Statistics & Decisions 2, 199–223 (1984)

SOME PROPERTIES OF WEIGHTED MULTIVARIATE EMPIRICAL PROCESSES

F.H. Ruymgaart and J.A. Wellner

Received: Revised version: June 2, 1983

Abstract. An exponential bound is obtained for the exceedance probability of the supremum of the variations over all rectangles, contained in a fixed rectangle R, of the multivariate empirical process. The probability mass of this rectangle R figures in the expression of the exponential bound. The bound is used in the study of the local behavior of weighted multivariate empirical processes. Applications are given to strong and weak convergence of such processes indexed both by points and by rectangles.

1. INTRODUCTION AND BASIC INEQUALITY

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random vectors in $\mathbb{R}^d$ defined on a probability space (Ω, A, P) with common d.f. F. *In this paper we shall exclusively deal with the case that the* $X_i = (X_{i1}, \ldots, X_{id})$ *take their values in* $I^d = [0,1]^d$ *with probability* 1 *and that the common d.f.* F *is continuous with Uniform* (0,1) *marginals*. For many purposes this is not too serious a restriction since it has been observed in Wichura (1973, p.293) that given a random vector Y in $\mathbb{R}^d$ with *arbitrary* d.f., there exists a random vector X in I^d with continuous d.f. F having

---

AMS 1980 subject classifications: Primary 60F05, 60G17; Secondary 62E20.

Key words and phrases: exponential inequalities, q-metrics, strong and weak convergence.

Uniform (0,1) marginals and a measurable function $\Psi : I^d \to \mathbb{R}^d$, such that $Y =_d \Psi(X)$; see also Philipp & Pinzur (1980, Lemma 1).

Adopting the notation of Orey & Pruitt (1973), we shall write $x = \langle x_1,\ldots,x_d\rangle = \langle x_j\rangle \in \mathbb{R}^d$ if it is desirable to display the coordinates of x. If $x_j = \xi$ for all j we simply write $\langle\xi\rangle$. For $\langle x_j\rangle,\langle y_j\rangle \in \mathbb{R}^d$ we define $\langle x_j\rangle \le \langle y_j\rangle$ to mean that $x_j \le y_j$ for all j; we write $\langle x_j\rangle < \langle y_j\rangle$ if $x_j \le y_j$ for all j and $x_j < y_j$ for at least one j. The half-open rectangles $(x_1,y_1] \times \ldots \times (x_d,y_d]$ will be preferably written as $R(x,y)$. In the present notation we have $(0,1]^d = (\langle 0\rangle,\langle 1\rangle]$. The class

$$(1.1) \qquad \mathcal{R} = \{R(x,y) : R(x,y) \subset I^d\}$$

of all half-open rectangles in I^d will play an important role. We will also write $|t| = t_1 \times \ldots \times t_d$, $|dt|$ for Lebesgue measure on I^d and $|B| = \int_{t\in B} |dt|$ for any $B \in \mathcal{B}^d$. Given any (random) function Λ on $\mathbb{R}^d$ that determines a finite (random, signed) measure on $(\mathbb{R}^d,\mathcal{B}^d)$, we write

$$(1.2) \qquad \Lambda\{B\} = \int_B d\Lambda, \quad B \in \mathcal{B}^d.$$

A *weight function* q is any function that satisfies

$$(1.3) \qquad \begin{cases} q : [0,1] \to [0,\infty) \textit{ and } q > 0 \textit{ on } (0,1]; \\ q \textit{ continuous and non-decreasing on } [0,1]; \\ t^{-\frac{1}{2}}q(t) \textit{ non-increasing for } t \in (0,1]. \end{cases}$$

Note that in particular for any $\delta \in [0,\frac{1}{2}]$ the function

$$(1.4) \qquad q_\delta(t) = t^{\frac{1}{2}-\delta}, \quad t \in [0,1],$$

is a weight function. Provided $q(0) = 0$, for values of t near 0 these functions q coincide with those introduced by Shorack & Wellner (1982). On the other hand we don't exclude here that $q \equiv 1$ on $[0,1]$.

The (reduced multivariate) *empirical process* (indexed by points) is defined by

(1.5) $\quad U_n(t) = n^{\frac{1}{2}}(\hat{F}_n(t)-F(t)), \quad t \in I^d,$

where the empirical d.f. $\hat{F}_n$ based on $X_1,\ldots,X_n$ is as usual defined by $n\hat{F}_n(t) = \#\{1 \le i \le n : X_i \in [<0>,<t_j>]\}$, $t = <t_j> \in I^d$. The process

(1.6) $\quad U_n(t)/q(F(t)), \quad t \in I^d \quad (0/0 = 0),$

is called a *weighted empirical process indexed by points*, and the process

(1.7) $\quad U_n\{R\}/q(F\{R\}), \quad R \in \mathcal{R} \quad (0/0 = 0),$

is called a *weighted empirical process indexed by rectangles.*

The local behavior of (1.6) will be investigated in Section 2. Some applications of these local properties to the speed of the Glivenko-Cantelli convergence, referred to as strong convergence, and to weak convergence will be given in Section 3, where we also consider a functional. Our results are related to those in Orey & Pruitt (1973) for multiparameter Wiener processes; see also Rüschendorf (1976, 1980). For d = 1 we may refer e.g. to the recent papers by O'Reilly (1974) and Stute (1982). In Section 3 we also show how some properties of (1.6) carry almost immediately over to (1.7). For d = 1 see Stute (1982) and Shorack & Wellner (1982). For arbitrary d and $q \equiv 1$ the empirical process indexed by the still more general class of all convex measurable subsets was studied by Stute (1977). For arbitrary d and non-i.i.d. observations we refer to van Zuijlen (1982) and Alexander (1982).

The remainder part of this section is devoted to a basic inequality dealing with an exponential bound for the supremum of the empirical process over rectangles contained in a fixed rectangle; it is related to Orey & Pruitt (1973, Lemma 1.2) and to Wichura (1969, Theorem 1). The proof is based on the representation of the empirical process as a conditioned Poisson process along with symmetrization of the latter process so that a Lévy-type inequality can be applied.

We shall also need an exponential bound for a Poisson random variable in a similar way as exponential bounds for binomial random variables might be used in a direct approach (see Remark 1.1). For any Binomial (n,p) random variable U we have

$$(1.8) \qquad P(n^{-\frac{1}{2}}|U-np| \geq \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{2p(1-p)} \psi\left(\frac{\lambda}{pn^{\frac{1}{2}}}\right)\right),$$

for $\lambda \geq 0$, $p \in (0,1)$, where ψ satisfies

$$(1.9) \qquad \begin{cases} \psi : [0,\infty] \to \mathbb{R} \textit{ decreasing to } 0,\ \psi(0) = 1; \\ \psi(\lambda) = 2\lambda^{-2} \int_0^\lambda \log(1+x)dx, \quad \lambda > 0. \end{cases}$$

The inequality (1.8) is due to Bennett (1962); see Shorack & Wellner (1982) for further details.

Similarly we have for a Poisson (τ) random variable Z that

$$(1.10) \qquad P(|Z-\tau| \geq \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{2\tau} \psi(\frac{\lambda}{\tau})\right),$$

for $\lambda \geq 0$, $\tau > 0$, and ψ as in (1.9). The proof of (1.10) is immediate from (1.8): relabel $\lambda n^{\frac{1}{2}}$ by λ in (1.8) and then use the fact that for $np = \tau$, $U \to_d Z$ as $n \to \infty$.

Let $N_n = \{N_n(t) : t \in [0,\infty)^d\}$ be a Poisson process with

$$(1.11) \qquad EN_n(t) = nF(t), \quad t \in [0,\infty)^d, \quad n \in \mathbb{N},$$

where F is the d.f. of the X_i. We shall use the fact that conditional on $N_n(<1>) = n$, the processes

$$(1.12) \qquad Z_n(t) = n^{-\frac{1}{2}}(N_n(t)-nF(t)), \quad t \in I^d,$$

are equal in law to the $U_n(t)$, $t \in I^d$, in (1.5) for each $n \in \mathbb{N}$.

LEMMA 1.1. *For any* $R \in \mathcal{R}$ *with* $F\{R\} < 1$ *and all* $\lambda > 0$ *we have*

$$(1.13) \qquad P(\sup_{S\subset R}|U_n\{S\}| \geq \lambda) \leq C(F\{R^c\})P(\sup_{S\subset R}|Z_n\{S\}| \geq \lambda),$$

where the suprema are taken over all $S \in \mathcal{R}$ *with* $S \subset R$, $C(\theta) \to 1$ *as* $\theta \to 1$,

$C(\theta) \le 2$ *for* $\theta \ge \frac{1}{2}$, *and* U_n *and* Z_n *are given in* (1.5) *and* (1.12) *respectively.*

PROOF. Let $A = [\sup_{S\subset R}|U_n\{S\}| \ge \lambda]$ and $B = [\sup_{S\subset R}|Z_n\{S\}| \ge \lambda]$. Then, since $U_n =_d (Z_n|N_n(<1>) = n)$ and using the independence of $N_n\{R\}$ and $N_n\{R^c\}$, we see that

$$\begin{aligned}P(A) &= P(B|N_n(<1>) = n) = \\ &= P(B \cap [N_n(<1>) = n])/P(N_n(<1>) = n) = \\ &= \Sigma_{k=0}^{n} P(B \cap [N_n\{R\} = k] \cap [N_n\{R^c\} = n-k])/P(N_n(<1>) = n) = \\ &= \Sigma_{k=0}^{n} (P(N_n\{R^c\} = n-k)/P(N_n(<1>) = n))P(B \cap [N_n\{R\} = k]) \le \\ &\le \frac{P(N_n\{R^c\} = [nF\{R^c\}])}{P(N_n(<1>) = n)} P(B),\end{aligned}$$

because $P(Z = k) \le P(Z = [\tau])$ if Z is a Poisson (τ) random variable ($[\tau]$ = largest integer $\le \tau$). An application of Stirling's formula completes the proof. Q.E.D.

THEOREM 1.1: *basic inequality. Let* $R \in \mathcal{R}$ *with* $F\{R\} < 1$. *Then for every* $\lambda > (8F\{R\})^{\frac{1}{2}}$ *we have*

$$(1.14)\quad P(\sup_{S\subset R}|U_n\{S\}| \ge \lambda) \le C(d)P(|Z_n\{R\}| \ge \tfrac{1}{2}\lambda) \le \le 2C(d)\exp\left(-\frac{\lambda^2}{32F\{R\}}\psi\left(\frac{\lambda}{4F\{R\}n^{\frac{1}{2}}}\right)\right),$$

where the supremum is taken over all $S \in \mathcal{R}$ *with* $S \subset R$, *and* $C(d) = 2^{2d+2}C(F\{R^c\}) \le 2^{2d+3}$ *if* $F\{R\} \le \frac{1}{2}$.

PROOF. Let N_n^* denote an independent Poisson process identically distributed as N_n. We now use the symmetrization inequalities of Loève (1977, p.256-260). For any $S \in \mathcal{R}$, write $\mu N_n\{S\}$ for the median of (the distribution of) $N_n\{S\}$, and set

$$Z_n^{\mu} = n^{-\frac{1}{2}}(N_n - \mu N_n)$$

and

$$Z_n^{s} = n^{-\frac{1}{2}}(N_n - N_n^*) = Z_n^{\mu} - Z_n^{*\mu}.$$

Note that the rv's $Z_n^s\{S\}$ are symmetrically distributed for all rectangles $S \subset I^d$. By Loève (1977, p.256,(a)) it follows that

$$|\mu N_n\{S\}-nF\{S\}| \le (2nF\{S\})^{\frac{1}{2}},$$

and hence, for $S \subset R$,

$$|Z_n\{S\}| \le |Z_n^{\mu}\{S\}|+(2F\{S\})^{\frac{1}{2}} \le$$
$$\le |Z_n^{\mu}\{S\}|+(2F\{R\})^{\frac{1}{2}}.$$

Thus by Lemma 1.1 it follows, with $C = C(F\{R^c\})$ from Lemma 1.1, that

$$P(\sup_{S\subset R}|U_n\{S\}| \ge \lambda) \le C(F\{R^c\})P(\sup_{S\subset R}|Z_n\{S\}| \ge \lambda) \le$$
$$\le CP(\sup_{S\subset R}|Z_n^{\mu}\{S\}| \ge \lambda-(2F\{R\})^{\frac{1}{2}}) \le$$
$$\le CP(\sup_{S\subset R}|Z_n^{\mu}\{S\}| \ge \tfrac{1}{2}\lambda)$$

since $\lambda > (8F\{R\})^{\frac{1}{2}}$ *implies* $-(2F\{R\})^{\frac{1}{2}} > -\frac{1}{2}\lambda$

$$\le 2CP(\sup_{S\subset R}|Z_n^s\{S\}| \ge \tfrac{1}{2}\lambda)$$

by Loève (1977, p.259, B(ii))

$$\le 2^{2d+1}CP(|Z_n^s\{R\}| \ge \tfrac{1}{2}\lambda)$$

by the same argument as given in the proof of Lemma 1.2 of Orey & Pruitt (1973) which uses only the symmetry and independent increment properties of the process

$$\le 2^{2d+2}CP(|Z_n\{R\}| \ge \tfrac{1}{4}\lambda)$$

since $|Z_n^s\{R\}| \le |Z_n\{R\}|+|Z_n^*\{R\}|$ *where* $Z_n^* =_d Z_n$
and this proves the first inequality in (1.14); the second one follows upon applying (1.10). Q.E.D.

<u>REMARK 1.1</u>: *a variation on the basic inequality.* A slightly weaker version of Theorem 1.1 can be proved for arbitrary F by conditioning on $\hat{F}_n\{R\}$ and using the exponential bound in Kiefer (1961, Theorem 1-m) together with (1.8). This proof is patterned on that of van Zwet's

lemma (see Ruymgaart (1974, Lemma 4.4)) and can be found in Ruymgaart & Wellner (1982).

2. LOCAL BEHAVIOR OF WEIGHTED EMPIRICAL PROCESSES

In this section we derive some local properties of weighted empirical processes. In particular we derive exponential bounds for the exceedance probability of the supremum and the oscillation of the weighted empirical process over a subrectangle. All results are immediate corollaries of Theorem 1.1. We shall occasionally use the properties

$$\alpha^2\psi(c\alpha) \geq \beta^2\psi(c\beta) \quad \forall\alpha \geq \beta \geq 0 \quad \forall c \in (0,\infty); \tag{2.1}$$

$$\psi(\alpha) \geq c\psi(c\alpha) \qquad \forall\alpha \geq 0 \quad \forall c \in (0,1]. \tag{2.2}$$

These inequalities are immediate from (1.9).

THEOREM 2.1: *van Zwet's lemma. For all* $R \in \mathcal{R}$ *and* $\lambda > 0$ *we have*

$$P(\sup_{S\subset R}|U_n\{S\}|/F^{\frac{1}{2}}\{R\} \geq \lambda) \leq 2^{2d+10}\lambda^{-2} \qquad (0/0 = 0), \tag{2.3}$$

where the supremum is taken over all $S \in \mathcal{R}$ *with* $S \subset R$.

PROOF. If $F\{R\} \leq \frac{1}{2}$, (2.3) follows directly from (1.14) by replacing λ by $\lambda F\{R\}^{\frac{1}{2}}$, applying Chebychev's inequality and noting that the constant $2^{2d+3}.16 = 2^{2d+7}$ works.

If $F\{R\} > \frac{1}{2}$, partition R into disjoint rectangles R_1 and R_2 so that $R = R_1 \cup R_2$ with $F\{R_i\} \leq \frac{1}{2}$. Then, using

$$\begin{aligned}\sup_{S\subset R}|U_n\{S\}| &\leq \sup_{S_1\subset R_1, S_2\subset R_2}|U_n\{S_1\}+U_n\{S_2\}| \leq \\ &\leq \sup_{S_1\subset R_1}|U_n\{S_1\}|+\sup_{S_2\subset R_2}|U_n\{S_2\}|\end{aligned}$$

and the elementary inequality $F\{R\}^{\frac{1}{2}} = (F\{R_1\}+F\{R_2\})^{\frac{1}{2}} \geq \frac{1}{2}(F\{R_1\}^{\frac{1}{2}}+F\{R_2\}^{\frac{1}{2}})$, we find that

$$\begin{aligned}P(\sup_{S\subset R}|U_n\{S\}| &\geq \lambda F\{R\}^{\frac{1}{2}}) \leq \\ &\leq P(\sup_{S_1\subset R_1}|U_n\{S_1\}| \geq \tfrac{1}{2}\lambda F\{R_1\}^{\frac{1}{2}})\end{aligned}$$

$$(2.15)\quad \frac{q_\delta(F(a))q_\delta(F(b))}{q_\delta(F(b))-q_\delta(F(a))} \geq \frac{F^{\frac12}(a)}{(F(b)-F(a))^\delta} \geq \frac{F^{\frac12}(a)}{\{\Sigma_{j=1}^d(b_j-a_j)\}^\delta},$$

for $a,b \in I^d$ with $a < b$.

In the sequel we shall use the numbers $C_i \in (0,\infty)$ $(i = 1,2,3)$ as generic constants that may only depend on the dimension d. (If γ was not fixed to $\frac12$ in (2.12), the constants C_2 and C_3 would depend on γ as well.)

COROLLARY 2.1. *Let* $a,b \in I^d$ *with* $a < b$ *and* $F(a) > 0$, *and let* $\delta \in [0,\frac12]$. *Suppose that condition* (2.12) *is fulfilled. Provided* $\lambda \geq 4$ *we have*

$$(2.16)\quad P(\sup_{t\in R(a,b)}|U_n(t)|/F^{\frac12-\delta}(t) \geq \lambda) \leq$$

$$\leq C_1 \exp\left(-\frac{C_2\lambda^2}{F^{2\delta}(b)}\,\psi\left(\frac{C_3\lambda}{n^{\frac12}F^{\frac12+\delta}(b)}\right)\right).$$

Provided $\lambda \geq 8d\{\Sigma_{j=1}^d(b_j-a_j)\}^\delta$ ($\lambda \geq 8d$ *suffices*) *we have*

$$(2.17)\quad P(\sup_{s,t\in R(a,b)}|U_n(s)/F^{\frac12-\delta}(s)-U_n(t)/F^{\frac12-\delta}(t)| \geq \lambda) \leq$$

$$\leq C_1 \exp\left(-\frac{C_2\lambda^2}{\{\Sigma_{j=1}^d(b_j-a_j)\}^{2\delta}}\,\psi\left(\frac{C_3\lambda}{n^{\frac12}\{\Sigma_{j=1}^d(b_j-a_j)\}^{\frac12+\delta}}\right)\right).$$

PROOF. Let us first consider (2.16). Because of (2.13) we have

$$\{8F(b)\}^{\frac12}/F^{\frac12-\delta}(a) \leq (8.2)^{\frac12}F^\delta(a),$$

so that the required condition on λ is satisfied according to Theorem 2.2. The bound itself follows easily from

$$\frac{C_2\lambda^2F^{1-2\delta}(a)}{F(b)}\,\psi\left(\frac{C_3\lambda F^{\frac12-\delta}(a)}{n^{\frac12}F(b)}\right) \geq \frac{C_2\lambda^2}{F^{2\delta}(b)}\,\psi\left(\frac{C_3\lambda}{n^{\frac12}F^{\frac12+\delta}(b)}\right);$$

for the inequality we use (2.13) and the monotonicity of ψ.

Let us next turn to (2.17) and write $\Sigma_{j=1}^d(b_j-a_j) = \Sigma$ for brevity. It is immediate from the inequalities

$$\frac{(F(b)-F(a))^{\frac{1}{2}}}{q_\delta(F(a))} \le \frac{\Sigma^{\frac{1}{2}}}{q_\delta(F(a))} \le \Sigma^\delta,$$

$$\frac{F^{\frac{1}{2}}(b)(q_\delta(F(b))-q_\delta(F(a)))}{q_\delta(F(a))q_\delta(F(b))} \le 2^{\frac{1}{2}}\Sigma^\delta,$$

that follow from (2.14) and (2.15) respectively, and from Theorem 2.3 that the required condition on λ is satisfied. As far as the bound itself is concerned we observe that

$$\frac{C_2\lambda^2 q_\delta^2(F(a))}{F(b)-F(a)}\,\psi\Big(\frac{C_3\lambda q_\delta(F(a))}{(F(b)-F(a))n^{\frac{1}{2}}}\Big) \ge$$

$$\ge \frac{C_2\lambda^2}{\Sigma^{2\delta}}\Big(\frac{F(a)}{\Sigma}\Big)^{1-2\delta}\,\psi\Big(\frac{C_3\lambda}{\Sigma^{\frac{1}{2}+\delta}n^{\frac{1}{2}}}\Big(\frac{F(a)}{\Sigma}\Big)^{\frac{1}{2}-\delta}\Big)$$

by (2.2) *with* $c = (F(b)-F(a))/\Sigma$

$$\ge \frac{C_2\lambda^2}{\Sigma^{2\delta}}\,\psi\Big(\frac{C_3\lambda}{n^{\frac{1}{2}}\Sigma^{\frac{1}{2}+\delta}}\Big)$$

by (2.14) *and* (2.1) *with* $\alpha = (F(a)/\Sigma)^{\frac{1}{2}-\delta}$ *and* $\beta = 1$

which yields the required upper bound for the first part of the expression on the right in (2.6), and we note that

$$\frac{C_2\lambda^2 q_\delta^2(F(a))q_\delta^2(F(b))}{F(b)(q_\delta(F(b))-q_\delta(F(a)))^2}\,\psi\Big(\frac{C_3\lambda q_\delta(F(a))q_\delta(F(b))}{F(b)(q_\delta(F(b))-q_\delta(F(a)))n^{\frac{1}{2}}}\Big) \ge$$

$$\ge \frac{C_2\lambda^2}{\Sigma^{2\delta}}\,\psi\Big(\frac{C_3\lambda}{n^{\frac{1}{2}}F^{\frac{1}{2}}(b)\Sigma^\delta}\Big)$$

by (2.15) *and* (2.1) *with* $\alpha = \dfrac{q_\delta(F(a))q_\delta(F(b))}{F^{\frac{1}{2}}(b)(q_\delta(F(b))-q_\delta(F(a)))}$ *and* $\beta = 1/(2^{\frac{1}{2}}\Sigma^\delta)$

$$\ge \frac{C_2\lambda^2}{\Sigma^{2\delta}}\,\psi\Big(\frac{C_3\lambda}{n^{\frac{1}{2}}\Sigma^{\frac{1}{2}+\delta}}\Big)$$

by (2.12) *and the monotonicity of* ψ

which yields the required upper bound for the second part of the expression on the right in (2.6). Since both upper bounds are of the same order, the result of the corollary follows immediately. Q.E.D.

REMARK 2.1: *van Zwet's lemma*. Although the constant on the right side of (2.3) is excessive, the bound is uniform in rectangles $R \in \mathcal{R}$ and continuous d.f's with Uniform (0,1) marginals. Hence for such d.f's Theorem 2.1 implies van Zwet's lemma; see also Remark 1.1.

REMARK 2.2: *the special choice* $q \equiv 1$. If the weight function q is identically equal to 1, the result of Theorem 2.3 simplifies to

$$(2.18)\qquad P(\sup_{s,t \in R(a,b)} |U_n(s)-U_n(t)| \geq \lambda) \leq$$
$$\leq C_1 \exp\left(-\frac{C_2\lambda^2}{F(b)-F(a)}\,\psi\left(\frac{C_3\lambda}{n^{\frac{1}{2}}(F(b)-F(a))}\right)\right).$$

It is easily seen that (2.18) remains true for $F(a) = 0$, provided $F(b)-F(a) < 1$.

REMARK 2.3: *improvement by additional conditions on* F. If we assume that F has a density with respect to Lebesgue measure, bounded away from both 0 and ∞, we may improve on the assertion of Theorem 2.2. See Ruymgaart & Wellner (1982) for details; see also Section 3.A.

3. APPLICATIONS TO STRONG AND WEAK CONVERGENCE

A. *Processes indexed by points*. One has to be careful when the weighted empirical process is studied over sets of time points, including points close to $\{t \in I^d : F(t) = 0\}$. This will be illustrated by the following example, related to what has been observed in Shorack & Wellner (1982, p.640).

Let $d = 2$ and take for F the d.f. that has mass 1 uniformly distributed over the line segment joining (0,1) and (1,0). Hence with probability 1 all the X_i lie on this line. Consider any such $X_i(\omega)$ in the open segment. By letting t converge to $X_i(\omega)$ from the "north-east" to the "south-west" it is clear that

$$(3.1)\qquad |U_{n,\omega}(t)|/q(F(t)) \uparrow \infty \qquad (t \swarrow X_i(\omega)),$$

$$+ P(\sup_{S_2 \subset R_2} |U_n\{S_2\}| \geq \tfrac{1}{2}\lambda F\{R_2\}^{\frac{1}{2}}) \leq$$
$$\leq 2^{2d+9}\lambda^{-2} + 2^{2d+9}\lambda^{-2} \leq 2^{2d+10}\lambda^{-2}$$

by applying (1.14) followed by Chebychev's inequality twice with $F\{R_i\} \leq \frac{1}{2}$ in each case. Q.E.D.

THEOREM 2.2. *Let* $a,b \in I^d$ *with* $a < b$ *and* $F(a) > 0$. *For* $\lambda \geq \{8F(b)\}^{\frac{1}{2}}/q(F(a))$ *we have*

$$(2.4) \qquad P(\sup_{t\in R(a,b)} |U_n(t)|/q(F(t)) \geq \lambda) \leq$$
$$\leq 2C(d) \exp\left(-\frac{\lambda^2 q^2(F(a))}{32F(b)}\, \psi\left(\frac{\lambda q(F(a))}{4F(b)n^{\frac{1}{2}}}\right)\right).$$

PROOF. From the monotonicity of q (see (1.3)) and because $R(a,b) \subset R(0,b)$, where 0 is short for <0>, it follows that

$$(2.5) \qquad P(\sup_{t\in R(a,b)} |U_n(t)|/q(F(t)) \geq \lambda) \leq$$
$$\leq P(\sup_{t\in R(0,b)} |U_n(t)| \geq \lambda q(F(a))).$$

Because $U_n(t) = U_n\{R(0,t)\}$ it is immediate from Theorem 1.1 that this last expression is bounded above by the quantity on the right in (2.4), provided λ satisfies the condition in the theorem. Q.E.D.

THEOREM 2.3. *Let* $a,b \in I^d$ *with* $a < b$ *and* $F(a) > 0$. *For* $\lambda \geq \max\{2^{\frac{1}{2}}4d(F(b)-F(a))^{\frac{1}{2}}/q(F(a)),\ 2^{\frac{1}{2}}4F^{\frac{1}{2}}(b)\ (q(F(b))-q(F(a)))/(q(F(a))q(F(b)))\}$ *we have*

$$(2.6) \qquad P(\sup_{s,t\in R(a,b)} |U_n(s)/q(F(s))-U_n(t)/q(F(t))| \geq \lambda) \leq$$
$$\leq 2C(d)\Big\{d \exp\left(-\frac{\lambda^2 q^2(F(a))}{128d^2(F(b)-F(a))}\, \psi\left(\frac{\lambda q(F(a))}{8d(F(b)-F(a))n^{\frac{1}{2}}}\right)\right)$$
$$+ \exp\left(-\frac{\lambda^2 q^2(F(a))q^2(F(b))}{128F(b)(q(F(b))-q(F(a)))^2}\, \psi\left(\frac{\lambda q(F(a))q(F(b))}{8F(b)(q(F(b))-q(F(a)))n^{\frac{1}{2}}}\right)\right)\Big\}.$$

PROOF. The probability in (2.6) is bounded above by p_1+p_2, where

$$(2.7) \qquad p_1 = P(\sup_{s,t\in R(a,b)} |U_n(s)-U_n(t)| \geq \tfrac{1}{2}\lambda q(F(a))),$$

(2.8) $p_2 = P\left(\sup_{t\in R(a,b)}|U_n(t)| \geq \frac{1}{2}\lambda \frac{q(F(a))q(F(b))}{q(F(b))-q(F(a))}\right).$

As to p_1 let us note that

(2.9) $\sup_{s,t\in R(a,b)}|U_n(s)-U_n(t)| \leq \Sigma_{j=1}^{d} \sup_{S\subset R_j}|U_n\{S\}|,$

where, for j = 1,...,d, the rectangle R_j is given by

(2.10) $R_j = (0,b_1]\times\ldots\times(0,b_{j-1}]\times(a_j,b_j]\times(0,b_{j+1}]\times\ldots\times(0,b_d],$

and where the supremum is taken over all $S \in R$ with $S \subset R_j$. It follows that $F\{R_j\} \leq F(b)-F(a)$. We see from (2.7) and (2.9) that

(2.11) $p_1 \leq \Sigma_{j=1}^{d}P(\sup_{S\subset R_j}|U_n\{S\}| \geq \lambda q(F(a))/(2d)).$

Using Theorem 1.1 along with (2.2) (choose $c = F\{R_j\}/(F(b)-F(a))$) it is clear that an upper bound for p_1 is given by the first part of the expression on the right in (2.6).

By the same argument as the one in (2.5) it follows that p_2 is bounded above by the second part of the expression on the right in (2.6). Q.E.D.

In most applications the rectangles will arise as elements of a partition. In such cases the points a and b will be close together. A reasonable assumption turns out to be that a and b satisfy

(2.12) $\Sigma_{j=1}^{d}(b_j-a_j) \leq \frac{1}{2}F(b).$

(The number $\frac{1}{2}$ could be replaced by any fixed $\gamma \in (0,1)$.) If (2.12) is fulfilled it follows that

(2.13) $F(a)/F(b) \geq (F(b)-\Sigma_{j=1}^{d}(b_j-a_j))/F(b) \geq \frac{1}{2},$

(2.14) $F(a)/\Sigma_{j=1}^{d}(b_j-a_j) \geq \frac{1}{2}F(b)/\Sigma_{j=1}^{d}(b_j-a_j) \geq 1.$

Restriction to the weight functions q_δ in (1.4), $\delta \in [0,\frac{1}{2}]$, leads to further simplifications. For such q_δ we have for any q with $q(0) = 0$. Hence for almost all ω the paths of the weighted empirical process are not in the function space $D(I^2)$ as defined e.g. in Bickel & Wichura (1971) and Neuhaus (1971). It is also clear that the supremum over all $t \in I^d$ of the absolute value of the weighted empirical process equals ∞ with probability 1.

For any $\gamma \in (0,1]$ by $\mathcal{R}(\gamma) \subset \mathcal{R}$ we understand a partition of $(<0>,<1>]$ into rectangles, obtained as a Cartesian product of partitions of the axes, such that

$$(3.2) \qquad \min_{j=1,\ldots,d} b_j - a_j \geq \gamma \qquad \forall R(a,b) \in \mathcal{R}(\gamma).$$

We shall, however, mostly use the partitions

$$(3.3) \qquad \mathcal{S}(\gamma) = \{R(<\frac{k(j)-1}{[1/\gamma]}>,<\frac{k(j)}{[1/\gamma]}>) : <k(j)> \in \mathbb{N}^d\},$$

that consist of squares ($[1/\gamma]$ = largest integer $\leq 1/\gamma$). Of course one has

(3.4) $\mathcal{S}(\gamma)$ *is one of the* $\mathcal{R}(\gamma)$.

Given $0 \leq \alpha \leq \beta \leq 1$ let us also introduce the subclass

$$(3.5) \qquad \mathcal{S}(\gamma;\alpha,\beta) = \{R(a,b) \in \mathcal{S}(\gamma) : F(b) \geq \alpha,\ F(a) \leq \beta\},$$

of all squares having a non-empty intersection with the set $\{t \in I^d : \alpha \leq F(t) \leq \beta\}$.

If we choose

$$(3.6) \qquad \gamma = \gamma(\alpha) = \alpha/(3d),$$

we see that

$$(3.7) \qquad \Sigma_{j=1}^{d}(b_j - a_j) = \frac{d}{[3d/\alpha]} \leq \tfrac{1}{2}\alpha \leq \tfrac{1}{2}F(b) \qquad \forall R(a,b) \in \mathcal{S}(\gamma(\alpha);\alpha,\beta),$$

so that condition (2.12) is fulfilled for all squares in $\mathcal{S}(\gamma(\alpha);\alpha,\beta)$. Note that (3.6) entails in particular that

$$(3.8) \qquad \{\alpha \leq F(t) \leq \beta\} \subset \cup_{R \in \mathcal{S}(\gamma(\alpha);\alpha,\beta)} R \subset \{\tfrac{1}{2}\alpha \leq F(t) \leq 2\beta\}.$$

THEOREM 3.1: *strong convergence. For each* $\varepsilon > 0$ *there exists* $C \in (0,\infty)$:

$$(3.9)\qquad \limsup_{n\to\infty} \sup_{t:F(t)\geq(\varepsilon \log n)/n} \frac{|U_n(t)|}{(CF(t) \log n)^{\frac{1}{2}}} \leq 1, \text{ a.s.}$$

PROOF. Let us choose $\alpha = \alpha_n = (\varepsilon \log n)/n$, $\beta = 1$, $\lambda = \lambda_n = (C \log n)^{\frac{1}{2}}$ for some $C \in (0,\infty)$, and $\gamma_n = \gamma(\alpha_n)$ as in (3.6). Using the corresponding partition (3.5) and its properties it is immediate from (2.16) and the monotonicity of ψ that

$$(3.10)\qquad P(\sup_{t:F(t)\geq\alpha_n} |U_n(t)|/F^{\frac{1}{2}}(t) \geq \lambda_n) \leq$$

$$\leq P(\max_{R\in S(\gamma_n;\alpha_n,1)} \sup_{t\in R}|U_n(t)|/F^{\frac{1}{2}}(t) \geq \lambda_n) \leq$$

$$\leq \Sigma_{R\in S(\gamma_n;\alpha_n,1)} C_1 \exp(-C_2\lambda_n^2\psi(C_3\lambda_n/(n\alpha_n)^{\frac{1}{2}})) \leq$$

$$\leq ([3dn/(\varepsilon \log n)])^d C_1 \exp(-C_2(C \log n)\psi(C_3(C/\varepsilon)^{\frac{1}{2}})) =$$

$$= O(n^{d-CC_2}), \text{ as } n \to \infty.$$

Since the sum over all $n \in \mathbb{N}$ of the numbers on the right in (3.10) is finite provided we choose C such that $CC_2 > d+1$, the conclusion of the theorem follows by the Borel-Cantelli lemma. Q.E.D.

THEOREM 3.2: *weak convergence. Let* $\delta \in (0,\frac{1}{2}]$ *and* $\alpha \in (0,1)$. *Provided the condition*

$$(3.11)\qquad \lim_{\alpha\downarrow 0} \limsup_{n\to\infty} P(\sup_{t:F(t)\leq\alpha} \frac{|U_n(t)|}{F^{\frac{1}{2}-\delta}(t)} \geq \lambda) = 0 \quad \forall\lambda > 0$$

is satisfied, the weighted empirical processes $\{U_n(t)/F^{\frac{1}{2}-\delta}(t),\ t \in I^d\}$ *converge weakly to a Gaussian process in the space* $D(I^d)$.

PROOF. The weak convergence of the finite dimensional distributions being immediate, we restrict ourselves to the behavior of the modulus of continuity. Although moment inequalities suffice, here we will use the exponential bound (2.17) of Corollary 2.1.

According to Bickel & Wichura (1971) we have to verify that

$$(3.12)\qquad \lim_{\alpha\downarrow 0}\ \limsup_{n\to\infty}\ P\left(w'_\alpha\left(\frac{U_n}{F^{\frac{1}{2}-\delta}}\right) \geq \lambda\right) = 0 \qquad \forall\lambda > 0,$$

where, for given α,

$$(3.13)\qquad w'_\alpha(U_n/F^{\frac{1}{2}-\delta}) = \\ = \inf_{R(\alpha)}\ \max_{R\in R(\alpha)}\ \sup_{s,t\in R}\left|\frac{U_n(s)}{F^{\frac{1}{2}-\delta}(s)} - \frac{U_n(t)}{F^{\frac{1}{2}-\delta}(t)}\right|.$$

Because of (3.4) we have

$$(3.14)\qquad w'_\alpha(U_n/F^{\frac{1}{2}-\delta}) \leq \\ \leq \max_{R\in S(\alpha)}\ \sup_{s,t\in R}\left|\frac{U_n(s)}{F^{\frac{1}{2}-\delta}(s)} - \frac{U_n(t)}{F^{\frac{1}{2}-\delta}(t)}\right|.$$

Hence, relabeling α by $\gamma(\alpha) = \alpha/(3d)$, see (3.6), and taking into account (3.11), we may prove as well

$$(3.15)\qquad \lim_{\alpha\downarrow 0}\ \limsup_{n\to\infty}\ P\left(w''_{\gamma(\alpha),\alpha}\left(\frac{U_n}{F^{\frac{1}{2}-\delta}}\right) \geq \lambda\right) = 0 \qquad \forall\lambda > 0,$$

where

$$(3.16)\qquad w''_{\gamma(\alpha),\alpha}\left(\frac{U_n}{F^{\frac{1}{2}-\delta}}\right) = \\ = \max_{R\in S(\gamma(\alpha);\alpha,1)}\ \sup_{s,t\in R}\left|\frac{U_n(s)}{F^{\frac{1}{2}-\delta}(s)} - \frac{U_n(t)}{F^{\frac{1}{2}-\delta}(t)}\right|.$$

For any $\lambda > 0$ we may apply (2.17) for α sufficiently small, since then $\lambda \geq 8d(d\gamma(\alpha))^\delta$, and application yields

$$(3.17)\qquad P\left(w''_{\gamma(\alpha),\alpha}\left(\frac{U_n}{F^{\frac{1}{2}-\delta}}\right) \geq \lambda\right) \leq \\ \leq \Sigma_{R\in S(\gamma(\alpha);\alpha,1)}\ P\left(\sup_{s,t\in R}\left|\frac{U_n(s)}{F^{\frac{1}{2}-\delta}(s)} - \frac{U_n(t)}{F^{\frac{1}{2}-\delta}(t)}\right| \geq \lambda\right) \leq \\ \leq \left([\frac{3d}{\alpha}]\right)^d C_1 \exp\left(-\frac{C_2\lambda^2}{\alpha^{2\delta}}\ \psi\left(\frac{C_3\lambda}{n^{\frac{1}{2}}\alpha^{\frac{1}{2}+\delta}}\right)\right) \to \\ \to C_1([\frac{3d}{\alpha}])^d \exp(-C_2\lambda^2\alpha^{-2\delta}), \quad \textit{as } n\to\infty,$$

since $\psi(0) = 1$. From this (3.15) follows at once. Q.E.D.

In the next theorem it is claimed that (3.11) holds true for the trivial choice $\delta = \frac{1}{2}$ without any further condition on F. Of course this result is not new, but we give a proof which displays once more the usefulness of Theorem 1.1.

THEOREM 3.3: *weak convergence, trivial weight function. If we choose* $\delta = \frac{1}{2}$ *(i.e. the weight function identically equal to* 1), *relation* (3.11) *holds true. Consequently the processes* $\{U_n(t),\ t \in I^d\}$ *converge to a Gaussian process in* $D(I^d)$.

PROOF. Let us first fix $\alpha \in (0,\frac{1}{2})$ and consider the partition $S(\alpha)$. It is clear that

$$P(\sup_{t:F(t)\le\alpha} |U_n(t)| \ge \lambda) \le$$

$$\le \Sigma_{R\in S(\alpha;0,\alpha)} P(\sup_{t\in R}|U_n(t)| \ge \lambda) \le$$

$$\le \left([\frac{1}{\alpha}]\right)^d C_1 \exp\left(-\frac{C_2\lambda^2}{F\{R\}}\, \psi\left(\frac{C_3\lambda}{n^{\frac{1}{2}}F\{R\}}\right)\right)$$

because of Theorem 1.1

$$\le \left([\frac{1}{\alpha}]\right)^d C_1\exp\left(-\frac{C_2\lambda^2}{\alpha}\, \psi\left(\frac{C_3\lambda}{n^{\frac{1}{2}}\alpha}\right)\right)$$

by application of (2.2) *with* $c = F\{R\}/(2d\alpha)$, *where* $c \le 1$ *because* $F\{R(a,b)\} \le F(b)-F(a) \le 2d\alpha$ *if* $R(a,b) \in S(\alpha)$ *and* $\alpha \in (0,\frac{1}{2})$

$$\to C_1([\tfrac{1}{\alpha}])^d \exp(-C_2\lambda^2\alpha^{-1}),\ \textit{as}\ n \to \infty.$$

This last upper bound obviously converges to zero as α decreases to zero. Q.E.D.

It will be clear from the introduction to this section that for non-trivial weight functions (3.11) is not satisfied without any restriction on F. A sufficient condition will turn out to be that F has a *density* f with respect to Lebesgue measure such that

(3.18) $\quad 0 < M_1 \le f(t) \le M_2 < \infty \qquad \forall t \in I^d,$

for some numbers M_1 and M_2. Without loss of generality the condition

THEOREM 3.4: *weak convergence, additional conditions on the d.f. Provided that* F *satisfies* (3.18), *condition* (3.11) *is satisfied for any* $\delta \in (0,\frac{1}{2}]$. *Consequently the weighted empirical processes* $\{U_n(t)/F^{\frac{1}{2}-\delta}(t),\ t \in I^d\}$ *converge weakly to a Gaussian process in the space* $D(I^d)$.

PROOF. Let $\lambda > 0$ be arbitrary but fixed throughout the proof and choose $\alpha \in (0,1)$. Let $\alpha_n = n^{-(1+\delta)}$. It is clear that the probability on the left in (3.11) is bounded above by

$$(3.29)\qquad P\Big(\sup_{t:F(t)\le\alpha_n} \frac{|U_n(t)|}{F^{\frac{1}{2}-\delta}(t)} \ge \lambda\Big) + P\Big(\sup_{t:\alpha_n\le F(t)\le\alpha} \frac{|U_n(t)|}{F^{\frac{1}{2}-\delta}(t)} \ge \lambda\Big).$$

Lemma 3.1 applies to the second term with α replaced by α_n and β by α, provided α is small enough to ensure that $\lambda > (8/\gamma)^{\frac{1}{2}}(\alpha/\gamma)^{\delta}$; this yields

$$(3.30)\qquad C_1 \int_{\{F(t)\le\alpha/\gamma\}} \frac{1}{|t|} \exp\Big(-\frac{C_2\lambda^2}{F^{\delta/2}(t)}\,\psi(C_3\lambda n^{-\delta^2/2})\Big)|dt| \le$$
$$\le C_1 \int_{\{F(t)\le\alpha/\gamma\}} \frac{1}{|t|} \exp\Big(-\frac{C_2\lambda^2}{|t|^{\delta/2}}\Big)|dt| \le$$
$$\le C_1 C(\lambda)\ |\{t \in I^d : |t| \le \alpha/(\gamma M_1)\}|,$$

for some constant $C(\lambda)$ depending on λ and C_2. This upper bound is independent of n and obviously converges to 0 as $\alpha \downarrow 0$.

Let us now turn to the first term in (3.29) and note that, provided $X_i \in \{t : F(t) \ge \alpha_n\}$ for $i = 1,\dots,n$, we have

$$(3.31)\qquad \frac{|U_n(t)|}{F^{\frac{1}{2}-\delta}(t)} = n^{\frac{1}{2}}F^{\frac{1}{2}+\delta}(t) \le n^{\frac{1}{2}}n^{-(1+\delta)(\frac{1}{2}+\delta)} < \lambda,$$

for $n \ge n_0(\lambda) \in \mathbb{N}$. It follows that

$$(3.32)\qquad \limsup_{n\to\infty} P\Big(\sup_{t:F(t)\le\alpha_n} \frac{|U_n(t)|}{F^{\frac{1}{2}-\delta}(t)} \ge \lambda\Big) \le$$
$$\le \limsup_{n\to\infty}\big(1 - P(\cap_{i=1}^{n}\{F(X_i) \ge \alpha_n\})\big) = 0,$$

that F has Uniform (0,1) marginals may be maintained, but this is not essential for this part of the subsection. See also Rüschendorf (1980) and Harel (1980).

The additional condition on F enables us to use a different kind of partitioning that is particularly well suited for the study of F near the "lower" boundary of its support, i.e. the set $\{F(t) = 0\}$ which is equal to the "lower" boundary of the unit cube I^d under this extra condition. For fixed $\theta \in (0,1)$ we first partition $[0,1]$ by

(3.19) $\quad 1 > \theta > \theta^2 > \ldots \downarrow 0.$

This infinite partition is used in Shorack & Wellner (1982). The induced infinite partition of I^d consists of the class of rectangles

(3.20) $\quad \mathcal{Q}(\theta) = \{R(<\theta^{k(j)}>, <\theta^{k(j)-1}>),\ <k(j)> \in \mathbb{N}^d\}.$

Given $0 \le \alpha \le \beta \le 1$ let us also introduce the subclass

(3.21) $\quad \mathcal{Q}(\theta;\alpha,\beta) = \{R(a,b) \in \mathcal{Q}(\theta) : F(b) \ge \alpha,\ F(a) \le \beta\}.$

Condition (3.18) entails that

(3.22) $\quad F(a)/F(b) \ge (M_1/M_2)\theta^d = \gamma \in (0,1) \quad \forall R(a,b) \in \mathcal{Q}(\theta),$

so that the essential part of (2.13) is satisfied with $\frac{1}{2}$ replaced by γ. In particular we have

(3.23) $\quad \{\alpha \le F(t) \le \beta\} \subset \cup_{R \in \mathcal{Q}(\theta;\alpha,\beta)} \subset \{\gamma\alpha \le F(t) \le \beta/\gamma\}.$

In the sequel C_1, C_2 and C_3 will be generic constants depending only on θ in (3.19), γ in (3.22) and the dimension d.

LEMMA 3.1. *Provided* F *satisfies* (3.18), *for each* $n \in \mathbb{N}$, $\delta \in (0,\frac{1}{2}]$, $0 < \alpha \le \beta \le 1$ *and* $\lambda \ge (8/\gamma)^{\frac{1}{2}}(\beta/\gamma)^{\delta}$ *we have*

(3.24) $$P(\sup_{t:\alpha \le F(t) \le \beta} |U_n(t)|/F^{\frac{1}{2}-\delta}(t) \ge \lambda) \le$$
$$\le C_1 \int_{\{\gamma\alpha \le F(t) \le \beta/\gamma\}} \frac{1}{\prod t} \exp\left(-\frac{C_2\lambda^2}{F^{\delta/2}(t)}\, \psi\left(\frac{C_3\lambda}{n^{\frac{1}{2}}\alpha^{\frac{1}{2}-\delta/2}}\right)\right)|dt|.$$

PROOF. We proceed as in the proof of Theorem 3.1 but here we use the partition $\mathcal{Q}(\theta;\alpha,\beta)$. In order to apply Theorem 2.2 to the rectangles in this partition let us observe that

$$\{8F(b)\}^{\frac{1}{2}}/F^{\frac{1}{2}-\delta}(a) = \{8F(b)/F(a)\}^{\frac{1}{2}}F^{\delta}(a) \leq (8/\gamma)^{\frac{1}{2}}(\beta/\gamma)^{\delta},$$

so that the condition on λ is satisfied. Application yields

$$(3.25)\quad P(\sup_{t:\alpha\leq F(t)\leq\beta}|U_n(t)|/F^{\frac{1}{2}-\delta}(t) \geq \lambda) \leq$$
$$\leq \Sigma_{R(a,b)\in\mathcal{Q}(\theta;\alpha,\beta)}\ C_1 \exp\left(-\frac{C_2\lambda^2F^{1-2\delta}(a)}{F(b)}\ \psi\left(\frac{C_3\lambda F^{\frac{1}{2}-\delta}(a)}{n^{\frac{1}{2}}F(b)}\right)\right).$$

Because of (3.22) the first factor in the exponent is bounded below as follows

$$(3.26)\quad \frac{C_2\lambda^2F^{1-2\delta}(a)}{F(b)} \geq \gamma C_2\lambda^2F^{-2\delta}(t) \geq C_2\lambda^2F^{-2\delta}(t) \qquad \forall t \in R(a,b).$$

Using (2.2) with $c = F^{3\delta/2}(t)$, (3.22) and the monotonicity of ψ the second factor may be bounded below by

$$(3.27)\quad \psi\left(\frac{C_3\lambda F^{\frac{1}{2}-\delta}(a)}{n^{\frac{1}{2}}F(b)}\right) \geq F^{3\delta/2}(t)\psi\left(\frac{C_3\lambda F^{3\delta/2}(t)}{n^{\frac{1}{2}}F^{\frac{1}{2}+\delta}(a)}\right) \geq$$
$$\geq F^{3\delta/2}(t)\psi\left(\frac{C_3\lambda F^{3\delta/2}(a)}{n^{\frac{1}{2}}F^{\frac{1}{2}+\delta}(a)\gamma^{3\delta/2}}\right) \geq$$
$$\geq F^{3\delta/2}(t)\psi\left(\frac{C_3\lambda}{n^{\frac{1}{2}}\alpha^{\frac{1}{2}-\delta/2}}\right) \qquad \forall t \in R(a,b).$$

Hence the product is bounded by the exponential on the right in (3.24).

Combination of (3.25)-(3.27) yields (3.24) when, in addition, we use

$$(3.28)\quad 1 = [(1-\theta)^d\, \Pi_{i=1}^d \theta^{k(j)-1}]^{-1} \int_{R(<\theta^{k(j)}>,<\theta^{k(j)-1}>)} |dt| \leq$$
$$\leq \frac{1}{(1-\theta)^d}\int_{R(<\theta^{k(j)}>,<\theta^{k(j)-1}>)} \frac{1}{|t|}|dt|,$$

at the transition from summation to integration. Q.E.D.

because under the present conditions on F it is clear that

$$(3.33)\quad P(\cap_{i=1}^{n}\{F(X_i) \geq \alpha_n\}) \geq \{P(|X_i| \geq \alpha_n/M_1)\}^n =$$
$$= \{1-P(|X_i| \leq \alpha_n/M_1)\}^n \geq$$
$$\geq \{1-M_2\int_{\{|t|\leq\alpha_n/M_1\}}|dt|\}^n.$$

It follows by elementary computations that this last bound in (3.33) converges to 1 as $n \to \infty$. Q.E.D.

REMARK 3.1: *the dependence on* d *only of the* C_i. It is clear from the expression for C(d) in Theorem 1.1 that C_1 varies with F{R(a,b)}. A uniform (for all $R \in \mathcal{R}$) upper bound for C_1 can be obtained, however, by splitting R into rectangles R_1 and R_2 with $F\{R_1\} = F\{R_2\} = \frac{1}{2}F\{R\} \leq \frac{1}{2}$, similar to what happens in the proof of Theorem 2.1. Then it is easily seen that Theorem 1.1 remains true for all $R \in \mathcal{R}$ with C(d) replaced by 2.2^{2d+3} and λ by $\lambda/2$. It follows that C_2 and C_3 are to be modified as well. It should, however, be also noted that in this section we do have $F\{R\} \leq \frac{1}{2}$ with only a few exceptions that could be easily taken care of by a straightforward modification of the argument.

B. Processes indexed by rectangles. Quite a few properties for processes indexed by rectangles may be obtained from related properties of processes indexed by points. For this purpose we associate with the i.i.d. random vectors $X_1,X_2,\ldots$ in I^d the i.i.d. sequence $Y_1,Y_2,\ldots$ of random vectors in I^{2d}, where $Y_i = (1-X_{i1},X_{i1},\ldots,1-X_{id},X_{id})$. The common d.f. of the Y_i will be denoted by G and the empirical d.f. of the first n by $\hat{G}_n$. The d.f. G obviously concentrates mass 1 on the set $\{t \in I^{2d} : t_{2j-1}+t_{2j} = 1;\ j = 1,\ldots,d\}$. Since F has Uniform (0,1) marginals, the same is true for G, so that our only condition on the underlying d.f. is still satisfied.

Let us consider an arbitrary half-open rectangle $R(s,t) \in \mathcal{R}$ and write $\bar{R}(s,t)$ for the corresponding closed rectangle. For this closed rectangle we have

$$(3.34)\quad U_n\{\bar{R}(s,t)\} = V_n(1-s_1,t_1,\ldots,1-s_d,t_d),$$

We can't, however, proceed as in that paper since we haven't derived all the properties on weak convergence of weighted empirical processes in the supremum-norm that are required. But we still may observe that for any *fixed* $\alpha \in (0,1)$,

$$(3.43) \quad S^2_{n,\alpha} \to_d \int_{\{G(u)\geq\alpha\}} \frac{V^2(u)}{G(u)}\, du, \text{ as } n \to \infty,$$

where V is the limiting Gaussian process of the V_n in $D(I^{2d})$. Relation (3.43) is immediate from Theorem 3.3, since 1/G is bounded by $1/\alpha$ on $\{G \geq \alpha\}$.

ACKNOWLEDGMENTS

The authors are grateful for the referee's comments leading to an improvement of the original version of this paper.

REFERENCES

[1] Alexander, K.S.: Some limit theorems and inequalities for weighted and non-identically distributed empirical processes. Ph.D. Thesis, M.I.T. (1982).

[2] Bennet, G.: Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc. 57*, 33-45 (1962).

[3] Bickel, P.J. & Wichura, M.J.: Convergence criteria for multi-parameter stochastic processes and some applications. *Ann. Math. Statist. 42*, 1656-1670 (1971).

[4] Harel, M.: Convergence en loi pour la topologie de Skorokhod du processus empirique multidimensionnel normalisé et semi-corrigé (étude au voisinage de la frontière inférieure de $[0,1]^{1+k}$). In: Statistique non Paramétrique Asymptotique, J.-P. Raoult ed. , *Lecture Notes in Math. 821*, 46-85, Springer, New York (1980).

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

[6] Loève, M.: *Probability Theory I* (4th ed.). Springer,New York (1977).

[7] Neuhaus, G.: On weak convergence of stochastic processes with multi-dimensional time parameter. *Ann. Math. Statist. 42*, 1285-1295 (1971).

[8] O'Reilly, N.E.: On the weak convergence of empirical processes in sup-norm metrics. *Ann. Probability 2*, 642-651 (1974).

[9] Orey, S. & Pruitt, W.E.: Sample functions of the N-parameter Wiener process. *Ann. Probability 1*, 138-163 (1973).

[10] Philipp, W. & Pinzur, L.: Almost sure approximation theorems for the multivariate empirical process. *Z. Wahrscheinlichkeitstheorie verw. Gebiete 54*, 1-13 (1980).

[11] Rüschendorf, L.: Asymptotic distributions of multivariate rank order statistics. *Ann. Statist. 3*, 912-923 (1976).

[12] Rüschendorf, L.: Weak convergence of the weighted multiparameter empirical process In: Statistique non Paramétrique Asymptotique, J.-P. Raoult ed., *Lecture Notes in Math. 821*, 86-94, Springer, New York (1980).

[13] Ruymgaart, F.H.: Asymptotic normality of nonparametric tests for independence. *Ann. Statist. 5,* 892-910 (1974).

[14] Ruymgaart, F.H.: A note on empirical distributions. Report 7721, Mathematisch Instituut, Katholieke Un., Nijmegen (1977).

[15] Ruymgaart, F.H. & Wellner, J.A.: Growth properties of multivariate empirical processes. Report 8202, Math. Inst., Katholieke Un., Nijmegen (1982).

[16] Shorack, G.R. & Wellner, J.A.: Limit theorems and inequalities for the uniform empirical process indexed by intervals. *Ann. Probability 10,* 639-652 (1982).

[17] Stute, W.: Convergence rates for the isotrope discrepancy. *Ann. Probability 5,* 707-723 (1977).

[18] Stute, W.: The oscillation behaviour of empirical processes. *Ann. Probability 10,* 86-107 (1982).

[19] Wichura, M.J.: Some Strassen-type laws of the iterated logarithm for multi parameter stochastic processes with independent increments. *Ann. Probability 1,* 272-296 (1973).

[20] Zuijlen,M.C.A. van: Properties of the empirical distribution function for independent non-identically distributed random vectors. *Ann. Probability 10,* 108-123 (1982).

F.H. Ruymgaart
Mathematisch Instituut
Kath. Un. Nijmegen
Toernooiveld
6525 ED Nijmegen
The Netherlands

J.A. Wellner
Department of Statistics
University of Rochester
Rochester, New York 14627
U.S.A.

where $V_n(u) = n^{\frac{1}{2}}(\hat{G}_n(u)-G(u))$, $u \in I^{2d}$. It is possible, however, to stay with the half-open rectangles, because

$$(3.35) \qquad U_n\{R(s,t)\} =_d U_n\{\bar{R}(s,t)\}.$$

This relation holds true because each hyperplane parallel to one of the sides of I^d has zero F-mass, due to the uniformity of the marginals of F.

THEOREM 3.5: *strong convergence. For any* $\varepsilon > 0$ *there exists* $C \in (0,\infty)$:

$$(3.36) \qquad \limsup_{n\to\infty} \sup_{R:F\{R\}\geq(\varepsilon \log n)/n} \frac{|U_n\{R\}|}{(CF\{R\}\log n)^{\frac{1}{2}}} \leq 1, \text{ a.s.}$$

PROOF. We only have to translate $U_n\{R\}$ in terms of $V_n(u)$, taking into account (3.35), and apply Theorem 3.1. Q.E.D.

REMARK 3.2: *intervals as a special case.* Specializing Theorem 3.5 to $d = 1$ we obtain relation (1.37) in Shorack & Wellner (1982) as a special case. Note that $F\{R\} = |R|$ when $d = 1$.

REMARK 3.3. In Ruymgaart (1977) an attempt has already been made to obtain properties of processes indexed by sets, more general than quadrants, from properties of processes indexed by points (= quadrants). This paper, however, contains an error due to the omission of a condition on the underlying d.f. In Ruymgaart & Wellner (1982) the same idea is used to describe the behavior of processes indexed by ordinate sets. By approximating such a set, one is faced with the necessity of letting the dimension of the vector representing this approximation grow to ∞ with n. Hence the dependence on d of the constants $C_i(d)$ places a serious restriction on the speed of the Glivenko-Cantelli convergence of processes indexed by ordinate sets.

C. On weak convergence of a functional. Motivated by Shorack & Wellner (1982, Section 2), let us consider the statistics

$$(3.37) \qquad S_n^2 = \iint_{\{s_j<t_j;j=1,\ldots,d\}} \frac{|U_n\{R(s,t)\}|^2}{|R(s,t)|} \, ds\, dt,$$

under the null hypothesis that the underlying d.f. F of the X_i is *uniform* on I^d, for arbitrary dimension d.

By letting G and V_n as in Section 3.B it follows that in the present special case

$$(3.38)\qquad G(u) = \begin{cases} \Pi_{j=1}^{d}(u_{2j-1}+u_{2j}-1), & \text{for } u_{2j-1}+u_{2j} > 1 \quad (j = 1,\ldots,d), \\ 0, & \text{otherwise}, \end{cases}$$

so that we may write as well $S_n^2 = S_{n,0,\alpha}^2 + S_{n,\alpha}^2$, where $\alpha \in (0,1)$ and

$$(3.39)\qquad S_{n,0,\alpha}^2 = \int_{\{u\in I^{2d}:0<G(u)\leq\alpha\}} \frac{V_n^2(u)}{G(u)}\,du,$$

$$(3.40)\qquad S_{n,\alpha}^2 = \int_{\{u\in I^{2d}:\alpha\leq G(u)\leq 1\}} \frac{V_n^2(u)}{G(u)}\,du.$$

Let us first consider $S_{n,0,\alpha}^2$. As we have seen in the beginning of Section 3.A the supremum of the weighted empirical processes will cause problems for such degenerate d.f.'s like the one in (3.38). Here we are dealing with an integral of the weighted empirical process, however, and then the problems don't arise. It follows simply from the Markov-inequality that

$$(3.41)\qquad S_{n,0,\alpha}^2 \to_p 0, \text{ as } \alpha \downarrow 0, \text{ uniformly in } n,$$

because we have

$$\sup_{n\in\mathbb{N}} ES_{n,0,\alpha}^2 \leq \int_{\{0<G(u)\leq\alpha\}} \frac{G(u)(1-G(u))}{G(u)}\,du \leq$$
$$\leq |\{0 < G(u) \leq \alpha\}| \downarrow 0, \text{ as } \alpha \downarrow 0.$$

Relation (3.26) entails that

$$(3.42)\qquad \begin{cases} S_n^2 \text{ is asymptotically equivalent with } S_{n,\alpha_n}^2, \text{ for any} \\ \text{sequence } \{\alpha_n\} \text{ with } \alpha_n \downarrow 0, \text{ as } n \to \infty. \end{cases}$$

Taking e.g. $\alpha_n = (\log n)/n$ brings us in a position similar to that of Shorack & Wellner (1982, beginning proof of Theorem 2.1).