PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 9.1 (1988), p. 25-76

AN INVARIANCE PRINCIPLE FOR PROCESSES INDEXED BY TWO PARAMETERS AND SOME STATISTICAL APPLICATIONS

BY

MANFRED DENKER* (GÖTTINGEN) AND MADAN L. PURI** (BLOOMINGTON)

Abstract. Let $D((0, 1]^2)$ denote the space of all functions on $(0, 1]^2$ with no discontinuities of the second kind. We prove weak invariance principles in the space $D((0, 1]^2)$ for processes of the form $\int h(H_{n+m}(t))dF_n(t)$, $m, n \geqslant 1$, where F_n and G_m are two independent empirical distribution functions of independent, identically distributed sequences of random variables,

$$H_{n+m} = (n+m+1)^{-1}(nF_n+mG_m),$$

and where h belongs to a certain class of functions on the open unit interval. The appropriate topology in $D((0, 1]^2)$ is given by uniform convergence on compact sets. This type of processes is central in nonparametric statistics having applications to two-sample linear rank statistics and signed rank statistics.

1. INTRODUCTION

In this paper we prove several weak invariance principles for a certain type of stochastic integrals which arise from ranking procedures in nonparametric statistics. These integrals can be written in the form

$$\int h(H_{n+m}(t))dF_n(t) \quad (n, m \geqslant 1) \tag{1.1}$$

(see section 2 for definitions), thus the corresponding random functions for the invariance principle are functions in two variables from $[0, 1]^2$. It turns out that the important aspect of the problem lies in the behavior of the functions h: $(0, 1) \to \mathbf{R}$. In view of the result of Dupač and Hájek [9], our final result in section 8 is optimal concerning the imposed condition on h. It is also worth noticing that Schulze–Pillot's [18] approach using the convergence in "weighted supremum norm" of the two-sample empirical

* This research was done during a visit at the Department of Mathematics at Indiana University.

** Work supported by the Office of Naval Research Contract N/00014/85/K/0648.

process cannot lead to a result of the same generality. Another difficulty arises from the behavior of the random functions near (0, 0), which is overcome by using a topology of uniform convergence on certain compact sets. In some cases of functions h, however, the usual topology of uniform convergence suffices.

Weak invariance principles, as proved in this paper, have not appeared in literature. Schulze–Pillot's work will remain unpublished, hence we shall briefly mention some of his results in sections 3 and 9. Though the motivation to the present investigation comes from statistics, most of the material presented here does not depend on it; in any case this motivation will be apparent for specialists. Sections 2 and 10 illustrate these connections.

The paper is organized as follows: Section 2 contains the definition of random functions based on the two-sample linear rank statistics. Section 3 gives a brief introduction to weak convergence in those function spaces which are necessary for the purpose of our study. The limit process which appears in subsequent sections is investigated in section 4. It is a Gaussian process, indexed by points in $\{(r, s), 0 < r, s \leqslant 1\} \cup \{(0, 0)\}$ with continuous paths. Section 5 contains a first invariance principle for (1.1) in $D([0, 1]^2)$ for certain differentiable functions h. In section 6 we derive an approximation theorem for general functions h and use this result together with the result of section 5 to derive an invariance principle for absolutely continuous h. Finally, in section 8, an invariance principle for general h is derived (like in [9]). As mentioned before, section 9 contains results of the unpublished work of Schulze–Pillot and section 10 is devoted to some applications.

2. TWO-SAMPLE LINEAR RANK STATISTICS

Let X_i, $i \geqslant 1$, and Y_i, $i \geqslant 1$, be two independent sequences of independent and identically distributed random variables with continuous distribution functions F and G, respectively. Since F and G are isomorphic to Lebesgue measure on (0, 1), without loss of generality we can assume that F and G are absolutely continuous and

$$(F(x)+G(x))/2 = x, \qquad 0 \leqslant x \leqslant 1. \tag{2.1}$$

Let F_n and G_m be the empirical distribution functions based on $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, respectively. Let $R(i, n, m)$, $1 \leqslant i \leqslant n$, be the rank of X_i among $(X_1, \ldots, X_n, Y_1, \ldots, Y_m)$. Let $n+m = N$, and consider the scores $a(i, n, m)$ satisfying

$$a(i, n, m) = h\big(i/(N+1)\big), \qquad 1 \leqslant i \leqslant N, \tag{2.2}$$

or

$$\sum_{i=1}^{N} \left|a(i, n, m) - h\big(i/(N+1)\big)\right| = o(N^{1/2}), \tag{2.3}$$

where $h\colon (0, 1) \to \boldsymbol{R}$ denotes some Lebesgue measurable function to be specified later. Denote the linear rank statistics $T^*_{n,m}$ by

$$T^*_{n,m} = \sum_{i=1}^{n} a\big(R(i, n, m), n, m\big), \tag{2.4}$$

where the scores $a(\cdot, n, m)$ satisfy (2.2) or (2.3). Write

$$\hat{H}_{r,s}(t) = \frac{1}{(r+s)}\Big(\sum_{i=1}^{r} I_{[X_i \leqslant t]} + \sum_{j=1}^{s} I_{[Y_j \leqslant t]}\Big) = \frac{rF_r(t)+sG_s(t)}{r+s}, \qquad r, s \in N, \tag{2.5}$$

and

$$H_\lambda(t) = \lambda F(t) + (1-\lambda)\, G(t), \qquad 0 \leqslant \lambda \leqslant 1. \tag{2.6}$$

Then, if the scores are given by (2.2), we can rewrite $T^*_{n,m}$ as

$$T^*_{n,m} = n\int_0^1 h\left(\frac{N}{N+1}\hat{H}_{n,m}(t)\right) dF_n(t). \tag{2.7}$$

Let

$$T_{n,m} = \int_0^1 h\left(\frac{N}{N+1}\hat{H}_{n,m}(t)\right) dF_n(t) - \int_0^1 h\big(H_{n/N}(t)\big)\, dF(t). \tag{2.8}$$

For $N \in N$ define the random functions $S_N(h)$ by

$$S_N(h, r, s) = \begin{cases} 0 & \text{if } r < 1/N, \\ N^{-1/2}\,[rN]\; T_{[rN],[sN]}(h) & \text{if } r \geqslant 1/N. \end{cases} \qquad (0 \leqslant r,\ s \leqslant 1)$$

By definition, $S_N(h)$ is a random function with values in $D([0, 1]^2)$, the space of all right continuous bounded functions in $[0, 1]^2$ having at most discontinuities of the first kind. We will be interested in the weak convergence of $S_N(h)$ with respect to the uniform topology in a special case, and with respect to the topology of uniform convergence on certain compact sets – in the general case.

We need to put some restrictions on the score functions h. We assume that h is right continuous and has bounded variation on every compact set contained in the open interval $(0, 1)$. We then can write $h = h_1 - h_2$, where h_1 and h_2 are non-decreasing right continuous functions having bounded variation on every compact set in $(0, 1)$. We also assume that $h_1(1/2) = h_2(1/2) = 0$ (the last assumption and the right continuity are really no restrictions).

If h_1 and h_2 are integrable with respect to $[t(1-t)]^{-1/2}\,dt$, we put

$$\|h\| = \int\big(|h_1(t)| + |h_2(t)|\big)\big(t(1-t)\big)^{-1/2}\, dt.$$

Denote by $\mathscr{H}$ the set of all functions h with $\|h\| < \infty$. It is clear that if $h \in \mathscr{H}$, then $\int h^2(t)\,dt < \infty$ and

$$\int [t(1-t)]^{1/2}\, dh(t) \leqslant \|h\| \ll \int [t(1-t)]^{1/2}\, dh(t). \tag{2.9}$$

In the sequel, we shall use the Vinogradov symbol "$\ll$" instead of Landau's "0". Thus $a \ll b$ means that a/b is bounded.

If $h \in \mathscr{H}$ is absolutely continuous, then h can be approximated (in the norm $\| \ \|$) by functions which have bounded continuous second derivatives. In general, every h can be approximated by functions of the form $hI_{[\varepsilon, 1-\varepsilon]}$ for $\varepsilon > 0$.

Since $h \in \mathscr{H}$ has bounded variation on each compact set and satisfies $h(1/2) = 0$, we can rewrite $T_{n,m}$ (see Pyke and Shorack [15]) as

$$(2.10) \qquad T_{n,m}(h) = \int_0^1 \left[F_n\left(\left(\frac{N}{N+1} \hat{H}_{n,m} \right)^{-1} (t) \right) - F\left(H^{-1}_{\lambda(n,m)}(t)\right) \right] dh(t),$$

where

$$(2.11) \qquad \lambda(n, m) = n/N.$$

3. FUNCTION SPACES OF UNIFORM CONVERGENCE ON COMPACT SETS

For $0 < \varkappa \leqslant 1/2$ let

$$(3.1) \qquad E_\varkappa = \left\{ (r, s) \in [0, 1]^2 | \varkappa \leqslant \frac{r}{r+s} \leqslant 1 - \varkappa \right\} \cup \{(0, 0)\}$$

and let

$$(3.2) \qquad E = \bigcup_{0 < \varkappa \leqslant 1/2} E_\varkappa = (0, 1]^2 \cup \{(0, 0)\}.$$

Note that, $E_\varkappa$ being a compact subset of $[0, 1]^2$, we can consider the space $D(E_\varkappa)$ of all functions (on $E_\varkappa$) which are "right continuous" and have at most discontinuities of the first kind (see [14]). $D(E)$ denotes the set of all "right continuous" functions having no discontinuities of the second kind. More precisely, if $f \in D(E_\varkappa)$ (resp. $D(E)$) whenever $r_k \downarrow r$ and $s_k \downarrow s$, then $\lim_{n \to \infty} f(r_k, s_k) = f(r, s)$ and otherwise $\lim_{k \to \infty} f(r_k, s_k)$ exists (if $r_k \to r$ and $s_k \to s$). Note that $f | E_\varkappa \in D(E_\varkappa)$ whenever $f \in D E)$.

Let $d_\varkappa$ denote the Skorohod metric on $D(E_\varkappa)$ and let $\delta_\varkappa$ denote the uniform metric on $D(E_\varkappa)$. These metrics extend to metrics d and δ on $D(E)$ by the following definition: for $g, f \in D(E)$ let $d_\varkappa(f, g) = d_\varkappa(f|E_\varkappa, g|E_\varkappa)$ and $\delta_\varkappa(f, g) = \delta_\varkappa(f|E_\varkappa, g|E_\varkappa)$. Then

$$(3.3) \qquad d(f, g) = \sum_{k=2}^{\infty} \min\left(2^{-k}, d_{1/k}(f, g)\right)$$

and

$$(3.4) \qquad \delta(f, g) = \sum_{k=2}^{\infty} \min\left(2^{-k}, \delta_{1/k}(f, g)\right).$$

Both the metrics d and δ are bounded; d metrizes the topology of Skorohod convergence on each of the $E_\varkappa$, and δ metrizes the topology on uniform convergence on every $E_\varkappa$. We note the following

LEMMA 3.1. *$D(E)$ is complete with respect to both metrics, and the space $C(E)$ of continuous functions is a closed separable subspace. Moreover, the space $(D(E), d)$ is separable.*

$D(E)$ can be described as the δ-closure of all functions which are finite linear combinations of indicator functions of the form $I_{E \cap A_1 \times A_2}$, where each A_i is either an interval of the form $[a_i, b_i)$ or $\{1\}$.

We now discuss weak convergence in $D(E)$ with respect to both metrics d and δ to the extent it is needed in the following discussion.

Let $\{Z_n, n \geqslant 1\}$ be a family of random functions with values in $D([0, 1]^2)$. Since the map $D([0, 1]^2) \to D(E)$, given by $f \to f|E$, is continuous with respect to the Skorohod topology on $D([0, 1]^2)$ and the metric d on $D(E)$, and also continuous with respect to the uniform metric on $D([0, 1]^2)$ and the metric δ on $D(E)$, weak convergence of $Z_n \to Z_0$ in $D([0, 1]^2)$ implies the weak convergence in $D(E)$ with respect to the corresponding metric. In particular, we shall show (in Section 5) the weak convergence in $D([0, 1]^2)$ of $S_N(h)$ to some Gaussian process when h satisfies certain conditions. Later, in Section 7, we shall show that the above fact implies weak convergence in $D(E)$ for a large class of score functions h. Also, as a consequence of the weak convergence of the empirical process $N^{-1/2}[rN](F_{[rN]}(t) - F(t))$ to a Kiefer process in $D([0, 1]^2)$, we infer its weak convergence also in $D(E)$.

We also note that the embedding $(D(E), \delta) \to (D(E), d)$ is continuous, so that the weak convergence in the δ-metric implies that in the d-metric. Conversely, we can argue as in [2] that a sequence Z_n converges weakly in $(D(E), \delta)$ to Z_0 if it converges weakly in $(D(E), d)$ and Z_0 has its support on $C_b(E)$, the space of bounded continuous functions. This follows from the general fact that if Z_0 has a separable range in $(D(E), d)$, then the Borel fields given by δ and d coincide when restricted to that separable range.

If $D_0 \subset D(E)$ is separable, then the Borel σ-field is generated by the projections (see [2], Section 18). In particular, a random function X with values in D_0 is measurable if $\Pi \circ X$ is measurable for every projection. The random functions $S_N(h)$ clearly have a separable range as they use only the points $(k/N, j/N)$ $(0 \leqslant k, j \leqslant N)$ in E, and hence they are measurable.

We shall always reduce the weak convergence in $D(E)$ to that in $D([0, 1]^2)$; this way we can avoid speaking of tightness. In order to do so, we shall use the following well known lemma:

LEMMA 3.2. *Let Z_n and U_n, $n \geqslant 1$, be random functions with values in $D(E)$ and suppose that U_n converges weakly to U_0 with respect to δ (resp. d). If, for any $\varepsilon > 0$, $\lim \mathrm{P}(\delta(Z_n, U_n) \geqslant \varepsilon) = 0$ (resp. $\lim \mathrm{P}(d(Z_n, U_n) \geqslant \varepsilon) = 0$), then Z_n converges weakly to U_0 with respect to δ (resp. d).*

If $\{P_n, n \geqslant 0\}$ is a family of probability measures on $D(E)$ such that $P_n(C(E)) = 1$ $(n \geqslant 1)$ and $P_n \to P_0$ weakly with respect to δ or d, then $P_0(C(E)) = 1$.

The following lemma is a straightforward extension of Lemma 3.2:

LEMMA 3.3. *Let Z_n be random functions in $D(E)$. Suppose that for each $\varepsilon \geqslant 0$ there exists a sequence U_n^ε, $n \geqslant 1$, of random functions and a random function U^ε such that*

$$U_n^\varepsilon \to U^\varepsilon \tag{3.5}$$

weakly in $D(E)$ with respect to δ and such that

$$\mathrm{P}\big(\delta(Z_n, U_n^\varepsilon) \geqslant \eta\big) \leqslant \eta^{-2} a(\varepsilon) \quad \text{for any } \eta \geqslant 0. \tag{3.6}$$

If there exists a random function Z such that $U^\varepsilon \to Z$ as $\varepsilon \to 0$ in probability and if $\lim_{\varepsilon \to 0} a(\varepsilon) = 0$, then Z_n converges weakly in $D(E)$ with respect to δ.

4. THE PROCESS $Z(h)$

Let the distribution functions F and G be fixed. Let $\hat{\mathscr{H}}$ be a class of score functions $h \in \mathscr{H}$ such that, for any $0 < \lambda < 1$, the functions $F \circ H_\lambda^{-1}$ and $G \circ H_\lambda^{-1}$ are almost surely differentiable with respect to $d|h|$ ($d|v|$ denotes the total variation measure given by the signed measure v). This assumption always holds if h is absolutely continuous.

We shall write dF/dH_λ and dG/dH_λ for versions of the Radon–Nikodym derivatives satisfying

$$\frac{dF}{dH_\lambda} \circ H_\lambda^{-1} = (F \circ H_\lambda^{-1})' \text{ and } \frac{dG}{dH_\lambda} \circ H_\lambda^{-1} = (G \circ H_\lambda^{-1})' \; d|h| \text{ a.s.}$$

For $h \in \hat{\mathscr{H}}$, define a process $Z(h)$ on E by $Z(h, 0, 0) = 0$ and

$$\begin{aligned} Z(h, r, s) = & -\frac{s}{r+s} \int_0^1 \frac{dG(t)}{dH_{\lambda(r,s)}} \mathscr{K}_1(r, t)\, dh\big(H_{\lambda(r,s)}(t)\big) + \\ & + \frac{r}{r+s} \int_0^1 \frac{dF(t)}{dH_{\lambda(r,s)}} \mathscr{K}_2(s, t)\, dh\big(H_{\lambda(r,s)}(t)\big) \quad (r, s \neq 0), \end{aligned} \tag{4.1}$$

where $\mathscr{K}_1$ and $\mathscr{K}_2$ are two independent Kiefer processes on $C([0, 1]^2)$ with expectations zero and the covariance structure given by

$$\mathrm{E}\, \mathscr{K}_1(r, t)\, \mathscr{K}_1(r', t') = \min(r, r') F(t)\big(1 - F(t')\big), \quad 0 \leqslant t \leqslant t' \leqslant 1, \; 0 \leqslant r, r' \leqslant 1$$

and

$$\mathrm{E}\, \mathscr{K}_2(s, t)\, \mathscr{K}_2(s', t') = \min(s, s') G(t)\big(1 - G(t')\big), \quad 0 \leqslant t \leqslant t' \leqslant 1, \; 0 \leqslant s, s' \leqslant 1$$

LEMMA 4.1. *The process $Z(h)$ has the same distribution as the process $-Z(h)$ and the process $Z^*(h)$, where*

$$(4.2)\quad Z^*(h,r,s) = -\frac{s}{r+s}\int_0^1 (G\circ H_{\lambda(r,s)}^{-1})'(t)\,\mathscr{K}_1\big(r, H_{\lambda(r,s)}^{-1}(t)\big)\,dh(t) + \frac{r}{r+s}\int_0^1 (F\circ H_{\lambda(r,s)}^{-1})'(t)\,\mathscr{K}_2\big(s, H_{\lambda(r,s)}^{-1}(t)\big)\,dh(t).$$

If h is absolutely continuous and h' is integrable, then $Z(h)$ has the same distribution as $Z_1^*(h)$ where

$$(4.3)\quad Z_1^*(h,r,s) = \int_0^1 \Big[h\big(H_{\lambda(r,s)}(u)\big) + \frac{r}{r+s}\int_t^1 h'\big(H_{\lambda(r,s)}(u)\big)\,dF(u)\Big]\,\mathscr{K}_1(r,dt) + \frac{r}{r+s}\int_0^1 h'(t)\,\mathscr{K}_2\big(s, H_{\lambda(r,s)}^{-1}(t)\big)\,dF\big(H_{\lambda(r,s)}^{-1}(t)\big)$$

Proof. Since for any Kiefer process $\mathscr{K}$ under consideration, $-\mathscr{K}$ has the same distribution, it follows that $Z(h)$ and $-Z(h)$ have the same distribution. That $Z^*(h)$ has the same distribution as $Z(h)$ follows by a change of variables in (4.1).

Finally, if h is absolutely continuous and if h' is Lebesgue integrable, then from (4.2) we obtain

$$Z(h,r,s) = -\frac{s}{r+s}\int_0^1 h'(t)\,\mathscr{K}_1\big(r, H_{\lambda(r,s)}^{-1}(t)\big)\,dG\big(H_{\lambda(r,s)}^{-1}(t)\big) + \frac{r}{r+s}\int_0^1 h'(t)\,\mathscr{K}_2\big(s, H_{\lambda(r,s)}^{-1}(t)\big)\,dF\big(H_{\lambda(r,s)}^{-1}(t)\big) =$$

$$= -\int_0^1 h'(t)\,\mathscr{K}_1\big(r, H_{\lambda(r,s)}^{-1}(t)\big)\,dt + \frac{r}{r+s}\int_0^1 h'(t)\,\mathscr{K}_1\big(r, H_{\lambda(r,s)}^{-1}(t)\big)$$

$$+\frac{r}{r+s}\int_0^1 h'(t)\,\mathscr{K}_2\big(s, H_{\lambda(r,s)}^{-1}(t)\big)\,dF\big(H_{\lambda(r,s)}^{-1}(t)\big)$$

$$= \int_0^1 \Big[h\big(H_{\lambda(r,s)}(t)\big) + \frac{r}{r+s}\int_t^1 h'\big(H_{\lambda(r,s)}(u)\big)\,dF(u)\Big]\,\mathscr{K}_1(r,dt) + \frac{r}{r+s}\int_0^1\int_t^1 h'\big(H_{\lambda(r,s)}(u)\big)\,dF(u)\,\mathscr{K}_2(s,dt),$$

which proves (4.3).

The integrals with respect to the Brownian bridge component of the Kiefer process were first defined by Filippova [10] for functions h satisfying $\int h^2(x)\,dF(x) < \infty$ (resp. $\int h^2(x)\,dG(x) < \infty$). Indeed, if h' is integrable and belongs to $\mathscr{H}$, then by (2.9) it is easy show that these integrals are finite in our special case. It is also clear that the processes of the form $\int f(t)\,\mathscr{K}(r, t)\,dv(t)$, where v is some measure and f some square integrable function, have continuous paths and are in fact Wiener processes with zero expectations and covariance

$$\mathrm{E}\int f(t)\,\mathscr{K}(r, t)\,dv(t)\int f(t)\,\mathscr{K}(r', t)\,dv(t)$$
$$= 2\iint_{u<v} f(u)f(v)\min(r, r')\,u(1-v)\,dv(u)\,dv(v)+$$
$$+\iint I_{\{u=v\}} f^2(u)\min(r, r')\,u(1-u)\,dv(u)\,dv(v).$$

Since, in (4.1), $H_{\lambda(r,s)}(t)$ depends continuously on r and s, the processes $Z(h)$ must have continuous paths. They are obviously Gaussian and centered at their expectations. Their covariance structure is given by the following

LEMMA 4.2. *We have*

$$\text{(4.4)}\qquad \mathrm{E}Z(h, r, s)Z(h, r', s') = \min(r, r')\frac{ss'}{(r+s)(r'+s')}\times$$
$$\times\iint (G\circ H^{-1}_{\lambda(r,s)})'(u)(G\circ H^{-1}_{\lambda(r',s')})'(v)\min\big(F(H^{-1}_{\lambda(r,s)}(u)),\, F(H^{-1}_{\lambda(r',s')}(v)\big)\times$$
$$\times\Big(1-\max\big(F(H^{-1}_{\lambda(r,s)}(u)),\, F(H^{-1}_{\lambda(r',s')}(v))\big)\Big)\,dh(u)\,dh(v)+$$
$$+\min(s, s')\frac{rr'}{(r+s)(r'+s')}\iint (F\circ H^{-1}_{\lambda(r,s)})'(u)(F\circ H^{-1}_{\lambda(r',s')})'(v)\times$$
$$\times\min\big(G(H^{-1}_{\lambda(r,s)}(u)),\, G(H^{-1}_{\lambda(r',s')}(v)\big)\times$$
$$\times\Big(1-\max\big(G(H^{-1}_{\lambda(r,s)}(u)),\, G(H^{-1}_{\lambda(r',s')}(v))\big)\Big)\,dh(u)\,dh(v)$$

and $|\mathrm{E}Z(h, r, s)Z(h, r', s')| \ll \|h\|^2$, *where* "$\ll$" *depends on* r, r', s *and* s'.

Proof. Let $r < r'$ and $s < s'$ (the other cases can be proved analogously). Then

$$\mathrm{E}Z(h, r, s)Z(h, r', s') = \mathrm{E}\left\{\frac{s}{r+s}\int_0^1 \frac{dG}{dH_{\lambda(r,s)}}(t)\,\mathscr{K}_1(r, t)\,dh\big(H_{\lambda(r,s)}(t)\big)-\right.$$
$$\left.-\frac{r}{r+s}\int_0^1 \frac{dF}{dH_{\lambda(r,s)}}(t)\,\mathscr{K}_2(s, t)\,dh\big(H_{\lambda(r,s)}(t)\big)\right\}\times$$
$$\times\left\{\frac{s'}{r'+s'}\int_0^1 \frac{dG}{dH_{\lambda(r,s)}}(t)\,\mathscr{K}_1(r', t)\,dh\big(H_{\lambda(r',s')}(t)\big)-\right.$$
$$\left.-\frac{r'}{r'+s'}\int_0^1 \frac{dF}{dH_{\lambda(r',s')}}(t)\,\mathscr{K}_2(s', t)\,dh\big(H_{\lambda(r',s')}(t)\big)\right\}$$

$$= \frac{ss'}{(r+s)(r+s')} \iint \frac{dG}{dH_{\lambda(r,s)}}(u) \frac{dG}{dH_{\lambda(r',s')}}(v) \times$$

$$\times \mathrm{E}\mathcal{K}_1(r, u)\, \mathcal{K}_1(r', v)\, dh\big(H_{\lambda(r,s)}(u)\big)\, dh\big(H_{\lambda(r',s')}(v)\big) +$$

$$+ \frac{rr'}{(r+s)(r'+s')} \iint \frac{dF}{dH_{\lambda(r,s)}}(u) \frac{dF}{dH_{\lambda(r',s')}}(v) \times$$

$$\times \mathrm{E}\mathcal{K}_2(s, u)\, \mathcal{K}_2(s', v)\, dh\big(H_{\lambda(r,s)}(u)\big)\, dh\big(H_{\lambda(r',s')}(v)\big).$$

Since $\mathrm{E}\mathcal{K}_1(r, u)\mathcal{K}_1(r', v) = r\min\big(F(u), F(v)\big)\big(1-\max\big(F(u), F(v)\big)\big)$ and $\mathrm{E}\mathcal{K}_2(s, u)\mathcal{K}_2(s', v) = s\, \min\big(G(u), G(v)\big)\big(1-\max\big(G(u), G(v)\big)\big)$, (4.4) follows by change of variables.

If h is non-decreasing, then both the integrands in (4.4) are bounded by $C(r, r', s, s')\|h\|^2$, where $C(r, r', s, s')$ is a constant, independent of h, but dependent on r, r', s, s'. Since the norm $\|\ \|$ of h on $\mathcal{H}$ is compatible with the Jordan decomposition of h, it follows that

$$|\mathrm{E}Z(h, r, s)\, Z(h, r', s')| \leqslant \|h\|^2.$$

If, in (4.1), $r/(r+s)$ is fixed, so are $H_{\lambda(r,s)}$ and $s/(r+s)$. Hence, for any fixed c $(0 < c < 1)$,

$$Z\left(h, r, \frac{1-c}{c}r\right) \tag{4.5}$$

is a sum of two independent Wiener processes with covariance structure given by Lemma 4.2. In particular, the variance is given by

$$\mathrm{E}Z\left(h, r, \frac{1-c}{c}r\right)^2 = r\sigma^2,$$

where

$$\sigma^2 = 2(1-c)^2 \iint_{u<v} \{(G\circ H_c^{-1})'(u)(G\circ H_c^{-1})'(v)(F\circ H_c^{-1})(u) \times$$

$$\times \big(1 - F\circ H_c^{-1}(v)\big)\}\, dh(u)\, dh(v) +$$

$$+ 2c(1-c) \iint_{u<v} \{(F\circ H_c^{-1})'(u)(F\circ H_c^{-1})'(v)(G\circ H_c^{-1})(u) \times$$

$$\times \big(1 - G\circ H_c^{-1}(v)\big)\}\, dh(u)\, dh(v),$$

provided dh has no point masses.

We also need the following

PROPOSITION 4.3. *If $h_1, h_2 \in \hat{\mathcal{H}}$, then*

$$\mathrm{E}\big[\delta\big(Z(h_1), Z(h_2)\big)\big]^2 \leqslant \|h_1 - h_2\|. \tag{4.6}$$

In particular, if $\{h_n\colon n \geqslant 1\}$ converges to h in norm, then $Z(h_n) \to Z(h)$ in probability with respect to δ.

Proof. We shall use the representation of $Z(h)$ given in Lemma 4.1. It is sufficient to show that $\mathrm{E}[\delta(Z(h), 0)]^2 \ll \|h\|$ for $h \in \hat{\mathscr{H}}$.

Fix $0 < \varkappa < \frac{1}{2}$ and denote by $I(z)$ the closed interval with endpoints $H_\varkappa^{-1}(z)$ and $H_{1-\varkappa}^{-1}(z)$, $0 < z < 1$. Assume first that – given a fixed $z \in (0, 1)$ $-F(H_\varkappa^{-1}(z)) \geqslant G(H_\varkappa^{-1}(z))$. In this case $I(z) = [H_{1-\varkappa}^{-1}(z), H_\varkappa^{-1}(z)]$ and it follows that

$$\max_{(r,s)\in E_\varkappa} |\mathscr{K}_2(s, H_{\lambda(r,s)}^{-1}(z))| \leqslant \max_{0<s\leqslant 1, y\in I(z)} |\mathscr{K}_2(s, y)|.$$

Since

$$\{\sup_{y\in I(z)} |\mathscr{K}_2(s, y)|^2 \colon 0 < s \leqslant 1\} \quad \text{and} \quad \left\{\frac{(\mathscr{K}_2(1, y))^2}{(1-G(y))^2} \colon y \in I(z)\right\}$$

are non-negative submartingales, applying Doob's inequality twice ([8], p. 314), it follows that

$$\mathrm{E}\max_{(r,s)\in E_\varkappa} |\mathscr{K}_2(s, H_{\lambda(r,s)}^{-1}(z))|^2 \leqslant 16\mathrm{E}(\mathscr{K}_2(1, H_\varkappa^{-1}(z)))^2 \left(\frac{1-G(H_{1-\varkappa}^{-1}(z))}{1-G(H_\varkappa^{-1}(z))}\right)^2.$$

(Note that $G(H_\varkappa^{-1}(z)) \neq 1$). Now, since

$$\mathrm{E}(\mathscr{K}_2(1, \mathscr{H}_\varkappa^{-1}(z))^2 = G(H_\varkappa^{-1}(z))(1-G(H_\varkappa^{-1}(z))) \leqslant (1-\varkappa)^{-1} z(1-z),$$

$$1-z = 1-\varkappa F(H_\varkappa^{-1}(z)) - (1-\varkappa) G(H_\varkappa^{-1}(z)) \leqslant 1-G(H_\varkappa^{-1}(z)),$$

$$\varkappa[1-G(H_{1-\varkappa}^{-1}(z))] \leqslant 1-(1-\varkappa) F(H_{1-\varkappa}^{-1}(z)) - \varkappa G(H_{1-\varkappa}^{-1}(z)) = 1-z,$$

therefore

(4.7) $$\mathrm{E}\max_{(r,s)\in E_\varkappa} |\mathscr{K}_2(s, H_{\lambda(r,s)}^{-1}(z))|^2 \leqslant 16\varkappa^{-2}(1-\varkappa)^{-1} z(1-z).$$

In order to prove the corresponding inequality for $\mathscr{K}_1$, let us first assume that $G(H_{1-\varkappa}^{-1}(t)) \leqslant F(H_{1-\varkappa}^{-1}(t))$ holds as well as $F(H_\varkappa^{-1}(z)) \geqslant G(H_\varkappa^{-1}(z))$.

Here we use the fact that $\{\mathscr{K}_1(1, y)^2/F(y)^2 \colon y \in I(z)\}$ also is a non-negative (backward) submartingale and we obtain

$$\mathrm{E}\max_{(r,s)\in E_\varkappa} |\mathscr{K}_1(r, H_{\lambda(r,s)}^{-1}(z)|^2 \leqslant 16\mathrm{E}(\mathscr{K}_1(1, H_{1-\varkappa}^{-1}(z)))^2 \left(\frac{F(H_\varkappa^{-1}(z))}{F(H_{1-\varkappa}^{-1}(z))}\right)^2.$$

Here $F(H_{1-\varkappa}^{-1}(z)) \neq 0$. Moreover,

$$\mathrm{E}(\mathscr{K}_1(1, H_{1-\varkappa}^{-1}(z)))^2 = F(H_{1-\varkappa}^{-1}(z))(1-F(H_{1-\varkappa}^{-1}(z))) \leqslant (1-\varkappa)^{-1} z(1-z),$$

$$z = (1-\varkappa) F(H_{1-\varkappa}^{-1}(z)) + \varkappa G(H_{1-\varkappa}^{-1}(z)) \leqslant F(H_{1-\varkappa}^{-1}(z)) \quad \text{and} \quad \varkappa F(H_\varkappa^{-1}(z)) \leqslant z.$$

Consequently,

(4.8) $$\mathrm{E}\max_{(r,s)\in E_\varkappa}\left|\mathscr{K}_1\left(r, H^{-1}_{\lambda(r,s)}(z)\right)\right|^2 \leqslant 16\varkappa^{-2}(1-\varkappa)^{-1}z(1-z).$$

Let now $G\left(H^{-1}_{1-\varkappa}(z)\right) > F\left(H^{-1}_{1-\varkappa}(z)\right)$. Then by continuity, there exists a $t_0\in[\varkappa, 1-\varkappa]$ with $G\left(H^{-1}_{t_0}(z)\right) = F\left(H^{-1}_{t_0}(z)\right)$. Put $I_1 = [H^{-1}_{1-\varkappa}(z), t_0]$ and $I_2 = [t_0, H^{-1}_\varkappa(z)]$. Then, on I_1, $G(t_0)\geqslant F(t_0)$ and the argument showing (4.7) (interchanging F with G and replacing $\mathscr{K}_2$ by $\mathscr{K}_1$) gives

$$\mathrm{E}\max_{0<r\leqslant 1, y\in I_1}|\mathscr{K}_1(r, y)|^2 \leqslant 16\varkappa^{-2}(1-\varkappa)^{-1}z(1-z).$$

On I_2 we can use the argument for (4.8) to show an analogous inequality. Hence

(4.9) $$\mathrm{E}\max_{(r,s)\in E_\varkappa}\left|\mathscr{K}_1\left(r, H^{-1}_{\lambda(r,s)}(z)\right)\right|^2 \leqslant 32\varkappa^{-2}(1-\varkappa)^{-1}z(1-z).$$

Finally, if $F\left(H^{-1}_\varkappa(z)\right)\leqslant G\left(H^{-1}_\varkappa(z)\right)$, we interchange in the preceding arguments F by G and $\mathscr{K}_1$ by $\mathscr{K}_2$ and obtain

$$\mathrm{E}\max_{(r,s)\in E_\varkappa}\left|\mathscr{K}_1\left(r, H^{-1}_{\lambda(r,s)}(z)\right)\right|^2 \leqslant 32\varkappa^{-2}(1-\varkappa)^{-1}z(1-z),$$

(4.10) $$\mathrm{E}\max_{(r,s)\in E_\varkappa}\left|\mathscr{K}_2\left(s, H^{-1}_{\lambda(r,s)}(z)\right)\right|^2 \leqslant 32\varkappa^{-2}(1-\varkappa)^{-1}z(1-z) \quad \text{for any } 0<z<1.$$

By our assumption, $(G\circ H^{-1}_{\lambda(r,s)})'$ and $(F\circ H^{-1}_{\lambda(r,s)})'$ are bounded functions if $(r, s)\in \mathrm{E}_\varkappa$. From this and (4.10) we obtain finally (if h is non-decreasing) that

$$\mathrm{E}\sup_{(r,s)\in E_\varkappa}\left(Z^*(h, r, s)\right)^2$$

$$\leqslant 2\sup_{(r,s)\in E_\varkappa}\frac{s^2}{(r+s)^2}\left(\int (G\circ H^{-1}_{\lambda(r,s)})'(z)\times\sqrt{32\varkappa^{-2}(1-\varkappa)^{-1}}\sqrt{z(1-z)}\,dh(z)+\right.$$

$$\left.+2\sup_{(r,s)\in E_\varkappa}\frac{r^2}{(r+s)^2}\left(\int (F\circ H^{-1}_{\lambda(r,s)})'(z)\sqrt{32\varkappa^{-2}(1-\varkappa)^{-1}}\sqrt{z(1-z)}\,dh(z)\leqslant C(\varkappa)\,\|h\|^2,\right.$$

where $C(\varkappa)$ denotes a constant with a polynomial dependence on $\varkappa^{-1}$.

Clearly, the same inequality holds for general h because of the definition of the norm $\|\ \|$. Choose $m\in N$ such

$$\sum_{k\leqslant m} C(1/k)^{1/2} \leqslant \|h\|^{-1/2} \leqslant 2^{m-1}.$$

Then

$$\left(\mathrm{E}\left(\delta(Z^*(h),0)\right)^2\right)^{1/2} = \left\{\mathrm{E}\left(\sum_{k=2}^{\infty}\min\left(2^{-k}, \sup_{(r,s)\in E_\varkappa} Z^*(h,r,s)\right)\right)^2\right\}^{1/2}$$

$$\leqslant \sum_{k=2}^{m} C(1/k)^{1/2}\|h\| + \sum_{k>m} 2^{-k} \leqslant \|h\|^{1/2} + 2^{-m+1} \leqslant 2\|h\|^{1/2},$$

proving (4.6).

5. AN INVARIANCE PRINCIPLE FOR C^2-SCORE

THEOREM 5.1. *Let the score function h of class C^2 satisfy*

$$|h''(t)| \leqslant K\left(t(1-t)\right)^{-1+\eta} \tag{5.1}$$

for some $\eta > 0$ and $K > 0$. Then, the processes $S_N(h)$, $N \geqslant 1$, converge weakly in $D([0,1]^2)$ with respect to the uniform metric to the process $Z(h)$ defined in (4.1). (Here $Z(h)$ can be considered on $C([0,1]^2)$).

The proof of this theorem is based on the following lemmas.

LEMMA 5.1. *We have*

$$\lim_{N\to\infty} \sup_{\substack{0\leqslant r,s\leqslant 1\\ r\geqslant 1/N}} [rN]N^{-1/2}\left|\int_0^1 \{h(H_{\lambda(r,s)}(t)) - h(H_{\lambda_{r,s}}(t))\}\,dF(t)\right| = 0, \tag{5.2}$$

(5.3)

$$\lim_{N\to\infty} \sup_{\substack{0\leqslant r,s\leqslant 1\\ r\geqslant 1/N}} [rN]N^{-1/2}\left|\int_0^1 \{h(H_{\lambda(r,s)}(t)) - h(H_{\lambda_{r,s}}(t))\}\,d\left(F_{[rN]}(t) - F(t)\right)\right| = 0,$$

where

$$\lambda_{r,s} = [rN]/N(r,s). \tag{5.4}$$

Proof. By the Mean Value Theorem, for some ξ in the interval given by $H_{\lambda(r,s)}(t)$ and $H_{\lambda_{r,s}}(t)$ we have

$$\left|h(H_{\lambda(r,s)}(t)) - h(H_{\lambda_{r,s}}(t))\right| = |h'(\xi)|\,|H_{\lambda(r,s)}(t) - H_{\lambda_{r,s}}(t)|$$

$$\leqslant |H_{\lambda(r,s)}(t) - H_{\lambda_{r,s}}(t)| \leqslant \left(\left|\frac{[rN]}{N(r,s)} - \frac{r}{r+s}\right| + \left|\frac{[sN]}{N(r,s)} - \frac{s}{r+s}\right|\right) \leqslant N(r,s)^{-1}.$$

(Note: $h'(\xi)$ is bounded.)

LEMMA 5.2. *We have*

$$\lim_{N\to\infty} \mathrm{E}\sup_{0\leqslant r,s\leqslant 1} \Big[[rN]N^{-1/2}\frac{r}{r+s}\int_0^1 \{h'(H_{\lambda(r,s)}(t)) - h'(H_{\lambda_{r,s}}(t))\} \times \tag{5.5}$$

$$\times \{F_{[rN]}(t) - F(t)\}\,dF(t)\Big]^2 = 0,$$

(5.6)
$$\lim_{N\to\infty} \mathrm{E} \sup_{0\leqslant r,s\leqslant 1} \Big[[sN] N^{-1/2} \frac{r}{r+s} \int_0^1 \{h'(H_{\lambda(r,s)}(t)) - h'(H_{\lambda_{r,s}}(t))\} \times \times \{G_{[sN]}(t) - G(t)\}\, dF(t) \Big]^2 = 0.$$

Proof. For $0 \leqslant t \leqslant 1$, $2/N \leqslant r \leqslant 1$ and $0 \leqslant s \leqslant 1$ there exists some ξ in the interval given by $H_{\lambda(r,s)}(t)$ and $H_{\lambda_{r,s}}(t)$ with

$$|h'(H_{\lambda(r,s)}(t)) - h'(H_{\lambda_{r,s}}(t))| = |h''(\xi)|\, |H_{\lambda(r,s)}(t) - H_{\lambda_{r,s}}(t)|$$
$$\leqslant \frac{2K}{[rN]} \Big(\Big\{ \frac{r}{r+s} (H_{\lambda(r,s)}(t)(1 - H_{\lambda(r,s)}(t)) \Big\}^{-1+\eta} + + \frac{[Nr]}{N(r,s)} \{H_{\lambda_{r,s}}(t)(1 - H_{\lambda_{r,s}}(t))\}^{-1+\eta} \Big).$$

Since

$$\mathrm{E} \sup_{1\leqslant n\leqslant N} (F_n(t) - F(t))^2 \leqslant \sum_{n=1}^{N} n^{-1} F(t)(1-F(t)) \ll (\log N) F(t)(1-F(t)),$$

it follows for $q = [\log N]^2$ that

$$\mathrm{E} \sup_{\substack{q/N\leqslant r\leqslant 1\\ 0\leqslant s\leqslant 1}} \Big[[rN] N^{-1/2} \frac{r}{r+s} \int_0^1 \{h'(H_{\lambda(r,s)}(t)) - h'(H_{\lambda_{r,s}}(t))\} \times \times \{F_{[rN]}(t) - F(t)\}\, dF(t) \Big]^2$$

$$\leqslant \Big(\int_0^1 \sup([rN] N^{-1/2})^2 |h'(H_{\lambda(r,s)}(u)) - h'(H_{\lambda_{r,s}}(u))| \times \times \{\mathrm{E} \max_{1\leqslant n\leqslant N} |F_n(u) - F(u)|^2\}^{1/2} dF(u) \Big)^2$$

$$\ll N^{-1} \log N \Big(\int_0^1 \sup_{r,s} \Big\{ (H_{\lambda(r,s)}(u)(1 - H_{\lambda(r,s)}(u)))^{-1+\eta} \Big(\frac{r}{r+s} F(u)(1-F(u) \Big) + + (H_{\lambda_{r,s}}(u)(1 - H_{\lambda_{r,s}}(u)))^{-1+\eta} \Big(\frac{[rN]}{N(r,s)} F(u)(1-F(u)) \Big)^{1/2} \Big\} dF(u) \Big)^2$$

$$\leqslant N^{-1} \log N \Big(\int_0^1 \sup_{r,s} \{ (H_{\lambda(r,s)}(t)(1 - H_{\lambda(r,s)}(t)))^{-1/2+\eta} + + (H_{\lambda_{r,s}}(t)(1 - H_{\lambda_{r,s}}(t)))^{-1/2+\eta} \} dF(t) \Big)^2$$

$$\ll N^{-1} \log N \Big(\int_0^1 (t(1-t))^{-1/2+\eta} dt \Big)^2 +$$

$$+N^{-1}\log N\,\Big\{\int_0^1 \Big(H_{\frac{q}{2(N+q)}}(t)\big(1-H_{\frac{q}{2(N+q)}}(t)\big)\Big)^{-1/2+\eta} dF(t)\Big\}^2$$

$$\leqslant [\log N]^{-1}\Big\{\int_0^1 \big(t(1-t)\big)^{-1/2+\eta} dt\Big\}^2$$

by the Cauchy–Schwarz inequality.

In the derivation of the above inequalities we have used the following facts:

(i) for $0 \leqslant \lambda \leqslant 1$, $H_\lambda(t)\big(1-H_\lambda(t)\big) \geqslant \lambda F(t)\big(1-F(t)\big)$;

(ii) if $q/N \leqslant r \leqslant 1$ and $0 \leqslant s \leqslant 1$, then:

(a) for $G(t) \leqslant F(t)$,

$$H_{\frac{q}{(q+N)}}(t) \leqslant H_{\lambda(r,s)}(t) \leqslant F(t) \quad \text{and} \quad H_{\frac{q}{2(q+N)}}(t) \leqslant H_{\lambda_{r,s}}(t) \leqslant F(t),$$

(b) for $G(t) \geqslant F(t)$,

$$F(t) \leqslant H_{\lambda(r,s)}(t) \leqslant H_{\frac{q}{q+N}}(t) \quad \text{and} \quad F(t) \leqslant H_{\lambda_{r,s}}(t) \leqslant H_{\frac{q}{2(q+N)}}(t).$$

The proof of (5.6) is analogous to that of (5.5) and is therefore omitted.

LEMMA 5.3. *If* $F(t) \geqslant G(t)$, *then for any* $r \leqslant r'$ *and* $s' \leqslant s$, $H_{\lambda(r,s)}(t) \leqslant H_{\lambda(r',s')}(t)$, *and if* $F(t) \leqslant G(t)$, *then for any* $r \leqslant r'$ *and* $s' \leqslant s$, $H_{\lambda(r,s)}(t) \geqslant H_{\lambda(r',s')}(t)$.

Since the proof is straightforward, it has been omitted.

LEMMA 5.4 *We have*

$$(5.7)\qquad \lim_{N\to\infty} \mathrm{E} \max_{1\leqslant n,m\leqslant N} \Big(N^{-1/2}(n+m)^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m}\big\{h'\big(H_{\frac{n}{n+m}}(X_i)\big)I_{[X_j\leqslant X_i]} - h'\big(H_{\frac{n}{n+m}}(X_i)\big)F(X_i) - \int_0^1 h'\big(H_{\frac{n}{n+m}}(t)\big)I_{[X_j\leqslant t]}\,dF(t) + \int_0^1 h'\big(H_{\frac{n}{n+m}}(t)\big)F(t)\,dF(t)\big\}\Big)^2 = 0,$$

$$(5.8)\qquad \lim_{N\to\infty} \mathrm{E} \max_{1\leqslant n,m\leqslant N} \Big(N^{-1/2}(n+m)^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m}\big\{h'\big(H_{\frac{n}{n+m}}(X_i)\big)I_{[Y_j\leqslant X_i]} - h'\big(H_{\frac{n}{n+m}}(X_i)\big)G(X_i) - \int_0^1 h'\big(H_{\frac{n}{n+m}}(t)\big)I_{[Y_j\leqslant t]}\,dF(t) + \int_0^1 h'\big(H_{\frac{n}{n+m}}(t)\big)G(t)\,dF(t)\big\}\Big)^2 = 0.$$

Proof. Write

$$f_{n,m}(u, v) = h'\big(H_{\frac{n}{n+m}}(u)\big)I_{[v\leqslant u]} - h'\big(H_{\frac{n}{n+m}}(u)\big)F(u) - \int_0^1 h'\big(H_{\frac{n}{n+m}}(t)\big)I_{[v\leqslant t]}\,dF(t) + \int_0^1 h'\big(H_{\frac{n}{n+m}}(t)\big)F(t)\,dt$$

and

$$W_{n,m} = \sum_{i \neq j = 1}^{n} f_{n,m}(X_i, X_j).$$

We first estimate $\mathrm{E} \max_{1 \leqslant n,m} N^{-1}(n+m)^{-2} W_{n,m}^2$. Define

$$\omega_{n,m} = W_{n,m} - W_{n-1,m} - W_{n,m-1} + W_{n-1,m-1} \qquad (n, m \geqslant 1),$$

where $W_{-j,k} = W_{j,-k} = 0$ for $j, k \geqslant 0$.

It is easy to check (by induction) that

$$W_{n,m} = \sum_{k=1}^{n} \sum_{j=1}^{m} \omega_{k,j}$$

and, for $1 \leqslant a < b \leqslant N$ and $1 \leqslant u < v \leqslant N$,

$$\begin{aligned}
\sum_{k=a+1}^{b} \sum_{j=u+1}^{v} \omega_{k,j} &= \sum_{k=a+1}^{b} \sum_{j=u+1}^{v} W_{k,j} - \sum_{k=a}^{b-1} \sum_{j=u+1}^{v} W_{k,j} - \\
&\qquad - \sum_{k=a+1}^{b} \sum_{j=u}^{v-1} W_{k,j} + \sum_{k=a}^{b-1} \sum_{j=u}^{v-1} W_{k,j} \\
&= \sum_{j=u+1}^{v} (W_{b,j} - W_{a,j}) + \sum_{j=u}^{v-1} (W_{a,j} - W_{b,j}) \\
&= W_{b,v} - W_{b,u} - W_{a,v} + W_{a,u} \\
&= \sum_{i \neq j=1}^{b} \big(f_{b,v}(X_i, X_j) - f_{b,u}(X_i, X_j)\big) - \\
&\qquad - \sum_{i \neq j=1}^{a} \big(f_{a,v}(X_i, X_j) - f_{a,u}(X_i, X_j)\big) \\
&= \sum_{i \neq j=1}^{a} (f_{b,v} - f_{b,u} - f_{a,v} + f_{a,u})(X_i, X_j) + \\
&\quad + \sum_{i=a+1}^{b} \sum_{j=1}^{a} (f_{b,v} - f_{b,u})(X_i, X_j) + \sum_{j=a+1}^{b} \sum_{i=1}^{a} (f_{b,v} - f_{b,u})(X_i, X_j) + \\
&\qquad + \sum_{\substack{i=a+1 \\ i \neq j}}^{b} \sum_{j=a+1}^{b} (f_{b,v} - f_{b,u})(X_i, X_j).
\end{aligned}$$

Because of independence, the square integral of the second and third summands is

$$\leqslant a(b-a) \int\int (f_{b,v} - f_{b,u})^2 (x, y)\, dF(x)\, dF(y) \tag{5.9}$$

and the square integral of the last summand is

$$\ll (b-a)^2 \int\int (f_{b,v}-f_{b,u})^2 (x, y)\, dF(x)\, dF(y). \tag{5.10}$$

To estimate the double integrals, it suffices to estimate

$$\int\int \left|h'\left(H_{\frac{b}{b+v}}(x)\right)-h'\left(H_{\frac{b}{b+u}}(x)\right)\right|^2 \left(I_{[y\leq x]}-F(x)\right)^2 dF(x)\, dF(y).$$

Considering different cases, it can be shown that

$$\begin{aligned}\left|h'\left(H_{\frac{b}{b+v}}(x)\right)-h'\left(H_{\frac{b}{b+u}}(x)\right)\right|^2 &\ll \Big[\left\{H_{\frac{b}{b+v}}(x)\left(1-H_{\frac{b}{b+v}}(x)\right)\right\}^{-2+2\eta}+\\ &\quad+\left\{H_{\frac{b}{b+u}}(x)\left(1-H_{\frac{b}{b+u}}(x)\right)\right\}^{-2+2\eta}\Big]\cdot\left|H_{\frac{b}{b+v}}(x)-H_{\frac{b}{b+u}}(x)\right|^2\\ &\leq (v-u)^2\Big[(b+u)^{-2}(b+v)^{-2}b^2\left\{H_{\frac{b}{b+v}}(x)\left(1-H_{\frac{b}{b+v}}(x)\right)\right\}^{-2+2\eta}+\\ &\quad+(b+u)^{-2}(b+v)^{-2}b^2\left\{H_{\frac{b}{b+u}}(x)\left(1-H_{\frac{b}{b+u}}(x)\right)\right\}^{-2+2\eta}\Big];\end{aligned}$$

and since

$$b(b+v)^{-1}F(x)\left(1-F(x)\right) \leq H_{\frac{b}{b+v}}(x)\left(1-H_{\frac{b}{b+v}}(x)\right),$$

and

$$b(b+u)^{-1}F(x)\left(1-F(x)\right) \leq H_{\frac{b}{b+u}}(x)\left(1-H_{\frac{b}{b+u}}(x)\right),$$

we obtain

$$\begin{aligned}\int\int (f_{bv}-f_{bu})^2(x, y)\, dF(x)\, dF(y) &\leq (v-u)^2(b+u)^{-2}\int [t(1-t)]^{-1+2\eta}\, dt\\ &\ll (v-u)\, vb^{-1}(b+u)^{-1}\int [t(1-t)]^{-1+2\eta}\, dt.\end{aligned} \tag{5.11}$$

The square integral of the first summand is bounded by

$$a^2 \int\int (f_{bv}-f_{bu}-f_{av}+f_{au})^2 (x, y)\, dF(x)\, dF(y).$$

We now fix x; the it is easy to check that

$$\left|H_{\frac{b}{b+u}}(x)-H_{\frac{a}{a+u}}(x)\right| \leq (b-a)\frac{u}{(b+u)(a+u)},$$

$$\left|H_{\frac{b}{b+v}}(x)-H_{\frac{a}{a+v}}(x)\right| \leq (b-a)\frac{v}{(b+v)(a+v)},$$

$$\left|H_{\frac{b}{b+v}}(x)-H_{\frac{b}{b+u}}(x)\right| \leq (v-u)\frac{b}{(b+v)(b+u)},$$

and

$$\left|H_{\frac{a}{a+v}}(x)-H_{\frac{a}{a+u}}(x)\right| \leq (v-u)\frac{a}{(a+v)(a+u)}.$$

Thus

$$[(h'(H_{\frac{b}{b+v}}(x))-h'(H_{\frac{b}{b+u}}(x))-h'(H_{\frac{a}{a+v}}(x))+h'(H_{\frac{a}{a+u}}(x))]^2$$

$$\leqslant (v-u)(b-a)\left[\frac{u}{(b+u)(a+u)}+\frac{v}{(b+v)(a+v)}\right]\left[\frac{a}{(a+v)(a+u)}+\right.$$

$$\left.+\frac{b}{(b+v)(b+u)}\right]\Big(\sum_j (H_j(x)(1-H_j(x))\Big)^{-2+2},$$

where the sum extends over $j=b/(b+u)$ and $a/(a+v)$ (cf. Lemma 5.3). Consequently,

$$a^2\int\int(f_{bv}-f_{bu}-f_{av}+f_{au})^2\,dF(x)\,dF(y)$$

$$\leqslant (v-u)(b-a)\Big(\int(t(1-t))^{-1+2\eta}\,dt\Big)\left(\max\left(\frac{a+v}{b+u},\frac{b+u}{a+v}\right)+1\right)$$

since

$$a^2\left[\frac{u}{(b+u)(a+u)}+\frac{v}{(b+v)(a+v)}\right]\left[\frac{b}{(b+v)(b+u)}+\frac{a}{(a+v)(a+u)}\right]$$

$$\leqslant\left(\frac{a}{a+v}\right)^2\left(\frac{a+v}{b+u}+1\right)\text{ and }\leqslant\left(\frac{b}{b+u}\right)^2\left(\frac{b+u}{a+v}+1\right),$$

we have, e.g.,

$$\frac{b}{b+u}\int(I_{[y\leqslant x]}-F(x))^2\,dF(y)=\frac{b}{b+u}F(x)(1-F(x))$$

and

$$\frac{b}{b+u}F(x)(1-F(x))\leqslant H_{\frac{b}{b+u}}(x)[1-H_{\frac{b}{b+u}}(x)].$$

Together with (5.9), (5.10) and (5.11), we obtain

$$\text{(5.12)}\qquad E\Big(\sum_{l=a+1}^{b}\sum_{j=u+1}^{v}\omega_{l,j}\Big)^2\leqslant (b-a)(v-u)\max\left(\frac{a+v}{b+u},\frac{b+u}{a+v}\right).$$

To estimate $E\max\limits_{\substack{1\leqslant n\leqslant N\\0\leqslant m\leqslant N}} N^{-1}(n+m)^{-2}W_{n,m}^2$ we apply Theorem 8 of Moricz [13].

If $2^{k-1}\leqslant a\leqslant b\leqslant 2^k$ and $2^{j-1}\leqslant u\leqslant v<2^j$, then

$$\max\left(\frac{a+u}{b+u},\frac{b+u}{a+v}\right)\leqslant 4.$$

Thus by (5.12) we have for $2^k,\ 2^j \leqslant N$,

$$E \max_{\substack{2^{k-1} \leqslant n < 2^k \\ 2^{j-1} \leqslant m < 2^j}} N^{-1}(n+m)^{-2} W_{n,m}^2 \ll 2^k \cdot 2^j \cdot (2^{k-1}+2^{j-1})^{-2} \cdot N^{-1} \log 2^k \log 2^j$$

$$\ll N^{-1}(\log N)^2.$$

Hence

$$\text{(5.13)} \qquad E \max_{\substack{1 \leqslant n \leqslant N \\ 0 \leqslant m \leqslant N}} N^{-1}(n+m)^{-2} W_{n,m}^2 \ll N^{-1}(\log N)^4.$$

To prove (5.7), it remains to estimate

$$E \max_{\substack{1 \leqslant n \leqslant N \\ 0 \leqslant m \leqslant N}} \Big(N^{-1/2}(n+m)^{-1} \sum_{i=1}^{n} \{h'(H_{\frac{n}{n+m}}(X_i))(1-F(X_i)) -$$

$$- \int h'(H_{\frac{n}{n+m}}(t)) I_{[X_i \leqslant t]} dF(t) + \int h'(H_{\frac{n}{n+m}}(t)) F(t) dF(t)\}\Big)^2.$$

But, since h' is bounded, this maximum is clearly bounded by N^{-1}, proving (5.7). The proof of (5.8) is similar to that of (5.13) and is therefore omitted.

Lemma 5.5. *We have*

$$\lim_{N \to \infty} \sup_{r,s \in [0,1]} [rN] N^{-1/2} \int_0^1 h''(\xi(t)) \left(\frac{N(r,s)}{N(r,s)+1} \hat{H}_{[rN],[sN]}(t) - H_{\lambda_{r,s}}(t)\right)^2 dF_{[rN]}(t)$$

$$= 0 \text{ in probability,}$$

where $\xi(t)$ is any point in the interval given by $H_{\lambda_{r,s}}(t)$ and $\frac{N(r,s)}{N(r,s)+1} \hat{H}_{[rN],[sN]}(t)$.

Proof. Using (5.1), we have

$$\text{(5.14)} \qquad \left| \int_0^1 h''(\xi(t)) \left(\frac{N(r,s)}{N(r,s)+1} \hat{H}_{[rN],[sN]}(t) - H_{\lambda_{r,s}}(t)\right)^2 dF_{[rN]}(t) \right|$$

$$\ll \int_0^1 \left[\left(H_{\lambda_{r,s}}(t)(1-H_{\lambda_{r,s}}(t))\right)^{-1+\eta} + \left(\frac{N(r,s)}{N(r,s)+1} \hat{H}_{[rN],[sN]}(t) \times \right.\right.$$

$$\left.\left. \times \left(1 - \frac{N(r,s)}{N(r,s)+1} \hat{H}_{[rN],[sN]}(t)\right)\right)^{1+2\eta} \right] \left(\frac{N(r,s)}{N(r,s)+1} \hat{H}_{[rN],[sN]}(t) - H_{\lambda_{r,s}}(t)\right)^2 dF_{[rN]}(t)$$

Next,

$$\text{(5.15)} \qquad \int_0^1 \left(\frac{N(r,s)}{N(r,s)+1} \hat{H}_{[rN],[sN]}(t) \left(1 - \frac{N(r,s)}{N(r,s)+1} \hat{H}_{[rN],[sN]}(t)\right)\right)^{-1+\eta} \times$$

$$\times\left(\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)-H_{\lambda_{r,s}}(t)\right)^2 dF_{[rN]}(t)$$

$$\leqslant \sup_t\left(\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)-H_{\lambda_{r,s}}(t)\right)^2 \int\left(\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)\times\right.$$

$$\left.\times\left(1-\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)\right)\right)^{-1+\eta} dF_{[rN]}(t)$$

$$\leqslant \sup_t\left(\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)-H_{\lambda_{r,s}}(t)\right)^2 \frac{N(r,s)+1}{[rN]}\int_0^1 \big(t(1-t)\big)^{-1+\eta}\,dt.$$

Now, from Theorem 4.1 in [7], we have for each r and s

$$\mathrm{E}\left(\sqrt{N(r,s)}\int_0^1 h''(\xi(t))\left(\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)-H_{\lambda_{r,s}}(t)\right)^2 dF_{[rN]}(t)\right)^2 \ll \frac{N(r,s)}{[rN]^2}.$$

Consequently,

$$\text{(5.16)}\quad \mathrm{E}\max_{\substack{1\leqslant [rN]\leqslant N^{1/2}/\log N\\ 0\leqslant [sN]\leqslant N^{1/2}/\log N}}\left([rN]\,N^{-1/2}\int_0^1 h''(\xi(t))\left(\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)-\right.\right.$$

$$\left.\left.-H_{\lambda_{r,s}}(t)\right)^2 dF_{[rN]}(t)\right)^2 \ll (\log N)^{-2}.$$

Now let $\varepsilon>0$ be fixed. Then, using (5.16), we obtain

$$\text{(5.17)}\quad \mathrm{P}\left(\sup_{\substack{1\leqslant n\leqslant N\\ 0\leqslant m\leqslant N}} nN^{-1/2}\int_0^1 |h''(\xi(t))|\left(\frac{n+m}{n+m+1}\hat{H}_{n,m}(t)-H_{\frac{n}{n+m}}(t)\right)^2 dF_n(t)\geqslant\varepsilon\right)$$

$$\ll \varepsilon^{-2}(\log N)^{-2}+\mathrm{P}\Big(\sup_{\substack{N^{1/2}/\log N\leqslant n\leqslant N\\ \text{or } N^{1/2}/\log N\leqslant m\leqslant N}} nN^{-1/2}\int_0^1 |h''(\xi)|\times$$

$$\times\left(\frac{n+m}{n+m+1}\hat{H}_{n,m}(t)-H_{\frac{n}{n+m}}(t)\right)^2 dF_n(t)\geqslant\varepsilon\Big).$$

Now let us fix intervals $[2^{k-1}, 2^k)$ and $[2^{j-1}, 2^j)$. Then, using Lemma 5.3, we have for each $2^{k-1}\leqslant n<2^k$ and $2^{j-1}\leqslant m<2^j$,

$$\left(H_{\frac{n}{n+m}}(t)\left(1-H_{\frac{n}{n+m}}(t)\right)\right)^{-1+\eta}\leqslant\left(H_{\frac{2^k}{2^k+2^j-1}}(t)\left(1-H_{\frac{2^k}{2^k+2^j-1}}(t)\right)\right)^{-1+\eta}+$$

$$+\left(H_{\frac{2^{k-1}}{2^{k-1}+2^j}}(t)\left(1-H_{\frac{2^{k-1}}{2^{k-1}+2^j}}(t)\right)\right)^{-1+\eta}$$

and so,

$$\text{(5.18)}\quad \mathrm{E}\max_{\substack{2^{k-1}\leqslant n\leqslant 2^k\\ 2^{j-1}\leqslant m\leqslant 2^j}}\sum_{i=1}^n\left(H_{\frac{n}{n+m}}(X_i)\left(1-H_{\frac{n}{n+m}}(X_i)\right)\right)^{-1+\eta}$$

$$\leqslant \mathrm{E}\sum_{i=1}^{2^k}\left(\mathrm{H}_{\frac{2^k}{2^k+2^{j-1}}}(X_i)\left(1-H_{\frac{2^k}{2^k+2^{j-1}}}(X_i)\right)\right)^{-1+\eta}+$$

$$+\mathrm{E}\sum_{i=1}^{2^k}\left(H_{\frac{2^{k-1}}{2^{k-1}+2^j}}(X_i)\left(1-H_{\frac{2^{k-1}}{2^{k-1}+2^j}}(X_i)\right)\right)^{-1+\eta}$$

$$\leqslant (2^k+2^j)\int\left(t\,(1-t)\right)^{-1+\eta}dt.$$

Let

$$B_{kj}(N)=\Big\{\max_{\substack{2^{k-1}\leqslant n<2^k\\ 2^{j-1}\leqslant m<2^j}}\sum_{i=1}^{n}\left(H_{\frac{n}{n+m}}(X_i)\left(1-H_{\frac{n}{n+m}}(X_i)\right)\right)^{-1+\eta}\geqslant (2^k+2^j)(\log N)^3\Big\}.$$

Then, using formula (5.18) and Markov's inequality, $\sum_{k,j\leqslant \log N}\mathrm{P}(B_{k,j}(N)) \leqslant (\log N)^{-1}$ and on $\bigcap_{k,j}(B_{kj}(N))^c$, we have for each n, m

$$n\int\left(H_{\frac{n}{n+m}}(t)\left(1-H_{\frac{n}{n+m}}(t)\right)\right)^{-1+\eta}dF_n(t)\leqslant 2(n+m)(\log N)^3. \tag{5.19}$$

Now let

$$A_N=\Big\{\sup_{N^{1/4}\leqslant n\leqslant N}\sup_t n^{3/8}|F_n(t)-F(t)|<1 \text{ and } \sup_{N^{1/4}\leqslant m\leqslant N}\sup_t m^{3/8}|G_m(t)-G(t)|<1\Big\}.$$

Then, by the law of the iterated logarithm for the empirical processes, we have $\lim_{N\to\infty}\mathrm{P}(A_N^c)=0$ and on A_N we have for $n\geqslant N^{1/2}/\log N$ or for $m \geqslant N^{1/2}/\log N$,

$$\left|\frac{n+m}{n+m+1}\hat{H}_{n,m}(t)-H_{\frac{n}{n+m}}(t)\right| \tag{5.20}$$

$$\leqslant\frac{1}{n+m}+\left|\frac{n}{n+m}(F_n(t)-F(t))+\frac{m}{n+m}(G_n(t)-G(t))\right|$$

$$\leqslant (n+m)^{-1}\left[1+n^{5/8}+m^{5/8}+N^{1/4}\right].$$

Using (5.14), (5.15), (5.19), and (5.20), we have on $A_N\cap[\bigcap_{k,j}B_{k,j}(N)^c]$ for $n \geqslant N^{1/2}/\log N$ or $m\geqslant N^{1/2}/\log N$,

$$\left|nN^{-1/2}\int_0^1 h''(\xi(t))\left(\frac{n+m}{n+m+1}\hat{H}_{n,m}(t)-H_{\frac{n}{n+m}}(t)\right)^2dF_n(t)\right|$$

$$\leqslant\frac{n+m+1}{\sqrt{N}}(n+m)^{-2}(1+n^{5/8}+m^{5/8}+N^{1/4})^2+$$

$$+N^{-1/2}(n+m)^{-2}(1+n^{5/8}+m^{5/8}+N^{1/4})^2(n+m)(\log N)^3\leqslant N^{-1/4}(\log N)^3.$$

Thus, with N sufficiently large, we have

$$\left| \sup_{\substack{N^{1/2}/\log N \leqslant n \leqslant N \\ \text{or } N^{1/2}/\log N \leqslant m \leqslant N}} \frac{n}{\sqrt{N}} \int_0^1 h''(\xi(t)) \left(\frac{n+m}{n+m+1} \hat{H}_{n,m}(t) - H_{\frac{n}{n+m}}(t) \right)^2 dF_n(t) \right| < \varepsilon$$

on the set $A_N \cap [\bigcap_{k,j} B_{k,j}(N)^c]$. Since the measure of its complement tends to zero, (5.17) tends to zero as $N \to \infty$. This completes the proof of the Lemma.

LEMMA 5.6. *We have*

$$\lim_{N\to\infty} \mathrm{E} \max_{0 \leqslant r,s \leqslant 1} \left(\frac{[rN]\, N^{-1/2}}{[rN]+[sN]+1} \int_0^1 h'(H_{\lambda_{r,s}}(t)) \hat{H}_{[rN],[sN]}(t)\, dF_{[rN]}(t) \right)^2 = 0.$$

Proof. Since h' is bounded, the proof follows trivially.

Proof of Theorem 5.1. Since h is of class C_2, we obtain from Taylor's theorem

$$\begin{aligned} S_N(h, r, s) &= [rN]\, N^{-1/2} \left[\int_0^1 h\left(\frac{N(r, s)}{N(r, s)+1} \hat{H}_{[rN],[sN]}(t) \right) dF_{[rN]}(t) - \right. \\ &\qquad \left. - \int_0^1 h(H_{\lambda_{r,s}}(t))\, dF(t) \right] \\ &= [rN]\, N^{-1/2} \int_0^1 h(H_{\lambda_{r,s}}(t))\, d(F_{[rN]}(t) - F(t)) + \\ &\quad + [rN]\, N^{-1/2} \int_0^1 h'(H_{\lambda_{r,s}}(t)) \left(\frac{N(r, s)}{N(r, s)+1} \hat{H}_{[rN],[sN]}(t) - H_{\lambda_{r,s}}(t) \right) \times \\ &\qquad \times dF_{[rN]}(t) + \\ &\quad + \frac{1}{2} [rN]\, N^{-1/2} \int_0^1 h''(\xi(t)) \left(\frac{N(r, s)}{N(r, s)+1} \hat{H}_{[rN],[sN]}(t) - H_{\lambda_{r,s}}(t) \right)^2 \times \\ &\qquad \times dF_{[rN]}(t), \end{aligned}$$

where $\xi(t)$ belongs to the interval given by $H_{\lambda_{r,s}}(t)$ and $\frac{N(r, s)}{N(r, s)+1} \hat{H}_{[rN],[sN]}(t)$.

Making some routine computations, we obtain

$$S_N(h, r, s) = \sum_{i=1}^{11} E_{Ni}(r, s),$$

where

$$E_{N1}(r, s) = \frac{r}{r+s} \int_0^1 h'(H_{\lambda(r,s)}(t))\, [rN]\, N^{-1/2} (F_{[rN]}(t) - F(t))\, dF(t) -$$

$$-\int_0^1 [rN]\, N^{-1/2} \big(F_{[rN]}(H^{-1}_{\lambda(r,s)}(t)) - F(H^{-1}_{\lambda(r,s)}(t))\big)\, dh(t),$$

$$E_{N2}(r,s) = \frac{r}{r+s}\int_0^1 h'(H_{\lambda(r,s)}(t))\,[sN]\,N^{-1/2}\big(G_{[sN]}(t)-G(t)\big)\,dF(t),$$

$$E_{N3}(r,s) = \frac{1}{2}[rN]\,N^{-1/2}\int_0^1 h''(\xi(t))\left(\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t) - \right.$$

$$\left. - H_{\lambda_{r,s}}(t)\right)^2 dF_{[rN]}(t),$$

$$E_{N4}(r,s) = [rN]\,N^{-1/2}\int_0^1 h'(H_{\lambda_{r,s}}(t))\frac{1}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)\,dF_{[rN]}(t),$$

$$E_{N5}(r,s) = [rN]\,N^{-1/2}\int_0^1 h'(H_{\lambda_{r,s}}(t))\frac{[rN]}{N(r,s)}\big(F_{[rN]}(t)-F(t)\big)\,d\big(F_{[rN]}(t)-F(t)\big),$$

$$E_{N6}(r,s) = [rN]\,N^{-1/2}\int_0^1 h'(H_{\lambda_{r,s}}(t))\frac{[sN]}{N(r,s)}\big(G_{[sN]}(t)-G(t)\big)\,d\big(F_{[rN]}(t)-F(t)\big),$$

$$E_{N7}(r,s) = [rN]\,N^{-1/2}\int_0^1 \big(h(H_{\lambda_{r,s}}(t)) - h(H_{\lambda(r,s)}(t))\big)\,d\big(F_{[rN]}(t)-F(t)\big),$$

$$E_{N8}(r,s) = [rN]\,N^{-1/2}\int_0^1 \big(h'(H_{\lambda_{r,s}}(t)) - h'(H_{\lambda(r,s)}(t))\big)\frac{r}{r+s}\big(F_{[rN]}(t) -$$

$$- F(t)\big)\,dF(t),$$

$$E_{N9}(r,s) = [sN]\,N^{-1/2}\int_0^1 \big(h'(H_{\lambda_{r,s}}(t)) - h'(H_{\lambda(r,s)}(t))\big)\frac{r}{r+s}\big(G_{[sN]}(t) -$$

$$- G(t)\big)\,dF(t),$$

$$E_{N10}(r,s) = [rN]\,N^{-1/2}\left(\frac{[rN]}{N(r,s)} - \frac{r}{r+s}\right)\int_0^1 h'(H_{\lambda_{r,s}}(t))\big(F_{[rN]}(t)-F(t)\big)\,dF(t),$$

and

$$E_{N11}(r,s) = [sN]\,N^{-1/2}\left(\frac{[rN]}{N(r,s)} - \frac{r}{r+s}\right)\int_0^1 h'(H_{\lambda_{r,s}}(t))\big(G_{[sN]}(t)-G(t)\big)\,dF(t).$$

Now, using Lemmas 5.5, 5.6, 5.4, 5.1, and 5.2, we notice that $E_{Ni} \to 0$ in probability with respect to the uniform metric for $i = 3$ to 9. Furthermore, since

$$[rN]\,N^{-1/2}\left|\frac{[rN]}{N(r,s)} - \frac{r}{r+s}\right| = O(N^{-1/2}),$$

it follows that E_{N10} as well as $E_{N11} \to 0$ in probability with respect to the uniform metric. Thus the weak convergence of $S_N(h)$ will follow from that of $E_{N1}+E_{N2}$.

Since E_{N1} and E_{N2} are independent random functions, we have to show that E_{N1} and E_{N2} converge to the corresponding integrals replacing the empirical processes by independent Kiefer processes. We consider only $E_{N1}(r, s)$ because for $E_{N2}(r, s)$ the same argument applies.

Consider the map ψ: $D([0, 1]^2) \to D([0, 1]^2)$ defined by

$$\psi[f](r, s) = \frac{r}{r+s}\int_0^1 h'\big(H_{\lambda(r,s)}(t)\big) f(r, t)\, dF(t) - \int_0^1 f\big(r, H^{-1}_{\lambda(r,s)}(t)\big)\, dh(t).$$

If f is a continuous function, then clearly ψ is continuous at f with respect to the uniform metric on $D([0, 1]^2)$ because $d|h|$ is a finite measure.

Let $\mathscr{K}$ be a Kiefer process on $C([0, 1]^2)$ with covariance $\mathrm{E}\mathscr{K}(u, v) = 0$ and $E\mathscr{K}(u, v)\mathscr{K}(u', v') = \min(u, u')F(v)\big(1-F(v')\big)$ for $0 \leqslant u$, $u' \leqslant 1$ and $0 \leqslant v \leqslant v' \leqslant 1$.

Since $\mathscr{K}$ has continuous paths and since $[rN]N^{-1/2}[F_{[rN]}(t)-F(t)]$ converges weakly in $D([0, 1]^2)$ to $\mathscr{K}(r, t)$ with respect to the uniform topology, Theorem 5.1 of Billingsley [2] shows that $\psi\big[\big([rN]N^{-1/2}(F_{[rN]}(t) - F(t)\big)\big]$ converges weakly in $D([0, 1]^2)$ to $\psi(\mathscr{K})$ with respect to the uniform topology:

$$\begin{aligned}\psi[\mathscr{K}](r, s) &= \frac{r}{r+s}\int_0^1 h'\big(H_{\lambda(r,s)}(t)\big)\mathscr{K}(r, t)\, dF(t) - \int_0^1 \mathscr{K}\big(r, H^{-1}_{\lambda(r,s)}(t)\big)\, dh(t)\\ &= -\frac{s}{r+s}\int_0^1 \frac{dG}{dH_{\lambda(r,s)}}(t)\,\mathscr{K}(r, t)\, dh\big(H_{\lambda(r,s)}(t)\big).\end{aligned}$$

COROLLARY 5.1. *Set*

$$(5.21)\qquad \bar{S}_N(h, r, s) = \begin{cases} 0 & \text{if } r < 1/N,\\ [rN]N^{-1/2}\left[\int_0^1 h\left(\frac{N(r, s)}{N(r, s)+1}\hat{H}_{[rN],[sN]}(t)\right)dF_{[rN]}(t) - \int_0^1 h\big(H_{\lambda(r,s)}(t)\big)\, dF(t)\right] & \text{otherwise.}\end{cases}$$

Then $\bar{S}_N(h)$ converges weakly in $D([0, 1]^2)$ with respect to the uniform metric to the process $Z(h)$ defined in (4.1).

COROLLARY 5.2. *Let the scores $a_N(i, n, m)$ satisfy* (2.3), *where h satisfies* (5.1). *Then the statistics*

$$(5.22)\qquad \hat{S}_N(h, r, s) = \begin{cases} 0 & \text{if } r < 1/N,\\ [rN]N^{-1/2}\big(T^*_{[rN],[sN]} - \mathrm{E}\,T^*_{[rN],[sN]}\big) & \text{otherwise,}\end{cases}$$

where T^* is defined in (2.4), converge weakly in $D(E)$ with respect to the uniform metric to $Z(h)$.

Proof. Set

$$h_{N,n,m}(t) = a_N(i, n, m) \quad \text{if } t \in \left[\frac{i}{n+m+1}, \frac{i+1}{n+m+1}\right), \; i = 0, \ldots, n+m.$$

Then, by (2.3) and the definition of the Riemann integral, it follows that

$$[rN]\, N^{-1/2} \int_0^1 \left[h\left(H_{\lambda(r,s)}(t)\right) - h_{N,n,m}\left(H_{\lambda(r,s)}(t)\right)\right] dF(t) \to 0$$

uniformly in r and s. Similarly,

$$[rN]\, N^{-1/2} \left(T^*_{[rN],[sN]} - \int_0^1 h\left(\frac{N(r, s)}{N(r, s)+1} \hat{H}_{[rN],[sN]}(t)\right) dF_{[rN]}(t)\right) \to 0$$

uniformly in r and s. The proof follows.

Replacing, in (5.22), $\mathrm{E}\, T^*_{[rN],[sN]}$ by $\int h_{N(r,s),[rN],[sN]}\left(H_{\lambda(r,s)}(t)\right) dF(t)$, the convergence of $\hat{S}_N(h)$ to $Z(h)$ is again in $D([0, 1]^2)$.

6. THE CONTINUITY THEOREM FOR S_N

The random functions $S_N(h, \cdot, \cdot)$ define operators S_N: $\mathscr{H} \to L_2(P, D(E_\varkappa))$ for every $N \geq 1$ and $0 < \varkappa < 1/2$. We shall show in this section that the family $\{S_N: N \geq 1\}$ is uniformly continuous for fixed $\varkappa$. This result will provide an easy argument to extend invariance principles for $S_N(h)$ from nice score functions h to more complicated ones.

THEOREM 6.1. *Let $0 < \varkappa < 1/2$. Then there exists a constant $C(\varkappa)$ such that for every score function $h \in \mathscr{H}$ and every $N \geq 1$*

$$\mathrm{E} \max_{(r,s) \in E_\varkappa} \left(S_N(h, r, s)\right)^2 \leq C(\varkappa) \|h\|^2. \tag{6.1}$$

Remark. This theorem is an extension of a theorem of Denker and Rösler [5], which says that $\mathrm{E}\left(S_N(h, r, s)\right)^2 \leq \text{const} \cdot \|h\|^2$ for every fixed r, $s \in (0, 1)$.

The theorem will be proven by a series of lemmas. If $r < 1/N$, then by definition, $S_N(h, r, s) = 0$ for every s. Also, if $s < 1/N$, then

$$S_{N(h,r,s)} = N^{-1/2} [rN] \left(\frac{1}{[rN]} \sum_{i=1}^{[rN]} h\left(\frac{i}{[rN]+1}\right) - \int_0^1 h(t)\, dt\right) \to 0$$

uniformly in $r \in (0, 1)$ (and $0 < s < 1/N$), since $h \in \mathscr{H}$. Therefore, it suffices to show (6.1) in the case where $r \geq 1/N$ and $s \geq 1/N$. We shall assume in the following lemmas that r, $s \geq 1/N$; especially $(r, s) \in E_\varkappa$ means, more precisely, that $(r, s) \in E_\varkappa$ and r, $s \geq 1/N$.

We begin with an easy observation (cf. Lemma 5.3):

LEMMA 6.1. *If* $(r, s) \in E_\varkappa$, *then*

$$H_{\frac{\varkappa}{2+\varkappa}}(t) \leqslant H_{\lambda_{r,s}}(t) \leqslant H_{\frac{2}{2+\varkappa}}(t) \quad \text{if } G(t) \leqslant F(t) \tag{6.2}$$

and

$$H_{\frac{2}{2+\varkappa}}(t) \leqslant H_{\lambda_{r,s}}(t) \leqslant H_{\frac{\varkappa}{2+\varkappa}}(t) \quad \text{if } F(t) \leqslant G(t), \tag{6.3}$$

where, as before, $\lambda_{r,s} = [rN]/N(r, s)$, $N(r, s) = [rN]+[sN]$.

Proof. Since $\varkappa \leqslant r/(r+s) \leqslant 1-\varkappa$,

$$\frac{[rN]+[sN]}{[rN]} = 1+\frac{[sN]}{[rN]} \leqslant 1+\frac{s}{r}\left(1+\frac{1}{[rN]}\right) \leqslant 2\varkappa^{-1}.$$

Using this estimate, where r and s are exchanged, we also have

$$\frac{[rN]}{[rN]+[sN]} = 1-\frac{[sN]}{[rN]+[sN]} \leqslant 1-\varkappa/2$$

and it follows that

$$\varkappa/2 \leqslant \frac{[rN]}{[rN]+[sN]} \leqslant 1-\varkappa/2. \tag{6.4}$$

Lemma 5.3 states that $F(t) \geqslant G(t)$, $r \leqslant r'$, $s' \leqslant s \Rightarrow H_{r/(r+s)}(t) \leqslant H_{r'/(r'+s')}(t)$ and $F(t) \leqslant G(t)$, $r \leqslant r'$, $s' \leqslant s \Rightarrow H_{r/(r+s)}(t) \geqslant H_{r'/(r'+s')}(t)$.

Assume now that $F(t) \geqslant G(t)$ (the case $F(t) \leqslant G(t)$ is similar). Choose $s' = \varkappa/2$, $r' = 1$. By (6.4), $[sN]/([rN]+[sN]) \geqslant 1-(1-\varkappa/2) = \varkappa/2$ and, consequently, by Lemma 5.3 we have

$$H_{\lambda_{r,s}}(t) \leqslant H_{\frac{r'}{r'+s'}}(t) = H_{\frac{2}{2+\varkappa}}(t).$$

On the other hand, choosing $s = 1$ and $r = \varkappa/2$, we obtain $H_{\varkappa/(2+\varkappa)}(t) \leqslant H_{\lambda_{r',s'}}(t)$.

LEMMA 6.2. *For every* $0 < z < 1$ *and every* $0 < \varkappa < 1/2$, *we have*

$$\max_{(r,s)\in E_\varkappa} \left|F_{[rN]}\left(H_{\lambda_{r,s}}^{-1}(z)\right) - F\left(H_{\lambda_{r,s}}^{-1}(z)\right)\right| \leqslant \max_{0<r\leqslant 1} \max_{y\in I_\varkappa(z)} |F_{[rN]}(y)-F(y)| \tag{6.5}$$

and

$$\max_{(r,s)\in E_\varkappa} \left|G_{[sN]}\left(H_{\lambda_{r,s}}^{-1}(z)\right) - G\left(H_{\lambda_{r,s}}^{-1}(z)\right)\right| \leqslant \max_{0<s\leqslant 1} \max_{y\in I_\varkappa(z)} |G_{[sN]}(y)-G(y)|, \tag{6.6}$$

where

$$I_\varkappa(z) = \left\{y\colon \min\{H^{-1}_{\frac{2}{2+\varkappa}}(z), H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z)\} \leqslant y \leqslant \max\{H^{-1}_{\frac{2}{2+\varkappa}}(z), H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z)\}\right\}.$$

Proof. Let $F(H^{-1}_{\varkappa/(2+\varkappa)}(z)) \geqslant G(H^{-1}_{\varkappa/(2+\varkappa)}(z))$. Then, by Lemma 6.1, $H_{\lambda_{r,s}}(H^{-1}_{\varkappa/(2+\varkappa)}(z)) \geqslant z$, i.e. $H^{-1}_{\varkappa/(2+\varkappa)}(z) \geqslant H^{-1}_{\lambda_{r,s}}(z)$. Since $H_{\lambda_{r,s}}(H^{-1}_{\lambda_{r,s}}(z)) \geqslant H_{\varkappa/(2+\varkappa)}(H^{-1}_{\lambda_{r,s}}(z))$, we have

$$\left(\frac{[rN]}{[rN]+[sN]}-\frac{\varkappa}{2+\varkappa}\right)F(H^{-1}_{\lambda_{r,s}}(z)) \geqslant \left(\frac{2}{2+\varkappa}-\frac{[sN]}{[rN]+[sN]}\right)G(H^{-1}_{\lambda_{r,s}}(z));$$

equivalently, $F(H^{-1}_{\lambda_{r,s}}(z)) \geqslant G(H^{-1}_{\lambda_{r,s}}(z))$, since by (6.4),

$$\frac{[rN]}{[rN]+[sN]}-\frac{\varkappa}{2+\varkappa} > 0.$$

It follows (by Lemma 6.1) that $H_{2/(2+\varkappa)}(H^{-1}_{\lambda_{r,s}}(z)) \geqslant z$ and $H^{-1}_{\lambda_{r,s}}(z) \geqslant H^{-1}_{2/(2+\varkappa)}(z)$.

LEMMA 6.3. *For every $0 < \varkappa < 1/2$ there exists a constant $C_0(\varkappa)$ such that, for every $0 < z < 1$,*

$$\text{(6.7)}\qquad \mathrm{E}\max_{(r,s)\in E_\varkappa} [Nr]^2\left(F_{[Nr]}(H^{-1}_{\lambda_{r,s}}(z)) - F(H^{-1}_{\lambda_{r,s}}(z))\right)^2 \leqslant C_0(\varkappa)Nz(1-z)$$

and

$$\text{(6.8)}\qquad \mathrm{E}\max_{(r,s)\in E_\varkappa} [Ns]^2\left(G_{[Ns]}(H^{-1}_{\lambda_{r,s}}(z)) - G(H^{-1}_{\lambda_{r,s}}(z))\right)^2 \leqslant C_0(\varkappa)Nz(1-z).$$

Proof. Assume that $G(H^{-1}_{\varkappa/(2+\varkappa)}(z)) \leqslant F(H^{-1}_{\varkappa/(2+\varkappa)}(z))$. We shall show that (6.7) and (6.8) hold in this case. (The other case will follow analogously). Under this assumption, we have $I_\varkappa(z) = [H^{-1}_{2/(2+\varkappa)}(z), H^{-1}_{\varkappa/(2+\varkappa)}(z)]$ (cf. Lemma 6.2) and so

$$\text{(6.9)}\qquad \mathrm{E}\max_{(r,s)\in E_\varkappa} [Ns]^2\left(G_{[Ns]}(H^{-1}_{\lambda_{r,s}}(z)) - G(H^{-1}_{\lambda_{r,s}}(z))\right)^2 \leqslant \mathrm{E}\max_{1\leqslant k\leqslant N}\max_{y\in I_\varkappa(z)} k^2(G_k(y)-G(y))^2.$$

Since $\{\max_{y\in I_\varkappa(z)} k^2(G_k(y)-G(y))^2 \colon 1 \leqslant k \leqslant N\}$ is a nonnegative submartingale, it follows from Doob's inequality ([8], p. 314) that

$$\text{(6.10)}\qquad \mathrm{E}\max_{1\leqslant k\leqslant N}\max_{y\in I_\varkappa(z)} k^2(G_k(y)-G(y))^2 \leqslant 4\,\mathrm{E}\max_{y\in I_\varkappa(z)} N^2(G_N(y)-G(y))^2.$$

Moreover, since

$$\left\{N\frac{G_N(y)-G(y)}{1-G(y)} \colon H^{-1}_{\frac{2}{2+\varkappa}}(z) \leqslant y \leqslant H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z)\right\}$$

is a martingale, Doob's inequality again implies that

$$\mathrm{E}\max_{y\in I_\varkappa(z)} N^2(G_N(y)-G(y))^2 \leqslant \mathrm{E}\max_{y\in I_\varkappa(z)} N^2\left(\frac{G_N(y)-G(y)}{1-G(y)}\right)^2\left(1-G(H^{-1}_{\frac{2}{2+\varkappa}}(z))\right)^2$$

$$\leqslant 4N^2\,\mathrm{E}\big(G_N(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))-G(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))\big)^2\left(\frac{1-G(H^{-1}_{2/(2+\varkappa)}(z))}{1-G(H^{-1}_{\varkappa/(2+\varkappa)}(z))}\right)^2$$

$$=4NG(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))\big(1-G(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))\big)\left(\frac{1-G(H^{-1}_{2/(2+\varkappa)}(z))}{1-G(H^{-1}_{\varkappa/(2+\varkappa)}(z))}\right)^2.$$

Note that $1-G(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))\neq 0$, since otherwise

$$z=\frac{\varkappa}{2+\varkappa}F(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))+\frac{2}{2+\varkappa}G(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))\geqslant G(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))=1$$

because of $F(H^{-1}_{\varkappa/(2+\varkappa)}(z))\geqslant G(H^{-1}_{\varkappa/(2+\varkappa)}(z))$.

Using

$$H_{\frac{\varkappa}{2+\varkappa}}(y)\big(1-H_{\frac{\varkappa}{2+\varkappa}}(y)\big)\geqslant\frac{2}{2+\varkappa}G(y)\big(1-G(y)\big),$$

$$1-z=1-\frac{\varkappa}{2+\varkappa}F(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))-\frac{2}{2+\varkappa}G(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))\leqslant 1-G(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z)),$$

and

$$\big(1-G(H^{-1}_{\frac{2}{2+\varkappa}}(z))\big)\frac{\varkappa}{2+\varkappa}\leqslant 1-z,$$

we note the left-hand side of (6.8) is bounded by

$$16N\frac{2+\varkappa}{2}z(1-z)\left(\frac{\frac{2+\varkappa}{\varkappa}(1-z)}{1-z}\right)^2=8\frac{(2+\varkappa)^3}{\varkappa^2}Nz(1-z).$$

The proof of (6.7) is more involved. Assume first that also $G(H^{-1}_{2/(2+\varkappa)}(z))\leqslant F(H^{-1}_{2/(2+\varkappa)}(z))$. Formula (6.10) remains valid replacing G by F. Then use the fact that

$$\left\{N\frac{1-F_N(y)-(1-F(y))}{F(y)}:\ H^{-1}_{\frac{2}{2+\varkappa}}(z)\leqslant y\leqslant H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z)\right\}$$

is a martingale to obtain

$$\mathrm{E}\max_{y\in I_\varkappa(z)}N^2\big(F_N(y)-F(y)\big)^2\leqslant \mathrm{E}\max_{y\in I_\varkappa(z)}N^2\frac{\big(1-F_N(y)-(1-F(y))\big)^2}{F^2(y)}F(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z))^2$$

$$\leqslant 4N^2E\Big(1-F_N(H^{-1}_{\frac{2}{2+\varkappa}}(z))-\big(1-F(H^{-1}_{\frac{2}{2+\varkappa}}(z))\big)\Big)^2\left(\frac{F(H^{-1}_{\varkappa/(2+\varkappa)}(z))}{F(H^{-1}_{2/(2+\varkappa)}(z))}\right)^2$$

$$=4NF(H^{-1}_{\frac{2}{2+\varkappa}}(z))\big(1-F(H^{-1}_{\frac{2}{2+\varkappa}}(z))\big)\left(\frac{F(H^{-1}_{\varkappa/(2+\varkappa)}(z))}{F(H^{-1}_{2/(2+\varkappa)}(z))}\right)^2.$$

(Note that $F(H^{-1}_{\frac{2}{2+\varkappa}}(z)) \neq 0$, since otherwise

$$z = \frac{2}{2+\varkappa} F(H^{-1}_{\frac{2}{2+\varkappa}}(z)) + \frac{\varkappa}{2+\varkappa} G(H^{-1}_{\frac{2}{2+\varkappa}}(z)) \leqslant F(H^{-1}_{\frac{2}{2+\varkappa}}(z)) = 0).$$

We also have

$$H_{\frac{2}{2+\varkappa}}(y)(1-H_{\frac{2}{2+\varkappa}}(y)) \geqslant \frac{2}{2+\varkappa} F(y)(1-F(y)),$$

$$z = \frac{2}{2+\varkappa} F(H^{-1}_{\frac{2}{2+\varkappa}}(z)) + \frac{\varkappa}{2+\varkappa} G(H^{-1}_{\frac{2}{2+\varkappa}}(z)) \leqslant F(H^{-1}_{\frac{2}{2+\varkappa}}(z))$$

$$\text{and} \quad \frac{\varkappa}{2+\varkappa} F(H^{-1}_{\frac{\varkappa}{2+\varkappa}}(z)) \leqslant z.$$

Consequently, the left-hand side of (6.7) is bounded by

$$16N\frac{2+\varkappa}{2} z(1-z)\left(\frac{\frac{2+\varkappa}{\varkappa} z}{z}\right)^2 = 8\frac{N(2+\varkappa)^3}{\varkappa^2} z(1-z).$$

We still have to prove formula (6.7) in the case where $G(H^{-1}_{2/(2+\varkappa)}(z)) > F(H^{-1}_{2/(2+\varkappa)}(z))$. The function $u \to (F-G)(H_u^{-1}(z))$ is continuous in u, where $\varkappa/(2+\varkappa) \leqslant u \leqslant 2/(2+\varkappa)$. By our present assumptions, there exists a u_0 such that $(F-G)(H_{u_0}^{-1}(z)) = 0$. Splitting the interval $I_\varkappa(z) = [H^{-1}_{2/(2+\varkappa)}(z), H^{-1}_{\varkappa/(2+\varkappa)}(z)]$ into two subintervals $J_1 = [H^{-1}_{2/(2+\varkappa)}(z), H^{-1}_{u_0}(z)]$ and $J_2 = [H^{-1}_{u_0}(z), H^{-1}_{\varkappa/(2+\varkappa)}(z)]$, we can argue on J_1 as previously in the proof of (6.8) for G and on J_2 as before. This completes the proof.

LEMMA 6.4. *For every* $0 < \varkappa < {}^1/_2$ *there exists a constant* $C_1(\varkappa)$ *such that for every* $h \in \mathscr{H}$ *and every* $N \geqslant 1$,

$$\mathrm{E}\left\{\max_{(r,s)\in E_\varkappa} \frac{[rN]}{\sqrt{N}}\left[\int_0^1 h\left(\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)\right) dF_{[rN]}(t) - \int_0^1 h(H_{\lambda_{r,s}}(t))\, dF_{[rN]}(t)\right]\right\}^2 \leqslant C_1(\varkappa)\|h\|^2.$$

Proof. We may assume that h is non-decreasing. Using

$$h(z) = \left\{\begin{array}{ll} \int I_{(1/2,z)}(y)\, dh(y) & (z > {}^1/_2) \\ -\int I_{(z,1/2)}(y)\, dh(y) & (z < {}^1/_2) \end{array}\right\} = \int_{1/2}^{z} dh(y),$$

we have

$$\int_0^1 \left[h\left(\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[Nr],[Ns]}(t)\right) - h(H_{\lambda_{r,s}}(t))\right] dF_{[rN]}(t)$$

$$= \int_0^1 \left\{ \int_{1/2}^{\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[Nr],[sN]}(t)} dh(z) - \int_{1/2}^{H_{\lambda_{r,s}}(t)} dh(z) \right\} dF_{[rN]}(t)$$

$$= \int_0^1 \{ I_{\{H_{\lambda_{r,s}}(t) \leqslant \frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)\}} \int_{H_{\lambda_{r,s}}(t)}^{\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)} dh(z) -$$

$$- I_{\{H_{\lambda_{r,s}}(t) \leqslant \frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)\}} \int_{\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)}^{H_{\lambda_{r,s}}(t)} dh(z) \} dF_{[rN]}(t)$$

$$= \int_0^1 \int_0^1 \{ I_{\{H_{\lambda_{r,s}}(t) \leqslant \frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)\}} I_{\{\hat{H}^{-1}_{[rN],[sN]}(\frac{N(r,s)+1}{N(r,s)}z) \leqslant t < H^{-1}_{\lambda_{r,s}}(z)\}} -$$

$$- I_{\{H_{\lambda_{r,s}}(t) > \frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)\}} I_{\{H^{-1}_{\lambda_{r,s}}(z) \leqslant t < \hat{H}^{-1}_{[rN],[sN]}(\frac{N(r,s)+1}{N(r,s)}z)\}} dF_{[rN]}(t)\, dh(z)$$

$$= I_1 + I_2,$$

where I_1 denotes the above integral with integration over $0 < z \leqslant 2N/(2N+1)$ and where I_2 denotes the above integral with integration over $2N/(2N+1) < z < 1$.

Let us first estimate I_2. Since $z > 2N/(2N+1)$, we have $\big((N(r,s)+1)/N(r,s)\big)z > 1$, and hence

$$I_{\{\hat{H}^{-1}_{[rN],[sN]}(\frac{N(r,s)+1}{N(r,s)}z) \leqslant t < H^{-1}_{\lambda_{r,s}}(z)\}} \equiv 0.$$

It follows that

$$|[rN]\, I_2| = \Big| - \int_{\frac{2N}{2N+1}}^{1} \int_0^1 I_{\{H^{-1}_{\lambda_{r,s}}(z) \leqslant t\}} [rN]\, dF_{[rN]}(t)\, dh(z) \Big|$$

$$\leqslant \int_{\frac{2N}{2N+1}}^{1} \int_0^1 I_{\{H^{-1}_{\lambda_{r,s}}(z) \leqslant t\}} dN(r,s)\, \hat{H}_{[rN],[sN]}(t)\, dh(z)$$

$$= N(r,s) \int_{\frac{2N}{2N+1}}^{1} \big[1 - \hat{H}_{[rN],[sN]}\big(H^{-1}_{\lambda_{r,s}}(z)\big)\big]\, dh(z)$$

$$\leqslant N(r,s) \int_{\frac{2N}{2N+1}}^{1} (1-z)\, dh(z) + N(r,s) \int_{\frac{2N}{2N+1}}^{1} \big|\hat{H}_{[rN],[sN]}\big(H^{-1}_{\lambda_{r,s}}(z)\big) - z\big|\, dh(z)$$

$$\leqslant \sqrt{2}\frac{N(r,s)}{\sqrt{N}} \|h\| + N(r,s) \int_{\frac{2N}{2N+1}}^{1} \big|\hat{H}_{[rN],[sN]}\big(H^{-1}_{\lambda_{r,s}}(z)\big) - z\big|\, dh(z),$$

since

$$1 - z \leqslant \frac{1}{\sqrt{2N}}\sqrt{1-z} \leqslant \frac{1}{\sqrt{N}}\sqrt{z(1-z)} \quad \text{for } z \geqslant \frac{2N}{2N+1}.$$

We now estimate I_1. Let

$$\varphi_{r,s}(t,z)=\begin{cases}1 & \text{if } H_{\lambda_{r,s}}(t)<z\leqslant \frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t),\\ -1 & \text{if } \frac{N(r,s)}{N(r,s)+1}\hat{H}_{[rN],[sN]}(t)<z\leqslant H_{\lambda_{r,s}}(t),\\ 0 & \text{otherwise.}\end{cases}$$

Then

$$|[rN]I_1|=\Big|\int_0^{\frac{2N}{2N+1}}\int_0^1\varphi_{r,s}(t,z)\,d[rN]F_{[rN]}(t)\,dh(z)\Big|$$

$$\leqslant N(r,s)\int_0^{\frac{2N}{2N+1}}\int_0^1|\varphi_{r,s}(t,z)|\,d\hat{H}_{[rN],[sN]}(t)\,dh(z)$$

$$=N(r,s)\int_0^{\frac{2N}{2N+1}}\Big|\hat{H}_{[rN],[sN]}(H^{-1}_{\lambda_{r,s}}(z))-$$

$$-\hat{H}_{[rN],[sN]}\Big(\hat{H}^{-1}_{[rN],[sN]}\Big(\frac{N(r,s)+1}{N(r,s)}z\Big)-\Big)\Big|\,dh(z).$$

It is easy to see that

$$\Big|\hat{H}_{[rN],[sN]}\Big(\hat{H}^{-1}_{[rN],[sN]}\Big(\frac{N(r,s)+1}{N(r,s)}z\Big)-\Big)-\frac{N(r,s)+1}{N(r,s)}z\Big|\leqslant 2\min\Big\{z,\frac{1}{N(r,s)}\Big\},$$

$$2\int_0^{\frac{2N}{2N+1}}N(r,s)\min\Big\{z,\frac{1}{N(r,s)}\Big\}dh(z)\leqslant 2\sqrt{5}\sqrt{N}\,\|h\|,$$

$$\int_0^{\frac{2N}{2N+1}}z\,dh(z)\leqslant\sqrt{5}\sqrt{N}\,\|h\|.$$

It follows that

$$|[rN]I_1|\leqslant 3\sqrt{5}\sqrt{N}\,\|h\|+N(r,s)\int_0^{\frac{2N}{2N+1}}\big|\hat{H}_{[rN],[sN]}(H^{-1}_{\lambda_{r,s}}(z))-z\big|\,dh(z),$$

and so

$$\Big|[rN]\int_0^1\Big[h\Big(\frac{N(r,s)}{N(r,s)+1}\hat{H}_{[Nr],[Ns]}(t)\Big)-h(H_{\lambda_{r,s}}(t))\Big]dF_{[rN]}(t)\Big|$$

$$\leqslant 4\sqrt{5}\sqrt{N}\,\|h\|+\int_0^1 N(r,s)\big|\hat{H}_{[rN],[sN]}(H^{-1}_{\lambda_{r,s}}(z))-z\big|\,dh(z).$$

By Lemma 6.3 it follows finally that

$$\Big\| \max_{(r,s)\in E_\varkappa} \frac{[rN]}{\sqrt{N}} \int_0^1 \Big[h\Big(\frac{N(r,s)}{N(r,s)+1} \hat{H}_{[Nr],[Ns]}(t)\Big) - h(H_{\lambda_{r,s}}(t))\Big] dF_{[rN]}(t) \Big\|_{L_2(P)}$$

$$\leqslant 4\sqrt{5}\,\|h\| + N^{-1/2} \Big\| \int_0^1 \max_{(r,s)\in E_\varkappa} |\hat{H}_{[rN],[sN]}(H_{\lambda_{r,s}}^{-1}(z)) - z|\, dh(z) \Big\|_{L_2(P)}$$

$$\leqslant 4\sqrt{5}\,\|h\| + \frac{1}{N}\int_0^1 \Big(\| \max_{(r,s)\in E_\varkappa} [rN](F_{[rN]}(H_{\lambda_{r,s}}^{-1}(z)) - F(H_{\lambda_{r,s}}^{-1}(z))\|_{L_2(P)} +$$

$$+ \| \max_{(r,s)\in E_\varkappa} [sN](G_{[sN]}(H_{\lambda_{r,s}}^{-1}(z)) - G(H_{\lambda_{r,s}}^{-1}(z))\|_{L_2(P)}\Big) dh(z)$$

$$\leqslant 4\sqrt{5}\,\|h\| + 2C_0^{1/2}(\varkappa) \int_0^1 \sqrt{z(1-z)}\, dh(z) = C_1^{1/2}(\varkappa)\,\|h\|.$$

LEMMA 6.5. *For every* $0 < \varkappa < 1/2$ *there exists a constant* $C_2(\varkappa)$ *such that for every* $h \in \mathscr{H}$ *and every* $N \geqslant 1$,

$$E\Big\{ \max_{(r,s)\in E_\varkappa} \frac{[rN]}{\sqrt{N}} \int_0^1 h(H_{\lambda_{r,s}}(t))\, d(F_{[rN]} - F)(t) \Big\}^2 \leqslant C_2(\varkappa)\,\|h\|^2. \tag{6.12}$$

Proof. Let h be increasing. Then, integrating by parts, we have

$$\int_0^1 h(H_{\lambda_{r,s}}(t))\, d(F_{[rN]} - F)(t) = -\int_0^1 \big(F_{[rN]}(H_{\lambda_{r,s}}^{-1}(t)) - F(H_{\lambda_{r,s}}^{-1}(t))\big)\, dh(z).$$

It follows from Lemma 6.3 that

$$\Big(\mathrm{E}\Big\{ \max_{(r,s)\in E_\varkappa} \frac{[rN]}{\sqrt{N}} \int_0^1 h(H_{\lambda_{r,s}}(t))\, d(F_{[rN]} - F)(t) \Big\}^2 \Big)^{1/2}$$

$$\leqslant \int_0^1 \Big\| \max_{(r,s)\in E_\varkappa} \frac{[rN]}{\sqrt{N}} (F_{[rN]}(H_{\lambda_{r,s}}^{-1}(t)) - F(H_{\lambda_{r,s}}^{-1}(t)))\Big\|_{L_2(P)} dh(z)$$

$$\leqslant C_2^{1/2}(\varkappa)\,\|h\|.$$

The proof of Theorem 6.1 follows from Lemmas 6.4 and 6.5, since

$$(\mathrm{E} \max_{(r,s)E_\varkappa} S_N^2(h,r,s))^{1/2} = \Big\| \max_{(r,s)\in E_\varkappa} \Big\{ \frac{[rN]}{\sqrt{N}} \int_0^1 \Big[h\Big(\frac{N(r,s)}{N(r,s)+1} \hat{H}_{[rN],[sN]}(t)\Big) -$$

$$- h(H_{\lambda_{r,s}}(t))\Big] dF_{[rN]}(t) + \frac{[rN]}{\sqrt{N}} \int_0^1 h(H_{\lambda_{r,s}}(t))\, d(F_{[rN]} - F)(t) \Big\} \Big\|_{L_2(P)}$$

$$\leqslant (C_1^{1/2}(\varkappa) + C_2^{1/2}(\varkappa))\,\|h\|.$$

7. AN INVARIANCE PRINCIPLE FOR ABSOLUTELY CONTINUOUS SCORES

In this section we list some (more or less) immediate corollaries to Theorems 5.1 and 6.1 Replacing Theorem 5.1 in the arguments below leads to general invariance principles to be proved in section 8.

THEOREM 7.1. *Let $h \in \mathscr{H}$ be an absolutely continuous score function. Then $S_N(h)$ $(N \geq 1)$ converges weakly in $(D(E), \delta)$ to the process $Z(h)$ defined in formula* (4.1).

Proof. As remarked in section 2, the score functions of class C^2 with bounded continuous second derivatives are dense in the space of all absolutely continuous score functions. By Theorem 5.1, for such a score function h, $S_N(h) \to Z(h)$ weakly in $D([0, 1]^2)$ equipped with the uniform topology. But then, the restrictions of $S_N(h)$ to E converge weakly with respect to δ (see section 3). Hence we have established the theorem for a dense subspace of absolutely continuous score functions.

Now let $h \in \mathscr{H}$ be an arbitrary absolutely continuous score function. Given $\varepsilon > 0$, choose a C^2-function h_ε with continuous bounded second derivative such that $\|h - h_\varepsilon\| < \varepsilon$. Since $S_N(h) - S_N(h_\varepsilon) = S_N(h - h_\varepsilon)$, Theorem 6.1 implies (cf. 3.4)) that

$$\begin{aligned}\left(\mathrm{E}\left[\delta\left(S_N(h), S_N(h_\varepsilon)\right)\right]^2\right)^{1/2} &= \left(\mathrm{E}\left[\sum_{k=2}^{\infty} \min\left(2^{-k}, \delta_{1/k}\left(S_N(h), S_N(h_\varepsilon)\right)\right)\right]^2\right)^{1/2} \\ &\leq \sum_{k=2}^{\infty} \min\left(2^{-k}, \left(\mathrm{E}\left[\delta_{1/k}\left(S_N(h), S_N(h_\varepsilon)\right)\right]^2\right)^{1/2}\right) \\ &\leq \sum_{k=2}^{\infty} \min\left(2^{-k}, \left(C(1/k)\|h - h_\varepsilon\|^2\right)^{1/2}\right) \\ &\leq \sum_{k=2}^{\infty} \min\left(2^{-k}, \left(C(1/k)\right)^{1/2} \varepsilon\right).\end{aligned}$$

Since the constants $C(1/k)$ do not depend on the functions in $\mathscr{H}$, there exists a $k(\varepsilon)$ with

$$\left(C(1/k)\right)^{1/2} \varepsilon < 2^{-k(\varepsilon)}\left[k(\varepsilon)\right]^{-1} \quad \text{for all } k \leq k(\varepsilon).$$

Hence

$$\left(\mathrm{E}\left[\delta\left(S_N(h), S_N(h_\varepsilon)\right)\right]^2\right)^{1/2} \leq 2^{-k(\varepsilon)} + \sum_{k > k(\varepsilon)} 2^{-k} = 2^{-k(\varepsilon)+1},$$

which tends to zero as $\varepsilon \to 0$.

Note that (3.6) is satisfied with $a(\varepsilon) = 2^{-2k(\varepsilon)+2}$ and (3.5) holds because of the previous discussion.

It follows from Proposition 4.3 that $Z(h_\varepsilon) \to Z(h)$ in probability. Therefore, by Lemma 3.3, $S_N(h)$ converge weakly to $Z(h)$.

The following result is an equivalent formulation of Theorem 1.

THEOREM 7.2. *For $n, m \in N$, define*

$$\tilde{S}_{n,m}(h, r, s)$$
$$= \frac{[rn]}{\sqrt{n+m}}\left(\int_0^1 h\left(\frac{[rn]+[sm]}{[rn]+[sm]+1}\hat{H}_{[rn],[sm]}(t)\right)dF_{[rn]}(t) - \int_0^1 h\left(H_{\frac{[rn]}{[rn]+[sm]}}(t)\right)dF(t)\right).$$

If $h \in \mathscr{H}$ (cf. section 4), then the weak convergence of $S_N(h)$ to $Z(h)$ in $(D(E), \delta)$ is equivalent to the weak convergence of $\tilde{S}_{n,m}(h)$ to $\tilde{Z}(h)$ in $(D(E), \delta)$ as $n/(m+n)$ converges to some $0 < \lambda < 1$, where the distribution of $\tilde{Z}(h)$ is given by

$$\left(\tilde{Z}(h, r, s)\right)_{0<r,s<1} \stackrel{D}{=} \left(Z(h, \lambda r, (1-\lambda)s)\right)_{0<r,s<1}.$$

Proof. Given $n, m \in N$, put $N = n+m$ and $\lambda_N = n/N$. Assume that $(S_N(h), \lambda_N) \to (Z(h), \lambda)$ weakly with respect to $\delta \times |\ |$. Since the map ψ: $D(E) \times [0, 1] \to D(E)$, defined by $\psi(f, u)(r, s) = f(ur, us)$, is continuous at every pair $(f, u) \in C(E) \times [0, 1]$, $\tilde{S}_{n,m}(h) = \psi(S_N(h), \lambda_N)$ converges weakly in $(D(E), \delta)$ to $\psi(Z(h), \lambda) = \tilde{Z}(h)$ (see Theorem 5.1 in [2]).

Conversely, if $\tilde{S}_{n,m}(h) \to \tilde{Z}(h)$ as $n/(m+n) \to \lambda$, for $N \in N$ choose n, m such that $(N/n) \to \lambda_1 \in (0, 1)$ and $(N/m) \to (\lambda\lambda_1/(1-\lambda))$. Similarly as before, it follows that

$$S_N(h, r, s) = \sqrt{\frac{n+m}{N}}\tilde{S}_{nm}\left(b, \frac{Nr}{n}, \frac{Ns}{m}\right) \to \frac{1}{\sqrt{\lambda_1 \lambda}} Z(h, \lambda_1 \lambda r, \lambda_1 \lambda s)$$

and the theorem follows since $a^{-1/2} Z(h, ar, as)$ has the same distribution as $Z(h)$.

COROLLARY 7.1. *Let $h \in \mathscr{H}$ be absolutely continuous and let $\bar{S}_N(h)$ be defined as in (5.21). Then $\bar{S}_N(h)$ converges weakly in $(D(E), \delta)$ to the process $Z(h)$ defined in (4.1).*

Proof. This follows from Theorem 7.1 and Lemma 3.2 together with the estimate on p. 65 of [5] for each $E_\varkappa$:

$$\sup_{r,s \in E_\varkappa} N^{-1/2}[rN]\left|\int_0^1 \left(h\left(H_{\lambda(r,s)}(t)\right) - h\left(H_{\lambda_{r,s}}(t)\right)\right)dF(t)\right| \leqslant \varkappa^{-2} N^{-1/2}\|h\|.$$

COROLLARY 7.2. *Let the scores $a_N(i, m, n)$ satisfy (2.3), where $h \in \mathscr{H}$ is absolutely continuous. Define the statistics $\hat{S}_N(h)$ as in (5.22). Then $\hat{S}_N(h)$ converges weakly in $(D(E), \delta)$ to $Z(h)$ defined in (4.1).*

Proof. The proof immediately follows from Theorem 7.1 together with the arguments used to prove Corollary 5.2.

8. WEAK INVARIANCE PRINCIPLES

In this section we prove similar to Theorem 7.1 invariance principles for certain score functions $h \in \mathscr{H}$. The proofs of the theorems appearing here are independent of sections 5 and 7. We just replace Theorem 5.1 by Proposition 8.3 below, which gives the invariance principle for score functions with finite total variation norm on (0, 1). Denote by $\hat{H}_{n,m}^{-1}$ the left continuous inverse of the distribution function $\hat{H}_{n,m}$ $(n, m \geqslant 1)$.

LEMMA 8.1. *Let $h \in \mathscr{H}$. Then for $(r, s) \in E_\varkappa$ with $r, s \geqslant 1/N$ we have*

$$(8.1)\quad S_N(h, r, s) = \frac{-[rN]}{\sqrt{N}} \int_0^1 \Big\{(1-\lambda_{r,s}) \frac{G(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - G(H_{\lambda_{r,s}}^{-1}(t))}{H_{\lambda_{r,s}}(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - t} \times$$
$$\times \big(F_{[rN]}(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - F(\hat{H}_{[rN],[sN]}^{-1}(\hat{t}))\big) -$$
$$-(1-\lambda_{r,s}) \frac{F(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - F(H_{\lambda_{r,s}}^{-1}(t))}{H_{\lambda_{r,s}}(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - t} \big(G_{[sN]}(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - G(\hat{H}_{[rN],[sN]}^{-1}(\hat{t}))\big) +$$
$$+ \frac{F(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - F(H_{\lambda_{r,s}}^{-1}(t))}{H_{\lambda_{r,s}}(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - t} (\hat{t} - t)\Big\} dh(t),$$

where

$$(8.2)\quad \hat{t} = \frac{i+1}{N(r, s)} \quad \text{if } \frac{i}{N(r, s)+1} < t \leqslant \frac{i+1}{N(r, s)+1} \quad (1 \leqslant i \leqslant N),\ (r, s) \in E_\varkappa,$$

and where the integrand is defined to be zero if

$$(8.3)\quad H_{\lambda_{r,s}}(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - t = 0.$$

Remark. The integrand (8.1) is a measurable function as an element in $D([0, 1]^3)$ equipped with the supremum metric due to using $\hat{t}$ instead of t. This will become clear in section 9.

Proof. By the definition of $S_N(h)$, (8.2) and integrating by parts we observe that

$$\frac{\sqrt{N}}{[rN]} S_N(h, r, s) = \int_0^1 h(t)\, dF_{[rN]}(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - \int_0^1 h(t)\, dF(H_{\lambda_{r,s}}^{-1}(t))$$
$$= -\int_0^1 \big(F_{[rN]}(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - F(H_{\lambda_{r,s}}^{-1}(t))\big) dh(t).$$

Use the identity

$$1 = \lambda_{r,s} \frac{F(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - F(H_{\lambda_{r,s}}^{-1}(t))}{H_{\lambda_{r,s}}(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - t} + (1-\lambda_{r,s}) \frac{G(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - G(H_{\lambda_{r,s}}^{-1}(t))}{H_{\lambda_{r,s}}(\hat{H}_{[rN],[sN]}^{-1}(\hat{t})) - t}$$

and the convention (8.3) to obtain

$$\frac{\sqrt{N}}{[rN]} S_N(h, r, s) = -\int_0^1 \Big\{(1-\lambda_{r,s}) \frac{G(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - G(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} \times$$

$$\times [F_{[rN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))] +$$

$$+\lambda_{r,s} \frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} [F_{[rN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))] +$$

$$+ F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(H^{-1}_{\lambda_{r,s}}(t))\Big\} dh(t)$$

$$= -\int_0^1 \Big\{(1-\lambda_{r,s}) \frac{G(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - G(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} [F_{[rN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) -$$

$$- F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))] +$$

$$+ \frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} \big[\lambda_{r,s}\big(F_{[rN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))\big) +$$

$$+ \lambda_{r,s}\big(F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F_{[rN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))\big) +$$

$$+ (1-\lambda_{r,s})\big(G(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - G_{[sN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))\big) + \hat{t} - t\big]\Big\} dh(t).$$

Here we used the fact that $\hat{H}_{[rN],[sN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) = \hat{t}$. So (8.1) follows.

Like the processes $Z(h)$, defined in section 4, can only be defined for special pairs (h, F) of score functions h and continuous distribution functions F, the invariance principle to be proved in this section needs a restriction on h and F which we state now as condition

(A) For every $0 < \lambda < 1$ the derivative $(F \circ H_\lambda^{-1})'$ exists, $d|h|$ a.e. (equivalently $h \in \mathscr{H}$), and for $d|h|$ a.e. $t \in (0, 1)$ $\{(F \circ H_\lambda^{-1})': 0 < \lambda < 1\}$ is uniformly (in λ) continuous at t.

This condition is satisfied if FH_λ^{-1} is differentiable on (0, 1) and has a continuous extension for some $\lambda_0 \in (0, 1)$. Indeed, we have

$$(FH_\lambda^{-1})' = \frac{(1-\lambda_0)(FH_{\lambda_0}^{-1})' \circ H_\lambda \circ H_\lambda^{-1}}{1-\lambda-(\lambda_0-\lambda)(FH_{\lambda_0}^{-1})' \circ H_{\lambda_0} \circ H_\lambda^{-1}}.$$

Thus $(F \circ H_\lambda^{-1})'$ is continuous, has a continuous extension to [0, 1] and for a.e. t it is uniformly continuous. It follows now that the condition for (h, F) is always satisfied if $F = G$ or if the following holds: F and G have continuous densities f and g and there exist intervals (α_i, β_i), $1 \leqslant i \leqslant n$, satisfying:

(i) $\{f > 0\} \cup \{g > 0\} = \bigcup_{i=1}^{n} (\alpha_i, \beta_i)$,

(ii) $\lim_{x \uparrow \beta_i} \frac{g(x)}{f(x)} = \lim_{x \downarrow \alpha_i} \frac{g(x)}{f(x)} \quad (1 \leq i < n-1)$,

(iii) $\lim_{x \uparrow \beta_n} \frac{g(x)}{f(x)}$ and $\lim_{x \downarrow \alpha_n} \frac{g(x)}{f(x)}$ exist.

We leave it as an exercise to show the implication. (These remarks are contained in [18]; cf. also section 9).

LEMMA 8.2. *Let* $0 < \varkappa < 1/2$. *If* h *and* F *satisfy condition* (A), *then*

$$\lim_{N \to \infty} \max_{\substack{(r,s) \in E_\varkappa \\ N(r,s) \geq \sqrt[4]{N}}} \left| \frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} - (F \circ H^{-1}_{\lambda(r,s)})'(t) \right| = 0 \tag{8.4}$$

for $d|h|$ *almost all* t, P *a.e. and*

$$\lim_{N \to \infty} \max_{\substack{(r,s) \in E_\varkappa \\ N(r,s) \geq \sqrt[4]{N}}} \left| \frac{G(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - G(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} - (G \circ H^{-1}_{\lambda(r,s)})'(t) \right| = 0 \tag{8.5}$$

for $d|h|$ *almost all* t, P *a.e.*

Moreover

$$\max_{(r,s) \in E_\varkappa} \left| \frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} - (F \circ H^{-1}_{\lambda(r,s)})'(t) \right| \leq 3\varkappa^{-1} \tag{8.6}$$

and

$$\max_{(r,s) \in E_\varkappa} \left| \frac{G(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - G(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} - (G \circ H^{-1}_{\lambda(r,s)})'(t) \right| \leq 3\varkappa^{-1} \tag{8.7}$$

for $d|h|$ *almost all* t.

Proof. We first show (8.6). The proof of (8.7) follows replacing F by G. By (2.1), $F(x)+G(x) = 2x$. Hence $F'(x) \leq 2$ whenever F is differentiable at x. Since $h \in \mathscr{H}$, $F \circ H^{-1}_{\lambda(r,s)}$ is $d|h|$ a.e. differentiable. Consequently,

$$(F \circ H^{-1}_{\lambda(r,s)})'(t) = \lim_{\varepsilon \downarrow 0} \frac{F(H^{-1}_{\lambda(r,s)}(t+\varepsilon)) - F(H^{-1}_{\lambda(r,s)}(t))}{\varepsilon}$$

$$\leq \lambda(r,s)^{-1} \lim_{\varepsilon \downarrow 0} \frac{\lambda(r,s)[F(H^{-1}_{\lambda(r,s)}(t+\varepsilon)) - F(H^{-1}_{\lambda(r,s)}(t))]}{\varepsilon} +$$

$$+ \frac{(1-\lambda(r,s))[G(H^{-1}_{\lambda(r,s)}(t+\varepsilon)) - G(H^{-1}_{\lambda(r,s)}(t))]}{\varepsilon} \leq \lambda(r,s)^{-1} \leq \varkappa^{-1}.$$

Similarly, if $u \geq v$, then $0 \leq \lambda_{r,s}(F(u)-F(v)) \leq H_{\lambda_{r,s}}(u) - H_{\lambda_{r,s}}(v)$ implies

$$\frac{F(u)-F(v)}{H_{\lambda_{r,s}}(u) - H_{\lambda_{r,s}}(v)} \leq \lambda_{r,s}^{-1} \leq 2\varkappa^{-1}$$

by (6.4). Formula (8.6) follows from both estimates.

We now prove (8.4) ((8.5) is similar). By (6.4) we have $[rN]^{-1} \leqslant 2\varkappa^{-1} N(r,s)^{-1} \leqslant 2\varkappa^{-1} N^{-1/4}$ and $[sN]^{-1} \leqslant 2\varkappa^{-1} N^{-1/4}$, and therefore

$$\left|1-\frac{r}{r+s}\frac{N(r,s)}{[rN]}\right| \leqslant \frac{1}{[rN]} \leqslant 2\varkappa^{-1} N^{-1/4},$$

$$\left|1-\frac{s}{r+s}\frac{N(r,s)}{[sN]}\right| \leqslant \frac{1}{[sN]} \leqslant 2\varkappa^{-1} N^{-1/4}.$$

It follows that

$$\begin{aligned}
&\left|H_{\lambda(r,s)}\big(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})\big)-t-H_{\lambda(r,s)}\big(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})\big)+H_{\lambda(r,s)}\big(H^{-1}_{\lambda_{r,s}}(t)\big)\right| \\
&\quad = \left|\left(1-\frac{r}{r+s}\frac{N(r,s)}{[rN]}\right)\frac{[rN]}{N(r,s)}\big(F\big(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})\big)-F\big(H^{-1}_{\lambda_{r,s}}(t)\big)\big)+\right. \\
&\qquad \left.+\left(1-\frac{s}{r+s}\frac{N(r,s)}{[sN]}\right)\frac{[sN]}{N(r,s)}\big(G\big(H^{-1}_{[rN],[sN]}(\hat{t})\big)-G\big(H^{-1}_{\lambda_{r,s}}(t)\big)\big)\right| \\
&\quad \leqslant 2\varkappa^{-1} N^{-1/4}\left|H_{\lambda_{r,s}}\big(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})\big)-t\right|
\end{aligned}$$

and

$$\left|1-\frac{H_{\lambda(r,s)}\big(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})\big)-H_{\lambda(r,s)}\big(H^{-1}_{\lambda_{r,s}}(t)\big)}{H_{\lambda_{r,s}}\big(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})\big)-t}\right| \leqslant 2\varkappa^{-1} N^{-1/4}.$$

Therefore, it suffices to show that

$$\text{(8.9)} \qquad \left|\frac{F\big(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})\big)-F\big(H^{-1}_{\lambda_{r,s}}(t)\big)}{H_{\lambda(r,s)}\big(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})\big)-H_{\lambda(r,s)}\big(H^{-1}_{\lambda_{r,s}}(t)\big)}-(F\circ H^{-1}_{\lambda(r,s)})'(t)\right| \to 0$$

uniformly in $(r,s)\in E_\varkappa$ such that $N(r,s) \geqslant N^{1/4}$.

Since h and F satisfy condition (A), we have for $d|h|$ a.e. t: Given $\varepsilon>0$ $\exists\eta>0$ such that for all $0<\lambda<1$ and all $|\delta|<\eta$ for which $(F\circ H_\lambda^{-1})'(t+\delta)$ is defined,

$$|(F\circ H_\lambda^{-1})'(t+\delta)-(F\circ H_\lambda^{-1})'(t)| < \varepsilon.$$

Since $F\circ H_\lambda^{-1}$ is absolutely continuous, it follows that

$$\begin{aligned}
&\left|\frac{F\big(H_\lambda^{-1}(t+\delta)\big)-F\big(H_\lambda^{-1}(t)\big)}{\delta}-(F\circ H_\lambda^{-1})'(t)\right| \\
&\qquad = \left|\delta^{-1}\int_t^{t+\delta}\big[(F\circ H_\lambda^{-1})'(u)-(F\circ H^{-1})'(t)\big]\,du\right| < \varepsilon,
\end{aligned}$$

and this holds uniformly in $0<\lambda<1$.

Since

$$\left|\frac{r}{r+s}-\frac{[rN]}{N(r,s)}\right| \leqslant N^{-1/4}$$

uniformly in $(r, s) \in E_\varkappa$ and $N(r, s) \geqslant N^{1/4}$, we have $H_{\lambda(r,s)}(H^{-1}_{\lambda_{r,s}}(t)) \to t$ as $N \to \infty$ (uniformly in $(r, s) \in E_\varkappa$) with $N(r, s) \geqslant N^{1/4}$ and hence, for $d|h|$ a.e. t,

$$\frac{F(H^{-1}_{\lambda_{r,s}}(t)) - F(H^{-1}_{\lambda(r,s)}(t))}{H_{\lambda(r,s)}(H^{-1}_{\lambda_{r,s}}(t)) - t} - (F \circ H^{-1}_{\lambda(r,s)})'(t) \to 0$$

uniformly in (r, s) as $N \to \infty$ $dJh|$ a.e.

Also, since $H_{\lambda(r,s)}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t \to 0$ uniformly in $(r, s) \in E_\varkappa$ with $N(r, s) \geqslant N^{1/4}$,

$$\frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(H^{-1}_{\lambda(r,s)}(t))}{H_{\lambda(r,s)}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} - (F \circ H^{-1}_{\lambda(r,s)})'(t) \to 0$$

uniformly in (r, s) as $N \to \infty$ $d|h|$ a.e.

It follows that (8.9) is

$$\Bigg|\left(\frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(H^{-1}_{\lambda(r,s)}(t))}{H_{\lambda(r,s)}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} - (F \circ H^{-1}_{\lambda(r,s)})'(t)\right) \times$$

$$\times \frac{H_{\lambda(r,s)}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t}{H_{\lambda(r,s)}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - H_{\lambda(r,s)}(H^{-1}_{\lambda_{r,s}}(t))} +$$

$$+ \left(\frac{F(H^{-1}_{\lambda(r,s)}(t)) - F(H^{-1}_{\lambda_{r,s}}(t))}{t - H_{\lambda(r,s)}(H^{-1}_{\lambda_{r,s}}(t))} - (F \circ H^{-1}_{\lambda(r,s)})'(t)\right) \times$$

$$\times \frac{t - H_{\lambda(r,s)}(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda(r,s)}(H^{-1}_{[rN],[sN]}(\hat{t})) - H_{\lambda(r,s)}(H^{-1}_{\lambda_{r,s}}(t))}\Bigg|$$

and tends to zero as $N \to \infty$ uniformly in $(r, s) \in E_\varkappa$ with $N(r, s) \geqslant N^{1/4}$, $d|h|$ a.e. This proves (8.4).

PROPOSITION 8.3. *Let h and F satisfy condition* (A) *and let h have bounded total variation norm on* (0, 1). *Assume that the underlying probability space is rich enough such that there exist two independent Kiefer processes $\mathscr{K}_1$ and $\mathscr{K}_2$ such that uniformly in (r, t), resp. (s, t),*

$$\frac{[rN]}{\sqrt{N}}[F_{[rN]}(t) - F(t)] \to \mathscr{K}_1(r, t) \quad \textit{and} \quad \frac{[sN]}{\sqrt{N}}[G_{[sN]}(t) - G(t)] \to \mathscr{K}_2(s, t)$$

in probability.

Then, for any $0 < \varkappa < 1/2$,

$$\max_{(r,s) \in E_\varkappa} \Bigg| S_N(h, s, r) + \frac{s}{r+s} \int_0^1 (G \circ H^{-1}_{\lambda(r,s)})'(t)\, \mathscr{K}_1(r, H^{-1}_{\lambda(r,s)}(t))\, dh(t) -$$

$$- \frac{r}{r+s} \int_0^1 (F \circ H^{-1}_{\lambda(r,s)})'(t)\, \mathscr{K}_2(s, H^{-1}_{\lambda(r,s)}(t))\, dh(t) \Bigg|$$

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

Proof. We shall use the representation (8.1) of $S_N(h)$ from Lemma 8.1. Observe first that, by (8.2) and the proof of (8.6),

(8.10) $$\left|\frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-F(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-t}(t-\hat{t})\right| \leqslant 2\varkappa^{-1}N(r,s)^{-1}.$$

Next write

(8.11)
$$A_N(r,s,t)=\left|\frac{[rN]}{\sqrt{N}}\left[F_{[r,N]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))\right]-\mathscr{K}_1(r,H^{-1}_{\lambda(r,s)}(t))\right|,$$

(8.12) $B_N(r,s,t)$
$$=\left|\frac{[sN]}{\sqrt{N}}\left[G_{[sN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-G(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))\right]-\mathscr{K}_2(s,H^{-1}_{\lambda(r,s)}(t))\right|.$$

Since all paths of $\mathscr{K}_1$ and $\mathscr{K}_2$ are uniformly continuous, $A_N(r,s,t)$, $B_N(r,s,t)\to 0$ as $N\to\infty$ uniformly in t and $(r,s)\in E_\varkappa$ with $N(r,s)\geqslant N^{1/4}$. In view of (8.6) and (8.7) it follows that

(8.13)
$$\max_{\substack{(r,s)\in E_\varkappa\\ N(r,s)\geqslant N^{1/4}}}\left|\frac{[rN]}{\sqrt{N}}\int_0^1(1-\lambda_{r,s})\left\{\frac{G(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-G(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-t}\times\right.\right.$$
$$\times\left[F_{[rN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))\right]-$$
$$\left.-\frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-F(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-t}\left[G_{[sN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-G(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))\right]\right\}dh(t)-$$
$$-\int_0^1\left\{(1-\lambda(r,s))\frac{G(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-G(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-t}\mathscr{K}_1(r,H^{-1}_{\lambda(r,s)}(t))-\right.$$
$$\left.\left.-\lambda(r,s)\frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-F(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-t}\mathscr{K}_2(s,H^{-1}_{\lambda(r,s)}(t))\right\}dh(t)\right|$$
$$\leqslant C(\varkappa)\left[\max A_N(r,s,t)+\max B_N(r,s,t)+N^{-1/4}\right]\to 0 \quad\text{as } N\to\infty,$$

where $C(\varkappa)$ denotes some constant depending on $\varkappa$ and the total variation norm $\|dh\|$ of dh.

By Lemma 8.2, (8.4), and (8.5), and the dominated convergence theorem,

$$\lim_{N\to\infty}\int_0^1\max_{\substack{(r,s)\in E_\varkappa\\ N(r,s)\geqslant N^{1/4}}}\left|\frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-F(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))-t}-(F\circ H^{-1}_{\lambda(r,s)})'(t)\right|d|h|(t)=0 \text{ a.s.}$$

and similarly for G.

Now, we set

$$R_N = \Big\{ C(\varkappa)\,[\max A_N(r,s,t) + \max B_N(r,s,t) + N^{-1/4}] < \varepsilon/3,$$

$$\max_{(u,v)\in[0,1]^2} |\mathscr{K}_i(u,v)| \leqslant M \quad (i=1,2),$$

$$\int \max_{\substack{N(r,s)\geqslant N^{1/4}\\ (r,s)\in E_\varkappa}} \left| \frac{F(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - F(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} - (F\circ H^{-1}_{\lambda(r,s)})'(t) \right| d|h|(t) < \frac{\varepsilon}{3M}$$

and

$$\int \max_{\substack{N(r,s)\geqslant N^{1/4}\\ (r,s)\in E_\varkappa}} \left| \frac{G(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - G(H^{-1}_{\lambda_{r,s}}(t))}{H_{\lambda_{r,s}}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t})) - t} - (G\circ H^{-1}_{\lambda(r,s)})'(t) \right| d|h|(t) < \frac{\varepsilon}{3M} \Big\}.$$

Here we may choose M, depending on N, in such a way that $M \to \infty$ and $P(R_N^c) \to 0$ as $N \to \infty$.

By (8.10) and (8.13) it follows that on R_N

$$\max_{\substack{(r,s)E_\varkappa\\ N(r,s)>N^{1/4}}} \Big| S_N(h,r,s) + (1-\lambda(r,s)) \int_0^1 (G\circ H^{-1}_{\lambda(r,s)})'(t) \times$$
$$\times \mathscr{K}_1(r, H^{-1}_{\lambda(r,s)}(t))\,dh(t) -$$
$$-\lambda(r,s) \int_0^1 (F\circ H^{-1}_{\lambda(r,s)})'(t)\, \mathscr{K}_2(s, H^{-1}_{\lambda(r,s)}(t))\,dh(t) \Big| \leqslant \varepsilon. \tag{8.14}$$

It is easy to see that we can extend the maximum in (8.14) over all $(r,s)\in E_\varkappa$ and this proves the proposition.

THEOREM 8.1. *Let h and F satisfy condition* (A) *and let h have bounded total variation on* $[0, 1]$. *Then $S_N(h)$ converges weakly in $(D(E), \delta)$ to the process $Z(h)$ defined in* (4.1).

Proof. We may assume that the probability space is rich enough. By Skorohod's theorem and the weak convergence of empirical processes to Kiefer processes, there exist independent Kiefer processes $\mathscr{K}_1$ and $\mathscr{K}_2$ such that the assumption of Proposition 8.3 is satisfied. Denote by $Z(h)$ the process defined in (4.1) with the two specified Kiefer processes. Then by Proposition 8.3 we have, for any $0<\varkappa<1/2$ and any $\varepsilon>0$,

$$\lim_N P\big(\max_{(r,s)\in E_\varkappa} |S_N(h,r,s) - Z(h,r,s)| \geqslant \varepsilon\big) = 0.$$

Let $\varepsilon>0$ and $\eta>0$ be given. Choose k_0 such that $\varepsilon - 2^{k_0} \geqslant \varepsilon/2$ and then N_0 such that, for $N\geqslant N_0$,

$$P\Big(\max_{(r,s)\in E_\varkappa} |S_N(h,r,s) - Z(h,r,s)| \geqslant \frac{\varepsilon}{2k_0}\Big) < \eta/k_0.$$

It follows that

$$P\big(\delta(S_N(h), Z(h)) \geqslant \varepsilon\big) = P\Big(\sum_{k=2}^{\infty} \min[2^{-k}, \delta_{1/k}(S_N(h), Z(h))] \geqslant \varepsilon\Big)$$

$$\leqslant P\Big(\sum_{k=2}^{k_0} \min[2^{-k}, \delta_{1/k}(S_N(h), Z(h))] \geqslant \frac{\varepsilon}{2}\Big) \leqslant \sum_{k=2}^{k_0} P\Big(\delta_{1/k}(S_N(h), Z(h)) \geqslant \frac{\varepsilon}{2k_0}\Big) < \eta.$$

The weak convergence follows now from Lemma 3.2.

THEOREM 8.2. *Let h and F satisfy condition* (A). *Then $S_N(h)$ converges weakly in $(D(E), \delta)$ to $Z(h)$ defined in* (4.1).

Proof. If $h \in \mathscr{H}$, then, for any $\theta > 0$,

$$h_\theta(t) = \begin{cases} h(t) & \text{if } \theta \leqslant t \leqslant 1-\theta, \\ 0 & \text{if } 0 < t < \theta \text{ or } 1-\theta < t < 1 \end{cases}$$

has bounded total variation. Moreover, if (h, F) satisfies condition (A), so does (h_θ, F). We can proceed as in the proof of Theorem 7.1 to prove our claim.

COROLLARY 8.1. *Let h and F satisfy condition* (A) *and let $\bar{S}_N(h)$ be defined as in* (5.21). *Then $\bar{S}_N(h)$ converges weakly in $(D(E), \delta)$ to $Z(h)$ defined in* (4.1).

Proof. This follows from Theorem 8.2 as in Corollary 7.1.

COROLLARY 8.2. *Let the scores $a_N(i, m, n)$ be defined by* (2.3), *where (h, F) is as before. Define the random function $\hat{S}_N(h)$ by* (5.22). *Then $\hat{S}_N(h)$ converges weakly in $(D(E), \delta)$ to $Z(h)$ defined in* (4.1).

Proof. This follows as Corollaries 5.3 or 7.2, using Theorem 8.2 instead of 5.1 or 7.1.

9. THE INVARIANCE PRINCIPLE OF T. SCHULZE-PILLOT

To the best of our knowledge, no invariance principle of the form given in sections 5, 7 and 8 has been proved except some of the results of Schulze–Pillot [18] in his Ph. D. thesis and, maybe, some very special cases. Since the results of Schulze–Pillot have never been published elsewhere, we shortly sketch his results.

Schulze–Pillot uses the approach of Pyke and Shorack [15] to prove limit theorems for the two-sample linear rank statistics. The linear rank statistic can be represented as an integral over the two-sample empirical process. We used this representation in (8.1), however in the present proof only in the case of score functions with bounded total variation on (0, 1). Contrary to this easy case, a general score function needs a more careful handling of the two-sample empirical process, that is, an invariance principle with respect to some dominating function q has to be proven.

This requires to enlarge E to $E\times[0, 1]$, to replace $D(E)$ by $D_q(E\times[0, 1])$ and to define a suitable metric on $D_q(E\times[0, 1])$. We will start with this description. Denote by $\mathscr{A}$ the set of all functions q: $[0, 1]\to \boldsymbol{R}_+$ which are continuous, non-decreasing on the interval $[0, {}^1/_2]$ and non-increasing on the interval $[{}^1/_2, 1]$, and (see [15]) which satisfy

$$\int_0^1 [q(t)]^{-2}\, dt < \infty. \tag{9.1}$$

As in section 3 let us define a metric d_q as

$$d_q(f, g) = \sum_{k=2}^{\infty} \min\left\{2^{-k};\ \sup_{\substack{(r,s)\in E_{1/k}\\ 0\leqslant t\leqslant 1}} \frac{|f(r, s, t)-g(r, s, t)|}{q(t)}\right\} \tag{9.2}$$

for two functions f and g defined on $E\times[0, 1]$. In order to make d_q well defined, use the convention $0/0 = 0$ and $a/0 = \infty$ for $0 < a < \infty$. Clearly, d_q defines the topology of uniform convergence with respect to q on each of the sets $E_\lambda\times[0, 1]$, $0 < \lambda \leqslant {}^1/_2$. This means that $d_q(f_n, f)\to 0$ if and only if, for every $0 < \lambda \leqslant {}^1/_2$, $[q(t)]^{-1}(f_n(r, s, t)-f(r, s, t))$ converges to zero uniformly in $(r, s)\in E_\lambda$ and $0\leqslant t\leqslant 1$.

In analogy to the definition of $D(E)$ denote by $D_q(E\times[0, 1])$, the d_q-closure of all simple functions of the form

$$f = \sum_{k=1}^{n} \alpha_k I_{[E\cap(A_1^k\times A_2^k)]\times A_3^k},$$

where A_i^k for $i = 1, 2$ is again an interval $[a_i^k, b_i^k)$ with $0\leqslant a_i^k, b_i^k\leqslant 1$ or $\{1\}$, and where $A_3^k = [a_3^k, b_3^k]$ with $0 < a_3^k, b_3^k \leqslant 1$. It is not hard to see that functions in $D_q(E\times[0, 1])$ have only discontinuities of the first kind and that $C_q(E\times[0, 1])$, the space of continuous functions on $E\times[0, 1]$, is a closed separable subset.

Let $\mathscr{D}_q$ denote the Borel-σ-field generated by d_q. If D^1 is a separable subspace, then $\mathscr{D}_q\cap D'$ is generated by the projections (cf. [2], § 18)). Thus, the random functions X taking value in $D^1\subset \mathscr{D}_q(E\times[0, 1])$ are measurable if, for every projection π: $D_q(E\times[0, 1])\to \boldsymbol{R}$, $\pi f = f(r, s, t)$, where $(r, s)\in E$ and $0\leqslant t\leqslant 1$, the function $\pi\circ X$ is measurable.

In particular, the two-sample empirical process appearing in (8.1) is measurable.

In the situation described in section 2, Pyke and Shorack [15] defined the two-sample empirical process as

$$N^{1/2}\left[F_n\left(\hat{H}_{n,m}^{-1}(t)\right)-F\left(H_{n/N}^{-1}(t)\right)\right]. \tag{9.3}$$

More generally, Schulze–Pillot defined the two sample sequential empirical process $\Lambda_N(r, s, t)$ by

(9.4) $$\Lambda_N(r, s, t) = \frac{[rN]}{\sqrt{N}}[F_{[rN]}\hat{H}^{-1}_{[rN],[sN]}(\hat{t}) - FH^{-1}_{\lambda_{r,s}}(t)]$$

for $N \in \mathbb{N}$, $(r, s, t) \in E \times [0, 1]$. Here, similarly to (8.2),

$$\hat{t} - \frac{i}{N} \text{ if } t \in \left[\frac{i}{N+1}, \frac{i+1}{N+1}\right), \; 0 \leqslant i < n+m, \quad \text{and} \quad \hat{t} = 1 \text{ if } t = 1.$$

Since $\hat{t}$ can assume only the values $i/(N+1)$, $F_{[rN]}(\hat{H}^{-1}_{[rN],[sN]}(\hat{t}))$ has a separable range and so has $F(H^{-1}_{\lambda_{r,s}}(t))$. It follows that $\Lambda_N(r, s, t)$ *is* $D_q(E \times [0, 1])$-measurable. We may, therefore, speak of weak convergence of Λ_N with respect to d_q.

Let $\mathscr{K}_1$ and $\mathscr{K}_2$ be two independent, standard Kiefer processes on $C[0, 1]^2$, i.e. $\mathrm{E}\mathscr{K}_i(s, t) = 0$ $(0 \leqslant s, t \leqslant 1)$ and $\mathrm{E}\mathscr{K}_i(s, t)\mathscr{K}_i(s', t') = t(1-t') \times \min(s, s')$ $(0 \leqslant t, t' \leqslant 1, 0 \leqslant s, s' \leqslant 1)$. For each $0 < \lambda \leqslant 1$, FH_λ^{-1} and GH_λ^{-1} are monotone and hence $a_\lambda(t) = (FH_\lambda^{-1})'(t)$ and $b_\lambda(t) = (GH_\lambda^{-1})'(t)$ exist a.e. with respect to t and satisfy $a_\lambda \geqslant 0$, $b_\lambda \geqslant 0$ and $\lambda a_\lambda + (1-\lambda) b_\lambda = 1$. We now put

(9.5) $$\Lambda(r, s, t) = (1-\lambda(r, s))\, b_{\lambda(r,s)}(t)\, \mathscr{K}_1(r, FH^{-1}_{\lambda(r,s)}(t)) - \lambda(r, s)\, a_{\lambda(r,s)}\, \mathscr{K}_2(s, GH^{-1}_{\lambda(r,s)}(t))$$

for all $(r, s, t) \in E \times [0, 1]$, where a_λ and b_λ are defined. (Note that integrating over Λ with respect to dh yields a random variable with the same distribution as in (4.2) provided $h \in \hat{\mathscr{H}}$.) For the hypothesis of the next theorem compare section 8.

THEOREM 9.1 [18]. *Suppose that, for every* $\lambda \in [0, 1]$, FH_λ^{-1} *is differentiable in the open interval* $(0, 1)$ *and that, for some* $0 < \lambda_0 < 1$, a_{λ_0} *has a continuous extension on* $[0, 1]$. *Then*:

(a) *for any* $q \in Q$, Λ *is* $D_q(E \times [0, 1])$-*measurable and has continuous paths*;

(b) Λ_N *converges weakly in* $(D_q(E \times [0, 1]), d_q)$ *to* Λ, *provided* $q \in Q$.

It is clear that Λ is a Gaussian process and its covariance structure is easily deduced from its definition.

Let $h \in \mathscr{H}$ and denote by $d|h|$ the total variation measure of h. Let $q \in Q$ such that $\int q\, d|h| < \infty$. Then every $f \in D_q(E \times [0, 1])$ is integrable with respect to $d|h|$ in its last coordinate, i.e., for every $(r, s) \in E$, $\int |f(r, s, t)|\, d|h|(t) < \infty$.

The map I_h: $D_q(E \times [0, 1]) \to D(E)$, given by $I_h(f)(r, s) = \int f(r, s, t)\, dh(t)$, is continuous and, by (4.2) and (9.5),

(9.6) $$I_h(\Lambda) = Z(h)$$

whenever this is well defined (cf. section 4, i.e. $h \in \hat{\mathscr{H}}$ in particular if h is absolutely continuous).

THEOREM 9.2 [18]. *Let $h \in \mathscr{H}$ and assume that there exists a $q \in Q$ with $\int q d|h| < \infty$. Then under the assumptions of Theorem 9.1, $S_N(h)$ converges weakly in $(D(E), \delta)$ to $Z(h)$ given in* (4.1).

If h can be written as a difference $h_1 - h_2$ of two monotone functions satisfying

$$|h_i| \leqslant K\big(t(1-t)\big)^{-1/2+\eta} \tag{9.7}$$

for some $K > 0$ and $\eta > 0$, then there exists a $q \in Q$ with $\int q(t) d|h|(t) < \infty$.

Formula (9.7) is not necessary to guarantee the existence of such a function $q \in Q$ as remarked in [15]. But not all functions $h = h_1 - h_2$ with

$$\int \frac{|h_i(t)|}{\sqrt{t(1-t)}} dt < \infty \qquad (i = 1, 2)$$

satisfy the above condition on q. This is one extension of Theorem 9.2 by Theorem 8.2. On the other hand, since we do not depend on the weak convergence of the two-sample sequential empirical process (9.4), we do not need the stronger assumptions on the differentiability in Theorem 9.1 for proving Theorems 8.1 and 8.2. Finally, we would like to remark that Theorem 5.1 seems to be completely new.

10. APPLICATIONS

Pyke and Shorack [16] have shown that $S_N(h, m_N, n_N)$ converges to a normal $N(0, \sigma^2)$ distribution for some specified σ^2 if m_N and n_N are random variables satisfying certain conditions. We first give a proof their result, using Theorems 8.2 and 7.1. This also extends a result of Schulze–Pillot [18].

THEOREM 10.1. *Let m_N and n_N be integer valued random variables ($N \in \mathbf{N}$) satisfying*

$$0 < m_0 \leqslant m_1 \leqslant \ldots \leqslant m_N \leqslant N, \tag{10.1}$$

$$0 < n_0 \leqslant n_1 \leqslant \ldots \leqslant n_N \leqslant N \qquad (N \geqslant 1)$$

and

$$\lim N^{-1} m_N = \lambda_0 \quad \textit{and} \quad \lim N^{-1} n_N \to 1 - \lambda_0 \tag{10.2}$$

in probability for some $0 < \lambda_0 < 1$.

If (h, F) satisfies the assumption of Theorem 8.2 or of Theorem 7.1, then $S_N(h, m_N/N, n_N/N)$ converges weakly to a normal distribution $N(0, \sigma^2)$, where σ^2 is given by

$$\sigma^2 = \mathrm{E}\,[Z(h, \lambda_0, 1 - \lambda_0)]^2, \tag{10.3}$$

i.e. the limiting distribution is given by $Z(h, \lambda_0, 1-\lambda_0)$ *which is the sum of two independent normally distributed random variables.*

Proof. We have $N^{-1}(m_N, n_N) \in E$ unless $m_N = 0$ or $n_N = 0$, and so the limit distribution for $S_N(h, m_N/N, n_N/N)$ can be obtained from $S_N(h)$ as follows. Let ψ: $D(E) \times E \to \boldsymbol{R}$ denote the map $\psi(f, r, s) = f(r, s)$. Since, by Theorem 8.2, $S_N(h) \to Z(h)$ weakly in $(D(E), \delta)$, we obtain (using Theorem 4.4 of [2]) that

$$\left(S_N(h), \frac{m_N}{N}, \frac{n_N}{N}\right) \to (Z(h), \lambda_0, 1-\lambda_0)$$

in $(D(E) \times E, \delta \times |\ | \times |\ |)$, where $|\ |$ denotes the usual metric in $\boldsymbol{R}$.

We observe that ψ is continuous (with respect to this metric) at all points $f \in D(E)$ which are themselves continuous and all points $(r, s,) \in (0, 1]^2$. Hence

$$S_N\left(h, \frac{m_N}{N}, \frac{n_N}{N}\right) = \psi\left(S_N(h), \frac{m_N}{N}, \frac{n_N}{N}\right) \to \psi(Z(h), \lambda_0, 1-\lambda_0) = Z(h, \lambda_0, 1-\lambda_0)$$

weakly (in $\boldsymbol{R}$) by Theorem 5.1 of [2]. Letting $c = \lambda_0 = r$ in (4.5) we see that $Z(h, \lambda_0, 1-\lambda_0)$ is normally distributed. The variance is explicitly given there.

In order to prove an invariance principle in the situation of the last theorem we need a bit stronger assumptions.

THEOREM 10.2. *Under the assumptions of Theorem* 10.1, *replacing* (10.2) *by*

$$\frac{m_N}{N} \to \lambda_0 \text{ and } \frac{n_N}{N} \to 1-\lambda_0 \text{ a.s.,} \tag{10.4}$$

it follows that $S_N(h, m_{[uN]}/N, n_{[uN]}/N)$ *converges weakly to a Wiener process* W *with variance given by* (10.3). *This convergence is in* $D([0, 1])$ *with respect to the uniform metric.*

Proof. We proceed as in [18]. Define a random function with values in E by

$$Y_N(u) = \left(\frac{m_{[uN]}}{N}, \frac{n_{[uN]}}{N}\right)$$

(i.e. Y_N has values in $D([0, 1], E)$, the space of functions v: $[0, 1] \to E$ having at most discontinuity of the first kind). By (10.4), $Y_N \to \varphi$ a.s. with respect to the sup-metric $\|\ \|_\infty$ in $D([0, 1], E)$, where $\varphi(u) = (u\lambda_0, u(1-\lambda_0))$.

Again by Theorem 4.4 in [2] we have

$$(S_N(h), Y_n) \to (Z(h), \varphi) \tag{10.5}$$

weakly with respect to the metric $\delta \times \|\ \|_\infty$.

Define now ψ: $D(E)\times D([0,1],E)\to D([0,1])$ by $\psi(f,\xi)(u) - f(\xi_1(u),\xi_2(u))$, where $f\in D(E)$, $\xi=(\xi_1,\xi_2)\in D([0,1],E)$, $0\leqslant u\leqslant 1$.

Unfortunately ψ is not continuous at points (f,φ), $f\in C(E)$, with respect to the uniform metric on $D([0,1])$ and the metric $\delta\times\|\ \|_\infty$ on $D(E)\times D([0,1],E)$. However, we can argue as follows. Let $k\in N$ be so large that $\lambda_0\in[k^{-1},1-k^{-1}]$. Define ψ_k: $D(E)\times D([0,1],E)\to D([0,1])$ by

$$\psi_k(f,\xi)=\begin{cases}\psi(f,\xi) & \text{if } \xi([0,1])\subset E_{1/k},\\ \psi(f,\varphi) & \text{otherwise.}\end{cases}$$

We claim that ψ_k is continuous on $C(E)\times\{\varphi\}$. Let $f_0\in C(E)$. Then, given $\varepsilon>0$, there is an $\eta>0$ such that

$$\sup_{\substack{|r-r'|<\eta\\ |s-s'|<\eta\\ (r,s)\in E_{1/2k}\\ (r',s')\in E_{1/2k}}} |f_0(r,s)-f_0(r',s')|<\varepsilon/2 \quad\text{and}\quad 1/k-\eta\geqslant\frac{1}{2k}.$$

Now let $f\in D(E)$ satisfy

$$\sup_{(r,s)\in E_{1/k}}|f(r,s)-f_0(r,s)|<\varepsilon/2$$

and let $\xi\in D([0,1],E)$ satisfy $\|\xi-\varphi\|_\infty<\eta$. Then, if $\xi([0,1])\subset E_{1/k}$, we have

$$\begin{aligned}&\sup_{0\leqslant u\leqslant 1}|\psi_k(f,\xi)(u)-\psi_k(f_0,\varphi)(u)|\\ &\quad\leqslant \sup_{0\leqslant u\leqslant 1}|\psi(f,\xi)(u)-\psi(f_0,\xi)(u)|+\sup_{0\leqslant u\leqslant 1}|\psi(f_0,\xi)(u)-\psi(f_0,\varphi)(u)|\\ &\quad=\sup_{0\leqslant u\leqslant 1}\left|f(\xi(u))-f_0(\xi(u))\right|+\sup_{0\leqslant u\leqslant 1}\left|f_0(\xi(u))-f_0(\varphi(u))\right|\\ &\quad\leqslant \sup_{(r,s)\in E_{1/k}}|f(r,s)-f_0(r,s)|+\sup_{\substack{(r,s),(r',s')\in E_{1/2k}\\ |r-s|,|r'-s'|<\eta}}|f_0(r,s)-f_0(r',s')|<\varepsilon.\end{aligned}$$

On the other hand, if $\xi([0,1])\not\subset E_{1/k}$, we have

$$\begin{aligned}\sup_{0\leqslant u\leqslant 1}|\psi_k(f,\xi)(u)-\psi_k(f_0,\varphi)(u)| &= \sup_{0\leqslant u\leqslant 1}|\psi(f,\varphi)(u)-\psi(f_0,\varphi)(u)|\\ &= \sup_{0\leqslant u\leqslant 1}\left|f(\varphi(u))-f_0(\varphi(u))\right|<\varepsilon/2\end{aligned}$$

since $\varphi(u)=(u\lambda_0,u(1-\lambda_0))\in E_{1/k}$ (because $u\lambda_0/(u\lambda_0+u(1-\lambda_0))=\lambda_0\in[k^{-1},1-k^{-1}]$). Together with (10.5), we conclude from Theorem 5.1 in [2] that $\psi_k(S_N(h),Y_N)\to\psi_k(Z(h),\varphi)=\psi(Z(h),\varphi)$ for every $k\in N$ with $\lambda_0\in[k^{-1},1-k^{-1}]$ weakly in $(D([0,1],\|\ \|_\infty)$, where $\|\ \|_\infty$ also is used for the supnorm on $D([0,1])$. Note that $\psi(Z(h),\varphi)(u)=Z(h,u\lambda_0,u(1-\lambda_0))$ will have the desired properties because of (4.5).

It, therefore, remains to show that for every $\varepsilon > 0$,

$$\lim_{k\to\infty}\overline{\lim_{N\to\infty}}\ \mathrm{P}\big(\|\psi_k(S_N(h),\ Y_N)-\psi(S_N(h),\ Y_N)\|_\infty \geqslant \varepsilon\big) = 0. \tag{10.6}$$

This follows from the estimate

$$\begin{aligned}
&\lim_{k\to\infty}\overline{\lim_{N\to\infty}}\ \mathrm{P}\big(\|\psi_k(S_N(h),\ Y_N)-\psi(S_N(h),\ Y_N)\|_\infty \geqslant \varepsilon\big)\\
&\leqslant \lim_{k\to\infty}\overline{\lim_{N\to\infty}}\ \mathrm{P}\big(Y_N([0,\ 1]) \not\subset E_{1/k}\big)\\
&= \lim_{k\to\infty}\overline{\lim_{N\to\infty}}\ \mathrm{P}\left(\exists u\in[0,\ 1] \text{ such that } \frac{m_{[uN]}}{m_{[uN]}+n_{[uN]}}\notin[k^{-1},\ 1-k^{-1}]\right)\\
&\leqslant \lim_{k\to\infty}\mathrm{P}\left(\exists l\in N \text{ such that } \frac{m_l}{m_l+n_l}\notin[k^{-1},\ 1-k^{-1}]\right)\\
&= \mathrm{P}\left(\bigcap_k\{\exists l\in N \text{ such that } \frac{m_l}{m_l+n_l}\notin[k^{-1},\ 1-k^{-1}]\right) = 0
\end{aligned}$$

since $m_l/l \to \lambda_0$ and $n_l/l \to 1-\lambda_0$ a.s.

THEOREM 10.3. *Let Y_N be random variables with values in* $(0,\ 1)$ *such that $Y_N \to \lambda_0$ a.s. Then $S_N(h,\ Y_N r,\ Y_N s)$ converges weakly in* $(D(E),\ \delta)$ *to the process* $Z(h,\ \lambda_0 r,\ \lambda_0 s)$, $(r,\ s)\in E$, *provided* $(h,\ F)$ *satisfies the assumption of Theorem* 8.2 *or Theorem* 7.1.

Proof. Again by Theorem 4.4 of [2], we get that $(S_N(h),\ Y_N) \to (Z(h),\ \lambda_0)$ weakly with respect to the metric $\delta\times|\ |$.

Define the map ψ: $D(E)\times[0,\ 1]\to D(E)$ by $\psi(f,\ u)(r,\ s) = f(ur,\ us)$. Clearly, ψ is continuous at every point $(f,\ u)\in C(E)\times[0,\ 1]$. Our claim follows again from Theorem 5.1 in [2], since $\psi(S_N(h),\ Y_N)(r,\ s) = S_N(h,\ Y_N r,\ Y_N s)$.

As in [16], tests for symmetry may be considered as a variation of the two-sample problem. Let $\xi_1, \xi_2, \ldots$ be independent, identically distributed random variables with common continuous distribution function π satisfying $0 < \pi(0) < 1$. Define, for $N\in N$,

$$Z_{N,i} = \begin{cases} 1 & \text{if the } i\text{-th smallest of } |\xi_1|, \ldots, |\xi_N| \text{ is from a positive } \xi, \\ 0 & \text{otherwise} \end{cases}$$

With $a_N(i)$, $1\leqslant i\leqslant N$, as scores, the statistic

$$T_N = \frac{1}{N}\sum_{i=1}^{N} a_N(i)\,Z_{Ni} \tag{10.7}$$

serves to test the symmetry of π.

In order to write this statistic as a two-sample linear rank statistic, define $X_k(\omega) = \xi_{i_k(\omega)}(\omega)$, $1 \leqslant k \leqslant m_N(\omega)$, where $i_k(\omega) = \min\{j > i_{k-1}(\omega) \mid \xi_j(\omega) \geqslant 0\}$ and $m_N(\omega) = |\{j | \xi_j(\omega) \geqslant 0\}|$. Similarly, Y_k is defined using the $\xi_j < 0$. Let $n_N(\omega) = N - m_N(\omega)$. The sequences $\{X_k\}_{k\geqslant 1}$, $\{Y_k\}_{k\geqslant 1}$ and $\{P_k\}_{k\geqslant 1}$ are independent, where $P_k = 1_{[0,\infty)} \circ \xi_k$. The common distribution of the X_k is given by

$$F(x) = \frac{\pi(x) - \pi(0)}{1 - \pi(0)}, \quad x \geqslant 0,$$

and that of the Y_k's is given by

$$G(x) = \frac{\pi(0) - \pi(-x)}{\pi(0)}, \quad x > 0,$$

i.e. the distributions are conditional distributions given $\xi_1 \geqslant 0$, resp. $\xi_1 < 0$. If $\pi_0 = \mathrm{P}(\xi_1 \geqslant 0)$, then $P_k (k \geqslant 1)$ are independent and identically distributed Bernoulli random variables with parameter π_0.

Clearly, $Z_{N,i} = 1$ if the i-th smallest among $X_1, \ldots, X_{m_N}, Y_1, \ldots, Y_{n_N}$ is an X-observation and then i is the rank of that observation. Therefore

$$T_N = \frac{m_N}{N} \int_0^1 h\left(\frac{N}{N+1} \hat{H}_{m_N, n_N}(t)\right) dF_{m_N}(t)$$

if the scores are given by (2.2) (cf. 2.4), and hence we may study the statistic

(10.8) $\quad \Sigma_N(u)$

$$= \frac{[uN]}{\sqrt{N}} \left(\frac{m_{[uN]}}{[uN]} \int_0^1 h\left(\frac{[uN]}{[uN]+1} \hat{H}_{m_{[uN]}, n_{[uN]}}(t) \right) dF_{m_{[uN]}}(t) - \pi_0 \int_0^1 h(H_{\pi_0}(t)) dF(t) \right).$$

We first prove the following

LEMMA 10.1. *If (h, F) satisfies the assumptions of section 8, then*

$$\frac{d}{d\lambda}\left[\int_0^1 F(H_\lambda^{-1}(t))\, dh(t)\right](\lambda_0) = \int_0^1 a_{\lambda_0}(t)\big(G(H_{\lambda_0}^{-1}(t)) - F(H_{\lambda_0}^{-1}(t))\big)\, dh(t),$$

where a_λ is defined as in (9.5), i.e. $a_\lambda = (F \circ H_\lambda^{-1})'$ whenever it exists.

Proof. We have

$$\frac{d}{d\lambda}\left[\int_0^1 F \circ H_\lambda^{-1}(t)\, dh(t)\right](\lambda_0) = \lim_{\lambda \to \lambda_0} \int_0^1 \frac{F \circ H_\lambda^{-1}(t) - F \circ H_{\lambda_0}^{-1}(t)}{\lambda - \lambda_0}\, dh(t)$$

as in the proof of Lemma 8.2. Hence it suffices to show that, for $d|h|$ almost all t,

$$\frac{F \circ H_\lambda^{-1}(t) - F \circ H_{\lambda_0}^{-1}(t)}{\lambda - \lambda_0} \to a_{\lambda_0}(t)\big(G \circ H_{\lambda_0}^{-1}(t) - F \circ H_{\lambda_0}^{-1}(t)\big).$$

Write

$$\frac{F\circ H_\lambda^{-1}(t)-F\circ H_{\lambda_0}^{-1}(t)}{\lambda-\lambda_0} = \frac{F\circ H_{\lambda_0}^{-1}\circ H_{\lambda_0}\circ H_\lambda^{-1}(t)-F\circ H_{\lambda_0}^{-1}(t)}{H_{\lambda_0}\circ H_\lambda^{-1}(t)-t}\,\frac{H_{\lambda_0}\circ H_\lambda^{-1}(t)-t}{\lambda-\lambda_0}.$$

The first factor converges $d|h|$ a.s. to $(F\circ H_{\lambda_0}^{-1})'(t) = a_{\lambda_0}(t)$ and the second factor can be written in the form

$$\frac{H_{\lambda_0}\circ H_\lambda^{-1}(t)-t}{\lambda-\lambda_0} = \frac{(\lambda_0-\lambda)F\big(H_\lambda^{-1}(t)\big)+(\lambda-\lambda_0)G\big(H_\lambda^{-1}(t)\big)}{\lambda-\lambda_0} = G\big(H_\lambda^{-1}(t)\big)-F\big(H_\lambda^{-1}(t)\big)$$

and converges to $G\big(H_{\lambda_0}^{-1}(t)\big)-F\big(H_{\lambda_0}^{-1}(t)\big)$ as $\lambda\to\lambda_0$ by continuity.

The following theorem extends a result of Schulze–Pillot [18].

THEOREM 10.4. *In the situation described above, let (h, F) satisfy the assumptions of section* 8. *Then, the statistics $\Sigma_N(u)$, defined by* (10.8), *converges weakly in $\big(D([0, 1]), \|\ \|_\infty\big)$ to the process*

$$\Sigma(u) = Z\big(h, u\pi_0, u(1-\pi_0)\big)-\sqrt{\pi_0(1-\pi_0)}\Big[\int_0^1 F\big(H_{\pi_0}^{-1}(t)\big)\,dh(t)+ +\pi_0\int_0^1 a_{\pi_0}(t)\Big(G\big(H_{\pi_0}^{-1}(t)\big)-F\big(H_{\pi_0}^{-1}(t)\big)\Big)\,dh(t)\Big]\,W(u),$$

where W denotes a standard Brownian motion independent of $Z(h)$, which is defined in (4.1).

Proof. We will proceed as in the proof of Theorem 10.2. Define the random functions $Y_N(u) = (m_{[uN]}/N,\ n_{[uN]}/N)$ as before and write

$$\lambda_N(u) = \frac{m_{[uN]}}{[uN]},\qquad V_N(u) = \frac{[uN]}{\sqrt{N}}\big(\lambda_N(u)-\pi_0\big)\qquad (0\leqslant u\leqslant 1) \tag{10.9}$$

and

$$l_N(u) = -\int F\big(H_{\lambda_N(u)}^{-1}(t)\big)\,dh(t) - -\pi_0\frac{\int_0^1 \Big(F\big(H_{\lambda_N(u)}^{-1}(t)\big)-F\big(H_{\pi_0}^{-1}(t)\big)\Big)\,dh(t)}{\lambda_N(u)-\pi_0}\quad \text{if } \lambda_N(u)-\pi_0\neq 0, \tag{10.10}$$

$$l_N(u) = 0 \text{ if } \lambda_N(u) = \pi_0\qquad (0\leqslant u\leqslant 1).$$

Integration by parts gives

$$\Sigma_N(u) = S_N\Big(h, \frac{m_{[uN]}}{N}, \frac{n_{[uN]}}{N}\Big)+V_N(u)\,l_N(u). \tag{10.11}$$

Let us define the map ψ: $D(E)\times D([0, 1])\times D([0, 1], E)\times D([0, 1]) \to D([0, 1])$ by $\psi(f, g, \varphi, \eta)(u) = f(\varphi_1(u), \varphi_2(u))+g(u)\eta(u)$ for $0\leqslant u\leqslant 1$, $f\in D(E)$, $g, \eta\in D([0, 1])$ and $\varphi = (\varphi_1, \varphi_2)\in D([0, 1], E)$. Here the difficulty of the non-continuity of ψ occurs again, but we can proceed as before.

Let k be so large that $\pi_0\in[k^{-1}, 1-k^{-1}]$ and define

$$\psi_k(f, g, \varphi, \eta) = \begin{cases}\psi(f, g, \varphi, \eta) & \text{if } \varphi([0, 1])\subset E_{1/k},\\ \psi(f, g, \varphi_0, \eta) & \text{otherwise,}\end{cases}$$

where $\varphi_0(u) = (u\pi_0, u(1-\pi_0))$, and

$$v_k(t) = \begin{cases}0, & 0\leqslant t\leqslant 1/2k,\\ 2kt-1, & 1/2k\leqslant t\leqslant 1/k,\\ 1, & 1/k\leqslant t\leqslant 1.\end{cases}$$

By Theorem 8.2 we have $S_N(h)\to Z(h)$ weakly in $(D(E), \delta)$. The strong law of large numbers implies that $\lambda_N(1)\to\pi_0$ a.s., consequently $Y_N\to\varphi_0$ uniformly in u a.s. The application of Donsker's theorem yields $V_N \to\sqrt{\pi_0(1-\pi_0)}\,W$ weakly in $(D([0, 1]), \|\ \|_\infty)$, and since $\{P_k\}_{k\geqslant 1}$ is independent of $(X_m, Y_n\colon n, m\geqslant 1)$, W is a standard Brownian motion, independent of $Z(h)$. Setting

$$\gamma = -\int_0^1 F(H_{\pi_0}^{-1}(t))\,dh(t)-\pi_0\int_0^1 a_{\pi_0}(t)\big(G(H_{\pi_0}^{-1}(t))-F(H_{\pi_0}^{-1}(t))\big)\,dh(t),$$

the strong law of large numbers and Lemma 10.1 imply that $l_N v_k\to\gamma v_k$ a.s. From Theorems 4.4 and 5.2. of [2] it follows that

$$(S_N(h), V_N, Y_N, l_N v_k)\to(Z(h), \sqrt{\pi_0(1-\pi_0)}\,W, \varphi_0, \gamma v_k)$$

weakly with respect to the metrics δ on $D(E)$, $\|\ \|_\infty$ on $D([0, 1])$ and on $D([0, 1], E)$. ψ_k is continuous at all points (f, g, φ_0, η) with $f\in C(E)$ and g, $\eta\in C([0, 1])$ (cf. the proof of Theorem 10.2). Using Theorem 5.1 of [2], we obtain, for $0\leqslant u\leqslant 1$,

$$\big(\psi_k(S_N(h), V_N, Y_N, l_N v_k)(u)\big)\to\big(Z(h, u\pi_0, u(1-\pi_0))+\sqrt{\pi_0(1-\pi_0)}\,\gamma v_k(u)\,W(u)\big)$$

weakly in $(D([0, 1]), \|\ \|_\infty)$.

Since

$$P\big(\|\Sigma-Z(h, \cdot\,\pi_0, \cdot(1-\pi_0))-\sqrt{\pi_0(1-\pi_0)}\,\gamma v_k(\cdot)\,W(\cdot)\|_\infty\geqslant\varepsilon\big)$$
$$= P\big(\sup_{0\leqslant u\leqslant k^{-1}}|\sqrt{\pi_0(1-\pi_0)}\,\gamma(1-v_k(u))\,W(u)|\geqslant\varepsilon\big)\to 0 \quad\text{as } k\to\infty$$

(since $\lim_{u\to 0} W(u) = 0$ a.s.), we obtain

$$\psi_k\big(Z(h, \cdot\,\pi_0, \cdot(1-\pi_0)), \sqrt{\pi_0(1-\pi_0)}\,W, \varphi_0, \gamma v_k\big)\to\Sigma$$

weakly in $(D([0, 1], \|\ \|_\infty)$ and hence it is left to show that $\psi_k(S_N(h), V_N, Y_N, l_N v_k) - \psi(S_N(h), V_N, Y_N, l_N)$ converges to zero in probability with respect to the metric $\|\ \|_\infty$ on $D([0, 1])$ (cf. the proof of Theorem 10.2).

Let $\varepsilon > 0$. Then

$$P\left(\|\psi_k(S_N(h), V_N, Y_N, l_N v_k) - \psi(S_N(h), V_N, Y_N, l_N)\|_\infty \geqslant \varepsilon\right)$$
$$\leqslant P\left(Y_N([0, 1]) \not\subset E_{k-1}\right) + P\left(\sup_{0 \leqslant u \leqslant 1} |V_N(u) l_N(u)(1 - v_k(u))| \geqslant \varepsilon\right).$$

The first term tends to zero using the same arguments as in the proof of Theorem 10.2. Since $1 - v_k(u) \to 0$ as $k \to \infty$ uniformly on compact sets in $(0, 1)$, by Lemma 10.1 and by the weak convergence of V_N to $\sqrt{\pi_0(1-\pi_0)}\, W$, we infer that $\max V_N(u) l_N(u)$ is bounded and

$$\max_{u \leqslant u_0} V_N(u) l_N(u) \to 0 \quad \text{as } u_0 \to 0.$$

Consequently, the second term tends to zero as well. Therefore

$$\lim_{k \to \infty} \overline{\lim_{N \to \infty}}\, P\left(\|\psi_k(S_N(h), V_N, Y_N, l_N v_k) - \psi(S_N(h), V_N, Y_N, l_N)\|_\infty \geqslant \varepsilon\right) = 0.$$

This completes the proof of the theorem.

REFERENCES

[1] P. Bickel and M. F. Wichura, *Convergence criteria for multiparameter stochastic processes and some applications*, Ann. Math. Statist. 42 (1971), p. 1656-1670.

[2] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York 1968.

[3] H. I. Braun, *Weak convergence of sequential linear rank statistics*, Ann. Statist. 4 (1976), p. 554-576.

[4] H. Chernoff and I. R. Savage, *Asymptotic normality and efficiency of certain nonparametric test statistic*, Ann. Math. Statist. 29 (1958), p. 972-994.

[5] M. Denker and U. Rösler, *Some contributions to Chernoff-Savage theorems*, Statistics and Decision 3 (1985), p. 49-75.

[6] — *A note on asymptotic normality for rank statistics*, Proc. Conf. on Limit Theorems in Prob. and Stat., Veszprém 1982, p. 405-425.

[7] M. Denker, M. L. Puri and U. Rösler, *A sharpening of the remainder term in the higher dimensional CLT for multivariate rank statistics*, J. Multiv. Anal. 17 (1985), p. 148–167.

[8] J. L. Doob, *Stochastic processes*, Wiley, New York 1953.

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

[10] A. A. Filippova, *Mises' theorem on the asymptotic behavior of functionals of empirical distribution functions and its statistical applications*, Teoria Veroyat. i ee Prim. 7 (1961), p. 24-57.

[11] J. Hájek, *Asymptotic normality of simple linear rank statistics under alternatives*, Ann. Math. Statist. 39 (1968), p. 325-346.

[12] T. L. Lai, *On Chernoff-Savage statistics and sequential rank tests*, Ann. Statist. 3 (1975), p. 825-845.

[13] F. Moricz, *Moment inequalities for the maximum of partial sums of random fields*, Acta Sci. Math. Hungar. 39 (1977), p. 353-366.

[14] G. Neuhaus, *On weak convergence of stochastic processes with multidimensional time parameter*, Ann. Math. Statist. 42 (1971), p. 1285-1295.

[15] R. Pyke and G. R. Shorack, *Weak convergence of a two sample empirical process and a new approach to Chernoff-Savage theorems*, ibidem 39 (1968), p. 755-771.

[16] – *Weak convergence and a Chernoff-Savage theorem for random sample sizes*, ibidem 39 (1968), p. 1675-1685.

[17] – *A note on Chernoff-Savage theorems*, ibidem 40 (1969), p. 1116-1119.

[18] T. Schulze-Pillot, *Über Invarianzprinzipien für lineare Rang-Statistiken*, Dissertation, Göttingen 1978.

Institut für Mathematische Stochastik
Universität Göttingen
Lotzestr. 13
3400 Göttingen, FRG

Dept. of Mathematics
Indiana University
Bloomington, Indiana 47405
USA

Received on 25. 1. 1987