AFDELING MATHEMATISCHE STATISTIEK

F.H. RUYMGAART, G.R. SHORACK and W.R. VAN ZWET
ASYMPTOTIC NORMALITY OF NONPARAMETRIC TESTS
FOR INDEPENDENCE

Prepublication
Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.
ASYMPTOTIC NORMALITY OF NONPARAMETRIC

TESTS FOR INDEPENDENCE

by

F.H. Ruymgaart, G.R. Shorack \(^1\) and W.R. van Zwet

Summary. 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.

\(^1\) This research was partially supported by NSF contract, number GP-13739.
1. Introduction. For each $n$ let $(X_1, Y_1), \ldots, (X_n, Y_n)$ 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 $i$-th 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 $(\Omega, \mathcal{A}, P)$.

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 $i = 1, \ldots, n$ (see Hájek and Šidák [6]). A suitably standardized version of $T_n$ will be (see also Bhuchongkul [1])

$$n^{1/2}(T_n - \mu) = n^{1/2} \left[ \int J_n(F_n) K_n(G_n) dH_n - \mu \right];$$

(1.1)

here

$$J_n(s) = a_n(i), K_n(s) = b_n(i),$$

(1.2)

for $(i-1)/n < s < i/n$ and $i = 1, \ldots, n$, and
\begin{equation}
\mu = \int \int J(F) K(G) dH,
\end{equation}

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

\begin{equation}
r = [I(1-I)]^{-1} \quad \text{on} \quad (0,1),
\end{equation}

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 \( r^{1/2 - \delta} \) for some \( \delta > 0 \) (see Hájek and Šidák [6]). 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 \( r^{1/4 - \delta} \).

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 \( r^{1/4 - \delta} \) in general and \( r^{1/2 - \delta} \) 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 \( a = (1/2-\delta)/p_0 \) and \( b = (1/2-\delta)/q_0 \) for some \( 0 < \delta < 1/2 \) 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

\[(\cdot.5) \quad \int \int |\phi(F)\psi(G)| dH \leq \left( \int |\phi|^p dI \right)^{1/p} \left( \int |\psi|^q dI \right)^{1/q}, \]

where \( \phi \) and \( \psi \) 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 \( a = b = 1/2 - \delta \) for some \( 0 < \delta < 1/2 \). The price for this is a condition on the df \( H \), keeping it in some sense similar to the null hypothesis. This condition is

\[(1.6) \quad dH \leq C[r(F) r(G)]^{\delta/2} dFdG, \]

with fixed constants \( C \geq 1 \) and \( 0 < \delta < 1/2 \). 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 \(-1 \leq \alpha \leq 1\) 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

\[ n^{1/2}(T_n - \mu) \xrightarrow{\text{d}} N(0, \sigma^2) \quad \text{as} \quad n \to \infty, \]

of (1.1); here \(\mu\) and \(\sigma^2\) are finite and are given by (1.3) and (3.10)
respectively.

Let \(H\) denote the class of all continuous bivariate dfs \(H\), and let
\(H_{\delta}^{c}\) denote the subclass that satisfies (1.6) for fixed \(C \geq 1\) and
\(0 < \delta < 1/2\).

To prove (2.1) for general \(H\) in \(H\) 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 \(|J_n| \leq Dr^a, |K_n| \leq Dr^b\)
and

\[ |J(i)| \leq Dr^{a+i}, \quad |K(i)| \leq Dr^{b+i} \quad \text{for} \quad i = c, 1, \]
where defined on (0,1). Here $D$ is a positive constant and $a$ and $b$ satisfy

$$a = (1/2-\delta)/p_0, \quad b = (1/2-\delta)/q_0$$

for some $0 < \delta < 1/2$ and some $p_0, q_0 > 1$ with $p_0^{-1} + q_0^{-1} = 1$.

In proving (2.1) for the more restrictive class $H_{0\delta}$ we only require a weak boundedness condition on the score functions.

**Assumption 2.2.** Assumption 2.1 holds with

$$a = b = 1/2-\delta$$

for some $0 < \delta < 1/2$.

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

$$B_{0n} = n^{1/2} \int_{\Delta_n} \left[ J_n(F_nK_n) - J(F_nK(G_n)) \right] dH_n,$$

$$B^*_{0n} = n^{1/2} \int_{\Delta_n} \left[ J_n(F_nK(G_n)) - J(F^*_nK(G^*_n)) \right] dH_n,$$

where
\( \Delta_n = \Delta_n \times \Delta_n \) with \( \Delta_n = [X_{1n}, X_{nn}] \) and \( \Delta_n = [Y_{1n}, Y_{nn}] \),

\[
F_n^* = \left[ n/(n+1) \right] F_n, \quad G_n^* = \left[ n/(n+1) \right] G_n.
\]

**Assumption 2.3.** Either (a) \( B_{0n} \to 0 \) as \( n \to \infty \), or (b) \( B_{0n}^* \to 0 \) 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 \( 1 \leq i \leq n \) for some functions \( J \) and \( K \), then Assumption 2.3 (b) holds uniformly for \( H \) in \( H \). (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 \( |J'(i)| \leq a^{i-s} \) and \( |K'(i)| \leq b^{i+1} \) 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 \( i \)-th 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 \( H \). (This statement generalizes Theorem 2 of [1] and the proof may be given in the same way. It relies mainly on the fact that \( \sum_{i=1}^{n-1} |a_n(i) - J(i/n)| = O(n^s) \) 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 \( a = a \) and \( a = b \) respectively.)
THEOREM 2.1. If $H$ is in $H$ and if Assumptions 2.1 and 2.3 are satisfied, then the asymptotic normality (2.1) holds. Given any subclass $H'$ of $H$ such that Assumption 2.3 holds uniformly for $H$ in $H'$ and such that $\sigma^2 = \sigma^2(H)$ is bounded away from 0 on $H'$, the convergence in (2.1) is uniform for $H$ in $H'$.

Note that (2.2) is satisfied if $a = b = 1/4 - \varepsilon$ for some $0 < \varepsilon < 1/4$ (take $p_0 = q_0 = 2$ and $\delta = 2\varepsilon$). Thus Theorem 2.1 allows a rate of growth $r^{1/4 - \varepsilon}$ for the score functions $J$ and $K$ and $r^{5/4 - \varepsilon}$ for their derivatives. In Theorem 1 of [1] these rates are $r^{1/8 - \varepsilon}$ 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 $c > 1$ and $0 < \delta < 1/2$. If $H$ is in $H_{C6}$ and if Assumptions 2.2 and 2.3 are satisfied, then the asymptotic normality (2.1) holds. Given any subclass $H'_{C6}$ of $H_{C6}$ such that Assumption 2.3 holds uniformly for $H$ in $H'_{C6}$ and such that $\sigma^2 = \sigma^2(H)$ is bounded away from 0 on $H'_{C6}$, the convergence in (2.1) is uniform for $H$ in $H'_{C6}$. 
3. Proof of the theorems: Asymptotic normality of the leading terms. Let $F^{-1}(s) = \inf \{x : F(x) > s\}$ and $G^{-1}(t) = \inf \{y : G(y) > t\}$; these definitions imply $F(F^{-1}) = G(G^{-1}) = 1$. 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 $F(X_{1}), \ldots, F(X_{n})$ and $G(Y_{1}), \ldots, G(Y_{n})$ respectively. Define the empirical processes $U_{n} = n^{1/2}[F_{n}(F^{-1}) - I]$ and $V_{n} = n^{1/2}[G_{n}(G^{-1}) - I]$ on $[0,1]$. With probability 1 these processes satisfy $U_{n}(F) = n^{1/2}(F_{n} - F)$ and $V_{n}(G) = n^{1/2}(G_{n} - G)$ on $(-\infty, \infty)$. All of the above remarks follow from the fact that

\begin{equation}
(3.1) \quad P(\Omega_{0}) = P((\omega: F_{n}(F^{-1}(F)) = F_{n}, G_{n}(G^{-1}(G)) = G_{n} \text{ for all } x, y \text{ and } n)) = 1.
\end{equation}

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 $\gamma$ define the sets

\begin{align}
S_{\gamma_{1}} &= [F^{-1}(\gamma), F^{-1}(s_{1} - \gamma)] \cup [F^{-1}(s_{1}, F^{-1}(1 - \gamma)), \\
S_{\gamma_{2}} &= [G^{-1}(\gamma), G^{-1}(t_{1} - \gamma)] \cup [G^{-1}(t_{1}, G^{-1}(1 - \gamma)), \\
(\gamma_{3}) \quad \Omega_{\gamma_{n}} &= \{\omega: \sup|F_{n} - F| < \gamma/2, \text{ sup}|G_{n} - G| < \gamma/2\}.
\end{align}

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 $\omega$ in $\Omega_{0} \cap \Omega_{\gamma_{n}}$ the mean value theorem gives
\[ n^{1/2} j(F_n) = n^{1/2} j(F) + u_n[F] j'(\phi_n) \]

for all \( x \) in \( \Delta_n \cap S_{r_1} \). In the above formula the function \( \phi_n \) 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))

\[ n^{1/2} (T_n - \mu) = \sum_{i=1}^{3} A_{in} + \sum_{i=0}^{2} B_{in} + \sum_{i=3}^{7} B_{yn} + B_{8n} + C_n, \]

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

\[ A_{in} = n^{1/2} \int_{H_n - H} J(F)K(G)dH, \]

\[ A_{2n} = \int U_n(F)j'(F)K(G)dH, \]

\[ A_{3n} = \int V_n(G)j(F)K'(G)dH, \]

\[ B_{in} = n^{1/2} \int_{\Delta_n} J_n(F_n)K(G_n)dH_n, \]

\[ B_{2n} = -n^{1/2} \int_{\Delta_n} J(F)K(G)dH_n, \]

\[ B_{3n} = \chi(\Omega_{yn}) \left( n^{1/2} \int_{\Delta_n} [J(F_n) - J(F)]K(G)dH_n - A_{2n} \right), \]

\[ B_{4n} = \chi(\Omega_{yn}) n^{1/2} \int_{\Delta_n \cap S_{r_1}} [J(F_n) - J(F)]K(G)dH_n, \]

\[ B_{5n} = \chi(\Omega_{yn}) \int_{\Delta_n \cap S_{r_1}} U_n(F)[J'(\phi_n) - J'(F)]K(G)dH_n, \]

\[ B_{6n} = \chi(\Omega_{yn}) \int_{\Delta_n \cap S_{r_1}} U_n(F)j'(F)K(G)dH_n, \]

\[ B_{7n} = -\chi(\Omega_{yn}) \int_{\Delta_n \cap S_{r_1}} U_n(F)j'(F)K(G)dH, \]

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

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

Let us note that...
\[ \sum_{i=3}^{7} B_{\gamma n} = n^{1/2} \int_{H_0} \int_{A_n} [J(F_n) - J(F)]K(G)dH_n - A_{2n}, \]

which is symmetric to \( B_{\gamma n} \). For this reason \( B_{\gamma n} \) 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 \( \varepsilon \) we can find numbers \( p_i, q_i > 1 \) satisfying \( p_i^{-1} + q_i^{-1} = 1 \) and

\[ (a+1/2+\delta/2)p_1 < 1, \quad b_q_1 < 1, \quad a_p_2 < 1, \quad (b-1/2+\delta/2)q_2 < 1. \]

As to the first pair of inequalities, we have \( a+1/2+\delta/2+b = 1-\delta/2 \) and consequently \( a+1/2+\delta/2 < 1-\delta/2 \) (the numbers \( a \) and \( b \) are strictly positive). Now choose \( p_1 = (a+1/2+3\delta/4)^{-1} \) and let \( q_1 = (1-p_1^{-1})^{-1}. \)

Then \( (a+1/2+\delta/2)p_1 < 1 \) and

\[ b_q_1 = (1/2-a-\delta)/(1/2-a-3\delta/4) < 1. \]

The second pair of inequalities can be obtained in the same way.

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

\[ A_{1n} = n^{-1/2} \sum_{i=1}^{n} A_{1in}, \]

where \( A_{1in} = J(F_i X_i)K(G_i Y_i) - \mu \) are independent and identically distributed (iid) 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 \( \mathcal{H}_{C^d} \) as may be seen by applying (1.6). Moreover this moment will be uniformly bounded above for \( H \) within \( \mathcal{H}(H_{C^d}) \).

Because

\[
U_n(F) = n^{-1/2} \sum_{i=1}^{n} (\phi_{X_i} - F),
\]

where

\[
(3.7) \quad \phi_{X_i}(x) = 0 \text{ if } x < X_i \text{ and } \phi_{X_i}(x) = 1 \text{ if } x \geq X_i,
\]

we have

\[
(3.8) \quad A_{2i} = n^{-1/2} \sum_{i=1}^{\ell} A_{2in},
\]

where the \( A_{2in} = \int (\phi_{X_i} - F) \nu_i(F)(G) d\mathcal{H} \) are iid with mean zero. Under Assumptions 2.1 or 2.2 we have

\[
|A_{2in}| \leq D^2 r^{1/2-\delta/4}(F_{X_i}) \int r^{a+1/2+\delta/4}(F) r^b G d\mathcal{H}.
\]

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 \( \mathcal{H} \).

Under Assumption 2.1 the non random integral is seen to be uniformly bounded above for \( H \) in \( \mathcal{H} \) 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 \( H_{C^d} \) holds under Assumption 2.2, as may be shown by application of (1.6).
Analogously we can write

\[(3.9) \quad A_{3n} = n^{-1/2} \sum_{i=1}^{n} A_{3in},\]

where $A_{3in} = \iint (\phi(x_i - G)J(F)K'(G) d\mathcal{H}$ are iid 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 $H(H_{C, \delta})$. 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 $\sum_{i=1}^{n} A_{3i} \to_d N(0, \sigma^2)$ as $n \to \infty$. The variance $\sigma^2$ is given by (see [1])

\[(3.10) \quad \sigma^2 = \text{Var}[J(F(X))K(G(Y))] + \iint (\phi(x - F)J'(F)K(G) d\mathcal{H})\]

\[+ \iint (\phi(y - G)J(F)K'(G) d\mathcal{H}),\]

with $\phi$ defined in (3.7).

Since we have shown that an absolute moment of order larger than 2 exists and is uniformly bounded on $H(H_{C, \delta})$, and because the variance is uniformly bounded away from zero on $H'(H_{C, \delta})$, the established convergence in distribution is uniform for $H$ in $H'(H_{C, \delta})$ by Esseen'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 $\xi > 0$ the function $r^\xi$ is symmetric about $1/2$, decreasing on $(0,1/2]$ and has the property that for each $\beta$ in $(0,1)$ there exists a constant $M = M_\beta$ such that $r^\xi(\beta s) \leq M r^\xi(s)$ for $0 < s \leq 1/2$ and $r^\xi(1-\beta(1-s)) \leq M r^\xi(s)$ for $1/2 < s < 1$.

**Proof.** On $(0,1/2]$ we have $r^\xi(\beta s) = (\beta s)^{-\xi}(1-\beta s)^{-\xi} \leq \beta^{-\xi}r^\xi(s)$.

A similar argument applies to the interval $(1/2,1)$. □

**Lemma 4.2.** For each $\omega$ let $\hat{\phi}_n = \hat{\phi}_{n\omega}$ and $\hat{\psi}_n = \hat{\psi}_{n\omega}$ be functions on $A_{n1} = A_{n1\omega}$ and $A_{n2} = A_{n2\omega}$ respectively (see (2.6)), satisfying

$$\min(F_n, F_n) \leq \hat{\phi}_n \leq \max(F_n, F_n)$$

and

$$\min(G_n, G_n) \leq \hat{\psi}_n \leq \max(G_n, G_n)$$

where defined. Then uniformly for $n = 1, 2, \ldots$ and $H \in \mathcal{H}$:

(i) $\sup_{A_{n1}} r^{\xi}(\hat{\phi}_n) r^{-\xi}(F) = O_p(1)$ for each $\xi > 0$;

(ii) $\sup_{A_{n2}} r^{\eta}(\hat{\psi}_n) r^{-\eta}(G) = O_p(1)$ for each $\eta > 0$;

(iii) $\sup_{(-\infty, \infty)} |U_n(F)| r^{1/2-\tau}(F) = O_p(1)$ for each $\tau > 0$.

**Proof.** (i) From formula (3.1) and e.g. from [11], Lemma 4.3 it follows that for each $\epsilon > 0$ there exists a constant $\beta = \beta_\epsilon$ in $(0,1)$ such that

$$P(\Omega_n) = P((\beta F < F_n < 1-\beta(1-F) \text{ on } A_{n})) > 1-\epsilon,$$
for all \( n \) and uniformly in all continuous \( F \). Because of the definition of \( \Phi_n \) we have \( \beta F \leq \Phi_n \leq 1 - \beta (1 - F) \) on \( \Delta_n \). By Lemma 4.1 this implies that for some constant \( M_{\xi \in C} \) we have \( r_{\xi}(\Phi_n) \leq M_{\xi \in C} r_{\xi}(F) \) for \( x \) in \( \Delta_n \) on the set \( \Omega_n \).

(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

\[
I_k(0) = 0, \quad I_k(s) = (i-1)/k \text{ for } (i-1)/k < s \leq i/k, \quad i = 1, \ldots, k.
\]

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

**Proof.** Note that \( \sup_{-\infty < x < \infty} |U_n(I_k(F)) - U_n(F)| = \sup_{0 \leq s \leq 1} |U_n(I_k) - U_n| \), 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 \( \widehat{U}_n \)- and \( \widehat{U}_0 \)-processes defined on a single new probability space \((\bar{\Omega}, \bar{A}, \bar{P})\) and having the same finite dimensional distributions as the original processes (see also Skorokhod [12]). These new processes satisfy \( \sup_{-\infty < x < \infty} |\widehat{U}_n(I_k(F)) - \widehat{U}_n(F)| \to 0 \) uniformly in \( k \), as \( n \to \infty \). Now \( \sup_{-\infty < x < \infty} |\widehat{U}_n(I_k) - \widehat{U}_n| \leq \sup_{0 \leq s \leq 1} |\widehat{U}_n - \widehat{U}_0| + \sup_{0 \leq s \leq 1} |\widehat{U}_0 - \widehat{U}_k| + \sup_{C} |\widehat{U}_C(I_k) - \widehat{U}_n(I_k)| \). For almost every \( \omega \) the function \( \widehat{U}_0 \) is uniformly continuous on \([0,1]\) so that \( \sup_{0 \leq s \leq 1} |\widehat{U}_0 - \widehat{U}_k| \to 0 \) as \( k \to \infty \). This proves that \( \sup_{-\infty < x < \infty} |\widehat{U}_n(I_k(F)) - \widehat{U}_n(F)| \to 0 \) for \( k, n \to \infty \). This last result implies the convergence in probability of the lemma. \( \square \)
Let \( \nu \) and \( \lambda \) be the random indices \( 1 \leq \nu(\omega), \lambda(\omega) \leq n \) such that

\[
(4.3) \quad X_\nu = X_{nn} \quad \text{and} \quad Y_\lambda = Y_{nn}.
\]

**Lemma 4.4.** As \( n \to \infty \), \( P(\{\alpha_n \leq F(X_\nu) \leq 1-\alpha_n\} \cap \{\alpha_n \leq G(Y_\nu) \leq 1-\alpha_n\}) \to 1 \)
uniformly for \( H \) in \( \mathcal{H} \) 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 \( \mathcal{H} \). \( \square \)

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 \( \mathcal{H}_C \).

**Proof.** \( P(\{Y_\nu = Y_{nn}\}) = P(\bigcup_{i=1}^{n} \{(X_i, Y_i) = (X_{nn}, Y_{nn})\}) = n \int_{\mathcal{H}} n^{H-1} dH. \)

Note that for all \( x, y \) we have \( H(x, y) \leq F(x) \) and \( H(x, y) \leq G(y) \). Letting \( n_0 = (n-1)/2 \) and applying (1.6) we obtain

\[
\int_{\mathcal{H}} n^{H-1}(x, y) dH(x, y) \leq \int_{\mathcal{H}} n_0^{n_0}(x) G^{n_0}(y) dH(x, y)
\]

\[
\leq C n \int_{0}^{1} r^{n_0-\delta/2} dI^2
\]

\[
= C n \int_{0}^{1} \Gamma(n_0-\delta/2+1) \Gamma(1-\delta/2) / \Gamma(n_0-\delta+2) dI^2
\]

\[
\leq C_1 n n_0^{\delta-2} = o(n^{\delta-1}) \to 0
\]

as \( n \to \infty \), because \( 0 < \delta < 1/2 \). Here \( C_1 \) is a constant depending on \( C \)
and \( \delta \) only; hence the convergence is uniform for \( H \) in \( \mathcal{H}_C \). \( \square \)
**Lemma 4.6.** As \( n \to \infty \), \( P(\{ \gamma_n \leq G(Y_n) \leq 1 - \gamma_n \}) \to 1 \) uniformly for \( H \) in \( H_{C^\delta} \) provided \( \gamma_n \leq an^{-\delta} \) for some positive constant \( a \).

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

\[
P(\{G(Y_n) < \gamma_n\}) = nP(\bigcap_{i=1}^{n-1}\{F(X_i) < F(X_n)\}) \cap \{G(Y_n) < \gamma_n\})
\]

\[
= n \int_{-\infty}^{\gamma_n} \int_{-\infty}^{\gamma_n} G^{-1}(\gamma_n) p_{n-1}(x_n) dH(x_n, y_n)
\]

\[
\leq C_n \left[ \int_{0}^{1} r^{n-1} \frac{dr}{r^\delta/2} \right] \left[ \int_{C} r^{\delta/2} dr \right]
\]

\[
= C_n \left[ \Gamma(n-\delta/2) / \Gamma(n+1-\delta) \right] n^{-\delta+\delta^2/2}
\]

\[
\leq C_2 n^{-1+\delta/2} n^{-\delta+\delta^2/2} = C_2 n^{-\delta/2+\delta^2/2} \to 0
\]

as \( n \to \infty \), because \(-\delta/2+\delta^2/2 < 0\) for \( 0 < \delta < 1/2 \). Here \( C_1 \) and \( C_2 \) are constants depending on \( C, \delta \) and \( a \) only; hence the convergence is uniform for \( H \) in \( H_{C^\delta} \). \( \square \)
5. Proof of the theorems: Asymptotic negligibility of the remainder terms under Assumption 2.3(a). Let us start with a further decomposition of $\mathcal{C}_n$, which can be seen to be the sum of

$$C_{\gamma 1n} = \chi(\Omega_{\gamma n}) n^{1/2} \int_{\Delta_n} [J(F_n) - J(F)] [K(G_n) - K(G)] dH_n,$$

$$C_{\gamma 2n} = \chi(\Omega_{\gamma n}) n^{1/2} \int_{\Delta_n \cap S_\gamma} [J(F_n) - J(F)] K(G_n) dH_n,$$

$$C_{\gamma 3n} = -\chi(\Omega_{\gamma n}) n^{1/2} \int_{\Delta_n \cap S_\gamma} [J(F_n) - J(F)] K(G) dH_n,$$

$$C_{\gamma 4n} = \chi(\Omega_{\gamma n}) \int_{\Delta_n \cap S_\gamma} U_n(F) J'(F_n) [K(G_n) - K(G)] dH_n.$$

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 $n \to \infty$, $B_{1n} \to \mathcal{C}$ uniformly for $H$ in $H(H_{\mathcal{C}})$.

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

$$B_{11n} = n^{1/2} J_n(1) \int_{(X_n, Y_n) \times \Delta_{n2}} |K_n(G_n(y))| dH_n(x, y),$$

$$B_{12n} = n^{1/2} J_n(1) K_n(1) \int_{(X_{nn}, Y_{nn})} dH_n(x, y),$$

$$B_{13n} = n^{1/2} K_n(1) \int_{\Delta_n \times (Y_{nn})} J_n(F_n(x)) dH_n(x, y).$$

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


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

$$\chi(\Omega_{2n}B_{11n} \leq DMn^{-1/2} |\gamma_n(1)| r^b(G(Y_n))),$$

where $\gamma$ is defined by (4.3). Set $\gamma_n = n^{-1/2} |\gamma_n(1)|$ and note that by (1.2) and Assumption 2.2 we have $\gamma_n \leq Dn^{-6}$ for some constant $D > 0$. Let $\Omega_{2n} = \{r_n \leq G(Y_n) \leq 1 - \gamma_n\}$. Then

$$\chi(\cap_{i=1}^{n^2} \Omega_{2n}B_{11n} \leq DMin^{-1/2}(1-\gamma_n)^{-b} = 0(n^{-\delta(1-b)} \to 0$$

as $n \to \infty$. Applying Lemma 4.6 we see that $P(\cap_{i=1}^{n^2} \Omega_{2n}) > 1-2\varepsilon$ for $n$ large enough, uniformly for $H$ in $H_{C_0}$. A symmetric argument can be given for $B_{3n}$.

For the rv $B_{12n}$ use Lemma 4.5 to see that the set on which this rv may assume a non-zero value has probability converging to zero as $n \to \infty$, uniformly for $H$ in $H_{C_0}$.

**Corollary 5.2.** As $n \to \infty$, $B_{2n} \to 0$ uniformly for $H$ in $H(H_{C_0})$.

**Proof.** The rv $B_{2n}$ is bounded by $\cap_{i=1}^{n^2} E_{2in}$ where
\[ B_{21n} = D^{2n-1/2} r^a(F(X_n)) r^b(G(Y_n)), \]
\[ B_{22n} = D^{2n-1/2} r^a(F(X_n)) r^b(G(Y_n)), \]

with \( \nu \) and \( \lambda \) defined by (4.3).

Under the assumptions of Theorem 2.1 consider
\[ \Omega_{1n} = \{a_n < F(X_n) < 1-a_n\} \cap \{a_n < G(Y_n) < 1-a_n\}, \]
with \( a_n = n^{a+b-3/2} \).

Note that \( n a_n \to 0 \). Then
\[ \chi(\Omega_{1n}) B_{21n} \leq D^{2n-1/2} \alpha_n^{-a-b} (1-\alpha_n)^{-a-b} \]
\[ = D^{2} (n \alpha_n)^{-1-a-b} (1-\alpha_n)^{-a-b} \to 0 \]
as \( n \to \infty \). Lemma 4.4 gives that \( P(\Omega_{1n}) \to 1 \) as \( n \to \infty \), uniformly for \( H \) in \( H \). The same argument applies for the rv \( B_{22n} \).

Under the assumptions of Theorem 2.2 consider
\[ \Omega_{2n} = \{b_n < F(X_n) < 1-b_n\} \cap \{\gamma_n < G(Y_n) < 1-\gamma_n\}, \]
with \( b_n = (n \log n)^{-1} \)
and \( \gamma_n = n^{-\delta} \). Then by (2,3)
\[ \chi(\Omega_{2n}) B_{21n} \leq D^{2n-1/2} \beta_n^{-a} \gamma_n^{-b} (1-\beta_n)^{-a} (1-\gamma_n)^{-b} \to 0 \]
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 \( H_{C,\delta} \). The rv \( B_{22n} \) can be treated in the same way.

**COROLLARY 5.3.** For fixed \( \gamma \), \( B_{\gamma n} \to 0 \) and \( C_{\gamma n} \to 0 \) as \( n \to \infty \), uniformly for \( H \) in \( H \).
PROOF. \( P(\omega^C_{\gamma n}) \to 0 \) uniformly for \( \omega \in H \) by the Glivenko-Cantelli theorem and because the distribution of \( \sup |F_n - F| \) does not depend on \( H \) in \( H \). \( \square \)

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

**PROOF.** According to Lemma 4.2(iii) with \( \tau = 1/2 \), for given \( \epsilon > 0 \) there exists a constant \( M \) such that \( \Omega_n = \{ \sup U_n(F) < M \} \) has probability larger than \( 1 - \epsilon \) for all \( n \) and \( H \) in \( H \). Also

\[
\chi(\Omega_n) |B_{\gamma 5n}| \leq M \sup_{\Delta_n \cap S_{\gamma 1}} |J'(\phi_n) - J'(F)| \sup_{S_{\gamma 2}} |K(G)|.
\]

The function \( K(G) \) is bounded on \( S_{\gamma 2} \) and the bound does not depend on \( H \) in \( H \). The function \( J' \) is uniformly continuous on \([\gamma/2, \gamma/2] \cup [s_{\gamma} + \gamma/2, 1 - \gamma/2] \). Since \( |\phi_n - F| \leq |F_n - F| \) where \( \phi_n \) is defined, the Glivenko-Cantelli theorem yields \( \sup_{\Delta_n \cap S_{\gamma 1}} |J'(\phi_n) - J'(F)| \to 0 \) uniformly for \( H \) in \( H \). A similar argument may be used for \( C_{\gamma 4n} \). \( \square \)

**COROLLARY 5.5.** For fixed \( \gamma \), \( B_{\gamma 6n} \to 0 \) as \( n \to \infty \), uniformly for \( H \) in \( H \).

**PROOF.** For arbitrary \( k \) we have (see (4.2))

\[
|B_{\gamma 6n}| \leq \sum_{i=1}^{3} B_{\gamma 6ikn}
\]

where
\[ B_{y61kn} = \int_{\Delta_n \cap S_y} |U_n(F)J'(F)K(G) - U_n(I_k(F))J'(I_k(F))K(I_k(G))| dH, \]
\[ B_{y62kn} = \int_{\Delta_n \cap S_y} |U_n(I_k(F))J'(I_k(F))K(I_k(G))d(H_n - H)|, \]
\[ B_{y63kn} = \int_{\Delta_n \cap S_y} |U_n(F)J'(F)K(G) - U_n(I_k(F))J'(I_k(F))K(I_k(G))| dH. \]

Let us first consider \( B_{y61kn} \) and \( B_{y63kn} \), which are both bounded by the supremum of the integrand over the set \( S_y \). 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 \( \Omega_{kn} = \{ \sup |U_n(F)| - U_n(I_k(F))| \leq \eta_{kn} \} \) has probability larger than \( 1-\varepsilon \) for all \( k,n \) and all \( H \) in \( H \). Note that on \( ([\gamma,s_1,1-\gamma]) \cup ([\gamma,s_1+\gamma,1-\gamma]) \times ([\gamma,t_1,1-\gamma]) \cup [t_1+\gamma,1-\gamma]) \) the function \( J'(s)K(t) \) is bounded, say by a constant \( M_\gamma \), and uniformly continuous. By Lemma 4.2(iii) with \( \tau = 1/2 \), there exists a constant \( M \) such that \( \Omega_n = \{ \sup |U_n(F)| \leq M \} \) has probability larger than \( 1-\varepsilon \). Let us finally write \( \zeta_{ky} = \max_{S_y} |J'(F)K(G) - J'(I_k(F))K(I_k(G))| \), which tends to zero as \( k \to \infty \), uniformly for \( H \) in \( H \). Hence for \( i = 1,3 \)

\[ \chi(\Omega_{kn} \cap \Omega_n)B_{y61kn} \leq \eta_{kn} M_\gamma + M_{\zeta_{ky}} \to 0 \]
as \( k,n \to \infty \) for fixed \( \gamma \). Because \( P(\Omega_{kn} \cap \Omega_n) > 1-2\varepsilon \) uniformly for \( H \) in \( H \), we may conclude that \( B_{y61kn} \to 0 \) and \( B_{y63kn} \to 0 \) uniformly for \( H \) in \( H \), as \( k,n \to \infty \).

Let us next consider \( B_{y62kn} \) for a fixed value \( k \). For each \( \omega \) in \( \Omega_n \) the integrand in the expression for this rv is a simple step function assuming a value \( a_{ijkn}(\omega) \) on the rectangle
\[ R_{ijn} = (F^{-1}(i/(i-1)/k), F^{-1}(i/k), G^{-1}((j-1)/k), G^{-1}(j/k)) \cap S \gamma \cap \Delta_n, \]

for \( i = 1, \ldots, k \) and \( j = 1, \ldots, k \). Because \( |a_{ijn}^k| \leq M(\gamma + \gamma \gamma) \) on \( \Omega_n \), we have

\[
\chi(\Omega_n)_{\gamma 2\gamma n} = \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ijn} \int_{R_{ijn}} d(H - H) \leq h^2 M(\gamma + \gamma \gamma) \sup |H_n - H| \to 0
\]

as \( n \to \infty \), uniformly for \( H \) in \( H \). Here Theorem 1-m of Kiefer [8] is used. The conclusion of the corollary follows by straightforward combination of these results. \( \Box \)

**COROLLARY 5.6.** As \( \gamma \to 0 \) and \( n \to \infty \), \( B_{\gamma n} \to 0 \) and \( C_{\gamma 2n} \to 0 \), uniformly for \( H \) in \( H(H_{C, \delta}) \).

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

\[
(5.1) \chi(\Omega_{1n})_{B_{\gamma n}} \leq d^2 M_1 \int_{\Delta_n \cup S \gamma} r^{a+1/2 - \delta/4}(F) \, r^p(G, dH).
\]
Next consider $C_{\gamma_{2n}}$. By Assumption 2.1 (Assumption 2.2) we have $|F(G_n)| \leq D \gamma(G_n)$ on $\Delta_{n2}$ and application of Lemma 4.2(ii) with $\psi = G_n$ and $n = b$ gives the existence of a constant $M_2$ such that $\Omega_{2n} = \{r^b(G_n) \leq M_2 \gamma(G) \text{ on } \Delta_{n2}\}$ has probability larger than $1-\epsilon$ for all $n$ and $H$ in $H$. Take an arbitrary $\omega$ in $\Omega$ and let us first consider those values of $x$ in $\Delta_{n1}$ 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 $n^{1/2}|J(F_n) - J(F)| = |U_n(F)J'(\phi_{on})| \leq D|U_n(F)|r^{a+1}(\phi_{on})$. For those values of $x$ in $\Delta_{n1}$ 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 $n^{1/2}|J(F_n) - J(F)| \leq |U_n(F)|\sum_{i=1}^2|J'(\phi_{in})| \leq D|U_n(F)|r_{i=1}^{a+1}(\phi_{in})$ by Assumption 2.1 (Assumption 2.2). Where defined on $\Delta_{n1}$, both $\phi_{0n}$ and $\phi_{1n}$, $\phi_{2n}$ lie between $F$ and $F_n$. By Lemma 4.2(i), taking $\xi = a+1$, there exists a constant $M_3$ such that $\Omega_{2n} = \{\max_{i=0,1,2} r^{a+1}(\phi_{in}) \leq M_3 \gamma^{a+1}(F) \text{ where defined on } \Delta_{n1}\}$ has probability larger than $1-\epsilon$ for all $n$ and $H$ in $H$. Combining these results we have

\begin{equation}
E(x(\gamma_{2n})|C_{\gamma_{2n}}|) \leq 2D^3 M_2 M_3 \int_{\gamma_{2n}} r^{a+1/2+\delta/4}(F) r^b(G) dH.
\end{equation}
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 \( \gamma \to 0 \) and \( n \to \infty \), uniformly for \( H \) in \( H_{Co}^\gamma \). For this purpose we start with the integral

\[
\int_S \int \mathcal{C}^{\alpha+1/2+\delta/4}(F) r^b(G) dH,
\]

and note that \( S_\gamma = (S_\gamma^1 \times (-\infty, \infty)) \cup ((-\infty, \infty) \times S_\gamma^2) \). 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 \( H \) by

\[
\frac{(\alpha+1/2+\delta/4)}{p_1} \int \left[ \int \int \mathcal{C}^{\alpha+1/2+\delta/4}(F) r^b(G) dH \right]^{1/p_1} \left[ \int \int \mathcal{C}^{\alpha+1/2+\delta/4}(F) r^b(G) dH \right]^{1/q_1}.
\]

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 \( \gamma \to 0 \). Under Assumption 2.2 and for \( H \) in \( H_{Co}^\gamma \), by an application of (1.6) we see that (5.3) is bounded uniformly for \( H \) in \( H_{Co}^\gamma \) by

\[
C \left[ \int \int \mathcal{C}^{\alpha+1/2+\delta/4}(F) r^b(G) dH \right]^{1/p_1} \left[ \int \int \mathcal{C}^{\alpha+1/2+\delta/4}(F) r^b(G) dH \right]^{1/q_1}.
\]

(5.5)

\[
+ C \left[ \int \int r^{1-\delta/4} dH \right] \left[ \int \int \mathcal{C}^{\alpha+1/2+\delta/4}(F) r^b(G) dH \right]^{1/p_1} \left[ \int \int \mathcal{C}^{\alpha+1/2+\delta/4}(F) r^b(G) dH \right]^{1/q_1},
\]
which by the dominated convergence theorem converges to zero as \( \gamma \to 0 \). Hence under the assumptions of Theorem 2.1 (Theorem 2.2) a value \( \hat{\gamma} \) of \( \gamma \) can be chosen such that (5.3) is smaller than \( \epsilon \) for all \( H \) in \( H(H_{C_0}) \) provided \( \gamma \leq \hat{\gamma} \). For this \( \hat{\gamma} \) there exists an index \( \tilde{n} = \tilde{n}_{\hat{\gamma}} \) such that \( P(\{\Delta_n \geq S_\gamma \}) > 1-\varepsilon \) uniformly for \( H \) in \( H \), provided \( n \geq \tilde{n} \). It follows that under the assumptions of Theorem 2.1 (Theorem 2.2) the integral on the right in (5.1) is smaller than \( \varepsilon \) with probability larger than \( 1-\varepsilon \) uniformly for \( H \) in \( H(H_{C_0}) \) for all \( \gamma \leq \hat{\gamma} \) and all \( n \geq \tilde{n} \). \( \Box \)

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 \( \gamma \) and an index \( n_1 \) to ensure \( P(\{|B_{\gamma n}|, |C_{\gamma n}| < \varepsilon\}) > 1-\varepsilon \) for all \( n > n_1 \). Next use Assumption 2.3(a) and Corollaries 5.1-5.5 to choose for the above fixed \( \gamma \) an index \( n_2 = n_{2\gamma} > n_1 \) such that \( P(\{|B_{i n}|, |B_{\gamma j n}|, |C_{\gamma k n}| < \varepsilon \text{ for } i = 0,1,2; j = 3,5,6; k = 1,4\}) > 1-\varepsilon \). This implies that the probability that the sum of all these second order terms does not exceed \( 10\varepsilon \), is larger than \( 1-2\varepsilon \) uniformly for \( H \) in \( H(H_{C_0}) \), 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

\[ U_n^*(F) = n^{1/2}(F_n^* - F), \quad V_n^*(G) = n^{1/2}(G_n^* - G). \]

Instead of the set \( \Omega \), we shall use \( \Omega_n = \{ \omega: \sup |F_n^* - F| < \gamma/2, \sup |G_n^* - G| < \gamma/2 \} \). The role of \( \Lambda_n \) will be taken over by its closure

\[ \overline{\Lambda}_n = \overline{\Lambda}_{n1} \times \overline{\Lambda}_{n2} = [X_{1n}, X_{nn}] \times [Y_{1n}, Y_{nn}]. \]

Because integration over \( \overline{\Lambda}_n \) with respect to \( dH_n \) is the same as integration over the entire plane, we now have the simpler decomposition

\[
6.1 \quad n^{1/2}(T_n - \mu) = \sum_{i=1}^{p^3} A_{i,n} + B_{0n}^* + E_{i=1}^5 B_{\gamma i n}^* + B_{6n}^* + C_n^*,
\]

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

\[
B_{\gamma 1n}^* = \chi(\Omega_{\gamma n}^*) n^{1/2} \int [J(F^*_n) - J(F)]K(G)dH_n - A_{2n},
\]

\[
B_{\gamma 2n}^* = \chi(\Omega_{\gamma n}^*) n^{1/2} \int_{S_{\gamma}} [J(F^*_n) - J(F)]K(G)dH_n,
\]

\[
B_{\gamma 3n}^* = \chi(\Omega_{\gamma n}^*) \int_{S_{\gamma}} U_n^*(F)[J'(F^*_n) - J'(F)]K(G)dH_n,
\]

\[
B_{\gamma 4n}^* = \chi(\Omega_{\gamma n}^*) \int_{S_{\gamma}} U_n^*(F)J'(F)K(G)dH_n - H_n,
\]

\[
B_{\gamma 5n}^* = \chi(\Omega_{\gamma n}^*) \int_{S_{\gamma}} U_n^*(F)J'(F)K(G)dH_n - A_{2n},
\]

\[
B_{6n}^* = n^{1/2} \int J(F)[K(G^*_n) - K(G)]dH_n - A_{3n},
\]

\[
C_n^* = n^{1/2} \int [J(F^*_n) - J(F)][K(G^*_n) - K(G)]dH_n.
\]
The function $\psi_n^*$ 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 $\omega$ 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, satisfying

$$\min(F, F_n^*) \leq \Phi_n^* \leq \max(F, F_n^*)$$

and

$$\min(G, G_n^*) \leq \Psi_n^* \leq \max(G, G_n^*)$$

where defined. Then, uniformly for $n = 1, 2, \ldots$ and $\omega \in \Omega$:

(i) \[ \sup_{\Lambda_{n1}} r^\tau(\Phi_n^*) R^{-\tau}(F) = O(1) \] for each $\tau \geq 0$;

(ii) \[ \sup_{\Lambda_{n2}} r^\tau(\Psi_n^*) R^{-\tau}(G) = O(1) \] for each $\tau \geq 0$.

**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

$$P\left( \left\{ \beta F \leq F_n^* \leq 1 - \beta (1 - F) \right\} \right) > 1 - \varepsilon,$$

for all $n$ and uniformly in all continuous $F$. By (4.1) and because

$$1/2 \leq n/(n+1) \leq 1,$$

we only have to prove that

$$P\left( \left\{ n/(n+1) \leq 1 - \beta [1 - F(X_{nn})] \right\} \right) > 1 - \varepsilon$$

for $\beta$ small enough. Because the $F(X_i)$ are independent uniform $(0, 1)$ rvs, this probability equals

$$1 - \left\{ 1 - 1/[\beta (n+1)] \right\}^n > 1 - \varepsilon$$

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) \( \sup_{n \in \mathbb{N}} |U_n^*(F) - U_n(F)|r^{1/2-\rho(F)} \leq \frac{c}{p} \) as \( n \to \infty \), for each \( p > 0 \);

(ii) \( \sup_{n \in \mathbb{N}} |U_n^*(F)|r^{1/2-\tau(F)} = O(1) \) uniformly for \( n = 1, 2, \ldots \), for each \( \tau > 0 \).

PROOF. (i) Note that \( |U_n^*(F) - U_n(F)|r^{1/2-\rho(F)} < n^{-1/2} r^{1/2-\rho(F)} \) and that for any fixed \( \beta \in (0, 1) \) we have \( r^{1/2-\rho(\beta/n)} = r^{1/2-\rho(1-\beta/n)} = O(n^{1/2-\rho}) \). Because the \( F(X_i) \) are independent uniform rvs, given an arbitrary \( \epsilon > 0 \) we can choose a \( \beta = \beta_\epsilon \) in \( (0, 1) \) such that

\[ P\left( \beta/n \leq F(X_{inh}) \leq F(X_{nn}) \leq 1 - \beta/n \right) > 1 - \epsilon \] 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


