The Annals of Statistics
1974, Vol. 2, No. 5, 892–910

ASYMPTOTIC NORMALITY OF NONPARAMETRIC TESTS FOR INDEPENDENCE[1]

BY F. H. RUYMGAART

Mathematisch Centrum, Amsterdam

Asymptotic normality of linear rank statistics for testing the hypothesis of independence is established both under fixed alternatives (or the null hypothesis) and under converging alternatives. The results of Ruymgaart, Shorack and van Zwet [14] are used to obtain a further weakening of the smoothness conditions on the score functions. In the present case the score functions are allowed to have a finite number of discontinuities of the first kind. The results of the present paper and of the paper [14] will be summarized in the author's thesis [13].

1. Introduction. For each n, let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be a set of independent identically distributed (i.i.d.) random vectors, with common continuous bivariate distribution function (df) $H(x, y)$ having marginal df's $F(x)$ and $G(y)$. The bivariate empirical df based on this sample is denoted by H_n. With respect to the n random variables (rv's) $X_i(Y_i)$ 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 $(\Omega, \mathscr{A}, P)$.

The rank statistics most commonly used to test the independence hypothesis $H = F \cdot G$, 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 [9]). A suitably standardized version of T_n will be (see also Bhuchongkul [2])

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

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

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

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.

This paper is a continuation of Ruymgaart, Shorack and van Zwet [14].

Received April 1972; revised July 1973.

[1] Report SW 15/72 of the Department of Mathematical Statistics, Mathematisch Centrum, Amsterdam.

AMS 1970 *subject classification.* 62G10.

Key words and phrases. Rank statistics for independence, smoothness and growth of (limiting) score functions, fixed alternatives, asymptotic normality.

Theorem 2.1 states the asymptotic normality of (1.1) both under the hypothesis and under fixed alternatives, and it covers [14], Theorems 2.1 and 2.2 under Assumption 2.3 (b) as special cases. The generalization lies in a further weakening of the smoothness conditions to be imposed on the score functions J and K on (0, 1). In the present case these functions are allowed to have a finite number of discontinuities of the first kind. This weakening of the smoothness conditions entails, as could be expected (see e.g. Dupač and Hájek [4]) a local differentiability condition on the underlying continuous df H. By a decomposition of the score functions J, K in their continuous parts J_c, K_c and their discontinuous parts J_d, K_d the method of [14] can be used to take care of the continuous part. This method is based on an application of the mean value theorem (Bhuchongkul [2] uses a Taylor-series expansion up to second order derivatives) and Lemma 2.2 of Pyke and Shorack [12]. For the discontinuous part we mainly need Lemma 4.4, which is similar to a bivariate form of Bahadur [1], Lemma 1 or Sen [15], Theorem 2.1. The results of Bahadur and Sen are for univariate df's only but stronger in the sense that they provide "almost sure" statements while our result gives a statement "in probability." On the other hand, Lemma 4.4 does not require any condition on the underlying bivariate df H, which need not even be continuous, and the conclusion of the lemma is uniform in all sequences of intervals in the plane. Similarly, Sen [16] utilizes his above result ([15], Theorem 2.1) for multivariate rank order statistics in the location problem, when purely discontinuous score functions with a finite number of jumps are used. More recently, among others, Ghosh [6] studied the above mentioned problem for univariate df's, initiated by Bahadur.

In Theorem 2.2 the case of converging alternatives is considered: the bivariate df H, from which the sample has been drawn, may now depend on the sample size n. Hence we write more explicitly $H_{(n)}$ instead of H, and $F_{(n)}$, $G_{(n)}$ for the marginals instead of F, G respectively. Under certain convergence conditions on the sequence of df's $H_{(1)}, H_{(2)}, \cdots$ asymptotic normality of

$$n^{\frac{1}{2}}(T_n - \mu_n) = n^{\frac{1}{2}}[\textstyle\iint J_n(F_n)K_n(G_n)\, dH_n - \mu_n] \tag{1.4}$$

is proved. Here

$$\mu_n = \mu(H_{(n)}) = \textstyle\iint J(F_{(n)})K(G_{(n)})\, dH_{(n)}\,. \tag{1.5}$$

2. Statement of the theorems. To formulate the assumptions needed for the theorems, we shall first introduce some notation. Attention will be restricted to the class $\mathscr{H}$ of all continuous bivariate df's H. Let further

$$\Lambda_n = \Lambda_{n\omega} = \Lambda_{n1\omega} \times \Lambda_{n2\omega}\,, \quad \text{with} \quad \Lambda_{n1} = [X_{1n}, X_{nn}]\,, \quad \Lambda_{n2} = [Y_{1n}, Y_{nn}]\,, \tag{2.1}$$

$$F_n^* = [n/(n+1)]F_n\,, \qquad G_n^* = [n/(n+1)]G_n\,. \tag{2.2}$$

For any pair of real numbers u, v the symbol $\delta_u(v)$ stands for

$$\delta_u(v) = 0 \quad \text{if} \quad v < u\,, \qquad \delta_u(v) = 1 \quad \text{if} \quad v \geqq u\,. \tag{2.3}$$

The assumption on the limit behaviour of J_n and K_n concerns

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

For ease of reference some definitions of Pyke and Shorack [12] and Shorack [17] will be copied.

DEFINITION 2.1. (Pyke and Shorack). Let $\mathscr{Q}$ denote the class of all functions q defined and continuous on $[0, 1]$, which are positive on $(0, 1)$, symmetric about $\frac{1}{2}$, increasing on $(0, \frac{1}{2}]$ and for which $\int_0^1 [q(u)]^{-2}\, du < \infty$. The members of $\mathscr{Q}$ will be called q-functions.

DEFINITION 2.2. (Shorack). A function r, defined and positive on $(0, 1)$, which is symmetric about $\frac{1}{2}$, will be called "u-shaped" if it is decreasing on $(0, \frac{1}{2}]$. If $0 < \beta < 1$ we introduce the notation r_β for the function satisfying

$$\begin{aligned} r_\beta(s) &= r(\beta s) \qquad \text{for} \quad 0 < s \leqq \tfrac{1}{2} \,, \\ r_\beta(s) &= r(1 - \beta(1 - s)) \qquad \text{for} \quad \tfrac{1}{2} < s < 1 \,. \end{aligned} \tag{2.5}$$

If for all positive β in a neighborhood of zero there exists a constant M_β such that $r_\beta \leqq M_\beta r$ on $(0, 1)$, then r will be called a "reproducing u-shaped function." The class of all reproducing u-shaped functions will be denoted by $\mathscr{R}$.

REMARK. Throughout this section the functions r_1, $\tilde r_1$, r_2, $\tilde r_2$ are members of $\mathscr{R}$. These functions and the points $0 < s_1 < \cdots < s_\lambda < 1$ and $0 < t_1 < \cdots < t_\nu < 1$ are supposed to be fixed.

ASSUMPTION 2.1. Let be given the subclass $\mathscr{H}' \subset \mathscr{H}$. As $n \to \infty$, $B_{0n}^* \to_p 0$ uniformly for $H \in \mathscr{H}'$.

ASSUMPTION 2.2. The functions J and K are defined on $(0, 1)$ and can be written as $J = J_c + J_d$ and $K = K_c + K_d$. Here $J_d = \sum_{i=1}^{\lambda} \alpha_i \delta_{s_i}$ and $K_d = \sum_{j=1}^{\nu} \beta_j \delta_{t_j}$ for arbitrary constants α_i, β_j and with δ_{s_i}, δ_{t_j} as defined in (2.3). Further J_c and K_c are continuous on $(0, 1)$ and have continuous derivatives $J_c' = J'$ and $K_c' = K'$ on the open intervals between the points $0, s_1, \cdots, s_\lambda, 1$ and $0, t_1, \cdots, t_\nu, 1$ respectively. As to the orders of magnitude of the above functions, where defined on $(0, 1)$ we have

$$|J| \leqq r_1 \,, \qquad |J'| \leqq \tilde r_1 \,, \qquad |K| \leqq r_2 \,, \qquad |K'| \leqq \tilde r_2 \,.$$

ASSUMPTION 2.3. Let be given the subclass $\mathscr{H}' \subset \mathscr{H}$. For some constant $\varepsilon \geqq 0$ and functions $q_1, q_2 \in \mathscr{Q}$ we have

$$\begin{gathered} \sup_{H \in \mathscr{H}'} \iint [r_1(F) r_2(G)]^{2+\varepsilon}\, dH < \infty \,, \\ \int_0^1 [q_1(s)]^{-2-\varepsilon}\, ds < \infty \,, \qquad \int_0^1 [q_2(t)]^{-2-\varepsilon}\, dt < \infty \,, \\ \sup_{H \in \mathscr{H}'} \iint [q_1(F) \tilde r_1(F) r_2(G)]^{1+\varepsilon}\, dH < \infty \,, \,. \\ \sup_{H \in \mathscr{H}'} \iint [q_2(G) r_1(F) \tilde r_2(G)]^{1+\varepsilon}\, dH < \infty \,. \end{gathered}$$

ASSUMPTION 2.4. Either (a) $J_d = K_d = 0$ on $(0, 1)$ in Assumption 2.2, or (b) the following holds for the subclass $\mathscr{H}' \subset \mathscr{H}$. There is an open set O_1 containing the points $s_1, \cdots, s_\lambda$ and an open set O_2 containing the points $t_1, \cdots, t_\nu$ such that for each $H \in \mathscr{H}'$ the density $h(s, t) = \partial^2 H(F^{-1}(s), G^{-1}(t))/\partial s\, \partial t$ exists and is continuous on $O_1 \times (0, 1) \cup (0, 1) \times O_2$. Moreover the subclass $\mathscr{H}'$ satisfies the equicontinuity conditions

$$\sup_{H \in \mathscr{H}'} |h(s, t) - h(s_i, t)| \to 0 \quad \text{as} \quad s \to s_i \quad \text{for all} \quad t \in (0, 1),\ i = 1, \cdots, \lambda\,,$$

$$\sup_{H \in \mathscr{H}'} |h(s, t) - h(s, t_j)| \to 0 \quad \text{as} \quad t \to t_j \quad \text{for all} \quad s \in (0, 1),\ j = 1, \cdots, \nu\,,$$

and has the property that there exist functions f and g on $(0, 1)$ such that

$$\sup_{H \in \mathscr{H}'} h(s, t) \leqq f(s) \quad \text{for all} \quad (s, t) \in (0, 1) \times O_2\,, \quad \text{with} \quad \textstyle\int_0^1 r_1(s) f(s)\, ds < \infty\,,$$

$$\sup_{H \in \mathscr{H}'} h(s, t) \leqq g(t) \quad \text{for all} \quad (s, t) \in O_1 \times (0, 1)\,, \quad \text{with} \quad \textstyle\int_0^1 r_2(t) g(t)\, dt < \infty\,.$$

We also need the following modification of Assumption 2.4.

ASSUMPTION 2.5. Let $H_{(n)} \in \mathscr{H}$ for $n = 0, 1, 2, \cdots$. As $n \to \infty$, $H_{(n)}(x, y) \to H_{(0)}(x, y)$ for all x, y. Moreover, either (a) $J_d = K_d = 0$ on $(0, 1)$ in Assumption 2.2, or (b) Assumption 2.4 (b) is satisfied with $\mathscr{H}' = \{H_{(0)}, H_{(1)}, H_{(2)}, \cdots\}$. In the latter case we further have $h_n(s_i, t) \to h_0(s_i, t)$ for all $t \in (0, 1)$, $i = 1, \cdots, \lambda$ and $h_n(s, t_j) \to h_0(s, t_j)$ for all $s \in (0, 1)$, $j = 1, \cdots, \nu$ as $n \to \infty$. Here h_n is the density corresponding to $H_{(n)}$, $n = 0, 1, 2, \cdots$.

THEOREM 2.1. (Hypothesis and fixed alternatives). *Suppose $(X_1, Y_1), \cdots, (X_n, Y_n)$ is a sample from a fixed* df *$H \in \mathscr{H}$ not depending on the sample size. If Assumptions* 2.1–2.4 *are satisfied with $\mathscr{H}' = \{H\}$ and $\varepsilon = 0$, then $n^{\frac{1}{2}}(T_n - \mu) \to_d$* $\mathrm{N}(0, \sigma^2)$ *as $n \to \infty$, with finite $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ given by* (1.3) *and* (3.5) *respectively.*

Suppose Assumptions 2.1–2.4 *are satisfied for some fixed subclass $\mathscr{H}' \subset \mathscr{H}$ and $\varepsilon > 0$. If $\sigma^2 = \sigma^2(H)$ is bounded away from zero on $\mathscr{H}'$, then the above convergence in distribution is uniform for $H \in \mathscr{H}'$.*

THEOREM 2.2. (Converging alternatives). *Suppose $(X_1, Y_1), \cdots, (X_n, Y_n)$ is a sample from a* df *$H_{(n)} \in \mathscr{H}$ that may depend on the sample size n. Let for some $H_{(0)} \in \mathscr{H}$ Assumptions* 2.1–2.3 *and* 2.5 *be satisfied with $\mathscr{H}' = \{H_{(0)}, H_{(1)}, H_{(2)}, \cdots\}$ and $\varepsilon > 0$. If in addition $\sigma_0^2 = \sigma^2(H_{(0)}) > 0$, then $n^{\frac{1}{2}}(T_n - \mu_n) \to_d \mathrm{N}(0, \sigma_0^2)$ as $n \to \infty$, with finite $\mu_n = \mu(H_{(n)})$ and σ_0^2 given by* (1.5) *and* (3.5) *respectively.*

In spite of their formidable appearance, the assumptions of the theorems are satisfied in many interesting situations. Two examples of the validity of the first theorem are provided by [14], Theorems 2.1 and 2.2. Suppose that $J_n(s) = J([n/(n+1)]s)$ and $K_n(t) = K([n/(n+1)]t)$. Thus Assumption 2.1 is trivially satisfied with $B_{0n}^* = 0$ for all n and $H \in \mathscr{H}$. Further suppose that Assumption

2.2 is satisfied with $J_d = K_d = 0$ on $(0, 1)$ (so that Assumption 2.4(a) holds) and $r_1(s) = D[s(1-s)]^{-a}$, $\tilde{r}_1(s) = D[s(1-s)]^{-a-1}$, $r_2(t) = D[t(1-t)]^{-b}$, $\tilde{r}_2(t) = D[t(1-t)]^{-b-1}$, where D is a positive constant and a and b are given real numbers. For $0 < \delta < \frac{1}{2}$, first let $a = (\frac{1}{2} - \delta)/p$ and $b = (\frac{1}{2} - \delta)/q$, where $p, q > 1$ with $p^{-1} + q^{-1} = 1$. Secondly let $a = b = \frac{1}{2} - \delta$; for this constant δ and a fixed constant C consider the subclass $\mathscr{H}_{C\delta} = \{H \in \mathscr{H}\colon dH \leqq C[F(1-F)G(1-G)]^{-\delta/2}\, dF\, dG\}$. Then for the above two choices of a, b Assumption 2.3 holds with $\mathscr{H}' = \mathscr{H}$ and $\mathscr{H}' = \mathscr{H}_{C\delta}$ respectively. In either case the assumption is satisfied for some $\varepsilon > 0$ (depending on a, b, δ) and for $q_1(s) = [s(1-s)]^{\frac{1}{2}-\delta/4}$, $q_2(t) = [t(1-t)]^{\frac{1}{2}-\delta/4}$. A third example is given by the quadrant test statistic for independence (see Hájek [8]), which is defined by the score functions $J_n(s) = \delta_{\frac{1}{2}+1/n}(s)$, $K_n(t) = \delta_{\frac{1}{2}+1/n}(t)$. Taking $J(s) = \delta_{\frac{1}{2}}(s)$, $K(t) = \delta_{\frac{1}{2}}(t)$ we see that $B_{0n}^* = O(n^{-\frac{1}{2}})$ uniformly for $H \in \mathscr{H}$, so that Assumption 2.1 is satisfied with $\mathscr{H}' = \mathscr{H}$. By the boundedness of the score functions Assumptions 2.2 and 2.3 are trivially satisfied for some $\varepsilon > 0$ and with $\mathscr{H}' = \mathscr{H}$.

However, in the latter case Assumption 2.4(b) must be fulfilled. Let us first note that Assumption 2.4(b) holds if for $\mathscr{H}'$ we take the class of all null-hypothesis df's in $\mathscr{H}$, since for such a df $H = F \cdot G$ the transformed df equals $s \cdot t$ on $(0, 1) \times (0, 1)$ with density identically equal to 1 on the unit square. By way of a further example let us verify this assumption in the case of bivariate normal df's $\Phi_\rho(x, y)$ with standard normal marginal df's $\Phi(x)$ and $\Phi(y)$ and correlation coefficient $-1 < \rho < 1$. The transformed df $\Phi_\rho(\Phi^{-1}(s), \Phi^{-1}(t))$ has a continuous density on $(0, 1) \times (0, 1)$ given by $\partial^2\Phi_\rho(\Phi^{-1}(s), \Phi^{-1}(t))/\partial s\, \partial t = (1-\rho^2)^{-\frac{1}{2}} \exp(-[(\rho\Phi^{-1}(s))^2 + (\rho\Phi^{-1}(t))^2 - 2\rho\Phi^{-1}(s)\Phi^{-1}(t)]/[2(1-\rho^2)])$. From this it follows that Assumption 2.4(b) is satisfied for any class $\mathscr{H}_\delta' = \{\Phi_\rho : -1 + \delta \leqq \rho \leqq 1 - \delta\}$ with $0 < \delta < 1$. In this case the assumption even holds with f and g constant on $(0, 1)$.

Theorem 2.2 is especially useful for the calculation of Pitman-efficiencies. Then we take $H_{(0)} = F_{(0)} \cdot G_{(0)}$, i.e. for $H_{(0)}$ we take a null-hypothesis df. If, moreover, in this case

$$n^{\frac{1}{2}}(\mu_n - \mu_0) \to e\,, \tag{2.6}$$

as $n \to \infty$ for some finite number e, Theorem 2.2 reduces to

$$n^{\frac{1}{2}}(T_n - \mu_0) \to_d \mathrm{N}(e, \sigma_0^2)\,, \tag{2.7}$$

as $n \to \infty$. Here μ_0 and σ_0^2 are the null-hypothesis mean and variance respectively.

For instance, consider the class of alternatives $H = FG[1 + \alpha(1-F)(1-G)]$ for some $-1 \leqq \alpha \leqq 1$, introduced by Gumbel [7]. The marginal df's of H are F and G. For some fixed $\alpha \neq 0$ and $F_{(0)}$, $G_{(0)}$ let us choose $H_{(n)} = F_{(0)}G_{(0)} \times [1 + n^{-\frac{1}{2}}\alpha(1 - F_{(0)})(1 - G_{(0)})]$ (more general alternatives of this form are considered e.g. by Puri, Sen and Gokhale [11]). It is not hard to see that $H_{(n)} \to F_{(0)}G_{(0)}$ and that the limit (2.6) exists as $n \to \infty$.

3. Proof of Theorem 2.1: Asymptotic normality of the leading terms. Let

$$F^{-1}(s) = \inf\{x : F(x) \geqq s\}\,, \qquad G^{-1}(t) = \inf\{y : G(y) \geqq t\}\,. \tag{3.1}$$

If F is continuous these definitions imply $F(F^{-1}(s)) = s$, $F(x) < s$ if and only if $x < F^{-1}(s)$, $F(x) \geqq s$ if and only if $x \geqq F^{-1}(s)$ and similar statements for continuous G. The random functions $F_n(F^{-1})$ and $G_n(G^{-1})$ are with probability 1 the empirical df's of the sets of independent uniform (0, 1) rv's $F(X_1), \cdots, F(X_n)$ and $G(Y_1), \cdots, G(Y_n)$ respectively. Define the empirical processes $U_n(s) = n^{\frac{1}{2}}[F_n(F^{-1}(s)) - s]$, $V_n(t) = n^{\frac{1}{2}}[G_n(G^{-1}(t)) - t]$ and the process $U_n^*(s) = n^{\frac{1}{2}}[F_n^*(F^{-1}(s)) - s]$, $V_n^*(t) = n^{\frac{1}{2}}[G_n^*(G^{-1}(t)) - t]$ for $s, t \in [0, 1]$ (see (2.2)). With probability 1 these processes satisfy $U_n(F) = n^{\frac{1}{2}}(F_n - F)$, $V_n(G) = n^{\frac{1}{2}}(G_n - G)$ and $U_n^*(F) = n^{\frac{1}{2}}(F_n^* - F)$, $V_n^*(G) = n^{\frac{1}{2}}(G_n^* - G)$ on $(-\infty, \infty)$. For a suitable decomposition of (1.1) we need the following lemma.

LEMMA 3.1. *Let for $H \in \mathscr{H}$ Assumption* 2.4 (b) *be satisfied with $\mathscr{H}' = \{H\}$. Let ϕ and ψ be functions on the unit interval such that $\int_0^1 |\phi(s)|\, ds$, $\int_0^1 |\phi(s)f(s)|\, ds$, $\int_0^1 |\psi(t)|\, dt$, $\int_0^1 |\psi(t)g(t)|\, dt < \infty$. Here f and g are defined in Assumption* 2.4 (b). *Then*

(i) *$E(\psi(G(Y)) \mid F(X) = s)$ has a version continuous on O_1, to be denoted by $E_H(\psi \mid s)$;*

(ii) *$E(\phi(F(X)) \mid G(Y) = t)$ has a version continuous on O_2, to be denoted by $E_H(\phi \mid t)$.*

PROOF. It suffices to prove (i). Since (X, Y) has df H, $(F(X), G(Y))$ has df $H(F^{-1}, G^{-1})$, so that the latter df has uniform (0, 1) marginals. Consequently the function $\int_0^1 \psi(t)h(s, t)\, dt$ is a version of the conditional expectation considered in (i), restricted to O_1. Moreover, this version is continuous on O_1, for let s, $s + \zeta \in O_1$, and consider $\int_0^1 \psi(t)[h(s + \zeta, t) - h(s, t)]\, dt$. By continuity of the function h we have $h(s + \zeta, t) - h(s, t) \to 0$ as $\zeta \to 0$ for each $t \in (0, 1)$. By the assumptions of the lemma, moreover, we have $|\psi(t)||h(s + \zeta, t) - h(s, t)| \leqq 2|\psi(t)g(t)|$ for each $t \in (0, 1)$, and $\int_0^1 |\psi(t)g(t)|\, dt < \infty$. Finally, by the dominated convergence theorem, we obtain $\int_0^1 \psi(t)[h(s + \zeta, t) - h(s, t)]\, dt \to 0$ as $\zeta \to 0$. □

At this point let us give the basic decomposition

$$n^{\frac{1}{2}}(T_n - \mu) = A_{0n} + \sum_{i=1}^{2} (A'_{in} + A_{in}) + B_{0n}^* + B_n' + B_n + \tilde{B}_n' + \tilde{B}_n + C_n' + C_n\,, \tag{3.2}$$

with probability 1. Here B_{0n}^* is defined in (2.4), and using the notation of Lemma 3.1, we further have

$$\begin{aligned}
A_{0n} &= n^{\frac{1}{2}} \textstyle\iint J(F)K(G)d(H_n - H)\,, \\
A'_{1n} &= \textstyle\iint U_n(F)J'(F)K(G)\, dH\,, \qquad A_{1n} = \sum_{i=1}^{\lambda} \alpha_i E_H(K \mid s_i)U_n(s_i)\,, \\
A'_{2n} &= \textstyle\iint V_n(G)J(F)K'(G)\, dH\,, \qquad A_{2n} = \sum_{j=1}^{\nu} \beta_j E_H(J \mid t_j)V_n(t_j)\,, \\
B_n' &= n^{\frac{1}{2}} \textstyle\iint [J_c(F_n^*) - J_c(F)]K(G)\, dH_n - A'_{1n}\,, \\
B_n &= n^{\frac{1}{2}} \textstyle\iint [J_d(F_n^*) - J_d(F)]K(G)\, dH_n - A_{1n}\,, \\
\tilde{B}_n' &= n^{\frac{1}{2}} \textstyle\iint J(F)[K_c(G_n^*) - K_c(G)]\, dH_n - A'_{2n}\,, \\
\tilde{B}_n &= n^{\frac{1}{2}} \textstyle\iint J(F)[K_d(G_n^*) - K_d(G)]\, dH_n - A_{2n}\,, \\
C_n' &= n^{\frac{1}{2}} \textstyle\iint [J_c(F_n^*) - J_c(F)][K(G_n^*) - K(G)]\, dH_n\,, \\
C_n &= n^{\frac{1}{2}} \textstyle\iint [J_d(F_n^*) - J_d(F)][K(G_n^*) - K(G)]\, dH_n\,.
\end{aligned}$$

Let us note that $\tilde{B}_n'$, $\tilde{B}_n$ are symmetric to B_n', B_n. Therefore $\tilde{B}_n'$ and $\tilde{B}_n$ will not be treated in the sequel.

In this section attention will be restricted to the asymptotic normality of the A-terms. As far as A_{0n}, A_{1n}' and A_{2n}' are concerned see also [14]. The rv A_{0n} may be written in the form

$$A_{0n} = n^{-\frac{1}{2}} \sum_{k=1}^{n} A_{0kn} , \tag{3.3}$$

where the $A_{0kn} = J(F(X_k))K(G(Y_k)) - \mu$ are i.i.d. with mean zero. For the fixed df H (the fixed subclass of df's $\mathscr{H}'$) the rv A_{0kn} has a finite moment of order 2 (a finite absolute moment of order larger than 2, bounded on $\mathscr{H}'$) by Assumption 2.3.

Note that for δ as defined in (2.3) with probability 1 we have $\delta_{X_k}(x) = \delta_{F(X_k)}(F(x))$ and $\delta_{X_k}(F^{-1}(s_i)) = \delta_{F(X_k)}(s_i)$. Thus with probability 1 we have $U_n(F) = n^{-\frac{1}{2}} \sum_{k=1}^{n} [\delta_{F(X_k)}(F) - F]$ and $U_n(s_i) = n^{-\frac{1}{2}} \sum_{k=1}^{n} [\delta_{F(X_k)}(s_i) - s_i]$. By this and similar expressions for $V_n(G)$ and $V_n(t_j)$ we obtain

$$\begin{aligned} A_{1n}' &= n^{-\frac{1}{2}} \textstyle\sum_{k=1}^{n} A_{1kn}' , \qquad & A_{1n} &= n^{-\frac{1}{2}} \textstyle\sum_{k=1}^{n} A_{1kn} , \\ A_{2n}' &= n^{-\frac{1}{2}} \textstyle\sum_{k=1}^{n} A_{2kn}' , \qquad & A_{2n} &= n^{-\frac{1}{2}} \textstyle\sum_{k=1}^{n} A_{2kn} , \end{aligned} \tag{3.4}$$

where $A_{1kn}' = \iint [\delta_{F(X_k)}(F) - F]J'(F)K(G)\, dH$, $A_{1kn} = \sum_{i=1}^{\lambda} \alpha_i[\delta_{F(X_k)}(s_i) - s_i] \times E_H(K \mid s_i)$, $A_{2kn}' = \iint [\delta_{G(Y_k)}(G) - G]J(F)K'(G)\, dH$, $A_{2kn} = \sum_{j=1}^{\nu} \beta_j[\delta_{G(Y_k)}(t_j) - t_j]E_H(J \mid t_j)$, $k = 1, \cdots, n$. Each of these four sets of rv's consists of n i.i.d. rv's with mean zero. As to the A_{1kn} and A_{2kn} the absolute moments of any order exist for fixed df H (are bounded on the fixed subclass of df's $\mathscr{H}'$). To see the existence of higher order moments of the A_{1kn}' and A_{2kn}' we need the following property of q-functions.

LEMMA 3.2. *Let for arbitrary s, $u \in (0, 1)$ the symbol $\delta_s(u)$ be defined as in* (2.3), *and let q be any function in $\mathscr{Q}$* (see Definition 2.1). *Then there exists a constant $M = M_q$ (depending on q only) such that $|\delta_s(u) - u| \leqq Mq(u)[q(s)]^{-1}$ for $0 < s$, $u < 1$.*

PROOF. Because of the properties of q-functions, there exists a number $\varepsilon = \varepsilon_q$ satisfying $0 < \varepsilon < \frac{1}{2}$, such that $s \leqq q(s)$ for $0 \leqq s \leqq \varepsilon$. For suppose such ε does not exist. Then there is a sequence $s_n \downarrow 0$ satisfying $q(s_n) < s_n$, and hence $[q(s_n)]^{-2} > s_n^{-2}$. The reciprocal of q is square integrable on the unit interval; on the other hand $\int_0^1 [q(s)]^{-2}\, ds \geqq s_n^{-1} \to \infty$ as $n \to \infty$, which yields a contradiction. (Similarly, sharper bounds for q in the neighborhood of zero may be obtained.)

Let us first consider pairs $u < s$. Then $|\delta_s(u) - u|[q(u)]^{-1} \leqq u[q(u)]^{-1}$. For $0 < u \leqq \varepsilon$, with ε as above, we find $u[q(u)]^{-1} \leqq 1 \leqq M_1[q(s)]^{-1}$, if $M_1 = \max_{s \in [0,1]} q(s)$. For $\varepsilon < u \leqq \frac{1}{2}$ and M_1 as above we have $u[q(u)]^{-1} \leqq [q(\varepsilon)]^{-1} \leqq M_1[q(\varepsilon)]^{-1}[q(s)]^{-1}$. Finally, for $\frac{1}{2} < u < 1$, we simply have $u[q(u)]^{-1} \leqq [q(s)]^{-1}$. Evidently for $u < s$ the lemma holds with $M = \max\{M_1, M_1[q(\varepsilon)]^{-1}, 1\}$. For pairs $u \geqq s$ the proof can be given in the same way. □

Lemma 3.2 applied with $q = q_1$, where q_1 is the function introduced in Assumption 2.3, guarantees the existence of a constant $M_1 = M_{q_1}$ such that for each ω

$$|A'_{1kn}| \leqq M_1[q_1(F(X_k))]^{-1} \int\int q_1(F)\bar{r}_1(F)r_2(G)\,dH\,,$$

$k = 1, \cdots, n$. By Assumption 2.3 for the fixed df H (the fixed subclass of df's $\mathscr{H}'$) the random part $[q_1(F(X_k))]^{-1}$ of this upper bound possesses a finite moment of order 2 (a finite absolute moment of order larger than 2, bounded on $\mathscr{H}'$). It is due to the same assumption that for the fixed df H (the fixed subclass of df's $\mathscr{H}'$) the non-random integral is finite (bounded on $\mathscr{H}'$). A similar argument deals with A'_{2kn}.

Combining (3.3) and (3.4) we see that, given the fixed df H, for $k = 1, \cdots, n$ the sums $A_{0kn} + A'_{1kn} + A_{1kn} + A'_{2kn} + A_{2kn}$ are i.i.d. with mean zero and finite variance depending on H, equal to $\sigma^2 = \sigma^2(H) = \text{Var}\,(A_{0kn} + A'_{1kn} + A_{1kn} + A'_{2kn} + A_{2kn})$. Hence application of the central limit theorem gives $n^{-\frac{1}{2}} \sum_{k=1}^{n} (A_{0kn} + A'_{1kn} + A_{1kn} + A'_{2kn} + A_{2kn}) = A_{0n} + A'_{1n} + A_{1n} + A'_{2n} + A_{2n} \to_d \text{N}(0, \sigma^2)$. Since, given the fixed subclass $\mathscr{H}'$ of df's, a finite absolute moment of order larger than 2 is bounded on $\mathscr{H}'$, and since moreover the variance is given to be bounded away from zero on $\mathscr{H}'$, by Esséen's theorem the above asymptotic normality is uniform on $\mathscr{H}'$.

The variance $\sigma^2 = \sigma^2(H)$ of the limiting normal distribution can be given a nice expression using the conditional expectations, introduced in Lemma 3.1, and Stieltjes-Lebesgue-integrals: thus we obtain

$$\text{(3.5)} \qquad \sigma^2 = \sigma^2(H) = \text{Var}\,[J(F(X))K(G(Y)) + \textstyle\int_0^1 [\delta_{F(X)}(s) - s]E_H(K\,|\,s)\,dJ(s) + \int_0^1 [\delta_{G(Y)}(t) - t]E_H(J\,|\,t)\,dK(t)]\,.$$

In Section 6 this expression for the variance is studied in more detail (see (6.1)).

4. Proof of Theorem 2.1: Some lemmas. First we shall give a lemma needed for the proof of the asymptotic negligibility of the second order terms B_n' and C_n' connected with the continuous part of the score function J. This lemma is based on Lemma 2.2 of Pyke and Shorack [12] and is only slightly more general than [14], Lemmas 6.1 and 6.2. The proof will therefore be omitted.

LEMMA 4.1. *For each ω let $\Phi_n = \Phi_{n\omega}$ and $\Psi_n = \Psi_{n\omega}$ be functions on $\Lambda_{n1} = \Lambda_{n1\omega}$ and $\Lambda_{n2} = \Lambda_{n2\omega}$ respectively* (*see* (2.1)), *satisfying* $\min\{F, F_n^*\} \leqq \Phi_n \leqq \max\{F, F_n^*\}$ *and* $\min\{G, G_n^*\} \leqq \Psi_n \leqq \max\{G, G_n^*\}$ *where defined* (*see* (2.2)). *Then, uniformly for $n = 1, 2, \cdots$ and $H \in \mathscr{H}$*:

(i) $\quad \sup_{\Lambda_{n1}} r(\Phi_n)/r(F) = O_p(1)\,, \qquad \sup_{\Lambda_{n2}} r(\Psi_n)/r(G) = O_p(1)\,,$

for each $r \in \mathscr{R}$ (see Definition 2.2);

(ii) $\quad \sup_{\Lambda_{n1}} |U_n^*(F)|/q(F) = O_p(1)\,,$

for each $q \in \mathscr{Q}$ (see Definition 2.1);

(iii) $\quad \sup_{\Lambda_{n1}} |U_n^*(F) - U_n(F)|/q(F) = o_p(1)\,,$

for each $q \in \mathscr{Q}$.

The remaining lemmas of this section are specific for the second order terms B_n and C_n connected with the discontinuous part of the score function J. Let us denote the binomial distribution for n trials with success probability s by Bi $(n; s)$. It is well known (see e.g. Dvoretzky, Kiefer and Wolfowitz [5]) that if Z is a Bi $(n; s)$ distributed rv we have the exponential bound

$$\Pr(|Z - ns| \geqq n\rho) = O(\exp(-2n\rho^2)) \tag{4.1}$$

as $n \to \infty$, uniformly for $s \in (0, 1)$ and $\rho \geqq 0$. This result entails a useful property of the function $p_n(a, b; s)$, for fixed constants a, b and for $s \in (0, 1)$ defined by

$$p_n(a, b; s) = \sum_{j=0}^{n} \binom{n}{j} s^j (1-s)^{n-j} |\delta_{s_1}((j+a)/(n+b)) - \delta_{s_1}(s)| \, . \tag{4.2}$$

Here the function δ_{s_1} is defined in (2.3) with $s_1 \in (0, 1)$.

Lemma 4.2. *Let a and b be fixed constants and let $p_n(a, b; s)$ be defined as in* (4.2). *Then*

(i) $$p_n(a, b; s) = O(\exp(-2n(s - s_1)^2)) \qquad as \quad n \to \infty \, ,$$

uniformly for $s, s_1 \in (0, 1)$;

(ii) $$\int_0^1 p_n(a, b; s)\, ds = O(n^{-\frac{1}{2}}) \qquad as \qquad n \to \infty \, .$$

Proof. (i) The function $p_n(a, b; s)$ is unequal to zero only if $s < s_1$ and $j \geqq (n + b)s_1 - a$, or $s \geqq s_1$ and $j < (n + b)s_1 - a$. Suppose $s < s_1$. Then $p_n(a, b; s) = \Pr(Z \geqq (n + b)s_1 - a)$, where Z is a Bi $(n; s)$ distributed rv. Since $(n + b)s_1 - a = n(s + [s_1 - s + (bs_1 - a)/n])$, we have by (4.1) since a and b are fixed

$$\begin{aligned}\Pr(Z \geqq (n + b)s_1 - a) &\leqq M_0 \exp(-2n[s_1 - s + (bs_1 - a)/n]^2) \\ &\leqq M_1 \exp(-2n(s_1 - s)^2) \, ,\end{aligned}$$

provided $s_1 - s + (bs_1 - a)/n \geqq 0$. Now consider the set $D = \{s : s_1 + (bs_1 - a)/n < s < s_1\}$. If D is empty there is nothing left to prove, hence suppose D is not empty. Then $\sup_{s \in D; n=1,2,\cdots} \exp(2n(s_1 - s)^2) \leqq \max_{n=1,2,\cdots} \exp(2(bs_1 - a)^2/n) = \exp(2(bs_1 - a)^2) = M_2$, say. Since p_n is a probability it is always bounded by 1 and hence by $M_2 \exp(-2n(s - s_1)^2)$ for all $s \in D$ and all $n = 1, 2, \cdots$. We thus have, letting $M = \max\{M_1, M_2\}$, that p_n is bounded by $M \exp(-2n(s - s_1)^2)$ for all $s < s_1$ and $n = 1, 2, \cdots$. This inequality can similarly be shown to hold for $s \geqq s_1$.

(ii) This follows at once from part (i) by

$$\int_0^1 p_n(a, b; s)\, ds \leqq M \int_{-\infty}^{\infty} \exp(-2n(s - s_1)^2)\, ds = O(n^{-\frac{1}{2}}) \qquad \text{as} \quad n \to \infty \, . \quad \square$$

Lemma 4.3. *Let ψ be a function on the unit interval with $\int_0^1 |\psi(t)|\, dt < \infty$. Then for any $H \in \mathscr{H}$ the following holds*

(i) $$\begin{aligned} &E \iint |[\delta_{s_1}(F_n^*) - \delta_{s_1}(F)]\psi(G)|\, dH \\ &\qquad \leqq \int_0^1 p_n(0, 1; s) E(|\psi(G(Y))| \mid F(X) = s)\, ds \, ;\end{aligned}$$

(ii) $$\begin{aligned} &E \iint |[\delta_{s_1}(F_n^*) - \delta_{s_1}(F)]\psi(G)|\, dH_n \\ &\qquad \leqq \int_0^1 p_{n-1}(1, 2; s) E(|\psi(G(Y))| \mid F(X) = s)\, ds \, .\end{aligned}$$

PROOF. (i) Because $P(\{(n+1)F_n^*(x) = j\}) = \binom{n}{j}F^j(x)(1-F(x))^{n-j}$ for $j = 0, 1, \cdots, n$ we obtain

$$\begin{aligned} E \textstyle\iint |[\delta_{s_1}(F_n^*) - \delta_{s_1}(F)]\psi(G)|\, dH \\ &\leqq \textstyle\iint \sum_{j=0}^n \binom{n}{j} F^j(1-F)^{n-j} |[\delta_{s_1}(j/(n+1)) - \delta_{s_1}(F)]\psi(G)|\, dH \\ &= \textstyle\int_0^1 p_n(0, 1; s) E(|\psi(G(Y))| \mid F(X) = s)\, ds\,. \end{aligned}$$

(ii) Similarly, since $P(\{(n+1)F_n^*(X_i) = j\} \mid F(X_i)) = \binom{n-1}{j-1}F^{j-1}(X_i)(1-F(X_i))^{n-j}$ for $j = 1, \cdots, n$, we have

$$\begin{aligned} E \textstyle\iint |[\delta_{s_1}(F_n^*) - \delta_{s_1}(F)]\psi(G)|\, dH_n \\ &\leqq E(E(|[\delta_{s_1}(F_n^*(X_i)) - \delta_{s_1}(F(X_i))]\psi(G(X_i))| \mid F(X_i), G(Y_i))) \\ &= E(|\psi(G(Y_i))| \cdot E(|\delta_{s_1}(F_n^*(X_i)) - \delta_{s_1}(F(X_i))| \mid F(X_i))) \\ &= \textstyle\iint \sum_{j=1}^n \binom{n-1}{j-1} F^{j-1}(1-F)^{n-j} |[\delta_{s_1}(j/(n+1)) - \delta_{s_1}(F)]\psi(G)|\, dH \\ &= \textstyle\int_0^1 p_{n-1}(1, 2; s) E(|\psi(G(Y))| \mid F(X) = s)\, ds\,. \qquad \square \end{aligned}$$

The last lemma is a corollary to Kiefer [10], Theorem 1-m; it is due to W. R. van Zwet. Like Kiefer's theorem, Lemma 4.4 can be formulated for m-dimensional random vectors. To avoid the introduction of additional notational conventions we shall restrict attention to the case where $m = 2$. One of the basic supports of Kiefer's theorem quoted above is a sharpening of the exponential bound (4.1); for $m = 2$ the theorem implies that for any fixed $\zeta > 0$

$$P(\{\sup_{-\infty<x,y<\infty} |H_n(x,y) - H(x,y)| \geqq \rho\}) = O(\exp(-(2-\zeta)n\rho^2))\,, \tag{4.3}$$

uniformly for all bivariate df's H (continuous or not) and uniformly for all $\rho \geqq 0$. For a comparison between Lemma 4.4 and related results of Bahadur [1], Sen [15] and Ghosh [6] see Section 1. For any Borel set D in the plane we shall write $\iint_D dH = H\{D\}$. By an interval I in the plane the product set of two intervals on the line will be understood.

LEMMA 4.4. (van Zwet). *Let $I_1, I_2, \cdots$ be a sequence of intervals in the plane and let $\mathscr{I}_n = \{I_n^* : I_n^*$ is an interval contained in $I_n\}$, $n = 1, 2, \cdots$. Then*

$$\sup_{I_n^* \in \mathscr{I}_n} |H_n\{I_n^*\} - H\{I_n^*\}| = O_p([H\{I_n\}/n]^{\frac{1}{2}})$$

as $n \to \infty$, uniformly in all sequences of intervals $I_1, I_2, \cdots$ and all bivariate df'*s H (continuous or not).*

PROOF. Given any $0 < \varepsilon < 1$, the existence of a number $M = M_\varepsilon$ must be established such that

$$P(\{\sup_{I_n^* \in \mathscr{I}_n} |H_n\{I_n^*\} - H\{I_n^*\}| \geqq M[H\{I_n\}/n]^{\frac{1}{2}}\}) \leqq \varepsilon\,, \tag{4.4}$$

for all n, uniformly in all sequences of intervals $I_1, I_2, \cdots$ and all bivariate df's H. If $H\{I_n\} = 0$ the lemma follows immediately. It proves to be convenient to consider the cases $0 < H\{I_n\} \leqq 8/(\varepsilon n)$ and $H\{I_n\} > 8/(\varepsilon n)$ separately.

First suppose that $0 < H\{I_n\} \leqq 8/(\varepsilon n)$ and choose $M = M_\varepsilon = (2/\varepsilon)^{\frac{3}{2}}$. It is always true that $\sup_{I_n^* \in \mathscr{I}_n} |H_n\{I_n^*\} - H\{I_n^*\}| \leqq \max\{H_n\{I_n\}, H\{I_n\}\}$. By our choice of M we have $M[H\{I_n\}/n]^{\frac{1}{2}} \geqq H\{I_n\}/\varepsilon$. Consequently we only have to prove the same inequality for $H_n\{I_n\}$. Since $nH_n\{I_n\}$ is a Bi $(n; H\{I_n\})$ distributed rv, application of Markov's inequality shows that the left side in (4.4) is bounded above by $P(\{\max\{H_n\{I_n\}, H\{I_n\}\} \geqq H\{I_n\}/\varepsilon\}) = P(\{H_n\{I_n\} \geqq H\{I_n\}/\varepsilon\}) \leqq \varepsilon$.

Next we suppose that $H\{I_n\} > 8/(\varepsilon n)$. Then for $k = 0, 1, \cdots, n$ we may define the conditional probabilities

$$\pi(k) = P(\{\sup_{I_n^* \in \mathscr{I}_n} |H_n\{I_n^*\} - H\{I_n^*\}| \geqq M[H\{I_n\}/n]^{\frac{1}{2}}\} \mid \{H_n\{I_n\} = k/n\}) .$$

The probability on the left in (4.4) can now be written as

$$\sum_{k < nH\{I_n\}/2} \pi(k) P(\{H_n\{I_n\} = k/n\}) + \sum_{k \geqq nH\{I_n\}/2} \pi(k) P(\{H_n\{I_n\} = k/n\}) . \tag{4.5}$$

By the Bienaymé–Chebyshev inequality we have

$$\begin{aligned} &\sum_{k < nH\{I_n\}/2} \pi(k) P(\{H_n\{I_n\} = k/n\}) \\ &\quad \leqq P(\{H_n\{I_n\} \leqq H\{I_n\}/2\}) \\ &\quad \leqq P(\{|H_n\{I_n\} - H\{I_n\}| > H\{I_n\}/2\}) \leqq 4/(nH\{I_n\}) < \varepsilon/2 , \end{aligned} \tag{4.6}$$

since by assumption $H\{I_n\} > 8/(\varepsilon n)$. In the second term in (4.5) only values $k \neq 0$ are involved. As $H\{I_n\} > 0$, we find that for any $k \neq 0$, we have, conditional on $H_n\{I_n\} = k/n$,

$$\begin{aligned} &\sup_{I_n^* \in \mathscr{I}_n} |H_n\{I_n^*\} - H\{I_n^*\}| \\ &\quad \leqq H\{I_n\}\left[\sup_{I_n^* \in \mathscr{I}_n} \left|\frac{H_n\{I_n^*\}}{H\{I_n\}} - \frac{H_n\{I_n^*\}}{H_n\{I_n\}}\right| + \sup_{I_n^* \in \mathscr{I}_n} \left|\frac{H_n\{I_n^*\}}{H_n\{I_n\}} - \frac{H\{I_n^*\}}{H\{I_n\}}\right|\right] \\ &\quad = |H_n\{I_n\} - H\{I_n\}| + H\{I_n\} \sup_{I_n^* \in \mathscr{I}_n} |\tilde{H}_k\{I_n^*\} - \tilde{H}\{I_n^*\}| . \end{aligned}$$

Here $\tilde{H}\{I_n^*\} = H\{I_n^*\}/H\{I_n\}$ is the conditional probability that the random vector (X, Y) is an element of $I_n^* \subset I_n$ under the hypothesis that it is an element of I_n. Given $H_n\{I_n\} = k/n$ with $k \neq 0$, the ratio $\tilde{H}_k\{I_n^*\} = H_n\{I_n^*\}/H_n\{I_n\}$ is distributed as the empirical df corresponding to $\tilde{H}$, based on $k \neq 0$ observations.

For $k \neq 0$ we have $\pi(k) \leqq \pi_1(k) + \pi_2(k)$, where

$$\begin{aligned} \pi_1(k) &= P(\{|H_n\{I_n\} - H\{I_n\}| \geqq M[H\{I_n\}/4n]^{\frac{1}{2}}\} \mid \{H_n\{I_n\} = k/n\}) , \\ \pi_2(k) &= P(\{\sup_{I_n^* \in \mathscr{I}_n} |\tilde{H}_k\{I_n^*\} - \tilde{H}\{I_n^*\}| \geqq M[4nH\{I_n\}]^{-\frac{1}{2}}\}) . \end{aligned}$$

Applying the Bienaymé–Chebyshev inequality once more we obtain

$$\begin{aligned} &\sum_{k \geqq nH\{I_n\}/2} \pi_1(k) P(\{H_n\{I_n\} = k/n\}) \\ &\qquad \leqq P(\{|H_n\{I_n\} - H\{I_n\}| \geqq M[H\{I_n\}/4n]^{\frac{1}{2}}\}) \leqq 4/M^2 . \end{aligned} \tag{4.7}$$

Finally we have to consider the summation involving the $\pi_2(k)$. For any interval I in the plane we have $|\tilde{H}_k\{I\} - \tilde{H}\{I\}| \leqq 4 \sup_{-\infty < x, y < \infty} |\tilde{H}_k(x, y) - \tilde{H}(x, y)|$. According to formula (4.3), applied to $\tilde{H}_k$ and $\tilde{H}$ with e.g. $\zeta = 1$, there exists a constant M_1 such that

$$\pi_2(k) \leqq M_1 \exp(-kM^2/(64nH\{I_n\})) ,$$

and hence

$$(4.8)\qquad \sum_{k\geq nH\{I_n\}/2} \pi_2(k)P(\{H_n\{I_n\} = k/n\}) \\ \leq M_1 \exp(-nH\{I_n\}M^2/(128nH\{I_n\})) = M_1 \exp(-M^2/128)\,.$$

Combining (4.6), (4.7) and (4.8) we see that for $H\{I_n\} > 8/(\varepsilon n)$ inequality (4.4) holds, provided M is chosen so large that both (4.7) and (4.8) are smaller than $\varepsilon/4$. Let us finally note that the argument is independent of the sequence $I_1, I_2, \cdots$ and the bivariate df H. □

5. Proof of Theorem 2.1: Asymptotic negligibility of the remainder terms. As has already been noted in Section 3, the rv's $\tilde{B}_n'$, $\tilde{B}_n$ are symmetric to B_n', B_n and hence need not be considered. Since J_c is continuous on (0, 1) and continuously differentiable on the open intervals between the points $0, s_1, \cdots, s_\lambda, 1$, the second order terms B_n' and C_n' can be dealt with in essentially the same way as the B_n^*—and C_n^*—terms in [14], Section 6. We only have to use Lemma 4.1 instead of [14], Lemmas 6.1 and 6.2. Although in the present case the function K is no longer continuous it is easily seen that this does not affect the argument, because the mean value theorem is applied only to J_c.

Therefore we may restrict attention to the terms B_n and C_n. It suffices to consider the case where (see Assumption 2.2)

$$J_d = \delta_{s_1}\,, \qquad K_d = \delta_{t_1}\,,$$

for fixed $s_1, t_1 \in (0, 1)$. Given any set D, $\bar{D}$ will denote its complement, $\chi(D)$ its indicator function and $\chi(D; x)$ the value of this function at the point x. For small positive γ we adopt the notation

$$(5.1)\qquad S_{\gamma 2} = [G^{-1}(\gamma), G^{-1}(t_1 - \gamma)] \cup [G^{-1}(t_1 + \gamma), G^{-1}(1 - \gamma)]\,,$$

where G^{-1} is defined in (3.1).

Since K satisfies the conditions of Lemma 3.1, the conditional expectation $E(K(G(Y))\,|\,F(X) = s)$ possesses a version which is continuous on the open set O_1 defined in Assumption 2.4. By convention this special version will be denoted by $E_H(K\,|\,s)$. Let us write B_n and C_n as

$$B_n = B_{1n} + \sum_{i=2}^{4} B_{\gamma in}\,, \qquad C_n = \sum_{i=1}^{3} C_{\gamma in}\,,$$

where

$$\begin{aligned}
B_{1n} &= n^{\frac{1}{2}} \iint [\delta_{s_1}(F_n^*) - \delta_{s_1}(F)]K(G)\,dH - U_n(s_1)E_H(K\,|\,s_1)\,,\\
B_{\gamma 2n} &= n^{\frac{1}{2}} \iint_{(-\infty,\infty)\times S_{\gamma 2}} [\delta_{s_1}(F_n^*) - \delta_{s_1}(F)]K(G)d(H_n - H)\,,\\
B_{\gamma 3n} &= -n^{\frac{1}{2}} \iint_{(-\infty,\infty)\times \bar{S}_{\gamma 2}} [\delta_{s_1}(F_n^*) - \delta_{s_1}(F)]K(G)\,dH\,,\\
B_{\gamma 4n} &= n^{\frac{1}{2}} \iint_{(-\infty,\infty)\times \bar{S}_{\gamma 2}} [\delta_{s_1}(F_n^*) - \delta_{s_1}(F)]K(G)\,dH_n\,,\\
C_{\gamma 1n} &= n^{\frac{1}{2}} \iint_{(-\infty,\infty)\times \bar{S}_{\gamma 2}} [\delta_{s_1}(F_n^*) - \delta_{s_1}(F)]K(G_n^*)\,dH_n\,,\\
C_{\gamma 2n} &= n^{\frac{1}{2}} \iint_{(-\infty,\infty)\times S_{\gamma 2}} [\delta_{s_1}(F_n^*) - \delta_{s_1}(F)][K(G_n^*) - K(G)]\,dH_n\,,\\
C_{\gamma 3n} &= -n^{\frac{1}{2}} \iint_{(-\infty,\infty)\times \bar{S}_{\gamma 2}} [\delta_{s_1}(F_n^*) - \delta_{s_1}(F)]K(G)\,dH_n\,.
\end{aligned}$$

From this we see that $B_{\gamma 4n}$ and $C_{\gamma 3n}$ cancel out. Throughout this section let

$\eta > 0$ be a fixed number small enough to ensure that $[s_1 - \eta, s_1 + \eta] \subset O_1$ (see Assumption 2.4), and let an arbitrary $\varepsilon > 0$ be given.

The asymptotic negligibility of B_{1n} and $B_{\gamma 2n}$ is mainly based on Lemma 4.4. Let $m = m(n)$ be the fixed sequence of natural numbers uniquely determined by

$$(5.2) \qquad (n+1)s_1 \leqq m < (n+1)s_1 + 1 .$$

If we define the function $\operatorname{sgn} x = -1$ for $x < 0$, $\operatorname{sgn} x = 0$ for $x = 0$, $\operatorname{sgn} x = 1$ for $x > 0$ we have

$$(5.3) \qquad \delta_{s_1}(F_n^*(x)) - \delta_{s_1}(F(x)) = \operatorname{sgn}(F^{-1}(s_1) - X_{mn})\chi(\Gamma_{n1}; x)$$

for each ω and all x. Here

$$(5.4) \qquad \Gamma_{n1} = [\min\{X_{mn}, F^{-1}(s_1)\}, \max\{X_{mn}, F^{-1}(s_1)\}] .$$

To verify the equality (5.3) we use that δ_{s_1} is continuous from the right in s_1 (see (2.3)) and we use the last two properties of F^{-1}, given below the definition in (3.1). From the properties of empirical df's and order statistics it follows that there exists a constant $M_0 = M_{0\varepsilon}$ such that

$$(5.5) \qquad \Omega_{0n} = \{F^{-1}(F_n^*(F^{-1}(s_1))), X_{mn} \in [F^{-1}(s_1 - M_0 n^{-\frac{1}{2}}), F^{-1}(s_1 + M_0 n^{-\frac{1}{2}})]\}$$

has probability $P(\Omega_{0n}) \geqq 1 - \varepsilon/2$ for all n and $H \in \mathscr{H}$. Let us further define

$$I_{n1} = [F^{-1}(s_1 - M_0 n^{-\frac{1}{2}}), F^{-1}(s_1 + M_0 n^{-\frac{1}{2}})] .$$

Applying Lemma 4.4 with $I_n = I_{n1} \times (-\infty, \infty)$, and thus with $H\{I_n\} = 2M_0 n^{-\frac{1}{2}}$, we find by (5.4) and (5.5) that there exists a constant $M_1 = M_{1\varepsilon}$ such that

$$(5.6) \qquad \Omega_{1n} = \Omega_{0n} \cap \{\sup_{I_{n2}^*} |H_n\{\Gamma_{n1} \times I_{n2}^*\} - H\{\Gamma_{n1} \times I_{n2}^*\}| \leqq M_1 n^{-\frac{3}{4}}\}$$

has probability $P(\Omega_{1n}) \geqq 1 - \varepsilon$ for all n and all $H \in \mathscr{H}$. Here the supremum is taken over all intervals $I_{n2}^* \subset (-\infty, \infty)$.

COROLLARY 5.1. *As $n \to \infty$, $B_{1n} \to_p 0$ uniformly for $H \in \mathscr{H}'$.*

PROOF. Let us consider only values of n large enough to ensure that $I_{n1} \subset [F^{-1}(s_1 - \eta), F^{-1}(s_1 + \eta)]$. Using the above notation and results we may write $B_{1n} = n^{\frac{1}{2}} \iint_{\Gamma_{n1} \times (-\infty,\infty)} \operatorname{sgn}(F^{-1}(s_1) - X_{mn}) K(G)\, dH - U_n(s_1) E_H(K \mid s_1) = \sum_{i=1}^{3} B_{1in}$, where

$$\begin{aligned} B_{11n} &= \chi(\bar{\Omega}_{1n}) B_{1n} , \\ B_{12n} &= \chi(\Omega_{1n})[n^{\frac{1}{2}} \textstyle\int_{2s_1 - F_n^*(F^{-1}(s_1))}^{s_1} E_H(K \mid s)\, ds - U_n(s_1) E_H(K \mid s_1)] , \\ B_{13n} &= \chi(\Omega_{1n}) n^{\frac{1}{2}} \textstyle\int_{F(X_{mn})}^{2s_1 - F_n^*(F^{-1}(s_1))} E_H(K \mid s)\, ds . \end{aligned}$$

By Assumption 2.4 we have

$$(5.7) \qquad \sup_{s \in [s_1 - \eta, s_1 + \eta]; H \in \mathscr{H}'} |E_H(K \mid s)| = M < \infty ,$$

and since $E_H(K \mid s)$ is continuous on $[s_1 - \eta, s_1 + \eta]$ the mean value theorem for integrals applies. We thus obtain, writing $\Phi_n(s_1)$ for a random point between

s_1 and $2s_1 - F_n^*(F^{-1}(s_1))$ and using (5.7),

$$|B_{12n}| \leqq \chi(\Omega_{1n})n^{\frac{1}{2}}|F_n^*(F^{-1}(s_1)) - s_1||E_H(K \mid \Phi_n(s_1)) - E_H(K \mid s_1)| \\ + \chi(\Omega_{1n})Mn^{\frac{1}{2}}|F_n(F^{-1}(s_1)) - F_n^*(F^{-1}(s_1))| .$$

By (5.6) and (5.5) for each $\omega \in \Omega_{1n}$ the random point $\Phi_n(s_1)$ satisfies $|\Phi_n(s_1) - s_1| \leqq M_0 n^{-\frac{1}{2}}$, so that the equicontinuity condition concerning the densities h corresponding to the $H \in \mathscr{H}'$ (see Assumption 2.4) yields that the first term in the bound for $|B_{12n}|$ converges to zero as $n \to \infty$, uniformly for all $H \in \mathscr{H}'$. The same holds for the second term in this bound, since $|F_n(F^{-1}(s_1)) - F_n^*(F^{-1}(s_1))| = 1/(n+1)$.

The rv B_{13n} is bounded by

$$\begin{aligned} |B_{13n}| &\leqq \chi(\Omega_{1n})Mn^{\frac{1}{2}}|F_n(F^{-1}(s_1)) - F_n(X_{mn}-) + F(X_{mn}) - s_1| \\ &\quad + \chi(\Omega_{1n})Mn^{\frac{1}{2}}|F_n^*(F^{-1}(s_1)) - F_n(F^{-1}(s_1)) + F_n(X_{mn}-) - s_1| \\ &\leqq \chi(\Omega_{1n})Mn^{\frac{1}{2}}|H_n\{\Gamma_{n1} \times (-\infty, \infty)\} - H\{\Gamma_{n1} \times (-\infty, \infty)\}| \\ &\quad + \chi(\Omega_{1n})Mn^{\frac{1}{2}}[1/(n+1) + |(m-1)/n - s_1|] \to 0 \end{aligned}$$

as $n \to \infty$, uniformly for $H \in \mathscr{H}'$, by (5.2), (5.6) and (5.7).

Since by (5.6) $P(\{B_{11n} \neq 0\}) \leqq \varepsilon$ for all n and all $H \in \mathscr{H}$, where $\varepsilon > 0$ is arbitrary, the conclusion of the corollary follows. □

COROLLARY 5.2. *For fixed* γ, $B_{\gamma 2n} \to_p 0$ *as* $n \to \infty$, *uniformly for* $H \in \mathscr{H}$.

PROOF. For each positive integer k we obtain the function K_k from the function K by

$$K_k(t) = K((i-1)/k) \qquad \text{for} \quad t \in (0,1) \cap [(i-1)/k, i/k], \qquad i = 1, \cdots, k .$$

For any such k, using (5.3), let us make the decomposition $B_{\gamma 2n} = B_{\gamma 21n} + \sum_{i=2}^4 B_{\gamma 2ikn}$, where

$$\begin{aligned} B_{\gamma 21n} &= \chi(\bar{\Omega}_{1n})B_{\gamma 2n} , \\ B_{\gamma 22kn} &= \chi(\Omega_{1n})n^{\frac{1}{2}} \textstyle\iint_{\Gamma_{n1}\times S_{\gamma 2}} \operatorname{sgn}(F^{-1}(s_1) - X_{mn})K_k(G)d(H_n - H) , \\ B_{\gamma 23kn} &= \chi(\Omega_{1n})n^{\frac{1}{2}} \textstyle\iint_{(-\infty,\infty)\times S_{\gamma 2}} [\delta_{s_1}(F_n^*) - \delta_{s_1}(F)][K(G) - K_k(G)]\, dH_n , \\ B_{\gamma 24kn} &= \chi(\Omega_{1n})n^{\frac{1}{2}} \textstyle\iint_{(-\infty,\infty)\times S_{\gamma 2}} [\delta_{s_1}(F_n^*) - \delta_{s_1}(F)][K_k(G) - K(G)]\, dH . \end{aligned}$$

For arbitrary fixed ω the integrand in the expression for $B_{\gamma 22kn}$ is a simple step function assuming the values $a_{\gamma jkn}(\omega)$ on the set $\Gamma_{n1} \times S_{\gamma j2}$, where

$$S_{\gamma j2} = [G^{-1}((j-1)/k), G^{-1}(j/k)] \cap S_{\gamma 2} ,$$

for $j = 1, \cdots, k$. Let $M_\gamma = \max_{S_{\gamma 2}} |K(G)|$, then by (5.6) we have for every ω

$$\begin{aligned} |B_{\gamma 22kn}| &= \chi(\Omega_{1n})n^{\frac{1}{2}}|\textstyle\sum_{j=1}^k a_{\gamma jkn} \iint_{\Gamma_{n1}\times S_{\gamma j2}} d(H_n - H)| \\ &\leqq \chi(\Omega_{1n})n^{\frac{1}{2}}M_\gamma \textstyle\sum_{j=1}^k |H_n\{\Gamma_{n1} \times S_{\gamma j2}\} - H\{\Gamma_{n1} \times S_{\gamma j2}\}| \\ &\leqq kM_\gamma M_1 n^{-\frac{1}{2}} \to 0 \end{aligned}$$

for fixed k as $n \to \infty$, uniformly for $H \in \mathscr{H}$.

Since $K(G)$ is bounded and continuous on $S_{\gamma 2}$ we have $\sup_{S_{\gamma 2}} |K(G) - K_k(G)| = \zeta_{\gamma k} \to 0$ for fixed γ as $k \to \infty$, uniformly for $H \in \mathscr{H}$. Application of Lemma 4.3(ii) and (i) with $\psi(G) = \zeta_{\gamma k}$ gives for the expectations of $|B_{\gamma 23kn}|$ and $|B_{\gamma 24kn}|$ the bounds (see also (4.2))

$$E(|B_{\gamma 23kn}|) \leqq n^{\frac{1}{2}} \zeta_{\gamma k} \int_0^1 p_{n-1}(1, 2; s)\, ds ,$$
$$E(|B_{\gamma 24kn}|) \leqq n^{\frac{1}{2}} \zeta_{\gamma k} \int_0^1 p_n(0, 1; s)\, ds$$

respectively. Since for fixed γ the sequence $\zeta_{\gamma k} \to 0$ as $k \to \infty$, application of Lemma 4.2(ii) leads to the conclusion that both expectations tend to zero for fixed γ as $k, n \to \infty$, uniformly for $H \in \mathscr{H}$.

As to $B_{\gamma 21n}$, by (5.6) $P(\{B_{\gamma 21n} \neq 0\}) \leqq \varepsilon$ for all n and all $H \in \mathscr{H}$, where $\varepsilon > 0$ is arbitrary. Combination of these partial results leads to the conclusion of the corollary. □

The rv's $B_{\gamma 3n}$ and $C_{\gamma 1n}$ concern the behavior of the functions $K(G(y))$ and $K(G_n^*(y))$ respectively for large values of $|y|$. Since by Assumption 2.2 the score function $|K| \leqq r_2$ on (0, 1) we have $|K(G)| \leqq r_2(G)$ and $|K(G_n^*)| \leqq r_2(G_n^*)$ on Λ_{n2} (see (2.1), (2.2)). By the reproducing u-shaped character of r_2 (see Definition 2.2), it is possible to replace the random argument G_n^* by the non-random argument G in the latter case, which may be seen from application of Lemma 4.1(i) with with $\Psi_n = G_n^*$. According to this lemma for each $\varepsilon > 0$ there exists a number $M_2 = M_{2\varepsilon}$ such that the set

$$\Omega_{2n} = \{r_2(G_n^*) \leqq M_2 r_2(G) \text{ on } \Lambda_{n2}\} \tag{5.8}$$

has probability $P(\Omega_{2n}) \geqq 1 - \varepsilon$ for all n and all $H \in \mathscr{H}$. Thus the asymptotic negligibility of the rv's $B_{\gamma 3n}$ and $C_{\gamma 1n}$ may be obtained essentially in the same way (note that the random measure dH_n restricts integration to the random set Λ_n). It is mainly based on Lemma 4.3. The asymptotic negligibility of $C_{\gamma 2n}$ is a simple application of the same lemma.

COROLLARY 5.3. *As* $\gamma \downarrow 0$ *and* $n \to \infty$, $B_{\gamma 3n} \to_p 0$ *and* $C_{\gamma 1n} \to_p 0$, *uniformly for* $H \in \mathscr{H}'$.

PROOF. For small positive γ, let us introduce the function

$$r_{2\gamma}(t) = r_2(t) \quad \text{for} \quad t \in (0, \gamma) \cup (t_1 - \gamma, t_1 + \gamma) \cup (1 - \gamma, 1) ,$$
$$r_{2\gamma}(t) = 0 \quad \text{elsewhere.}$$

Because by Assumption 2.4 the functions $r_{2\gamma}$ satisfy the conditions of Lemma 3.1 for such values of γ, the conditional expectations $E(r_{2\gamma}(G(Y)) \mid F(X) = s)$ have versions continuous on the open set O_1, by convention denoted by $E_H(r_{2\gamma} \mid s)$. Since $r_{2\gamma} \downarrow 0$ on (0, 1) as $\gamma \downarrow 0$, by the dominated convergence theorem and Assumption 2.4 we have as $\gamma \downarrow 0$

$$\sup_{s \in [s_1 - \eta, s_1 + \eta]; H \in \mathscr{H}'} E_H(r_{2\gamma} \mid s) = \zeta_\gamma \leqq \int_0^1 r_{2\gamma}(t) g(t)\, dt \to 0 . \tag{5.9}$$

As to $B_{\gamma 3n}$, application of Lemma 4.3 (i) yields (see also (4.2))

$$E(|B_{\gamma 3n}|) \leqq n^{\frac{1}{2}} \int_0^1 p_n(0, 1; s) E_H(r_{2\gamma} \mid s)\, ds\,. \tag{5.10}$$

As to $C_{\gamma 1n}$, using Lemma 4.3 (ii) and (5.8) we find

$$E(\chi(\Omega_{2n})|C_{\gamma 1n}|) \leqq M_2 n^{\frac{1}{2}} \int_0^1 p_{n-1}(1, 2; s) E_H(r_{2\gamma} \mid s)\, ds\,. \tag{5.11}$$

Because of the similarity between the right-hand sides of (5.10) and (5.11) and because $P(\Omega_{2n}) \geqq 1 - \varepsilon$ for all n and $H \in \mathscr{H}$, it suffices to investigate the right-hand side of (5.10). By Lemma 4.2 and (5.9) for that expression we find the bound

$$n^{\frac{1}{2}}[\sup_{s \in [0, s_1 - \eta] \cup [s_1 + \eta, 1]} p_n(0, 1; s)][\int_0^1 r_2(t)\, dt] + n^{\frac{1}{2}}[\int_0^1 p_n(0, 1; s)\, ds]\zeta_\gamma \to 0$$

as $\gamma \downarrow 0$ and $n \to \infty$, uniformly for $H \in \mathscr{H}'$. □

COROLLARY 5.4. *For fixed* γ, $C_{\gamma 2n} \to_p 0$ *as* $n \to \infty$, *uniformly for* $H \in \mathscr{H}$.

PROOF. As dH_n restricts integration to Λ_n, application of Lemma 4.3 (ii) with $\psi(G) = 1$ gives

$$|C_{\gamma 2n}| \leqq \sup_{\Lambda_{n2} \cap S_{\gamma 2}} |K(G_n^*) - K(G)| n^{\frac{1}{2}} \int_0^1 p_{n-1}(1, 2; s)\, ds\,.$$

The function K is uniformly continuous on $[\gamma/2, t_1 - \gamma/2] \cup [t_1 + \gamma/2, 1 - \gamma/2]$ and $|G_n^* - G| \leqq 1/(n + 1) + |G_n - G|$. Hence by the Glivenko-Cantelli theorem we have $\sup_{\Lambda_{n2} \cap S_{\gamma 2}} |K(G_n^*) - K(G)| \to_p 0$ as $n \to \infty$, uniformly for $H \in \mathscr{H}$. The proof may be concluded by applying Lemma 4.2 (ii). □

In order to show that $B_n + C_n \to_p 0$ as $n \to \infty$, uniformly for $H \in \mathscr{H}'$, given an arbitrary $\varepsilon > 0$, first use Corollary 5.3 to choose a fixed $\tilde{\gamma}$ and an index n_0 such that $P(\{|B_{\tilde{\gamma}3n}|, |C_{\tilde{\gamma}1n}| \leqq \varepsilon\}) \geqq 1 - \varepsilon$ for all $n \geqq n_0$ and all $H \in \mathscr{H}'$. Next application of Corollaries 5.1, 5.2 and 5.4 with the above fixed $\tilde{\gamma}$ gives the existence of an index $n_1 = n_1(\tilde{\gamma}) > n_0$, such that $P(\{|B_{1n}|, |B_{\tilde{\gamma}2n}|, |C_{\tilde{\gamma}2n}| \leqq \varepsilon\}) \geqq 1 - \varepsilon$ for all $n \geqq n_1$ and all $H \in \mathscr{H}'$. Hence $P(\{|B_n + C_n| \leqq 5\varepsilon\}) \geqq 1 - 2\varepsilon$ for all $n \geqq n_1$ and $H \in \mathscr{H}'$.

6. Proof of Theorem 2.2. It suffices to prove that $\sigma_n^2 \to \sigma_0^2$ as $n \to \infty$. For then we may ascertain that $\sigma_n^2 \geqq \sigma_0^2/2 > 0$ for $n \geqq n_0$ and all the conditions, necessary for the application of the part of Theorem 2.1 concerning the uniformity with $\mathscr{H}' = \{H_{(n_0)}, H_{(n_0+1)}, \cdots\}$, are covered by the conditions of Theorem 2.2. So we may conclude that the convergence $n^{\frac{1}{2}}(T_n - \mu(H)) \to_d \mathrm{N}(0, \sigma^2(H))$ is uniform on the above subclass $\mathscr{H}'$, and hence that $n^{\frac{1}{2}}(T_n - \mu_n)/\sigma_n \to_d \mathrm{N}(0, 1)$ as $n \to \infty$. But if $\sigma_n^2 \to \sigma_0^2$ the weak convergence of $\mathrm{N}(0, \sigma_n^2)$ to $\mathrm{N}(0, \sigma_0^2)$ follows, and thus we finally obtain $n^{\frac{1}{2}}(T_n - \mu_n) \to_d \mathrm{N}(0, \sigma_0^2)$ as $n \to \infty$.

As in Section 5 let us assume that $J_d = \delta_{s_1}$ and $K_d = \delta_{t_1}$ for fixed $s_1, t_1 \in (0, 1)$. For a function $\phi(F_{(n)}(x), G_{(n)}(y))$, integrable with respect to $H_{(n)}$, we have $\iint \phi(F_{(n)}(x), G_{(n)}(y))\, dH_{(n)}(x, y) = \iint \phi(s, t)\, d\bar{H}_{(n)}(s, t)$, where $\bar{H}_{(n)}(s, t) = H_{(n)} \times (F_{(n)}^{-1}(s), G_{(n)}^{-1}(t))$. Note that $\bar{H}_{(n)}$ has uniform (0, 1) marginal df's. Using the above transformation and writing the square of an integral as a repeated integral, we arrive at the following alternative expression for the variance (see (3.5))

$$\begin{aligned}\sigma_n^2 &= \iint \{J(s)K(t) - \iint J(u)K(v)\, d\bar{H}_{(n)}(u, v) \\ &\quad + \iint [\delta_s(u) - u]J'(u)K(v)\, d\bar{H}_{(n)}(u, v) + [\delta_{s_1}(s) - s_1]E_{H_{(n)}}(K \mid s_1) \\ &\quad + \iint [\delta_t(v) - v]J(u)K'(v)\, d\bar{H}_{(n)}(u, v) \\ &\quad + [\delta_{t_1}(t) - t_1]E_{H_{(n)}}(J \mid t_1)\}^2\, d\bar{H}_{(n)}(s, t) \\ &= \textstyle\sum_{i=1}^6 \sum_{j=1}^6 \iiiint\iint \phi_i(s, t, u, v)\phi_j(s, t, u', v') \\ &\quad \times d\bar{H}_{(n)}(u, v)\, d\bar{H}_{(n)}(u', v')\, d\bar{H}_{(n)}(s, t)\,,\end{aligned} \tag{6.1}$$

for $n = 0, 1, 2, \cdots$. Here, s, t, u, v, u', v' are restricted to $(0, 1)$ and

$$\begin{aligned}
&\phi_1(s, t, u, v) = J(s)K(t)\,, && |\phi_1| \leqq r_1(s)r_2(t)\,, \\
&\phi_2(s, t, u, v) = J(u)K(v)\,, && |\phi_2| \leqq r_1(u)r_2(v)\,, \\
&\phi_3(s, t, u, v) = [\delta_s(u) - u]J'(u)K(v)\,, && |\phi_3| \leqq M_1[q_1(s)]^{-1}q_1(u)\bar{r}_1(u)r_2(v)\,, \\
&\phi_4(s, t, u, v) = [\delta_{s_1}(s) - s_1]E_{H_{(n)}}(K \mid s_1)\,, && |\phi_4| \leqq \textstyle\int_0^1 r_2(t)g(t)\, dt\,, \\
&\phi_5(s, t, u, v) = [\delta_t(v) - v]J(u)K'(v)\,, && |\phi_5| \leqq M_2[q_2(t)]^{-1}q_2(v)r_1(u)\bar{r}_2(v)\,, \\
&\phi_6(s, t, u, v) = [\delta_{t_1}(t) - t_1]E_{H_{(n)}}(J \mid t_1)\,, && |\phi_6| \leqq \textstyle\int_0^1 r_1(s)f(s)\, ds\,.
\end{aligned}$$

The bounds for the absolute values of the ϕ_i follow from Assumptions 2.2, 2.5 and Lemma 3.2 (M_i depends on q_i only, $i = 1, 2$).

Let us first note that the convergence $H_{(n)}(x, y) \to H_{(0)}(x, y)$ for all x, y (see Assumption 2.5) entails the convergence $\bar{H}_{(n)}(u, v)\bar{H}_{(n)}(u', v')\bar{H}_{(n)}(s, t) \to \bar{H}_{(0)}(u, v)\bar{H}_{(0)}(u', v')\bar{H}_{(0)}(s, t)$ as $n \to \infty$ in all continuity points of the latter product of df's. A further application of Assumption 2.5 combined with the dominated convergence theorem yields

$$E_{H_{(n)}}(K \mid s_1) \to E_{H_{(0)}}(K \mid s_1)\,, \qquad E_{H_{(n)}}(J \mid t_1) \to E_{H_{(0)}}(J \mid t_1)\,, \qquad \text{as } n \to \infty\,. \tag{6.2}$$

Convergence of each of the summands in (6.1) suffices to prove the convergence of the variances. The functions ϕ_4 and ϕ_6, which actually depend on n through multiplicative constants, do not disturb the applicability of Billingsley [3] Theorem 5.4, since by (6.2) these multiplicative constants converge properly. It thus remains to show that for some $\zeta > 0$

$$\sup_{n=1,2\cdots} \iiiint\iint |\phi_i(s, t, u, v)\phi_j(s, t, u', v')|^{1+\zeta} \times d\bar{H}_{(n)}(u, v)\, d\bar{H}_{(n)}(u', v')\, d\bar{H}_{(n)}(s, t) < \infty \tag{6.3}$$

for $1 \leqq i \leqq j \leqq 6$. By the nature of the bounds for the $|\phi_i|$, the fact that we are dealing with a product measure, and the similarity between ϕ_3 and ϕ_5 it follows that we only have to verify (6.3) for $i = j = 1, 2, 3$.

From now on let us choose $\zeta = \varepsilon/2 > 0$ and let us first take $i = j = 1$. Since ϕ_1 is a function of s and t only, the supremum in (6.3) is bounded by

$$\sup_{n=1,2,\cdots} \iint [r_1(s)r_2(t)]^{2+2\zeta}\, d\bar{H}_{(n)}(s, t) < \infty\,,$$

by Assumption 2.3. The function ϕ_2 does not depend on s, t so that for $i = j = 2$ the supremum (6.3) is bounded by

$$\sup_{n=1,2,\cdots} [\iint [r_1(u)r_2(v)]^{1+\zeta}\, d\bar{H}_{(n)}(u, v)]^2 < \infty\,,$$

by Assumption 2.3. Finally for $i = j = 3$ we see that the supremum in (6.3) is bounded by

$$\sup_{n=1,2,\cdots} M_1^{2+2\zeta} \int\int\int\int\int\int [q_1(s)]^{-2-2\zeta}[q_1(u)\tilde{r}_1(u)r_2(v)]^{1+\zeta}$$
$$\times [q_1(u')\tilde{r}_1(u')r_2(v')]^{1+\zeta}\, d\bar{H}_{(n)}(u, v)\, d\bar{H}_{(n)}(u', v')\, d\bar{H}_{(n)}(s, t)$$
$$\leqq \sup_{n=1,2,\cdots} M_1^{2+2\zeta} \int_0^1 [q_1(s)]^{-2-2\zeta}\, ds$$
$$\times [\int\int [q_1(u)\tilde{r}_1(u)r_2(v)]^{1+\zeta}\, d\bar{H}_{(n)}(u, v)]^2 < \infty ,$$

again by Assumption 2.3. This concludes the proof of Theorem 2.2.

7. Application and extension. An application of Theorem 2.1 in the case where the score functions are simple step-functions lies in the treatment of ties. Let us suppose that the sample has been drawn from a df H which is no longer continuous but, on the contrary, is entirely concentrated on a finite lattice of points in the plane. As has been pointed out in Hájek [8], there are two possible techniques for adjusting the original rank statistic to this situation where necessarily ties will occur. The first technique is referred to as the method of randomizing the ranks, and the second as the method of averaging the scores. By the former technique, which represents a purely theoretical approach to the problem, asymptotic normality of the resulting rank statistic follows immediately from Theorem 2.1. When we restrict our attention to the null hypothesis, a slight generalization of this result concerning the asymptotic normality for randomized ranks may be used to obtain conditional asymptotic normality for averaged scores, given the marginal ties. Under the alternative, however, another approach is appropriate when averaged scores are used. More general results for the regression problem are given by Vorličková [18].

Finally it should be noted that the restriction to linear rank statistics for which the score functions factorize and can be written as a product $J(s)K(t)$ is inessential. No new difficulties will be encountered when developping the theory for more general score functions $J(s, t)$, as long as the functions that bound $J(s, t)$ and its first partial derivatives $\partial J(s, t)/\partial s$, $\partial J(s, t)/\partial t$ still factorize as products of functions of the arguments s and t separately.

Acknowledgments. I wish to thank G. R. Shorack for suggesting the problem of discontinuous score functions and local alternatives, and W. R. van Zwet for his kind and essential help throughout the preparation of this work.

REFERENCES

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

[2] Bhuchongkul, S. (1964). A class of nonparametric tests for independence in bivariate populations. *Ann. Math. Statist.* **35** 138–149.

[3] Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[4] Dupač, V. and Hájek, J. (1969). Asymptotic normality of simple linear rank statistics under alternatives II. *Ann. Math. Statist.* **40** 1992–2017.

[5] 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.

[6] GHOSH, J. K. (1971). A new proof of the Bahadur representation of quantiles and an application. *Ann. Math. Statist.* **42** 1957–1961.

[7] GUMBEL, E. J. (1958). Multivariate distributions with given margins and analytical examples. *Bull. Inst. Internat. Statist.* **37-3** 363–373.

[8] HÁJEK, J. (1969). *A Course in Nonparametric Statistics.* Holden-Day, San Francisco.

[9] HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[10] KIEFER, J. (1961). On large deviations of the empiric d.f. of vector chance variables and a law of the iterated logarithm. *Pacific J. Math.* **11** 649–660.

[11] PURI, M. L., SEN, P. K. and GOKHALE, D. V. (1970). On a class of rank order tests for independence in multivariate distributions. *Sankhyā Ser. A* **32** 271–298.

[12] PYKE, R. and SHORACK, G. R. (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorems. *Ann. Math. Statist.* **39** 755–771.

[13] RUYMGAART, F. H. (1973). Asymptotic theory of rank tests for independence. Mathematical Centre Tracts 43, Amsterdam.

[14] RUYMGAART, F. H., SHORACK, G. R. and VAN ZWET, W. R. (1972). Asymptotic normality of nonparametric tests for independence. *Ann. Math. Statist.* **43** 1122–1135.

[15] SEN, P. K. (1968). Asymptotic normality of sample quantiles for *m*-dependent processes. *Ann. Math. Statist.* **39** 1724–1730.

[16] SEN, P. K. (1970). Asymptotic distribution of a class of multivariate rank order statistics. *Calcutta Statist. Assoc. Bull.* **19** 23–31.

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

[18] VORLIČKOVÁ, D. (1970). Asymptotic properties of rank tests under discrete distributions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **14** 275–289.

STICHTING MATHEMATISCH CENTRUM
2E BOERHAAVESTRAAT 49
AMSTERDAM (OOST), NETHERLANDS