The almost sure behaviour of the oscillation modulus of the multivariate empirical process

Statistics & Probability Letters 6 (1987) 87–96
North-Holland

November 1987

THE ALMOST SURE BEHAVIOR OF THE OSCILLATION MODULUS OF THE MULTIVARIATE EMPIRICAL PROCESS

J.H.J. EINMAHL and F.H. RUYMGAART

Katholieke Universiteit Nijmegen, The Netherlands

Received February 1987
Revised May 1987

Abstract: Let ω_n denote the oscillation modulus of the uniform multivariate empirical process, defined as the variation of the process over multi-dimensional intervals with Lebesgue measure not exceeding $a_n \in (0, 1)$. The a.s. limiting behavior of ω_n is established for sequences $\{a_n\}$ of five different orders of magnitude that constitute an essentially complete spectrum of possibilities. Extensions to processes with underlying d.f. more general than the uniform are indicated. The results may have applications in density estimation and in the theory of multivariate spacings. For related results in the univariate case we refer in particular to Mason, Shorack and Wellner (1983) and Mason (1984), and for a general setting to Alexander (1984).

Keywords: multivariate empirical process, oscillation modulus, almost sure limits.

1. Introduction

The oscillation modulus of an empirical process is not only a theoretically interesting concept, but has direct applications in e.g. density estimation and the theory of spacings. In this paper we establish the a.s. behavior of the oscillation modulus for multivariate empirical processes of arbitrary finite dimension d: the complete spectrum of interesting possibilities is considered. Even when specializing to $d = 1$ part (e) of Theorem 3.1 seems to be new, and part (d) of that theorem solves open question 2 on p. 545 in the monograph by Shorack and Wellner (1986). Interesting related results in a general setting may be found in Alexander (1984). In Section 3 we present the main result. Section 4 is devoted to a brief outline of certain applications. In the remainder part of this section some notation and definitions are introduced.

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random vectors on the same probability space taking values in I^d, with $I = [0, 1]$ and $d \in N$. To display the coordinates of $t \in \mathbb{R}^d$ we write $t = \langle t_1, \ldots, t_d \rangle = \langle t_j \rangle$ and, given $s \in \mathbb{R}^d$, we define $s \leqslant t$ to mean that $s_j \leqslant t_j$ for $1 \leqslant j \leqslant d$. For any $s,\ t \in I^d$ with $s \leqslant t$ we briefly write

$$[s_1, t_1] \times \cdots \times [s_d, t_d] = [s, t], \qquad |t - s| = \prod_{i=1}^{d} |t_i - s_i|. \tag{1.1}$$

The empirical d.f. at stage $n \in \mathbb{N}$ is as usual defined by

$$F_n(t) = \frac{1}{n} \cdot \#\{1 \leqslant i \leqslant n\colon X_i \leqslant t\}, \quad t \in I^d, \tag{1.2}$$

and the empirical process with underlying d.f. F by

$$U_{n,F}(t) = n^{1/2}(F_n(t) - F(t)), \quad t \in I^d. \tag{1.3}$$

For the definition of the oscillation modulus it will be convenient to also introduce the notation

$$U_{n,F}([s, t]) = n^{1/2}(F_n([s, t]) - F([s, t])), \tag{1.4}$$

where

$$F([s, t]) = P(X_i \in [s, t]) \text{ and } F_n([s, t]) = n^{-1} \cdot \#\{1 \leqslant i \leqslant n\colon X_i \in [s, t]\}.$$

In this paper we will focus on the oscillation modulus

$$\omega_{n,F}(a) = \sup_{s \leqslant t\colon |t-s| \leqslant a} |U_{n,F}([s, t])|, \quad a \in (0, 1), \tag{1.5}$$

and on $\omega_{n,F}^{+}$ and $\omega_{n,F}^{-}$ which are defined as $\omega_{n,F}$ with absolute value replaced by $+$ and $-$ respectively.

Although the oscillation modulus will be considered for fairly general F, undue technicalities may and will be avoided by occasionally restricting to the special case where F is the uniform distribution, i.e.

$$F(t) = |t|, \quad t \in I^d. \tag{1.6}$$

We will simply write

$$U_{n,F} = U_n \text{ and } \omega_{n,F} = \omega_n, \quad \text{if } F \text{ satisfies (1.6)}. \tag{1.7}$$

Some variations on the definition of oscillation modulus are possible. First one might consider

$$\tilde{\omega}_{n,F}(a) = \sup_{s \leqslant t\colon |t-s| = a} |U_{n,F}([s, t])|, \quad a \in (0, 1). \tag{1.8}$$

The corresponding $\tilde{\omega}_{n,F}^{+}$ and $\tilde{\omega}_{n,F}^{-}$ are considered by Mason (1984) for $d = 1$ who derives interesting results for spacings from the properties of these moduli; the modulus $\tilde{\omega}_{n,F}$ has applications in the theory of density estimation. See also Section 4. A second variation, that boils down to (1.5) for $d = 1$, is given by

$$\omega_{n,F}'(h) = \sup_{s \leqslant t\colon |t_j - s_j| \leqslant h\ \forall j} |U_{n,F}([s, t])|, \quad h \in (0, 1); \tag{1.9}$$

thanks to the condition on the lengths of the sides of the intervals the problem is essentially reduced to a one dimensional problem. Therefore the present modulus (1.5) is mathematically somewhat harder to handle but seems to be more natural. For arbitrary d, Stute (1984) considers the a.s. behavior of $\omega_{n,F}'(h_n)$ for the particular sequence $\{h_n\}$ corresponding to case (b) of Theorem 3.1, and Einmahl and Ruymgaart (1985) consider the a.s. behavior of $\omega_{n,F}'(h_n)$ for F uniform and the particular sequence $\{h_n\}$ corresponding to case (c) of Theorem 3.1. Since properties of $\tilde{\omega}_{n,F}$, $\tilde{\omega}_{n,F}^{+}$, $\tilde{\omega}_{n,F}^{-}$ and $\omega_{n,F}'$ are immediate from those of $\omega_{n,F}$, $\omega_{n,F}^{+}$ and $\omega_{n,F}^{-}$ we will not pay further attention to these moduli here.

We conclude this section with a preparatory remark. It will be convenient to rewrite $U_n([s, t])$ as an ordinary empirical process over the unit square I^{2d} of doubled dimension $2d$. For this purpose we introduce the sequence of $2d$-dimensional random vectors

$$\bar{X}_i = \langle 1 - X_{i1}, \ldots, 1 - X_{id}, X_{i1}, \ldots, X_{id} \rangle, \quad i \in \mathbb{N}. \tag{1.10}$$

The common d.f. of the $\bar{X}_i$ will be denoted by $\bar{F}$, the empirical d.f. of the first n by $\bar{F}_n$ and the corresponding empirical process by

$$U_{n,\bar{F}}(\langle u, t \rangle) = n^{1/2}\left(\bar{F}_n(\langle u, t \rangle) - \bar{F}(\langle u, t \rangle)\right), \quad \langle u, t \rangle \in I^d \times I^d. \tag{1.11}$$

Note that if F is the uniform d.f. in (1.6) the d.f. $\bar{F}$ reduces to

$$\bar{F}(\langle u, t \rangle) = \begin{cases} |t - s|, & \text{for } \langle u, t \rangle \in I^d \times I^d \text{ with } \langle 1 - u_j \rangle = \langle s_j \rangle \leqslant \langle t_j \rangle, \\ 0, & \text{otherwise.} \end{cases} \tag{1.12}$$

We will simply write

$$U_{n,\bar{F}} = \bar{U}_n, \quad \text{if } \bar{F} \text{ satisfies (1.12)}. \tag{1.13}$$

Let us now observe that

$$\overline{U}_n(\langle u, t\rangle) = U_n([s, t]), \quad \text{with } \langle s_j\rangle = \langle 1 - u_j\rangle, \tag{1.14}$$

so that the process U_n indexed by rectangles as needed in definition (1.5) behaves as the ordinary empirical process $\overline{U}_n$ indexed by points.

Some fundamental tools that might be of independent interest are collected in Section 2. Since we derived these tools for arbitrary F and d they apply in particular to the empirical process in (1.13).

2. Fundamental inequalities

Throughout this section the underlying d.f. F is completely arbitrary. Fix $v \in \mathbb{N}^d$ and, for $t \in \mathbb{N}^d$ with $t \leqslant v$, let ξ_t be a real valued r.v. with mean 0 and variance $\sigma_t^2 < \infty$. We assume these r.v.'s to be independent. Let us next define the partial sums

$$S_u = \sum_{t \leqslant u} \xi_t, \quad \text{for } u \in \mathbb{N}^d \text{ with } u \leqslant v, \tag{2.1}$$

and let us write

$$\sigma^2 = \operatorname{Var}(S_v) = \sum_{t \leqslant v} \sigma_t^2. \tag{2.2}$$

The following inequality may be found in Klesov (1983) and is essentially contained in Orey and Pruitt (1973, proof of Lemma 1.2). For a simple proof we refer to Einmahl (1987).

Inequality 2.1. *For all $\lambda \in \mathbb{R}$ we have*

$$P\Big(\max_{u \leqslant v} S_u \geqslant \lambda\Big) \leqslant 2^d P\Big(S_v \geqslant \lambda - d\sqrt{2\sigma^2}\Big). \tag{2.3}$$

This inequality is particularly useful when path properties are studied for processes with independent increments. The inequality applies to Poisson processes as these processes do have independent increments. Let $N_{n,F}$ be a Poisson process on I^d with $E(N_{n,F}(t)) = nF(t)$ for all $t \in I^d$, and consider the standardized process

$$Z_{n,F}(t) = n^{-1/2}(N_{n,F}(t) - nF(t)), \quad t \in I^d. \tag{2.4}$$

The following is immediate from Inequality 2.1.

Inequality 2.2. *For any $t \in I^d$, $\lambda \in \mathbb{R}$ and either choice of sign we have*

$$P\Big(\sup_{s \leqslant t} \pm Z_{n,F}(s) \geqslant \lambda\Big) \leqslant 2^d P\Big(\pm Z_{n,F}(t) \geqslant \lambda - d\sqrt{2F(t)}\Big). \tag{2.5}$$

Conditioned on $Z_{n,F}(\langle 1, \dots, 1\rangle) = 0$ the process $Z_{n,F}$ is equal in law to $U_{n,F}$ in (1.3). By a standard argument (see e.g. Pyke and Shorack (1968) or Ruymgaart and Wellner (1984)) we find the following result. For any $t \in I^d$ with $F(t) \leqslant \frac{1}{2}$, all $\lambda \in \mathbb{R}$ and either choice of sign we have

$$P\Big(\sup_{s \leqslant t} \pm U_{n,F}(s) \geqslant \lambda\Big) \leqslant 2P\Big(\sup_{s \leqslant t} \pm Z_{n,F}(s) \geqslant \lambda\Big). \tag{2.6}$$

It remains to find an upper bound for the tail probabilities of the standardized Poisson r.v. $Z_{n,F}(t)$ for given $t \in I^d$. For this purpose let us introduce the function

$$\begin{cases} \psi(\lambda) = 2\lambda^{-2}\{(1+\lambda)\log(1+\lambda) - \lambda\}, & \lambda \in (-1, 0) \cup (0, \infty), \\ \psi(0) = 1, \quad \psi(-1) = 2. \end{cases} \tag{2.7}$$

For $t \in I^d$ with $F(t) > 0$ the usual moment generating function technique yields

$$P(\pm Z_{n,F}(t) \geqslant \lambda) \leqslant \exp\left(\frac{-\lambda^2}{2F(t)}\psi\left(\frac{\pm\lambda}{n^{1/2}F(t)}\right)\right), \quad \begin{cases} \lambda \geqslant 0 & (+) \\ 0 \leqslant \lambda \leqslant n^{1/2}F(t) & (-) \end{cases}. \tag{2.8}$$

Combining (2.5) – (2.8) we eventually obtain the inequalities that are of basic importance in the study of empirical processes.

Inequality 2.3. *Let $t \in I^d$ with $0 < F(t) \leqslant \frac{1}{2}$. For any $\epsilon \in (0, 1)$ we have*

$$P\Big(\sup_{s \leqslant t} U_{n,F}(s) \geqslant \lambda\Big)0 \leqslant C^+ \exp\left(\frac{-(1-\epsilon)\lambda^2}{2F(t)}\psi\left(\frac{\lambda}{n^{1/2}F(t)}\right)\right), \quad \lambda \geqslant 0, \tag{2.9}$$

$$P\Big(\sup_{s \leqslant t} - U_{n,F}(s) \geqslant \lambda\Big) \leqslant C^- \exp\left(\frac{-(1-\epsilon)\lambda^2}{2F(t)}\psi\left(\frac{-(1-\epsilon)\lambda}{n^{1/2}F(t)}\right)\right), \quad 0 \leqslant \lambda \leqslant n^{1/2}F(t), \tag{2.10}$$

$$P\Big(\sup_{s \leqslant t} |U_{n,F}(s)| \geqslant \lambda\Big) \leqslant C \exp\left(\frac{-(1-\epsilon)\lambda^2}{2F(t)}\psi\left(\frac{\lambda}{n^{1/2}F(t)}\right)\right), \quad \lambda \geqslant 0, \tag{2.11}$$

where C^+, C^-, $C \in (0, \infty)$ depend only on d and ϵ.

3. Application to the oscillation modulus

We start this section with two inequalities in the case of the uniform distribution. First we will give a global version of (2.11) and next a convenient maximal inequality. The following representation of the oscillation modulus,

$$\omega_n(a) = \sup_{\bar{F}(\langle u, t\rangle) \leqslant a} |\bar{U}_n(\langle u, t\rangle)|, \tag{3.1}$$

which is immediate from (1.7), (1, 12) and (1.14), will be used.

Inequality 3.1. *Let $\epsilon \in (0, \frac{1}{2}]$ and $a \in (0, \frac{1}{4})$. Then*

$$P\big(\omega_n(a) \geqslant \lambda a^{1/2}\big) \leqslant C\frac{1}{a}\left(\log\frac{1}{a}\right)^{d-1} \exp\left(\frac{-(1-\epsilon)\lambda^2}{2}\psi\left(\frac{\lambda}{(na)^{1/2}}\right)\right), \quad \lambda \geqslant 0, \tag{3.2}$$

where $C = C(d, \epsilon) \in (0, \infty)$.

Proof. Let us define θ by $\theta^{2d+1} = 1 - \epsilon$ and $l \in \mathbb{N}$ by $\theta^{l+1} < a \leqslant \theta^l$ and observe that $l \leqslant \log(1/a)/\log(1/\theta)$. For $k \in \mathbb{N}_0$ let us introduce

$$A(\theta, k) = \left\{v \in I^d\colon \langle v_j\rangle = \langle\theta^{k(j)}\rangle \text{ with } k(j) \in \mathbb{N}_0 \text{ and } \sum_{j=1}^d k(j) = k\right\}, \tag{3.3}$$

and for $\beta \in (1, 2)$ and $x \in (0, 1)$ define

$$B(\beta, x) = \begin{cases} \left\{ \langle 1 - k(\beta-1)x,\ \beta x + k(\beta - 1)x \rangle \in I^2 \colon k \in \mathbb{N}_0,\ k \leqslant \left[\dfrac{1 - \beta x}{(\beta - 1)x} \right] \right\} \\ \quad \cup \{ \langle \beta x, 1 \rangle \} \quad \text{if } \beta x < 1, \\ \{ \langle 1, 1 \rangle \} \qquad\quad \text{if } \beta x \geqslant 1, \end{cases} \tag{3.4}$$

where $[z]$ denotes the greatest integer $\leqslant z$. It is easily seen that

$$\# A(\theta, k) = \binom{k + d - 1}{d - 1} \quad \text{and} \quad \# B(\beta, x) \leqslant 1/((\beta - 1)x).$$

According to (2.11) and the fact that the function $\lambda \mapsto \lambda\psi(\lambda)$ is increasing for $\lambda \geqslant -1$, we have for $\langle \tilde{u}, \tilde{t} \rangle \in I^d \times I^d$ with $\bar{F}(\langle \tilde{u}, \tilde{t} \rangle) \leqslant a/(1 - \epsilon)$

$$P\left(\sup_{\langle u, t \rangle \leqslant \langle \tilde{u}, \tilde{t} \rangle} | \bar{U}_n(\langle u, t \rangle) | \geqslant \lambda a^{1/2} \right) \leqslant C \exp\left(\frac{-(1-\epsilon)^2 \lambda^2}{2} \psi\left(\frac{\lambda}{(na)^{1/2}} \right) \right). \tag{3.5}$$

Using (3.1) and (3.3)–(3.4) we find

$$\begin{aligned} P\left(\omega_n(a) \geqslant \lambda a^{1/2} \right) &\leqslant P\left(\sup_{\bar{F}(\langle u, t \rangle) \leqslant \theta^l} | \bar{U}_n(\langle u, t \rangle) | \geqslant \lambda a^{1/2} \right) \\ &\leqslant P\left(\max_{v \in A(\theta, l-d)} \sup_{\langle t_j + u_j - 1 \rangle \leqslant v} | \bar{U}_n(\langle u, t \rangle) | \geqslant \lambda a^{1/2} \right) \\ &\leqslant \sum_{v \in A(\theta, l-d)} P\left(\sup_{\langle t_j + u_j - 1 \rangle \leqslant v} | \bar{U}_n(\langle u, t \rangle) | \geqslant \lambda a^{1/2} \right) \\ &\leqslant \sum_{v \in A(\theta, l-d)} P\left(\max_{\langle \tilde{u}, \tilde{t} \rangle \in \prod_{j=1}^{d} B(1/\theta,\ v_j)} \sup_{\langle u, t \rangle \leqslant \langle \tilde{u}, \tilde{t} \rangle} | \bar{U}_n(\langle u, t \rangle) | \geqslant \lambda a^{1/2} \right), \end{aligned} \tag{3.6}$$

where $\prod_{j=1}^{d} B(1/\theta, v_j)$ is the Cartesian product of the $B(1/\theta, v_j)$, noting that $1/\theta \in (1, 2)$. For $\langle \tilde{u}, \tilde{t} \rangle \in I^d \times I^d$ we take the j-th coordinate $\tilde{u}_j$ of $\tilde{u}$ and the j-th coordinate $\tilde{t}_j$ of $\tilde{t}$ pairwise together to form I^2, for each $1 \leqslant j \leqslant d$. Also notice that when $\langle \tilde{u}, \tilde{t} \rangle \in \prod_{j=1}^{d} B(1/\theta, v_j)$ we have

$$\bar{F}(\langle \tilde{u}, \tilde{t} \rangle) = \prod_{j=1}^{d} \left((v_j/\theta) \wedge 1 \right) \leqslant \theta^{l - 2d} \leqslant a/(1 - \epsilon).$$

Hence by (3.5) the last expression in (3.6) is in turn bounded by

$$\begin{aligned} &\binom{l-1}{d-1} \prod_{j=1}^{d} \left(\frac{\theta}{(1-\theta) v_j} \right) C \exp\left(\frac{(1-\epsilon)^2 \lambda^2}{2} \psi\left(\frac{\lambda}{(na)^{1/2}} \right) \right) \\ &\quad \leqslant C \left(\frac{1}{\log(1/\theta)} \right)^{d-1} \left(\frac{\theta}{1-\theta} \right)^{d} \frac{1}{a} \left(\log \frac{1}{a} \right)^{d-1} \exp\left(\frac{-(1-\epsilon)^2 \lambda^2}{2} \psi\left(\frac{\lambda}{(na)^{1/2}} \right) \right). \end{aligned}$$

Relabeling $(1 - \epsilon)^2$ by $(1 - \epsilon)$, which entails $\theta^{4d+2} = 1 - \epsilon$, and $C(\log(1/\theta))^{1-d} \cdot (\theta/(1 - \theta))^d$ by C we obtain the right hand side of (3.2). $\square$

In order to obtain precise a.s. limits we use a standard blocking argument. The success of this method hinges on a maximal inequality in the following form. For $d=1$ we refer to James (1975, Lemma 2.3) and Shorack (1980, Lemma 3).

Inequality 3.2. *Let $\epsilon \in (0, 1)$, $a \in (0, 1]$ and write $n_k = [(1+\frac{1}{2}\epsilon)^k]$, $k \in \mathbb{N}$. For all $k \in \mathbb{N}$ and $\lambda > 2(a/\epsilon)^{1/2}$ we have*

$$P\Big(\max_{n_k < n \leqslant n_{k+1}} \omega_n(a) \geqslant \lambda\Big) \leqslant 2P\big(\omega_{n_{k+1}}(a) \geqslant (1-\epsilon)\lambda\big). \tag{3.7}$$

Proof. Note that it suffices to prove (3.7) with $\geqslant \lambda$ replaced by $> \lambda$. Let $\{r_i\colon i \in \mathbb{N}\}$ be an enumeration of the set

$$\{\langle u, t\rangle \in I^d \times I^d\colon \bar{F}(\langle u, t\rangle) \leqslant a\} \cap \mathbb{Q}^{2d}$$

and let us write

$$S_n(\langle u, t\rangle) = n^{1/2}\bar{U}_n(\langle u, t\rangle), \quad \langle u, t\rangle \in I^d \times I^d. \tag{3.8}$$

According to the representation (3.1) the probability on the left in (3.7) (with $\geqslant \lambda$ replaced by $> \lambda$) is less than or equal to

$$P\Big(\sup_{n_k < n \leqslant n_{k+1};\, i \in \mathbb{N}} |S_n(r_i)| > \lambda(n_k+1)^{1/2}\Big). \tag{3.9}$$

For $n_k < n \leqslant n_{k+1}$ and $i \in \mathbb{N}$ let us introduce the sets

$$A_{ni} = \big\{ |S_n(r_i)| > \lambda(n_k+1)^{1/2}\big\}, \tag{3.10}$$

$$B_{ni} = \big\{ |S_{n_{k+1}}(r_i) - S_n(r_i)| \leqslant \lambda(n_k+1)^{1/2} \cdot \tfrac{1}{2}\epsilon\big\}. \tag{3.11}$$

An application of the "lemma for events", see e.g. Loève (1977, Section 18.1, p. 258) generalized to the countable well-ordered index set $\{\langle n, i\rangle\colon n_k < n \leqslant n_{k+1},\ i \in \mathbb{N}\}$ with the lexicographical ordering, cf. Einmahl (1987), yields

$$\{\text{the probability in (3.9)}\} \cdot \inf_{n_k < n \leqslant n_{k+1};\, i \in \mathbb{N}} P(B_{ni})$$
$$\leqslant P\Big(\sup_{i \in \mathbb{N}} |S_{n_{k+1}}(r_i)| \geqslant (1-\tfrac{1}{2}\epsilon)\lambda(n_k+1)^{1/2}\Big). \tag{3.12}$$

It is easily seen that

$$P\Big(\sup_{i \in \mathbb{N}} |S_{n_{k+1}}(r_i)| \geqslant (1-\tfrac{1}{2}\epsilon)\lambda(n_k+1)^{1/2}\Big) \leqslant P\Big(\sup_{\bar{F}(\langle u, t\rangle) \leqslant a} |\bar{U}_{n_{k+1}}(\langle u, t\rangle)| \geqslant (1-\epsilon)\lambda\Big), \tag{3.13}$$

so that it remains to prove that $P(B_{ni}) \geqslant \frac{1}{2}$ for all $n_k < n \leqslant n_{k+1}$ and all $i \in \mathbb{N}$. Application of Chebyshev's inequality yields

$$P(B_{ni}^c) \leqslant \frac{4\bar{F}(r_i)(1-\bar{F}(r_i))(n_{k+1}-(n_k+1))}{\lambda^2(n_k+1)\epsilon^2} \leqslant \frac{2\epsilon a}{\lambda^2\epsilon^2} \leqslant \tfrac{1}{2}, \tag{3.14}$$

for $\lambda > 2(a/\epsilon)^{1/2}$ as required. □

Theorem 3.1. *Suppose that F has a continuous density f on I^d with maximum $M < \infty$. In each of the following five cases $\{a_n\}$ ($a_n \in (0, 1)$) is a sequence such that $a_n \downarrow 0$ as $n \to \infty$.*

(a) *If* $\log(1/a_n)/\log\log n \to c \in [0, \infty)$ *and if* $na_n \uparrow$ *as* $n \to \infty$, *we have*

$$\limsup_{n\to\infty} \frac{\omega_{n,F}(a_n)}{(a_n \log\log n)^{1/2}} = (2(1+c)M)^{1/2}, \quad a.s. \tag{3.15}$$

(b) *If* $\log(1/a_n)/\log\log n \to \infty$, $na_n/\log n \to \infty$ *and* $na_n \uparrow$ *as* $n \to \infty$, *we have*

$$\lim_{n\to\infty} \frac{\omega_{n,F}(a_n)}{(a_n \log(1/a_n))^{1/2}} = (2M)^{1/2}, \quad a.s. \tag{3.16}$$

(c) *If* $na_n/\log n \to c \in (0, \infty)$ *as* $n \to \infty$ *we have*

$$\lim_{n\to\infty} \frac{n^{1/2}\omega_{n,F}(a_n)}{\log n} = cM(\beta_{cM}^+ - 1), \quad a.s. \tag{3.17}$$

where β_{cM}^+ *is defined by the conditions* $\beta_{cM}^+(\log \beta_{cM}^+ - 1) + 1 = 1/cM$ *and* $\beta_{cM}^+ > 1$.

(d) *If* $na_n/\log n \to 0$, $\log(1/a_n)/\log n \to 1$ *and* $n^2a_n \uparrow$ *as* $n \to \infty$ *we have*

$$\lim_{n\to\infty} \frac{n^{1/2}\omega_{n,F}(a_n)\log((\log n)/(na_n))}{\log n} = 1, \quad a.s. \tag{3.18}$$

(e) *If* $\log(1/a_n)/\log n = c \in (1, \infty)$ (*i.e.* $a_n = n^{-c}$) *we have*

$$\limsup_{n\to\infty} n^{1/2}\omega_{n,F}(a_n) = [c/(c-1)], \quad a.s. \tag{3.19}$$

If in addition $c/(c-1) \notin \mathbb{N}$, *then c may be replaced by* $c_n \to c$ *and* limsup *by* lim.

Proof. Without loss of generality we restrict ourselves in the proof to the case where F is uniform. Let us first prove that the numbers on the right are upper bounds for the respective expressions on the left almost surely. These results follow easily by an application of Inequalities 3.1 and 3.2. Because of the similarity of the proofs for all the cases, we will restrict ourselves to a proof of case (a).

By the Borel–Cantelli lemma and Inequality 3.2 it suffices to prove that for every small $\epsilon > 0$ we have $\sum_{k=1}^{\infty} P(A_k) < \infty$, where

$$A_k = \left\{ \omega_{n_{k+1}}(a_{n_k}) \geqslant \left(2(1+2\epsilon)(1-\epsilon)^{-1}(1+c)\, a_{n_{k+1}} \log\log n_k\right)^{1/2} \right\}, \tag{3.21}$$

and $n_k = [(1+\frac{1}{2}\epsilon)^k]$. Application of Inequality 3.1 yields

$$P(A_k) \leqslant C\frac{1}{a_{n_k}}\left(\log\frac{1}{a_{n_k}}\right)^{d-1} \exp\left(-(1+2\epsilon)(1+c)\frac{a_{n_{k+1}}}{a_{n_k}}(\log\log n_k) \cdot \psi\left(\frac{\left(2(1+2\epsilon)(1-\epsilon)^{-1}(1+c)\, a_{n_{k+1}} a_{n_k}^{-1} \log\log n_k\right)^{1/2}}{(n_{k+1}a_{n_k})^{1/2}} \right)\right). \tag{3.22}$$

The conditions in (a) entail that the argument of the function ψ tends to 0 as $k \to \infty$ so that, for large k, we have

$$P(A_k) \leqslant C\frac{1}{a_{n_k}}\left(\log\frac{1}{a_{n_k}}\right)^{d-1} \exp(-(1+\epsilon)(1+c)\log\log n_k)$$
$$= C\frac{1}{a_{n_k}}\left(\log\frac{1}{a_{n_k}}\right)^{d-1}\left(\frac{1}{\log n_k}\right)^{(1+\epsilon)(1+c)} \tag{3.23}$$

Using once more the conditions under (a) it follows from (3.23) that the $P(A_k)$ are summable, which completes this part of the proof.

Let us next turn to the proof that these same numbers are lower bounds for the expressions on the left almost surely. Due to the fact that all the results in the theorem are independent of the dimension, we can use the one-dimensional lower bounds as far as available. These lower bounds can indeed be found in the literature for the cases (a), (b), (c) and (d), namely in Mason, Shorack and Wellner (1983, Theorem 3), Stute (1982, Lemma 2.9), Komlós, Major and Tusnády (1975, (3.11)) and Alexander (1984, Theorem 3.1(C) and its proof) respectively. Hence it remains to consider the lower bounds in case (e).

Let us first consider the case $c/(c-1) \in \mathbb{N}$. For n, $m \in \mathbb{N}$ with $m \geqslant n$ let us write $(m-n)F_{n,m} = mF_m - nF_n$. Since $na_n \to 0$ it suffices to prove

$$\limsup_{k\to\infty} \sup_{|t-s| \leqslant a_{n_{k+1}}} (n_{k+1}-n_k)F_{n_k,n_{k+1}}([s, t]) \geqslant c/(c-1), \quad \text{a.s.,} \tag{3.24}$$

with $n_k = 2^k$. For $a \in (0, 1)$ let $\mathscr{P}_a$ be a family of $[1/a]$ rectangles $[s, t] \subset I^d$ with disjoint interiors and with $|t-s| = a$. By straightforward computation it may be seen that for a Binomial (n, p) r.v. Y with $np \leqslant \frac{1}{2}$ and $n \geqslant 2m$ for some $m \in \mathbb{N}_0$, there exists a number $C(m) \in (0, \infty)$ depending only on m such that

$$P(Y \geqslant m) \geqslant C(m)(np)^m. \tag{3.25}$$

Hence, for large k, we have

$$P\Big(\sup_{|t-s| \leqslant a_{n_{k+1}}} (n_{k+1}-n_k)F_{n_k,n_{k+1}}([s, t]) \geqslant c/(c-1)\Big)$$
$$\geqslant P\Big(\max_{[s,t]\in\mathscr{P}_{a_{n_{k+1}}}} (n_{k+1}-n_k)F_{n_k,n_{k+1}}([s, t]) \geqslant c/(c-1)\Big)$$
$$\geqslant 1 - \prod_{[s,t]\in\mathscr{P}_{a_{n_{k+1}}}} \Big(1 - P\big((n_{k+1}-n_k)F_{n_k,n_{k+1}}([s, t]) \geqslant c/(c-1)\big)\Big)$$
$$\geqslant 1 - \left(1 - C\Big(\frac{c}{c-1}\Big)2^{\{k-(k+1)c\}\{c/(c-1)\}}\right)^{[2^{(k+1)c}]}$$
$$\to 1 - \exp\Big(-C\Big(\frac{c}{c-1}\Big)2^{-c/(c-1)}\Big), \quad \text{as } k \to \infty. \tag{3.26}$$

In the second inequality we use a result in Mallows (1968) concerning the negative dependence of the components in a multinomium. Now (3.24) follows by an application of the Borel–Cantelli lemma.

Finally we consider the case $c/(c-1) \notin \mathbb{N}$. Then we have $\log(1/a_n)/\log n \to c$, which is equivalent with $a_n = n^{-c+o(1)}$. Since again $na_n \to 0$ it suffices to prove

$$\liminf_{n\to\infty} \sup_{|t-s| \leqslant a_n} nF_n([s, t]) \geqslant [c/(c-1)], \quad \text{a.s.} \tag{3.27}$$

This follows if we can prove that

$$\sum_{n=1}^{\infty} P\Big(\sup_{|t-s| \leqslant a_n} nF_n([s,\, t]) \leqslant [c/(c-1)] - 1 \Big) < \infty. \tag{3.28}$$

Defining $\mu \in (0, 1)$ by $\mu = c/(c-1) - [c/(c-1)]$, and using again $\mathscr{P}_a$, Mallows (1968) and (3.25), we have for large n that

$$\begin{aligned} &P\Big(\sup_{|t-s| \leqslant a_n} nF_n([s,\, t]) \leqslant [c/(c-1)] - 1 \Big) \\ &\quad \leqslant P\Big(\max_{[s,t] \in \mathscr{P}_{a_n}} nF_n([s,\, t]) \leqslant [c/(c-1)] - 1 \Big) \\ &\quad \leqslant \prod_{[s,t] \in \mathscr{P}_{a_n}} \big(1 - P\big(nF_n([s,\, t]) \geqslant [c/(c-1)]\big)\big) \leqslant \Big\{ 1 - C\Big(\Big[\frac{c}{c-1}\Big]\Big) n^{-(c-1)[c/(c-1)] + o(1)} \Big\}^{n^{c+o(1)}} \\ &\quad \leqslant \exp\Big(-C\Big(\Big[\frac{c}{c-1}\Big]\Big) n^{1/2(c-1)\mu} \Big). \end{aligned} \tag{3.29}$$

These last numbers in (3.29) are summable since $\frac{1}{2}(c-1)\mu > 0$. This proves (3.27). □

Theorem 3.2. *Under the conditions of the previous theorem $\omega_{n,F}^{+}$ has the same a.s. limits as $\omega_{n,F}$ in each of the five cases. In cases* (a) *and* (b) *also $\omega_{n,F}^{-}$ has the same a.s. limits as $\omega_{n,F}$, but in case* (c) *we have to replace the limit by $cM(1 - \beta_{cM}^{-})$ (β_{cM}^{-} is defined by the conditions $\beta_{cM}^{-}(\log \beta_{cM}^{-} - 1) + 1 = 1/cM$ and $\beta_{cM}^{-} < 1$ for $cM > 1$, and $\beta_{cM}^{-} = 0$ for $cM \leqslant 1$). In cases* (d) *and* (e) *the modulus $\omega_{n,F}^{-}$ is degenerate.*

Proof. Since $\omega_{n,F}^{+} \leqslant \omega_{n,F}$ and because the lower bounds for $\omega_{n,F}$ are based on $\omega_{n,F}^{+}$ anyway, the assertions about $\omega_{n,F}^{+}$ are immediate.

Applying the same method of proof along with inequality (2.10) rather than (2.11) proves the statements concerning $\omega_{n,F}^{-}$. □

4. Brief outline of applications

4.1. Multivariate spacings

For $d = 1$ Mason (1984) exploited a certain duality between the oscillation moduli $\tilde{\omega}_{n,F}^{\pm}$ (see (1.8)) and spacings to derive strong convergence properties for the latter from those obtained for the moduli. When we restrict to spacings of the multidimensional interval type a similar approach is possible for arbitrary $d \in \mathbb{N}$ using the results of Section 3. In Deheuvels et al (1987), however, results for spacings of more general shape can be obtained as well, using a different method.

4.2. Multivariate density estimation

In proving uniform a.s. convergence of density estimators – in particular naive density estimators – to the true density, inequality (3.2) is a key result. We will not dwell on this application, since it is well-known: see e.g. Stute (1984).

4.3. Relation with emp. difference process

For $d = 1$, a highly technical proof regarding the strong limiting behavior of the difference of the empirical and the quantile process in Kiefer (1970) can be reduced to a few lines when results for the oscillation modulus are used (Stute (1982), Shorack and Wellner (1986, pp. 589, 590)). A result by Watts (1980) on the tails of this so called empirical difference process will be further extended and generalized along these lines in a forthcoming paper by Einmahl and Mason (1987).

References

Alexander, K.S. (1984), Rates of growth and sample moduli for weighted empirical processes indexed by sets, Tech. rept., Dept. of Mathematics, University of Washington, Seattle.

Deheuvels, P., J.H.J. Einmahl, D.M. Mason and F.H. Ruymgaart (1987), The almost sure behaviour of maximal and minimal multivariate k_n-spacings, to appear in *J. Mult. Analysis*.

Einmahl, J.H.J. (1987), *Multivariate Empirical Processes*, CWI Tract 32, Amsterdam.

Einmahl, J.H.J. and D.M. Mason (1987), Strong limit theorems for weighted quantile processes, Tech. report, Un. of Delaware, Dept. of Math. Sc. No. 87-2.

Einmahl, J.H.J. and F.H. Ruymgaart (1985), A strong law for the oscillation modulus of the multivariate empirical process, *Statistics & Decisions* **3**, 357–362.

James, B.R. (1975), A functional law of the iterated logarithm for weighted empirical distributions, *Ann. Probability* **3**, 762–772.

Kiefer, J. (1970), Deviations between the sample quantile process and the sample DF, in: M.L. Puri ed., *Nonparametric Techniques in Statistical Inference* (Cambridge), 299–319.

Klesov, O.I. (1983), The law of the iterated logarithm for multiple sums, *Th. Probability Math. Stat.* **27**, 65–72.

Komlós, J., P. Major and G. Tusnáday (1975), Weak convergence and embedding, in: *Colloquia Mathematica Societatis János Bolyai, Limit Theorems of Probability Theory (North-Holland, Amsterdam), 149–165.*

Loève, M. (1977), *Probability Theory, Part I* (Springer, New York, 4th ed.).

Mallows, C.L. (1968), An inequality involving multinomial probabilities, *Biometrika* **55**, 422–424.

Mason, D.M. (1984), A strong limit theorem for the oscillation modulus of the uniform empirical quantile process, *Stoch. Processes Appl.* **17**, 127–136.

Mason, D.M., G.R. Shorack and J.A. Wellner (1983), Strong limit theorems for oscillation moduli of the uniform empirical process, *Probability Th. Rel. Fields* **65**, 83–97.

Orey, S. and W.E. Pruitt (1973), Sample functions of the N-parameter Wiener process, *Ann. Probability* **1**, 138–163.

Pyke, R. and G.R. Shorack (1968), Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorems, *Ann. Math. Stat.* **39**, 755–771.

Ruymgaart, F.H. and J.A. Wellner (1984), Some properties of weighted multivariate empirical processes, *Statistics & Decisions* **2**, 199–223.

Shorack, G.R. (1980), Some law of the iterated logarithm type results for the empirical process, *Austral. J. Stat.* **22**, 50–59.

Shorack, G.R. and J.A. Wellner (1986), *Empirical Processes with Applications to Statistics* (Wiley, New York).

Stute, W. (1982), The oscillation behaviour of empirical processes, *Ann. Probability* **10**, 86–107.

Stute, W. (1984), The oscillation behaviour of empirical processes: the multivariate case, *Ann. Probability* **12**, 361–379.

Watts, V. (1980), The almost sure representation of intermediate order statistics, *Probability Th. Rel. Fields* **54**, 281–285.