The Annals of Statistics
1975, Vol. 3, No. 4, 862–873

ESTIMATION OF SHIFT AND CENTER OF SYMMETRY BASED ON KOLMOGOROV–SMIRNOV STATISTICS[1]

BY P. V. RAO, EUGENE F. SCHUSTER[2] AND RAMON C. LITTELL

University of Florida, University of Texas at El Paso, and University of Florida

A point estimator and a set of confidence intervals based on the Kolmogorov-Smirnov statistic are proposed for the shift parameter in the two-sample problem. Asymptotic distibution of the etimator as well as asymptotic bounds for the lengths of the intervals are derived. The two-sample results are then adapted to the one-sample problem to define an estimator and a set of confidence intervals for the center of a symmetric population.

1. Introduction. Let $X_1, X_2, \cdots, X_m$ and $Y_1, Y_2, \cdots, Y_n$ be independent random samples from populations with continuous cumulative distribution functions (cdf's) $F(t)$ and $G(t) = F(t - \Delta_0)$ respectively. Let F_m and G_n denote the empirical cdf's of the X's and Y's respectively. The Kolmogorov-Smirnov test for the null hypothesis $H_0: \Delta_0 = a$, where a is a specified number is based on the statistic

$$D_{m,n}(a) = \sup_t |F_m(t) - G_n(t + a)| .$$

The test rejects H_0 at the level α if $D_{m,n}(a)$ is greater than or equal to the α-level critical value $\gamma_{m,n,\alpha}$. Since small values of $D_{m,n}(a)$ favor $H_0: \Delta_0 = a$, it seems reasonable to choose any value of a that minimizes $D_{m,n}(a)$ as an estimator for Δ_0. In Section 2 we will show that there exists an interval of values of a which minimize $D_{m,n}(a)$ and propose a unique value in this interval as an estimator, $\hat{\Delta}_{m,n}$, for Δ_0. Several properties of $\hat{\Delta}_{m,n}$, including an explicit computational formula are presented in Section 2.

Confidence intervals for Δ_0 based on $D_{m,n}(a)$ are also considered. It is shown that the $100(1 - \alpha)\%$ confidence set for Δ_0 given by $\{a: D_{m,n}(a) \leqq \gamma_{m,n,\alpha}\}$, is in fact an interval $(\hat{\Delta}_L, \hat{\Delta}_U)$, and an asymptotic upperbound to $(m + n)^{\frac{1}{2}}(\hat{\Delta}_U - \hat{\Delta}_L)$ is presented. This upperbound is used to compare these confidence intervals with the Lehmann intervals ([6]) based on the Wilcoxon statistic.

In Section 3 the results of the two-sample problem are used to define point estimators and confidence intervals in the corresponding one-sample problem of estimating the center of symmetry, θ_0, of a continuous symmetric distribution. It turns out that the resulting estimators and confidence intervals are based on

Received April 1973; revised December 1974.

[1] Florida Agricultural Experiment Station Journal Series No. 5865.

[2] This author's research was supported in part by The University of Texas at El Paso summer research grant 083-50-794-09.

Key words and phrases. Shift parameter, center of symmetry, Kolmogorov–Smirnov statistics, empirical distribution function, estimation, confidence interval, asymptotic distribution.

the Butler statistic ([2]),

$$D_n(a) = \sup_t |G_n(t) + G_n((2a - t)^-) - 1| .$$

Schuster and Narvarte [8] have proposed the center of the interval of all a which minimize $D_n(a)$ as an estimator for θ_0. This estimator is similar to the one proposed in Section 3. However, the results of Section 3 go beyond the Schuster-Narvarte paper in that the asymptotic lengths of the confidence intervals as well as the asymptotic distribution of the estimator are also presented there. In addition, Bickel's conjecture mentioned in the Schuster-Narvarte paper is also verified.

Section 4 contains proofs of the results stated in Sections 2 and 3.

2. The two-sample problem. Assume that the observations are ordered within each sample so that $X_1 < X_2 < \cdots < X_m$ and $Y_1 < Y_2 < \cdots < Y_n$. Let

$$\begin{aligned} D^+_{m,n}(a) &= \sup_t [G_n(t+a) - F_m(t)] \\ D^-_{m,n}(a) &= \sup_t [F_m(t) - G_n(t+a)] . \end{aligned} \tag{2.1}$$

Then it was shown by Rao and Littell [7] (also see Schuster and Narvarte [8]) that $D^+_{m,n}(a)$ ($D^-_{m,n}(a)$) is a left-continuous non-decreasing (right-continuous non-increasing) step-function of a taking jumps of $(mn)^{-1}$ at the points

$$\begin{aligned} \Delta_r^+ &= \min_{r \leqq k \leqq mn} \{Y_{[(k-1)/m]+1} - X_{[(k-r)/n]+1}\} \\ (\Delta_r^- &= \max_{r \leqq k \leqq mn} \{Y_{[(k-r)/m]+1} - X_{[(k-1)/n]+1}\}) \end{aligned} \tag{2.2}$$

for $r = 1, 2, \cdots, mn$, where $[x]$ denotes the greatest integer less than or equal to x. Note that if $m = n$, equations (2.2) greatly simplify, and $D^+_{n,n}(a)$ and $D^-_{n,n}(a)$ will have only n jump points. Now $\Delta^-_{mn} \leqq \cdots \leqq \Delta_1^-$, $\Delta_1^+ \leqq \cdots \leqq \Delta^+_{mn}$, $\Delta^-_{mn} < \Delta^+_{mn}$ and $\Delta_1^+ < \Delta_1^-$. Thus $D_{m,n}(a) = \max\{D^+_{m,n}(a), D^-_{m,n}(a)\}$ attains its minimum value of $(r_0 - 1)(mn)^{-1}$ for any a in the interval $[\Delta^-_{r_0}, \Delta^+_{r_0}]$, where $r_0 = \min\{r : \Delta_r^- \leqq \Delta_r^+\}$. As stated in Section 1, any value in this interval may be taken as an estimator for Δ_0. We propose a unique estimator $\hat{\Delta}_{m,n}$ defined by

$$\hat{\Delta}_{m,n} = \tfrac{1}{2}(\Delta^* + \Delta^{**}) , \tag{2.3}$$

where $\Delta^* = \inf A$, $\Delta^{**} = \sup A$ and $A = \{a : D^+_{m,n}(a) = D^-_{m,n}(a)\}$. It is easy to see that $[\Delta^*, \Delta^{**}]$ is a subset of $[\Delta^-_{r_0}, \Delta^+_{r_0}]$ and that

$$\Delta^* = \max\{\Delta^-_{r_0}, \Delta^+_{r_0-1}\} , \qquad \Delta^{**} = \min\{\Delta^-_{r_0-1}, \Delta^+_{r_0}\} . \tag{2.4}$$

The following two properties of $\hat{\Delta}_{m,n}$ are easily proved (see [7]):

(1) $\hat{\Delta}_{m,n}$ is translation invariant; that is, for any real number c,

$$\hat{\Delta}(X_1, \cdots, X_m; Y_1 + c, \cdots, Y_n + c) = \hat{\Delta}(X_1, \cdots, X_m; Y_1, \cdots, Y_n) + c .$$

(2) If $m = n$, then the probability distribution of $\hat{\Delta}_{m,n}$ is symmetric about Δ_0. Furthermore, if one notes that

$$\begin{aligned} \hat{\Delta}_{m,n}(2\Delta_0 - X_1, \cdots, 2\Delta_0 - X_m; 2\Delta_0 - Y_1, \cdots, 2\Delta_0 - Y_n) \\ = -\hat{\Delta}_{m,n}(X_1, \cdots, X_m; Y_1, \cdots, Y_n) \end{aligned}$$

and takes $\varepsilon_N = \sup_x [G(b_N + \Delta_0 + x) - F(x)]$ in the proof of Theorem 4 of [8], one can easily modify the proof there to show that

(3) If $N = \min(m, n)$ and $\nu_N = O(N^{\frac{1}{2}-\delta})$ for some $\delta > 0$, then $\nu_N(\hat{\Delta}_{m,n} - \Delta_0) \to 0$ as $N \to \infty$ with probability 1.

In Theorem 1 the asymptotic distribution of $\hat{\Delta}_{m,n}$ is expressed in terms of certain functionals of a Brownian bridge.

THEOREM 1. *Suppose F is absolutely continuous and $f = F'$ is uniformly continuous on $S = \{x : 0 < F(x) < 1\}$. If $N = m + n$ and $0 < \lambda < 1$, then as $N \to \infty$ and $nN^{-1} \to \lambda$,*

$$\Pr[N^{\frac{1}{2}}(\hat{\Delta}_{m,n} - \Delta_0) \leqq y] \quad \text{tends to}$$

$$\begin{aligned}\Pr[\sup\nolimits_{0<u<1}\{\beta(u) + y(\lambda(1-\lambda))^{\frac{1}{2}} f(F^{-1}(u))\} \\ + \inf\nolimits_{0<u<1}\{\beta(u) + y(\lambda(1-\lambda))^{\frac{1}{2}} f(F^{-1}(u))\} \geqq 0]\end{aligned} \tag{2.5}$$

where $\{\beta(u) : 0 \leqq u \leqq 1\}$ is a Brownian bridge and $F^{-1}(u) = \inf\{y : F(y) \geqq u\}$.

REMARK 1. In the special case when f is the uniform density on (0, 1) we have $f(F^{-1}(u)) = 1$, $0 < u < 1$, and the asymptotic distribution of $N^{\frac{1}{2}}(\hat{\Delta}_{m,n} - \Delta_0)$ is given by

$$\Pr[\sup\nolimits_{0<u<1} \beta(u) + \inf\nolimits_{0<u<1} \beta(u) \geqq -2y(\lambda(1-\lambda))^{\frac{1}{2}}],$$

an expression which depends upon the known joint distribution of $\sup \beta(u)$ and $\inf \beta(u)$ given in [4].

As stated in Section 1, the set $\{a : D_{m,n}(a) \leqq \gamma_{m,n,\alpha}\}$ will be a $100(1-\alpha)\%$ confidence set for Δ_0. From (2.1) and (2.2) it follows that this set is the interval $\{a : \hat{\Delta}_L \leqq a \leqq \hat{\Delta}_U\}$, where

$$\begin{aligned}\hat{\Delta}_L &= \max\nolimits_{k\in K}(Y^*_{k-mn\gamma_{m,n,\alpha}+1} - X_k^*) \\ \hat{\Delta}_U &= \min\nolimits_{k\in K}(Y_k^* - X^*_{k-mn\gamma_{m,n,\alpha}+1}), \\ K &= \{mn\gamma_{m,n,\alpha}, mn\gamma_{m,n,\alpha}+1, \cdots, mn\}, \\ X_i^* &= X_{[(i-1)/n]+1}, \\ Y_j^* &= Y_{[(j-1)/m]+1}.\end{aligned} \tag{2.6}$$

and

It must be noted that for a specified α, the interval may be empty (corresponding to the case $\hat{\Delta}_U < \hat{\Delta}_L$), but such intervals are obtained in a portion of events for which a true null hypothesis will be rejected.

Theorem 2 gives almost sure asymptotic upper bounds to the length $N^{\frac{1}{2}}(\hat{\Delta}_U - \hat{\Delta}_L)$.

THEOREM 2. *Let ζ be a point for which (1) $F(\zeta) = p$, $0 < p < 1$; (2) F is twice differentiable in a neighborhood of ζ, and (3) $f(\zeta) = F'(\zeta) > 0$, and f' is bounded in a neighborhood of ζ. Then as $N = m + n \to \infty$ and $nN^{-1} \to \lambda$, $0 < \lambda < 1$,*

$$\limsup N^{\frac{1}{2}}(\hat{\Delta}_U - \hat{\Delta}_L) \leqq \frac{2d_\alpha}{f(\zeta)(\lambda(1-\lambda))^{\frac{1}{2}}} \quad \text{w.p. } 1 \tag{2.7}$$

where $d_\alpha = \lim_{N\to\infty}(mn/N)^{\frac{1}{2}}\gamma_{m,n,\alpha}$. Further, if ζ_0 is a point and $\{\zeta_k\}$ is a sequence *of points such that* $\zeta_k \to \zeta_0$, $f(\zeta_k) \to f(\zeta_0)$, *and* (1), (2) *and* (3) *are satisfied at each* ζ_k, *then* $f(\zeta)$ *in the right-hand side of* (2.7) *may be replaced by* $f(\zeta_0)$.

REMARK 2. The best bound in (2.7) is obtained if ζ can be taken as the mode of (a unimodal) density f. For many standard densities this will be the case since even when the mode ζ will not satisfy (1), (2) and (3), it is often possible to find a sequence $\{\zeta_k\}$ converging to ζ and having the required properties. The double exponential density is an example of this situation.

Theorem 3 is a stronger version of a result by Lehmann [6] in that the convergence here is with probability one whereas Lehmann established convergence in probability. Sen and Ghosh [9], page 195, obtained the corresponding result with probability 1 for the one-sample case. The present result is in a specific setting to serve our immediate needs and the proof is therefore much shorter than the proof in [9]. A crucial step in both cases utilizes a generalization of a result by Bahadur which was proved by Sen and Ghosh [9], page 193. Let $(\tilde{\Delta}_L, \tilde{\Delta}_U)$ denote the $100(1-\alpha)\%$ Lehmann [6] confidence interval for Δ_0.

THEOREM 3. *Assume* F *has a bounded density* f *with a bounded derivative* f'. *Then as* $N \to \infty$, $nN^{-1} \to \lambda$, $0 < \lambda < 1$,

$$N^{\frac{1}{2}}(\tilde{\Delta}_U - \tilde{\Delta}_L) \to \frac{K_{\alpha/2}}{(3\lambda(1-\lambda))^{\frac{1}{2}} \int f^2(x)\,dx} \quad \text{w.p. } 1\,,$$

where K_p *is the* $(1-p)$*th quantile of the standard normal distribution.*

The comparison of confidence intervals based on the ratio of their asymptotic lengths was employed by Lehmann [6]. Since the remark after Lemma 3 in Section 4 shows that $N^{\frac{1}{2}}(\hat{\Delta}_U - \hat{\Delta}_L)$ does not converge to a constant, we consider in this paper the asymptotic upper bound for the ratio of lengths obtained from Theorems 2 and 3, namely,

$$\limsup_{N\to\infty} \frac{(\hat{\Delta}_U - \hat{\Delta}_L)}{(\tilde{\Delta}_U - \tilde{\Delta}_L)} \leqq \phi = \frac{(12)^{\frac{1}{2}} d_\alpha \int f^2(x)\,dx}{K_{\alpha/2} f(\xi)}$$

where ξ is taken as the mode of f. Following Lehmann's interpretation [6] page 1510, modified for the present situation, if the intervals $(\tilde{\Delta}_L, \tilde{\Delta}_U)$ are based on $m + n = N$ observations and the intervals $(\hat{\Delta}_L, \hat{\Delta}_U)$ are based on $m' + n' = N' = N\phi$ observations, then with probability one, the length $L' = \hat{\Delta}_U - \hat{\Delta}_L$ will be less than or equal to the length $L = \tilde{\Delta}_U - \tilde{\Delta}_L$ for large N.

TABLE 1
Values of ϕ *for selected choices of* α *and* f

α	Normal	Double Exponential	Cauchy
.10	1.8166	1.2846	.4089
.05	1.6996	1.2018	.3825
.025	1.6184	1.1443	.3642
.01	1.5500	1.0960	.3489
.005	1.5080	1.0664	.3394

Table 1 gives the values of ϕ for some selected densities f and critical levels α. From an inspection of Table 1, it is seen that the confidence interval proposed in this paper does almost as well as Lehmann's for double exponential populations and does extremely well for Cauchy populations. Thus we may conclude that the confidence interval proposed here may be preferable over Lehmann's if the underlying population has heavy tails.

3. One-sample problem. Let $Y_1, Y_2, \cdots, Y_n$ be the order statistics of a random sample of size n from a population with cdf $G(y) = F(y - \theta_0)$, where F is continuous and $F(x) = 1 - F(-x)$. The estimation procedures of Section 2 are readily adapted to the problem of estimating θ_0 by setting $X_1 = -Y_n$, $X_2 = -Y_{n-1}, \cdots, X_n = -Y_1$ and letting F_n and G_n be the empirical cdf's of the X's and Y's respectively. Then since the X's and Y's are samples from $F(x + \theta_0)$ and $F(x - \theta_0)$ respectively, and since $F_n(x) = 1 - G_n(-x^-)$, an estimator, $\hat{\theta}_n$, of θ_0 may be chosen to minimize

$$\begin{aligned} D_n(a) &= \sup_t |F_n(t) - G_n(t + 2a)| \\ &= \sup_t |1 - G_n(-t^-) - G_n(t + 2a)| \\ &= \sup_t |G_n(t) + G_n((2a - t)^-) - 1| . \end{aligned}$$

From Section 2 it follows that θ_r^+ and θ_r^- are given by

$$\begin{aligned} \theta_r^+ &= \tfrac{1}{2} \min_{r \leqq k \leqq n} (Y_k + Y_{n-k+r}) , \\ \theta_r^- &= \tfrac{1}{2} \max_{r \leqq k \leqq n} (Y_{k-r+1} + Y_{n-k+1}) , \end{aligned} \tag{3.1}$$

for $r = 1, 2, \cdots, n$, and that any point in $[\theta_{r_0}^-, \theta_{r_0}^+]$ where $r_0 = \min\{r : \theta_r^- \leqq \theta_r^+\}$ will minimize $D_n(a)$. As in the two-sample case, we propose estimation of θ_0 by $\hat{\theta}_n$ defined by

$$\hat{\theta}_n = \tfrac{1}{2}(\theta^* + \theta^{**}) , \tag{3.2}$$

where $\theta^* = \inf B$, $\theta^{**} = \sup B$ and $B = \{a : D_n^+(a) = D_n^-(a)\}$.

Schuster and Narvarte [8] have proposed $\frac{1}{2}(\theta_{r_0}^- + \theta_{r_0}^+)$ as an estimator for θ_0 and derived expressions equivalent to (3.1). They have also established some consistency and symmetry properties for their estimator (see Theorems 2, 3 and 4 of [8]). All these properties are satisfied by $\hat{\theta}_n$ of (3.2) as can be easily seen by following the proofs of the corresponding properties (1), (2) and (3) in the two-sample case. The asymptotic distribution of $\hat{\theta}_n$ can be obtained in a manner analogous to the proof of Theorem 1 and is given in Theorem 4.

THEOREM 4. *If F satisfies conditions of Theorem* 1 *and* $F(t) = 1 - F(-t)$, *then as* $n \to \infty$,

$$\begin{aligned} &\Pr[n^{\frac{1}{2}}(\hat{\theta}_n - \theta_0) \leqq y] \quad \textit{tends to} \\ &\Pr[\sup_{0<u\leqq\frac{1}{2}} \{W(u) + (2)^{\frac{1}{2}} y f(F^{-1}(u))\} \\ &\qquad + \inf_{0<u\leqq\frac{1}{2}} \{W(u) + (2)^{\frac{1}{2}} y f(F^{-1}(u))\} \geqq 0] , \end{aligned}$$

where $W(u)$ is a standardized Weiner process.

REMARK 3. As in the two-sample case, the uniform density leads to a simplification of the asymptotic distribution of $\hat{\theta}_n$. In fact, if f is the uniform density on $(\theta_0 - c, \theta_0 + c)$, then the asymptotic distribution of $n^{\frac{1}{2}}(\hat{\theta}_n - \theta_0)$ is the same as the distribution of $c/2[\sup_{0<u<1} W(u) + \inf_{0<u<1} W(u)]$, a quantity which can be evaluated by the known joint distribution of $\sup W(u)$ and $\inf W(u)$ (See [5], page 329.)

REMARK 4. (Bickel's conjecture). Suppose the random variable $\tilde{\theta}$ is chosen so as to minimize the function

$$\delta(y) = \sup_{0 \leqq u \leqq \frac{1}{2}} |W(u) + 2^{\frac{1}{2}} y f(F^{-1}(u))| .$$

Now

$$\delta^+(y) = \sup_{0 \leqq u \leqq \frac{1}{2}} (W(u) + 2^{\frac{1}{2}} y f(F^{-1}(u)))$$

and

$$\delta^-(y) = -\inf_{0 \leqq u \leqq \frac{1}{2}} (W(u) + 2^{\frac{1}{2}} y f(F^{-1}(u)))$$

are continuous monotone functions and $\delta(y) = \max\{\delta^+(y), \delta^-(y)\}$. Since

$$\Pr[\delta^+(y) = \delta^-(y)] = 0$$

by Lemma 1 in Section 4, it is easy to see that

$$\Pr[\tilde{\theta} \leqq y] = \Pr[\delta^+(y) \geqq \delta^-(y)] .$$

Thus the asymptotic distribution of $n^{\frac{1}{2}}(\hat{\theta} - \theta_0)$ is the same as that of $\tilde{\theta}$, and hence Theorem 4 indicates that Bickel's conjecture, given in Remark 3 of Schuster and Narvarte [8], holds for the minimax estimator $\hat{\theta}$. In the two-sample case the corresponding result would say that $N^{\frac{1}{2}}(\hat{\Delta}_{m,n} - \Delta_0)$ has the same asymptotic distribution as $\tilde{\Delta}$ where $\tilde{\Delta}$ is chosen so as to minimize (see Lemma 1)

$$\delta(y) = \sup_{0 \leqq u \leqq 1} |\beta(u) + y(\lambda(1 - \lambda))^{\frac{1}{2}} f(F^{-1}(u))| .$$

As in the two-sample case, confidence intervals for θ_0 may be based on $D_n(a)$. A $100(1 - \alpha)\%$ confidence interval is given by

$$(\hat{\theta}_L, \hat{\theta}_U) = \{a : D_n(a) \leqq \gamma^*_{n,\alpha}\} ,$$

where $\gamma^*_{n,\alpha}$ is the α-level critical value for the statistic $D_n(\theta)$ for testing $H_0 : \theta_0 = \theta$. The values of $\gamma^*_{n,\alpha}$ can be obtained from Chatterjee and Sen [3]. It is readily seen that

$$\hat{\theta}_L = \theta^-_{r_\alpha} , \qquad \hat{\theta}_U = \theta^+_{r_\alpha} , \tag{3.3}$$

where $r_\alpha = n^{\frac{1}{2}} \gamma^*_{n,\alpha}$.

Theorem 5 is the one-sample version of Theorem 2 and can be proved in a similar manner.

THEOREM 5. *Under conditions of Theorem* 2, *as* $n \to \infty$,

$$\limsup n^{\frac{1}{2}}(\hat{\theta}_U - \hat{\theta}_L) \leqq \frac{2^{\frac{1}{2}} d_\alpha^*}{f(\zeta)} \quad \text{w.p. } 1 ,$$

where $d_\alpha^* = \lim_{n\to\infty} n^{\frac{1}{2}} \gamma^*_{n,\alpha}$.

Now Sen and Ghosh ([9], page 195) have shown that if f satisfies the conditions of Theorem 3 and is symmetric, then the Lehmann interval $(\tilde{\theta}_L, \tilde{\theta}_U)$ satisfies

$$\lim_{n\to\infty} n^{\frac{1}{2}}(\tilde{\theta}_U - \tilde{\theta}_L) = \frac{K_{\alpha/2}}{3^{\frac{1}{2}} \int f^2(x)\, dx} \quad \text{w.p. } 1 .$$

Therefore,

$$\limsup_{n\to\infty} \frac{(\hat{\theta}_L - \hat{\theta}_U)}{(\tilde{\theta}_L - \tilde{\theta}_U)} \leqq \frac{6^{\frac{1}{2}} d_\alpha^* \int f^2(x)\, dx}{K_{\alpha/2} f(\xi)} \quad \text{w.p. } 1 , \tag{3.4}$$

where ξ may be taken as the mode of f, if the mode satisfies the conditions of Theorem 5. The right-hand side of (3.4) provides a measure of the relative efficiency of the two procedures.

TABLE 2

Values of the right-hand side of (3.4) *for selected* f

α	Normal	Double Exponential	Cauchy
.05	1.97	1.39	0.44
.01	1.89	1.34	0.42

From an inspection of Table 2 it appears that the confidence intervals of this paper may be preferable to the Lehmann confidence intervals if the underlying distribution has heavy tails.

4. Proof of theorems.

LEMMA 1. *Under the conditions of Theorem* 1,

$$\text{(i)} \quad \Pr[\sup_{0\leqq u\leqq 1}\{\beta(u) + y(\lambda(1-\lambda))^{\frac{1}{2}} f(F^{-1}(u))\} = -\inf_{0\leqq u\leqq 1}\{\beta(u) + y(\lambda(1-\lambda))^{\frac{1}{2}} f(F^{-1}(u))\}] = 0$$

and under the conditions of Theorem 4,

$$\text{(ii)} \quad \Pr[\sup_{0\leqq u\leqq \frac{1}{2}}\{W(u) + 2^{\frac{1}{2}} y f(F^{-1}(u))\} = -\inf_{0\leqq u\leqq \frac{1}{2}}\{W(u) + 2^{\frac{1}{2}} y f(F^{-1}(u))\}] = 0 .$$

PROOF. We shall first prove (ii). Let $h(u) = 2^{\frac{1}{2}} y f(F^{-1}(u))$, and let $\xi(u) = W(u) + h(u)$. Without loss of generality, assume $y > 0$. Now

$$\Pr[\sup_{0\leqq u\leqq \frac{1}{2}} \xi(u) = -\inf_{0\leqq u\leqq \frac{1}{2}} \xi(u)] \leqq \sum_{(r,s)} \Pr[\sup_{0\leqq u\leqq r} |\xi(u)| = \sup_{s\leqq u\leqq \frac{1}{2}} |\xi(u)|]$$

where the summation extends over all rational pairs (r, s) with $0 \leqq r < s \leqq \frac{1}{2}$. Hence it suffices to show that

$$\Pr[\sup_{0\leqq u\leqq r} \pm \xi(u) = \sup_{s\leqq u\leqq \frac{1}{2}} \pm \xi(u)] = 0$$

for each rational pair (r, s), with $0 \leqq r < s \leqq \frac{1}{2}$.

There are four cases to be considered. We shall show that

$$\Pr[\sup_{0\leqq u\leqq r} \xi(u) = \sup_{s\leqq u\leqq \frac{1}{2}} (-\xi(u))] = 0 .$$

The other cases can be handled in similar fashion. In this direction we observe that

$$\begin{aligned} Y &= \sup_{0\leq u\leq r}(\xi(u)) - \sup_{s\leq u\leq \frac{1}{2}}(-\xi(u)) \\ &= \{\sup_{0\leq u\leq r}(W(u)) + h(u) + W(r))\} \\ &\quad + \{-\sup_{s\leq u\leq \frac{1}{2}} - (W(u) - W(s) + h(u))\} + \{W(s) - W(r)\} \\ &= X_1 + X_2 + X_3 , \end{aligned}$$

where X_1, X_2, and X_3 are the first, second, and third terms, respectively, in the above expression. Now X_1 depends on $W(u)$ only to time r, and since X_2 and X_3 depend on independent increments after r, X_1, X_2, and X_3 are independent. Since X_3 is absolutely continuous, so is Y, and hence $\Pr[Y = 0] = 0$, which concludes the proof of (ii).

We now prove (i). In this case, let $h(u) = y(\lambda(1-\lambda))^{\frac{1}{2}}f(F^{-1}(u))$, let $\xi(u) = \beta(u) + h(u)$, and assume $y > 0$. The proof is more involved and does not depend on the proof of (ii). Take

$$\delta^+ = \sup_{0\leq u\leq 1}\xi(u) \qquad \text{and} \qquad \delta^- = \sup_{0\leq u\leq 1}(-\xi(u)) . \tag{4.0}$$

Since $\beta(1-u)$ is also a Brownian bridge we can without loss of generality assume $h(0) \geqq h(1)$. We will also assume the process β is of the form $\beta(u) = W(u) - uW(1)$ where W is a Brownian movement process. For each positive integer n we define functions d_n^+ and d_n^- on the reals by

$$d_n^+(z) = \sup_{0\leq u\leq 1-1/n}(\xi(u) - zu) \tag{4.1}$$

and

$$d_n^-(z) = \sup_{0\leq u\leq 1-1/n} -(\xi(u) - zu) . \tag{4.2}$$

We note that $-d_n^+$ and d_n^- are non-decreasing functions (of z) and with probability one the function

$$H_n(\cdot) = d_n^+(\cdot) - d_n^-(\cdot) \tag{4.3}$$

will have unique zeroes. The first mentioned property is easily checked. At least one of the two functions in (4.1) and (4.2) is strictly monotone at each z for which one of the two sups is attained in $(0, 1 - 1/n]$. If z_0 is a zero of H_n and both sups in (4.1) and (4.2) are attained at zero, then $h(0) = d_n^+(z_0) = d_n^-(z_0) = -h(0)$ which says $h(0) = 0$ and $W(u) + h(u) - z_0u = 0$ for all u in $(0, 1 - 1/n]$. Since this event has probability zero, $H_n(\cdot)$ will (w.p.1) have unique zeroes.

From $h(0) \geqq h(1)$ it follows that

$$\Pr[\delta^+ = \delta^-] \leqq \sum_{n=2}^{\infty} \Pr[d_n^+(W(1)) = d_n^-(W(1))] \tag{4.4}$$

(this inequality results in effect from the observation that (w.p.1) neither sup in (4.0) is attained only at $u = 1$). Hence our proof will be complete if we show

$$a = \Pr[d_n^+(W(1)) = d_n^-(W(1))] = \Pr[B_n(W)]$$

equals zero for each n. In this direction, for each rational r in $(1 - 1/n, 1]$ we define a Brownian motion V_r by reflecting W at r. Then $a = \Pr[B_n(V_r)]$ for each rational r in $(1 - 1/n, 1]$. Now

$$\begin{aligned} c_{n,r} &= B_n(V_r) \\ &= \{\sup_{0 \leq u \leq 1-1/n} (W(u) + h(u) - u(2W(r) - W(1))) \\ &\qquad = \sup_{0 \leq u \leq 1-1/n} -(W(u) + h(u) - u(2W(r) - W(1)))\} . \end{aligned}$$

If s is also a rational in $(1 - 1/n, 1]$ different from r then $W \in C_{n,r} \cap C_{n,s}$ implies $2W(r) - W(1)$ and $2W(s) - W(1)$ are both zeroes of H_n. Since zeroes of H_n are almost surely unique and $\{2W(r) - W(1) = 2W(s) - W(1)\}$ is an event of probability zero, we conclude $P_r[C_{n,r} \cap C_{n,s}] = 0$ for each pair of distinct rationals in $(1 - 1/n, 1]$, a must be zero and the proof is complete.

LEMMA 2. *Let $X_1, X_2, \cdots, X_m$ and $Y_1, Y_2, \cdots, Y_n$ be random samples (not necessarily independent) from a continuous population F. Let F_m and G_n denote their respective empirical* cdf's. *If $S_{m,n}^+(a) = \sup_{t \in S(a)} [G_n(t + a) - F_m(t)]$ and $S_{m,n}^-(a) = \sup_{t \in S(a)} [F_m(t) - G_n(t + a)]$, where $S(a) = \{x : x \in S, x + a \in S\}$, $a > 0$, and $S = \{x : 0 < F(x) < 1\}$, then*

(i) $D_{m,n}^+(a) = S_{m,n}^+(a)$ (w.p. 1)
(ii) $D_{m,n}^-(a) = \max\{0, S_{m,n}^-(a)\}$ (w.p. 1)
(iii) $\Pr\{D_{m,n}^+(a) \geqq D_{m,n}^-(a)\} = \Pr\{S_{m,n}^+(a) \geqq S_{m,n}^-(a)\}$,

and

(iv) $\Pr\{D_{m,n}^+(a) > D_{m,n}^-(a)\} \geqq \Pr\{S_{m,n}^+(a) > S_{m,n}^-(a)\} - \Pr\{S_{m,n}^+(a) = 0\}$.

PROOF. Let $S = (c, d)$. Then $t \in S(a)$ implies that $c < t < d - a$. If $t \leqq c$, then $\sup_{t \leqq c}\{G_n(t + a) - F_m(t)\} = G_n(c + a)$ and $\sup_{t \geqq d-a}\{G_n(t + a) - F_m(t)\} = 1 - F_m(d - a)$ (w.p.1). Thus (i) follows since

$$\lim_{t \to c^+} [G_n(t + a) - F_m(t)] = G_n(c + a) \quad \text{(w.p. 1)}$$

and

$$\lim_{t \to (d-a)^-} [G_n(t + a) - F_m(t)] = 1 - F_m((d - a)^-) \quad \text{(w.p. 1)} .$$

To establish (ii) note that (w.p. 1)

$$\sup\{F_m(t) - G_n(t + a) : t \leqq c \ \text{ or } \ t \geqq d - a\} = 0 .$$

(iii) follows from (i) and (ii). (iv) is an easy consequence of the fact that $D_{m,n}^+(a) = D_{m,n}^-(a) = 0$ implies that $S_{m,n}^+(a) = 0$ and $D_{m,n}^+(a) = D_{m,n}^-(a) > 0$ implies that $D_{m,n}^-(a) = S_{m,n}^-(a)$ and hence that $S_{m,n}^-(a) = S_{m,n}^+(a)$.

PROOF OF THEOREM 1. We shall prove the theorem for $y > 0$. The case $y \leqq 0$ is similar. Since $\hat{\Delta}_{m,n}$ is translation invariant we may take $\Delta_0 = 0$. From (2.3) we get $\Delta^* \leqq \hat{\Delta}_{m,n} \leqq \Delta^{**}$ and (2.4) implies that $\hat{\Delta}_{m,n}$ is a continuous random variable. Also,

$$(\Delta^{**} < y/N^{\frac{1}{2}}) \Rightarrow (D_{m,n}^+(y/N^{\frac{1}{2}}) > D_{m,n}^-(y/N^{\frac{1}{2}})) ,$$

and

$$(D_{m,n}^+(y/N^{\frac{1}{2}}) \geqq D_{m,n}^-(y/N^{\frac{1}{2}})) \Rightarrow (\Delta^* \leqq y/N^{\frac{1}{2}}) .$$

Thus

$$\Pr[D^+_{m,n}(y/N^{\frac{1}{2}}) > D^-_{m,n}(y/N^{\frac{1}{2}})] \leqq \Pr[\hat{\Delta}_{m,n} \leqq (y/N^{\frac{1}{2}})] \leqq \Pr[D^+_{m,n}(y/N^{\frac{1}{2}}) \geqq D^-_{m,n}(y/N^{\frac{1}{2}})] .$$

Hence from Lemma 2 it follows that

$$\Pr[S^+_{m,n}(y/N^{\frac{1}{2}}) > S^-_{m,n}(y/N^{\frac{1}{2}})] - \Pr[S^+_{m,n}(y/N^{\frac{1}{2}}) = 0] \leqq \Pr[\hat{\Delta}_{m,n} \leqq y/N^{\frac{1}{2}}) \leqq \Pr[S^+_{m,n}(y/N^{\frac{1}{2}}) \geqq S^-_{m,n}(y/N^{\frac{1}{2}})] .$$

Therefore, in view of Lemma 1, it suffices to show that (in some probability space) as $N \to \infty$ and $nN^{-1} \to \lambda$,

$$\text{(1)} \qquad \left(\frac{nm}{N}\right)^{\frac{1}{2}} S^+_{m,n}(y/N^{\frac{1}{2}}) = \sup_{t\in S}\{\beta(F(t)) + (\lambda(1-\lambda))^{\frac{1}{2}} yf(t)\} + o_p(1)$$

and

$$\text{(2)} \qquad \left(\frac{mn}{N}\right)^{\frac{1}{2}} S^-_{m,n}(y/N^{\frac{1}{2}}) = -\inf_{t\in S}\{\beta(F(t)) + (\lambda(1-\lambda))^{\frac{1}{2}} yf(t)\} + o_p(1) .$$

To establish (1) and (2), we follow Shorack's approach (see [10], Appendix) and use the representation $X_i = F^{-1}(\xi_i)$ and $Y_j = F^{-1}(\eta_j)$, where $\xi_1, \cdots, \xi_m$ and $\eta_1, \cdots, \eta_n$ are independent uniformly distributed random variables on (0, 1). If $F_m{}^*$ and $G_n{}^*$ are empirical cdf's of the ξ's and η's respectively, and $U_m(t) = m^{\frac{1}{2}}[F_m{}^*(t) - t]$, $V_n(t) = n^{\frac{1}{2}}[G_n{}^*(t) - t]$, then there exist independent Brownian bridges $\{U(t) : 0 \leqq t \leqq 1\}$ and $\{V(t) : 0 \leqq t \leqq 1\}$ on some probability space such that

$$\sup_{0<t<1} |U_m(t) - U(t)| \to 0 ,$$

and

$$\sup_{0<t<1} |V_n(t) - V(t)| \to 0 ,$$

where the convergence holds for every sample path of the U_m and V_n processes. Now let $\beta(t) = (1-\lambda)^{\frac{1}{2}} V(t) - \lambda^{\frac{1}{2}} U(t)$. Using the fact that

$$\lim_{N\to\infty} \sup_{t\in S(y/N^{\frac{1}{2}})} |N^{\frac{1}{2}}\{F(t + y/N^{\frac{1}{2}}) - F(t)\} - yf(t)| \to 0 ,$$

it is easy to see that, with probability one,

$$\begin{aligned}(N\lambda(1-\lambda))^{\frac{1}{2}} S^+_{m,n}(y/N^{\frac{1}{2}}) &= \sup_{t\in S(y/N^{\frac{1}{2}})} \{(1-\lambda)^{\frac{1}{2}} V(F(t + y/N^{\frac{1}{2}})) - \lambda^{\frac{1}{2}} U(F(t)) \\ &\qquad + (\lambda(1-\lambda))^{\frac{1}{2}} yf(t)\} + o(1) \\ &= \sup_{t\in S(y/N^{\frac{1}{2}})} \{\beta(F(t)) + (\lambda(1-\lambda))^{\frac{1}{2}} yf(t)\} + o(1) \\ &= \sup_{t\in S} \{\beta(F(t)) + (\lambda(1-\lambda))^{\frac{1}{2}} yf(t)\} + o(1) ,\end{aligned}$$

which establishes (1). In a similar fashion one can show that (2) holds.

Proof of Theorem 4. Since $\hat{\theta}_n$ is translation invariant we may take $\theta_0 = 0$. Furthermore, by letting $X_i = -Y_i$, $i = 1, 2, \cdots, n$, we obtain $X_1, \cdots, X_n$ and $Y_1, \cdots, Y_n$ as random samples from a symmetric population F. If $\hat{\Delta}_{n,n}$ denotes the estimator (2.3) for these two samples, we have $\hat{\Delta}_{n,n} = 2\hat{\theta}_n$, so that

$\Pr[n^{\frac{1}{2}}\hat{\theta}_n \leqq y] = \Pr[N^{\frac{1}{2}}\hat{\Delta}_{n,n} \leqq 2\cdot 2^{\frac{1}{2}}y]$, where $N = n + n$. Since Lemma 2 applies even for dependent samples, we can follow the proof of Theorem 1 with $\lambda = \frac{1}{2}$ and $N = 2n$. The U and V processes are now dependent with $U(t) = -V(1-t)$. Writing $W(t) = (1-\lambda)^{\frac{1}{2}}U(t) - \lambda^{\frac{1}{2}}V(t) = (V(t) + V(1-t))/2^{\frac{1}{2}}$, we see that $W(t)$ is a standardized Weiner process on $(0, \frac{1}{2})$. Hence $\Pr[N^{\frac{1}{2}}\hat{\Delta}_{n,n} \leqq 2\cdot 2^{\frac{1}{2}}y]$ tends to $\Pr[\sup_{t\in S}\{W(F(t)) + 2^{\frac{1}{2}}yf(t)\} + \inf_{t\in S}\{W(F(t)) + 2^{\frac{1}{2}}yf(t)\} \geqq 0]$. The desired conclusion follows by noting that $W(F(t)) + 2^{\frac{1}{2}}yf(t)$ is an even function.

We shall now state and prove Lemma 3 used in the proofs of Theorems 2 and 5.

LEMMA 3. *Let $N = m + n$ and assume $nN^{-1} \to \lambda$ as $N \to \infty$, $0 < \lambda < 1$. If ξ_1 and ξ_2 are two points satisfying conditions* (1), (2) *and* (3) *of Theorem* 2, *then*

$$N^{\frac{1}{2}}(\hat{\Delta}_U - \hat{\Delta}_L) \leqq N^{\frac{1}{2}}\left[\gamma_{m,n,\alpha}\left(\frac{1}{f(\xi_1)} + \frac{1}{f(\xi_2)}\right)\right] + Z_{m,n}(\xi_1, \xi_2) + R_N\,,$$

where

$$Z_{m,n}(\xi_1, \xi_2) = N^{\frac{1}{2}}\left[\frac{F_m(\xi_1) - G_n(\xi_1)}{f(\xi_1)} + \frac{G_n(\xi_2) - F_m(\xi_2)}{f(\xi_2)}\right] \quad \textit{and} \quad R_N \to 0 \quad \text{w.p. } 1\,.$$

PROOF. Assume without loss of generality that $\Delta_0 = 0$, and let $p_1 = F(\xi_1)$ and $p_2 = F(\xi_2)$. For sufficiently large m and n, $\gamma_{m,n,\alpha} \leqq \min(p_1, p_2)$. Thus from (2.6),

$$\hat{\Delta}_L \geqq Y^*_{[mnp_1]-mn\gamma_{m,n,\alpha}+1} - X^*_{[mnp_1]}\,, \qquad Y^*_{[mnp_1]-mn\gamma_{m,n,\alpha}+1} \geqq Y_{[n(p_1-\gamma_{m,n,\alpha})]}\,,$$

and

$$X^*_{[mnp_1]} \leqq X_{[mp_1]+1}\,.$$

Hence

$$\hat{\Delta}_L \geqq Y_{[n(p_1-\gamma_{m,n,\alpha})]} - X_{[mp_1]+1}\,.$$

Similarly,

$$\hat{\Delta}_U \leqq Y_{[np_2]+1} - X_{[m(p_2-\gamma_{m,n,\alpha})]}\,,$$

so that

$$\hat{\Delta}_U - \hat{\Delta}_L \leqq (Y_{[np_2]+1} - X_{[m(p_2-\gamma_{m,n,\alpha})]}) + (X_{[mp_1]+1} - Y_{[n(p_1-\gamma_{m,n,\alpha})]})\,.$$

Thus Bahadur's quantile representation ([1], Lemma 3) yields

$$\begin{aligned} N^{\frac{1}{2}}(\hat{\Delta}_U - \hat{\Delta}_L) \leqq{}& N^{\frac{1}{2}}\left[\left(\xi_1 + \frac{p_1 - G_n(\xi_1)}{f(\xi_1)}\right) - \left(\xi_1 + \frac{p_1 - \gamma_{m,n,\alpha} - F_m(\xi_1)}{f(\xi_1)}\right)\right] \\ &+ N^{\frac{1}{2}}\left[\left(\xi_2 + \frac{p_2 - F_m(\xi_2)}{f(\xi_2)}\right) - \left(\xi_2 + \frac{p_2 - \gamma_{m,n,\alpha} - G_n(\xi_2)}{f(\xi_2)}\right)\right] \\ &+ O(N^{-\frac{1}{4}}\log N) \quad \text{w.p. } 1\,, \end{aligned}$$

which proves the lemma.

Note that since with $\xi_1 \neq \xi_2$, $Z_{m,n}(\xi_1, \xi_2)$ is asymptotically normal with mean zero and variance of order 1, it follows that $N^{\frac{1}{2}}(\hat{\Delta}_U - \hat{\Delta}_L)$ will not converge (in probability) to a constant.

PROOF OF THEOREM 2. Follows by noting that $\xi_1 = \xi_2$ implies $Z_{m,n}(\xi_1, \xi_2) = 0$ in Lemma 3.

Theorem 5 can be proved in a manner analogous to the proof of Theorem 2.

Outline of the proof of Theorem 3. Set up the following quantities:

(1) $S_{mn}(a) =$ sum of ranks of X's in pooled sample of X's and $(Y - a)$'s,

(2) $H_{m,n}(x, a) = N^{-1}[mF_m(x) + nG_n(x + a)]$;

(3) $T_{mn}(a) = \int H_{mn}(x, a)\, dF_m(x)$;

(4) $W_{mn}(a) = N^{\frac{1}{2}}[T_{mn}(a) - T_{mn}(0) - a(1 - \lambda) \int f^2(x)\, dx]$.

The first step is to prove the asymptotic linearity of $T_{mn}(a)$, i.e., to prove $\sup_{|a|<N^{\frac{1}{2}} \log N} W_{mn}(a) \to 0$ with probability one. This follows in a straightforward manner by utilizing the generalized Bahadur result.

The desired conclusion can then be established by observing $S_{mn}(a) = m(m + n)T_{mn}(a)$ and employing the asymptotic normality of $S_{mn}(0)$.

Acknowlegment. The authors wish to thank Professors P. J. Bickel and J. Shuster for their helpful comments.

REFERENCES

[1] BAHADUR, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37** 577–580.

[2] BUTLER, C. C. (1969). A test for symmetry using the sample distribution function. *Ann. Math. Statist.* **40** 2209–2210.

[3] CHATTERJEE, S. K. and SEN, P. K. (1973). On Kolmogorov-Smirnov-type tests for symmetry. *Ann. Inst. Statist. Math.* **25** 287–300.

[4] DOOB, J. L. (1949). Heuristic approach to Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **20** 393–403.

[5] FELLER, WILLIAM (1966). *An Introduction to Probability Theory and Its Applications*, **2.** Wiley, New York.

[6] LEMANN, E. L. (1963). Nonparametric confidence intervals for a shift parameter. *Ann. Math. Statist.* **34** 1507–1512.

[7] RAO, P. V. and LITTELL, R. C. (1971–72). Univ. of Florida Department of Statistics Technical Reports No. 30–30A, Gainesville.

[7] SCHUSTER, E. F. and NARVARTE, J. A. (1973). A new nonparametric estimator of the center of a symmetric distribution. *Ann. Statist.* **1** 1906–2104.

[9] SEN, P. K. and GHOSH, M. (1971). On bounded length sequential confidence intervals based on one-sample rank-order statistics. *Ann. Math. Statist.* **42** 189–203.

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

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TEXAS
EL PASO, TEXAS

DEPARTMENT OF STATISTICS
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA 32611