*The Annals of Statistics*
1975, Vol. 3, No. 4, 825–845

# ON CHERNOFF–SAVAGE STATISTICS AND SEQUENTIAL RANK TESTS[1]

BY TZE LEUNG LAI

*Columbia University*

In this paper, we shall represent a generalized Chernoff–Savage statistic as the sum of i.i.d. random variables plus a remainder term and analyze the order of magnitude of the remainder term. While Chernoff and Savage have proved that the remainder term, when suitably normalized, converges to 0 in probability, we obtain a stronger form of convergence in this paper. Our result gives an invariance principle and a law of the iterated logarithm for generalized Chernoff–Savage statistics. We also use our result to obtain asymptotic approximations for the stopping rules of certain sequential rank tests.

**1. Introduction.** In [4], to prove the asymptotic normality of a class of linear rank statistics, Chernoff and Savage have expressed this kind of statistics as the sum of i.i.d. random variables plus a remainder term and have demonstrated that the remainder term, when normalized by an appropriate factor, converges to zero in probability. In certain applications involving linear rank statistics, however, we need to have a stronger result concerning the order of magnitude of the normalized remainder term than simply convergence to zero in probability. In Section 4 below, we shall prove a stronger form of convergence which we shall need in the study of sequential rank tests in Section 5. As an immediate corollary of our result, we also obtain an invariance principle and a law of the iterated logarithm for generalized Chernoff–Savage statistics. To prove our representation theorem in Section 4, certain results concerning the large deviation probability for the tails of the empirical distribution function will be needed. Section 2 deals with this problem of large deviation probabilities.

In Section 5, we shall study the stopping times of certain sequential rank tests. Suppose $X_1, X_2, \cdots$ are i.i.d. with a continuous distribution function $F$, and are independent of $Y_1, Y_2, \cdots$ which are i.i.d. with a continuous distribution function $G$. In [13], Savage and Sethuraman have examined the rank-order sequential probability ratio test of the null hypothesis $H_0: F = G$ versus the Lehmann alternative $H_1: F = G^A$ where $0 < A \neq 1$ is a known constant. Let $F_n(x) = n^{-1} \sum_1^n I_{[X_i \leq x]}$, $G_n(x) = n^{-1} \sum_1^n I_{[Y_i \leq x]}$ and $W_n(x) = F_n(x) + AG_n(x)$.

Received November 1973; revised December 1974.

[1] Research supported by the Public Health Service under Grant Contract No. 5-R01-GM-16895-05.

*AMS* 1970 *subject classifications*. Primary 62E20, 62L10, 62G10.

*Key words and phrases*. Chernoff–Savage theorem, sequential rank tests, Lehmann alternatives, Wilcoxon tests, empirical distribution function, large deviation probabilities, invariance principle, last time.

Define

$$(1.1)\qquad \begin{aligned} l_n &= \log((2n)!/n^{2n}) - \textstyle\sum_1^n \{\log W_n(X_i) + \log W_n(Y_i) - \log A\} \\ &= -\textstyle\sum_1^n \{\log W_n(X_i) + \log W_n(Y_i) - \log 4A + 2\} + \frac{1}{2}\log n + O(1)\,. \end{aligned}$$

The rank-order SPRT stops at stage

$$(1.2)\qquad N = \inf\{n \geqq 1 : l_n \notin (-a, b)\} \qquad (a, b > 0)$$

(cf. [13]). In [13], Savage and Sethuraman have shown that given $\varepsilon > 0$, there exists $0 < \rho < 1$ such that

$$(1.3)\qquad P[|n^{-1}l_n - S(A, F, G)| \geqq \varepsilon] = O(\rho^n)$$

where

$$(1.4)\qquad S(A, F, G) = \log 4A - 2 - \int \log(F(x) + AG(x))(dF(x) + dG(x))\,.$$

From (1.3), it is easy to see that if $S(A, F, G) \neq 0$, then $Ee^{tN} < \infty$ for $t \leqq \theta$ $(\theta > 0)$ and as $\min(a, b) \to \infty$,

$$(1.5)\qquad \begin{aligned} EN^\beta &\sim (b/S(A, F, G))^\beta \qquad \text{if} \quad S(A, F, G) > 0\,; \\ EN^\beta &\sim (a/|S(A, F, G)|)^\beta \qquad \text{if} \quad S(A, F, G) < 0 \end{aligned}$$

for any $\beta > 0$ (cf. [1] for the case $\beta = 1$). The situation in the case $S(A, F, G) = 0$ is much harder. Sethuraman [15] has shown that the stopping time $N$ still remains exponentially bounded in this case. In Section 5, by making use of our results in Section 4, we find the asymptotic distribution and the asymptotic moments of $N$ when $S(A, F, G) = 0$. We shall also examine a sequential two-sample Wilcoxon test in Section 5.

**2. Large deviation probabilities for the tails of the empirical distribution function.** Let $X_1, X_2, \cdots$ be i.i.d. random variables with a common continuous distribution function $F$. Let $F_n(x) = n^{-1}\sum_1^n I_{[X_i \leqq x]}$ denote the empirical distribution function. Large deviation probabilities for the Kolmogorov–Smirnov statistic $||F_n - F|| = \max_x |F_n(x) - F(x)|$ are well known; in fact, Dvoretzky, Kiefer and Wolfowitz [6] have proved that there exists a universal constant $C$ such that

$$(2.1)\qquad P[n^{\frac{1}{2}}||F_n - F|| \geqq u] \leqq Ce^{-2u^2}\,, \qquad n = 1, 2, \cdots, u \geqq 0\,.$$

The following theorem deals with certain large deviation probabilities for the tails of the empirical distribution function, which will be useful in the analysis of linear rank statistics.

THEOREM 1. (i) *Given any* $c > 0$, $0 < \alpha < 1$, *there exist positive constants* $k_1, k_2$ *such that for all* $u \geqq 1$, $n = 1, 2, \cdots$

$$(2.2)\qquad \begin{aligned} &P[\max_{F_n(x) \leqq cn^{-\alpha}} |F_n(x) - F(x)| \geqq n^{-(1+\alpha)/2}u] \\ &\quad + P[\max_{F_n(x) \geqq 1-cn^{-\alpha}} |F_n(x) - F(x)| \geqq n^{-(1+\alpha)/2}u] \leqq k_1 e^{-k_2 u}\,; \end{aligned}$$

$$(2.3)\qquad \begin{aligned} &P[\max_{F(x) \leqq cn^{-\alpha}} |F_n(x) - F(x)| \geqq n^{-(1+\alpha)/2}u] \\ &\quad + P[\max_{F(x) \geqq 1-cn^{-\alpha}} |F_n(x) - F(x)| \geqq n^{-(1+\alpha)/2}u] \leqq k_1 e^{-k_2 u}\,. \end{aligned}$$

(ii) *Given any $c > 0$, $0 < \alpha < 1$ and $\delta > 1$, there exists a positive constant $\lambda$ such that*

$$(2.4)\qquad \begin{aligned} &P[\max\nolimits_{F_n(x) \geqq cn^{-\alpha}} F(x)/F_n(x) > \delta] \\ &\quad + P[\max\nolimits_{F_n(x) \leqq 1-cn^{-\alpha}} (1 - F(x))/(1 - F_n(x)) > \delta] \\ &\quad = O(\exp(-\lambda n^{1-\alpha}))\,; \end{aligned}$$

$$(2.5)\qquad \begin{aligned} &P[\max\nolimits_{F(x) \geqq cn^{-\alpha}} F(x)/F_n(x) > \delta] \\ &\quad + P[\max\nolimits_{F(x) \leqq 1-cn^{-\alpha}} (1 - F(x))/(1 - F_n(x)) > \delta] \\ &\quad = O(\exp(-\lambda n^{(1-\alpha)/2}))\,. \end{aligned}$$

(iii) *Given $c > 0$, $1 \geqq \alpha > 0$ and $\beta > 0$, there exists $\lambda > 0$ such that letting $\gamma = \min\{2\beta, \beta + \frac{1}{2}(1 - \alpha)\}$, we have*

$$(2.6)\qquad P[|F_n F^{-1}(cn^{-\alpha}) - cn^{-\alpha}| \geqq n^{\beta+\frac{1}{2}(1-\alpha)-1}] = O(\exp(-\lambda n^{\gamma}))$$

$$(2.7)\qquad P[|F_n F^{-1}(1 - cn^{-\alpha}) - (1 - cn^{-\alpha})| \geqq n^{\beta+\frac{1}{2}(1-\alpha)-1}] = O(\exp(-\lambda n^{\gamma}))$$

*where $F^{-1}(t)$ can be taken to be any number $x$ such that $F(x) = t$.* (In our applications below, we sometimes take $F^{-1}(t)$ to be $\sup\{x : F(x) = t\}$, and at other times take $F^{-1}(t)$ to be $\inf\{x : F(x) = t\}$.)

(iv) *Given any $c > 0$, $1 \geqq \alpha > 0$, $\beta > 0$ and $\delta > \alpha - \beta$,*

$$(2.8)\qquad \begin{aligned} &P[\max\nolimits_{1 \leqq n \leqq m} n^{\delta} F_n F^{-1}(cn^{-\alpha}) \geqq m^{\delta-\alpha+\beta}] \\ &\quad + P[\max\nolimits_{1 \leqq n \leqq m} n^{\delta}\{1 - F_n F^{-1}(1 - cn^{-\alpha})\} \geqq m^{\delta-\alpha+\beta}] \\ &\quad = o(\exp(-m^{\beta})) \qquad \text{if} \quad \delta \geqq \alpha \\ &\quad = o(\exp(-m^{\delta-\alpha+\beta})) \qquad \text{if} \quad \delta < \alpha\,. \end{aligned}$$

PROOF. To prove (2.2), since $\max_{F_n(x) \geqq 1-cn^{-\alpha}} |F_n(x) - F(x)|$ has the same distribution as $\max_{F_n(x) \leqq cn^{-\alpha}} |F_n(x) - F(x)|$, it suffices to consider only the lower tail of the empirical distribution function. The same remark also applies to the other parts of Theorem 1. Since $F$ is continuous, we can write

$$\max\nolimits_{F_n(x) \leqq cn^{-\alpha}} |F_n(x) - F(x)| \leqq \max\nolimits_{k \leqq cn^{1-\alpha}} |U_k^{(n)} - (k/n)| + n^{-1}$$

where $U_1^{(n)}, \cdots, U_n^{(n)}$ are the order statistics of the uniform distribution (cf. [2], page 285). Let $W_1, W_2, \cdots$ be i.i.d. random variables having the negative exponential distribution with mean 1, and let $S_n = W_1 + \cdots + W_n$. Since $U_k^{(n)}$, $k = 1, \cdots, n$, have the same joint distribution as $S_k/S_{n+1}$, $k = 1, \cdots, n$, we obtain that for $u \geqq 1$ and $n \geqq n_0$,

$$\begin{aligned} &P[\max\nolimits_{F_n(x) \leqq cn^{-\alpha}} |F_n(x) - F(x)| \geqq n^{-\frac{1}{2}(1+\alpha)}u] \\ &\quad \leqq P[(1/S_{n+1}) \max\nolimits_{k \leqq cn^{1-\alpha}} |(S_k - k) - (k/n)(S_{n+1} - n)| \geqq \tfrac{1}{2}n^{-\frac{1}{2}(1+\alpha)}u] \\ &\quad \leqq P[S_{n+1} < \tfrac{1}{2}n] + P[\max\nolimits_{k \leqq cn^{1-\alpha}} |S_k - k| \geqq \tfrac{1}{8}n^{\frac{1}{2}(1-\alpha)}u] \\ &\qquad + P[|S_{n+1} - n| \geqq (8c)^{-1}n^{\frac{1}{2}(1+\alpha)}u] = A_n + B_n + C_n\,, \quad \text{say.} \end{aligned}$$

By a theorem of Chernoff [3], $A_n \leqq k_1 \exp(-k_2 n)$ for some $k_1, k_2 > 0$.

To give an upper bound for $B_n$, let $m = [cn^{1-\alpha}]$. Then since $\{\exp(\theta(S_k - k)),$

$k = 1, \cdots, m\}$ is a submartingale for $0 < \theta < 1$, it follows from the submartingale inequality that for $\varepsilon > 0$,

$$P[\max_{k \leqq m} (S_k - k) \geqq \varepsilon] \leqq e^{-\theta\varepsilon} E \exp(\theta(S_m - m)) = \exp\{-\theta\varepsilon - m(\theta + \log(1-\theta))\}. \tag{2.9}$$

Now $|\theta + \log(1-\theta)| \leqq \theta^2$ for $|\theta| \leqq \theta_0$. Hence setting $\theta = m^{-\frac{1}{2}}$ and $\varepsilon = \frac{1}{8} n^{\frac{1}{2}(1-\alpha)} u$ in (2.9), we obtain for $n \geqq n_1 \geqq n_0$ and $u \geqq 1$ that for some $k_1, k_2 > 0$,

$$P[\max_{k \leqq cn^{1-\alpha}} (S_k - k) \geqq \varepsilon] \leqq \tfrac{1}{2} k_1 e^{-k_2 u}.$$

Replacing $S_k - k$ by $-(S_k - k)$ in the above argument, we can easily see that $B_n \leqq k_1 \exp(-k_2 u)$ for $n \geqq n_1$. In a similar way, we can show that $C_n \leqq k_1 \exp(-k_2 u)$ for $n \geqq n_1$. Hence there exist $k_1, k_2 > 0$ such that for $n \geqq n_1$ and $n \geqq u \geqq 1$,

$$A_n + B_n + C_n \leqq k_1 e^{-k_2 n} + 2k_1 e^{-k_2 u} \leqq 3k_1 e^{-k_2 u}.$$

If $u \geqq n$, then it follows from (2.1) that

$$\begin{aligned} P[\|F_n - F\| \geqq n^{-\frac{1}{2}(1+\alpha)} u] &\leqq P[\|F_n - F\| \geqq n^{-\frac{1}{2}\alpha} u^{\frac{1}{2}}] \\ &\leqq P[n^{\frac{1}{2}} \|F_n - F\| \geqq u^{\frac{1}{2}}] \leqq Ce^{-2u}. \end{aligned}$$

Therefore we have proved (2.2) for $n \geqq n_1$. By (2.1), we can choose $k_1, k_2$ such that (2.2) also holds for $1 \leqq n \leqq n_1$.

To prove (2.4), we let $\delta = \delta_1 \delta_2$ with $\delta_1 > 1$, $\delta_2 > 1$ and note that by Chernoff's theorem,

$$\begin{aligned} &P[\max_{F_n(x) \geqq cn^{-\alpha}} F(x)/F_n(x) > \delta] \\ &\quad \leqq P[U_k^{(n)} > \delta(k-1)/n \quad \text{for some} \quad k \geqq cn^{1-\alpha}] \\ &\quad = P[S_k/S_{n+1} > \delta(k-1)/n \quad \text{for some} \quad k \geqq cn^{1-\alpha}] \\ &\quad \leqq P[S_{n+1} < n/\delta_1] + P[S_k > \delta_2(k-1) \quad \text{for some} \quad k \geqq cn^{1-\alpha}] \\ &\quad \leqq \zeta(\rho^n + \textstyle\sum_{k \geqq cn^{1-\alpha}} \rho^k) \qquad \text{for some} \quad 0 < \rho < 1 \\ &\quad = O(\exp(-\lambda n^{1-\alpha})) \qquad \text{for some} \quad \lambda > 0. \end{aligned}$$

To prove (2.5), we use (2.2) and (2.4) to obtain that

$$\begin{aligned} &P[\max_{F(x) \geqq cn^{-\alpha}} F(x)/F_n(x) > \delta] \\ &\quad \leqq P[\max_{F_n(x) \leqq \frac{1}{2}cn^{-\alpha}} |F_n(x) - F(x)| \geqq \tfrac{1}{2}cn^{-\alpha}] \\ &\qquad + P[\max_{F_n(x) \leqq \frac{1}{2}cn^{-\alpha}} |F_n(x) - F(x)| < \tfrac{1}{2}cn^{-\alpha}, \\ &\qquad\quad \max_{F(x) \geqq cn^{-\alpha}} F(x)/F_n(x) > \delta] \\ &\quad = O(\exp(-\lambda n^{\frac{1}{2}(1-\alpha)})) + P[\max_{F_n(x) \geqq \frac{1}{2}cn^{-\alpha}} F(x)/F_n(x) > \delta] \\ &\quad = O(\exp(-\lambda n^{\frac{1}{2}(1-\alpha)})) \qquad \text{for some} \quad \lambda > 0. \end{aligned}$$

We now prove (2.8) by making use of Bernstein's inequality (cf. [17], pages 204–205). We note that

$$\textstyle\sum_{n=1}^m P[|nF_n F^{-1}(cn^{-\alpha}) - cn^{1-\alpha}| \geqq m^{\delta-\alpha+\beta} n^{1-\delta}] \leqq 2 \sum_{n=1}^m \exp(-h_n)$$

where $h_n = m^{2(\delta-\alpha+\beta)}n^{2(1-\delta)}/\{2[cn^{1-\alpha}(1-cn^{-\alpha}) + \frac{1}{3}m^{\delta-\alpha+\beta}n^{1-\delta}\max(cn^{-\alpha}, 1-cn^{-\alpha})]\}$. For $\delta \geqq \alpha$, $m^{2(\delta-\alpha+\beta)}n^{2(1-\delta)}/n^{1-\alpha} \geqq m^{2\beta}$ and so the desired conclusion follows. For $\delta < \alpha$, we have $1-\delta > 1-\alpha \geqq 0$ and the desired conclusion is obvious. Likewise using Bernstein's inequality, we can prove (2.6) and (2.7).

It remains to prove (2.3). Without loss of generality, we can assume that $F$ is the distribution function of the uniform distribution on $[0, 1]$. Let $X_1, X_2, \cdots$ be i.i.d. uniform random variables and let $X_i(t) = I_{[X_i \leqq t]}$, $t \in [0, 1)$. Then $\{(X_i(t)-t)/(1-t), 0 \leqq t < 1\}$ is a martingale (cf. [8], page 7) and so $\{\sum_1^n (X_i(t)-t)/(1-t), 0 \leqq t < 1\}$ is also a martingale. Set $\varepsilon = n^{\frac{1}{2}(1-\alpha)}u$, $\theta = \frac{1}{2}n^{-\frac{1}{2}(1-\alpha)}$. Then using the submartingale inequality, we obtain that for $n \geqq n_0$,

$$\begin{aligned} P[\max_{t \leqq cn^{-\alpha}} |F_n(t) - t| \geqq n^{-\frac{1}{2}(1+\alpha)}u] &\leqq P[\max_{t \leqq cn^{-\alpha}} |\textstyle\sum_1^n (X_i(t)-t)/(1-t)| \geqq \varepsilon] \\ &\leqq e^{-\theta\varepsilon} E\exp(\theta|\textstyle\sum_1^n (X_i(cn^{-\alpha}) - cn^{-\alpha})/(1-cn^{-\alpha})|) \\ &\leqq e^{-\frac{1}{2}u} E\exp(2\theta|\textstyle\sum_1^n (X_i(cn^{-\alpha}) - cn^{-\alpha})|) \,. \end{aligned}$$

We note that by Bernstein's inequality,

$$\begin{aligned} E\exp(2\theta|\textstyle\sum_1^n (X_i(cn^{-\alpha}) - cn^{-\alpha})|) &= 1 + \int_0^\infty e^x P[2\theta|\textstyle\sum_1^n (X_i(cn^{-\alpha}) - cn^{-\alpha})| \geqq x]\, dx \\ &= 1 + \int_0^\infty e^x P[|\textstyle\sum_1^n (X_i(cn^{-\alpha}) - cn^{-\alpha})| \geqq xn^{\frac{1}{2}(1-\alpha)}]\, dx \\ &\leqq 1 + 2\int_0^\infty e^x \exp(-g_n(x))\, dx \end{aligned}$$

where

$$\begin{aligned} g_n(x) &= x^2 n^{1-\alpha}/\{2[cn^{1-\alpha}(1-cn^{-\alpha}) + \tfrac{1}{3}xn^{\frac{1}{2}(1-\alpha)}\max(cn^{-\alpha}, 1-cn^{-\alpha})]\} \\ &\geqq 2x \qquad \text{for} \quad x \geqq x_0 \quad \text{and} \quad n \geqq n_1 \geqq n_0 \,. \end{aligned}$$

Therefore (2.3) holds for $n \geqq n_1$. In view of (2.1), we can choose $k_2$ such that (2.3) also holds for $1 \leqq n \leqq n_1$.

**3. Some preliminary lemmas.** Suppose $X_1, X_2, \cdots$ are i.i.d. random variables with a common continuous distribution function $F$ and are independent of $Y_1, Y_2, \cdots$ which are i.i.d. with a common distribution function $G$. Let $F_n(x) = n^{-1}\sum_1^n I_{[X_i \leqq x]}$, $G_m(x) = m^{-1}\sum_1^m I_{[Y_i \leqq x]}$ be the empirical distribution functions. In this section, we shall prove some lemmas which we shall use in Section 4 below.

LEMMA 1. *Suppose $u_n : R \to R$ satisfies $\max_x |u_n(x)| \leqq K_n < \infty$ and*

$$U_{m,n} = \int_{-\infty}^\infty (G_m(x) - G(x))u_n(x)\,d(F_n(x) - F(x)) \,. \tag{3.1}$$

*Then for any $p \geqq 1$, there exists an absolute constant $A_p > 0$ depending only on $p$ such that*

$$E|U_{m,n}|^{2p} \leqq A_p K_n^{2p}(mn)^{-p} \,. \tag{3.2}$$

*Consequently, if $K_n = O(n^\theta)$ for some $\theta \geqq 0$ and $(m_n)$ is a sequence of positive integers such that* $\liminf_{n\to\infty} n^{-1}m_n > 0$, *then given any $\zeta > \theta - 1$,*

$$P[|U_{m_n,n}| > n^\zeta] = o(n^{-p}) \qquad \textit{for all } p > 0 \,. \tag{3.3}$$

PROOF. Set $g_{m,n}(x) = (G_m(x) - G(x))u_n(x)$ and $\lambda_{m,n} = \int g_{m,n}(x)\,dF(x) = E(g_{m,n}(X_1) \mid Y_1, \cdots, Y_m)$. Then $U_{m,n} = n^{-1}\sum_1^n (g_{m,n}(X_i) - \lambda_{m,n})$, and so the conditional distribution of $U_{m,n}$ given $(Y_1, \cdots, Y_m)$ is that of the average of i.i.d. random variables with mean 0. Hence by the Marcinkiewicz-Zygmund inequality (cf. [10]), there exists a universal constant $C_p > 0$ such that

$$\begin{aligned} E(|U_{m,n}|^{2p} \mid Y_1, \cdots, Y_m) &\leqq n^{-2p}C_p E\{\textstyle\sum_1^n (g_{m,n}(X_i) - \lambda_{m,n})^2 \mid Y_1, \cdots, Y_m\}^p \\ &\leqq n^{-2p}C_p(4n||g_{m,n}||^2)^p \end{aligned}$$

where $||g_{m,n}|| = \max_x |g_{m,n}(x)|$. Therefore

$$E|U_{m,n}|^{2p} \leqq 4^p C_p n^{-p} E||g_{m,n}||^{2p} .$$

Now $||g_{m,n}|| \leqq K_n||G_m - G||$. By (2.1), there exists an absolute constant $B_p > 0$ such that $E(m^{\frac{1}{2}}||G_m - G||)^{2p} \leqq B_p$. Hence we have proved (3.2), and (3.3) follows easily from (3.2) and the Markov inequality.

LEMMA 2. *Let $H = \frac{1}{2}(F + G)$, and let $0 < \alpha < 1$, $\tau > 0$ and $\eta > 1$. Let $(m_n)$ be a sequence of positive integers satisfying $\liminf_{n\to\infty} n^{-1}m_n > 0$.*

(i) *There exist positive constants c and d such that*

$$\begin{aligned} &P[\max_{H(x)\geqq n^{-\alpha}} H(x)/\max(F_n(x), G_{m_n}(x)) > \eta] \\ (3.4)\qquad &\quad + P[\max_{H(x)\leqq 1-n^{-\alpha}} (1 - H(x))/\max(1 - F_n(x), 1 - G_{m_n}(x)) > \eta] \\ &\quad = O(\exp(-cn^{\frac{1}{2}(1-\alpha)})\,; \end{aligned}$$

$$\begin{aligned} (3.5)\qquad &P[\max_{H(x)\leqq n^{-\alpha}} |F_n(x) - F(x)| \geqq n^{-\frac{1}{2}(1+\alpha)+\tau}] \\ &\quad + P[\max_{H(x)\geqq 1-n^{-\alpha}} |F_n(x) - F(x)| \geqq n^{-\frac{1}{2}(1+\alpha)+\tau}] = O(\exp(-dn^{\tau})) . \end{aligned}$$

(ii) *Given any $\lambda > \alpha - \tau$,*

$$\begin{aligned} &P[\max_{1\leqq n\leqq m} n^{\lambda} F_n H^{-1}(n^{-\alpha}) \geqq m^{\lambda-\alpha+\tau}] \\ (3.6)\qquad &\quad + P[\max_{1\leqq n\leqq m} n^{\lambda}\{1 - F_n H^{-1}(1 - n^{-\alpha})\} \geqq m^{\lambda-\alpha+\tau}] \\ &\quad = o(\exp(-m^{\tau})) \qquad \text{if } \lambda \geqq \alpha \\ &\quad = o(\exp(-m^{\lambda-\alpha+\tau})) \qquad \text{if } \lambda < \alpha . \end{aligned}$$

PROOF. We note that $H \leqq \max(F, G)$, $1 - H \leqq \max(1 - F, 1 - G)$, and so (3.4) follows easily from (2.5). Since $F \leqq 2H$ and $1 - F \leqq 2(1 - H)$, (2.3) implies (3.5). From the relation $FH^{-1}(t) + GH^{-1}(t) = 2t$, it follows that $FH^{-1}(t) \leqq 2t$ and $1 - FH^{-1}(t) \leqq 2(1 - t)$. Hence we can make use of Bernstein's inequality to prove (3.6) in the same way as our proof of (2.8).

LEMMA 3. *Let $Z_1, Z_2, \cdots$ be any sequence of random variables. For any $\varepsilon > 0$ and any real number $\zeta$, set $\tau(\zeta, \varepsilon) = \sup\{n \geqq 1 : |Z_n| \geqq \varepsilon n^{\zeta}\}$ $(\sup \varnothing = 0)$. Let $\alpha > 0$ and $p > 0$.*

(i) *If $\sum_1^\infty n^p P[|Z_n| \geqq \varepsilon n^{\zeta}] < \infty$, then $E\tau^p(\zeta, \varepsilon) < \infty$.*
(ii) *If $\sum_1^\infty n^{p-1} P[\max_{j\leqq n} |Z_j| \geqq \frac{1}{4}\varepsilon n^{\alpha}] < \infty$, then $E\tau^p(\alpha, \varepsilon) < \infty$.*
(iii) *If $\sum_1^\infty n^{p-1} P[\max_{j\leqq n} j^{\alpha-\zeta}|Z_j| \geqq \frac{1}{4}\varepsilon n^{\alpha}] < \infty$, then $E\tau^p(\zeta, \varepsilon) < \infty$.*

PROOF. (i) follows easily from the fact that $P[\tau(\zeta, m) \geqq m] \leqq \sum_{n=m}^{\infty} P[|Z_n| \geqq \varepsilon n^{\zeta}]$. (ii) is known (cf. Lemma 2 of [5]), and noting that $\tau(\zeta, \varepsilon) = \sup\{n \geqq 1 : n^{\alpha-\zeta}|Z_n| \geqq \varepsilon n^{\alpha}\}$. (iii) follows from (ii).

LEMMA 4. *Let $f\colon [0, 1] \times [0, 1] \to R$ be twice continuously differentiable except possibly at the points $(0, 0)$ and $(1, 1)$. We shall write $f^{(0)} = |f|$, $f^{(1)} = |\partial f/\partial x| + |\partial f/\partial y|$, $f^{(2)} = |\partial^2 f/\partial x^2| + |\partial^2 f/\partial y^2| + |\partial^2 f/\partial x\, \partial y|$, and we shall let $a \vee b$ denote* $\max(a, b)$. *Suppose there exists $0 < \delta < \frac{5}{2}$ such that for $i = 2$,*

$$f^{(i)}(x, y) \leqq K(\min\{x \vee y, (1-x) \vee (1-y)\})^{-i-\frac{1}{2}+\delta} \quad \text{for some } K > 0 \text{ and all } 0 < x,\ y < 1 . \tag{3.7}$$

*Then* (3.7) *also holds for $i = 0, 1$ if $\delta < \frac{1}{2}$, while in the case $\delta = \frac{1}{2}$,* (3.7) *holds for $i = 1$ and as* $\max(x, y) \to 0$ *or as* $\max(1-x, 1-y) \to 0$,

$$f^{(0)}(x, y) = O(|\log(x \vee y)| + |\log((1-x) \vee (1-y))|) . \tag{3.8}$$

*As to the case $\frac{1}{2} < \delta < \frac{5}{2}$, there exist functions $g$, $h$ such that $f = g + h$, $g$ is twice continuously differentiable on* $[0, 1] \times [0, 1]$ (hence $g^{(2)}$ is bounded) *and $h$ satisfies* (3.7) (with $h^{(i)}$ replacing $f^{(i)}$) *for $i = 0, 1, 2$ if $\delta \neq \frac{3}{2}$, while in the case $\delta = \frac{3}{2}$,* (3.8) *holds with $h^{(1)}$ replacing $f^{(0)}$, and*

$$\begin{aligned} h^{(0)}(x, y) &= O((x \vee y)|\log(x \vee y)|) && \text{as } x \vee y \to 0 ; \\ &= O(((1-x) \vee (1-y))|\log((1-x) \vee (1-y))|) && \\ & && \text{as } (1-x) \vee (1-y) \to 0 . \end{aligned} \tag{3.9}$$

PROOF. We note that if $\varphi$ is continuously differentiable on $[a, b] \times [c, d]$, then

$$\varphi(b, d) - \varphi(a, c) = \int_a^b \frac{\partial \varphi}{\partial u}(u, c)\, du + \int_c^d \frac{\partial \varphi}{\partial v}(b, v)\, dv . \tag{3.10}$$

Hence if $0 < \delta \leqq \frac{1}{2}$, then (3.7) also holds for $i = 0, 1$ when $\delta \neq \frac{1}{2}$, while (3.8) holds when $\delta = \frac{1}{2}$. If $\frac{1}{2} < \delta < \frac{3}{2}$, using (3.10), we see that (3.7) also holds for $i = 1$, and so $\lim_{(x,y)\to(0,0)} f(x, y) = L$, $\lim_{(x,y)\to(1,1)} f(x, y) = L'$ both exist and are finite, and

$$f(x, y) - L = \int_{0+}^{x} \frac{\partial f}{\partial u}(u, 0)\, du + \int_{0+}^{y} \frac{\partial f}{\partial v}(x, v)\, dv \tag{3.11}$$

for $0 \leqq x, y \leqq \frac{1}{2}$ such that $(x, y) \neq (0, 0)$, with a similar expression for $L' - f(x, y)$. By choosing $h$ equal to the right-hand side of (3.11) in a deleted neighborhood of $(0, 0)$, we can easily construct $h$ and $g$. The case $\delta = \frac{3}{2}$ is similar, while the case $\frac{3}{2} < \delta < \frac{5}{2}$ can be treated by repeated use of an argument similar to (3.11).

LEMMA 5. *Let $(m_n)$ be a sequence of positive integers and let $\gamma_n = n/(n + m_n)$. Suppose $\gamma_n = \gamma + o(n^{-\rho})$ for some $\rho > 0$ and $0 < \gamma < 1$. Let $u\colon [0, 1] \to R$ be continuously differentiable on the open interval* $(0, 1)$ *and let $J\colon [0, 1] \times [0, 1] \to R$ be defined by $J(x, y) = u(\gamma x + (1-\gamma)y)$.*

(i) *Suppose there exist* $0 < \lambda < 2$ *and* $K > 0$ *such that*

$$|u'(t)| \leqq K(t(1-t))^{-\lambda}, \qquad 0 < t < 1. \tag{3.12}$$

*Define* $J_n : \{0, 1/n, \cdots, 1\} \times \{0, 1/m_n, \cdots, 1\} \to R$ *by* $J_n(x, y) = u(\gamma_n x + (1-\gamma_n)y)$. *Then*

$$n^{-1} \sum_{i=1}^{n} \sup_{y \in \{0,1/m_n,\cdots,1\}} |J_n(i/n, y) - J(i/n, y)| = o(n^{-\rho}) \qquad \text{as } n \to \infty. \tag{3.13}$$

(ii) *Suppose in* (i) *we define* $J_n$ *by* $J_n(x, y) = u((nx + m_n y)/(n + m_n + 1))$. *If* $\rho < \min(1, 2-\lambda)$, *then* (3.13) *still holds.*

(iii) *Suppose u is twice continuously differentiable on the open interval* (0, 1), *and there exist* $K > 0$ *and* $0 < \delta < \frac{1}{2}$ *such that*

$$|u''(t)| \leqq K(t(1-t))^{-\frac{5}{2}+\delta}, \qquad 0 < t < 1. \tag{3.14}$$

*Let* $u_n(j/(n + m_n)) = Eu(U_{j,n})$, *where* $U_{j,n}$ *is the* $j$th *order statistic of a sample of size* $(n + m_n)$ *from the uniform distribution on* (0, 1). *Define* $J_n(0, 0) = 0$ *and* $J_n(x, y) = u_n(\gamma_n x + (1-\gamma_n)y)$ *if* $(x, y) \neq (0, 0)$, $x = 0, 1/n, \cdots, 1$, $y = 0, 1/m_n, \cdots, 1$. *If* $\rho < \frac{1}{2} + \delta$, *then* (3.13) *still holds.*

PROOF. Let $M_n$ denote the set $\{0, m_n^{-1}, 2m_n^{-1}, \cdots, 1\}$. To prove (i), we obtain by (3.12) and the mean value theorem that for $\frac{1}{2}n \leqq i \leqq n-1$,

$$|J_n(i/n, y) - J(i/n, y)| \leqq c|(\gamma - \gamma_n)(i/n - y)|(1 - Q(n; i, y))^{-\lambda}, \tag{3.15}$$

where $Q(n; i, y)$ lies between $\gamma_n(i/n) + (1-\gamma_n)y$ and $\gamma(i/n) + (1-\gamma)y$. Since $|i/n - y| = |(1 - i/n) - (1 - y)| \leqq 1 - i/n$ and $Q(n; i, y) \leqq \max\{\gamma_n(i/n) + (1-\gamma_n), \gamma(i/n) + (1-\gamma)\}$ for $y \in M_n$, it follows from (3.15) that

$$\begin{aligned} \sum_{n/2 \leqq i \leqq n-1} \sup_{y \in M_n} &|J_n(i/n, y) - J(i/n, y)| \\ &\leqq c|\gamma - \gamma_n| \sum_{n/2 \leqq i \leqq n-1} (1 - i/n)\{(\gamma_n^{-\lambda} + \gamma^{-\lambda})(1 - i/n)^{-\lambda}\} \\ &\leqq c_1 n|\gamma - \gamma_n|, \qquad \text{since } \lambda < 2. \end{aligned} \tag{3.16}$$

An obvious modification of the above argument leads to

$$\sum_{1 \leqq i < n/2} \sup_{y \in M_n} |J_n(i/n, y) - J(i/n, y)| \leqq c_2 n|\gamma - \gamma_n|. \tag{3.17}$$

Since $J_n(1, 1) = u(1) = J(1, 1)$, it follows from the mean value theorem that

$$\begin{aligned} \sup_{y \in M_n} &|J_n(1, y) - J(1, y)| \\ &\leqq c|\gamma - \gamma_n| \sup_{y \in \{0, m_n^{-1}, \cdots, 1-m_n^{-1}\}} [(1-y)\{(1-\gamma_n)^{-\lambda} \\ &\qquad + (1-\gamma^{-\lambda})\}(1-y)^{-\lambda}] \\ &\leqq c_3 n|\gamma - \gamma_n|, \qquad \text{since } \lambda < 2. \end{aligned} \tag{3.18}$$

From (3.16), (3.17) and (3.18), the desired conclusion (3.13) follows.

We now prove (ii). Let $\zeta_n = |\gamma - (n+1)/(n + m_n + 1)|$ and $\theta_n = 1/(n+m_n+1)$. The mean value theorem in this case gives that for $\frac{1}{2}n \leqq i \leqq n-1$,

$$|J_n(i/n, y) - J(i/n, y)| \leqq c\{\zeta_n|i/n - y| + \theta_n i/n\}(1 - Q_1(n; i, y))^{-\lambda}, \tag{3.19}$$

where $Q_1(n; i, y)$ lies between $(i + m_n y)/(n + m_n + 1)$ and $\gamma(i/n) + (1-\gamma)y$. We note that

$$\begin{aligned}
&\theta_n \textstyle\sum_{n/2\le i\le n-1} \sup_{y\in M_n} (1 - Q_1(n; i, y))^{-\lambda} \\
&\qquad \le \theta_n \textstyle\sum_{n/2\le i\le n-1} \{\gamma^{-\lambda} + ((n+1)/(n+m_n+1))^{-\lambda}\}(1 - i/n)^{-\lambda} \\
&\qquad \le c_4\theta_n \textstyle\int_{n/2-\frac12}^{n-\frac12} (1 - t/n)^{-\lambda}\, dt \\
&\qquad \le c_5 n^{\lambda}\theta_n = o(n^{1-\rho}), \qquad \text{since} \quad \rho < 2 - \lambda .
\end{aligned}$$

As in Lemma 4, the condition (3.12) implies that as $t \uparrow 1$, $u(t) = O((1-t)^{-(\lambda-1)})$ if $\lambda > 1$, $u(t) = O(\log(1-t))$ if $\lambda = 1$ and $u(t) = O(1)$ if $\lambda < 1$. Therefore

$$\begin{aligned}
&\sup_{y\in M_n} |J_n(1, y) - J(1, y)| \\
&\qquad \le |u(1)| + |u((n+m_n)/(n+m_n+1))| \\
&\qquad\qquad + \sup_{y\in\{0, m_n^{-1}, \cdots, 1-m_n^{-1}\}} (|J_n(1, y)| + |J(1, y)|) \\
&\qquad = o(n^{1-\rho}), \qquad \text{since} \quad \rho < \min(1, 2-\lambda) .
\end{aligned}$$

The rest of the proof of (ii) proceeds in the same way as in (i).

To prove (iii), we note that as in Lemma 4, condition (3.14) implies that (3.12) holds with $\lambda = \frac{3}{2} - \delta$ and $|u(t)| \le K_1(t(1-t))^{-\frac12+\delta}$ for $0 < t < 1$. Chernoff and Savage ([4] pages 991–993) have shown that there exists a constant $C$ such that $|u_n(1)| \le Cn^{\frac12-\delta}$, and in general for $1 \le j \le \frac12(n+m_n)$,

$$\begin{aligned}
&|u_n((n+m_n)^{-1}j) - u((n+m_n)^{-1}j)| \\
&\qquad + |u_n(1-(n+m_n)^{-1}j) - u(1-(n+m_n)^{-1}j)| \\
&\qquad\qquad \le Cn^{\frac12-\delta}\{\Phi(-j^{\frac12}/C) + n^{-1} + j^{-\frac32+\delta}\} ,
\end{aligned}$$

where $\Phi$ is the standard normal distribution function. Hence

$$\begin{aligned}
(3.20)\qquad & n^{-1} \textstyle\sum_{i=1}^{n} \sup_{y\in M_n} |u_n(\gamma_n i/n + (1-\gamma_n)y) - u(\gamma_n i/n + (1-\gamma_n)y)| \\
&\qquad \le C'n^{-\frac12-\delta} = o(n^{-\rho}), \qquad \text{since} \quad \rho < \tfrac12 + \delta .
\end{aligned}$$

The desired conclusion (3.13) then follows easily from (i) and (3.20).

**4. A representation theorem, an invariance principle and a law of the iterated logarithm for generalized Chernoff–Savage statistics.** Let $X_1, X_2, \cdots, Y_1, Y_2, \cdots, F_n, G_m, F, G$ be as in Section 3. Suppose $J: [0, 1] \times [0, 1] \to R$ is twice continuously differentiable except possibly at the points (0, 0) and (1, 1). With the same notation as in Lemma 4, we shall assume that $J$ satisfies Assumption $(A_\delta)$ for some $0 \le \delta < \frac52$ described below:

ASSUMPTION $(A_0)$. There exists $K$ such that $J^{(2)}(x, y) \le K$, $0 < x, y < 1$.

ASSUMPTION $(A_\delta)$ (with $0 < \delta < \frac52$). There exists $K$ such that

$$(4.1)\qquad J^{(2)}(x, y) \le K(\{\max(x, y)\}^{-\frac52+\delta} + \{\max(1-x, 1-y)\}^{-\frac52+\delta}) , \qquad 0 < x, y < 1 .$$

Let $(m_n)$ be a non-decreasing sequence of positive integers and let $J_n: \{0, 1/n, 2/n, \cdots, 1\} \times \{0, 1/m_n, 2/m_n, \cdots, 1\} \to R$ be a sequence of functions such that for some $\rho > 0$, the following Assumption $(B_\rho)$ is satisfied:

ASSUMPTION $(B_\rho)$. As $n \to \infty$,

$$n^{-1} \sum_{i=1}^n \sup_{y \in \{0, 1/m_n, \cdots, 1\}} |J_n(i/n, y) - J(i/n, y)| = o(n^{-\rho}) . \tag{4.2}$$

We shall call the statistic

$$T_n = \int_{-\infty}^{\infty} J_n(F_n(x), G_{m_n}(x)) \, dF_n(x) \tag{4.3}$$

a *generalized Chernoff–Savage statistic*. To give some examples, let $\gamma_n = n/(n + m_n)$ and assume that $\gamma_n = \gamma + o(n^{-\rho})$ for some $0 < \gamma < 1$ and $0 < \rho < 1$. First define $u : [0, 1] \to R$ by $u(0) = u(1) = 0$ and $u(t) = \Phi^{-1}(t)$ if $0 < t < 1$, where $\Phi$ is the distribution function of the standard normal distribution. If we set $J_n(x, y) = u((nx + m_n y)/(n + m_n + 1))$ and $J(x, y) = u(\gamma x + (1 - \gamma)y)$, then $\int_{-\infty}^{\infty} J_n(F_n, G_{m_n}) \, dF_n$ is the van der Waerden statistic. It is easy to see that $J$ satisfies Assumption $(A_\delta)$ with $\delta = \frac{1}{2}$. By Lemma 5(ii), Assumption $(B_\rho)$ is also satisfied. For another example, take the normal scores statistic $\int_{-\infty}^{\infty} u_n(\gamma_n F_n + (1 - \gamma_n)G_{m_n}) \, dF_n$, where $u_n$ is defined from $u$ as in Lemma 5(iii), and again Assumptions $(A_{\frac{1}{2}})$ and $(B_\rho)$ are satisfied. More generally, if $u : [0, 1] \to R$ is twice continuously differentiable on $(0, 1)$ and (3.14) is satisfied for some $0 < \delta \leqq \frac{5}{2}$, then the statistic $\int_{-\infty}^{\infty} u(\gamma_n F_n + (1 - \gamma_n)G_{m_n}) \, dF_n$ is a generalized Chernoff–Savage statistic satisfying Assumptions $(A_\delta)$ and $(B_\rho)$ (see Lemma 5(i)).

In the following theorem, we shall represent $nT_n$ as the partial sum of i.i.d. random variables plus a remainder term, whose magnitude we shall describe in terms of the finiteness of moments of the last time its absolute value exceeds a square-root boundary, or more generally, a boundary of the form $n^{1-\mu}$ for some $0 < \mu < 1$. This stronger notion than almost everywhere convergence was introduced by Strassen ([16], page 316) and is needed in our study of the stopping times of sequential rank tests in Section 5.

THEOREM 2. *Suppose the generalized Chernoff–Savage statistic $T_n$ of* (4.3) *is written as*

$$\begin{aligned} T_n = \int_{-\infty}^{\infty} J(F(x), G(x)) \, dF(x) &+ n^{-1} \textstyle\sum_1^n (\psi(X_i) - E\psi(X_i)) \\ &+ m_n^{-1} \textstyle\sum_1^{m_n} (\psi^*(Y_i) - E\psi^*(Y_i)) + R_n \end{aligned} \tag{4.4}$$

*where we define*

$$\psi(u) = J(F(u), G(u)) - \int_{u_0}^{u} \frac{\partial J}{\partial x} (F(t), G(t)) \, dF(t)$$

$$\psi^*(u) = -\int_{u_0}^{u} \frac{\partial J}{\partial y} (F(t), G(t)) \, dF(t) .$$

*Assume that there exist positive constants $\lambda_1 < \lambda_2$ such that*

$$n\lambda_1(1 + o(1)) \leqq m_n \leqq n\lambda_2(1 + o(1)) . \tag{4.5}$$

*For $\mu > 0$, define $L(\mu, \varepsilon) = \sup\{n \geqq 1 : |R_n| \geqq \varepsilon n^{-\mu}\}$ $(\sup \emptyset = 0)$.*

(i) *If $0 < \mu < 1$, then under Assumptions $(A_0)$ and $(B_\rho)$ with $\rho \geqq \mu$, $EL^\gamma(\mu, \varepsilon) < \infty$ for all $\gamma > 0$ and $\varepsilon > 0$.*

(ii) *If* $0 < \delta \leqq \frac{1}{2}$ *and* $0 < \mu < \frac{1}{2} + \delta$, *then under Assumptions* $(A_\delta)$ *and* $(B_\rho)$ *with* $\rho \geqq \mu$, $EL^\gamma(\mu, \varepsilon) < \infty$ *for all* $\varepsilon > 0$ *and* $0 < \gamma < (\frac{1}{2} + \delta) - \mu$.

(iii) *Suppose* $\frac{1}{2} < \delta < \frac{5}{2}$. *Let* $\mu(\delta) = (1 + 2\delta)(9 - 2\delta)/2(17 - 2\delta)$. *Then* $\mu(\delta)$ *is increasing in* $\delta$ *for* $\delta$ *belonging to the range specified above, with* $\lim_{\delta\to\frac{1}{2}} \mu(\delta) = \frac{1}{2}$ *and* $\lim_{\delta\to\frac{5}{2}} \mu(\delta) = 1$. *Let* $\rho \geqq \mu > 0$ *and suppose that Assumptions* $(A_\delta)$ *and* $(B_\rho)$ *both hold. If* $\mu < \mu(\delta)$, *then* $EL^\gamma(\mu, \varepsilon) < \infty$ *for all* $\gamma > 0$ *and* $\varepsilon > 0$. *If* $\mu(\delta) \leqq \mu < 1$, *then* $EL^\gamma(\mu, \varepsilon) < \infty$ *for all* $\varepsilon > 0$ *and* $0 < \gamma < 1 - \mu$.

COROLLARY. *Let* $0 < \mu < 1$. *Assume* $(B_\mu)$, (4.5) *and either* $(A_0)$ *or* $(A_\delta)$ *with* $\delta + \frac{1}{2} > \mu$ $(0 < \delta < \frac{5}{2})$. *Then* $\lim_{n\to\infty} n^\mu R_n = 0$ a.e. *Consequently, if* $\lim_{n\to\infty} n/m_n = \lambda$ $(> 0)$ *and Assumptions* $(A_\delta)$ *and* $(B_{\frac{1}{2}})$ *hold for some* $0 \leqq \delta < \frac{5}{2}$, *then the following conclusions* (i) *and* (ii) *both hold*:

(i) *Invariance principle for* $T_n$. *Setting* $V_n = n(T_n - \int_{-\infty}^{\infty} J(F, G)\, dF)$, *then* $n^{-\frac{1}{2}} V_{[nt]}/\sigma$, $0 \leqq t \leqq 1$, *converges weakly to the standard Wiener process, where*

$$\sigma^2 = \operatorname{Var} \psi(X_1) + \lambda \operatorname{Var} \psi^*(Y_1)\,. \tag{4.6}$$

(ii) *Law of the iterated logarithm.*

$$\limsup_{n\to\infty} n^{\frac{1}{2}}(T_n - \textstyle\int_{-\infty}^{\infty} J(F, G)\, dF)/(2 \log\log n)^{\frac{1}{2}} = \sigma \quad \text{a.e.} \tag{4.7}$$

REMARK. We note that $\sigma < \infty$. In fact, under Assumption $(A_0)$ or Assumption $(A_\delta)$ for $\frac{5}{2} > \delta \geqq \frac{1}{2}$, $E|\psi(X_1)|^p < \infty$ and $E|\psi^*(Y_1)|^p < \infty$ for all $p > 0$. If $0 < \delta < \frac{1}{2}$, then for $\delta' > 0$ such that $(2 + \delta')(-\frac{1}{2} + \delta) > -1$, $E|\psi(X_1)|^{2+\delta'} < \infty$ and $E|\psi^*(Y_1)|^{2+\delta'} < \infty$ under Assumption $(A_\delta)$ (cf. [4], page 977). Related to (i) and (ii), Sen and Ghosh [14] have given an invariance principle and a law of the iterated logarithm for two-sample linear rank statistics when $F = G$.

PROOF OF THEOREM 2. To show that $EL^\gamma(\mu, \varepsilon) < \infty$, as the same argument works for any $\varepsilon > 0$, we shall for simplicity consider $L(\mu, 1)$ and write $L(\mu)$ instead of $L(\mu, 1)$. In view of the Assumption $(B_\rho)$ with $\rho \geqq \mu$, we shall without loss of generality assume that $J_n = J$ for all $n$.

Given any $0 < \mu < 1$, we can choose $\frac{1}{2} < \delta < \frac{5}{2}$ such that $\mu < \mu(\delta) = (1 + 2\delta)(9 - 2\delta)/2(17 - 2\delta)$. Since Assumption $(A_0)$ obviously implies $(A_\delta)$, the conclusion in (i) follows immediately from that of (iii). However, in our proof of (iii), we shall need the fact that the conclusion in (i) holds with the following slightly stronger Assumption $(A_0{}^*)$ replacing $(A_0)$:

ASSUMPTION $(A_0{}^*)$. $J$ is twice continuously differentiable on the whole of $[0, 1] \times [0, 1]$.

Under Assumption $(A_0{}^*)$, we can apply Taylor's expansion to $J$ and write

$$\begin{aligned}\textstyle\int_{-\infty}^{\infty} J(F_n, G_{m_n})\, dF_n = \int_{-\infty}^{\infty} J(F, G)\, dF &+ \textstyle\int_{-\infty}^{\infty} J(F, G)\, d(F_n - F)\\ &+ \int_{-\infty}^{\infty} (F_n - F) \frac{\partial J}{\partial x}(F, G)\, dF\\ &+ \int_{-\infty}^{\infty} (G_{m_n} - G) \frac{\partial J}{\partial y}(F, G)\, dF + \textstyle\sum_{i=1}^{5} R_{in}\end{aligned}$$

where

$$R_{1n} = \int_{-\infty}^{\infty} (F_n - F) \frac{\partial J}{\partial x} (F, G)\, d(F_n - F)$$

$$= -\frac{1}{2}\left[\int_{-\infty}^{\infty} (F_n - F)^2 \frac{\partial^2 J}{\partial x^2} (F, G)\, dF\right.$$

$$\left. + \int_{-\infty}^{\infty} (F_n - F)^2 \frac{\partial^2 J}{\partial x\, \partial y} (F, G)\, dG - \frac{1}{n} \int_{-\infty}^{\infty} \frac{\partial J}{\partial x} (F, G)\, dF_n \right],$$

$$R_{2n} = \int_{-\infty}^{\infty} (G_{m_n} - G) \frac{\partial J}{\partial y} (F, G)\, d(F_n - F),$$

$$R_{3n} = \tfrac{1}{2} \int_{-\infty}^{\infty} (F_n - F)^2 \frac{\partial^2 J}{\partial x^2} (\hat{F}, \hat{G})\, dF_n,$$

$$R_{4n} = \tfrac{1}{2} \int_{-\infty}^{\infty} (G_{m_n} - G)^2 \frac{\partial^2 J}{\partial y^2} (\hat{F}, \hat{G})\, dF_n,$$

$$R_{5n} = \int_{-\infty}^{\infty} (F_n - F)(G_{m_n} - G) \frac{\partial^2 J}{\partial x\, \partial y} (\hat{F}, \hat{G})\, dF_n,$$

and $(\hat{F}(x), \hat{G}(x))$ above denotes a point (given by Taylor's expansion) lying on the line segment joining $(F(x), G(x))$ and $(F_n(x), G_{m_n}(x))$. Now $R_n = \sum_1^5 R_{in}$ and noting that $\int_{-\infty}^{\infty} d(F_n - F) = 0$, we have

$$\int_{-\infty}^{\infty} J(F, G)\, d(F_n - F) + \int_{-\infty}^{\infty} (F_n - F) \frac{\partial J}{\partial x} (F, G)\, dF$$

$$+ \int_{-\infty}^{\infty} (G_{m_n} - G) \frac{\partial J}{\partial y} (F, G)\, dF$$

$$= n^{-1} \sum_1^n (\psi(X_i) - E\psi(X_i)) + m_n^{-1} \sum_1^{m_n} (\psi^*(Y_i) - E\psi^*(Y_i)).$$

Under Assumption $(A_0^*)$, $J$, $\partial J/\partial x$, $\partial^2 J/\partial x^2$, etc., are all bounded, and so by (2.1) and Lemma 1, $EL^\gamma(\mu) < \infty$ for all $0 < \mu < 1$ and $\gamma > 0$.

We now prove (ii). Let $0 < \delta < \frac{1}{2}$, $0 < \mu < \frac{1}{2} + \delta$ and take $1 > \alpha > 2\mu/(1 + 2\delta)$. Set $H = \frac{1}{2}(F + G)$. Define

$$\tilde{L} = \sup\{n \geq 1 : \max_{H(x) \geq n^{-\alpha}} H(x)/\max(F_n(x), G_{m_n}(x)) > 2 \text{ or } \max_{1-H(x) \geq n^{-\alpha}} (1 - H(x))/\max(1 - F_n(x), 1 - G_{m_n}(x)) > 2\}. \tag{4.8}$$

By (3.4), $E\tilde{L}^\gamma < \infty$ for all $\gamma > 0$. When $n > \tilde{L}$, $n^{-\alpha} \leq H(x) \leq 1 - n^{-\alpha}$ implies that $(F_n(x), G_{m_n}(x)) \notin \{(0, 0), (1, 1)\}$, and therefore writing

$$\int_{-\infty}^{\infty} J(F_n, G_{m_n})\, dF_n = \int_{H < n^{-\alpha}} + \int_{n^{-\alpha} \leq H \leq 1 - n^{-\alpha}} + \int_{H > 1 - n^{-\alpha}},$$

we can use Taylor's expansion for $J(F_n, G_{m_n})$ in the middle integral and obtain:

$$\begin{aligned} &\int_{-\infty}^{\infty} J(F_n, G_{m_n})\, dF_n \\ &\quad = \int_{-\infty}^{\infty} J(F, G)\, dF + \int_{-\infty}^{\infty} J(F, G)\, d(F_n - F) \\ &\qquad + \int_{-\infty}^{\infty} (F_n - F) \frac{\partial J}{\partial x} (F, G)\, dF + \int_{-\infty}^{\infty} (G_{m_n} - G) \frac{\partial J}{\partial y} (F, G)\, dF \\ &\qquad + \sum_{i=1}^4 D_{in} + \sum_{i=1}^4 H_{in} + \sum_{i=1}^5 Q_{in}, \end{aligned} \tag{4.9}$$

where

$$D_{1n} = \int_{H<n^{-\alpha}} J(F_n, G_{m_n})\, dF_n\,, \qquad D_{2n} = \int_{1-H<n^{-\alpha}} J(F_n, G_{m_n})\, dF_n\,,$$
$$D_{3n} = -\int_{H<n^{-\alpha}} J(F, G)\, dF_n\,, \qquad D_{4n} = -\int_{1-H<n^{-\alpha}} J(F, G)\, dF_n\,,$$
$$H_{1n} = -\int_{H<n^{-\alpha}} (F_n - F) \frac{\partial J}{\partial x}(F, G)\, dF\,,$$
$$H_{2n} = -\int_{1-H<n^{-\alpha}} (F_n - F) \frac{\partial J}{\partial x}(F, G)\, dF\,,$$
$$H_{3n} = -\int_{H<n^{-\alpha}} (G_{m_n} - G) \frac{\partial J}{\partial y}(F, G)\, dF\,,$$
$$H_{4n} = -\int_{1-H<n^{-\alpha}} (G_{m_n} - G) \frac{\partial J}{\partial y}(F, G)\, dF\,,$$
$$Q_{1n} = \int_{n^{-\alpha}\leqq H\leqq 1-n^{-\alpha}} (F_n - F) \frac{\partial J}{\partial x}(F, G)\, d(F_n - F)\,,$$
$$Q_{2n} = \int_{n^{-\alpha}\leqq H\leqq 1-n^{-\alpha}} (G_{m_n} - G) \frac{\partial J}{\partial y}(F, G)\, d(F_n - F)\,,$$
$$Q_{3n} = \tfrac{1}{2}\int_{n^{-\alpha}\leqq H\leqq 1-n^{-\alpha}} (F_n - F)^2 \frac{\partial^2 J}{\partial x^2}(\hat{F}, \hat{G})\, dF_n\,,$$
$$Q_{4n} = \tfrac{1}{2}\int_{n^{-\alpha}\leqq H\leqq 1-n^{-\alpha}} (G_{m_n} - G)^2 \frac{\partial^2 J}{\partial y^2}(\hat{F}, \hat{G})\, dF_n\,,$$
$$Q_{5n} = \int_{n^{-\alpha}\leqq H\leqq 1-n^{-\alpha}} (F_n - F)(G_{m_n} - G) \frac{\partial^2 J}{\partial x\, \partial y}(\hat{F}, \hat{G})\, dF_n\,.$$

We shall let $L(D_i; \mu) = \sup\{n \geqq 1 : |D_{in}| \geqq n^{-\mu}\}$ and define $L(H_i; \mu)$, $L(Q_1; \mu)$, $L(Q_2; \mu)$ similarly. For $j = 3, 4, 5$, let $L(Q_j; \mu) = \sup\{n > \tilde{L} : |Q_{jn}| \geqq n^{-\mu}\}$. We note that $\sup\{n \geqq 1 : |R_n| \geqq 13n^{-\mu}\} \leqq \tilde{L} + 1 + \sum_{i=1}^4 L(D_i; \mu) + \sum_{i=1}^4 L(H_i; \mu) + \sum_{i=1}^5 L(Q_i; \mu)$.

By Lemma 4, we can assume that for $i = 0, 1, 2$,

$$|J^{(i)}(F, G)| \leqq K\{\min(H, 1 - H)\}^{-i-\frac{1}{2}+\delta} \qquad \text{if} \quad 0 < H < 1\,. \tag{4.10}$$

In our argument below, we shall frequently use the following fact:

$$F \leqq 2H\,, \qquad G \leqq 2H\,, \qquad dF \leqq 2dH\,, \qquad dG \leqq 2dH\,. \tag{4.11}$$

We note that

$$\begin{aligned} |D_{1n}| &\leqq K \int_{H<n^{-\alpha}} (F_n(1 - F_n))^{-(\frac{1}{2}-\delta)}\, dF_n \\ &\leqq K_1(F_n H^{-1}(n^{-\alpha}))^{\frac{1}{2}+\delta}(1 - F_n H^{-1}(n^{-\alpha}))^{-(\frac{1}{2}-\delta)}\,. \end{aligned} \tag{4.12}$$

Since $\alpha > 2\mu/(1 + 2\delta)$, it is easy to see from (3.6) (where we set $\lambda = \alpha$ and $\tau = \alpha - 2\mu(1 + 2\delta)^{-1}$) that $EL^{\gamma}(D_1; \mu) < \infty$ for all $\gamma > 0$. Likewise we can show that $EL^{\gamma}(D_2; \mu) < \infty$ for all $\gamma > 0$.

We now consider $L(D_3; \mu)$. Noting that $n \int f\, dF_n$ is increasing in $n$ for any nonnegative function $f$, we have

$$
\begin{aligned}
&P[\max_{1\leqq n\leqq m} n \textstyle\int_{H<n^{-\alpha}} |J(F,G)|\, dF_n \geqq \frac{1}{4}m^{1-\mu}] \\
&\qquad \leqq P[\max_{1\leqq n\leqq m} n \textstyle\int_{m^{-\alpha}\leqq H<n^{-\alpha}} |J(F,G)|\, dF_n \geqq \frac{1}{8}m^{1-\mu}] \\
&\qquad\qquad + P[m \textstyle\int_{H<m^{-\alpha}} |J(F,G)|\, dF_m \geqq \frac{1}{8}m^{1-\mu}] \\
&\qquad \leqq P[\max_{1\leqq n\leqq m} K' m^{\alpha(\frac{1}{2}-\delta)} n F_n H^{-1}(n^{-\alpha}) \geqq \frac{1}{8}m^{1-\mu}] \\
&\qquad\qquad + 8m^{\mu} E \textstyle\int_{H<m^{-\alpha}} |J(F,G)|\, dF_m\,.
\end{aligned} \tag{4.13}
$$

The first term above can be handled using (3.6), noting that $(1-\mu)-\alpha(\frac{1}{2}-\delta) > 1-\alpha$, while for the second term, we have by (4.10) and (4.11),

$$E \int_{H<m^{-\alpha}} |J(F,G)|\, dF_m = \int_{H<m^{-\alpha}} |J(F,G)|\, dF = O(m^{-\alpha(\frac{1}{2}+\delta)})\,.$$

Hence by Lemma 3(iii), $EL^{\gamma}(D_3;\mu) < \infty$ for $0<\gamma<\alpha(\frac{1}{2}+\delta)-\mu$. The same conclusion obviously also holds for $L(D_4;\mu)$.

We now analyze $L(H_3;\mu)$ in a similar way:

$$
\begin{aligned}
&P\left[\max_{1\leqq n\leqq \nu}\left|n \int_{H<n^{-\alpha}} (G_{m_n}-G)\frac{\partial J}{\partial y}(F,G)\, dF\right| \geqq \tfrac{1}{4}\nu^{1-\mu}\right] \\
&\quad \leqq P\left[\max_{1\leqq n\leqq \nu}\left|\mathrm{m}_n \int_{H<\nu^{-\alpha}} (G_{m_n}-G)\frac{\partial J}{\partial y}(F,G)\, dF\right| \geqq \varepsilon_1\nu^{1-\mu}\right] \\
&\qquad + P\left[\max_{1\leqq n\leqq \nu} n \int_{\nu^{-\alpha}\leqq H\leqq n^{-\alpha}} G_{m_n}\left|\frac{\partial J}{\partial y}(F,G)\right| dF \geqq \varepsilon_2\nu^{1-\mu}\right] \\
&\qquad + P\left[\max_{1\leqq n\leqq \nu} n \int_{\nu^{-\alpha}\leqq H\leqq n^{-\alpha}} G\left|\frac{\partial J}{\partial y}(F,G)\right| dF \geqq \varepsilon_3\nu^{1-\mu}\right] \\
&\quad = \xi_\nu^{(1)} + \xi_\nu^{(2)} + \xi_\nu^{(3)}\,, \quad \text{say.}
\end{aligned} \tag{4.14}
$$

Using the martingale inequality, we obtain that

$$
\begin{aligned}
\xi_\nu^{(1)} &\leqq (\varepsilon_1\nu^{1-\mu})^{-1} m_\nu \int_{H<\nu^{-\alpha}} E|G_{m_\nu}-G|\left|\frac{\partial J}{\partial y}(F,G)\right| dF \\
&= O(\nu^{\mu-(\frac{1}{2}+\alpha\delta)})\,.
\end{aligned} \tag{4.15}
$$

The last relation above follows from (4.10), (4.11), together with the fact that $E|G_{m_\nu}-G| \leqq E^{\frac{1}{2}}|G_{m_\nu}-G|^2 = \{G(1-G)/m_\nu\}^{\frac{1}{2}}$. Since $\frac{1}{2}-\delta<1-\mu$, we can choose $\delta_1>\frac{1}{2}-\delta$ such that $\delta_1\alpha<1-\mu$. Letting $\delta_1+\delta_2=\frac{3}{2}-\delta$, then $\delta_2<1$ and using (4.10), (4.11) and (3.6), we have

$$
\begin{aligned}
\xi_\nu^{(2)} &\leqq P[\max_{1\leqq n\leqq \nu} K'\nu^{\delta_1\alpha} n G_{m_n} H^{-1}(n^{-\alpha}) \textstyle\int_{H\leqq n^{-\alpha}} H^{-\delta_2}\, dH \geqq \varepsilon_2\nu^{1-\mu}] \\
&= P[\max_{1\leqq n\leqq \nu} n^{1-\alpha(1-\delta_2)} G_{m_n} H^{-1}(n^{-\alpha}) \geqq \varepsilon_4\nu^{1-\mu-\delta_1\alpha}] \\
&= o(\exp(-\nu^p)) \qquad \text{for some} \quad p>0\,.
\end{aligned} \tag{4.16}
$$

By (4.10) and (4.11), we obtain

$$\int_{\nu^{-\alpha}\leqq H\leqq n^{-\alpha}} G\left|\frac{\partial J}{\partial y}(F,G)\right| dF \leqq K' n^{-\alpha(\frac{1}{2}+\delta)}\,.$$

Since $\alpha(\frac{1}{2}+\delta)>\mu$, it follows that $\xi_\nu^{(3)}=0$ for all $\nu$ large. Hence from (4.14), (4.15) and (4.16), we obtain using Lemma 3(iii) that $EL^\gamma(H_3;\mu) < \infty$ for $0 < \gamma < \frac{1}{2} + \alpha\delta - \mu$. Similarly we can show that the same conclusion also holds for $L(H_4;\mu)$, $L(H_1;\mu)$ and $L(H_2;\mu)$.

To analyze $L(Q_1;\mu)$, as in [4], we can write

$$Q_{1n} = \frac{1}{2}\left[\int_{n^{-\alpha}\leq H\leq 1-n^{-\alpha}} \frac{\partial J}{\partial x}(F, G)\, d(F_n - F)^2 + \frac{1}{n}\int_{n^{-\alpha}\leq H\leq 1-n^{-\alpha}} \frac{\partial J}{\partial x}(F, G)\, dF_n\right].$$

Hence using integration by parts, we need only show that for $1 \leq i \leq 5$,

$$EL^\gamma(V_i;\mu) < \infty \qquad \text{for} \quad 0 < \gamma < 1 - \mu - \alpha(\tfrac{1}{2} - \delta)\,, \tag{4.17}$$

where

$$\begin{aligned} V_{1n} &= \int_{n^{-\alpha}\leq H\leq 1-n^{-\alpha}} (F_n - F)^2 \left|\frac{\partial^2 J}{\partial x^2}(F, G)\right| dF\,,\\ V_{2n} &= \int_{n^{-\alpha}\leq H\leq 1-n^{-\alpha}} (F_n - F)^2 \left|\frac{\partial^2 J}{\partial x\,\partial y}(F, G)\right| dG\,,\\ V_{3n} &= K' n^{\alpha(\frac{3}{2}-\delta)}\{F_n H^{-1}(n^{-\alpha}) - FH^{-1}(n^{-\alpha})\}^2\,,\\ V_{4n} &= K' n^{\alpha(\frac{3}{2}-\delta)}\{F_n H^{-1}(1 - n^{-\alpha}) - FH^{-1}(1 - n^{-\alpha})\}^2\,,\\ V_{5n} &= n^{-1}\int_{n^{-\alpha}\leq H\leq 1-n^{-\alpha}} \left|\frac{\partial J}{\partial x}(F, G)\right| dF_n\,. \end{aligned} \tag{4.18}$$

Since $\{n^2 V_{1n},\ 1 \leq n \leq m\}$ is a submartingale, we obtain using (4.10) and (4.11) that

$$\begin{aligned} &P[\max_{1\leq n\leq m} n^2 V_{1n} \geq \tfrac{1}{4} m^{2-\mu}]\\ &\qquad \leq 4m^\mu EV_{1m}\\ &\qquad = 4m^{\mu-1}\int_{m^{-\alpha}\leq H\leq 1-m^{-\alpha}} F(1 - F)\left|\frac{\partial^2 J}{\partial x^2}(F, G)\right| dF = O(m^{\mu+\alpha(\frac{1}{2}-\delta)-1})\,. \end{aligned}$$

Hence by Lemma 3(iii), (4.17) holds for $i = 1$, and we can similarly show that (4.17) holds for $i = 2$. From (3.5), it is easy to see that (4.17) also holds for $i = 3, 4$. Since $n^2 V_{5n}$ is increasing in $n$, we obtain by (4.10) and (4.11) that

$$\begin{aligned} &P[\max_{1\leq n\leq m} n^2 V_{5n} \geq \tfrac{1}{4} m^{2-\mu}]\\ &\qquad \leq 4m^\mu EV_{5m}\\ &\qquad = 4m^{\mu-1}\int_{m^{-\alpha}\leq H\leq 1-m^{-\alpha}} \left|\frac{\partial J}{\partial x}(F, G)\right| dF = O(m^{\mu+\alpha(\frac{1}{2}-\delta)-1})\,, \end{aligned}$$

and so (4.17) also holds for $i = 5$.

Now consider $L(Q_2;\mu)$. We note that

$$\begin{aligned} &P[\max_{1\leq n\leq \nu} n^2|Q_{2n}| \geq \tfrac{1}{4}\nu^{2-\mu}]\\ &\quad \leq P\left[\max_{1\leq n\leq \nu} nm_n\left|\int_{\nu^{-\alpha}\leq H\leq 1-\nu^{-\alpha}} (G_{m_n} - G)\frac{\partial J}{\partial y}(F, G)\, d(F_n - F)\right| \geq \varepsilon_1\nu^{2-\mu}\right]\\ &\qquad + P\left[\max_{1\leq n\leq \nu} n^2\int_{\nu^{-\alpha}\leq H\leq n^{-\alpha}} (G_{m_n} + G)\left|\frac{\partial J}{\partial y}(F, G)\right| d(F_n + F) \geq \varepsilon_2\nu^{2-\mu}\right]\\ &\qquad + P\left[\max_{1\leq n\leq \nu} n^2\int_{\nu^{-\alpha}\leq 1-H\leq n^{-\alpha}} (G_{m_n} + G)\left|\frac{\partial J}{\partial y}(F, G)\right| d(F_n + F) \geq \varepsilon_3\nu^{2-\mu}\right]\\ &\quad = a_\nu + b_\nu + c_\nu\,, \quad \text{say.} \end{aligned}$$

By the martingale inequality,

$$\begin{aligned} a_\nu &\leqq O(\nu^{2\mu})E\left|\int_{\nu^{-\alpha}\leqq H\leqq 1-\nu^{-\alpha}} (G_{m_\nu} - G)\frac{\partial J}{\partial y}(F, G)\,d(F_\nu - F)\right|^2 \\ &= O(\nu^{2\mu-2})\left\{\int_{\nu^{-\alpha}\leqq H\leqq 1-\nu^{-\alpha}} G(1-G)\left(\frac{\partial J}{\partial y}(F, G)\right)^2 dF \right. \\ &\quad - 2\iint_{s<t,\,\nu^{-\alpha}\leqq H(s),H(t)\leqq 1-\nu^{-\alpha}} G(s)(1-G(t))\frac{\partial J}{\partial y}(F(s), G(s)) \\ &\quad \left. \times \frac{\partial J}{\partial y}(F(t), G(t))\,dF(s)\,dF(t)\right\} \\ &= O(\nu^{2\mu-2+\alpha(1-2\delta)}), \quad \text{by (4.10) and (4.11).} \end{aligned}$$

Noting that

$$\begin{aligned} n^2 \int_{\nu^{-\alpha}\leqq H\leqq n^{-\alpha}} (G_{m_n} + G)\left|\frac{\partial J}{\partial y}(F, G)\right| d(F_n + F) \\ \leqq K'\nu^{\alpha(\frac{3}{2}-\delta)}\{n(G_{m_n} + G)(H^{-1}(n^{-\alpha}))\}\{n(F_n + F)(H^{-1}(n^{-\alpha}))\}, \end{aligned}$$

and that $\alpha(\frac{3}{2} - \delta) + 2(1 - \alpha) < 2 - \mu$, we can handle $b_\nu$ using (3.6). Likewise we can treat $c_\nu$. Therefore $EL^r(Q_2; \mu) < \infty$ for $0 < \gamma < 2\{1 - \mu - \alpha(\frac{1}{2} - \delta)\}$.

To consider $L(Q_3; \mu)$, we note that if $n > \tilde{L}$ (where $\tilde{L}$ is defined in (4.8)), then $n^{-\alpha} \leqq H(x) \leqq 1 - n^{-\alpha}$ implies that $(F_n(x) \vee G_{m_n}(x)) \geqq \frac{1}{2}H(x)$ and $(1 - F_n(x)) \vee (1 - G_{m_n}(x)) \geqq \frac{1}{2}(1 - H(x))$. Since $(\hat{F}(x), \hat{G}(x))$ lies on the line segment joining $(F(x), G(x))$ and $(F_n(x), G_{m_n}(x))$, it then follows from (4.1) that

$$\left|\frac{\partial^2 J}{\partial x^2}(\hat{F}(x), \hat{G}(x))\right| \leqq K_1\{H(1 - H)\}^{-\frac{5}{2}+\delta}.$$

Hence $L(Q_3; \mu) \leqq L(Q_3^*; \mu) = \sup\{n \geqq 1 : |Q_{3n}^*| \geqq n^{-\mu}\}$, where

$$Q_{3n}^* = K_1 \int_{n^{-\alpha}\leqq H\leqq 1-n^{-\alpha}} (F_n - F)^2\{H(1 - H)\}^{-\frac{5}{2}+\delta}\,dF_n. \tag{4.19}$$

By the submartingale inequality,

$$\begin{aligned} &P[\max_{1\leqq n\leqq m} K_1 n^3 \textstyle\int_{n^{-\alpha}\leqq H\leqq 1-n^{-\alpha}} (F_n - F)^2\{H(1 - H)\}^{-\frac{5}{2}+\delta}\,dF_n \geqq \frac{1}{4}m^{3-\mu}] \\ &\quad \leqq 4K_1 m^\mu E \textstyle\int_{m^{-\alpha}\leqq H\leqq 1-m^{-\alpha}} (F_m - F)^2\{H(1 - H)\}^{-\frac{5}{2}+\delta}\,dF_m \\ &\quad = 4K_1 m^{\mu-1} \textstyle\int_{m^{-\alpha}\leqq H\leqq 1-m^{-\alpha}} \{F(1 - F) - m^{-1}F(1 - 2F)\}\{H(1 - H)\}^{-\frac{5}{2}+\delta}\,dF \\ &\quad = O(m^{\mu-1+\alpha(\frac{1}{2}-\delta)}), \quad \text{by (4.11).} \end{aligned}$$

Therefore by Lemma 3(iii), $EL^r(Q_3^*, \mu) < \infty$ and so $EL^r(Q_3; \mu) < \infty$ for $0 < \gamma < 1 - \mu - \alpha(\frac{1}{2} - \delta)$. In a similar way, we can show that the same conclusion also holds for $L(Q_4; \mu)$ and $L(Q_5: \mu)$.

From the above analysis, we see that in the case $\delta < \frac{1}{2}$, if $0 < \mu < \frac{1}{2} + \delta$, then under the Assumptions $(A_\delta)$ and $(B_\rho)$ with $\rho \geqq \mu$, $EL^r(\mu) < \infty$ for $0 < \gamma < \alpha(\frac{1}{2} + \delta) - \mu$. Since $2\mu/(1 + 2\delta) < \alpha < 1$ is arbitrary, $EL^r(\mu) < \infty$ for $0 < \gamma < \frac{1}{2} + \delta - \mu$. The same conclusion obviously still holds when $\delta = \frac{1}{2}$ since in this case, Assumption $(A_{\delta'})$ holds for any $\delta' < \frac{1}{2}$.

We now prove (iii). Since $\frac{1}{2} < \delta < \frac{5}{2}$, Assumption $(A_\delta)$ implies Assumption $(A_{\frac{1}{2}})$ and therefore for $0 < \mu < 1$, under Assumptions $(A_\delta)$ and $(B_\rho)$ with $\rho \geqq \mu$, we have $EL^\gamma(\mu) < \infty$ for $0 < \gamma < 1 - \mu$. Let us now consider the case $0 < \mu < \mu(\delta) = (1 + 2\delta)(9 - 2\delta)/2(17 - 2\delta)$. First we note that since $0 < \mu < \mu(\delta)$, $(\frac{5}{2} - \delta)(2\mu/(1 + 2\delta)) < (1 - \mu)\{1 + 2(\frac{5}{2} - \delta)^{-1}\}$, and in the case $\frac{1}{2} < \delta < \frac{3}{2}$, we also have $(\frac{3}{2} - \delta)(2\mu/(1 + 2\delta)) < 1 - \mu$. Hence we can choose $\alpha, \beta > 0$ such that

$$2\mu/(1 + 2\delta) < \alpha < 1 \quad \text{and in the case} \quad \delta < \tfrac{3}{2}, \quad \alpha(\tfrac{3}{2} - \delta) < 1 - \mu ; \tag{4.20}$$

$$\alpha(\tfrac{5}{2} - \delta) < 1 - \mu + 2\beta , \qquad \beta(\tfrac{5}{2} - \delta) < 1 - \mu . \tag{4.21}$$

Since $(A_{\frac{3}{2}})$ implies $(A_{\delta'})$ for any $\delta' < \frac{3}{2}$, we shall assume below that $\delta \neq \frac{3}{2}$. By Lemma 4, we can write $J = g + h$ where $g$ satisfies Assumption $(A_0{}^*)$ and $h$ satisfies (3.7) (with $h^{(i)}$ replacing $f^{(i)}$) for $i = 0, 1, 2$. Hence without loss of generality, we can assume that $J$ satisfies (4.10) for $i = 0, 1, 2$. As in our proof of (ii), we express $\int_{-\infty}^{\infty} J(F_n, G_{m_n})\, dF_n$ by (4.9) and obtain that $EL^\gamma(D_j; \mu) < \infty$ for all $\gamma > 0$, $j = 1, 2$. We note that

$$|D_{3n}| \leqq K \int_{H<n^{-\alpha}} H^{\delta-\frac{1}{2}}\, dF_n \leqq K_1 n^{-\alpha(\delta-\frac{1}{2})} F_n H^{-1}(n^{-\alpha}) .$$

Hence by (3.6), $EL^\gamma(D_3; \mu) < \infty$ for all $\gamma > 0$. Likewise the same conclusion also holds for $L(D_4; \mu)$. To see that the same conclusion also holds for $H_{3n}$, say, we note that $\frac{3}{2} - \delta < 1$ and

$$|H_{3n}| \leqq K_1\{\max_{H(x)<n^{-\alpha}} |G_{m_n}(x) - G(x)|\} \int_{H<n^{-\alpha}} H^{-(\frac{3}{2}-\delta)}\, dF ,$$

and so by (3.5) and (4.11), we obtain the desired conclusion.

To analyze $L(Q_2; \mu)$, we make use of Lemma 1. Setting $u_n(x) = (\partial J/\partial y)(F(x), G(x))$ if $n^{-\alpha} \leqq H(x) \leqq 1 - n^{-\alpha}$ and $u_n(x) = 0$ if otherwise, we have by (4.10) that $\max_x |u_n(x)| \leqq K_1 n^{\alpha(\frac{3}{2}-\delta)^+}$. Since $\alpha(\frac{3}{2} - \delta)^+ < 1 - \mu$ by (4.20), it follows from Lemma 1 that $EL^\gamma(Q_2; \mu) < \infty$ for all $\gamma > 0$.

Now consider $L(Q_1; \mu)$. As before we shall show that $EL^\gamma(V_i; \mu) < \infty$ for all $\gamma > 0$ and $i = 1, \cdots, 5$, where $V_{in}$ is defined by (4.18). The case for $V_{3n}$ and $V_{4n}$ can be handled using (3.5) as before. We note that $|V_{5n}| \leqq K_1 n^{\alpha(\frac{3}{2}-\delta)^+ - 1} = o(n^{-\mu})$ by our choice of $\alpha$. To analyze $V_{1n}$ (and in a similar way $V_{2n}$ as well), we note that

$$\begin{aligned} &\int_{n^{-\alpha}\leqq H\leqq n^{-\beta}} (F_n - F)^2 \left|\frac{\partial^2 J}{\partial x^2}(F, G)\right| dF \\ &\qquad \leqq K_1 n^{\alpha(\frac{5}{2}-\delta)-\beta} \max_{H(x)\leqq n^{-\beta}} (F_n(x) - F(x))^2 ; \end{aligned} \tag{4.22}$$

$$\int_{n^{-\beta}<H<1-n^{-\beta}} (F_n - F)^2 \left|\frac{\partial^2 J}{\partial x^2}(F, G)\right| dF \leqq K_1 n^{\beta(\frac{5}{2}-\delta)} \|F_n - F\|^2 , \tag{4.23}$$

$$\begin{aligned} &\int_{n^{-\alpha}\leqq 1-H\leqq n^{-\beta}} (F_n - F)^2 \left|\frac{\partial^2 J}{\partial x^2}(F, G)\right| dF \\ &\qquad \leqq K_1 n^{\alpha(\frac{5}{2}-\delta)-\beta} \max_{1-H(x)\leqq n^{-\beta}} (F_n(x) - F(x))^2 . \end{aligned} \tag{4.24}$$

Using (2.1), (3.5) and (4.21), it is easy to see that $EL^{\gamma}(V_1; \mu) < \infty$ for all $\gamma > 0$.

As shown before, we can consider $L(Q_3^*; \mu)$ instead of $L(Q_3; \mu)$, where $Q_{3n}^*$ is defined by (4.19). A similar analysis as in (4.22), (4.23) and (4.24) together with an application of (3.6) to deal with $F_n H^{-1}(n^{-\beta})$ and $1 - F_n H^{-1}(1 - n^{-\beta})$ shows that $EL^{\gamma}(Q_3^*; \mu) < \infty$ for all $\gamma > 0$. The same method can also be used to analyze $L(Q_4; \mu)$ and $L(Q_5; \mu)$.

**5. Moments of the stopping rule of certain sequential rank tests.** Let us first consider the rank-order SPRT of the null hypothesis $H_0: F = G$ against the Lehmann alternative $H_1: F = G^A$ described in Section 1. Sethuraman [15] has proved the asymptotic normality of the log likelihood ratio $l_n$ defined by (1.1). Letting $J(x, y) = \log(x + Ay)$, $0 < x, y \leqq 1$, then $J$ satisfies Assumption $(A_\delta)$ for $\delta = \frac{1}{2}$. Defining $S(A, F, G)$ as in (1.4), we note that

$$
\begin{aligned}
l_n - nS(A, F, G) &= -n\{(\textstyle\int J(F_n, G_n)\, dF_n - \int J(F, G)\, dF) \\
&\quad + (\textstyle\int J(F_n, G_n)\, dG_n - \int J(F, G)\, dG)\} \\
&\quad + \tfrac{1}{2}\log n + O(1) \\
&= -\textstyle\sum_{i=1}^n (\psi_1(X_i) - E\psi_1(X_i)) \\
&\quad - \textstyle\sum_{i=1}^n (\psi_2(Y_i) - E\psi_2(Y_i)) + nR_n\,,
\end{aligned} \tag{5.1}
$$

where choosing $u_0$ such that $F(u_0) + AG(u_0) = 1$, we define

$$
\begin{aligned}
\psi_1(u) &= J(F(u), G(u)) - \int_{u_0}^{u} \frac{\partial J}{\partial x}(F(t), G(t))(dF(t) + dG(t)) \\
&= (A - 1) \int_{u_0}^{u} dG(t)/(F(t) + AG(t))\,; \\
\psi_2(u) &= J(F(u), G(u)) - \int_{u_0}^{u} \frac{\partial J}{\partial y}(F(t), G(t))(dF(t) + dG(t)) \\
&= -(A - 1) \int_{u_0}^{u} dF(t)/(F(t) + AG(t))\,.
\end{aligned}
$$

By Theorem 2, $EL^{\gamma}(\mu, \varepsilon) < \infty$ for all $\varepsilon > 0$, $\frac{1}{2} < \mu < 1$ and $0 < \gamma < 1 - \mu$, where $L(\mu, \varepsilon) = \sup\{n \geqq 1 : |R_n| \geqq \varepsilon n^{-\mu}\}$. This implies that $\lim_{n\to\infty} n^{\mu} R_n = 0$ a.e. for any $\mu < 1$ and consequently we have the asymptotic normality, and what is more, the invariance principle and the law of the iterated logarithm for $l_n$. Another implication of this result is an asymptotic approximation for the stopping rule $N$ defined by (1.2) in the case $S(A, F, G) = 0$. In Section 1, we have mentioned an asymptotic expression for $EN^{\gamma}$ when $S(A, F, G) \neq 0$. The following theorem considers the case $S(A, F, G) = 0$.

THEOREM 3. *Let $N$ be defined by* (1.2) *and let* $S(A, F, G) = 0$. *Let $\Phi$ denote the distribution function of the standard normal distribution and let*

$$
\begin{aligned}
\sigma^2 = 2(A - 1)^2\{&\textstyle\iint_{x<y} [G(x)(1 - G(y))/W(x)W(y)]\, dF(x)\, dF(y) \\
&+ \textstyle\iint_{x<y} [F(x)(1 - F(y))/W(x)W(y)]\, dG(x)\, dG(y)\}
\end{aligned} \tag{5.2}
$$

*where $W = F + AG$. Let $0 < \nu < 1$. Then as $a \to \infty$ and $b \to \infty$ such that*

$a/(a+b) \to \nu$,

$$\begin{aligned} &\forall t > 0\,, \quad P[N > \sigma^{-2}(a+b)^2 t] \\ (5.3) \qquad &\quad \to \sum_{k=-\infty}^{\infty} \{\Phi(t^{-\frac{1}{2}}(2k+1-\nu)) - \Phi(t^{-\frac{1}{2}}(2k-\nu)) \\ &\qquad - \Phi(t^{-\frac{1}{2}}(2k+1+\nu)) + \Phi(t^{-\frac{1}{2}}(2k+\nu))\} \\ &\quad = U(t;\nu)\,, \quad say, \end{aligned}$$

$$(5.4) \qquad EN^\gamma \sim \gamma(a+b)^{2\gamma}\sigma^{-2\gamma} \int_0^\infty t^{\gamma-1} U(t;\nu)\,dt\,, \quad 0 < \gamma < \tfrac{1}{2}\,.$$

PROOF. We note that $E|\psi_1(X_1)|^p < \infty$ and $E|\psi_2(Y_1)|^p < \infty$ for all $p > 0$ (see the remark to the corollary in Section 4) and that $\sigma^2 = \text{Var}\,\psi_1(X_1) + \text{Var}\,\psi_2(Y_1)$ (cf. [15], page 1332). Since $S(A, F, G) = 0$, Theorem 3 follows easily from the representation (5.1) and Lemma 6 below. The proof of Lemma 6 is presented in [9]. It is well known (cf. [7], page 329) that $U(t;\nu) = P[T > t]$ where $T$ is defined in Lemma 6.

LEMMA 6. *Suppose* $Z_1, Z_2, \cdots$ *are* i.i.d. *random variables such that* $EZ_1 = 0$, $EZ_1^2 = \sigma^2 > 0$. *Let* $S_n = Z_1 + \cdots + Z_n$, *and let* $V_n$ *be a sequence of random variables. Define* $N = \inf\{n \geq 1 : S_n + V_n \notin (-a, b)\}$ $(\inf \emptyset = \infty)$ *and let* $0 < \nu < 1$.

(i) *If* $\lim_{n\to\infty} n^{-\frac{1}{2}} V_n = 0$ a.e., *then as* $a \to \infty$ *and* $b \to \infty$ *such that* $a/(a+b) \to \nu$, $\sigma^2(a+b)^{-2}N$ *converges in distribution to* $T = \inf\{t \geq 0 : B(t) \notin (-\nu, 1-\nu)\}$, *where* $B(t)$, $t \geq 0$, *stands for the standard Wiener process.*

(ii) *Suppose* $E|Z_1|^{2+\eta} < \infty$ *for some* $\eta > 0$ *and* $EL^\gamma(\zeta) < \infty$ *for some* $\gamma > 0$ *and* $\zeta < \frac{1}{2}$, *where* $L(\zeta) = \sup\{n \geq 1 : |V_n| \geq n^\zeta\}$ $(\sup \emptyset = 0)$. *Then* $\{(\sigma^2(a+b)^{-2}N)^\gamma : a, b \geq 1\}$ *is uniformly integrable and so as* $a \to \infty$ *and* $b \to \infty$ *such that* $a/(a+b) \to \nu$,

$$E(\sigma^2(a+b)^{-2}N)^\gamma \to ET^\gamma\,.$$

In [11], Miller has introduced a sequential Wilcoxon test in the one-sample case. Let $X_1, X_2, \cdots$ be i.i.d. with a common continuous distribution function $F$. To test the null hypothesis $H$ that $X_1$ is symmetric, Miller ([11], page 99) proposes the following truncated sequential Wilcoxon test. Let $SR_n$ be the Wilcoxon signed rank statistic based on the first $n$ observations. Continue sampling if and only if $|SR_n| \leq cn$ and $n < k$. If $|SR_n| > cn$ for some $n \leq k$, then reject $H$; otherwise accept $H$. The constants $c$ and $k$ are so chosen that $P_H(\text{reject } H) \leq \alpha$, where $\alpha$ is a preassigned constant. In computing $P_H(\text{reject } H)$ and the expected sample size, Miller [11] has shown that the Wiener process approximation agrees well with the Monte Carlo estimates in certain numerical studies. In [12], Miller and Sen have established an invariance principle for $U$-statistics and this gives an asymptotic justification of the Wiener process approximation for $P_H(\text{reject } H)$. An asymptotic justification of the approximation for the moments of the stopping rule, however, needs more than the invariance principle. Here using our results in Section 4, we shall consider the asymptotic approximation for the corresponding two-sample problem.

Suppose $X_1, X_2, \cdots, Y_1, Y_2, \cdots$ are independent with continuous distribution functions $F$ and $G$ respectively. To test the null hypothesis $H_0: F = G$, consider the following test. At the $n$th stage, observe $(X_n, Y_n)$ and compute the Wilcoxon statistic $W_n =$ sum of ranks of $X_1, \cdots, X_n$ in the ordered sample of $2n$ observations. Under $H_0$, $EW_n = \frac{1}{2}n(2n+1)$. Therefore in analogy with Miller's one-sample case, we continue sampling as long as $|W_n - \frac{1}{2}n(2n+1)| \leqq cn$ and $n < k$. If $|W_n - \frac{1}{2}n(2n+1)| > cn$ for some $n \leqq k$, then reject $H_0$; otherwise accept $H_0$. The constants $c$ and $k$ are so chosen that $P_{H_0}(\text{reject } H_0) \leqq \alpha$.

THEOREM 4. *With the same notation as in the preceding paragraph, let* $M = \inf\{n : 1 \leqq n \leqq k, |W_n - \frac{1}{2}n(2n+1)| > cn\}$ ($\inf \emptyset = k$). *Suppose* $F = G$. *Then as* $c \to \infty$ *and* $k \to \infty$ *such that* $k/c^2 \to \zeta > 0$,

(i) $P[M > k\theta] \to U(\zeta\theta/24; \frac{1}{2})$ *for all* $\theta \in (0, 1)$ *where* $U(t; \nu)$ *is defined by* (5.3);
(ii) $EM^\gamma \sim \gamma(24c^2)^\gamma \int_0^{\zeta/24} t^{\gamma-1}U(t; \frac{1}{2})\, dt$ for all $\gamma > 0$.

PROOF. We can write $W_n = n^2 \int_{-\infty}^{\infty} J(F_n, G_n)\, dF_n$ where $J(x, y) = x + y$. Then since $F = G$, $\int_{-\infty}^{\infty} J(F, G)\, dF = 1$. Obviously $J$ satisfies Assumption $(A_0)$, and so by Theorem 2, if we write

$$\begin{aligned} n \int_{-\infty}^{\infty} J(F_n, G_n)\, dF_n &= n + \textstyle\sum_1^n (F(X_i) - \frac{1}{2}) - \sum_1^n (F(Y_i) - \frac{1}{2}) + nR_n \\ &= \tfrac{1}{2}(2n+1) + \textstyle\sum_1^n (F(X_i) - \frac{1}{2}) - \sum_1^n (F(Y_i) - \frac{1}{2}) + V_n , \end{aligned}$$

then $EL^\gamma(\mu, \varepsilon) < \infty$ for all $\gamma > 0$, $\varepsilon > 0$ and $\frac{1}{2} < \mu < 1$, where $L(\mu, \varepsilon) = \sup\{n \geqq 1 : |V_n| \geqq \varepsilon n^{1-\mu}\}$. Letting $S_n = \sum_{i=1}^n (F(X_i) - \frac{1}{2}) - \sum_1^n (F(Y_i) - \frac{1}{2})$, we note that $M = \min(N, k)$ where $N = \inf\{n \geqq 1 : |S_n + V_n| > c\}$. Hence by Lemma 6, as $c \to \infty$ and $k \to \infty$ such that $k/c^2 \to \zeta > 0$, $\sigma^2(2c)^{-2}M$ converges in distribution to $\min(T, \zeta\sigma^2/4)$ and $E(\sigma^2(2c)^{-2}M)^\gamma \to E(\min(T, \zeta\sigma^2/4))^\gamma$ for all $\gamma > 0$, where $T$ is as defined in Lemma 6 and $\sigma^2 = E(F(X_1) - \frac{1}{2})^2 + E(F(Y_1) - \frac{1}{2})^2 = \frac{1}{6}$. Hence the desired conclusion follows immediately.

## REFERENCES

[1] BERK, R. (1973). Some asymptotic aspects of sequential analysis. *Ann. Statist.* **1** 1126–1138.

[2] BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading.

[3] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

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

[5] CHOW, Y. S. and LAI, T. L. (1975). Some one-sided theorems on the tail distribution of sample sums with applications to the last time and largest excess of boundary crossings. To appear in *Trans. Amer. Math. Soc.*

[6] DVORETZKY, A., KIEFER, J. and WOLFOWITZ, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27** 642–669.

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

[8] KIEFER, J. (1972). Skorohod embedding of multivariate *RV*'s and the sample DF. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **24** 1–35.

[9] LAI, T. L. (1975). A note on first exit times with applications to sequential analysis. *Ann. Statist.* **3** 999–1005.

[10] MARCINKIEWICZ, J. et ZYGMUND, A. (1937). Sur les fonctions indépendantes. *Fund. Math.* **29** 60–90.

[11] MILLER, R. G. JR. (1972). Sequential rank tests—one sample case. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 97–108.

[12] MILLER, R. G., JR. and SEN, P. K. (1972). Weak convergence of $U$-statistics and von Mises' differentiable statistical functions. *Ann. Math. Statist.* **43** 31–41.

[13] SAVAGE, I. R. and SETHURAMAN, J. (1966). Stopping time of a rank-order sequential test based on Lehmann alternatives. *Ann. Math. Statist.* **37** 1154–1160.

[14] SEN, P. K. and GHOSH, M. (1972). On strong convergence of regression rank statistics. *Sankhyā Ser. A* **34** 335–348.

[15] SETHURAMAN, J. (1970). Stopping time of a rank-order sequential probability ratio test based on Lehmann alternatives II. *Ann. Math. Statist.* **41** 1322–1333.

[16] STRASSEN, V. (1967). Almost sure behavior of sums of independent random variables and martingales. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **2** 315–343.

[17] USPENSKY, J. V. (1937). *Introduction to Mathematical Probability*. McGraw-Hill, New York.

DEPARTMENT OF MATHEMATICAL STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027