*The Annals of Statistics*
1974, Vol. 2, No. 4, 661–675

# RANDOM MEANS[1]

By Galen R. Shorack

*University of Washington*

General theorems on asymptotic normality of randomly trimmed and Winsorized means are obtained. A new small sample studentization is proposed. Many examples are presented. The two sample problem is also considered.

**1. Introduction.** Let $X_{n1} \leqq \cdots \leqq X_{nn}$ denote the order statistics of a random sample from $F_\theta = F(\cdot - \theta)$ where $F$ is an unknown df. Let $g = F^{-1}$ denote the left continuous inverse of $F$. Let $\alpha \equiv \alpha(F)$ and $\beta \equiv \beta(F)$ denote numbers (possibly unknown values of $F(\cdot)$, possibly known fixed numbers, etc.) satisfying $0 \leqq \alpha < \beta \leqq 1$. Let $A \equiv A(F) = g(\alpha)$ and $B \equiv B(F) = g(\beta)$. (The above notation means that $\alpha$, $\beta$, $A$ and $B$ may depend on $F$.)

Let $\alpha_n$ and $\beta_n$ be integer valued random functions of $X_{n1}, \cdots, X_{nn}$ for which $0 \leqq \alpha_n < \beta_n \leqq n$. Let

$$T_n = \textstyle\sum_{\alpha_n+1}^{\beta_n} X_{ni}/(\beta_n - \alpha_n) \tag{1}$$

denote the general *randomly trimmed mean*. Let

$$T_n^* = [\alpha_n X_{n\alpha_n+1} + \textstyle\sum_{\alpha_n+1}^{\beta_n} X_{ni} + (n - \beta_n)X_{n\beta_n}]/n \tag{1*}$$

denote the general *randomly Winsorized mean*. We will call $\alpha_n/n$ and $(n - \beta_n)/n$ the *random adjustment percentages*.

We will concentrate initially on the case

$$F \text{ is symmetric about } 0\,, \qquad \beta = 1 - \alpha \qquad \text{and} \qquad A = -B\,. \tag{S}$$

In this situation it seems natural that the adjustment percentages should be symmetric in some sense.

*Method* 1. $T_n$ and $T_n^*$ will be called *nearly symmetric random means or symmetric random means* according as

$$\alpha_n - (n - \beta_n) = o_p(n^{\frac{1}{2}}) \tag{S1}$$

or

$$\alpha_n = n - \beta_n\,. \tag{S2}$$

*Method* 2. Let $\alpha_n$ and $n - \beta_n$ denote the number of observations less than $\hat\theta_0 - \hat B$ and greater than $\hat\theta_0 + \hat B$ respectively, where the preliminary estimate $\hat\theta_0 \equiv \hat\theta_{0,n}$ of $\theta$ and the positive rv $\hat B \equiv \hat B_n$ satisfy

$$\hat\theta_0 - \theta = O_p(n^{-\frac{1}{2}}) \tag{2}$$

Received September 1971; revised July 1973.

[1] This research was supported by National Science Foundation Grant GP31361X.

*AMS* 1970 *subject classification.* Primary 62G35, 62E20.

*Key words and phrases.* Randomly trimmed means, small sample robustness.

and

$$\hat{B} - B = O_p(n^{-\frac{1}{2}}) . \tag{3}$$

The resulting estimates will be called *metrically symmetrized random means.*

Asymptotic normality of Methods 1 and 2 random means is established in Theorems 1 and 2 respectively. Studentization of the more promising of these random means is considered in Corollaries 1 and 2. In particular, it is conjectured that the studentization of certain estimates $\tau_n$ provided by Corollary 2 should be robust in small samples.

The two-sample problem is considered in Section 6.

We will require of the random adjustment percentages that for some (possibly unknown) $\alpha$ and $\beta$ either

(A1) $\alpha_n/n = \alpha + o_p(1)$ and $\beta_n/n = \beta + o_p(1)$,

(A2) $\alpha_n/n = \alpha + o_p(n^{-\frac{1}{4}})$ and $\beta_n/n = \beta + o_p(n^{-\frac{1}{4}})$,

(A3) $\alpha_n/n = \alpha + O_p(n^{-\frac{1}{2}})$ and $\beta_n/n = \beta + O_p(n^{-\frac{1}{2}})$, or

(A4) $\alpha_n/n = \alpha + o_p(n^{-\frac{1}{2}})$ and $\beta_n/n = \beta + o_p(n^{-\frac{1}{2}})$.

We will also require that $F$ satisfy either

(F1) $g$ is continuous at $\alpha$ and $\beta$,

(F2) $g$ has a derivative at $\alpha$ and $\beta$,

(F3) $g$ satisfies a Lipschitz condition in neighborhoods of $\alpha$ and $\beta$, or

(F4) $F$ has a strictly positive continuous derivative $f$ in neighborhoods of $A$ and $B$.

We remark that $g$ is continuous at $t$ in $(0, 1)$ if and only if there is at most one $x$ for which $F(x) = t$. Note that (F1)—(F3) do not imply that $F$ is continuous at $A$ and $B$.

**2. Representations of trimmed and Winsorized means.** For purposes of our proofs, we suppose that $X_{ni} = g(\xi_{ni})$ for $1 \leqq i \leqq n$ where $0 < \xi_{n1} < \cdots < \xi_{nn} < 1$ are the special Uniform (0, 1) order statistics described in the appendix of Shorack (1972). Let $\Gamma_n$ denote the empirical df of these $\xi_{ni}$'s and let $U_n(t) = n^{\frac{1}{2}}[\Gamma_n(t) - t]$ for $0 \leqq t \leqq 1$ denote their *empirical process.* These $\xi_{ni}$'s are special because they satisfy $\rho(U_n, U) \equiv \sup_{0 \leqq t \leqq 1} |U_n(t) - U(t)| \to 0$ as $n \to \infty$ for all points in the probability space; here $U$ is a special Brownian bridge on [0, 1] having continuous sample paths. Also the *quantile process* $V_n(t) = n^{\frac{1}{2}}[\Gamma_n^{-1}(t) - t]$ for $0 \leqq t \leqq 1$ satisfies $\rho(V_n, V) \to 0$ for all points in the probability space; here $V = -U$ is also a Brownian bridge. Note that

$$Z(t) = -[U(t) + U(1 - t)]/2^{\frac{1}{2}} \tag{4}$$

is Brownian motion for $0 \leqq t \leqq \frac{1}{2}$.

This special construction will enable us to represent the limits in distribution of $T_n$ and $T_n^*$ as certain functionals of $U$ that are being converged to in the strong sense of $\to_p$.

LEMMA 1. (*The trimmed mean*). (a) (*General F*). *If* (A3) *and* (F1) *hold or if* (A2) *and either* (F2) *or* (F3) *hold, then*

$$n^{\frac{1}{2}}(T_n - \mu) = -\{\textstyle\int_\alpha^\beta U\,dg + (A-\mu)n^{\frac{1}{2}}[\alpha_n/n - \alpha] - (B-\mu)n^{\frac{1}{2}}[\beta_n/n - \beta]\}/(\beta-\alpha) + \varepsilon_n$$

*where* $\mu = \int_\alpha^\beta g(t)\,dt/(\beta-\alpha)$ *and where* $\varepsilon_n \to_p 0$ *as* $n \to \infty$.

(b) (*Symmetric F*). *If* (S), (S1), (F3) *and* (A1) *hold, then*

$$n^{\frac{1}{2}}(T_n - \theta) = -\textstyle\int_\alpha^{1-\alpha} U\,dg/(1-2\alpha) + \varepsilon_n$$

*where* $\varepsilon_n \to_p 0$ *as* $n \to \infty$.

LEMMA 2. (*The Winsorized mean*). (a) (*General F*). *If* (F2) *and* (A2) *hold, then*

$$n^{\frac{1}{2}}(T_n{}^* - \mu^*) = -[\textstyle\int_\alpha^\beta U\,dg + \alpha g'(\alpha)U(\alpha) + (1-\beta)g'(\beta)U(\beta)] + \alpha g'(\alpha)n^{\frac{1}{2}}[\alpha_n/n - \alpha] + (1-\beta)g'(\beta)n^{\frac{1}{2}}[\beta_n/n - \beta] + \varepsilon_n$$

*where* $\mu^* = \alpha A + \int_\alpha^\beta g(t)\,dt + (1-\beta)B$ *and where* $\varepsilon_n \to_p 0$ *as* $n \to \infty$.

(b) (*Symmetric F*). *If* (S), (S1), (F2), (F3) *and* (A1) *hold, then*

$$n^{\frac{1}{2}}(T_n{}^* - \theta) = -\{\textstyle\int_\alpha^{1-\alpha} U\,dg + \alpha g'(\alpha)[U(\alpha) + U(1-\alpha)]\} + \varepsilon_n$$

*where* $\varepsilon_n \to_p 0$ *as* $n \to \infty$.

Note from (4) that when (S) holds, $-\int_\alpha^\beta U\,dg = 2^{\frac{1}{2}}\int_\alpha^{\frac{1}{2}} Z\,dg$ provided $g$ is continuous at $\frac{1}{2}$.

These fundamental lemmas are proved in Section 7.

**3. The main results for symmetric** $F$. Throughout this section we suppose that the symmetry condition (S) holds.

We define

$$\sigma^2 \equiv \sigma^2(\alpha) = [\textstyle\int_{-B}^{B} x^2\,dF(x) + 2\alpha B^2]/(1-2\alpha)^2 \tag{5}$$

and

$$\sigma_*{}^2 \equiv \sigma_*{}^2(\alpha) = \textstyle\int_{-B}^{B} x^2\,dF(x) + 2\alpha[B + \alpha g'(\alpha)]^2 . \tag{5*}$$

THEOREM 1. (*Asymptotic normality of nearly symmetric random means*).

(i) *If* (S), (S1), (F3) *and* (A1) *hold, then*

$$n^{\frac{1}{2}}(T_n - \theta) \to_d N(0, \sigma^2) .$$

(ii) *If* (S), (S1), (F2), (F3) and (A1) *hold, then*

$$n^{\frac{1}{2}}(T_n{}^* - \theta) \to_d N(0, \sigma_*{}^2) .$$

PROOF. (i) It is immediate from Lemma 1(b) that $n^{\frac{1}{2}}(T_n - \theta) \to_p M_\alpha$ where

$$M_\alpha \equiv -\textstyle\int_\alpha^\beta U\,dg/(\beta-\alpha) . \tag{6}$$

This limiting rv is $N(0, \sigma^2)$ by Lemma 3 below; let $K(t)$ equal $A$, $g(t)$, $B$ according as $0 < t < \alpha$, $\alpha \leqq t \leqq \beta$, $\beta < t < 1$. Jaeckel's (1971) development of Example 6 below implies a proof of this result under the condition that $F$ has a density that is strictly positive and continuous on the interval of support.

(ii) Likewise, from Lemma 2(b) we obtain $n^{\frac{1}{2}}(T_n{}^* - \theta) \to_p M_\alpha{}^*$ where

$$(6^*) \qquad M_\alpha{}^* \equiv -[\alpha g'(\alpha)U(\alpha) + \textstyle\int_\alpha^\beta U\,dg + (1-\beta)g'(\beta)U(\beta)]\,.$$

This limiting rv is seen to be $N(0, \sigma_*{}^2)$ by Lemma 3 and some tedious book-keeping. □

Let $\hat\theta_0$ denote any consistent estimate of $\theta$. Denote the *Winsorized sample variance* about $\hat\theta_0$ by

$$(7) \qquad V_n{}^2 = \frac{n[\alpha_n(X_{n\alpha_n+1} - \hat\theta_0)^2 + \sum_{\alpha_n+1}^{\beta_n}(X_{ni} - \hat\theta_0)^2 + (n-\beta_n)(X_{n\beta_n} - \hat\theta_0)^2]}{(\beta_n - \alpha_n)(\beta_n - \alpha_n - 1)}\,.$$

COROLLARY 1. (*Studentization of* $T_n$). *If* (S), (S1), (F3) *and* (A1) *hold, then*

$$n^{\frac{1}{2}}(T_n - \theta)/V_n \quad \textit{is approximately distributed as} \quad t_{\beta_n-\alpha_n-1}\,.$$

PROOF. Replace $\hat\theta_0$ by $\theta$ in (7) and call the result $\bar V_n{}^2$. Now $\bar V_n{}^2 \to \sigma^2$ by the proof of Lemma 2(b) with $g$ replaced by $g^2 = (g_\theta - \theta)^2$; note that (F2) is not needed if we only claim this much. Also $V_n{}^2 - \bar V_n{}^2 \to_p 0$ since $\hat\theta_0 \to_p \theta$. Thus $V_n \to_p \sigma$, which combined with Theorem 1 gives $n^{\frac{1}{2}}(T_n - \theta)/V_n \to_d N(0, 1)$.

The use of $t_{\beta_n-\alpha_n-1}$ instead of $N(0, 1)$ was proposed by Tukey and McLaughlin (1963); and illuminating comment is found in Huber (1970). □

EXAMPLE 1. (Ordinary trimmed and Winsorized means). Let $0 \leqq \alpha < \frac{1}{2}$ be a fixed known number, and let $\alpha_n = n - \beta_n = [n\alpha]$ equal the greatest integer in $n\alpha$. In this important special case we denote (1) and (1*) by $T_n(\alpha)$ and $T_n{}^*(\alpha)$ respectively. Note that (A3) is trivially true with orders of magnitude to spare. Thus under (S) we have from Lemma 1(a) and Lemma 2(a) respectively, that $n^{\frac{1}{2}}[T_n(\alpha) - \theta] \to_d N(0, \sigma^2)$ if (F1) holds and $n^{\frac{1}{2}}[T_n{}^*(\alpha) - \theta] \to_d N(0, \sigma_*{}^2)$ if (F2) holds. (See also Bickel (1965) and Huber (1969). See Corollaries 3 and 4 below for the case of asymmetric $F$ that depend on $n$.) □

We now turn from Method 1 random means and consider Method 2 random means. To phrase our theorem we will need the following example.

EXAMPLE 2. (The truncated mean). Let $Y_i$ equal $X_i$ or 0 according as $|X_i - \theta| \leqq B$ or as $|X_i - \theta| > B$. Let $\bar T_n = \sum_1^n Y_i/n$ denote the *truncated mean*; and note that it is unobservable. By setting $\hat\theta_0 = \theta$ and $\hat B = B$ in Theorem 2(i) below we find that

$$(8) \qquad n^{\frac{1}{2}}(\bar T_n - \theta) \to_d \bar M_\alpha \equiv [-AU(\alpha) - \textstyle\int_\alpha^\beta U\,dg + BU(\beta)]/(\beta - \alpha)$$

provided (S) and (F4) hold. The limiting rv $\bar M_\alpha$ is $N(0, \bar\sigma^2)$ where $\bar\sigma^2 = \int_A^B x^2\,dF(x)/(\beta-\alpha)^2$, as follows from Lemma 3 below. (Asymptotic normality of $\bar T_n$ also follows from the ordinary central limit theorem.) □

THEOREM 2. (*Asymptotic normality of metrically symmetrized random means*). *Suppose* (S), (2), (3) *and* (F4) *hold.* (i) *Then*

$$n^{\frac{1}{2}}(T_n - \theta) = \bar M_\alpha + [2Bf(B)/(1-2\alpha)]n^{\frac{1}{2}}(\hat\theta_0 - \theta) + \varepsilon_n$$

*where* $\bar{M}_\alpha$ *is defined in* (8) *and where* $\varepsilon_n \to_p 0$ *as* $n \to \infty$. (ii) *Also*

$$n^{\frac{1}{2}}(T_n^* - \theta) = (1 - 2\alpha)M_\alpha + 2\alpha n^{\frac{1}{2}}(\hat{\theta}_0 - \theta) + \varepsilon_n$$

*where* $M_\alpha$ *is defined in* (6) *and where* $\varepsilon_n \to_p 0$ *as* $n \to \infty$.

PROOF. Apply Lemmas 1(a) and 2(a) to $F_\theta$ noting that $\mu_\theta = \theta$, $dg_\theta = dg$, $A_\theta - \mu_\theta = A = -B$, $B_\theta - \mu_\theta = B$, $\mu_\theta^* = \theta$, $g_\theta'(\alpha) = g_\theta'(\beta) = g'(\alpha) = 1/f(B)$. It remains only to consider $\alpha_n$ and $\beta_n$ in the formulas of the lemmas. Let $\hat{\alpha} = F_\theta(\hat{\theta}_0 + \hat{A})$ and $\hat{\beta} = F_\theta(\hat{\theta}_0 + \hat{B})$ where $\hat{A} = -\hat{B}$; then

$$\begin{aligned} n^{\frac{1}{2}}(\alpha_n/n - \alpha) &= U_n(\hat{\alpha}) + n^{\frac{1}{2}}(\hat{\alpha} - \alpha) + o_p(1) \\ &= U_n(\hat{\alpha}) + n^{\frac{1}{2}}[(\hat{\theta}_0 + \hat{A}) - (\theta + A)][F_\theta(\hat{\theta}_0 + \hat{A}) - F_\theta(\theta + A)] \\ &\qquad \div [(\hat{\theta}_0 + \hat{A}) - (\theta + A)] + o_p(1) \\ &= U(\alpha) + n^{\frac{1}{2}}[(\hat{\theta}_0 + \hat{A}) - (\theta + A)]f_\theta(\theta + A) + o_p(1), \end{aligned}$$

while

$$\begin{aligned} n^{\frac{1}{2}}(\beta_n/n - \beta) &= U_n(\hat{\beta}) + n^{\frac{1}{2}}(\hat{\beta} - \beta) \\ &= U(\beta) + n^{\frac{1}{2}}[(\hat{\theta}_0 + \hat{B}) - (\theta + B)]f_\theta(\theta + B) + o_p(1). \end{aligned}$$

Note that $f_\theta(\theta + A) = f_\theta(\theta + B)$ leads to cancellation of the $\hat{A}$ and $\hat{B}$ terms. This completes the proof.

Note that if assumption (S) is dropped we still have

$$\begin{aligned} n^{\frac{1}{2}}(T_n^* - \mu^*) &= (\beta - \alpha)M_{\alpha\beta} + (\alpha + 1 - \beta)n^{\frac{1}{2}}(\hat{\theta}_0 - \theta) \\ &\quad + \alpha n^{\frac{1}{2}}(\hat{A} - A) + (1 - \beta)n^{\frac{1}{2}}(\hat{B} - B) + o_p(1), \end{aligned} \tag{9}$$

where we now use $M_{\alpha\beta}$ to denote the expression in (6). This expression will be used in Section 6. □

EXAMPLE 3. (A special random mean, $\tau_n$). Suppose the random variables $\bar{\alpha}_n = n - \bar{\beta}_n$ satisfy (A3); and let $T_n$ denote the trimmed mean of (1) based on these choices $\bar{\alpha}_n$ and $\bar{\beta}_n$. Examine the $n$ residuals $|X_{ni} - T_n|$ and reject the $2\bar{\alpha}_n$ observations whose residuals are largest; let $\alpha_n$ (let $n - \beta_n$) denote the number of observations rejected whose residuals were negative (positive). Let $\tau_n$ denote the Winsorized mean of (1*) based on these $\alpha_n$ and $\beta_n$; and let $\nu_n^2$ denote the variance of (7) based on $\alpha_n$, $\beta_n$ and $\hat{\theta}_0 = \tau_n$. (Note that $\tau_n$ is a metrically Winsorized mean about the preliminary symmetrically trimmed random mean $T_n$; and $\nu_n^2$ is the matching Winsorized sample variance about $\tau_n$ divided by an estimate of $(\beta - \alpha)^2$.) □

COROLLARY 2. (*Studentization of* $\tau_n$). *If* (S) *holds, if* $\bar{\alpha}_n = n - \bar{\beta}_n$ *satisfy* (A3) *for some* $0 < \alpha < \frac{1}{2}$, *and if* $F$ *satisfies* (F4), *then*

$$n^{\frac{1}{2}}(\tau_n - \theta)/\nu_n \quad \textit{is approximately distributed as} \quad t_{\beta_n - \alpha_n - 1}.$$

PROOF. (The following proof is valid with $\alpha_n = n - \beta_n$ replaced by (S1).) From Theorem 2, and then Theorem 1, we obtain

$$\begin{aligned} n^{\frac{1}{2}}(\tau_n - \theta) &= (1 - 2\alpha)M_\alpha + 2\alpha n^{\frac{1}{2}}(T_n - \theta) + o_p(1) \\ &= (1 - 2\alpha)M_\alpha + 2\alpha M_\alpha + o_p(1) = M_\alpha + o_p(1); \end{aligned}$$

so that $n^{\frac{1}{2}}(\tau_n - \theta) \to_d N(0, \sigma^2)$. Also $\nu_n \to_p \sigma$ as in the proof of Corollary 1, since $\alpha_n$ and $\beta_n$ satisfy (A3). □

REMARK. In Theorem 2 and in (9) we can replace (2) and (3) by the weaker hypotheses $\hat{\theta}_0 - \theta = o_p(n^{-\frac{1}{4}})$ and $\hat{B} - B = o_p(n^{-\frac{1}{4}})$, provided we require in addition to (F4) that

(F5) $[f(A + \varepsilon) - f(A)]/\varepsilon$ and $[f(B + \varepsilon) - f(B)]/\varepsilon$ are bounded for $\varepsilon$ in some neighborhood of $0$.

Thus (A2) may replace (A3) in Corollary 2 if (F5) is added.

**4. Examples.** Throughout this section we suppose (S) holds. We first give some examples relating to Method 1.

EXAMPLE 4. Let $\hat{\theta}_0$ denote some consistent preliminary estimate of location $\theta$; and let $\hat{d}$ denote some estimate of scale or dispersion that satisfies $\hat{d} \to_p d$ as $n \to \infty$, for some constant $d$. Let $2\alpha_n$ denote the number of observations $X$ satisfying $|X - \hat{\theta}_0| \geqq k\hat{d}$, where $k$ is a fixed constant. Let $T_n$ denote the trimmed mean of (1) with $\alpha_n = n - \beta_n$. If (F3) holds and if $F$ is continuous at $-kd$, then we conclude from Theorem 1(i) that $n^{\frac{1}{2}}(T_n - \theta) \to_d M_\alpha$ with $\alpha = F(-kd)$.

We have forced (S2) to hold in a somewhat unnatural fashion; but the consequences of this are that the limiting distribution is independent of which preliminary consistent estimate $\hat{\theta}_0$ is used, and the final estimate behaves asymptotically like a trimmed mean.

For this case of forced symmetry, it would seem that overtrimming is better than undertrimming.

If we let $\hat{\theta}_0$ denote the median and $\hat{d}$ denote the median of the values $|X_{ni} - \hat{\theta}_0|$, then the breakdown point (see Hampel (1971)) of this randomly trimmed mean is $\frac{1}{2}$.

The more natural asymmetric adjustment is considered in Examples 9 and 10 below. Note that it does not lead to the two properties cited in paragraph two of this example.

Alternatively, we could replace the constant $k$ above by $\psi(\hat{K})$ where $\hat{K}$ is the sample kurtosis and $\psi$ is a suitable smooth function whose range is $(0, \infty)$. In this case $\alpha = F(-\psi(K)d)$ where $K$ is the true kurtosis. See also Hogg (1967). □

EXAMPLE 5. Let $H_n \equiv H_n(X_{n1}, \cdots, X_{nn})$ denote the ratio of the mean deviation to the standard deviation. Let $p_n \equiv \psi(H_n)$ where $\psi$ maps $[0, \infty)$ into $[0, \frac{1}{2})$. The idea is that $H_n$ tells us how heavy the tails of $F$ look; then $\psi$ calibrates this to a percentage that we will trim. The actual numbers we will trim are $\alpha_n = n - \beta_n = np_n$. If (F3) holds and if $F$ has a finite variance and $\psi$ is suitably smooth, then Theorem 1(i) gives $n^{\frac{1}{2}}(T_n - \theta) \to_d M_\alpha$ with $\alpha = \psi(H)$; here $H$ is the ratio of the true mean deviation to the true standard deviation.

Clearly, a multitude of other examples along this same line is possible. □

EXAMPLE 6. (Jaeckel). Let $V_n^2(\alpha)$ denote the Winsorized sample variance (see

(7)) about the ordinary trimmed mean $T_n(\alpha)$ of Example 1. Let $\hat{\alpha}$ minimize $V_n^2(\alpha)$ over a fixed range $0 < \alpha_0 \leqq \alpha \leqq \alpha_1 < \frac{1}{2}$, let $\alpha_n = n - \beta_n$ denote the greatest integer in $n\hat{\alpha}$, and denote the resulting trimmed mean by $T_n(\hat{\alpha})$. If $\sigma^2(\alpha)$ has a unique minimum in $[\alpha_0, \alpha_1]$ at $\alpha_{\min}$ and if $F$ has a strictly positive continuous density on an open interval containing $[g(\alpha_0), g(1-\alpha_0)]$, then Jaeckel (1971) show that (A1) holds. Hence Theorem 1(i) also yields Jaeckel's result that $n^{\frac{1}{2}}(T_n(\hat{\alpha}) - \theta) \to_d M_{\alpha_{\min}}$.

Jaeckel established certain large sample minimax properties for this adaptive estimate. However, its small sample performance in Andrews, *et al.* (1972) was not too impressive. □

EXAMPLE 7. (Johns). Let $0 < \alpha_1 < \cdots < \alpha_\kappa < \frac{1}{2}$. Let $\hat{c}_1, \cdots, \hat{c}_\kappa$ satisfy $\hat{c}_i \to_p c_i$ for $1 \leqq i \leqq \kappa$ and $\sum_1^\kappa \hat{c}_i = 1$. If (F1) holds at $\alpha_1, \cdots, \alpha_\kappa$, then

$$n^{\frac{1}{2}}(\textstyle\sum_1^\kappa \hat{c}_i T_n(\alpha_i) - \theta) \to_d \sum_1^\kappa c_i M_{\alpha_i} ,$$

Johns (1971) shows how the $c_i$'s can be chosen so that the variance of $\sum_1^\kappa c_i M_{\alpha_i}$ is uniformly close to the Cramér–Rao bound over a large class of rather smooth $F$. □

We now turn to examples of Method 2. We will consider the $T_n$ and $T_n^*$ of Theorem 2 for various choices of $\hat{\theta}_0$ and $\hat{B}$. We suppose throughout that (S) and (F4) hold, and all our choices for $\hat{\theta}_0$ and $\hat{B}$ will satisfy (2) and (3) respectively.

EXAMPLE 8. (Metrically symmetrizing about the trimmed mean with percentages fixed). Let $0 < \alpha < \frac{1}{2}$ be fixed. Let $\hat{\theta}_0$ denote the ordinary trimmed mean with $\alpha_n = n - \beta_n = [n\alpha]$; thus $n^{\frac{1}{2}}(\hat{\theta}_0 - \theta) \to_p M_\alpha$. Let $T_n$ ($T_n^*$) denote the metrically trimmed (Winsorized) mean about $\hat{\theta}_0$ in which the $2\alpha_n$ observations whose residuals about $\hat{\theta}_0$ are largest are trimmed (Winsorized). Then

$$n^{\frac{1}{2}}(T_n - \theta) \to_d \bar{M}_\alpha + [2Bf(B)/(1-2\alpha)]M_\alpha$$

and

$$n^{\frac{1}{2}}(T_n^* - \theta) \to_d M_\alpha .$$

(This $T_n^*$ is really the simplest special case of the $\tau_n$ of Example 3.) □

EXAMPLE 9. (Metrically symmetrizing about the median). We now add the assumption that $F$ has a strictly positive continuous derivative $f$ in a neighborhood of 0. Letting $\hat{\theta}_0$ denote the median, it is then easy to show that $n^{\frac{1}{2}}(\hat{\theta}_0 - \theta) \to_p -U(\frac{1}{2})/f(0)$. Suppose now that $\hat{B}$ satisfies (3), and let $\alpha = F(-B)$. Then

$$n^{\frac{1}{2}}(T_n - \theta) \to_d \bar{M}_\alpha - [2Bf(B)/(1-2\alpha)]U(\tfrac{1}{2})/f(0)$$

and

$$n^{\frac{1}{2}}(T_n^* - \theta) \to_d (1-2\alpha)M_\alpha - 2\alpha U(\tfrac{1}{2})/f(0) .$$

By way of illustration, we could let $\hat{B} = k\hat{d}$ for some constant $k$ and some estimate of dispersion $\hat{d}$ that satisfies $\hat{d} - d = O_p(n^{-\frac{1}{2}})$ for some constant $d$. In this case $\alpha = F(-kd)$. Compare this to Example 4.

Note that this approach is sufficient for treating the "one-step estimators" described on page 13 of Andrews, *et al.* (1972). □

EXAMPLE 10. (Metrically symmetrizing about trimmed and Winsorized means). Let $0 < \alpha_0 < \frac{1}{2}$ be a fixed initial adjustment percentage. We now add the assumption that (F4) holds at $B_0 \equiv -g(\alpha_0)$. We suppose $\hat{B}$ satisfies (3) for some $B$.

If $\hat{\theta}_0$ denotes the $\alpha_0$-trimmed mean, then

$$n^{\frac{1}{2}}(T_n - \theta) \to_d \bar{M} + [2Bf(B)/(1-2\alpha)]M_{\alpha_0}$$

and

$$n^{\frac{1}{2}}(T_n^* - \theta) \to_d (1-2\alpha)M_\alpha + 2\alpha M_{\alpha_0}\,.$$

(Compare the result for $T_n^*$ to Example 7.)

If $\hat{\theta}_0$ denotes the $\alpha_0$-Winsorized mean, then merely replace $M_{\alpha_0}$ by $M_{\alpha_0}^*$ in the two formulas above.

It seems useful to record that

$$\begin{aligned}\operatorname{Cov}[M_\alpha, M_{\alpha_0}] &= (1-2\bar{\alpha})\sigma^2(\bar{\alpha})/(1-2\underline{\alpha}) \\ &\quad + 2\underline{B}(\underline{\alpha}\bar{B} - \bar{\alpha}\underline{B} + \textstyle\int_{\underline{B}}^{\bar{B}} x\,dF(x))/((1-2\alpha)(1-2\alpha_0))\end{aligned}$$

where $\underline{\alpha} = \alpha \wedge \alpha_0$, $\bar{\alpha} = \alpha \vee \alpha_0$, $\underline{B} = B \wedge B_0$, $\bar{B} = B \vee B_0$ and $\sigma^2(\alpha)$ is defined in (5).

For $\hat{B} = k\hat{d}$ and $\alpha = F(-kd)$ this can again be compared to Example 4. □

EXAMPLE 11. (A Huber type of metrically symmetrized means). Let $H_\alpha$ ($H_\alpha^*$) denote the estimate with the property that if the $2[n\alpha]$ observations having the largest residuals about $H_\alpha$ ($H_\alpha^*$) are trimmed (Winsorized), then the resulting estimate is again $H_\alpha$ ($H_\alpha^*$). That $H_\alpha$ ($H_\alpha^*$) is well defined and satisfies (2) is shown in Huber (1967) (in Huber (1964)).

Thus from Theorem 2(i) we obtain

$$\begin{aligned}n^{\frac{1}{2}}(H_\alpha - \theta) &= D + c[D + c[D + c[\cdots] + \varepsilon_n] + \varepsilon_n] + \varepsilon_n \\ &= (D + \varepsilon_n)(1 + c + c^2 + \cdots) \\ &\to_d D/(1-c) \\ &= (1-2\alpha)\bar{M}_\alpha/[1 - 2\alpha - 2Bf(B)]\end{aligned}$$

where $D = \bar{M}_\alpha$ and $c = 2Bf(B)/(1-2\alpha)$.

Likewise, from Theorem 2(ii) we obtain

$$\begin{aligned}n^{\frac{1}{2}}(H_\alpha^* - \theta) &= D + 2\alpha[D + 2\alpha[D + 2\alpha[\cdots] + \varepsilon_n] + \varepsilon_n] + \varepsilon_n \\ &= (D + \varepsilon_n)[1 + 2\alpha + (2\alpha)^2 + \cdots] \\ &\to_d D/(1-2\alpha) = M_\alpha\end{aligned}$$

where $D = (1-2\alpha)M_\alpha$.

These limiting distributions agree with those established by Huber. The present approach seems interesting. □

**5. Remarks.**

1. $T_n(\alpha)$ seems to be generally regarded as being preferable to $T_n^*(\alpha)$ because of its greater efficiency for heavy-tailed $F$, because estimating its variance does not require estimating the troublesome $f(B)$, and because the simple estimate $V_n$

of $\sigma$ is available. These reasons for prefering trimming to Winsorizing carry over to the general situation of Method 1.

2. For Method 2 the situation is less clear; but seems to be reversed. In Examples 8—11 the metrically Winsorized estimates have simpler asymptotic forms than do the metrically trimmed estimates. With insight or luck in choosing $\hat{\theta}_0$, we may obtain a particularly simple asymptotic form; this was the case with the estimates $\tau_n$ of Examples 3 and 8.

3. The most promising estimates seem to be the Method 1 symmetrically trimmed means $T_n$ and the special Method 2 metrically Winsorized means $\tau_n$. Note that both of these classes of estimates behave asymptotically like a trimmed mean $M_\alpha$ in which $\alpha = \alpha(F)$ may be determined by the data.

4. Which is best, $T_n$ or $\tau_n$?

4a. The one whose studentization (see Corollaries 1 and 2) is the more robust in small samples. I do not know which is more robust; but on the basis of "balance" or "matching" of the numerator and denominator, I would conjecture $\tau_n$. (The results of Levene (1960) and Efron (1969) do not apply directly here, but they seem encouraging to $\tau_n$.)

4b. The one which is more powerful in small samples. Recall that $\tau_n$ is formed by metrically Winsorizing about $T_n$. For data sets that are rather symmetric, the estimates will be nearly the same; while for data sets that are heavily skewed, $\tau_n$ will lead to asymmetric adjustment percentages. For this reason, I conjecture $\tau_n$. (See also page 253 of Andrews, *et al.* (1972).)

4c. The one which generalize most readily to more complicated problems. Bickel (1971) has generalized $T_n(\alpha)$ to the linear model; the procedure leads to omitting observations with large residuals, and may thus destroy the balance or equal spacing of a design. The $\tau_n$ procedure causes observations to be modified, not omitted; this generalization is currently being worked on.

4d. $T_n$ is slightly easier to compute.

5. Some observations on Theorem 2.

5a. All estimators $\hat{B}$ of $B$ satisfying (3) lead to equivalent asymptotic results.

5b. To obtain asymptotic distributions of $T_n$ and $T_n{}^*$ we need to represent the limiting form of $n^{\frac{1}{2}}(\hat{\theta}_0 - \theta)$ in terms of the $U$ process. We did this for trimmed means, Winsorized means and medians in the examples. Let me note that for the Hodges–Lehmann estimate $\hat{\theta}_0 \equiv \text{median}\,\{(X_i + X_j)/2 \colon 1 \leqq i, j \leqq n\}$, the limiting form of $n^{\frac{1}{2}}(\hat{\theta}_0 - \theta)$ is $-\int_0^1 U(t)\,dt/\int F'\,dF$.

5c. $T_n{}^*$ behaves asymptotically like a weighted average of the preliminary estimate and an $\alpha$-trimmed mean with $\alpha$ determined by the data. This makes sense intuitively, provided the preliminary estimator is a good estimator for heavy-tailed $F$.

Suppose we use $\hat{\theta}_0$ as a preliminary estimator. If $F$ has heavy tails, then presumably our choice of $\hat{B}$ is such that $\alpha = F(-B)$ is "near $\frac{1}{2}$." Thus Theorem 2(ii) tells us that $T_n{}^*$ is practically the same as $\hat{\theta}_0$; and that is bad if $\hat{\theta}_0$ is geared to light tails.

6. If we are led to suspect "light-tailed" $F$ we might well throw away the middle order statistics and average the "trimmings"; see Hogg (1967). The representation of the limiting rv is $-\int_0^\alpha [U(t) + U(1-t)]\, dg/2\alpha$.

**6. The two-sample problem for general $F$.** Let $X_{m1} \leqq \cdots \leqq X_{mm}$ and $Y_{n1} \leqq \cdots \leqq Y_{nn}$ denote the order statistics of independent samples from $F$ and $F(\cdot - \theta)$ respectively. Let $\alpha_m$, $m - \beta_m$ and $\alpha_n$, $n - \beta_n$ denote the random adjustments for the two samples. We will denote trimmed and Winsorized means from the two samples by $T_m$, $T_m^*$ and $T_n$, $T_n^*$. Again, we consider two methods of making random adjustments.

*Method* 3. Our estimates will be called *equalized random means* if

$$(\alpha_m/m) - (\alpha_n/n) = o_p((mn/(m+n))^{-\frac{1}{2}}), \tag{10}$$

and if the $\beta$'s satisfy the analogous condition.

*Method* 4. Let $\alpha_m$ $(\alpha_n)$ and $m - \beta_m$ $(n - \beta_n)$ denote the number of observations less than $\hat{c}_X + \hat{A}$ $(\hat{c}_Y + \hat{A})$ and greater than $\hat{c}_X + \hat{B}$ $(\hat{c}_Y + \hat{B})$ respectively, where the preliminary estimates of location $\hat{c}_X \equiv \hat{c}_{X,m}$ and $\hat{c}_Y \equiv \hat{c}_{Y,n}$ and the rv's $\hat{A} \equiv \hat{A}_{X,Y,m,n}$ and $\hat{B} \equiv \hat{B}_{X,Y,m,n}$ satisfy

$$\hat{c}_X - c = O_p(n^{-\frac{1}{2}}), \qquad \hat{c}_Y - (c+\theta) = O_p(n^{-\frac{1}{2}}) \tag{11}$$

for some constant $c$ and

$$\hat{A} - A = O_p(n^{-\frac{1}{2}}), \qquad \hat{B} - B = O_p(n^{-\frac{1}{2}}) \tag{12}$$

for some constants $A$ and $B$. The resulting estimates will be called *metrically equalized random means.*

EXAMPLE 12. (The special random means $\tau_m$, $\tau_n$). Suppose $\bar{\alpha}_m$, $m - \bar{\beta}_m$ and $\bar{\alpha}_n$, $n - \bar{\beta}_n$ satisfy (A3) and (10); and let $T_m$ and $T_n$ denote trimmed means based on these adjustments. Winsorize the $\bar{\alpha}_m + \bar{\alpha}_n$ $(m - \bar{\beta}_m + n - \bar{\beta}_n)$ observations whose residuals $|X_i - T_m|$ and $|Y_j - T_n|$ are smallest (largest). Let $\alpha_m$, $m - \beta_m$, $\alpha_n$, $n - \beta_n$ be the new adjustment percentages; and let $\tau_m$, $\tau_n$ be the new Winsorized means. We define

$$\sigma^2(\alpha, \beta) = [\alpha A^2 + \textstyle\int_A^B x^2\, dF(x) + (1-\beta)B^2]/(\beta-\alpha)^2\,. \tag{13}$$ □

THEOREM 3. (*Asymptotic normality of equalized trimmed means*). *If* (10), (F1) *and* (A3) *hold in the location model, then*

$$(mn/(m+n))^{\frac{1}{2}}(T_n - T_m - \theta) \to_d N(0, \sigma^2(\alpha, \beta)) \qquad as \quad m \wedge n \to \infty\,.$$

PROOF. Let $U_X$ and $U_Y$ denote independent Brownian bridges associated, as in Section 2, with the two samples. Then $U \equiv (m/(m+n))^{\frac{1}{2}}U_Y - (n/(m+n))^{\frac{1}{2}}U_X$ is also a Brownian bridge. Using this and (10), the conclusion follows immediately from Lemma 1(a). □

THEOREM 4. (*Asymptotic normality of special metrically equalized Winsorized means*). *If* $\bar{\alpha}_m$, $\bar{\beta}_m$, $\bar{\alpha}_n$, $\bar{\beta}_n$ *satisfy* (10) *and* (A3) *and if* $F$ *satisfies* (F4), *then*

$$(mn/(m+n))^{\frac{1}{2}}(\tau_n - \tau_m - \theta) \to_d N(0, \sigma^2(\alpha, \beta)) \qquad as \quad m \wedge n \to \infty\,.$$

PROOF. We again use the $U$, $U_X$, $U_Y$ of the proof of Theorem 3. We use (9) (which is Lemma 2(a) in disguise) and then Theorem 3 to find

$$\begin{aligned}(mn/(m+n))^{\frac{1}{2}}(\tau_n - \tau_m - \theta) &= (m/(m+n))^{\frac{1}{2}}[-\textstyle\int_\alpha^\beta U_Y\,dg + (\alpha+1-\beta)n^{\frac{1}{2}}(T_n - \mu - \theta)] \\ &\quad - (n/(m+n))^{\frac{1}{2}}[-\textstyle\int_\alpha^\beta U_X\,dg + (\alpha+1-\beta)m^{\frac{1}{2}}(T_m - \mu)] \\ &= -\textstyle\int_\alpha^\beta U\,dg - (\alpha+1-\beta)\textstyle\int_\alpha^\beta U\,dg/(\beta-\alpha) \\ &= -\textstyle\int_\alpha^\beta U\,dg/(\beta-\alpha)\,.\end{aligned}$$

This proves the present theorem; which is the analog of a special case of Theorem 2.

The analog of Theorem 2 itself is

$$\begin{aligned}(mn/(m+n))^{\frac{1}{2}}&(T_n{}^* - T_m{}^* - \theta) \\ &= -\textstyle\int_\alpha^\beta U\,dg + (\alpha+1-\beta)(mn/(m+n))^{\frac{1}{2}}(\hat{c}_Y - \hat{c}_X - \theta) + o_p(1)\,;\end{aligned}\tag{14}$$

where we are now considering the general case of Method 4, and not the special case of Example 12.

Theorem 4 and (14) may be generalized as in the Remark at the end of Section 3. □

The natural analog of (7) to use in studentizing these estimates of $\theta$ seems to be

$$V_{m,n}^2 = [(m-1)V_m{}^2 + (n-1)V_n{}^2]/(m+n-2)\tag{15}$$

where $V_m{}^2$ and $V_n{}^2$ are obtained from (7). (Of course, we would introduce the notation $\nu_{m,n}$, $\nu_m$, $\nu_n$ when studentizing $\tau_n - \tau_m$.) The appropriate number of degrees of freedom seems to be $\beta_m - \alpha_m + \beta_n - \alpha_n - 2$.

REMARK. Note that both Methods 3 and 4 use the combined sample to determine the adjustments in the individual samples. If this is not done, then the random adjustment percentages typically contribute extra terms to the limiting rv. In particular, this would destroy the simplicity of Theorems 3 and 4.

## 7. Proofs of lemmas.

PROOF OF LEMMA 1. Now, a.s., none of the rv's $\xi_{ni}$ take values corresponding to discontinuities of $g$; thus

$$\textstyle\int_{[\xi_{n\alpha_n+1},\xi_{n\beta_n})} \psi(\Gamma_n)\,dg = \sum_{\alpha_n+1}^{\beta_n-1} \psi(i/n)[g(\xi_{ni+1}) - g(\xi_{ni})]$$

where $\psi(t) \equiv \int_t^1 ds$. When summed by parts this gives

$$\begin{aligned}S_n &\equiv n^{-1}\textstyle\sum_{\alpha_n+1}^{\beta_n} g(\xi_{ni}) \\ &= \psi(\alpha_n/n)g(\xi_{n\alpha_n+1}) + \textstyle\int_{\xi_{n\alpha_n+1}}^{\xi_{n\beta_n}} \psi(\Gamma_n)\,dg - \psi(\beta_n/n)g(\xi_{n\beta_n})\,.\end{aligned}$$

Integration by parts gives

$$\textstyle\int_\alpha^\beta g(t)\,dt = \psi(\alpha)g(\alpha) + \int_\alpha^\beta \psi\,dg - \psi(\beta)g(\beta)\,.$$

Thus

$$n^{\frac{1}{2}}[S_n - \textstyle\int_\alpha^\beta g(t)\,dt] = -\textstyle\int_{\xi_{n\alpha_n+1}}^{\xi_{n\beta_n}} U_n\,dg + \gamma_{n1} + \gamma_{n2}\,;$$

where

$$\gamma_{n1} = n^{\frac{1}{2}}[-\int_\alpha^{\xi_{n\alpha_n+1}} \psi\, dg + \psi(\alpha_n/n)g(\xi_{n\alpha_n+1}) - \psi(\alpha)g(\alpha)]$$
$$= -n^{\frac{1}{2}} \int_\alpha^{\xi_{n\alpha_n+1}} g(t)\, dt + g(\xi_{n\alpha_n+1})n^{\frac{1}{2}}(\xi_{n\alpha_n+1} - \alpha_n/n)$$
$$= -g(\alpha)n^{\frac{1}{2}}(\alpha_n/n - \alpha) + [g(\xi_{n\alpha_n+1}) - g(\alpha)]n^{\frac{1}{2}}(\xi_{n\alpha_n+1} - \alpha_n/n)$$
$$- n^{\frac{1}{2}} \int_\alpha^{\xi_{n\alpha_n+1}} [g(t) - g(\alpha)]\, dt\,, \qquad \text{and}$$

$$\gamma_{n2} = n^{\frac{1}{2}}[-\int_{\xi_{n\beta_n}}^\beta \psi\, dg - \psi(\beta_n/n)g(\xi_{n\beta_n}) + \psi(\beta)g(\beta)]$$
$$= -n^{\frac{1}{2}} \int_{\xi_{n\beta_n}}^\beta g(t) - g(\xi_{n\beta_n})n^{\frac{1}{2}}(\xi_{n\beta_n} - \beta_n/n)$$
$$= g(\beta)n^{\frac{1}{2}}(\beta_n/n - \beta) - [g(\xi_{n\beta_n}) - g(\beta)]n^{\frac{1}{2}}(\xi_{n\beta_n} - \beta_n/n)$$
$$+ n^{\frac{1}{2}} \int_\beta^{\xi_{n\beta_n}} [g(t) - g(\beta)]\, dt\,, \qquad \text{and}$$

$$-\int_{\xi_{n\alpha_n+1}}^{\xi_{n\beta_n}} U_n\, dg = -\int_\alpha^\beta U\, dg + \int_\alpha^\beta (U - U_n)\, dg + \int_\alpha^{\xi_{n\alpha_n+1}} U_n\, dg - \int_\beta^{\xi_{n\beta_n}} U_n\, dg\,.$$

In case (a), when (A3) holds we have

$$n^{\frac{1}{2}}|\xi_{n\beta_n} - \beta| \leqq n^{\frac{1}{2}}|\xi_{n\beta_n} - \beta_n/n| + n^{\frac{1}{2}}|\beta_n/n - \beta|$$
$$\leqq \rho(V_n, 0) + n^{\frac{1}{2}}|\beta_n/n - \beta| = O_p(1)\,;$$

so that from (F1) we have

$$n^{\frac{1}{2}}[S_n - \int_\alpha^\beta g(t)\, dt] = -\int_\alpha^\beta U\, dg - g(\alpha)n^{\frac{1}{2}}(\alpha_n/n - \alpha) + g(\beta)n^{\frac{1}{2}}(\beta_n/n - \beta) + \bar{\varepsilon}_n \tag{16}$$

where $\bar{\varepsilon}_n \to_p 0$. Note that (A2) and either (F2) or (F3) also imply (16).

In case (b), we use instead the middle expressions for $\gamma_{n1}$ and $\gamma_{n2}$ to get from (S) and (S1) that

$$\gamma_{n1} + \gamma_{n2} = n^{\frac{1}{2}}g(\xi_{n\alpha_n+1})(\xi_{n\alpha_n+1} - \alpha_n/n) - n^{\frac{1}{2}}g(1 - \xi_{n\beta_n})[(1 - \xi_{n\beta_n}) - \alpha_n/n]$$
$$+ n^{\frac{1}{2}} \int_{\xi_{n\alpha_n+1}}^{1-\xi_{n\beta_n}} g(t)\, dt + o_p(1)$$
$$= n^{\frac{1}{2}}(\alpha_n/n)[g(1 - \xi_{n\beta_n}) - g(\xi_{n\alpha_n+1})] - n^{\frac{1}{2}} \int_{\xi_{n\alpha_n+1}}^{1-\xi_{n\beta_n}} t\, dg(t) + o_p(1)$$
$$= -n^{\frac{1}{2}} \int_{\xi_{n\alpha_n+1}}^{1-\xi_{n\beta_n}} [t - \alpha_n/n]\, dg(t) + o_p(1)\,;$$

and thus by (A1), (F3) and (S1)

$$|\gamma_{n1} + \gamma_{n2}| \leqq \sup |t - \alpha_n/n| n^{\frac{1}{2}}|g(1 - \xi_{n\beta_n}) - g(\xi_{n\alpha_n+1})|$$
$$= o_p(1)O_p(1) = o_p(1)\,.$$

Thus (16) holds in case (b) also.

From (16) we get

$$n^{\frac{1}{2}}(T_n - \mu) = n^{\frac{1}{2}}\left[\frac{n}{\beta_n - \alpha_n}(S_n - \int_\alpha^\beta g(t)\, dt) + \int_\alpha^\beta g(t)\, dt \left(\frac{n}{\beta_n - \alpha_n} - \frac{1}{\beta - \alpha}\right)\right]$$
$$= \left(\frac{n}{\beta_n - \alpha_n}\right)[n^{\frac{1}{2}}(S_n - \int_\alpha^\beta g(t)\, dt) - \mu n^{\frac{1}{2}}(\beta_n/n - \beta) + \mu n^{\frac{1}{2}}(\alpha_n/n - \alpha)]$$
$$= -\frac{\int_\alpha^\beta U\, dg + (g(\alpha) - \mu)n^{\frac{1}{2}}(\alpha_n/n - \alpha) - (g(\beta) - \mu)n^{\frac{1}{2}}(\beta_n/n - \beta)}{\beta - \alpha}$$
$$+ \varepsilon_n$$

where $\varepsilon_n \to_p 0$. □

PROOF OF LEMMA 2. Now

$$n^{\frac{1}{2}}[(\alpha_n/n)g(\xi_{n\alpha_n+1}) - \alpha g(\alpha)]$$
$$= g(\alpha)n^{\frac{1}{2}}(\alpha_n/n - \alpha) + [g(\xi_{n\alpha_n+1}) - g(\alpha)]n^{\frac{1}{2}}(\alpha_n/n - \alpha)$$
$$+ \alpha n^{\frac{1}{2}}[g(\xi_{n\alpha_n+1}) - g(\alpha)]$$

and

$$n^{\frac{1}{2}}[((n - \beta_n)/n)g(\xi_{n\beta_n}) - (1 - \beta)g(\beta)]$$
$$= -g(\beta)n^{\frac{1}{2}}(\beta_n/n - \beta) - [g(\xi_{n\beta_n}) - g(\beta)]n^{\frac{1}{2}}(\beta_n/n - \beta)$$
$$+ (1 - \beta)n^{\frac{1}{2}}[g(\xi_{n\beta_n}) - g(\beta)] .$$

In either case (a) or case (b) we can combine these with (16) to get

$$n^{\frac{1}{2}}(T_n{}^* - \mu_n{}^*) = -\int_\alpha^\beta U\,dg + \alpha n^{\frac{1}{2}}[g(\xi_{n\alpha_n+1} - g(\alpha)]$$
$$+ (1 - \beta)n^{\frac{1}{2}}[g(\xi_{n\beta_n}) - g(\beta)] + o_p(1) ;$$

and finally, we note that

$$\alpha n^{\frac{1}{2}}[g(\xi_{n\alpha_n+1}) - g(\alpha)] = \alpha([g(\xi_{n\alpha_n+1}) - g(\alpha)]/[\xi_{n\alpha_n+1} - \alpha])$$
$$\times [V_n((\alpha_n + 1)/n) + n^{\frac{1}{2}}((\alpha_n + 1)/n - \alpha)]$$
$$= \alpha g'(\alpha)[-U(\alpha) + n^{\frac{1}{2}}(\alpha_n/n - \alpha)] + o_p(1)$$

and

$$(1 - \beta)n^{\frac{1}{2}}[g(\xi_{n\beta_n}) - g(\beta)] = (1 - \beta)([g(\xi_{n\beta_n}) - g(\beta)]/[\xi_{n\beta_n} - \beta])$$
$$\times [V_n(\beta_n/n) + n^{\frac{1}{2}}(\beta_n/n - \beta)]$$
$$= (1 - \beta)g'(\beta)[-U(\beta) + n^{\frac{1}{2}}(\beta_n/n - \beta)] + o_p(1) . \quad \square$$

LEMMA 3. *Let $K$ denote a non-decreasing left continuous function on* $(0, 1)$. *If* $\sigma^2 \equiv \int_0^1 \int_0^1 (s \wedge t - st)\, dK(s)\, dK(t) < \infty$, *then*

$$\sigma^2 = \int_0^1 K^2(t)\, dt - (\int_0^1 K(t)\, dt)^2 .$$

PROOF. Let $t_1, t_2, \cdots$ be an enumeration of the discontinuities of $K$, let $p_i = K(t_i + 0) - K(t_i)$ for $i \geqq 1$ and let

$$\Delta = \sum_i t_i(1 - t_i)p_i^2 = \sum_i p_i^2 \iint_{0<u\leqq t_i\leqq \nu<1} du\, d\nu .$$

From ($\nu$) on page 419 of Hewitt and Stromberg (1965) we find

$$K^2(\nu + 0) - K^2(u) = \int_{[u,\nu]} [K(t + 0) + K(t)]\, dK(t)$$
$$= 2 \int_{[u,\nu]} K(t)\, dK(t) + \int_{[u,\nu]} [K(t + 0) - K(t)]\, dK(t)$$
$$= 2 \int_{[u,\nu]} K\, dK + \sum_{t_i \in (u,\nu]} p_i^2 + [K^2(u + 0) - K^2(u)] .$$

Using the idea on page 978 of Chernoff and Savage (1958), we write

$$\sigma^2 = 2 \iint_{0<s<t<1} s(1 - t)\, dK(s)\, dK(t) + \Delta$$
$$= 2 \iiiint_{0<u\leqq s<t\leqq \nu<1} dK(s)\, dK(t)\, du\, d\nu + \Delta$$
$$= 2 \iiint_{0<u<t\leqq \nu<1} [K(t) - K(u)]\, dK(t)\, du\, d\nu + \Delta$$

$$= \iint_{0<u<\nu<1} \{[K^2(\nu + 0) - K^2(u + 0) - \textstyle\sum_{t_i \in (u,\nu]} p_i^2]$$
$$- 2K(u)[K(\nu + 0) - K(u + 0)]\}\, du\, d\nu + \Delta$$
$$= \iint_{0<u<\nu<1} \{[K(\nu) - K(u)]^2 - \textstyle\sum_{t_i \in [u,\nu]} p_i^2\}\, du\, d\nu + \Delta$$
$$= \tfrac{1}{2} \textstyle\int_0^1 \int_0^1 [K(\nu) - K(u)]^2\, du\, d\nu - \Delta + \Delta$$
$$= \textstyle\int_0^1 K^2(\nu)\, d\nu - (\int_0^1 K(\nu)\, d\nu)^2 .$$

Use Fubini's theorem to obtain the $-\Delta$ in the next to the last line. □

Suppose now that the true df $F_n$ is indexed by $n$. Let $g_n = F_n^{-1}$; and let $\mu_n$ and $\mu_n^*$ be as in Lemmas 1 and 2, but with $g$ replaced by $g_n$. Note that $g$ below is not assumed to be symmetric.

COROLLARY 3. *Let $\alpha_n$ and $\beta_n$ satisfy* (A4). *Then*

$$n^{\frac{1}{2}}(T_n - \mu_n) \to_d -\textstyle\int_\alpha^\beta U\, dg/(\beta - \alpha)$$

*provided the family of functions $g_n$ is uniformly equicontinuous in open neighborhoods about each of $\alpha$ and $\beta$, and provided $g_n$ converges weakly to some function $g$ on some open interval in* $[0, 1]$ *containing* $[\alpha, \beta]$. *(If $g_n = g$ for all $n$, we require only* (F1).) *The limiting* rv *has variance $\int_0^1 K^2(t)\, dt - (\int_0^1 K(t)\, dt)^2$ where $K(t)$ equals $A$, $g(t)$, $B$ for $0 < t < \alpha$, $\alpha < t < \beta$, $\beta < t < 1$.*

PROOF. After applying Lemma 1(a), we need only replace $\int_\alpha^\beta U\, dg_n$ by $\int_\alpha^\beta U\, dg$. But this can be done by the weak convergence of $g_n$ to $g$ and the continuity of the sample paths of $U$. The variance expression comes from Lemma 3. □

COROLLARY 4. *Let $\alpha_n$ and $\beta_n$ satisfy* (A4). *Then*

$$n^{\frac{1}{2}}(T_n^* - \mu_n^*) \to_d -[\alpha U(\alpha) g'(\alpha) + \textstyle\int_\alpha^\beta U\, dg + (1 - \beta) U(\beta) g'(\beta)]$$

*provided the family of functions $g_n'$ is uniformly equicontinuous in open neighborhoods about each of $\alpha$ and $\beta$, provided $g_n$ converges weakly to some function $g$ on some open interval in* $[0, 1]$ *containing* $[\alpha, \beta]$, *and provided $g_n'(\alpha) \to g'(\alpha)$ and $g_n'(\beta) \to g'(\beta)$ as $n \to \infty$. (If $g_n = g$ for all $n$, we require only* (F2).) *The limiting* rv *has variance $\int_0^1 K_*^2(t)\, dt - (\int_0^1 K_*(t)\, dt)^2$ where $K_*(t)$ equals $A - \alpha g'(\alpha)$, $g(t)$, $B + (1 - \beta) g'(\beta)$ for $0 < t < \alpha$, $\alpha < t < \beta$, $\beta < t < 1$.*

## REFERENCES


[1] ANDREWS, D. F., *et al.* (1972). *Robust Estimates of Location.* Princeton Univ. Press.

[2] BICKEL, P. (1965). On some robust estimates of location. *Ann. Math. Statist.* **36** 847–858.

[3] BICKEL, P. (1971). On some analogs to linear combinations of order statistics in the linear model. Unpublished technical report.

[4] CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29** 972–994.

[5] EFRON, B. (1969). Student's *t*-test under symmetry conditions. *J. Amer. Statist. Assoc.* **64** 1278–1302.

[6] HAMPEL, F. (1971). A general qualitative definition of robustness. *Ann. Math. Statist.* **42** 1887–1896.

[7] Hewitt, E. and Stromberg, K. (1965). *Real and Abstract Analysis*. Springer-Verlag, Berlin.

[8] Hogg, R. (1967). Some observations on robust estimation. *J. Amer. Statist. Assoc.* **62** 1179–1186.

[9] Huber, P. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35** 73–101.

[10] Huber, P. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *Proc. Fifth Berkeley Symp. Math. Statist. and Prob.* **1** 221–233.

[11] Huber, P. (1969). *Theorie de l'inference Statistique Robuste*. Les Presses de l'Université de Montréal.

[12] Huber, P. (1970). Studentizing robust estimates. *Nonparametric Techniques in Statistical Inference*, (M. Puri, ed.). Cambridge Univ. Press. 453–464.

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

[14] Johns, M. V. Jr. (1971). Nonparametric estimation of location. Unpublished technical report.

[15] Levene, H. (1960). Robust tests for equality of variances. *Contributions to Probability and Statistics*, (I. Olkin, *et al.*, eds.). Stanford Univ. Press. 278–292.

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

[17] Tukey, J. and McLaughlin, D. (1963). Less vulnerable confidence and significance procedures for location based on a single sample: Trimming/Winsorizing 1. *Sankhyā Ser. A* **25** 331–352.

Department of Mathematics
University of Washington
Seattle, Washington 98195