GENERALIZATIONS OF THEOREMS OF CHERNOFF AND SAVAGE ON THE ASYMPTOTIC NORMALITY OF TEST STATISTICS

Z. GOVINDARAJULU, L. LeCAM, and M. RAGHAVACHARI
University of California, Berkeley

1. Introduction

The purpose of the present paper is to generalize the results obtained by Chernoff and Savage [5] on the asymptotic normality of a large class of two-sample nonparametric test statistics.

The assumptions made in [5] involve a certain function J which is assumed to possess two derivatives satisfying boundedness restrictions. However, certain test statistics, for instance those proposed by Ansari and Bradley [1] and Siegel and Tukey [15], do not satisfy the regularity conditions imposed by Chernoff and Savage. In particular, the *first* derivative of the appropriate function J fails to exist at certain points, so that the arguments of Chernoff and Savage are no longer directly applicable.

It will be shown here that the basic asymptotic normality result of [5] remains valid without any assumptions whatsoever or the existence of second derivatives. The assumption of existence of the first derivative is replaced by an assumption of absolute continuity. It should be noted that even this assumption is somewhat too stringent if one is willing to impose restrictions on the couple (F, G). However, the discussion of such possibilities remains beyond the purview of the present paper.

Section 2 of the paper gives a number of definitions which will be used throughout. Section 3 summarizes some properties of the set of functions J which will be used later. The main results are a lemma (lemma 2) on uniform square integrability and a continuity theorem (lemma 3) for the variances of the normal approximations to the distributions of the Chernoff-Savage statistics. Section 4 gives an account of convergence properties of empirical cumulative distributions and of their inverse functions.

The tails of the Chernoff-Savage statistics are bounded in section 5, and the main asymptotic normality theorem appears in section 6. Finally, natural extensions to the c-sample situation are provided in section 7.

Prepared with the partial support of the Mathematics Division of the Air Force Office of Scientific Research under Grant No. AF-AFOSR-741-65 and with the partial support of the U. S. Army Research Office (Durham) Grant DA-31-124-AROD-548.

2. Standing assumptions and notations

Let $X_1, X_2, \cdots, X_m$ and $Y_1, Y_2, \cdots, Y_n$ be random samples of sizes m and n drawn from populations with cumulative distribution functions (= c.d.f.) F and G respectively. Let $N = m + n$ and let $N\lambda_N = m$. *It will be assumed throughout that there is a* $\lambda_0 > 0$ *such that* $0 < \lambda_0 \leq \lambda_N \leq 1 - \lambda_0$ *and that the distribution functions* F *and* G *have no common discontinuities.*

The function $H = \lambda_N F + (1 - \lambda_N)G$ will be called the combined population cumulative distribution function. Let F_m and G_n be the empirical c.d.f.'s of the X's and Y's respectively. The function $H_N = \lambda_N F_m + (1 - \lambda_N)G_n$ is called the combined empirical c.d.f. It will be assumed that (F, G) and λ_N vary with m and n. However, to avoid an excess of indices the notation suppresses this fact. Another reason for this simplified notation is that the following theorems are 'uniform' and are valid whether the distributions are constant, tend to a limit, or vary rather arbitrarily with N.

Define

$$Z_{N,i} = \begin{cases} 1 & \text{if the } i\text{-th smallest in the combined sample is an } X, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

Then, we will be concerned with statistics of the form

$$mT_N = \sum_{i=1}^{N} E_{N,i} Z_{N,i}, \tag{2.2}$$

where the $E_{N,i}$ are given constants. Many statistics occurring in nonparametric statistical inference can be reduced to the form (2.2). For examples the reader is referred to Chernoff and Savage [5]. We will, as Chernoff and Savage did, use the following representation:

$$T_N = \int_{-\infty}^{\infty} J_N\left(\frac{N}{N+1} H_N\right) dF_m(x). \tag{2.3}$$

The representations (2.2) and (2.3) are equivalent when $E_{N,i} = J_N(i/(N+1))$. Although J_N need be defined at $1/(N+1), 2/(N+1), \cdots, N/(N+1)$, we can conveniently extend its domain of definition to $(0, 1)$ by letting J_N be constant on $(i/(N+1), (i+1)/(N+1))$, $(i = 0, 1, 2, \cdots, N)$. Our J_N is slightly different from that used by Chernoff and Savage [5]. In (2.3) they use $J_N(H_N)$. Consequently, their J_N need be defined at $1/N, 2/N, \cdots, N/N$, that is, in $(0, 1]$. Our main purpose in slightly changing the J_N function is to avoid asymmetry and eliminate the possibility that F_m gives mass at points where the argument of J_N is unity. The implication of this symmetry will be clear in the statements of the main theorems, in which one of the assumptions of [5] can be dispensed with. The problem of asymmetry had also been recognized by J. Pratt and I. R. Savage who informed one of the authors via personal communication.

It is easily verified that one could replace the assumption that F and G have no common discontinuities by the apparently stronger requirement that both F

and G be continuous. However, the more general case reduces immediately to the continuous one as we shall now show.

If F has a jump of size α at a point t, remove the point t from the real line and insert in its place a closed interval of length α. Distribute the probability mass α uniformly over this interval. The cumulative distribution G is kept constant over the inserted interval. Proceed similarly for the jumps of G. The new cumulative distributions F^* and G^* so obtained are continuous. For samples obtained from F^* and G^* the relative order relations between X's and Y's have the same probability distribution as if the samples were obtained from F and G.

If F and G had common discontinuities, ties would occur with positive probability. The definition of the variables $Z_{N,i}$ would no longer be complete. Thus, taking into account the possibility of "continuization" as performed above, the assumption of continuity of both F and G is equivalent to the assumption that ties between X's and Y's occur with probability equal to zero.

As a further reduction, let us show that there is in fact no loss of generality in assuming that the following assumption holds.

Assumption (A). *For each integer N the cumulative distributions (F, G) and the number λ_N are such that $H = \lambda_N F + (1 - \lambda_N) G$ is the cumulative $H(x) \equiv x$, $x \in [0, 1]$, of the uniform distribution and $0 < \lambda_0 \leq \lambda_N \leq 1 - \lambda_0$.*

To show this, note that if after removal of discontinuities, the function H^* remains constant over certain intervals, no observations will occur in these intervals. Thus, these intervals can be deleted from the line without affecting the order of the observations. This will leave us with a continuous strictly increasing cumulative distribution which can now be transformed to the uniform cumulative $H(x) \equiv x$ by a strictly increasing continuous transformation.

In view of this we shall assume throughout that assumption A holds and, if necessary, indicate the original distributions before transformation by $(\check{F}, \check{G})$ instead of (F, G).

When assumption A is satisfied the measure dF induced by F possesses a density φ with respect to the Lebesgue measure dH on $[0, 1]$. The inequality $0 \leq \varphi \leq \lambda_0^{-1}$ will play an important role in the sequel.

To describe a class of functions to which our asymptotic normality results will apply, it is convenient to introduce the following definition.

Definition 1. *A function f, $f \geq 1$ defined on the interval $(0, 1)$ will be said to belong to the class $\mathfrak{U}_1$ (respectively $\mathfrak{U}_2$) if it is integrable (respectively square integrable) for the Lebesgue measure and if in addition there is some $\alpha \in (0, 1)$ such that f is monotone decreasing in $(0, \alpha]$ and monotone increasing in $[\alpha, 1)$.*

Let b denote a constant $0 < b < \infty$. Let f_0, f, and g be three nonnegative functions defined on $(0, 1)$. Assume that f_0 is Lebesgue integrable, and that $f \in \mathfrak{U}_1$ and that $g \in \mathfrak{U}_2$.

Consider functions J defined by integrals of the type $J(x) = \int_{1/2}^{x} J'(\xi)\, d\xi$. We shall say that J belongs to the class $\mathcal{S}_0$ if $|J'| \leq fg$, that $J \in \mathcal{S}$ if $|J'| \leq f_0 + fg$ and that $J \in \mathcal{S}_1$ if $J' = J_1' + J_2'$ with $|J_2'| \leq fg$ and $\int |J_1'(x)|\, dx \leq b$.

One could also introduce functions J which differ from the integrals

$\int_{1/2}^{x} J'(\xi)\,d\xi$ by a constant. However, this will not change the difference studied below. Thus the consideration of functions of the type $a + \int_{1/2}^{x} J'(\xi)\,d\xi$ is left to the care of the reader.

In the sequel the product fg will play essentially the same role as the function $f(x)g(x) = K[x(1-x)]^{-1+\delta}[x(1-x)]^{-(1/2)+\delta} = K[x(1-x)]^{-(3/2)+2\delta}$ of Chernoff and Savage. For this special choice of product fg one can also prove the following result. Let $\xi_{N,k}$ be the k-th order statistic in a sample of size N from the uniform distribution on $[0, 1]$. For any function J let $\bar{J}_N$ be the function defined on $(0, 1)$ as follows. If $y = (N+1)^{-1}k$, $k = 1, 2, \cdots, N$, let

$$\bar{J}_N(y) = EJ(\xi_{N,k}) = \int J(x)\beta_N(x, k)\,dx \tag{2.4}$$

where

$$\beta_N(x, k) = \frac{\Gamma(N+1)}{\Gamma(k)\Gamma(N+1-k)} x^{k-1}(1-x)^{N-k} \tag{2.5}$$

is the density of $\xi_{N,k}$.

Complete the definition of $\bar{J}_N$ by interpolating linearly between successive values $\{k/(N+1), (k+1)/(N+1)\}$ and leaving $\bar{J}_N$ constant below $1/(N+1)$ and above $N/(N+1)$.

LEMMA 1. *Assume that there exists a constant K and a δ, $0 < \delta < \frac{1}{2}$, such that $|J'(x)| \le K[x(1-x)]^{-(3/2)+\delta}$. Then, there exists a constant K_1 and an N_0 such that $N \ge N_0$ implies*

$$|\bar{J}_N'(x)| \le K_1[x(1-x)]^{-(3/2)+\delta}. \tag{2.6}$$

Furthermore, if $\{J_\nu'\}$ is a sequence such that J_ν' converges to J' in Lebesgue measure and $|J_\nu'| \le K[x(1-x)]^{-(3/2)+\delta}$, for all ν, then $\bar{J}_{\nu,N}' - J_\nu'$ converges to zero in Lebesgue measure as $N \to \infty$ uniformly in the index ν.

PROOF. Decompose J' into its positive and negative parts and then separate each of these into two pieces, one of which vanishes on $(0, \frac{1}{2}]$ and the other on $(\frac{1}{2}, 1)$. If the results hold for each of these four parts separately, they will hold for J' itself. Owing to the symmetry of the situation, it will be sufficient to prove the result for a function J' such that $J' \le 0$ and $J'(x) = 0$ for $x > \frac{1}{2}$ with the additional restriction $|J'| \le x^{-(3/2)+\delta}$. Let $J(x) = \int_x^1 |J'(u)|\,du$. The slope of the function $\bar{J}_N$ between two successive points $k/(N+1)$ and $(k+1)/(N+1)$ is given by the expression

$$\begin{aligned} s_N(k) &= (N+1)\int_0^1 J(x)[\beta_N(x, k) - \beta_N(x, k+1)]\,dx, \\ &= (N+1)\int_0^1 \left[J(x) - J\left(\frac{k}{N}\right)\right][\beta_N(x, k) - \beta_N(x, k+1)]\,dx. \end{aligned} \tag{2.7}$$

In this expression the integrand is nonnegative, since J is decreasing. This implies in particular that the slope $s_N(k)$ is smaller than the slope obtainable from the function $J'(x) = -x^{-(3/2)+\delta}$. Thus

$$s_N(k) \le \frac{2(N+1)}{1-2\delta}\int_0^1 x^{-(1/2)+\delta}[\beta_N(x, k) - \beta_N(x, k+1)]\,du. \tag{2.8}$$

This last expression is easily expressible in terms of gamma functions and the first result follows by direct computation for $k = 1$ and by application of Stirling's formula for $k > 1$.

To prove the second result note that the slope $s_N(k)$ can also be written in the form

$$s_N(k) = \int \frac{N[J(x) - J(k/N)]}{k - Nx} B_N(x, k)\, dx, \tag{2.9}$$

with $B_N(x, k)$ equal to the probability density

$$B_N(x, k) = \frac{N + 1}{N} \frac{|k - Nx|^2}{k(1 - x)} \beta_N(x, k). \tag{2.10}$$

For every $\epsilon > 0$ and every $y \in (0\ 1)$, if $(k/N) \to y$, then

$$\int_{|x-y| \geq \epsilon} \frac{N[J(x) - J(k/N)]}{k - Nx} B_N(x, k)\, dx \leq \int_{|x-y| \geq \epsilon} \frac{[x^{-(1/2)+\delta} - y^{-(1/2)+\delta}]}{y - x} B_N(x, k)\, dx. \tag{2.11}$$

This quantity tends to zero as $N \to \infty$. Thus it is sufficient to consider the behavior of the integral taken for $|x - y| < \epsilon$. For this purpose note first that when y is fixed, $y \in (0, 1)$ and $0 < y - \epsilon$, then the ratio $N[J(x) - J(y)][y - x]^{-1}$ remains bounded, independently of the choice of J, in the interval $[y - \epsilon, y + \epsilon]$. Therefore, taking for k_N the integer part of $(N + 1)y$, one can select a number $c < \infty$ and an N_0 such that

$$\int_{\sqrt{N}|x-y| \geq c} \frac{N[J(x) - J(k_N/N)]}{k_N - Nx} B_N(x, k_N)\, dx < \epsilon \tag{2.12}$$

for every J and every $N \geq N_0$.

Suppose then that the sequence $\{J_\nu'\}$ converges in measure to a limit J'. Taking a subsequence, if necessary, one can assume that $J_\nu' \to J'$ almost everywhere. In this case, for every $\alpha > 0$ there exist a compact subset S of the interval $(0, 1)$ such that S^c has a Lebesgue measure inferior to α and such that the J_ν' are continuous when restricted to S and such that $J_\nu'(x)$ converges to $J'(x)$ uniformly for $x \in S$. Suppose that y is a point of density of the set S and consider the integrals

$$I_{N,\nu} = \int \left[\frac{N[J_\nu(x) - J_\nu(k_N/N)]}{k_N - Nx} - J_\nu'(y)\right] B_N(x, k_N)\, dx, \tag{2.13}$$

taken over the set $S_N = \{x\colon x \in S \text{ and } \sqrt{N}|x - y| \leq c\}$. A simple change of variable $x = y + \xi/\sqrt{N}$ will show immediately that $I_{N,\nu}$ converges to zero uniformly in ν. Furthermore, an analogous integral taken over the set $\{x\colon x \in S^c$ and $\sqrt{N}|x - y| \leq c\}$ must tend to zero, since the point y is assumed to be a point of density of S. Taking into account the fact that almost all points of S are points of density, the result follows.

One could also apply the same argument to functions J_N^* obtained by the formula

$$J_N^*(\xi) = \int J_N(x)\beta_N[x, (N+1)\xi]\, dx, \tag{2.14}$$

for all values of ξ such that $1 \le (N+1)\xi \le N$.

3. Properties of functions which belong to $\mathcal{S}$

In this section we shall assume that the functions f_0, f and g are fixed and derive certain boundedness and integrability properties for the elements of the corresponding set $\mathcal{S}$ of functions.

LEMMA 2. *There is a number b_0 such that* $\sup \{\int J^2(u)\, du; J \in \mathcal{S}\} < b_0$. *Furthermore, for every $\epsilon > 0$ there is a number b such that*

$$\int_{|J(u)|>b} J^2(u)\, du < \epsilon \tag{3.1}$$

for every $J \in \mathcal{S}$.

PROOF. If $J' \in \mathcal{S}$, so are its positive and negative parts. Thus, it is sufficient to prove the result assuming $J' \ge 0$. In addition, the part $J_1' = \min[f_0, J']$ contributes a bounded term to the indefinite integral J. Therefore, it is sufficient to prove the lemma assuming $0 \le J' \le fg$. Take α so small that both f and g are monotone decreasing in $(0, \alpha]$. For every $\xi \in (0, \alpha]$ one can write $\xi f(\xi) \le \int_0^\xi f(u)\, du$ and $\xi g^2(\xi) \le \int_0^\xi g^2(u)\, du$. Let $c^2(\alpha)$ be the number

$$c^2(\alpha) = \max\left\{\int_0^\alpha g^2(u)\, du, \left[\int_0^\alpha f(u)\, du\right]^2\right\}. \tag{3.2}$$

Let $\varphi(u) = \int_u^\alpha J'(\xi)\, d\xi$. One can write

$$\begin{aligned} \int_0^\alpha \varphi^2(u)\, du &\le \int_0^\alpha \int_0^\alpha f(\xi)g(\xi)f(y)g(y) \min(\xi, y)\, d\xi\, dy \\ &= 2\int_0^\alpha \left\{\int_0^y f(\xi)g(\xi)\xi\, d\xi\right\} f(y)g(y)\, dy \\ &\le 2\int_0^\alpha c^2(\alpha)\left\{\int_0^y \frac{1}{\sqrt{\xi}}\, d\xi\right\} f(y)g(y)\, dy \\ &\le 4c^3(\alpha)\int_0^\alpha f(y)\, dy \le 4c^4(\alpha). \end{aligned} \tag{3.3}$$

A similar argument applies to the interval $[1-\alpha, 1)$ for α sufficiently small. If a is the maximum of $f(x)g(x)$ for $x \in [\alpha, 1-\alpha]$, the term $\int_\alpha^{1-\alpha} J'(\xi)\, d\xi$ remains bounded by $\int f_0(\xi)\, d\xi + a$. Hence the result.

Assuming as usual that H is the cumulative distribution of the Lebesgue measure on $[0, 1]$, let φ be the density $\varphi = [dF/dH]$ and let $\psi = [dG/dH]$. By assumption, $\lambda_N\varphi + [1-\lambda_N]\psi$ is identically unity on $[0, 1]$.

Let L and M be the functions defined by the equalities

$$L(x) = \int_{1/2}^{x} J'(\xi)\, dF(\xi) = \int_{1/2}^{x} J'(\xi)\varphi(\xi)\, d\xi, \tag{3.4}$$

and

$$M(x) = \int_{1/2}^{x} J'(\xi)\, dG(\xi) = \int_{1/2}^{x} J'(\xi)\psi(\xi)\, d\xi. \tag{3.5}$$

If the function J belongs to $\mathcal{S}$, then both L and M belong to the set $\lambda_0^{-1}\mathcal{S} = \{v: \lambda_0 v \in \mathcal{S}\}$. Therefore, the preceding lemma applies to L and M as well as to J.

The remainder of the present section is devoted to continuity theorems which are easily proved under the assumption $H(x) \equiv x$ for $x \in [0, 1]$. However, to make them more directly applicable they will be stated for distributions on the line. For this purpose let $\mathfrak{D}$ be the set of pairs $(\check{F}, \check{G})$ of distributions on the real line subject to the only restriction that $\check{F}$ and $\check{G}$ have no common discontinuities. One could topologize $\mathfrak{D}$ as usual by the requirement that $(\check{F}_\nu, \check{G}_\nu) \to (\check{F}, \check{G})$ if $\check{F}_\nu(x) \to \check{F}(x)$ and $\check{G}_\nu(x) \to \check{G}(x)$ at every point of continuity. This topology can also be induced by the BL-norm (for Bounded Lipschitz) defined as follows (see [8]). If P and Q are two finite signed measures on the line, then

$$\|P - Q\|_{BL} = \sup_h \left| \int h\, dP - \int h\, dQ \right| \tag{3.6}$$

where the supremum is taken over all functions h such that $|h| \le 1$ and $|h(x) - h(y)| \le |x - y|$.

The space $\mathcal{S}' = \{J': |J'| \le f_0 + fg\}$ will be topologized by the topology of convergence in Lebesgue measure. This topology can be induced by the metric

$$\text{dist}\,(J_1', J_2') = \int_0^1 \frac{|J_1'(x) - J_2'(x)|}{1 + |J_1'(x) - J_2'(x)|}\, dx. \tag{3.7}$$

To each pair $(\check{F}, \check{G}) \in \mathfrak{D}$ and $J' \in \mathcal{S}'$ and each $\lambda \in [\lambda_0, 1 - \lambda_0]$ corresponds a pair (L, M) of functions defined on the interval $[0, 1]$. This pair is obtained by first reducing $H = \lambda F + (1 - \lambda)G$ to be uniform on $[0, 1]$ as explained in the introduction and then defining $L(x) = \int_{1/2}^{x} J'(\xi)\, dF(\xi)$, and so forth.

Let us say that $L_\nu \to L$ if $\int |L_\nu(x) - L(x)|^2\, dx \to 0$ and if $\sup \{|L_\nu(x) - L(x)|;\ x \in S\} \to 0$ for every compact subset S of the open interval $(0, 1)$ and similarly for M.

Lemma 3. *The map which makes correspond to* $[(F, G), J', \lambda] \in \mathfrak{D} \times \mathcal{S}' \times [\lambda_0, 1 - \lambda_0]$ *the pair* (L, M) *is jointly continuous for the topologies defined above.*

Proof. Since the topologies in question are all metrizable, it is sufficient to show that whenever a sequence $\{((\check{F}_k, \check{G}_k), J_k', \lambda_k)\}$ converges to a limit $((\check{F}, \check{G}), J', \lambda)$, then the corresponding pairs (L_k, M_k) converge to the appropriate pair (L, M).

Let $\check{H}_k = \lambda_k \check{F}_k + (1 - \lambda_k)\check{G}_k$ and let (F_k, G_k) be the pair obtained by the process described in the introduction. Let ξ be a number $\xi \in (0, 1)$. Consider the graph Γ_k of $\check{H}_k$ augmented by inserting vertical lines at jumps. If the horizontal line at the ordinate ξ meets Γ_k at a point which is not a jump of $\check{F}_k$, then $F_k(\xi) =$

$\check{F}_k[\check{H}_k(x)]$ for any point x such that $\check{H}_k(x-0) \leq \xi \leq \check{H}_k(x+0)$. It follows that F_k converges to F and that G_k converges to G. Thus we can assume that $(\check{F}_k, \check{G}_k)$ has been replaced by (F_k, G_k). However, in this case $\varphi_k = dF_k/dH_k$ is a bounded measurable function, $0 \leq \varphi_k \leq \lambda_0^{-1}$. Therefore, convergence of F_k to F implies that

$$\int \varphi_k(x)v(x)\,dx \to \int \varphi(x)v(x)\,dx \tag{3.8}$$

for every integrable function v. This, in turn, implies that $\int \varphi_k(x)v_k(x)\,dx \to \int \varphi(x)v(x)\,dx$ whenever $\int |v_k(x) - v(x)|\,dx \to 0$. Therefore, $L_k(x)$ converges to $L(x)$ uniformly on every interval of values of x which is bounded away from zero and unity. The convergence in quadratic mean follows from this and from the uniform integrability asserted by lemma 2. This proves the desired result.

A simple consequence of lemmas 2 and 3 which will be used in section 6 is the following. Let B_N be the random variable

$$\begin{aligned} B_N &= \int J(x)\,d(F_m - F) - \int L(x)\,d[H_N(x) - H(x)] \\ &= (1-\lambda_N)\left\{\int M(x)\,d(F_m - F) - \int L(x)\,d[G_n(x) - G(x)]\right\}. \end{aligned} \tag{3.9}$$

This expression is equivalent to the formula

$$\begin{aligned} \sqrt{N}B_N = {} & \frac{1-\lambda_N}{\sqrt{\lambda_N}}\frac{1}{\sqrt{m}}\sum_{j=1}^{m}[M(X_j) - EM(X_j)] \\ & + \sqrt{1-\lambda_N}\frac{1}{\sqrt{n}}\sum_{j=1}^{m}[L(Y_j) - EL(Y_j)], \end{aligned} \tag{3.10}$$

where the variables X_j and Y_j are all independent, and each X_j has distribution F, whereas each Y_j has distribution G. The variance of $\sqrt{N}B_N$ is given by

$$\begin{aligned} \sigma_N^2[F, G, J, \lambda_N] = {} & \frac{(1-\lambda_N)^2}{\lambda_N} \quad \text{variance} \quad M(X_1) \\ & + (1-\lambda_N) \quad \text{variance} \quad L(Y_1). \end{aligned} \tag{3.11}$$

PROPOSITION 1. *Let P_N be the distribution of $\sqrt{N}B_N$ and let Q_N be the normal distribution which has variance $\sigma_N^2[F, G, J, \lambda_N]$ and expectation zero. For every $\epsilon > 0$ there exists an $N(\epsilon)$ such that $N \geq N(\epsilon)$ implies $\|P_N - Q_N\|_{BL} < \epsilon$ for every $J \in \mathcal{S}$, and every triple $[(F, G), \lambda_N]$.*

In addition, there is an $N(\epsilon, a)$ such that $N \geq N(\epsilon, a)$ and $\sigma_N^2[F, G, J, \lambda_N] \geq a$ implies $\sup_x |P_N\{(-\infty, x]\} - Q_N\{(-\infty, x]\}| < \epsilon$ *for all $J \in \mathcal{S}$ and all triples $[(F, G), \lambda_N]$.*

PROOF. The first statement follows immediately from the usual central limit theorem and the uniform integrability asserted by lemma 2. The second statement follows from the first by the simple procedure of considering $\sqrt{N}B_N/\sigma_N$ instead of $\sqrt{N}B_N$. Hence the result.

In this connection the following lemma is of some interest.

LEMMA 4. *The equality* $\sigma_N^2[F, G, J, \lambda_N] = 0$ *implies that* $J'(x)\varphi(x)\psi(x) = 0$ *almost everywhere on the interval* $(0, 1)$. *However the identity* $J'\varphi\psi \equiv 0$ *is not sufficient to imply* $\sigma_N^2 = 0$.

PROOF. If, for instance, variance $M(X) = 0$, then M is almost everywhere constant on the set $E = \{x: \varphi(x) > 0\}$. Therefore, the derivative $M'(x) = J'(x)\psi(x)$ must be equal to zero at all points of density of the set E. This implies the stated result.

Let $\mathfrak{D} \times \mathcal{S}' \times [\lambda_0, 1 - \lambda_0]$ be topologized by the product topology used in lemma 3. For every $a \geq 0$ the set of triples $((F, G), J', \lambda)$ such that $\sigma_N^2 > a$ is an open subset of $\mathfrak{D} \times \mathcal{S}' \times [\lambda_0, 1 - \lambda_0]$. This implies the following corollary.

COROLLARY. *Assume that* J *is not constant. If* $(\check{F}_0, \check{G}_0) \in \mathfrak{D}$ *is a pair such that* $\varphi_0(x)\psi_0(x) > 0$ *almost everywhere for some* $\lambda \in [\lambda_0, 1 - \lambda_0]$, *then there is an* $a > 0$ *and an open neighborhood of* $(\check{F}_0, \check{G}_0)$ *such that* $\sigma_N^2 > a$ *for every pair* $(\check{F}, \check{G})$ *in this neighborhood.*

PROOF. The condition $\varphi_0(x)\psi_0(x) > 0$ almost everywhere with respect to the Lebesgue measure is equivalent to the condition that the measures induced by $\check{F}_0$ and $\check{G}_0$ are mutually absolutely continuous. Thus it is independent of the choice of λ. Since σ_N^2 is continuous in λ, the values $\sigma_N^2[(F_0, G_0), J, \lambda]$ attain their minimum as λ varies in $[\lambda_0, 1 - \lambda_0]$. According to lemma 4, this minimum value is a positive number, say, $2a > 0$. For each $\lambda \in [\lambda_0, 1 - \lambda_0]$ let V_λ be a neighborhood of $(\check{F}_0, \check{G}_0)$ and let W_λ be a neighborhood of λ such that $\sigma_N^2 > a$ for $(\check{F}, \check{G}) \in V_\lambda$ and $\xi \in W_\lambda$. There is a finite system $\{W_{\lambda_i}\}$ which covers $[\lambda_0, 1 - \lambda_0]$. If $(\check{F}, \check{G}) \in \bigcap_i V_{\lambda_i}$, one has $\sigma_N^2 > a$ for every $\lambda \in [\lambda_0, 1 - \lambda_0]$. Hence the result.

More specifically, the following lemma holds.

LEMMA 5. *Assume that* J *is not constant. Let* $\{J_k\}$ *be a sequence such that* $J_k \in \mathcal{S}$ *and such that* $J_k' \to J'$ *in Lebesgue measure. Let* $\{(\check{F}_k, \check{G}_k)\}$ *be a sequence of pairs converging to a pair* $(\check{F}, \check{G})$ *at all points of continuity of the pair* $(\check{F}, \check{G})$. *Then if* $\check{F} = \check{G}$,

$$\frac{\lambda_N}{1 - \lambda_N} \sigma_N^2[\check{F}_k, \check{G}_k, J_k, \lambda_N] \tag{3.12}$$

converges uniformly in N to $\int_0^1 J^2(u)\,du - [\int_0^1 J(u)\,du]^2 > 0$.

PROOF. It is sufficient to apply lemma 3 and compute the limiting value of σ_N^2. This limit is equal to

$$\begin{aligned} \frac{(1 - \lambda_N)^2}{\lambda_N} \text{ variance } M(X_1) + (1 - \lambda_N) \text{ variance } L(Y_1) \\ = \left[\frac{(1 - \lambda_N)^2}{\lambda_N} + (1 - \lambda_N)\right] \text{ variance } J(X_1), \end{aligned} \tag{3.13}$$

since $L \equiv M$, and since X_1 and Y_1 have the same distribution. The result follows.

4. Certain properties of empirical distribution functions

For this section we shall derive several inequalities and limit theorems which can be used to show that the higher order random terms occurring in theorem 4.1

tend to zero as N tends to infinity. The first results are inequalities on the tails of empirical distribution functions and a sharpened form of a theorem of Donsker [7]. For simplicity of notation the results are given for the uniform distribution [0, 1]. There is no difficulty in rewording them to apply to arbitrary continuous distributions.

A convenient tool in the derivation of these results is a replacement of binomial variables by Poisson variables which can be described as follows.

Let $\epsilon \in (0, 1)$. Let $\{u_j, j = 1, 2, \cdots\}$ be a sequence of independent random variables which are uniformly distributed on $(0, \epsilon)$. Let (r, s) be a pair of integer-valued random variables independent of the u_j. Assume that the joint distribution of (r, s) is such that marginally r has a binomial distribution, $B(m, \epsilon)$, corresponding to m trials with probability of success ϵ. Assume also that s has a Poisson distribution with expectation $m\epsilon$. Let U_m and V_m be the processes defined for $t \in (0, \epsilon)$ by taking $mU_m(t)$ equal to the number of u_j's such that $u_j \leq t$ and $j \leq r$ and taking $mV_m(t)$ equal to the number of u_j's such that $u_j \leq t$ and $j \leq s$.

LEMMA 6. *There is a joint distribution for the pair* (r, s) *such that*

$$P\{U_m(t) \equiv V_m(t) \text{ all } t \in (0, \epsilon)\} \geq 1 - 2\epsilon. \tag{4.1}$$

PROOF. It is sufficient to select a joint distribution for (r, s) such that $P[r \neq s] \leq 2\epsilon$. The possibility of such a selection results from a theorem of Prohorov [12].

Note that if F_m is the empirical cumulative obtained from m uniformly distributed independent variables on [0, 1], then the two processes $\{F_m(t);\ t \in (0, \epsilon)\}$ and $\{U_m(t);\ t \in (0, \epsilon)\}$ have identical distributions.

LEMMA 7. *Let* g *be a positive nonincreasing function defined on* $(0, \epsilon)$. *Then*

$$P\{\sup_{0<t<\epsilon} \sqrt{m}g(t)|U_m(t) - t| \geq 1\} \leq 2\epsilon + \int_0^\epsilon g^2(u)\,du. \tag{4.2}$$

PROOF. According to lemma 6, it is sufficient to show that

$$P\{\sup_{0<t<\epsilon} \sqrt{m}g(t)|V_m(t) - t| \geq 1\} \leq \int_0^\epsilon g^2(u)\,du. \tag{4.3}$$

This follows immediately from the remark that V_m has independent increments such that $EV_m(t) = t$ and $Em[V_m(t) - t]^2 = t$. The process $Z(t) = m|V_m(t) - t|^2$ is a semimartingale to which the Hájek-Rényi inequalities [10], or their generalization by Birnbaum and Marshall [3], can be applied. This gives the stated result.

After this was written, we became aware of results of B. Rosen [14] which give similar inequalities without using the Poisson approximation. Also, certain deeper results of D. M. Čibisov [6] could be used to obtain sharper inequalities.

Another result needed in the sequel is the following lemma.

LEMMA 8. *Let* U_m *be the empirical cumulative distribution obtained from* m *independent observations on the uniform distribution on* [0, 1]. *For every* $\epsilon > 0$ *there exists a* $\beta > 0$ *such that*

$$P\left\{\sup_t \frac{U_m(t)}{t} > \frac{1}{\beta}\right\} < \epsilon \tag{4.4}$$

and

$$P\{U_m(t) \geq \beta t \text{ for every } t \text{ such that } U_m(t) > 0\} > 1 - \epsilon. \tag{4.5}$$

PROOF. For $t > \delta > 0$ this follows, for instance, from Donsker's theorem, or from the Kolmogorov-Smirnov theorems. For t small one can again reduce the problem to an equivalent one concerning the standard Poisson processes. For the standard Poisson process the result is well known and easily verifiable.

Consider now two cumulative distribution functions F and G and two integers m and n. Let $N = m + n$ and let $\lambda_N = m/N$. Assume $0 < \lambda_0 < \lambda_N < 1 - \lambda_0 < 1$. Assume also that $H(t) = \lambda_N F(t) + (1 - \lambda_N)G(t)$ is identical to t for $t \in [0, 1]$. If $H_N = \lambda_N F_m + (1 - \lambda_N)G_n$ is the combined sample cumulative obtainable from m observations with distribution F and n observations with distribution G, one can obtain bounds on H_N from the bounds on the component cumulative distributions F_m and G_n. Further information can also be obtained as follows.

Let $\varphi = dF/dH$ be the density of F with respect to the Lebesgue measure on $[0, 1]$. Let S be the set $S = (0, \delta] \cup [1 - \delta, 1)$ with $0 < 2\delta < 1$. Classify points to be placed on the interval $(0, 1)$ in four categories, according to whether they are in S or S^c and according to whether they are labeled X or Y. For the pair (F_m, G_n) this gives a matrix $\nu = \{(\nu_{i,j}); i = 1, 2; j = 1, 2\}$ with $\nu_{1,1} + \nu_{2,1} = m$ and $\nu_{1,2} + \nu_{2,2} = n$. Let p be the probability S for F and let q be the probability of S for G. One can form another matrix ν^* such that $\nu^*_{1,1}$ and $\nu^*_{1,2}$ are independent Poisson variables with expectations $E\nu^*_{1,1} = E\nu_{1,1} = mp$ and $E\nu^*_{1,2} = E\nu_{1,2} = nq$. Taking $\nu^*_{2,1} = \nu_{2,1}$ and $\nu^*_{2,2} = \nu_{2,2}$, Prohorov's theorem insures the existence of a joint distribution such that $P[\nu \neq \nu^*] \leq 2(p + q)$.

Consider also another matrix $\tilde{\nu}$ whose distribution is given by a multinomial distribution with N trials and probabilities $p_{1,1} = \lambda_N p$ and $p_{1,2} = (1 - \lambda_N)q$ and $p_{2,1} = \lambda_N(1 - p)$, and finally $p_{2,2} = (1 - \lambda_N)(1 - q)$. One could find a joint distribution such that $P\{(\nu^*_{1,1}, \nu^*_{1,2}) \neq (\tilde{\nu}_{1,1}, \tilde{\nu}_{1,2})\} \leq 2(p + q)$. Therefore, one can find a joint distribution such that $P\{(\nu_{1,1}, \nu_{1,2}) \neq (\tilde{\nu}_{1,1}, \tilde{\nu}_{1,2})\} \leq 4(p + q)$. For such a joint distribution, one can construct the second row of the matrix $\tilde{\nu}$ by selecting $\tilde{\nu}_{2,1}$ from a binomial distribution with probability of success $[\lambda_N(1 - p)]\,[\lambda_N(1 - p) + (1 - \lambda_N)(1 - q)]^{-1}$ and $[N - (\tilde{\nu}_{1,1} + \tilde{\nu}_{1,2})]$ trials.

Another matrix ν' can be constructed with $(\nu'_{1,1}, \nu'_{1,2}) = (\tilde{\nu}_{1,1}, \tilde{\nu}_{1,2})$ and $\nu'_{1,1} + \nu'_{2,1} = m$ and $\nu'_{1,2} + \nu'_{2,2} = n$. Then $P[\nu \neq \nu'] \leq 4(p + q)$. Given the matrix $\tilde{\nu}$ one can place independently $\tilde{\nu}_{1,1}$ points in S and $\tilde{\nu}_{2,1}$ points in S^c, according to the distribution F. One can also place independently $\tilde{\nu}_{1,2}$ points in S and $\tilde{\nu}_{2,2}$ points in S^c, according to the distribution G. It is easily verified that the system so obtained has exactly the same distribution as the system of points obtainable by the following procedure. First select points $\{\xi_j, j = 1, 2, \cdots, N\}$ independently according to the Lebesgue measure on $[0, 1]$. Then, for each ξ_j and independently of the rest, with probability $\lambda_N \varphi(\xi_j)$ label it X, and with probability $1 - \lambda_N \varphi(\xi_j)$, label it Y.

Since the combined cumulative H_N ignores the distinction between X and Y, the above argument shows that, except for cases having total probability at most $4(p + q)$, the behavior of H_N on S will be the same as that of a sample cumulative from N independent uniformly distributed observations. To apply the preceding lemmas to the study of H_N, let us introduce the following notation. If v is a numerical function defined on (0, 1) and g is an element of $\mathcal{U}_2$, the g-norm of v is the number

$$\|v\|_g = \sup_t \{|v(t)g(t)|; t \in (0, 1)\}. \tag{4.6}$$

Let $B(g)$ be the set of functions v which are defined on (0, 1) and have a finite g norm. Let $S = \{s_j; j = 0, 1, 2, \cdots, h\}$ be a finite subset of the interval [0, 1] such that $0 = s_0 < s_1 < s_2 < \cdots < s_{h-1} < s_h = 1$. To such a set S associate a projection Π_S of $B(g)$ into itself by the requirements that $(\Pi_S v)(s_j) = v(s_j)$ for every $s_j \in S$ and that

$$\{[\Pi_S v](s) - (\Pi_S v)s_j\} = \frac{s - s_j}{s_{j+1} - s_j} [v(s_{j+1}) - v(s_j)] \tag{4.7}$$

for $s \in (s_j, s_{j+1})$.

Let $\{W_N(t); t \in [0, 1]\}$ be the process $W_N(t) = \sqrt{N}[H_N(t) - t]$.

PROPOSITION 2. *Let g be an element of $\mathcal{U}_2$. For every $\epsilon > 0$ there is a $\delta > 0$ and an integer N_0 depending on ϵ and g only such that $N \geq N_0$ implies*

$$P\{\|W_N - \Pi_S W_N\|_g > \epsilon\} < \epsilon \tag{4.8}$$

for every pair (F, G) of continuous distribution functions and every set $S = \{s_j; j = 0, 1, 2, \cdots, k\}$; $0 = s_0 < s_1 < \cdots < s_{k-1} < s_k = 1$, such that $s_{j+1} - s_j < \delta$ for every $j = 0, 1, \cdots, k - 1$.

PROOF. Let α be a number so small that g becomes monotone in $(0, \alpha]$ and $[1 - \alpha, 1)$ and that

$$16\left\{\alpha + \frac{1}{\epsilon^2}\int_0^\alpha g^2(\xi)\,d\xi + \frac{1}{\epsilon^2}\int_{1-\alpha}^1 g^2(\xi)\,d\xi\right\} < \epsilon. \tag{4.9}$$

Since $tg^2(t) \leq \int_0^t g^2(\xi)\,d\xi \leq \int_0^\alpha g^2(\xi)\,d\xi$ for $t \leq \alpha$, it follows from lemma 7 that

$$P\{\sup_t [g(t)|W_N(t) - tW_N(\alpha)|; t \in (0, \alpha]] > \epsilon\} \leq 2\alpha + \frac{4}{\epsilon^2}\int_0^\alpha g^2(\xi)\,d\xi. \tag{4.10}$$

A similar inequality holds for values of t belonging to $[1 - \alpha, 1)$. Therefore, it will be sufficient to prove the assertion for the process $\{W_N(t); t \in (\alpha, 1 - \alpha)\}$ and a function g which is bounded. The process W_N can be written

$$W_N(t) = \sqrt{\frac{m}{N}}\sqrt{m}[F_m(t) - F(t)] + \sqrt{\frac{n}{N}}\sqrt{n}[G_n(t) - G(t)]. \tag{4.11}$$

According to the argument of Donsker [5], there exist an N_0 and a $\delta > 0$ such that $m \geq \lambda_0 N_0$ and $\sup_j [F(s_j) - F(s_{j-1})] < \delta$ implies that if Ω'_j is the oscillation of $\sqrt{m}[F_m(t) - F(t)]$ in the interval $[s_{j-1}, s_j]$, then

$$\sum_j P\left[\Omega_j' > \frac{\epsilon}{2\|g\|}\right] < \frac{\epsilon}{4}. \tag{4.12}$$

The same result can be applied to $\sqrt{n}[G_n(t) - G(t)]$. Thus, if $[s_j - s_{j-1}] < \delta\lambda_0$ and $N \geq N_0$, and if Ω_j is the oscillation of W_N in the interval $[s_{j-1}, s_j]$, one can write $\sum_j P[\Omega_j > \epsilon/\|g\|] < \epsilon/2$. This implies the desired result.

COROLLARY. *For every $g \in \mathfrak{U}_2$ and every $\epsilon > 0$ there is an $N_0 < \infty$ and a finite set $\{v_j; j = 1, 2, \cdots, k\}$ of continuous functions defined on* $[0, 1]$ *such that*

$$P\{\inf_j \|W_N - v_j\|_g > \epsilon\} < \epsilon \tag{4.13}$$

for every $N \geq N_0$ and every pair (F, G) of distributions having no common discontinuities.

Consider the process Z_m defined by $Z_m(t) = \sqrt{m}\{F_m(t) - F(t)\}$, where F_m is the empirical cumulative from a sample of size m drawn from the distribution F. Assume as usual that $\lambda_N F + (1 - \lambda_N)G = H$ is the uniform distribution on $[0, 1]$ and that $0 < \lambda_0 \leq \lambda_N \leq 1 - \lambda_0 < 1$.

LEMMA 9. *For each integer m let K_m be a random process defined on the interval* $[0, 1]$. *Let $Z_m^*(x) = Z_m[K_m(x)]$. If $P\{\sup_t |K_m(t) - t| > \epsilon\} \to 0$ for every $\epsilon > 0$, then*

$$P^*\{\sup_x |Z_m^*(x) - Z_m(x)| > \epsilon\} \tag{4.14}$$

tends to zero for every $\epsilon > 0$.

PROOF. According to proposition 2, or according to Donsker's theorem, for every $\epsilon > 0$ there exists an $N(\epsilon) < \infty$ and a finite set of continuous functions $\{v_j: j = 1, 2, \cdots, k\}$ such that

$$P\{\inf_j \|Z_m - v_j\| > \epsilon/3\} < \epsilon \tag{4.15}$$

for every $m \geq N(\epsilon)$. Let $\gamma_j(z) = 1$ if the *first* index i such that $\|z - v_i\| \leq \epsilon/3$ is precisely equal to j. Let $\gamma_j(z) = 0$ otherwise. According to the above inequality, if $\bar{Z}_m = \sum_j \gamma_j(Z_m)v_j$, then $P\{\|\bar{Z}_m - Z_m\| > \epsilon/3\} < \epsilon$. Therefore, eliminating cases having probability at most ϵ, one can also write $\sup_x |\bar{Z}_m[K_m(x)] - Z_m[K_m(x)]| \leq \epsilon/3$.

Furthermore, there exists a δ such that $|s - t| < \delta$ implies $|v_j(t) - v_j(s)| < \epsilon/3$ for every $j = 1, 2, \cdots, k$. Therefore, if $P\{\sup_t |K_m(t) - t| \geq \delta\} < \epsilon$, one can write $\|\bar{Z}_m[K_m] - \bar{Z}_m\| < \epsilon/3$, except in cases having probability at most 2ϵ. The result follows.

The preceding lemma 9 can be used under the following circumstances. Let H_N be the combined empirical cumulative. Let K_N be the function defined by $K_N(x) = \inf\{t: H_N(t) \geq x\}$. Assume as usual that $H(u) = u$ for $u \in [0, 1]$. Since $\|H_N - H\| \to 0$ in probability, the difference $\sup_x |K_N(x) - x|$ must also tend to zero in probability. It follows that $\|Z_m[K_N] - Z_m\|$ tends to zero in probability.

For values of x of the type $x = j/N$, the variable $Z_m[K_N(x)]$ is simply equal to $\sqrt{m}[F_m(\xi_j) - F(\xi_j)]$ where ξ_j is the order statistic of rank j in the combined sample. In other words, the number $mF_m[K_N(j/N)]$ is the number of X_j's whose rank is inferior or equal to j.

For the next proposition, it is convenient to introduce the space $\mathfrak{M}$ of all finite signed measures on the interval (0, 1) and their indefinite integrals. If $\mu \in \mathfrak{M}$, let $J_\mu(x) = \mu\{(0, x]\}$ and let $\|\mu\|$ be the total mass of μ. The functions J_μ are simply those functions of bounded variation on [0, 1] which are right continuous and vanish at zero.

PROPOSITION 3. *For every $\epsilon > 0$ there exists an $N(\epsilon)$ such that $N \geq N(\epsilon)$ implies*

$$P\left\{\sqrt{N}\left|\int [J_\mu(H_N) - J_\mu(H)]\, d(F_m - F)\right| > \epsilon\|\mu\|\right\} < \epsilon \tag{4.16}$$

for every $\mu \in \mathfrak{L}$ and every pair (F, G) of distribution functions having no common discontinuities.

REMARK. In the above proposition one could replace H_N by $(N/(N+1))H_N$, since $\sqrt{N}\int [J_\mu((N/(N+1))H) - J_\mu(H)]\, dF$ is of order $1/\sqrt{N}$.

PROOF. The integral

$$I_m = \sqrt{m}\int J_\mu(H_N)\, d(F_m - F) = \int J_\mu[H_N(x)]\, dZ_m(x) \tag{4.17}$$

can also be written

$$I_m = \int \{Z_m(1) - Z_m[K_N(\xi) - 0]\}\mu(d\xi). \tag{4.18}$$

Therefore,

$$\begin{aligned} \left|\sqrt{m}\int [J_\mu(H_N) - J_\mu(H)]\, d(F_m - F)\right| &\\ = \left|\int \{Z_m(\xi - 0) - Z_m[K_N(\xi) - 0]\}\mu(d\xi)\right. &\\ \leq \|\mu\| \sup_x |Z_m(x) - Z_m[K_N(x)]|. & \end{aligned} \tag{4.19}$$

This implies the desired result by application of lemma 9. Another result which may be useful in the investigation of the Chernoff-Savage statistics (but will not be needed for our purposes) is a theorem relative to the behavior of the quantile function K_N defined on the interval [0, 1] by the formula

$$K_N(u) = \inf\{x\colon H_N(x) \geq u\}. \tag{4.20}$$

For the present purposes the assumptions that $H(x) \equiv x$ for $x \in (0, 1)$ and that $H = \lambda_N F + (1 - \lambda_N)G$ with $0 < \lambda_0 \leq \lambda_N \leq 1 - \lambda_0$ are rather important. Consider, under these conditions, the process Z_N defined by $Z_N(u) = \sqrt{N}[K_N(u) - u]$, $u \in [0, 1]$.

Since the process $\sqrt{N}(H_N - H)$ is asymptotically Gaussian, it follows from the equivalence $\{K_N(u) > x\} \Leftrightarrow \{H_N(x) < u\}$ that for any finite set $0 = u_0 < u_1 < \cdots < u_{r-1} < u_r = 1$ the distribution of the vector $\{Z_N(u_j); j = 1, 2, \cdots, r\}$ is also asymptotically normal, with the same covariance function

(4.21) $$C_N(u, v) = \frac{m}{N} F(u)[1 - F(v)] + \frac{n}{N} G(u)[1 - G(v)]$$

for $u \leq v$ as the process $\sqrt{N}(H_N - H)$ itself.

The following proposition strengthens this result. Let S be a finite set $S = \{u_j; j = 0, 1, 2, \cdots, r\}$, with $0 = u_0 < u_1 < \cdots < u_r = 1$. Let $Z_{N,S}(u)$ be defined by $Z_{N,S}(u) = Z_N(u)$ if $u \in S$. If u is between two consecutive points u_j and u_{j+1} of S, define $Z_{N,S}(u)$ by linear interpolation.

PROPOSITION 4. *For every $\epsilon > 0$ there exists a finite set S and an integer $N(\epsilon)$ such that $N \geq N(\epsilon)$ implies*

(4.22) $$P\{\|Z_{N,S} - Z_N\| > \epsilon\} < \epsilon$$

for every pair (F, G).

PROOF. Let ϵ be a positive number $0 < \epsilon < 1$ and let b and r be positive integers. Select b such that $2P\{\sqrt{N}\|H_N - H\| > b\} < \epsilon$. For any function h defined on $[0, 1]$ let Πh be the function obtained by taking $(\Pi h)(j/r) = h(j/r)$ for $j = 0, 1, 2, \cdots, r$ and interpolating linearly between successive values. One can find a number r and an integer N_0 such that $N \geq N_0$ implies $2P\{\sqrt{N}\|\Pi H_N - H_N\| \geq \epsilon(1 - \epsilon)\} < \epsilon$ and $2r[b(\epsilon) + 1]^2 < \epsilon\sqrt{N}$.

Let $H_N^* = H_N$ if $\sqrt{N}\|H_N - H\| < b$ and $\sqrt{N}\|\Pi H_N - H_N\| < \epsilon(1 - \epsilon)$. Let $H_N^* = H$ otherwise. It follows that $N \geq N_0$ implies $P\{H_N^* \neq H_N\} < \epsilon$. Furthermore, $\sqrt{N}\|\Pi H_N^* - H_N^*\| < \epsilon(1 - \epsilon)$ and $\sqrt{N}\|H_N^* - H\| < b$ without exception. The second inequality implies that the segments of lines which compose ΠH_N^* have slopes which differ from unity by no more than $\alpha < [2rb(\epsilon)]/\sqrt{N} < \epsilon$. This implies in particular that ΠH_N^* is increasing. Further, let K_N^* be the function related to H_N^* by the equation $K_N^*(u) = \inf\{x: H_N^*(x) \geq u\}$ and let $\tilde{K}_N$ be the corresponding function relative to $\tilde{H}_N = \Pi H_N^*$.

Since the slope of ΠH_N^* is always larger than $1 - \epsilon$, the inequality $\sqrt{N}\|\Pi H_N^* - H_N^*\| < \epsilon(1 - \epsilon)$ implies $\sqrt{N}\|\tilde{K}_N - \tilde{K}_N^*\| < \epsilon$. Thus, it will be sufficient to prove the result for $\tilde{K}_N$, instead of K_N^* or K_N. Interpolate $\tilde{K}_N$ just as before. That is, let $\bar{K}_N(u) = \tilde{K}_N(u)$ if $ru = j$, $j = 0, 1, 2, \cdots, r$ and interpolate linearly between these values. Suppose that $\sqrt{N}[\tilde{H}_N(j/r) - H(j/r)] = z_1 > 0$ and that $\sqrt{N}[\tilde{H}_N((j + 1)/r) - H((j + 1)/r)] = z_2 < 0$. Since the slopes of the segments composing $\tilde{H}_N$ are between $1 - \alpha$ and $1 + \alpha$, this implies $\tilde{K}_N(j/r) \geq j/r - (1 - \alpha)z_1/\sqrt{N}$ and $\tilde{K}_N((j + 1)/N) \leq (j + 1)/r + (1 + \alpha)|z_2|/\sqrt{N}$. It follows from this that

(4.23) $$\begin{aligned} \sqrt{N}\left\{\bar{K}_N\left[\frac{j}{r} + \frac{z_1}{\sqrt{N}}\right] - \frac{j}{r}\right\} &\leq r\left\{(1 + \alpha)\frac{|z_2|}{\sqrt{N}} + (1 - \alpha)\frac{z_1}{\sqrt{N}} + \frac{1}{r}\right\} z_1 - (1 - \alpha)z_1 \\ &\leq z_1\left\{\alpha + (1 + \alpha)r\left[\frac{z_1 + |z_2|}{\sqrt{N}}\right]\right\}. \end{aligned}$$

Since $|z_i| \leq b$ and since $2rb^2 < \epsilon\sqrt{N}$, this is smaller than $b\alpha + (1 + \alpha)\epsilon < 3\epsilon$. The desired result is an immediate consequence of these inequalities.

REMARK. The result of proposition (4) is well known for the case where $F = G$. In fact, one can reverse the chain of arguments leading to the proof of proposition (4) to obtain a simple and rather elementary proof of Donsker's theorem. For the case where H is not the uniform distribution and for applications, see [2] and [9].

A consequence of proposition (4) is the following result.

COROLLARY. *There exist joint distributions for pairs of processes* (Z'_N, W_N) *such that*

(i) *the distribution of* Z'_N *is the same as that of* $\sqrt{N}[K_N - K]$;

(ii) *the process* W_N *is Gaussian, with mean zero and covariance*

$$EW_N(s)W_N(t) = \frac{m}{N} F(s)[1 - F(t)] + \frac{n}{N} G(s)[1 - G(t)] \tag{4.24}$$

for $s \leq t$;

(iii) *for every* $\epsilon > 0$ *there is an* $N(\epsilon)$ *such that* $P\{\|Z'_N - W_N\| > \epsilon\} < \epsilon$ *if* $N \geq N(\epsilon)$.

This follows immediately from proposition (4) and a theorem of V. Strassen [16].

Propositions (3) and (4) can be used to investigate the asymptotic properties of statistics of the Chernoff-Savage type as follows.

Let $\mathfrak{M}$ be the space of finite signed measures on the interval [0, 1]. For each $\mu \in M$ let $J_\mu = \mu\{(0, x]\}$. Let T_N be the expression

$$T_N = \sqrt{N}\left\{\int J_\mu(H_N)\, dF_m - \int J_\mu(H)\, dF\right\}. \tag{4.25}$$

Introduce the pair of stochastic processes (Z_N, Z^*_N) by the equalities

$$Z_N = \sqrt{N}[K_N - K], \qquad Z^*_N(\xi) = \sqrt{m}[F_m(\xi - 0) - F(\xi)]. \tag{4.26}$$

Let T^*_N be the expression

$$\begin{aligned} T^*_N &= \sqrt{N} \int [J_\mu(H_N) - J_\mu(H)]\, dF + \sqrt{N} \int J_\mu(H)\, d(F_m - F) \\ &= \sqrt{N} \int \left[F(\xi) - F\left[\xi + \frac{1}{\sqrt{N}} Z_N(\xi)\right]\right] \mu(d\xi) - \sqrt{\frac{N}{m}} \int Z^*_N(\xi)\mu(d\xi). \end{aligned} \tag{4.27}$$

According to proposition 3, the difference $T^*_N - T_N$ converges to zero in probability, uniformly for $\|\mu\|$ bounded.

According to propositions 2 and 4, for $\epsilon > 0$, both Z_N and Z^*_N admit linear interpolations of bounded rank which differ from them by less than ϵ, except in cases of probability ϵ. Therefore, one could find a suitable probability space and pairs (W_N, W^*_N) of Gaussian processes with appropriate covariances such that for every $\epsilon > 0$,

$$(4.28)\qquad P\{\|Z_N - W_N\| + \|Z_N^* - W_N^*\| > \epsilon\} < \epsilon$$

for $N \geq N(\epsilon)$ and for every pair (F, G). The functions

$$(4.29)\qquad T_N^*(W, W^*) = \sqrt{N} \int \left[F(\xi) - F\left(\xi + \frac{W(\xi)}{\sqrt{N}}\right)\right] \mu(d\xi) - \sqrt{\frac{N}{m}} \int W^*(\xi)\mu(d\xi)$$

satisfies a Lipschitz condition

$$(4.30)\qquad |T_N^*(u, u^*) - T_N^*(v, v^*)| \leq [\|u - v\| + \|u^* - v^*\|] \frac{1}{\lambda_0} \|\mu\|,$$

since $\lambda_N F$ has a derivative bounded by unity and since $\lambda_0 \leq \lambda_N$. As a consequence, one can state the following corollary.

PROPOSITION 5. *Let T_N, W_N, and W_N^* be the objects defined above. Let P_N be the distribution of T_N and let Q_N be the distribution of*

$$(4.31)\qquad \sqrt{N} \int \left\{F(\xi) - F\left[\xi + \frac{W_N}{\sqrt{N}}\right]\right\} \mu(d\xi) - \sqrt{\frac{N}{m}} \int W_N^*(\xi)\mu(d\xi).$$

For every $\epsilon > 0$ there exists an $N(\epsilon)$ such that $N \geq N(\epsilon)$ implies $\|P_N - Q_N\|_{BL} \leq \epsilon\|\mu\|$ for every $\mu \in \mathfrak{M}$ and every pair (F, G).

REMARK. The above proposition remains valid if T_N is modified by replacing H_N by $(N/(N + 1))H_N$. It is easily checked that such a replacement amounts to a slight modification of the measure μ and the introduction of terms which are at most of order $[\|\mu\|/\sqrt{N}]$.

5. Bounds for the tails of the Chernoff-Savage statistics

For the purposes of the present section, let f and g be the functions defining the set $\mathcal{S}$ and let $\mathcal{S}_0$ be the set of indefinite integrals of the type $J(x) = \int_{1/2}^{x} J'(\xi)\, d\xi$ with $|J'| \leq fg$. If τ is a number such that $0 < 2\tau < 1$, let

$$(5.1)\qquad \Delta_N^*(J, \tau) = \sqrt{N} \int_A \left\{J\left(\frac{N}{N + 1} H_N\right) - J(H)\right\} dF_m,$$

where the integral is taken over the set $A = (0, \tau] \cup [1 - \tau, 1)$.

PROPOSITION 6. *For every $\epsilon > 0$ there exists a number τ_0 such that*

$$(5.2)\qquad P\{\sup [|\Delta_N^*(J, \tau)|; 0 < \tau \leq \tau_0, J \in \mathcal{S}_0] > \epsilon\} < \epsilon$$

for every N and every pair (F, G).

PROOF. If one reverses the order of the observations by changing x to $1 - x$, the part of $\Delta_N^*(J, \tau)$ arising from the integral over $[1 - \tau, 1)$ is transformed into a similar integral, for a different function J, over the interval $(0, \tau]$. Thus, it will be sufficient to bound the part relative to $(0, \tau]$. Let δ be a number $0 < 8\delta < \lambda_0^{-1}\epsilon$ such that both f and g are monotone decreasing in the interval

$(0, \delta]$. One can assume throughout that $\tau \leq \delta$. In this case, replacing binomial distributions by Poisson distributions as explained in section 4, one can bound instead of Δ_N^* the simpler expression

$$S_N^*(J, \tau) = \frac{\sqrt{N}}{m} \sum_{\xi_i \leq \tau} \left| J\left(\frac{i}{N+1}\right) - J(\xi_i) \right| Z_{N,i} \tag{5.3}$$

where the ξ_i are the order statistics from a sample of size N taken from the uniform distribution. The $Z_{N,i}$ are independently selected with conditional probability of being equal to unity given by

$$P[Z_{N,i} = 1 | \xi_1, \xi_2, \cdots, \xi_N] = 1 - P[Z_{N,i} = 0 | \xi_1, \cdots, \xi_N] = \lambda_N \varphi(\xi_i). \tag{5.4}$$

Instead of using the above representation, one can also introduce independent random variables $\{U_j\}$, $j = 1, 2, \cdots, N$ which are uniformly distributed on $[0, 1]$ and their ranks R_j. The whole system $\{(U_j, R_j); j = 1, 2, \cdots, N\}$ will be denoted by the letter W. Taking this possibility into account and the fact that $\lambda_N \geq \lambda_0$, it will be sufficient to bound

$$S_N(J, \tau) = \frac{1}{\sqrt{N}} \sum_{U_i \leq \tau} \left| J\left(\frac{R_i}{N+1}\right) - J(U_i) \right|. \tag{5.5}$$

According to lemma 8, there is a number $\beta > 0$ such that

$$P\left\{\beta U_i \leq \frac{R_i}{N+1} \leq \frac{1}{\beta} U_i \text{ for all } i\right\} > 1 - \epsilon/4. \tag{5.6}$$

Let f_β be the function $f_\beta(x) = f(\beta x)$ defined for $x \in (0, \delta]$. Define g_β similarly. Once β is chosen, lemma 7 implies the existence of a number c_1 such that

$$P\left\{\sup_i \left[\sqrt{N} g_\beta(U_i) \left|\frac{R_i}{N} - U_i\right|; U_i \leq \delta\right] \geq c_1\right\} < \epsilon/4. \tag{5.7}$$

In addition, if $\beta(N+1)U_i \leq R_i \leq \beta^{-1}(N+1)U_i$, one can write

$$\max_i \left\{\frac{1}{\sqrt{N}} g_\beta(U_i) \frac{R_i}{N+1}\right\} \leq \frac{1}{\beta\sqrt{N}} \max_i U_i g_\beta(U_i). \tag{5.8}$$

If the maximum is restricted to those values i such that $U_i \leq \tau$, this last term is not larger than

$$\frac{1}{\beta\sqrt{N}} \sup_{x \leq \tau} x g_\beta(x) = \frac{1}{\beta\sqrt{N}} \sup_{x \leq \tau} x g(\beta x). \tag{5.9}$$

Since $xg(x) \to 0$ as $x \to 0$, there is a τ_1 and a c such that $P[W \in \mathcal{R}] > 1 - \epsilon$, if $\mathcal{R}$ is the set of systems $W = \{(U_i, R_i)\}$ which satisfy

(i) $(N+1)\beta U_i \leq R_i \leq (N+1)\beta^{-1} U_i$

and

(ii) $\sup_i \left\{\sqrt{N} g_\beta(U_i) \left|\frac{R_i}{N+1} - U_i\right|; U_i \leq \tau_1\right\} \leq c.$

However, if $W \in \mathcal{R}$, then $|J'(u)| \leq f_\beta(U_i) g_\beta(U_i)$ for every point u belonging to the interval between U_i and $R_i/(N+1)$.

It follows that

$$(5.10)\qquad S_N(J, \tau) = \frac{1}{\sqrt{N}} \sum_{U_i \le \tau} \left| J\left(\frac{R_i}{N+1}\right) - J(U_i)\right| \le \frac{c}{N} \sum_{U_i \le \tau} f_\beta(U_i).$$

The integral $\int_0^\tau f_\beta(u)\, du$ is equal to $(1/\beta) \int_0^{\tau/\beta} f(v)\, dv$. Therefore there exists a $\tau_0 \le \tau_1$ such that $c \int_0^{\tau_0} f_\beta(u)\, du \le \epsilon^2$. The desired result follows by application of Markov's inequality.

The quantity which appears in the study of Chernoff-Savage statistics is not exactly equal to $\Delta_N^*(J, \tau)$ but to

$$(5.11)\qquad \Delta_N(J, \tau) = \sqrt{N} \left\{ \int_A J\left(\frac{N}{N+1} H_N\right) dF_m - \int_A J(H)\, dF \right\},$$

where A is again equal to $(0, \tau] \cup [1 - \tau, 1)$. Clearly,

$$(5.12)\qquad \Delta_N = \Delta_N^* + \sqrt{N} \int_A J(H)\, d(F_m - F).$$

The difference term $\Delta_N^* - \Delta_N$ is a normalized sum with expectation zero and a variance bounded by expressions of the type $4\lambda_0^{-1} c^4(\tau)$ where $c(\tau)$ is a function described in the proof of lemma 2. Thus for every $\epsilon > 0$ there is a τ_0 such that $P\{|\Delta_N^* - \Delta_N| > \epsilon\} < \epsilon$ for $J \in \mathcal{S}_0$ and $\tau \le \tau_0$. In other words, the following corollary holds.

COROLLARY. *Let A be the set $A = (0, \tau] \cup [1 - \tau, 1)$ and let*

$$(5.13)\qquad \Delta_N(J, \tau) = \sqrt{N} \left| \int_A J\left(\frac{N}{N+1} H_N\right) dF_m - \int_A J(H)\, dF \right|.$$

For every $\epsilon > 0$ there is a number $\tau_0 > 0$ such that $\tau < \tau_0$ and $J \in \mathcal{S}_0$ implies

$$(5.14)\qquad P\{|\Delta_N(J, \tau)| > \epsilon\} < \epsilon$$

for every N and every pair (F, G).

For some purposes it is convenient to eliminate a few terms in the tails of the Chernoff-Savage statistics. In this connection, let us mention the following easy result. Suppose that $|J|$ is monotone decreasing in the interval $(0, \delta]$. Let k be an integer and let $y = k/(N + 1)$. Then

$$(5.15)\qquad \frac{\sqrt{N}}{m} \left| \sum_{i=1}^k J\left(\frac{i}{N+1}\right)\right| \le \frac{\sqrt{N}}{m} (N + 1) \int_0^y |J(x)|\, dx$$
$$\le \frac{N+1}{m} \sqrt{\frac{Nk}{N+1}} \left\{ \int_0^y J^2(x)\, dx \right\}^{1/2}.$$

Therefore, whenever J stays in a family of monotone functions which are uniformly square integrable the sum $(\sqrt{N}/m) \sum_{i=1}^k |J(i/(N + 1))|$ tends to zero for each fixed k as $N \to \infty$. It follows that for each fixed $y \in (0, \infty)$ the terms

$$(5.16)\qquad \sqrt{N} \int_{Nx \le y} \left| J\left(\frac{N}{N+1} H_N(x)\right)\right| dF_m(x)$$

tend to zero in probability as $N \to \infty$.

6. The asymptotic behavior of the Chernoff-Savage statistics

For the purposes of the present section it is convenient to consider the set $\mathcal{S}_1 \supset \mathcal{S}$ of functions J which are indefinite integrals of functions J' such that $J' = J_1' + J_2'$ with $|J_2'| \leq fg$ and $\int |J_1'(x)|\,dx \leq b$. The set $\mathcal{S}_1$ will be topologized as follows. A sequence $\{J_k\}$ converges to J if $\int_A |J_k'(x) - J'(x)|\,dx \to 0$ for every interval $A = [\tau, 1 - \tau]$ with $0 < 2\tau < 1$. A subset S of $\mathcal{S}_1$ is called relatively compact if every sequence $\{J_k\} \subset S$ admits a convergent subsequence.

Triples $\{F, G, \lambda\}$ such that $H = \lambda F + (1 - \lambda)G$ is the uniform distribution on [0, 1] will be topologized by requiring that when $[F_\nu, G_\nu, \lambda_\nu] \to \{F, G, \lambda]$, the densities $\lambda_\nu \varphi_\nu = \lambda_\nu (dF_\nu)/dH$ converges in measure. Of course, it is still assumed that $0 < \lambda_0 \leq \lambda \leq 1 - \lambda_0$.

Consider the functions L and M defined in section 3 by $L(x) = \int^x J'(\xi)\,dF(\xi)$ and $M(x) = \int^x J'(\xi)\,dG(\xi)$. Let σ_N^2 be the variance $\sigma_N^2[(F, G), J, \lambda_N]$ introduced in section 3.

THEOREM 1. *Let J be an element of $\mathcal{S}_1$ and let*

$$(6.1)\qquad T_N = \sqrt{N}\left\{\int J\left[\frac{N}{N+1} H_N\right] dF_m - \int J(H)\,dF\right\}.$$

Let P_N be the distribution of T_N and let Q_N be the normal distribution which has expectation zero and variance $\sigma_N^2[(F, G), J, \lambda_N)]$. If S is a relatively compact subset of $\mathcal{S}_1$, then for every $\epsilon > 0$ there is an $N(\epsilon)$ such that $N \geq N(\epsilon)$ implies $\|P_N - Q_N\|_{BL} < \epsilon$ for every $J \in S$ and every triple $\{(F, G), \lambda_N\}$. Similarly, if $\mathfrak{F}$ is a relatively compact set of triples $[(F, G), \lambda]$, then for every $\epsilon > 0$ there exists an $N(\epsilon)$ such that $N \geq N(\epsilon)$ implies $\|P_N - Q_N\|_{BL} < \epsilon$ for every $J \in \mathcal{S}$ and every triple $\{(F, G), \lambda\} \in \mathfrak{F}$.

On sets such that σ_N^2 stays bounded away from zero the bounded Lipschitz norm $\|P_N - Q_N\|_{BL}$ may be replaced by the Kolmogorov vertical distance.

PROOF. Suppose $J' = J_1' + J_2'$ with $|J_2'| \leq fg$. According to proposition 6, for any given $\epsilon > 0$ there is a number $\tau > 0$ and an $N_0(\epsilon)$ such that if A is the set $(0, \tau] \cup [1 - \tau, 1)$ and if T_N^τ is the expression

$$(6.2)\qquad T_N^\tau = \sqrt{N}\left\{\int_A J_2\left[\frac{N}{N+1} H_N\right] dF_m - \int_A J_2(H)\,dF\right\},$$

then $P\{|T_N^\tau| > \epsilon\} < \epsilon$ for every J_2' and every triple $\{(F, G), \lambda\}$. Since the function $|J_2'|$ is integrable and bounded by $\sup\{f(x)g(x), \tau \leq x \leq 1 - \tau\}$ on the interval $[\tau, 1 - \tau]$, it will be sufficient to prove the theorem for integrable functions J' such that $\int |J'(x)|\,dx \leq b$.

In this case, according to proposition 3 or proposition 5, one can replace the variable T_N by $T_N^* = \sqrt{N} B_N + R_N$ with

$$(6.3)\qquad \sqrt{N} B_N = \sqrt{N} \int J(H)\,d(F_m - F) + \sqrt{N} \int [H_N - H] J'(H)\,dF,$$

$$(6.4)\qquad R_N = \sqrt{N} \int [J(H_N) - J(H) - (H_N - H)J'(H)]\,dF.$$

For every $\epsilon > 0$ there is an $N_1(\epsilon)$ such that $N \geq N_1(\epsilon)$ implies $P\{|T_N^* - T_N| > \epsilon\} < \epsilon$, whatever may be J and whatever may be F, G, and λ_N. Since $\sqrt{N}B_N$ is precisely the term introduced in section 3, the results claimed in the statement of the theorem will depend on the evaluation of appropriate bounds for R_N.

First note that given $\epsilon > 0$ there is a $c < \infty$ such that $P\{\sqrt{N}\|H_N - H\| \geq c\} < \epsilon$. Let $\tilde{H}_N = H_N$ if $\sqrt{N}\|H_N - H\| < c$ and let $\tilde{H}_N = H$ otherwise. Let

$$\tilde{R}_{N,1} = \sqrt{N} \int [J(\tilde{H}_N) - J(H)]\, dF \tag{6.5}$$

and let

$$\tilde{R}_{N,2} = \int \sqrt{N}[\tilde{H}_N - H]J'(H)\, dF. \tag{6.6}$$

If $\varphi = dF/dH$ and if $\tilde{K}_N(x) = \inf\{t\colon \tilde{H}_N(t) \geq x\}$, one can also write $\tilde{R}_{N,1}$ and $\tilde{R}_{N,2}$ in the form

$$\tilde{R}_{N,1} = \sqrt{N} \int \{F(\xi) - F[\tilde{K}_N(\xi)]\}J'(\xi)\, d\xi, \tag{6.7}$$

$$\tilde{R}_{N,2} = \int \sqrt{N}[\tilde{H}_N(\xi) - H(\xi)]J'(\xi)\varphi(\xi)\, d\xi. \tag{6.8}$$

Since $0 \leq \lambda_0\varphi \leq 1$ and since $\sqrt{N}\|\tilde{H}_N - H\| \leq c$ implies $\sqrt{N}\|\tilde{K}_N(\xi) - \xi\| \leq c$, both $\tilde{R}_{N,1}$ and $\tilde{R}_{N,2}$ satisfy Lipschitz conditions in J' for the norm $\|J'\| = \int |J'(s)|\, dx$. If $\tilde{R}_N[F, J'] = \tilde{R}_{N,1} - \tilde{R}_{N,2}$, this implies $|\tilde{R}_N(F, J')| \leq (2c/\lambda_0)\|J'\|$. Therefore, the first statement of the theorem, with uniformity of the convergence on compact subsets of $\mathcal{S}_1$ will follow if we show that for a fixed J' the term $\tilde{R}_N[F, J']$ converges to zero uniformly in F.

Let $\rho_N(x)$ be the ratio

$$\rho_N(x) = \sup_{|\xi| \leq c} \sqrt{N}\left|J\left[x + \frac{\xi}{\sqrt{N}}\right] - J(x) - \frac{\xi}{\sqrt{N}}J'(x)\right|. \tag{6.9}$$

By definition of the derivative, this converges to zero for almost every x. In addition, $\int \rho_N(x)\, dx$ converges to zero. However,

$$|\tilde{R}_N[F, J']| \leq \int \rho_N(x)\, dF(x) \leq \frac{1}{\lambda_0} \int \rho_N(x)\, dx. \tag{6.10}$$

The first statement follows.

For the second statement, note that if $J \in \mathcal{S}$ and if the part of J' carried by the set $(0, \tau] \cup [1 - \tau, 1)$ has been removed, the remaining part of J' is smaller than a certain integrable function $\omega = a + f_0$, with a equal to $\sup\{f(x)g(x);\ \tau \leq x \leq 1 - \tau\}$. The term $\tilde{R}_{N,2}$ can be written

$$\tilde{R}_{N,2}(\varphi) = \int \sqrt{N}[\tilde{H}_N(\xi) - H(\xi)]\left[\frac{J'(\xi)}{\omega(\xi)}\right]\omega(\xi)\, \varphi(\xi) d\xi. \tag{6.11}$$

This satisfies a Lipschitz condition,

(6.12) $$|\tilde{R}_{N,2}(\varphi)| \leq c\|\varphi\|,$$

for the norm $\|\varphi\| = \int \omega(\xi)|\varphi(\xi)|\,d\xi$. Similarly,

(6.13) $$\tilde{R}_{N,1}(\varphi) = \int \sqrt{N}[J(\tilde{H}_N) - J(H)]\varphi(\xi)\,d\xi$$

satisfies the condition

(6.14) $$\tilde{R}_{N,1}(\varphi) \leq \int \gamma_N(\xi)|\varphi(\xi)|\,d\xi$$

with

(6.15) $$\gamma_N(\xi) = \sup_{|x| \leq c} \sqrt{N}\left|\Omega\left[\xi + \frac{x}{\sqrt{N}}\right] - \Omega(\xi)\right|,$$

(6.16) $$\Omega(x) = \int_0^x \omega(\xi)\,d\xi.$$

Suppose then that $\varphi_\nu \to \varphi$ in measure; then

(6.17) $$\int \omega(\xi)|\varphi_\nu(\xi) - \varphi(\xi)|\,d\xi \to 0.$$

Furthermore,

(6.18) $$\int \gamma_N(\xi)|\varphi_\nu(\xi) - \varphi(\xi)|\,d\xi \to 0$$

uniformly in N since the functions γ_N are uniformly integrable. The result follows by the usual argument. This completes the proof of the theorem.

REMARK 1. The convergence is *not* uniform on the set of systems $\{(F, G), J, \lambda\}$ such that $J \in \mathcal{S}$. In fact, suppose that $J'_N(x)$ is equal to -1 for $2k/2^N < x \leq (2k+1)/2^N$ and to $+1$ for $(2k+1)/2^N < x \leq (2k+2)/2^N$, $k = 0, 1, 2, \cdots, 2^N$. Then $|J_N| \leq 2^{-N}$. Suppose that $2\lambda_N = 1$ and that F_N has a density φ_N equal to 2 for $2k/2^N < x \leq (2k+1)/2^N$ and to zero otherwise. Then the function L_N differs little from $L(x) = -x$ and M_N differs little from $M(x) = +x$. The expression $|T_N|$ is smaller than $2^{-N}\sqrt{N}$, but $4\sigma^2$ is approximately equal to unity.

In the paper of Chernoff and Savage, it is assumed that the second derivative J'' satisfies a restriction of the type $|J''(x)| \leq K[x(1-x)]^{-(5/2)+\delta}$. This implies in particular that the available family $\{J'\}$ is relatively compact for uniform convergence on the compact intervals of (0, 1). Thus theorem 1 asserts uniformity of the convergence on that class.

REMARK 2. One particular case in which the uniformity asserted in the theorem may be of interest is the following.

Suppose that for each N the distribution $\check{F}$ of the original observations labeled X is given by a distribution function Ψ which admits a density. Suppose also that the distribution $\check{G}$ of the variables Y is given by $\Psi\{(x - \theta)/\beta\}$ with $\beta > 0$. In this case if $\theta_\nu \to \theta_0$ and $\beta_\nu \to \beta_0 \neq 0$, the density of $\Psi[(x - \theta_\nu)/\beta_\nu]$ converges in measure to that of $\Psi[(x - \theta_0)/\beta_0]$. It follows that the corresponding distributions (F, G) reduced to the interval [0, 1] converge in the same sense. The convergence $\|P_N - Q_N\|_{BL} \to 0$ asserted in the theorem is therefore uniform for $J \in \mathcal{S}$, $\lambda_N \in [\lambda_0, 1 - \lambda_0]$ and every bounded set of values $[\theta, \log \beta]$. In addition, if for the integer N the value of (θ, β) is (θ_N, β_N) and $\theta_N \to 0$ and $\beta_N \to 1$, then the Kolmogorov distance $|P_N - Q_N|$ also tends to zero uniformly for any set $S \subset \mathcal{S}$ such that $J \in S$ implies $\int J^2(u)\,du - [\int J(u)\,du]^2 \geq \alpha > 0$. This follows immediately from lemma 5.

In many cases the functions J are obtained using expectations of suitable order statistics. In this connection the following theorem may be of interest.

For each integer N let J'_N be a nonnegative function such that $0 \leq J'_N(x) \leq K[x(1-x)]^{-(3/2)+\delta}$ for some fixed $K < \infty$ and some fixed δ, $0 < 2\delta < 1$. Let J_N be an integral of J'_N and let $\bar{J}_N(i/(N+1))$ be the expected value $E[J_N(\xi_{N,i})]$ where $\xi_{N,i}$ is the i-th smallest order statistic in a sample of size N from the uniform distribution on $[0, 1]$. Complete the definition of $\bar{J}_N$ by linear interpolation between successive integers i and by leaving $\bar{J}_N$ constant below $(1/(N+1))$ and above $(N/(N+1))$.

THEOREM 2. *Suppose that the relation* $|J'_N(x)| \leq K[x(1-x)]^{-(3/2)+\delta}$ *is satisfied and that* J'_N *converges in Lebesgue measure to a limit* J'. *Let* T_N *be the expression*

$$T_N = \sqrt{N}\left\{\int \bar{J}_N\left(\frac{N}{N+1}H_N\right)dF_m - \int J_N(H)\,dF\right\}. \tag{6.19}$$

Let P_N *be the distribution of* T_N *and let* Q_N *be the normal approximation of theorem* 1. *Then* $\|P_N - Q_N\|_{BL}$ *converges to zero, uniformly in* $[(F, G), \lambda]$ *as* $N \to \infty$.

PROOF. Let $T^*_N = \sqrt{N}\{\int \bar{J}_N(N/(N+1)H_N)\,dF_m - \int \bar{J}_N(H)\,dF\}$. According to lemma 1, the function $\bar{J}'_N$ satisfies the conditions of theorem 1, with $f = K[x(1-x)]^{-1+(\delta/2)}$ and $g^2 = [x(1-x)]^{-(1-\delta)}$. In addition, J'_N converges in measure to a limit J'; hence, $\bar{J}'_N$ converges to J' according to lemma 1. It follows from this that theorem 2 would be proved if T_N was replaced by T^*_N. To complete the proof, it will be sufficient to bound the difference $T_N - T^*_N = \sqrt{N}\{\int [\bar{J}_N(H) - J_N(H)]\,dF$. However, this is smaller than $\lambda_N^{-1}\int \sqrt{N}|\bar{J}_N(x) - J_N(x)|\,dx$.

Since, in this last integral, the terms in the absolute value sign are linear functions of J'_N and since J'_N can be split into positive and negative parts and each of these into parts supported by $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ respectively, it will be sufficient to prove the result under the assumption $J_N(x) = \int_x^1 h_N(\xi)\,d\xi$, $0 \leq h_N(x) \leq x^{-(3/2)+\delta}$, $h_N(x) = 0$ for $x > \frac{1}{2}$. The parts relative to the interval $[\frac{1}{2}, 1]$ are handled by changing x into $1 - x$.

With this definition, note that

$$\sqrt{N}\int_0^x J_N(\xi)\,d\xi \leq K\sqrt{N}x^{-(1/2)+\delta}, \tag{6.20}$$

and that a similar inequality holds for $\bar{J}_N$. Thus it will be sufficient to prove that $\sqrt{N}\int_{\epsilon_N}^1 |\bar{J}_N(x) - J_N(x)|\,dx$ tends to zero for $N\epsilon_N = N^{\delta/2}$.

We shall proceed by showing that $\sqrt{N}|\bar{J}_N(x) - J_N(x)|$ stays bounded by an integrable function and that $\sqrt{N}|\bar{J}_N(x) - J_N(x)| \to 0$ in measure. For the first part it is convenient to divide the range $[\epsilon_N, 1]$ into two parts $[\epsilon_N, \tau]$ and $(\tau, 1]$ with τ fixed but $\tau < 1$.

Consider first the part relative to the interval (ϵ_N, τ). Let $\beta_N(x, k)$ be the density of the k-th order statistics from a uniform sample of size N, and let

$$\beta_N(x, k) = \frac{\Gamma(N+1)}{\Gamma(k)\Gamma(N+1-k)} x^{k-1}(1-x)^{N-k} \tag{6.21}$$

even for noninteger values of the symbols N and k. Let $\xi = (k/(N+1))$ and let $I_1(\xi)$ be the integral

$$I_1(\xi) = \int_0^{\xi/4} J_N(x)\beta_N(x, k)\, dx. \tag{6.22}$$

Since $0 \leq J_N(x) \leq x^{-\alpha}$, $\alpha = \frac{1}{2} - \delta$, one can bound $I_1(\xi)$ by the integral

$$I_2(\xi) = \int_0^{\xi/4} \frac{\Gamma(N+1)\Gamma(k-\alpha)}{\Gamma(k)\Gamma[N+1-\alpha]} \beta_{N-\alpha}[x, k-\alpha]\, dx. \tag{6.23}$$

Consider also the function

$$\gamma_N(x, k-\alpha) = \frac{(N-\alpha)^{k-\alpha}}{\Gamma(k-\alpha)} x^{k-\alpha-1} e^{-(N-\alpha)x}. \tag{6.24}$$

A simple application of Stirling's formula shows that

$$\beta_N[x, k-\alpha] \leq c\left[1 - \frac{\alpha}{N}\right]^{1/2}\left[1 - \frac{k}{N}\right]^{-1/2} \gamma_N[x, k-\alpha] \tag{6.25}$$

for a certain constant c and for all values of N and k such that $N - k - \alpha \geq 1$.

In addition, $\int_0^{\xi/4} \gamma_N[x, k-\alpha]\, dx$ can be bounded by Markov's inequality as follows. If $\mu = (k-\alpha)/(N-\alpha)$, then, for $s > 0$,

$$\int e^{-2s[N-\alpha][x-\mu/2]}\gamma_N[x,\ k-\alpha]\, dx = \left[\frac{e^s}{1+2s}\right]^{k-\alpha}. \tag{6.26}$$

Therefore, $\int_0^{\mu/2} \gamma_N[x,\ k-\alpha]\, dx \leq \rho^{k-\alpha}$, with $\rho = \inf\{(e^s/(1+2s)),\ s \geq 0\} < 1$. Since $\xi/4 \leq \mu/2$ for $N \geq N_0$ and $k \geq N_0^{\delta/2}$, this implies

$$\begin{aligned} I_2(\xi) &\leq c\rho^{k-\alpha}\left(1 - \frac{\alpha}{N}\right)^{1/2}\left(1 - \frac{k}{N}\right)^{-1/2} \frac{\Gamma(N+1)\Gamma(k-\alpha)}{\Gamma(k)\Gamma(N+1-\alpha)} \\ &\leq c'\rho^k\left(\frac{N-\alpha}{k-\alpha-1}\right)^{\alpha}. \end{aligned} \tag{6.27}$$

Therefore, given $\epsilon > 0$, there exists an $N_1(\epsilon)$ such that $N \geq N_1$ and $k \geq N^{\delta/2}$ implies $\sqrt{N} I_2(\xi) \leq \epsilon\xi^{(1/2)-\delta}$ for $\epsilon_N < \xi \leq \tau$. Thus this term will become negligible. Consider the term

$$\begin{aligned} I_3(\xi) &= \sqrt{N}\int_{\xi/4}^1 [J_N(x) - J_N(\xi)]\beta_N(x, k)\, dx \\ &= \int_{\xi/4}^1 \frac{J_N(x) - J_N(\xi)}{|x-\xi|}\sqrt{N}|x-\xi|\beta_N[x, (N+1)\xi]\, dx. \end{aligned} \tag{6.28}$$

Note that

$$\sqrt{N}\int |x-\xi|\beta_N[x, (N+1)\xi]\, dx \leq \sqrt{\frac{N}{N+2}}\sqrt{\xi(1-\xi)}. \tag{6.29}$$

Furthermore, in I_3 the differential ratio involving J_N is, by assumption, smaller than the same ratio involving the function $\Omega(x) = x^{-1/2+\delta}$. For the latter function, the maximum value of the ratio is obtained at $x = \xi/4$ giving

$$I_3(\xi) \leq (\tfrac{1}{4}\xi)^{-(3/2)+\delta}\sqrt{\xi(1-\xi)} = c_1\sqrt{1-\xi}\,\xi^{-(1/2)+\delta}. \tag{6.30}$$

This is an integrable function.

To show that $I_3(\xi)$ does in fact tend to zero for almost all ξ, if J'_N converges to a function J', it is sufficient to repeat an argument similar to the argument of lemma 1. Note also that $J^*_N(\xi) = \int J_N(x)\beta_N[x, (N+1)\xi]\,dx$ provides a decreasing interpolation of the function $\overline{J}_N$. From this one concludes that $\sqrt{N}|J^*_N(\xi) - J_N(\xi)| \to 0$ for almost every ξ, and therefore $\sqrt{N}|\overline{J}_N(\xi) - J_N(\xi)| \to 0$ for almost every ξ. Furthermore, for $N \geq N_0$ one has $\sqrt{N}|\overline{J}_N(\xi) - J_N(\xi)| \leq c_2\xi^{-1/2+\delta}$ for every $\xi \geq \epsilon_N$. The result follows.

REMARK. The preceding theorem 2 corresponds to theorem 2 of Chernoff and Savage. As a particular application, let us mention the following corollary.

COROLLARY. *Let k be a fixed integer and let a_j, $j = 1, 2, \cdots, k$ be bounded constants. Let*

$$J_N\left(\frac{i}{N+1}\right) = \sum_{j=1}^{k} a_j E(\xi^j_{N,i}) \tag{6.31}$$

where $\xi^j_{N,i}$ is the i-th order statistic in a sample of size N from a population whose cumulative distribution is the inverse of a function S_j.

If $|(dS^j(x)/dx)| \leq K[x(1-x)]^{-(3/2)+\delta}$ for $j = 1, 2, \cdots, k$, then the functions J_N satisfy the conditions of theorem 1.

PROOF. This follows from the linearity of the transformation $J \rightsquigarrow \overline{J}_N$ used to define the functions which occur in theorem 2.

For the case where $a_1 = 0$, $a_2 = 1$, $k = 2$, and $S = \Phi^{-1}$ the resultant test statistic is the one considered by Capon [4] and Klotz [11].

7. The c-sample case

In this section we shall extend the results of section 6 to c-sample situations. Without additional assumptions Puri [13] had extended the Chernoff-Savage results to c-sample cases. Our theorems 3 and 4 are direct extensions of Puri's lemma 5.1 and theorem 6.1.

Let $X_{j,k}$, $k = 1, 2, \cdots, n_j$ be a random sample from a population having a continuous cumulative distribution function $F^{(j)}$. Assume that the c-samples obtained for $j = 1, 2, \cdots, c$ are independent. Let $N = \sum n_j$ and let $\lambda_j = n_j/N$. Assume that there is a $\lambda_0 > 0$ such that $\lambda_0 \leq \lambda_j \leq 1 - \lambda_0$ for every $j = 1, 2, \cdots, c$ and every N. Let $H = \sum_j \lambda_j F^{(j)}$ and $H_N = \sum_j \lambda_j F^{(j)}_{n_j}$ be respectively the combined cumulative and the combined empirical cumulative based on the samples $\{X_{j,k}\}$.

Let $T_{N,j} = n_j^{-1}\sum_{i=1}^{N} E_{N,i,j}Z_{N,i,j}$, $j = 1, 2, \cdots, c$, where $Z_{N,i,j} = 1$ if the i-th smallest observation from the combined sample of size N belongs to j-th sample and where $Z_{N,i,j}$ is equal to zero otherwise.

The $E_{N,i,j}$ are given constants. Following the notation of section 2, one can represent $T_{N,j}$ in an integral form

$$T_{N,j} = \int_{-\infty}^{+\infty} J_{N,j}\left[\frac{N}{N+1} H_N\right] dF_{n_j}^{(j)}(x). \tag{7.1}$$

For simplicity we have assumed that the functions $F^{(j)}$ are continuous and state a result analogous to theorem 1 in a form similar to the form of theorem 1 of Chernoff and Savage, for the original distribution functions.

THEOREM 3. *Assume that for all* $j = 1, 2, \cdots, c$ *the following conditions hold:*

(1) $J_j(H) = \lim J_{N,j}(H)$ *exists for* $0 < H < 1$, *and this limit is not a constant and it is absolutely continuous on* $(0, 1)$;

(2) $\int_{0<H_N\leq 1} \left\{J_{N,j}\left(\frac{N}{N+1} H_N\right) - J_j\left(\frac{N}{N+1} H_N\right)\right\} dF_{n_j}^{(j)}(x) = o_p(N^{-1/2}).$

(3) $\left|\frac{dJ_j(x)}{dx}\right| \leq K[x(1-x)]^{-(3/2)+\delta}$ with $0 < 2\delta < 1$.

Let $\mu_{N,j} = \int_{-\infty}^{+\infty} J_j[H(x)]\, dF^{(j)}(x)$ *and*

$$\begin{aligned}\sigma_{N,j}^2 = \sum_{i\neq j} 2\lambda_i \iint_{-\infty<x<y<+\infty} & F^{(i)}(x)[1 - F^{(i)}(y)]J_j'[H(x)]J_j'[H(y)] \\ & \cdot dF^{(j)}(x)dF^{(j)}(y) \\ + \frac{2}{\lambda_j} \iint_{x<y} & F^{(j)}(x)[1 - F^{(j)}(y)]J_j'[H(x)]J_j'[H(y)] \\ & \cdot d[H(x) - \lambda_j F^{(j)}(x)]\, d[H(y) - \lambda_j F^{(j)}(y)].\end{aligned} \tag{7.2}$$

Then, if $\liminf_{N\to\infty} \sigma_{N,j}^2 > 0$, *one has*

$$\lim_{N\to\infty} P\left\{\sqrt{N}[T_{N,j} - \mu_{N,j}] < t\sigma_{N,j}\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2}\, dx. \tag{7.3}$$

PROOF. Conditions (1) and (2) of the theorem imply that one may consider instead of $(T_{N,j} - \mu_{N,j})\sqrt{N}$ the expression

$$\tilde{T}_{N,j} = \sqrt{N}\left\{\int J_j\left[\frac{N}{N+1} H_N\right] dF_{n_j}^{(j)} - \int J_j(H)\, dF^{(j)}\right\}, \tag{7.4}$$

which is similar to the expression covered by theorem 1. One can proceed exactly as in theorem 1, along the following sequence of steps. First one can show that an integral of the type

$$\Delta_N[J_j, \tau] = \sqrt{N}\left\{\int_A J_j\left[\frac{N}{N+1} H_N\right] dF_{n_j}^{(j)} - \int_A J_j(H)\, dF^{(j)}\right\} \tag{7.5}$$

with $A = (0, \tau] \cup [1 - \tau, 1)$ can be made small by selecting τ small enough. Second, removing an appropriate term from J_j' (on the set A), one is left with functions J_j' which are bounded, and one shows, by an argument similar to that of proposition 3, that terms of the type

$$\sqrt{N} \int \left[J_j\left[\frac{N}{N+1} H_N\right] - J_j(H)\right] d[F_{n_i}^{(j)} - F^{(j)}] \tag{7.6}$$

can also be neglected. This replaces our $\tilde{T}_{N,j}$ by a term of the type

$$T_{N,j}^* = \sqrt{N} B_{N,j} + R_{N,j} \tag{7.7}$$

with

$$\sqrt{N} B_{N,j} = \sqrt{N} \int J_j(H)\, d(F_{n_j}^{(j)} - F^j) + \sqrt{N} \int [H_N - H] J_j'(H)\, dF^{(j)}, \tag{7.8}$$

and

$$R_{N,j} = \sqrt{N} \int [J_j(H_N) - J_j(H) - (H_N - H) J_j'(H)]\, dF. \tag{7.9}$$

An argument similar to that of theorem 1 shows that $R_{N,j}$ also tends to zero. Thus, the only nonnegligible term left is the term $\sqrt{N} B_{N,j}$, which can be written as sums of independent variables. The result is then obtainable by appropriate algebra and the central limit theorem.

To proceed through these steps, the necessary tools are the appropriate versions of lemmas 8 and 9 and propositions 2 and 3. Both lemmas 8 and 9 are proved there by substituting to binomial variables appropriate Poisson variables. This is still possible here. The difference in probabilities will be less than $2 \sum_{j=1}^{c} p_j$ where p_j is the probability attached to a set $(0, \epsilon] \cup [1 - \epsilon, 1)$ by the measure $F^{(j)}$ (reduced to the interval $(0, 1)$ as before). Thus, in the tails, $\sqrt{N}\, [H_N - H]$ and H_N will still behave essentially as if they were obtained by taking N observations from the uniform distribution. Proposition 2 involves an interpolation which is feasible simply because none of the $\sqrt{N}[F_{n_j}^{(j)} - F^{(j)}]$ oscillates much on intervals which have small probability for the parent distribution. Since the derivatives $[dF^{(j)}/dH] = \varphi_j$ are still bounded, this is possible. Proposition 3 depends only on the behavior of $\sqrt{N}\, \|H_N - H\|$ and $\sqrt{N}\, \|F_{n_j}^{(j)} - F^{(j)}\|$. Hence it is still valid here. One could also extend proposition 4 to the present case, since its proof depends only on the validity of proposition 2 and on the fact that the interpolation formula of proposition 2, when applied to H_N itself, gives functions whose slope is arbitrarily close to unity. However, proposition 4 is not even needed for the proof of theorem 1.

Since in the present case one may have to consider the joint distribution of the statistics $T_{N,j}$, $j = 1, 2, \cdots, c$, it appears proper to mention that the random vector

$$T_N = \sqrt{N}\{(T_{N,1} - \mu_{N,1}), (T_{N,2} - \mu_{N,2}), \cdots, (T_{N,c} - \mu_{N,c})\} \tag{7.10}$$

has a joint limiting normal distribution, provided that the relevant covariances converge. This is the purpose of the following theorem.

THEOREM 4. *Let assumptions* (1), (2), *and* (3) *of theorem* 3 *be satisfied, and let* T_N *be the random vector* $T_N = \{\sqrt{N}(T_{N,j} - \mu_{N,j});\ j = 1, 2, \cdots, c\}$. *Let* P_N *be the distribution of* T_N *and let* Q_N *be the normal distribution which has expectation zero and covariance matrix* Γ_N. *Then* $\|P_N - Q_N\|_{BL} \to 0$ *provided* Γ_N *be given by* $\Gamma_N = ((\sigma_{N,i,j}))$ *with* $\sigma^2_{N,j}$ *equal to the quantity used in theorem* 3 *and*

$$
\begin{aligned}
(7.11)\qquad \sigma_{N,i,j} = \sum_{\substack{k=1\\k\neq i,j}}^{c} \lambda_k \iint_{x<y} & F^{(k)}(x)[1 - F^{(k)}(y)]J'_i(H(x))J'_j(H(y)) \\
&\qquad \cdot dF^{(i)}(x)\, dF^{(j)}(y) \\
+ \iint_{x<y} & F^{(k)}(x)[1 - F^{(k)}(y)]J'_i(H(y))J'_j(H(x)) \\
&\qquad \cdot dF^{(i)}(y)\, dF^{(j)}(x) \\
- \iint_{x<y} & F^{(i)}(x)[1 - F^{(i)}(y)]J'_i(H(y))J'_j(H(x)) \\
&\qquad \cdot dF^{(j)}(x)\, d[H(y) - \lambda_i F^{(i)}(y)] \\
- \iint_{x<y} & F^{(i)}(x)[1 - F^{(i)}(y)]J'_i(H(x))J'_j(H(y)) \\
&\qquad \cdot dF^{(j)}(y)\, d[H(x) - \lambda_i F^{(i)}(x)] \\
- \iint_{x<y} & F^{(j)}(x)[1 - F^{(j)}(y)]J'_i(H(x))J'_j(H(y)) \\
&\qquad \cdot dF^{(i)}(x)\, d[H(y) - \lambda_j F^{(j)}(y)] \\
- \iint_{x<y} & F^{(j)}(x)[1 - F^{(j)}(y)]J'_i(H(y))J'_j(H(x)) \\
&\qquad \cdot dF^{(i)}(y)\, d[H(x) - \lambda_j F^{(j)}(x)].
\end{aligned}
$$

PROOF. The argument of theorem 3 shows that each term $\sqrt{N}\,[T_{N,j} - \mu_{N,j}]$ is asymptotically equivalent to a term $\sqrt{N}B_{N,j}$ defined by

$$
(7.12)\qquad \sqrt{N}B_{N,j} = \sqrt{N}\int J_j(H)\, d(F_{n_j} - F) + \sqrt{N}\int (H_N - H)J'_j(H)\, dF.
$$

Let L_j be the function $L_j(x) = \int_{1/2}^{x} J'_j(\xi)\, dF^j(\xi)$. Integrate by parts and separate the components of $[H_N - H]$. This gives

$$
\begin{aligned}
(7.13)\qquad \sqrt{N}B_{N,j} &= \sqrt{N}\int [J_j - \lambda_j L_j]\, d[F^{(j)}_{n_j} - F^{(j)}] \\
&\quad - \sqrt{N}\sum_{k\neq j}\int \lambda_k L_j\, d(F^{(k)}_{n_k}).
\end{aligned}
$$

This can also be written in the form

$$
\begin{aligned}
(7.14)\qquad \sqrt{N}B_{N,j} &= \frac{1}{\sqrt{\lambda_j}}\frac{1}{\sqrt{n_j}}\sum_{\nu}^{n_j} [J_j[X_{j,\nu}] - EJ_j(X_{j,\nu})] \\
&\quad - \sum_k \sqrt{\lambda_k}\sum_{\nu=1}^{n_k} \{L_j[X_{k,\nu}] - EL_j[X_{k,\nu}]\}.
\end{aligned}
$$

The central limit theorem applies to the sums, giving the result of theorem 4 upon evaluation of the covariance matrix. The explicit form given in the statement of the theorem is obtained by writing $\sqrt{N}B_{N,j}$ in still another form as follows.

Let $W_{N,j}$ be the process $W_{N,j} = \sqrt{n_j}\,[F_{n_j}^{(j)} - F^{(j)}]$ which has expectation zero and a covariance function

$$C_j(s, t) = F^{(j)}(s)[1 - F^{(j)}(t)] \tag{7.15}$$

for $s \leq t$. Then, the expression $\sqrt{N}B_{N,j}$ becomes

$$\sqrt{N}B_{N,j} = \sum_k \sqrt{\lambda_k} \int W_{N,k}(x)J_j'[H(x)]\, dH(x) - \frac{1}{\sqrt{\lambda_j}} \int W_{N,j}(x)J_j'[H(x)]\, dH(x). \tag{7.16}$$

The formula is obtained by taking the expectation of the product $\sqrt{N}B_{N,i}\sqrt{N}B_{N,j}$ in this form and using the fact that $EW_{N,k}(x)W_{N,\nu}(y) = 0$ for $k \neq \nu$, since then the processes $W_{N,k}$ and $W_{N,\nu}$ are independent.

Let us also note the following result.

THEOREM 5. *The conclusion of theorems* 3 *and* 4 *is still valid if the assumption* (3) *of theorem* 3 *is replaced by the condition that for each value of* $j = 1, 2, \cdots, c$ *the function* J_j' *is of the form* $J_{j,1}' + J_{j,2}'$ *with* $\int |J_{j,1}'(x)|\, dx < \infty$ *and* $|J_{j,2}'| \leq fg$ *with* $f \in \mathfrak{U}_1$ *and* $g \in \mathfrak{U}_2$. *The uniformity statements of theorem* 1 *extend to this case.*

PROOF. The proof is the same as that of theorem 3, with uniformity of the convergence depending on Lipschitz-type conditions as detailed in the proof of theorem 1.

Finally, let us note that the remark concerning location and scale families which follows theorem 1 is still applicable here and that theorem 2 provides a class of functions J_N which satisfy the assumptions (1) and (2) of theorem 3. The arguments which lead to this last statement do not in any way depend on the observations but only on properties of order statistics of the uniform distribution.

◇ ◇ ◇ ◇ ◇

ACKNOWLEDGMENTS

This paper was prepared during the Spring 1965 while the three authors were gathered at the University of California, Berkeley. The brief mention of functions of bounded variation (in proposition 5) was suggested by a remark of H. Chernoff.

In preparing the present version, one of the authors had the benefit of several conversations with J. Hájek to whom we owe several suggestions and several references. In addition, Hájek has informed one of us that he obtained recently a bound on the variances of Chernoff-Savage statistics which would allow drastic simplifications in the proofs of theorems such as theorem 1 and may also lead to extensions of its domain of validity.

REFERENCES

[1] A. R. ANSARI and R. A. BRADLEY, "Rank-sum tests for dispersions," *Ann. Math. Statist.*, Vol. 31 (1960), pp. 1174–1189.

[2] P. BICKEL, "Some contributions to the theory of order statistics," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley and Los Angeles, 1966, Vol. I, pp. 575–591.

[3] Z. W. BIRNBAUM and A. W. MARSHALL, "Some multivariate Chebyshev inequalities with extensions to continuous parameter processes," *Ann. Math. Statist.*, Vol. 32 (1961), pp. 687–703.

[4] JACK CAPON, "Asymptotic efficiency of certain locally most powerful rank tests," *Ann. Math. Statist.*, Vol. 32 (1961), pp. 88–100.

[5] HERMANN CHERNOFF and I. R. SAVAGE, "Asymptotic normality and efficiency of certain non-parametric test statistics," *Ann. Math. Statist.*, Vol. 29 (1958), pp. 972–994.

[6] D. M. ČIBISOV, "Some theorems on the limiting behavior of an empirical distribution functions," *Trudy Mat. Inst. Steklov.*, Vol. 71 (1964), pp. 104–112.

[7] MONROE D. DONSKER, "Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems," *Ann. Math. Statist.*, Vol. 23 (1952), pp. 277–281.

[8] R. M. DUDLEY, "Convergence of Baire measures," to appear in *Studia Math.*

[9] J. HÁJEK, "Extension of the Kolmogorov-Smirnov test to regression alternatives," *Bernoulli-Bayes-Laplace Anniversary Volume*, Berlin-Heidelberg-New York, Springer Verlag, 1965.

[10] J. HÁJEK and A. RÉNYI, "Generalization of an inequality of Kolmogorov," *Acta Math. Acad. Sci. Hungar.*, Vol. 6 (1955), pp. 281–283.

[11] JEROME KLOTZ, "Nonparametric tests for scale," *Ann. Math. Statist.*, Vol. 33 (1962), pp. 498–512.

[12] YU. V. PROHOROV, "Asymptotic behavior of the binomial distribution," *Uspehi. Mat. Nauk SSSR*, Vol. 8 (1953), pp. 135–142.

[13] MADAN LAL PURI, "Asymptotic efficiency of a class of *c*-sample tests," *Ann. Math. Statist.*, Vol. 35 (1964), pp. 102–121.

[14] BENGT ROSEN, "Limit theorems for sampling from finite populations," *Ark. Mat.*, Vol. 5 (1965), pp. 383–424.

[15] SIDNEY SIEGEL and JOHN W. TUKEY, "A nonparametric sum of ranks procedure for relative spread in unpaired samples," *J. Amer. Statist. Assoc.*, Vol. 55 (1960), pp. 429–444. (Errata, *ibid.* (1961), p. 1005.)

[16] V. STRASSEN, "The existence of probability measures with given marginals," *Ann. Math. Statist.*, Vol. 36 (1965), pp. 423–439.