} |\alpha - \alpha_0| < C_1 n^{-1/4}\Big) > 1 - \xi \tag{18}$$

for all n. Define U_{ni} as in the proof of Lemma 2, and observe that for $\varepsilon > 0$ there exists C_2 so large that for all $n \geqq N(\varepsilon)$,

$$\mathrm{P}(|U_{ni} - i(n+1)^{-1}| \leqq C_2 n^{-1/2} \text{ for } \varepsilon n \leqq i \leqq n(1-\varepsilon)) > 1 - \xi$$

(see Jaeckel [7], Lemma 1). If ε is small then for all $n \geqq N(\varepsilon)$ we have with probability $> 1 - \xi$ that for all i with $|i - [n\alpha_0]| \leqq 2C_1 n^{3/4}$,

$$|U_{ni} - i(n+1)^{-1}| + |i(n+1)^{-1} - \alpha_0| \leqq C_2 n^{-1/2} + 2C_1 n^{-1/4} + 2/n .$$

Let $\delta = \delta(n) = 4C_1 n^{-1/4}$. For large n,

$$\begin{aligned} 1 - \xi &< \mathrm{P}(|U_{ni} - \alpha_0| < \delta \text{ for all } |i - [n\alpha_0]| \leqq 2C_1 n^{3/4}) \\ &= \mathrm{P}(-\xi_{\alpha_0 - \delta} < X_{ni} < -\xi_{\alpha_0 + \delta} \text{ for all } |i - [n\alpha_0]| \leqq (1/2)n\delta) . \end{aligned} \tag{19}$$

Similarly $\mathrm{P}(\xi_{\alpha_0+\delta} < X_{ni} < \xi_{\alpha_0-\delta}$ for all $|i - n + [n\alpha_0]| \leqq (1/2)n\delta) > 1 - \xi$. Combining this with (18) and (19), and defining Y_{ni}^{δ}, Z_{ni}^{δ}, G_{δ} and H_{δ} as in the proof of Lemma 2, with α_0 replacing α, we see that for large n, with probability $> 1 - 3\xi$,

$$\sup_{\alpha \in \mathscr{A}_n(a)} |\bar{y} - \bar{y}_{\alpha_0}| \leqq \sup_{|\alpha - \alpha_0| < \delta/4} |\bar{y} - \bar{y}_{\alpha_0}| ,$$

$X_{ni} = Y_{ni}^{\delta}$ for $|i - [n\alpha_0]| \leqq (1/2)n\delta$, and $X_{ni} = Z_{ni}^{\delta}$ for $|i - n + [n\alpha_0]| \leqq (1/2)n\delta$. Arguing as in proof of Lemma 2 we deduce that with probability $>$ $1 - 3\xi$,

$$\begin{aligned} \sup_{\alpha \in \mathscr{A}_n(a)} |\bar{y} - \bar{y}_{\alpha_0}| &\leqq \Delta_{n1} + \Delta_{n2} \\ &\leqq 2\int_0^1 d\{G_{\delta}^{-1}(t) + H_{\delta}^{-1}(t)\}\{\sup_t |G_n(t) - 1|\} + O(n^{-1}) . \end{aligned}$$

(Here Δ_{n1} is for $\alpha > \alpha_0$ and Δ_{n2} for $\alpha < \alpha_0$.) Since $\int_0^1 d\{G_{\delta}^{-1}(t) + H_{\delta}^{-1}(t)\} = O(\delta) = O(n^{-1/4})$ then the right side is $O(n^{-3/4})$ in probability. Finally, using (14) with $(\alpha, \alpha + \varepsilon)$ replaced (α_0, α) we see that the desired result follows from (17) and the fact that $\bar{y}_{\alpha_0} = O(n^{-1/2})$ in probability.

Acknowledgement

I am most grateful to the referee for his detailed and pertinent remarks. These enabled me to clarify my exposition.

AUSTRALIAN NATIONAL UNIVERSITY

REFERENCES

[1] Andrews, D. F., Bickel, P. J., Hampel, F. R., Huber, P. J., Rogers, W. H. and Tukey, J. W. (1972). *Robust Estimates of Location*, Princeton University Press.

[2] Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

[3] Breiman, L. (1968). *Probability*, Addison-Wesley, Reading, Mass.

[4] Bloch, D. A. and Gastwirth, J. L. (1968). On a simple estimate of the reciprocal of the density function, *Ann. Math. Statist.*, **39**, 1083–1085.

[5] David, H. A. (1970). *Order Statistics*, Wiley, New York.

[6] Doob, J. L. (1953). *Stochastic Processes*, Wiley, New York.

[7] Jaeckel, L. B. (1971). Some flexible estimates of location, *Ann. Math. Statist.*, **42**, 1540–1552.

[8] James, B. R. (1975). A functional law of the iterated logarithm for weighted empirical distributions, *Ann. Prob.*, **3**, 762–772.

[9] Prescott, P. (1978). Selection of trimming proportions for robust adaptive trimmed means, *J. Amer. Statist. Ass.*, **73**, 133–139.

[10] Shorack, G. R. (1972). Functions of order statistics, *Ann. Math. Statist.*, **43**, 412–427.

[11] Stigler, S. M. (1973). The asymptotic distribution of the trimmed mean, *Ann. Statist.*, **1**, 472–477.

[12] Wellner, J. A. (1977). A law of the iterated logarithm for functions of order statistics, *Ann, Statist.*, **5**, 481–494.