The Annals of Mathematical Statistics
1972, Vol. 43, No. 4, 1122-1135

ASYMPTOTIC NORMALITY OF NONPARAMETRIC TESTS FOR INDEPENDENCE[1]

BY F. H. RUYMGAART, G. R. SHORACK[2] AND W. R. VAN ZWET

Mathematisch Centrum, Amsterdam and University of Leiden

Asymptotic normality of linear rank statistics for testing the hypothesis of independence is established under fixed alternatives. A generalization of a result of Bhuchongkul [1] is obtained both with respect to the conditions concerning the orders of magnitude of the score functions and with respect to the smoothness conditions on these functions.

1. Introduction. For each n let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be a random sample from a continuous bivariate distribution function (df) $H(x, y)$ having marginal dfs $F(x)$ and $G(y)$. The bivariate empirical df based on this sample is denoted by H_n. With respect to the n random variables (rvs) $X_i(Y_i)$ corresponding to the first (second) coordinates, the empirical df is denoted by $F_n(G_n)$, the ith order statistic by $X_{in}(Y_{in})$ and the rank of $X_i(Y_i)$ by $R_i(Q_i)$. All samples are defined on a single probability space $(\Omega, \mathcal{S}, P)$.

The rank statistics most commonly used to test the independence hypothesis $H = F.G$, are of the linear type

$$T_n = n^{-1} \sum_{i=1}^{n} a_n(R_i) b_n(Q_i) ,$$

where $a_n(i)$, $b_n(i)$ are real numbers for $i = 1, \cdots, n$ (see Hájek and Šidák [6]). A suitably standardized version of T_n will be (see also Bhuchongkul [1])

$$n^{\frac{1}{2}}(T_n - \mu) = n^{\frac{1}{2}}[\int\int J_n(F_n) K_n(G_n)\, dH_n - \mu] ; \tag{1.1}$$

here

$$J_n(s) = a_n(i) , \qquad K_n(s) = b_n(i) , \tag{1.2}$$

for $(i - 1)/n < s \leqq i/n$ and $i = 1, \cdots, n$, and

$$\mu = \int\int J(F) K(G)\, dH , \tag{1.3}$$

for some functions J and K on $(0, 1)$ that can be thought of as limits of the score functions J_n and K_n.

In order to summarize the main results of this paper let us introduce the function

$$r = [I(1 - I)]^{-1} \quad \text{on} \quad (0, 1) , \tag{1.4}$$

where I is the identity function on the unit interval. Under the hypothesis and under contiguous alternatives, asymptotic normality of (1.1) may be proved for score functions J and K of order $r^{\frac{1}{2}-\delta}$ for some $\delta > 0$ (see Hájek and Šidák [6]).

Received September 25, 1970; revised October 29, 1971.

[1] Report SW 5a/71 of the Department of Mathematical Statistics, Mathematisch Centrum, Amsterdam.

[2] This research was partially supported by NSF Contract No. GP-13739.

S. van de Geer and M. Wegkamp (eds.), *Selected Works of Willem van Zwet*, Selected Works in Probability and Statistics, DOI 10.1007/978-1-4614-1314-1_5, © Springer Science+Business Media, LLC 2012

Jogdeo [7] establishes asymptotic normality under the hypothesis of a statistic more general than T_n; the growth condition on his score functions in the case of T_n is $r^{\frac{1}{4}-\delta}$. By an approach analogous to that of Chernoff and Savage [3] for the two-sample problem, Bhuchongkul [1] proves asymptotic normality under fixed alternatives provided the score functions are of the order log r (see Section 2). The main purpose of this paper is to relax these conditions to $r^{\frac{1}{4}-\delta}$ in general and $r^{\frac{1}{2}-\delta}$ for a special class of dfs H.

In Theorem 2.1 the asymptotic normality of (1.1) is established for rather smooth score functions with orders of magnitude not exceeding r^a and r^b, where the numbers a and b satisfy the relations $a = (\frac{1}{2} - \delta)/p_0$ and $b = (\frac{1}{2} - \delta)/q_0$ for some $0 < \delta < \frac{1}{2}$ and some $p_0, q_0 > 1$ with $p_0^{-1} + q_0^{-1} = 1$. No condition other than continuity is imposed on the df H. The theorem is stronger than Theorem 1 of Bhuchongkul [1]. The proof is based on Hölder's inequality in the form

$$\int\int |\phi(F)\psi(G)|\, dH \leqq [\int |\phi|^p\, dI]^{1/p}[\int |\psi|^q\, dI]^{1/q}\,, \tag{1.5}$$

where ϕ and ψ are functions on $(0, 1)$, dI denotes Lebesgue measure restricted to the unit interval and $p, q > 1$ satisfy $p^{-1} + q^{-1} = 1$.

Theorem 2.2 gives asymptotic normality of (1.1) under much weaker conditions on the score functions. Here these functions are allowed to be of order r^a and r^b, where $a = b = \frac{1}{2} - \delta$ for some $0 < \delta < \frac{1}{2}$. The price for this is a condition on the df H, keeping it in some sense similar to the null hypothesis. This condition is

$$dH \leqq C[r(F)r(G)]^{\delta/2}\, dF\, dG\,, \tag{1.6}$$

with fixed constants $C \geqq 1$ and $0 < \delta < \frac{1}{2}$. Mathematically, (1.6) allows a direct factorization of the left-hand integral in (1.5) which is more efficient than Hölder's inequality. Intuitively, this condition prevents the large (small) X's from occurring in the same pair as large (small) Y's with too high a probability. Condition (1.6) trivially holds under the null hypothesis. More generally it is also satisfied if H can be written as a polynomial in its marginals F and G. This class of distributions was introduced by Lehmann [9] and the special case where $H = FG[1 + \alpha(1 - F)(1 - G)]$ for $-1 \leqq \alpha \leqq 1$ was considered by Gumbel [5]. Finally (1.6) holds for all bivariate normal distributions with a sufficiently small correlation coefficient (use Lemma 2 on page 166 of Feller [4] to see that (1.6) holds for a correlation coefficient between $-\delta/(2 - \delta)$ and $\delta/(2 - \delta)$).

2. Statement of the theorems. Each of the theorems below establishes the asymptotic normality

$$n^{\frac{1}{2}}(T_n - \mu) \to_d N(0, \sigma^2) \qquad \text{as} \quad n \to \infty\,, \tag{2.1}$$

of (1.1); here μ and σ^2 are finite and are given by (1.3) and (3.10) respectively.

Let $\mathscr{H}$ denote the class of all continuous bivariate dfs H, and let $\mathscr{H}_{C\delta}$ denote the subclass that satisfies (1.6) for fixed $C \geqq 1$ and $0 < \delta < \frac{1}{2}$.

To prove (2.1) for general H in $\mathscr{H}$ we require a strong boundedness condition on the score functions.

ASSUMPTION 2.1. The functions J and K are continuous on $(0, 1)$; each is differentiable except at an at most finite number of points, and in the open intervals between these points the derivatives are continuous. The functions J_n, K_n, J, K satisfy $|J_n| \leqq Dr^a$, $|K_n| \leqq Dr^b$ and

$$|J^{(i)}| \leqq Dr^{a+i}, \qquad |K^{(i)}| \leqq Dr^{b+i} \qquad \text{for} \quad i = 0, 1,$$

where defined on $(0, 1)$. Here D is a positive constant and a and b satisfy

$$a = (\tfrac{1}{2} - \delta)/p_0, \qquad b = (\tfrac{1}{2} - \delta)/q_0 \tag{2.2}$$

for some $0 < \delta < \frac{1}{2}$ and some $p_0, q_0 > 1$ with $p_0^{-1} + q_0^{-1} = 1$.
In proving (2.1) for the more restrictive class $\mathscr{H}_{C\delta}$ we only require a weak boundedness condition on the score functions.

ASSUMPTION 2.2. Assumption 2.1 holds with

$$a = b = \tfrac{1}{2} - \delta \tag{2.3}$$

for some $0 < \delta < \frac{1}{2}$.

We also need a condition on the convergence of J_n, K_n to J, K. Define

$$B_{0n} = n^{\frac{1}{2}} \int\int_{\Delta_n} [J_n(F_n)K_n(G_n) - J(F_n)K(G_n)]\, dH_n, \tag{2.4}$$

$$B_{0n}^* = n^{\frac{1}{2}} \int\int [J_n(F_n)K_n(G_n) - J(F_n{}^*)K(G_n{}^*)]\, dH_n, \tag{2.5}$$

where

$$\Delta_n = \Delta_{n1} \times \Delta_{n2} \qquad \text{with} \quad \Delta_{n1} = [X_{1n}, X_{nn}) \quad \text{and} \quad \Delta_{n2} = [Y_{1n}, Y_{nn}), \tag{2.6}$$

$$F_n{}^* = [n/(n+1)]F_n, \qquad G_n{}^* = [n/(n+1)]G_n, \tag{2.7}$$

ASSUMPTION 2.3. Either (a) $B_{0n} \to_p 0$ as $n \to \infty$, or (b) $B_{0n}^* \to_p 0$ as $n \to \infty$. This assumption is very general, but may occasionally be difficult to verify. However, most examples are special cases of Remarks 2.1 and 2.2 below.

REMARK 2.1. If the scores of (1.2) satisfy $a_n(i) = J(i/(n+1))$ and $b_n(i) = K(i/(n+1))$ for $1 \leqq i \leqq n$ for some functions J and K, then Assumption 2.3 (b) holds uniformly for H in $\mathscr{H}$. (In this case $B_{0n}^* = 0$ for all n.)

REMARK 2.2. Suppose that J and K are increasing and twice differentiable on $(0, 1)$, and that $|J^{(i)}| \leqq Dr^{a+i}$ and $|K^{(i)}| \leqq Dr^{b+i}$ for $i = 0, 1, 2$ where $D > 0$ and a and b satisfy (2.2). Let the scores $a_n(i)$ and $b_n(i)$ of (1.2) be the expectations of the ith order statistics of samples of size n from populations whose dfs are the inverse functions of J and K respectively. Then Assumption 2.1 holds and Assumption 2.3 (a) holds uniformly for all H in $\mathscr{H}$. (This statement generalizes Theorem 2 of [1] and the proof may be given in the same way. It relies mainly on the fact that $\sum_{i=1}^{n-1} |a_n(i) - J(i/n)| = O(n^a)$ and $\sum_{i=1}^{n-1} |b_n(i) - K(i/n)| = O(n^b)$, which follows from formulas (7.14) and (7.24) of [3] with $\alpha = a$ and $\alpha = b$ respectively.)

THEOREM 2.1. *If H is in $\mathscr{H}$ and if Assumptions* 2.1 *and* 2.3 *are satisfied, then the asymptotic normality* (2.1) *holds. Given any subclass $\mathscr{H}'$ of $\mathscr{H}$ such that As-*

sumption 2.3 *holds uniformly for* H *in* $\mathscr{H}'$ *and such that* $\sigma^2 = \sigma^2(H)$ *is bounded away from* 0 *on* $\mathscr{H}'$, *the convergence in* (2.1) *is uniform for* H *in* $\mathscr{H}'$.

Note that (2.2) is satisfied if $a = b = \frac{1}{4} - \varepsilon$ for some $0 < \varepsilon < \frac{1}{4}$ (take $p_0 = q_0 = 2$ and $\delta = 2\varepsilon$). Thus Theorem 2.1 allows a rate of growth $r^{\frac{1}{4}-\varepsilon}$ for the score functions J and K and $r^{\frac{3}{4}-\varepsilon}$ for their derivatives. In Theorem 1 of [1] these rates are $r^{\frac{1}{8}-\varepsilon}$ and r respectively; in fact the latter condition reduces the rate for J and K to $\log r$. Moreover in [1] the score functions are assumed to be twice differentiable throughout the unit interval.

THEOREM 2.2. *Fix* $C \geqq 1$ *and* $0 < \delta < \frac{1}{2}$. *If* H *is in* $\mathscr{H}_{C\delta}$ *and if Assumptions* 2.2 *and* 2.3 *are satisfied, then the asymptotic normality* (2.1) *holds. Given any subclass* $\mathscr{H}'_{C\delta}$ *of* $\mathscr{H}_{C\delta}$ *such that Assumption* 2.3 *holds uniformly for* H *in* $\mathscr{H}'_{C\delta}$ *and such that* $\sigma^2 = \sigma^2(H)$ *is bounded away from* 0 *on* $\mathscr{H}'_{C\delta}$, *the convergence in* (2.1) *is uniform for* H *in* $\mathscr{H}'_{C\delta}$.

3. Proof of the theorems: Asymptotic normality of the leading terms. Let $F^{-1}(s) = \inf\{x\colon F(x) \geqq s\}$ and $G^{-1}(t) = \inf\{y\colon G(y) \geqq t\}$; these definitions imply $F(F^{-1}) = G(G^{-1}) = I$. The random functions $F_n(F^{-1})$ and $G_n(G^{-1})$ are with probability 1 the empirical dfs of the sets of independent uniform (0, 1) rvs $F(X_1), \cdots, F(X_n)$ and $G(Y_1), \cdots, G(Y_n)$ respectively. Define the empirical processes $U_n = n^{\frac{1}{2}}[F_n(F^{-1}) - I]$ and $V_n = n^{\frac{1}{2}}[G_n(G^{-1}) - I]$ on $[0, 1]$. With probability 1 these processes satisfy $U_n(F) = n^{\frac{1}{2}}(F_n - F)$ and $V_n(G) = n^{\frac{1}{2}}(G_n - G)$ on $(-\infty, \infty)$. All of the above remarks follow from the fact that

$$P(\Omega_0) = P(\{\omega\colon F_n(F^{-1}(F)) = F_n,\ G_n(G^{-1}(G)) = G_n \text{ for all } x, y \text{ and } n\}) = 1\,. \tag{3.1}$$

Without loss of generality we shall prove Theorems 2.1 and 2.2 in the case where both J and K fail to have a derivative at just one point, say at s_1 and t_1 respectively. For small positive γ define the sets

$$S_{\gamma 1} = [F^{-1}(\gamma), F^{-1}(s_1 - \gamma)] \cup [F^{-1}(s_1 + \gamma), F^{-1}(1 - \gamma)]\,. \tag{3.2}$$
$$S_{\gamma 2} = [G^{-1}(\gamma), G^{-1}(t_1 - \gamma)] \cup [G^{-1}(t_1 + \gamma), G^{-1}(1 - \gamma)]\,,$$

$$\Omega_{\gamma n} = \{\omega\colon \sup|F_n - F| < \gamma/2,\ \sup|G_n - G| < \gamma/2\}\,. \tag{3.3}$$

Let $S_\gamma = S_{\gamma 1} \times S_{\gamma 2}$ be the product set in the plane and let $\chi(\Omega_{\gamma n})$ denote the indicator function of $\Omega_{\gamma n}$. For ω in $\Omega_0 \cap \Omega_{\gamma n}$ the mean value theorem gives

$$n^{\frac{1}{2}}J(F_n) = n^{\frac{1}{2}}J(F) + U_n(F)J'(\Phi_n)$$

for all x in $\Delta_{n1} \cap S_{\gamma 1}$. In the above formula the function Φ_n is defined by $\Phi_n = F + \theta(F_n - F)$, where $\theta = \theta(\omega, x, n)$ is a number between 0 and 1. Thus with probability 1 (using Assumption 2.3 (a))

$$n^{\frac{1}{2}}(T_n - \mu) = \textstyle\sum_{i=1}^{3} A_{in} + \sum_{i=0}^{2} B_{in} + \sum_{i=3}^{7} B_{\gamma in} + B_{8n} + C_n\,, \tag{3.4}$$

where B_{0n} is defined in (2.4) and where

$$A_{1n} = n^{\frac{1}{2}} \int\int J(F)K(G)d(H_n - H) ,$$

$$A_{2n} = \int\int U_n(F)J'(F)K(G)\, dH , \qquad A_{3n} = \int\int V_n(G)J(F)K'(G)\, dH ,$$

$$B_{1n} = n^{\frac{1}{2}} \int\int_{\Delta_n{}^c} J_n(F_n)K_n(G_n)\, dH_n , \qquad B_{2n} = -n^{\frac{1}{2}} \int\int_{\Delta_n{}^c} J(F)K(G)\, dH_n ,$$

$$B_{\gamma 3n} = \chi(\Omega_{\gamma n}^c)\{n^{\frac{1}{2}} \int\int_{\Delta_n} [J(F_n) - J(F)]K(G)dH_n - A_{2n}\} ,$$

$$B_{\gamma 4n} = \chi(\Omega_{\gamma n})n^{\frac{1}{2}} \int\int_{\Delta_n \cap S_\gamma{}^c} [J(F_n) - J(F)]K(G)\, dH_n ,$$

$$B_{\gamma 5n} = \chi(\Omega_{\gamma n}) \int\int_{\Delta_n \cap S_\gamma} U_n(F)[J'(\Phi_n) - J'(F)]K(G)\, dH_n ,$$

$$B_{\gamma 6n} = \chi(\Omega_{\gamma n}) \int\int_{\Delta_n \cap S_\gamma} U_n(F)J'(F)K(G)d(H_n - H) ,$$

$$B_{\gamma 7n} = -\chi(\Omega_{\gamma n}) \int\int_{\Delta_n{}^c \cup S_\gamma{}^c} U_n(F)J'(F)K(G)\, dH ,$$

$$B_{8n} = n^{\frac{1}{2}} \int\int_{\Delta_n} J(F)[K(G_n) - K(G)]dH_n - A_{3n} ,$$

$$C_n = n^{\frac{1}{2}} \int\int_{\Delta_n} [J(F_n) - J(F)][K(G_n) - K(G)]\, dH_n .$$

Let us note that

$$\sum_{i=3}^{7} B_{\gamma in} = n^{\frac{1}{2}} \int\int_{\Delta_n} [J(F_n) - J(F)]K(G)dH_n - A_{2n} ,$$

which is symmetric to B_{8n}. For this reason B_{8n} will not be treated in the sequel.

We now proceed to prove the asymptotic normality of the A-terms. Let us start with the very useful remark that if a and b satisfy (2.2), then for $i = 1$ and 2 we can find numbers $p_i, q_i > 1$ satisfying $p_i{}^{-1} + q_i{}^{-1} = 1$ and

$$(3.5) \qquad (a + \tfrac{1}{2} + \delta/2)p_1 < 1 , \qquad bq_1 < 1 , \qquad ap_2 < 1 , \qquad (b + \tfrac{1}{2} + \delta/2)q_2 < 1 .$$

As to the first pair of inequalities, we have $a + \frac{1}{2} + \delta/2 + b = 1 - \delta/2$ and consequently $a + \frac{1}{2} + \delta/2 < 1 - \delta/2$ (the numbers a and b are strictly positive). Now choose $p_1 = (a + \frac{1}{2} + 3\delta/4)^{-1}$ and let $q_1 = (1 - p_1{}^{-1})^{-1}$. Then $(a + \frac{1}{2} + \delta/2)p_1 < 1$ and $bq_1 = (\frac{1}{2} - a - \delta)/(\frac{1}{2} - a - 3\delta/4) < 1$. The second pair of inequalities can be obtained in the same way.

The rv A_{1n} can be written in the form

$$(3.6) \qquad A_{1n} = n^{-\frac{1}{2}} \sum_{i=1}^{n} A_{1in} ,$$

where $A_{1in} = J(F(X_i))K(G(Y_i)) - \mu$ are independent and identically distributed (i.i.d.) with mean zero. Under Assumption 2.1 application of (1.5) with $p = p_0$ and $q = q_0$ shows that the rv A_{1in} has a finite absolute moment of order $2 + \delta_0$ for some $\delta_0 > 0$. The same conclusion holds under Assumption 2.2 for H in $\mathscr{H}_{C\delta}$ as may be seen by applying (1.6). Moreover this moment will be uniformly bounded above for H within $\mathscr{H}(\mathscr{H}_{C\delta})$.

Because

$$U_n(F) = n^{-\frac{1}{2}} \sum_{i=1}^{n} (\phi_{X_i} - F) ,$$

where

$$(3.7) \qquad \phi_{X_i}(x) = 0 \quad \text{if} \quad x < X_i \qquad \text{and} \qquad \phi_{X_i}(x) = 1 \quad \text{if} \quad x \geqq X_i ,$$

we have

$$(3.8) \qquad A_{2n} = n^{-\frac{1}{2}} \sum_{i=1}^{n} A_{2in} ,$$

where the $A_{2in} = \int\int (\phi_{X_i} - F)J'(F)K(G)\,dH$ are i.i.d. with mean zero. Under Assumptions 2.1 or 2.2 we have

$$|A_{2in}| \leqq D^2 r^{\frac{1}{2}-\delta/4}(F(X_i)) \int\int r^{a+\frac{1}{2}+\delta/4}(F) r^b(G)\,dH\,.$$

For some $\delta_1 > 0$ the random part of this upper bound possesses an absolute moment of order $2 + \delta_1$ which is uniformly bounded above for H in $\mathscr{H}$. Under Assumption 2.1 the nonrandom integral is seen to be uniformly bounded above for H in $\mathscr{H}$ by application of (1.5) with $p = p_1$ and $q = q_1$ as in (3.5). Uniform boundedness of this integral for H in $\mathscr{H}_{C\delta}$ holds under Assumption 2.2, as may be shown by application of (1.6).

Analogously we can write

$$A_{3n} = n^{-\frac{1}{2}} \textstyle\sum_{i=1}^{n} A_{3in}\,, \tag{3.9}$$

where $A_{3in} = \int\int (\phi_{Y_i} - G)J(F)K'(G)\,dH$ are i.i.d. with mean zero. Again for $\delta_1 > 0$ this rv has a finite absolute moment of order $2 + \delta_1$ which is uniformly bounded for H in $\mathscr{H}(\mathscr{H}_{C\delta})$. This time use (1.5) with $p = p_2$ and $q = q_2$ as in (3.5).

Combining (3.6), (3.8) and (3.9) we get $\sum_{i=1}^{3} A_{in} \to_d N(0, \sigma^2)$ as $n \to \infty$. The variance σ^2 is given by (see [1])

$$\begin{aligned}\sigma^2 = \operatorname{Var}[&J(F(X))K(G(Y)) + \int\int (\phi_X - F)J'(F)K(G)\,dH \\ &+ \int\int (\phi_Y - G)J(F)K'(G)\,dH]\,,\end{aligned} \tag{3.10}$$

with ϕ defined in (3.7).

Since we have shown that an absolute moment of order larger than 2 exists and is uniformly bounded on $\mathscr{H}(\mathscr{H}_{C\delta})$, and because the variance is uniformly bounded away from zero on $\mathscr{H}'(\mathscr{H}'_{C\delta})$, the established convergence in distribution is uniform for H in $\mathscr{H}'(\mathscr{H}'_{C\delta})$ by Esséen's theorem (see e.g. [3], Section 4).

4. Some lemmas. We start with a number of lemmas to be used in the proofs of both Theorem 2.1 and Theorem 2.2.

LEMMA 4.1. *For any $\zeta \geqq 0$ the function r^ζ is symmetric about $\frac{1}{2}$, decreasing on $(0, \frac{1}{2}]$ and has the property that for each β in $(0, 1)$ there exists a constant $M = M_\beta$ such that $r^\zeta(\beta s) \leqq M r^\zeta(s)$ for $0 < s \leqq \frac{1}{2}$ and $r^\zeta(1 - \beta(1 - s)) \leqq M r^\zeta(s)$ for $\frac{1}{2} < s < 1$.*

PROOF. On $(0, \frac{1}{2}]$ we have $r^\zeta(\beta s) = (\beta s)^{-\zeta}(1 - \beta s)^{-\zeta} \leqq \beta^{-\zeta} r^\zeta(s)$. A similar argument applies to the interval $(\frac{1}{2}, 1)$. □

LEMMA 4.2. *For each ω let $\tilde{\Phi}_n = \tilde{\Phi}_{n\omega}$ and $\tilde{\Psi}_n = \tilde{\Psi}_{n\omega}$ be functions on $\Delta_{n1} = \Delta_{n1\omega}$ and $\Delta_{n2} = \Delta_{n2\omega}$ respectively (see (2.6)), satisfying* $\min(F, F_n) \leqq \tilde{\Phi}_n \leqq \max(F, F_n)$ *and* $\min(G, G_n) \leqq \tilde{\Psi}_n \leqq \max(G, G_n)$ *where defined. Then uniformly for $n = 1, 2, \cdots$ and $H \in \mathscr{H}$:*

(i) $\sup_{\Delta_{n1}} r^\zeta(\tilde{\Phi}_n) r^{-\zeta}(F) = O_p(1)$ *for each* $\zeta \geqq 0$;
(ii) $\sup_{\Delta_{n2}} r^\eta(\tilde{\Psi}_n) r^{-\eta}(G) = O_p(1)$ *for each* $\eta \geqq 0$;
(iii) $\sup_{(-\infty,\infty)} |U_n(F)| r^{\frac{1}{2}-\tau}(F) = O_p(1)$ *for each* $\tau > 0$.

PROOF. (i) From formula (3.1) and e.g. from [11], Lemma A.3 it follows that for each $\varepsilon > 0$ there exists a constant $\beta = \beta_\varepsilon$ in (0, 1) such that

$$P(\Omega_n) = P(\{\beta F \leqq F_n \leqq 1 - \beta(1 - F) \text{ on } \Delta_{n1}\}) > 1 - \varepsilon\,, \tag{4.1}$$

for all n and uniformly in all continuous F. Because of the definition of $\tilde{\Phi}_n$ we have $\beta F \leqq \tilde{\Phi}_n \leqq 1 - \beta(1 - F)$ on Δ_{n1}. By Lemma 4.1 this implies that for some constant M_{ζ_ε} we have $r^\zeta(\tilde{\Phi}_n) \leqq M_{\zeta_\varepsilon} r^\zeta(F)$ for x in Δ_{n1} on the set Ω_n.

(ii) This is analogous to (i).

(iii) This follows immediately from Lemma 2.2 of Pyke and Shorack [10]. □

For each positive integer k we define a function I_k on [0, 1] by

$$I_k(0) = 0\,, \qquad I_k(s) = (i-1)/k \quad \text{for} \quad (i-1)/k < s \leqq i/k\,, \quad i = 1, \cdots, k\,, \tag{4.2}$$

LEMMA 4.3. *As $k, n \to \infty$, $\sup_{(-\infty,\infty)} |U_n(I_k(F)) - U_n(F)| \to_p 0$ uniformly in all continuous F.*

PROOF. Note that $\sup_{-\infty<x<\infty} |U_n(I_k(F)) - U_n(F)| = \sup_{0\leqq s\leqq 1} |U_n(I_k) - U_n|$, which is no longer dependent on F. The U_n-processes converge weakly to a tied-down Wiener process U_0 (see e.g. Billingsley [2]). In Pyke and Shorack [10] these U_n- and U_0-processes are replaced by $\tilde{U}_n$- and $\tilde{U}_0$-processes defined on a single new probability space $(\tilde{\Omega}, \tilde{\mathscr{A}}, \tilde{P})$ and having the same finite dimensional distributions as the original processes (see also Skorokhod [12]). These new processes satisfy $\sup |\tilde{U}_n - \tilde{U}_0| \to_{\text{a.s.}} 0$ and hence also $\sup |\tilde{U}_n(I_k) - \tilde{U}_0(I_k)| \to_{\text{a.s.}} 0$ uniformly in k, as $n \to \infty$. Now $\sup |\tilde{U}_n(I_k) - \tilde{U}_n| \leqq \sup |\tilde{U}_n - \tilde{U}_0| + \sup |\tilde{U}_0 - \tilde{U}_0(I_k)| + \sup |\tilde{U}_0(I_k) - \tilde{U}_n(I_k)|$. For almost every $\tilde{\omega}$ the function $\tilde{U}_0$ is uniformly continuous on [0, 1] so that $\sup |\tilde{U}_0 - \tilde{U}_0(I_k)| \to_{\text{a.s.}} 0$ as $k \to \infty$. This proves that $\sup |\tilde{U}_n(I_k) - \tilde{U}_n| \to_{\text{a.s.}} 0$ for $k, n \to \infty$. This last result implies the convergence in probability of the lemma. □

Let ν and λ be the random indices $1 \leqq \nu(\omega), \lambda(\omega) \leqq n$ such that

$$X_\nu = X_{nn} \qquad \text{and} \qquad Y_\lambda = Y_{nn}\,. \tag{4.3}$$

LEMMA 4.4. *As $n \to \infty$, $P(\{\alpha_n \leqq F(X_\nu) \leqq 1 - \alpha_n\} \cap \{\alpha_n \leqq G(Y_\nu) \leqq 1 - \alpha_n\}) \to 1$ uniformly for H in $\mathscr{H}$ provided only $\alpha_n = o(n^{-1})$.*

PROOF. The probability of the complementary event is bounded above by $2{\alpha_n}^n + 2[1 - (1 - \alpha_n)^n] \to 0$ as $n \to \infty$, independently of H in $\mathscr{H}$. □

We conclude this section with some lemmas needed for Theorem 2.2.

LEMMA 4.5. *As $n \to \infty$, $P(\{Y_\nu = Y_{nn}\}) \to 0$ uniformly for H in $\mathscr{H}_{C\delta}$.*

PROOF. $P(\{Y_\nu = Y_{nn}\}) = P(\bigcup_{i=1}^n \{(X_i, Y_i) = (X_{nn}, Y_{nn})\}) = n \iint H^{n-1}\, dH$. Note that for all x, y we have $H(x, y) \leqq F(x)$ and $H(x, y) \leqq G(y)$. Letting $n_0 = (n-1)/2$ and applying (1.6) we obtain

$$\begin{aligned} n \iint H^{n-1}(x, y)\, dH(x, y) &\leqq n \iint F^{n_0}(x) G^{n_0}(y)\, dH(x, y) \\ &\leqq Cn[\textstyle\int_0^1 I^{n_0} r^{\delta/2}\, dI]^2 \\ &= Cn[\Gamma(n_0 - \delta/2 + 1)\Gamma(1 - \delta/2)/\Gamma(n_0 - \delta + 2)]^2 \\ &\leqq C_1 n n_0^{\delta-2} = O(n^{\delta-1}) \to 0 \end{aligned}$$

as $n \to \infty$, because $0 < \delta < \frac{1}{2}$. Here C_1 is a constant depending on C and δ only; hence the convergence is uniform for H in $\mathscr{H}_{C\delta}$. □

LEMMA 4.6. *As* $n \to \infty$, $P(\{\gamma_n \leqq G(Y_\nu) \leqq 1 - \gamma_n\}) \to 1$ *uniformly for* H *in* $\mathscr{H}_{C\delta}$ *provided* $\gamma_n \leqq an^{-\delta}$ *for some positive constant* a.

PROOF. This probability equals $1 - P(\{G(Y_\nu) < \gamma_n\}) - P(\{G(Y_\nu) > 1 - \gamma_n\})$ for n larger than $(2a)^{1/\delta}$. Because of the independence of the sample elements, application of (1.6) gives

$$\begin{aligned} P(\{G(Y_\nu) < \gamma_n\}) &= nP([\textstyle\bigcap_{i=1}^{n-1} \{F(X_i) \leqq F(X_n)\}] \cap \{G(Y_n) < \gamma_n\}) \\ &= n \int_{-\infty}^{\infty} \int_{-\infty}^{G^{-1}(\gamma_n)} F^{n-1}(x_n)\, dH(x_n, y_n) \\ &\leqq Cn[\textstyle\int_0^1 I^{n-1} r^{\delta/2}\, dI][\int_0^{\gamma_n} r^{\delta/2}\, dI] \\ &= C_1 n[\Gamma(n - \delta/2)/\Gamma(n + 1 - \delta)]n^{-\delta+\delta^2/2} \\ &\leqq C_2 nn^{-1+\delta/2}n^{-\delta+\delta^2/2} = C_2 n^{-\delta/2+\delta^2/2} \to 0 \end{aligned}$$

as $n \to \infty$, because $-\delta/2 + \delta^2/2 < 0$ for $0 < \delta < \frac{1}{2}$. Here C_1 and C_2 are constants depending on C, δ and a only; hence the convergence is uniform for H in $\mathscr{H}_{C\delta}$. □

5. Proof of the theorems: Asymptotic negligibility of the remainder terms under Assumption 2.3(a). Let us start with a further decomposition of C_n, which can be seen to be the sum of

$$\begin{aligned} C_{\gamma 1n} &= \chi(\Omega_{\gamma n}^c)n^{\frac{1}{2}} \iint_{\Delta_n} [J(F_n) - J(F)][K(G_n) - K(G)]\, dH_n\,, \\ C_{\gamma 2n} &= \chi(\Omega_{\gamma n})n^{\frac{1}{2}} \iint_{\Delta_n \cap S_\gamma{}^c} [J(F_n) - J(F)]K(G_n)\, dH_n\,, \\ C_{\gamma 3n} &= -\chi(\Omega_{\gamma n})n^{\frac{1}{2}} \iint_{\Delta_n \cap S_\gamma{}^c} [J(F_n) - J(F)]K(G)\, dH_n\,, \\ C_{\gamma 4n} &= \chi(\Omega_{\gamma n}) \iint_{\Delta_n \cap S_\gamma} U_n(F)J'(\Phi_n)[K(G_n) - K(G)]\, dH_n\,. \end{aligned}$$

From this we see that $B_{\gamma 4n}$ and $C_{\gamma 3n}$ cancel out. The asymptotic negligibility of the other B- and C-terms will be given as corollaries to the lemmas of the previous section.

COROLLARY 5.1. *As* $n \to \infty$, $B_{1n} \to_p 0$ *uniformly for* H *in* $\mathscr{H}(\mathscr{H}_{C\delta})$.

PROOF. The rv B_{1n} is bounded by $\sum_{i=1}^{3} B_{1in}$ where

$$\begin{aligned} B_{11n} &= n^{\frac{1}{2}}|J_n(1)| \iint_{\{X_{nn}\} \times \Delta_{n2}} |K_n(G_n(y))|\, dH_n(x, y)\,, \\ B_{12n} &= n^{\frac{1}{2}}|J_n(1)K_n(1)| \iint_{\{(X_{nn}, Y_{nn})\}} dH_n(x, y)\,, \\ B_{13n} &= n^{\frac{1}{2}}|K_n(1)| \iint_{\Delta_{n1} \times \{Y_{nn}\}} |J_n(F_n(x))|\, dH_n(x, y)\,. \end{aligned}$$

Under the assumptions of Theorem 2.1 we have at once that the sum of these terms is of order $O(n^{-\frac{1}{2}+a+b}) = O(n^{-\delta}) \to 0$ as $n \to \infty$, uniformly for H in $\mathscr{H}$.

Under the assumptions of Theorem 2.2 first consider B_{11n}. By Assumption 2.2, $|K_n(G_n(y))| \leqq Dr^b(G_n(y))$. Application of Lemma 4.2 (ii) with $\Psi_n = G_n$ and $\eta = b$ gives the existence of a constant M such that $\Omega_{1n} = \{r^b(G_n) \leqq Mr^b(G)$ on $\Delta_{n2}\}$ has probability larger than $1 - \varepsilon$ uniformly for $n = 1, 2, \cdots$ and all continuous H. Also

$$\chi(\Omega_{1n})B_{11n} \leqq DMn^{-\frac{1}{2}}|J_n(1)|r^b(G(Y_\nu)) ,$$

where ν is defined by (4.3). Set $\gamma_n = n^{-\frac{1}{2}}|J_n(1)|$ and note that by (1.2) and Assumption 2.2 we have $\gamma_n \leqq D_1 n^{-\delta}$ for some constant $D_1 \geqq D$. Let $\Omega_{2n} = \{\gamma_n \leqq G(Y_\nu) \leqq 1 - \gamma_n\}$. Then

$$\chi(\bigcap_{i=1}^2 \Omega_{in})B_{11n} \leqq DM\gamma_n^{1-b}(1 - \gamma_n)^{-b} = O(n^{-\delta(1-b)}) \to 0$$

as $n \to \infty$. Applying Lemma 4.6 we see that $P(\bigcap_{i=1}^2 \Omega_{in}) > 1 - 2\varepsilon$ for n large enough, uniformly for H in $\mathscr{H}_{C\delta}$. A symmetric argument can be given for B_{13n}.

For the rv B_{12n} use Lemma 4.5 to see that the set on which this rv may assume a nonzero value has probability converging to zero as $n \to \infty$, uniformly for H in $\mathscr{H}_{C\delta}$. □

COROLLARY 5.2. *As $n \to \infty$, $B_{2n} \to_p 0$ uniformly for H in $\mathscr{H}(\mathscr{H}_{C\delta})$.*

PROOF. The rv B_{2n} is bounded by $\sum_{i=1}^2 B_{2in}$ where

$$B_{21n} = D^2 n^{-\frac{1}{2}} r^a(F(X_\nu))r^b(G(Y_\nu)) ,$$
$$B_{22n} = D^2 n^{-\frac{1}{2}} r^a(F(X_\lambda))r^b(G(Y_\lambda)) ,$$

with ν and λ defined by (4.3).

Under the assumptions of Theorem 2.1 consider $\Omega_{1n} = \{\alpha_n \leqq F(X_\nu) \leqq 1 - \alpha_n\} \cap \{\alpha_n \leqq G(Y_\nu) \leqq 1 - \alpha_n\}$, with $\alpha_n = n^{a+b-\frac{3}{2}}$. Note that $n\alpha_n \to 0$. Then

$$\begin{aligned}\chi(\Omega_{1n})B_{21n} &\leqq D^2 n^{-\frac{1}{2}}\alpha_n^{-a-b}(1 - \alpha_n)^{-a-b} \\ &= D^2(n\alpha_n)^{1-a-b}(1 - \alpha_n)^{-a-b} \to 0\end{aligned}$$

as $n \to \infty$. Lemma 4.4 gives that $P(\Omega_{1n}) \to 1$ as $n \to \infty$, uniformly for H in $\mathscr{H}$. The same argument applies for the rv B_{22n}.

Under the assumptions of Theorem 2.2 consider

$$\Omega_{2n} = \{\beta_n \leqq F(X_\nu) \leqq 1 - \beta_n\} \cap \{\gamma_n \leqq G(Y_\nu) \leqq 1 - \gamma_n\} ,$$

with $\beta_n = (n \log n)^{-1}$ and $\gamma_n = n^{-\delta}$. Then by (2.3)

$$\chi(\Omega_{2n})B_{21n} \leqq D^2 n^{-\frac{1}{2}}\beta_n^{-a}\gamma_n^{-b}(1 - \beta_n)^{-a}(1 - \gamma_n)^{-b} \to 0$$

as $n \to \infty$. By Lemmas 4.4 and 4.6 we see that $P(\Omega_{2n}) \to 1$ as $n \to \infty$, uniformly for H in $\mathscr{H}_{C\delta}$. The rv B_{22n} can be treated in the same way. □

COROLLARY 5.3. *For fixed γ, $B_{\gamma 3n} \to_p 0$ and $C_{\gamma 1n} \to_p 0$ as $n \to \infty$, uniformly for H in $\mathscr{H}$.*

PROOF. $P(\Omega_{\gamma n}^c) \to 0$ uniformly for H in $\mathscr{H}$ by the Glivenko-Cantelli theorem and because the distribution of $\sup|F_n - F|$ does not depend on H in $\mathscr{H}$. □

COROLLARY 5.4. *For fixed γ, $B_{\gamma 5n} \to_p 0$ and $C_{\gamma 4n} \to_p 0$ as $n \to \infty$, uniformly for H in $\mathscr{H}$.*

PROOF. According to Lemma 4.2 (iii) with $\tau = \frac{1}{2}$, for given $\varepsilon > 0$ there exists

a constant M such that $\Omega_n = \{\sup |U_n(F)| \leqq M\}$ has probability larger than $1 - \varepsilon$ for all n and H in $\mathscr{H}$. Also

$$\chi(\Omega_n)|B_{\gamma 5n}| \leqq M \sup_{\Delta_{n1} \cap S_{\gamma 1}} |J'(\Phi_n) - J'(F)| \sup_{S_{\gamma 2}} |K(G)| .$$

The function $K(G)$ is bounded on $S_{\gamma 2}$ and the bound does not depend on H in $\mathscr{H}$. The function J' is uniformly continuous on $[\gamma/2, s_1 - \gamma/2] \cup [s_1 + \gamma/2, 1 - \gamma/2]$. Since $|\Phi_n - F| \leqq |F_n - F|$ where Φ_n is defined, the Glivenko-Cantelli theorem yields $\sup_{\Delta_{n1} \cap S_{\gamma 1}} |J'(\Phi_n) - J'(F)| \to_p 0$ uniformly for H in $\mathscr{H}$. A similar argument may be used for $C_{\gamma 4n}$. □

COROLLARY 5.5. *For fixed* γ, $B_{\gamma 6n} \to_p 0$ *as* $n \to \infty$, *uniformly for* H *in* $\mathscr{H}$.

PROOF. For arbitrary k we have (see (4.2)) $|B_{\gamma 6n}| \leqq \sum_{i=1}^{3} B_{\gamma 6ikn}$, where

$$B_{\gamma 61kn} = \int\int_{\Delta_n \cap S_\gamma} |U_n(F)J'(F)K(G) - U_n(I_k(F))J'(I_k(F))K(I_k(G))| \, dH_n ,$$
$$B_{\gamma 62kn} = |\int\int_{\Delta_n \cap S_\gamma} U_n(I_k(F))J'(I_k(F))K(I_k(G)) \, d(H_n - H)| ,$$
$$B_{\gamma 63kn} = \int\int_{\Delta_n \cap S_\gamma} |U_n(F)J'(F)K(G) - U_n(I_k(F))J'(I_k(F))K(I_k(G))| \, dH .$$

Let us first consider $B_{\gamma 61kn}$ and $B_{\gamma 63kn}$, which are both bounded by the supremum of the integrand over the set S_γ. Let an arbitrary $\varepsilon > 0$ be given. Application of Lemma 4.3 gives the existence of constants $\eta_{kn} \to 0$ as $k, n \to \infty$, such that $\Omega_{kn} = \{\sup |U_n(F) - U_n(I_k(F))| \leqq \eta_{kn}\}$ has probability larger than $1 - \varepsilon$ for all k, n and all H in $\mathscr{H}$. Note that on $([\gamma, s_1 - \gamma] \cup [s_1 + \gamma, 1 - \gamma]) \times ([\gamma, t_1 - \gamma] \cup [t_1 + \gamma, 1 - \gamma])$ the function $J'(s)K(t)$ is bounded, say by a constant M_γ, and uniformly continuous. By Lemma 4.2 (iii) with $\tau = \frac{1}{2}$, there exists a constant M such that $\Omega_n = \{\sup |U_n(F)| \leqq M\}$ has probability larger than $1 - \varepsilon$. Let us finally write $\zeta_{k\gamma} = \max_{S_\gamma} |J'(F)K(G) - J'(I_k(F))K(I_k(G))|$, which tends to zero as $k \to \infty$, uniformly for H in $\mathscr{H}$. Hence for $i = 1, 3$

$$\chi(\Omega_{kn} \cap \Omega_n)B_{\gamma 6ikn} \leqq \eta_{kn} M_\gamma + M\zeta_{k\gamma} \to 0$$

as $k, n \to \infty$ for fixed γ. Because $P(\Omega_{kn} \cap \Omega_n) > 1 - 2\varepsilon$ uniformly for H in $\mathscr{H}$ we may conclude that $B_{\gamma 61kn} \to_p 0$ and $B_{\gamma 63kn} \to_p 0$ uniformly for H in $\mathscr{H}$, as $k, n \to \infty$.

Let us next consider $B_{\gamma 62kn}$ for a fixed value k. For each ω in Ω_n the integrand in the expression for this rv is a simple step function assuming a value $a_{ijkn}(\omega)$ on the rectangle

$$R_{ijkn} = (F^{-1}((i-1)/k), F^{-1}(i/k)] \times (G^{-1}((j-1)/k), G^{-1}(j/k)] \cap S_\gamma \cap \Delta_n ,$$

for $i = 1, \cdots, k$ and $j = 1, \cdots, k$. Because $|a_{ijkn}| \leqq M(M_\gamma + \zeta_{k\gamma})$ on Ω_n, we have

$$\chi(\Omega_n)B_{\gamma 62kn} = |\sum_{i=1}^{k} \sum_{j=1}^{k} a_{ijkn} \int\int_{R_{ijkn}} d(H_n - H)|$$
$$\leqq 4k^2 M(M_\gamma + \zeta_{k\gamma}) \sup |H_n - H| \to_p 0$$

as $n \to \infty$, uniformly for H in $\mathscr{H}$. Here Theorem 1-*m* of Kiefer [8] is used. The conclusion of the corollary follows by straightforward combination of these results. □

COROLLARY 5.6. *As* $\gamma \downarrow 0$ *and* $n \to \infty$, $B_{\gamma 7n} \to_p 0$ *and* $C_{\gamma 2n} \to_p 0$, *uniformly for* H *in* $\mathscr{H}(\mathscr{H}_{C\delta})$.

PROOF. Let $\varepsilon > 0$ be given and let us first consider $B_{\gamma 7n}$. By Lemma 4.2 (iii), taking $\tau = \delta/4$, there exists a constant M_1 such that $\Omega_{1n} = \{|U_n(F)| \leqq M_1 r^{-\frac{1}{2}+\delta/4}(F)\}$ has probability larger than $1 - \varepsilon$ for all n and H in $\mathscr{H}$. From Assumption 2.1 (Assumption 2.2) it may be seen that

$$\chi(\Omega_{1n})|B_{\gamma 7n}| \leqq D^2 M_1 \iint_{\Delta_n{}^c \cup S_\gamma{}^c} r^{a+\frac{1}{2}+\delta/4}(F) r^b(G)\, dH\,. \tag{5.1}$$

Next consider $C_{\gamma 2n}$. By Assumption 2.1 (Assumption 2.2) we have $|K(G_n)| \leqq Dr^b(G_n)$ on Δ_{n2} and application of Lemma 4.2 (ii) with $\tilde{\Psi}_n = G_n$ and $\eta = b$ gives the existence of a constant M_2 such that $\Omega_{2n} = \{r^b(G_n) \leqq M_2 r^b(G)$ on $\Delta_{n2}\}$ has probability larger than $1 - \varepsilon$ for all n and H in $\mathscr{H}$. Take an arbitrary ω in Ω and let us first consider those values of x in Δ_{n1} for which the open random interval between the points $F(x)$ and $F_n(x)$ does not contain s_1. Then by continuity of J on the closed and differentiability on the open interval, the mean value theorem can be applied; it follows from Assumption 2.1 (Assumption 2.2) that

$$n^{\frac{1}{2}}|J(F_n) - J(F)| = |U_n(F)J'(\Phi_{0n})| \leqq D|U_n(F)|r^{a+1}(\Phi_{0n})\,.$$

For those values of x in Δ_{n1} for which the open random interval between the points $F(x)$ and $F_n(x)$ does contain s_1, the mean value theorem can be applied stepwise, since J is continuous on the closed interval and differentiable on the two open intervals between $F(x)$, $F_n(x)$ and s_1. We thus get the estimate

$$n^{\frac{1}{2}}|J(F_n) - J(F)| \leqq |U_n(F)| \sum_{i=1}^{2} |J'(\Phi_{in})| \leqq D|U_n(F)| \sum_{i=1}^{2} r^{a+1}(\Phi_{in})$$

by Assumption 2.1 (Assumption 2.2). Where defined on Δ_{n1}, both Φ_{0n} and Φ_{1n}, Φ_{2n} lie between F and F_n. By Lemma 4.2 (i), taking $\zeta = a + 1$, there exists a constant M_3 such that $\Omega_{3n} = \{\max_{i=0,1,2} r^{a+1}(\Phi_{in}) \leqq M_3 r^{a+1}(F)$ where defined on $\Delta_{n1}\}$ has probability larger than $1 - \varepsilon$ for all n and H in $\mathscr{H}$. Combining these results we have

$$E(\chi(\textstyle\bigcap_{i=1}^{3} \Omega_{in})|C_{\gamma 2n}|) \leqq 2D^2 M_1 M_2 M_3 \iint_{S_\gamma{}^c} r^{a+\frac{1}{2}+\delta/4}(F) r^b(G)\, dH\,. \tag{5.2}$$

From (5.1) and (5.2) it is clear that the corollary is proved if we show that the integral on the right in (5.1) converges to zero as $\gamma \downarrow 0$ and $n \to \infty$, uniformly for H in $\mathscr{H}(\mathscr{H}_{C\delta})$. For this purpose we start with the integral

$$\iint_{S_\gamma{}^c} r^{a+\frac{1}{2}+\delta/4}(F) r^b(G)\, dH\,, \tag{5.3}$$

and note that $S_\gamma{}^c \subset (S_{\gamma 1}^c \times (-\infty, \infty)) \cup ((-\infty, \infty) \times S_{\gamma 2}^c)$. Under Assumption 2.1, by application of (1.5) with $p = p_1$ and $q = q_1$ as in (3.5), we find that (5.3) is bounded uniformly for H in $\mathscr{H}$ by

$$\begin{aligned}[\textstyle\int_{(0,\gamma)\cup(s_1-\gamma,s_1+\gamma)\cup(1-\gamma,1)} r^{(a+\frac{1}{2}+\delta/4)p_1}\, dI]^{1/p_1} [\textstyle\int r^{bq_1}\, dI]^{1/q_1} \\ + [\textstyle\int r^{(a+\frac{1}{2}+\delta/4)p_1}\, dI]^{1/p_1} [\textstyle\int_{(0,\gamma)\cup(t_1-\gamma,t_1+\gamma)\cup(1-\gamma,1)} r^{bq_1}\, dI]^{1/q_1}\,.\end{aligned} \tag{5.4}$$

Since by (3.5) both exponents of the function r are smaller than 1, the dominated

convergence theorem implies convergence of (5.4) to zero as $\gamma \downarrow 0$. Under Assumption 2.2 and for H in $\mathscr{H}_{C\delta}$, by an application of (1.6) we see that (5.3) is bounded uniformly for H in $\mathscr{H}_{C\delta}$ by

$$(5.5)\qquad C[\int_{(0,\gamma)\cup(s_1-\gamma,s_1+\gamma)\cup(1-\gamma,1)} r^{1-\delta/4}\,dI][\int r^{\frac{1}{2}-\delta/2}\,dI] + C[\int r^{1-\delta/4}\,dI][\int_{(0,\gamma)\cup(t_1-\gamma,t_1+\gamma)\cup(1-\gamma,1)} r^{\frac{1}{2}-\delta/2}\,dI],$$

which by the dominated convergence theorem converges to zero as $\gamma \downarrow 0$. Hence under the assumptions of Theorem 2.1 (Theorem 2.2) a value $\tilde{\gamma}$ of γ can be chosen such that (5.3) is smaller than ε for all H in $\mathscr{H}(\mathscr{H}_{C\delta})$ provided $\gamma \leqq \tilde{\gamma}$. For this $\tilde{\gamma}$ there exists an index $\tilde{n} = \tilde{n}_{\tilde{\gamma}}$ such that $P(\{\Delta_n \supset S_{\tilde{\gamma}}\}) > 1 - \varepsilon$ uniformly for H in $\mathscr{H}$, provided $n \geqq \tilde{n}$. It follows that under the assumptions of Theorem 2.1 (Theorem 2.2) the integral on the right in (5.1) is smaller than ε with probability larger than $1 - \varepsilon$ uniformly for H in $\mathscr{H}(\mathscr{H}_{C\delta})$ for all $\gamma \leqq \tilde{\gamma}$ and all $n \geqq \tilde{n}$. □

In order to show how the results of these corollaries can be combined to complete the proof of Theorems 2.1 and 2.2, let an arbitrary $\varepsilon > 0$ be given. First use Corollary 5.6 to choose a fixed γ and an index n_1 to ensure $P(\{|B_{\gamma 7n}|, |C_{\gamma 2n}| \leqq \varepsilon\}) > 1 - \varepsilon$ for all $n > n_1$. Next use Assumption 2.3 (a) and Corollaries 5.1–5.5 to choose for the above fixed γ an index $n_2 = n_{2\gamma} > n_1$ such that $P(\{|B_{in}|, |B_{\gamma jn}|, |C_{\gamma kn}| \leqq \varepsilon$ for $i = 0, 1, 2$; $j = 3, 5, 6$; $k = 1, 4\}) > 1 - \varepsilon$ for $n > n_2$. This implies that the probability that the sum of all these second order terms does not exceed 10ε is larger than $1 - 2\varepsilon$ uniformly for H in $\mathscr{H}(\mathscr{H}_{C\delta})$, as $n > n_2$.

6. Replacing Assumption 2.3(a) **by Assumption** 2.3(b). We shall now suppose that Assumption 2.3 (b) holds. Again the theorems will be considered only in the case where J and K fail to have a derivative at one point, s_1 and t_1 respectively. The proof is based on an analogue of (3.4). We shall need both the empirical processes and the processes $U_n^*(F) = n^{\frac{1}{2}}(F_n^* - F)$, $V_n^*(G) = n^{\frac{1}{2}}(G_n^* - G)$. Instead of the set $\Omega_{\gamma n}$ we shall use $\Omega_{\gamma n}^* = \{\omega : \sup|F_n^* - F| < \gamma/2, \sup|G_n^* - G| < \gamma/2\}$. The role of Δ_n will be taken over by its closure $\bar{\Delta}_n = \bar{\Delta}_{n1} \times \bar{\Delta}_{n2} = [X_{1n}, X_{nn}] \times [Y_{1n}, Y_{nn}]$. Because integration over $\bar{\Delta}_n$ with respect to dH_n is the same as integration over the entire plane, we now have the simpler decomposition

$$(6.1)\qquad n^{\frac{1}{2}}(T_n - \mu) = \sum_{i=1}^{3} A_{in} + B_{0n}^* + \sum_{i=1}^{5} B_{\gamma in}^* + B_{6n}^* + C_n^*,$$

with probability 1. Here B_{0n}^* is defined in (2.5), the A-terms are as given in Section 3 and

$$\begin{aligned}
B_{\gamma 1n}^* &= \chi(\Omega_{\gamma n}^{*c})\{n^{\frac{1}{2}} \iint [J(F_n^*) - J(F)]K(G)\,dH_n - A_{2n}\},\\
B_{\gamma 2n}^* &= \chi(\Omega_{\gamma n}^*)n^{\frac{1}{2}} \iint_{S_\gamma{}^c} [J(F_n^*) - J(F)]K(G)\,dH_n,\\
B_{\gamma 3n}^* &= \chi(\Omega_{\gamma n}^*) \iint_{S_\gamma} U_n^*(F)[J'(\Phi_n^*) - J'(F)]K(G)\,dH_n,\\
B_{\gamma 4n}^* &= \chi(\Omega_{\gamma n}^*) \iint_{\bar{\Delta}_n \cap S_\gamma} U_n^*(F)J'(F)K(G)\,d(H_n - H),\\
B_{\gamma 5n}^* &= \chi(\Omega_{\gamma n}^*)\{\iint_{\bar{\Delta}_n \cap S_\gamma} U_n^*(F)J'(F)K(G)\,dH - A_{2n}\},\\
B_{6n}^* &= n^{\frac{1}{2}} \iint J(F)[K(G_n^*) - K(G)]\,dH_n - A_{3n},\\
C_n^* &= n^{\frac{1}{2}} \iint [J(F_n^*) - J(F)][K(G_n^*) - K(G)]\,dH_n.
\end{aligned}$$

The function Φ_n^* arises from application of the mean value theorem and lies strictly between F and F_n^* where defined. The analogues of B_{1n} and B_{2n} are missing in this decomposition; this essentially simplifies the proof of the theorems. However, if one tries to prove the validity of Assumption 2.3 (b) when Assumption 2.3 (a) is given to hold, problems similar to those connected with B_{1n} and B_{2n} recur.

Only the second order terms differ from those in (3.4). For their asymptotic negligibility we need the following modifications of Lemma 4.2.

LEMMA 6.1. *For each ω let $\tilde{\Phi}_n^* = \tilde{\Phi}_{n\omega}^*$ and $\tilde{\Psi}_n^* = \tilde{\Psi}_{n\omega}^*$ be functions on $\bar{\Delta}_{n1} = \bar{\Delta}_{n1\omega}$ and $\bar{\Delta}_{n2} = \bar{\Delta}_{n2\omega}$ respectively, satisfying $\min(F, F_n^*) \leqq \tilde{\Phi}_n^* \leqq \max(F, F_n^*)$ and $\min(G, G_n^*) \leqq \tilde{\Psi}_n^* \leqq \max(G, G_n^*)$ where defined. Then, uniformly for $n = 1, 2, \cdots$ and $H \in \mathscr{H}$:*

(i) $\sup_{\bar{\Delta}_{n1}} r^{\zeta}(\tilde{\Phi}_n^*) r^{-\zeta}(F) = O_p(1)$ *for each* $\zeta \geqq 0$;
(ii) $\sup_{\bar{\Delta}_{n2}} r^{\eta}(\tilde{\Psi}_n^*) r^{-\eta}(G) = O_p(1)$ *for each* $\eta \geqq 0$.

PROOF. It suffices to prove (i). Let us first show that for each $\varepsilon > 0$ there exists a $\beta = \beta_\varepsilon$ in (0, 1) such that $P(\{\beta F \leqq F_n^* \leqq 1 - \beta(1 - F) \text{ on } \bar{\Delta}_{n1}\}) > 1 - \varepsilon$, for all n and uniformly in all continuous F. By (4.1) and because $\frac{1}{2} \leqq n/(n+1) \leqq 1$, we only have to prove that $P(\{n/(n+1) \leqq 1 - \beta[1 - F(X_{nn})]\}) > 1 - \varepsilon$ for β small enough. Because the $F(X_i)$ are independent uniform (0, 1) rvs, this probability equals $1 - \{1 - 1/[\beta(n+1)]\}^n > 1 - \varepsilon$ for all n and uniformly in all continuous F, provided $\beta = \beta_\varepsilon$ is chosen sufficiently small. The proof can be concluded in the same way as that of Lemma 4.2. □

LEMMA 6.2. *Uniformly in all continuous F we have:*

(i) $\sup_{\bar{\Delta}_{n1}} |U_n^*(F) - U_n(F)| r^{\frac{1}{2}-\rho}(F) \to_p 0$ *as* $n \to \infty$, *for each* $\rho > 0$;
(ii) $\sup_{\bar{\Delta}_{n1}} |U_n^*(F)| r^{\frac{1}{2}-\tau}(F) = O_p(1)$ *uniformly for* $n = 1, 2, \cdots$, *for each* $\tau > 0$.

PROOF. (i) Note that $|U_n^*(F) - U_n(F)| r^{\frac{1}{2}-\rho}(F) < n^{-\frac{1}{2}} r^{\frac{1}{2}-\rho}(F)$ and that for any fixed $\beta \in (0, 1)$ we have $r^{\frac{1}{2}-\rho}(\beta/n) = r^{\frac{1}{2}-\rho}(1 - \beta/n) = O(n^{\frac{1}{2}-\rho})$. Because the $F(X_i)$ are independent uniform rvs, given an arbitrary $\varepsilon > 0$ we can choose a $\beta = \beta_\varepsilon$ in (0, 1) such that $P(\{\beta/n \leqq F(X_{1n}) \leqq F(X_{nn}) \leqq 1 - \beta/n\}) > 1 - \varepsilon$ for all n and uniformly for all continuous F. Part (i) follows from a combination of these results. (ii) follows from (i) and Lemma 4.2 (iii). □

The proof that the sum of the B^*- and C^*-terms converges in probability to zero can be given by a method quite similar to that of Section 5, by using Lemmas 6.1, 6.2 instead of Lemma 4.2.

REFERENCES

[1] BHUCHONGKUL, S. (1964). A class of nonparametric tests for independence in bivariate populations. *Ann. Math. Statist.* **35** 138-149.
[2] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
[3] CHERNOFF, H. and I. R. SAVAGE (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29** 972-994.

[4] FELLER, W. (1957). *An Introduction to Probability Theory and its Applications* **1**. Wiley, New York.
[5] GUMBEL, E. J. (1958). Multivariate distributions with given margins and analytical examples. *Bull. Inst. Internat. Statist.* **37-3** 363-373.
[6] HÁJEK, J. and Z. ŠIDÁK (1967). *Theory of Rank Tests*. Academic Press, New York.
[7] JOGDEO, K. (1968). Asymptotic normality in nonparametric methods. *Ann. Math. Statist.* **39** 905-922.
[8] KIEFER, J. (1961). On large deviations of the empiric df of vector chance variables and a law of the iterated logarithm. *Pacific J. Math.* **11** 649-660.
[9] LEHMANN, E. L. (1953). The power of rank tests. *Ann. Math. Statist.* **24** 23-43.
[10] PYKE, R. and G. R. SHORACK (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorems. *Ann. Math. Statist.* **39** 755-771.
[11] SHORACK, G. R. (1972). Functions of order statistics. *Ann. Math. Statist.* **43** 412-427.
[12] SKOROKHOD, A. V. (1956). Limit theorems for stochastic processes. *Theor. Probability Appl.* **1** 261-290.