Sankhyā : The Indian Journal of Statistics
1986, Volume **48**, Series A, Pt. 3, pp. 393-403

SOME PROPERTIES OF WEIGHTED COMPOUND MULTIVARIATE EMPIRICAL PROCESSES

By J. H. J. EINMAHL and F. H. RUYMGAART
Katholieke Universiteit Nijmegen

SUMMARY. Exploiting a relation with the compound multivariate Poisson process a fundamental fluctuation theorem is established for the compound multivariate empirical process, where the jumps of the empirical d.f. involved are allowed to be random variables. This fluctuation theorem is applied to obtain strong and weak convergence for the weighted compound multivariate processes. It provides, moreover, the tool for proving strong convergence properties of regression function estimators.

1. Introduction

In this paper we consider a generalization of the weighted multivariate empirical process by allowing the jumps of the empirical d.f. involved to be random variables. The mean value function of this generalized process is, in a sense to be specified below, a cumulative version of the regression function; accordingly, properties of this process may be applied in regression function estimation. Since these processes are related with compound Poisson processes they will be referred to as *weighted compound multivariate empirical processes.*

Marcus and Zinn (1984) deal with the univariate non-i.i.d. case using a method quite different from ours and derive e.g. a law of the iterated logarithm. Cox and Mason (1984) establish an embedding result in the univariate i.i.d. case in order to obtain sharper results for the distribution of a statistic, based on so called induced order statistics, introduced by Bhattacharya (1974) for testing against a specific regression function. In this paper we consider the weighted multivariate i.i.d. case. An exponential probability bound is derived for the local fluctuations of the process. This fluctuation inequality, derived in Section 1, provides the basis for some strong and weak convergence properties of the process, given in Section 2, as well as for consistency of regression function estimators, brieffly sketched in Section 3.

The relation with the compound Poisson process enables us to almost copy the proof in Ruymgaart and Wellner (1984), for short RW (1984), where the relation between the ordinary empirical and Poisson process is exploited. The method of proof in RW (1984) can be streamlined, however, since we use stronger conditions almost throughout. Formally, we don't exclude the trivial case where the random jumps are degenerate random variables, and hence the present results generalize some of the results in RW (1984).

AMS (1980) *subject classification :* 60G55, 62G05 ;
Key words and phrases : Compound imperical and Poisson process, nonparametric regression function estimation.

2. Basic inequality and assumptions

Let $(X_1, Y_1), (X_2, Y_2), \ldots$ be a sequence of i.i.d. $(d+1)$-dimensional random vectors. For the present purposes there is no less of generality in assuming that the X_i take their values in the d-dimensional unit square I^d $(I = (0, 1])$; their common d.f. is denoted by F. The random variables Y_i take their values in $\boldsymbol{R}$. Hence the (X_i, Y_i) are random vectors with common d.f. concentrating mass 1 on $I^d \times \boldsymbol{R}$. Typically X_i and Y_i are dependent.

We will write $x = <x_1, \ldots, x_d> = <x_j> = <x(j)> \in \boldsymbol{R}^d$ if we wish to display the coordinates of x. If $x_j = \xi \in \boldsymbol{R}$ for all j we write $x = <\xi>$. The half-open rectangles $(x_1, y_1] \times \ldots \times (x_d, y_d]$ will be preferably written as $R = R(x, y) = R(< x_j >, < y_j >)$. The class

$$\mathcal{R} = \{R(< s_j >, < t_j >) \; ; \; 0 \leqslant s_j < t_j \leqslant 1 \; \forall \, j\} \qquad \ldots \quad (1.1)$$

of all half-open rectangles in I^d will play an important role. We will briefly write

$$R(< 0 >, < t_j >) = R(< t_j >) = R(t), \; t \in I^d. \qquad \ldots \quad (1.2)$$

The Lebesgue measure of any $R \in \mathcal{R}$ is denoted by $|R|$, $|t| = t_1 \times \ldots \times t_d$ and $|dt|$ denotes Lebesgue measure restricted to I^d. Given any (random) function $L : \boldsymbol{R}^d \to \boldsymbol{R}$, $R = R(x, y)$ and disjoint rectangles $R_1, R_2, \ldots$ we define

$$L\{R\} = L\{R(x, y)\} = \Delta_x^y L, \; L\{\textstyle\bigcup_j R_j\} = \Sigma_j \, L\{R_j\}, \qquad \ldots \quad (1.3)$$

where Δ_x^y is the usual difference operator on $\boldsymbol{R}^d$; e.g. the probability of $R(s, t)$ is given by $\Delta_s^t F = F\{R(s, t)\}$.

The compound empirical d.f. based on the first n random vectors in the sequence is defined by

$$\hat{M}_n(t) = n^{-1} \sum_{i=1}^{n} Y_i \, 1_{R(t)}(X_i), \; t \in I^d. \qquad \ldots \quad (1.4)$$

It will be assumed throughout that

the Y_i are bounded and that F is continuous. $\qquad \ldots \quad (1.5)$

For a proper standardization let us introduce

$$E Y_i \, 1_{R(t)}(X_i) = M(t), \; t \in I^d. \qquad \ldots \quad (1.6)$$

The compound empirical process will now be defined as

$$U_n^*(t) = n^{1/2}(\hat{M}_n(t) - M(t)), \; t \in I^d. \qquad \ldots \quad (1.7)$$

For a proper weighing of the process let us consider

$$E Y_i^2 \, 1_{R(t)}(X_i) = V(t), \; t \in I^d. \qquad \ldots \quad (1.8)$$

The weighted compound empirical processes considered in this paper are of the form

$$U_n^*(t)/V^{\frac{1}{2}-\delta}(t),\ t \in I^d,\ \textit{for some}\ \delta \in [0, \tfrac{1}{2}]. \qquad \ldots \quad (1.9)$$

According to (1.3) we may write

$$EY_i\, 1_R(X_i) = M\{R\},\ EY_i^2\, 1_R(X_i) = V\{R\},\ R \in \mathcal{R}. \qquad \ldots \quad (1.10)$$

Also note that when $P(Y_i = 1) = 1$ the compound processes reduce to the ordinary processes; in this case we have $M = V = F$.

To establish the relation with the compound Poisson process let $\nu(n)$ be a Poisson (n) random variable which is independent of the sequence (X_1, Y_1), $(X_2, Y_2), \ldots$ and let

$$\hat{F}_n(t) = n^{-1} \sum_{i=1}^{n} 1_{R(t)}(X_i),\ t \in I^d, \qquad \ldots \quad (1.11)$$

be the ordinary empirical d.f. of the $X_1, \ldots, X_n$. It is well-known that

$$N_n(t) = \nu(n)\hat{F}_{\nu(n)}(t),\ t \in I^d, \qquad \ldots \quad (1.12)$$

is a Poisson process with mean values $nF(t)$. Analogously it turns out that

$$N_n^*(t) = \nu(n)\hat{M}_{\nu(n)}(t),\ t \in I^d, \qquad \ldots \quad (1.13)$$

is a compound Poisson process with mean values $nM(t)$. A standardized version of (1.13) is given by

$$Z_n^*(t) = n^{-1/2}(N_n^*(t) - nM(t)),\ t \in I^d. \qquad \ldots \quad (1.14)$$

The relation with the compound empirical process is given by

$$\{U_n^*(t), t \in I^d\} \overset{d}{=} \{Z_n^*(t), t \in I^d \mid \nu(n) = n\}. \qquad \ldots \quad (1.15)$$

An important property of the process Z_n^* is the independence of its increments. We will need the slightly more general property that the $\boldsymbol{R}^2$-valued process

$$(N_n, Z_n^*)\ \textit{has independent increments}, \qquad \ldots \quad (1.16)$$

with N_n defined in (1.12) and Z_n^* in (1.14). Due to (1.16) a property is quite often relatively easy to prove for Z_n^*. Thanks to relation (1.15) it may then in general be subsequently shown for U_n^*. Let us also note that

$$EN_n^*(t) = nM(t),\ \text{var}\ N_n^*(t) = nV(t),\ t \in I^d, \qquad \ldots \quad (1.17)$$

where M and V have already been defined in (1.6) and (1.8) respectively. This makes $V^{1/2-\delta}$ a rather natural weight function.

In order to prepare for a convenient formulation of the basic inequality and its corollaries let us make the convention that *throughout this paper the* *symbols A, B, C will be used to denote generic constants in* $(0, \infty)$ *that are independent of all the relevant parameters.* In Theorem 1.1 below in particular A, B, C are independent of F, M, V, λ, R and n (but C depends on d).

Theorem 1.1 : *Fluctuation inequality. Let* (1.5) *be fulfilled. For all* $R \in \mathcal{R}$, $n \in \mathbf{N}$ *and* $\lambda \geqslant 0$ *we have*

$$P\left(\sup_{S \subset R} |U_n^*\{S\}| \geqslant \lambda\right) \leqslant C \exp\left(\frac{-A\lambda^2}{V\{R\}} \psi\left(\frac{B\lambda}{n^{1/2}V\{R\}}\right)\right), \quad \ldots \quad (1.18)$$

where $S \in \mathcal{R}$ *and* ψ *is a continuous decreasing function defined by*

$$\psi(\lambda) = 2\lambda^{-2} \int_0^\lambda \log(1+x)dx,\ \lambda > 0;\ \psi(0) = 1. \quad \ldots \quad (1.19)$$

Proof : Representation (1.15) enables us to follow the pattern of the proof of Theorem 1.1 in RW (1984). As in that paper let us choose R such that $F\{R\} \leqslant 1/2$ throughout step 1—step 4.

Step 1. Following the proof of Lemma 1.1 in RW (1984) it may be easily seen that (1.15) and (1.16) entail

$$P\left(\sup_{S \subset R} |U_n^*\{S\}| \geqslant \lambda\right) \leqslant 2P\left(\sup_{S \subset R} |Z_n^*\{S\}| \geqslant \lambda\right),\ \lambda \geqslant 0. \quad \ldots \quad (1.20)$$

Step 2. Using the independent increments property of Z_n^* along with symmetrization it follows that

$$P\left(\sup_{S \subset R} |Z_n^*\{S\}| \geqslant \lambda\right) \leqslant 2^{2d+2} P\left(|Z_n^*\{R\}| \geqslant \frac{1}{4}\lambda\right), \quad \ldots \quad (1.21)$$

provided that

$$\lambda \geqslant (8V\{R\})^{1/2}. \quad \ldots \quad (1.22)$$

The proof is almost identical to the first part of the proof of Theorem 1.1 in RW (1984) and is patterned on the proof of Lemma 1.2 in Orey and Pruitt (1973).

Step 3. A probability inequality for binomial random variables due to Bennett (1962) leads to the inequality (1.10) in RW (1984) for Poisson random variables. In a similar way the inequality

$$P(|Z_n^*\{R\}| \geqslant \lambda) \leqslant 2 \exp\left(\frac{-A\lambda^2}{V\{R\}} \psi\left(\frac{B\lambda}{n^{1/2}V\{R\}}\right)\right),\ \lambda \geqslant 0, \quad \ldots \quad (1.23)$$

may be obtained for the compound Poisson random variable $Z_n^*\{R\}$.

Step 4. Combination of (1.20)–(1.23) leads to the following upper bound for the probability on the left in (1.18) :

$$2^{2d+4} \exp\left(\frac{-A\lambda^2}{V\{R\}}\,\psi\left(\frac{B\lambda}{n^{1/2}V\{R\}}\right)\right), \quad \lambda \geqslant (8V\{R\})^{1/2}. \qquad \ldots \quad (1.24)$$

For $0 \leqslant \lambda < (8V\{R\})^{1/2}$, however, the bound is larger than 1, which entails that (1.24) is an upper bound for all $\lambda \geqslant 0$, still for $R \in \mathcal{R}$ with $F\{R\} \leqslant \frac{1}{2}$.

Step 5. Let us next take any $R \in \mathcal{R}$. Then we can always write R as $R = R_1 \bigcup R_2$ for $R_1, R_2 \in \mathcal{R}$ with $R_1 \bigcap R_2 = \phi$ and $F\{R\} \leqslant \frac{1}{2}$ $(i = 1, 2)$. It is clear that

$$\sup_{S \subset R} |U_n^*\{S\}| \leqslant \sum_{i=1}^{2} \sup_{S \subset R_i} |U_n^*\{S\}|, \qquad \ldots \quad (1.25)$$

and consequently we have

$$P\left(\sup_{S \subset R} |U_n^*\{S\}| \geqslant \lambda\right) \leqslant \sum_{i=1}^{2} P\left(\sup_{S \subset R_i} |U_n^*\{S\}| \geqslant \frac{1}{2}\lambda\right)$$

$$\leqslant C \sum_{i=1}^{2} \exp\left(\frac{-A\lambda^2}{V\{R_i\}}\,\psi\left(\frac{B\lambda}{n^{1/2}V\{R_i\}}\right)\right), \qquad \ldots \quad (1.26)$$

according to what has been obtained so far. It follows from direct computation that

$$\psi(\lambda) \geqslant \varepsilon\, \psi(\varepsilon\, \lambda), \; 0 \leqslant \varepsilon \leqslant 1, \; \lambda \geqslant 0, \qquad \ldots \quad (1.27)$$

and hence

$$\frac{A\lambda^2}{V\{R_i\}}\,\psi\left(\frac{B\lambda}{n^{1/2}V\{R_i\}}\right) \geqslant \frac{A\lambda^2}{V\{R\}}\,\frac{V\{R\}}{V\{R_i\}}\,\frac{V\{R_i\}}{V\{R\}}\,\psi\left(\frac{V\{R_i\}}{V\{R\}}\,\frac{B\lambda}{n^{1/2}V\{R_i\}}\right)$$

$$= \frac{A\lambda^2}{V\{R\}}\,\psi\left(\frac{B\lambda}{n^{1/2}V\{R\}}\right). \qquad \ldots \quad (1.28)$$

It is immediate from (1.24), (1.26) and (1.28) that, using A, B, C as generic constants, the inequality (1.17) holds true for all $R \in \mathcal{R}$. Q.E.D.

Both in the weighted process (1.9) and in the exponential bound (1.18) the same function V plays a role. This means that we are in the same position as RW (1984) where this common function is F rather than V. Hence the remaining results and proofs in RW (1984) carry over almost immediately to our case, provided that we consistently replace F by V. Throughout the remainder of this paper let the following assumption be fulfilled.

Assumption 1.1 : Condition (1.5) is satisfied. The finite signed measure M and the finite positive measures V and F are absolutely continuous with respect to Lebesgue measure with densities satisfying

$$\left.\begin{cases} 0 \leqslant \left| \dfrac{dM(t)}{|dt|} \right| \leqslant c_0 < \infty \\ 0 < c_1 \leqslant \dfrac{dV(t)}{|dt|} \leqslant c_2 < \infty,\ 0 < c_3 \leqslant \dfrac{dF(t)}{|dt|} \leqslant c_4 < \infty \end{cases} \quad \forall t \in I^d. \right\} \quad \ldots (1.29)$$

2. Strong and weak convergence

We first present two corollaries of Theorem 1.1. Let us fix a $\gamma \in (0, 1)$ and define $\mathcal{R}_\gamma$ to be the subclass of $\mathcal{R}$ of all rectangles $R(a, b)$ such that

$$V(a)/V(b) \geqslant \gamma \in (0, 1). \quad \ldots \quad (2.1)$$

For $T \subset I^d$ let us, for $\lambda \geqslant 0$, define

$$p_{n,\delta}(T; \lambda) = P\Big(\sup_{t \in T} \frac{|U_n^*(t)|}{V^{\frac{1}{2}-\delta}(t)} \geqslant \lambda \Big) \quad \ldots \quad (2.2)$$

and

$$q_{n,\delta}(T; \lambda) = P\Big(\sup_{s, t \in T} \left| \frac{U_n^*(s)}{V^{\frac{1}{2}-\delta}(s)} - \frac{U_n^*(t)}{V^{\frac{1}{2}-\delta}(t)} \right| \geqslant \lambda \Big). \quad \ldots \quad (2.3)$$

Corollary 2.1 : *Let* $\lambda \geqslant 0$. *For any* $R(a, b) \in \mathcal{R}_\gamma$ *and* $0 \leqslant \delta \leqslant \frac{1}{2}$ *we have*

$$p_{n,\delta}(R(a, b); \lambda) \leqslant C \, exp \left(\frac{-A\lambda^2}{(V^{\delta/2}(b)} \psi\Big(\frac{B\lambda}{n^{\frac{1}{2}} V^{(1-\delta)/2}(b)} \Big) \right) \quad \ldots \quad (2.4)$$

and

$$q_{n,\delta}(R(a, b); \lambda) \leqslant C \, exp \left(\frac{-A\lambda^2}{(V(b) - V(a))^{\delta/2}} \psi\Big(\frac{B\lambda}{n^{\frac{1}{2}}(V(b) - V(a))^{(1-\delta)/2}} \Big) \right), \quad \ldots \quad (2.5)$$

provided that Assumption 1.1 *is satisfied.* (*The generic constants* A, B, C *are also independent of the* a *and* b *but* A *and* B *do depend on* γ.)

The proof is similar to a combination of the proofs of Theorem 2.2, Theorem 2.3 and Corollary 2.1 in RW (1984) and will be omitted.

Corollary 2.2 : *For any* $0 < \varepsilon_1 < \varepsilon_2 \leqslant c_2$, $0 \leqslant \delta \leqslant 1/2$ *and* $\lambda \geqslant 0$ *we have*

$$p_{n,\delta}(\{\varepsilon_1 \leqslant V(t) \leqslant \varepsilon_2\}; \lambda)$$
$$\leqslant C \int_{\{(\varepsilon_1/c_2)\gamma \leqslant |t| \leqslant \varepsilon_2/(\gamma c_1)\}} \frac{1}{|t|} \; exp\Big(\frac{-A\lambda^2}{|t|^{\delta/2}} \psi\Big(\frac{B\lambda}{n^{1/2} \varepsilon_1^{(1-\delta)/2}} \Big) \Big) |dt|, \quad \ldots \quad (2.6)$$

provided that Assumption 1.1 *is satisfied.*

Proof: Let us choose $\vartheta \in (0, 1)$ and consider the countably infinite partition $\mathcal{P} \subset \mathcal{R}$ of I^d defined by the rectangles

$$\{R(<\vartheta^{k(j)}>, <\vartheta^{k(j)-1}>), <k(j)> \in N^d\}. \qquad \dots \quad (2.7)$$

Since for each $R \in \mathcal{P}$ we have

$$\frac{V(<\vartheta^{k(j)}>)}{V(<\vartheta^{k(j)-1}>)} \geqslant \frac{c_1}{c_2}\vartheta^d, \qquad \dots \quad (2.8)$$

it follows that

$$\mathcal{P} \subset \mathcal{R}_\gamma, \text{ for } \gamma = (c_1/c_2)\,\vartheta^d \in (0, 1). \qquad \dots \quad (2.9)$$

Let $\mathcal{P}(\varepsilon_1, \varepsilon_2) \subset \mathcal{P}$ consist of all $R(a, b) \in \mathcal{P}$ with $V(a) \leqslant \varepsilon_2$ and $V(b) \geqslant \varepsilon_1$. Together with (1.30) relation (2.9) implies

$$\{\varepsilon_1 \leqslant V(t) \leqslant \varepsilon_2\} \subset \bigcup_{R \in \mathcal{P}(\varepsilon_1, \varepsilon_2)} R \subset \{(\varepsilon_1/c_2)\gamma \leqslant |t| \leqslant \varepsilon_2/(\gamma c_1)\}. \qquad \dots \quad (2.10)$$

It is immediate from Corollary 2.1 that for each $R(a, b) \subset \mathcal{P}(\varepsilon_1, \varepsilon_2)$ we have

$$p_{n,\delta}(R(a, b); \lambda) \leqslant C \exp\left(\frac{-A\lambda^2}{|a|^{\delta/2}}\,\psi\left(\frac{B\lambda}{n^{1/2}V^{(1-\delta)/2}(b)}\right)\right)$$

$$\leqslant C \exp\left(\frac{-A\lambda^2}{|t|^{\delta/2}}\,\psi\left(\frac{B\lambda}{n^{1/2}\,\varepsilon_1^{(1-\delta)/2}}\right)\right) \forall\ t \in R(a, b). \qquad \dots \quad (2.11)$$

The required upper bound is obtained by noting that

$$p_{n,\delta}(\{\varepsilon_1 \leqslant V(t) \leqslant \varepsilon_2\}; \lambda) \leqslant \sum_{R \in \mathcal{P}(\varepsilon_1, \varepsilon_2)} p_{n,\delta}(R; \lambda)$$

and by applying (2.11) to each term in this sum. The transition to the integral is due to the particular choice of $\mathcal{P}$ and is identical to the proof of the corresponding part of Lemma 3.1 in RW (1984). Q.E.D.

Theorem 2.1: *Strong convergence. Let Assumption* 1.1 *be fulfilled. For each* $\varepsilon > 0$ *there exists* $c \in (0, \infty)$ *such that*

$$\limsup_{n \to \infty} \sup_{\{V(t) \geqslant (\varepsilon \log n)/n\}} \frac{|U_n^*(t)|}{(V(t) \log n)^{1/2}} \leqslant c, \ a.s. \qquad \dots \quad (2.12)$$

Proof: The proof is immediate from Corollary 2.2. Note that $V(t) \leqslant c_2$ for all t. Using the upper bounds in (2.11) it turns out that

$$\sum_{n=1}^{\infty} p_{n,\delta}(\{(\varepsilon \log n)/n \leqslant V(t) \leqslant c_2\}; c(\log n)^{1/2}) < \infty, \qquad \dots \quad (2.13)$$

for $c \in (0, \infty)$ sufficiently large. Q.E.D.

Theorem 2.2 : *Weak convergence. Let Assumption 1.1 be fulfilled. For each $0 < \delta \leqslant 1/2$ we have that*

$$\frac{U_n^*}{V^{1/2-\delta}} \xrightarrow{d} G_\delta, \text{ as } n \to \infty, \text{ on } D(I^d), \qquad \dots \quad (2.14)$$

where $D(I^d)$ is endowed with the usual metric and G_δ is a continuous Gaussian process on $D(I^d)$.

Proof : The convergence of the finite dimensional distributions is immediate. To prove the tightness let, for $\alpha \in (0, 1)$, $\mathcal{P}_\alpha$ denote the partition of I^d defined by the squares

$$\{R(< l(k(j)-1)>, < lk(j) >), <k(j)> \in \mathbf{N}^d, k(j) \leqslant \frac{1}{l} \ \forall_j\}, \frac{1}{l} = \left[\frac{2c_2 d}{(1-\gamma)\alpha}\right], \qquad \dots \quad (2.15)$$

for some $\gamma \in (0, 1)$. Let $\mathcal{P}_\alpha(\alpha) \subset \mathcal{P}_\alpha$ be the subclass of all $R(a, b) \in \mathcal{P}_\alpha$ with $V(b) \geqslant \alpha$. It follows that

$$V(a)/V(b) \geqslant (V(b)-dl)/V(b) \geqslant 1-c_2 dl/\alpha \geqslant \gamma \in (0, 1), \qquad \dots \quad (2.16)$$

for each $R(a, b) \in \mathcal{P}_\alpha(\alpha)$, and hence

$$\mathcal{P}_\alpha(\alpha) \subset \mathcal{R}_\gamma \quad \forall \ \alpha \in (0, 1). \qquad \dots \quad (2.17)$$

It suffices to prove that

$$\lim_{\alpha \downarrow 0} \limsup_{n \to \infty} P\Big(\max_{R \in \mathcal{P}_\alpha} \sup_{s, t \in R} \Big| \frac{U_n^*(s)}{V^{1/2-\delta}(s)} - \frac{U_n^*(t)}{V^{1/2-\delta}(t)} \Big| \geqslant \lambda\Big) = 0 \ \forall \lambda > 0 \dots \quad (2.18)$$

Let us note, however, that the probability in (2.18) is bounded above by

$$\sum_{R \in \mathcal{P}_\alpha(\alpha)} q_{n,\delta}(R ; \lambda) + p_{n,\delta}(\{0 \leqslant V(t) \leqslant \alpha\}; \lambda/2). \qquad \dots \quad (2.19)$$

For each rectangle $R(a, b) \in \mathcal{P}_\alpha(\alpha)$ we have

$$dc_1\Big(\Big[\frac{2c_2 d}{(1-\gamma)\alpha}\Big]\Big)^{-1} \leqslant V(b)-V(a) \leqslant dc_2\Big(\Big[\frac{2c_2 d}{(1-\gamma)\alpha}\Big]\Big)^{-1}. \qquad \dots \quad (2.20)$$

It is immediate from (2.20) and Corollary 2.1 that

$$\lim_{\alpha \downarrow 0} \limsup_{n \to \infty} \sum_{R \in \mathcal{P}_\alpha(\alpha)} q_{n,\delta}(R; \lambda)$$

$$\leqslant \lim_{\alpha \downarrow 0} \limsup_{n \to \infty} C\Big(\frac{1}{\alpha}\Big)^d \exp\Big(\frac{-A\lambda^2}{\alpha^{\delta/2}} \ \psi\Big(\frac{B\lambda}{n^{1/2}\alpha^{(1-\delta)/2}}\Big)\Big)$$

$$= \lim_{\alpha \downarrow 0} C\Big(\frac{1}{\alpha}\Big)^d \exp\Big(\frac{-A\lambda^2}{\alpha^{\delta/2}}\Big) = 0. \qquad \dots \quad (2.21)$$

Let us notice, moreover, that

$$p_{n,\delta}(\{0 \leqslant V(t) \leqslant \alpha\}; \lambda/2) \leqslant p_{n,\delta}(\{0 \leqslant V(t) \leqslant \alpha_n\}; \lambda/2)$$
$$+p_{n,\delta}(\{\alpha_n \leqslant V(t) \leqslant \alpha\}; \lambda/2), \text{ with } \alpha_n = n^{-(1+\delta)}. \quad \ldots \quad (2.22)$$

It is immediate from Corollary 2.2 that

$$\lim_{\alpha \downarrow 0} \limsup_{n \to \infty} p_{n,\delta}(\{\alpha_n \leqslant V(t) \leqslant \alpha\}; \lambda/2) = 0. \quad \ldots \quad (2.23)$$

In order to deal with $p_{n,\delta}(\{0 \leqslant V(t) \leqslant \alpha_n\}; \lambda/2)$ first note that

$$\frac{|U_n^*(t)|}{V^{1/2-\delta}(t)} = \frac{n^{1/2}|M(t)|}{V^{1/2-\delta}(t)} \leqslant \frac{n^{1/2}c_0|t|}{(c_1|t|)^{1/2-\delta}}$$
$$\leqslant c_0 c_1^{\delta-1/2} n^{1/2} n^{-(1+\delta)(1/2+\delta)} \;\forall\; t : V(t) \leqslant \alpha_n, \quad \ldots \quad (2.24)$$

provided that $\min_{1 \leqslant i \leqslant n} V(X_i) > \alpha_n$.

The upper bound in (2.24) converges to 0 as $n \to \infty$, and hence to prove

$$\limsup_{n \to \infty} p_{n,\delta}(\{0 \leqslant V(t) \leqslant \alpha_n\}; \lambda/2) = 0, \quad \ldots \quad (2.25)$$

it remains to show that

$$P\Big(\min_{1 \leqslant i \leqslant n} V(X_i) > \alpha_n\Big) \to 1, \text{ as } n \to \infty. \quad \ldots \quad (2.26)$$

This last relation is immediate from Assumption 1.1. See also the proof of Theorem 3.4 in RW (1984). Q.E.D.

3. Estimation of the regression function

The regression function to be estimated is defined by

$$m(t) = E(Y_i | X_i = t),\ t \in I^d. \quad \ldots \quad (3.1)$$

It is related with the function M according to

$$M(t) = \int_{R(t)} m(s)dF(s), \text{ or } m(t) = \frac{dM(t)}{dF(t)},\ t \in I^d. \quad \ldots \quad (3.2)$$

Let us define

$$R_{n,t} = R(<t_j - l_n/2>, <t_j + l_n/2>),\ l_n \in (0, 1). \quad \ldots \quad (3.3)$$

Using a histogram type kernel determined by these squares a natural estimator appears to be

$$\hat{m}_n(t) = \frac{\hat{M}_n\{R_{n,t}\}}{\hat{F}_n\{R_{n,t}\}} = \frac{\hat{\mu}_n(t)}{\hat{f}_n(t)},\ t \in I^d, \quad \ldots \quad (3.4)$$

where

$$\hat{\mu}_n(t) = \frac{\hat{M}_n\{R_{n,t}\}}{|R_{n,t}|}, \; \widehat{f_n(t)} = \frac{\hat{F}_n\{R_{n,t}\}}{|R_{n,t}|}, \; t \in I^d. \quad \ldots \quad (3.5)$$

Convergence properties of $\hat{m}_n$ may be derived from convergence properties of $\hat{\mu}_n$ and $\hat{f}_n$ separately. Let us focus on $\hat{\mu}_n$ and introduce

$$\mu(t) = \frac{dM(t)}{|dt|}, \; t \in I^d. \quad \ldots \quad (3.6)$$

As in the proof for density estimators, we consider the two parts

$$\|\hat{\mu}_n - \mu_n\| \quad and \quad \|\mu_n - \mu\|, \quad \ldots \quad (3.7)$$

where $\|\cdot\|$ denotes the supremum norm and

$$\mu_n(t) = E\hat{\mu}_n(t) = \frac{M\{R_{n,t}\}}{|R_{n,t}|}, \; t \in I^d. \quad \ldots \quad (3.8)$$

We will restrict ourselves to the convergence of the random part in (3.7) at the usual rate.

Theorem 3.1 : *Choose an arbitrary* $0 < \delta \leqslant 1/2$ *and let*

$$l_n^d = n^{-\delta}. \quad \ldots \quad (3.9)$$

Provided that Assumption 1.1 *is satisfied we have*

$$n^{1/2-\delta} \; \| \hat{\mu}_n - \mu_n \| \xrightarrow{\text{a.s.}} 0, \; as \; n \to \infty. \quad \ldots \quad (3.10)$$

Proof : This theorem follows almost immediately from Theorem 1.1. Let $\mathcal{P}_n$ be the partition of I^d defined by the squares

$$R\Big(<\frac{k(j)-1}{m_n}>, <\frac{k(j)}{m_n}>\Big), \; m_n = \lceil 1/l_n \rceil, \; k(j) \in \boldsymbol{N} \; and \; k(j) \leqslant m_n \; \forall j. \ldots \quad (3.11)$$

Notice that

$$\# \mathcal{P}_n = m_n^d \leqslant l_n^{-d}, \; |R| = m_n^{-d} \; for \; R \in \mathcal{P}_n, \quad \ldots \quad (3.12)$$

and that any $R_{n,t}$ as in (3.3) intersects at most 2^d adjacent rectangles of the partition $\mathcal{P}_n$.

Hence it is clear that

$$\begin{aligned} n^{1/2-\delta} \| \hat{\mu}_n - \mu_n \| &\leqslant 2^d \, n^{1/2-\delta} \, l_n^{-d} \max_{R \in \mathcal{P}_n} \sup_{S \subset R} | \hat{M}_n\{S\} - M\{S\} | \\ &= 2^d \, n^{-\delta} \, l_n^{-d} \max_{R \in \mathcal{P}_n} \sup_{S \subset R} | U_n^*\{S\} |. \quad \ldots \quad (3.13) \end{aligned}$$

Condition (1.30) entails $c_1 m_n^{-d} \leqslant V\{R\} \leqslant c_2 m_n^{-d}$ for each $R \in \mathscr{P}_n$. Jointly with (3.13) and Theorem 1.1 this entails

$$\begin{aligned} & P(n^{1/2-\delta} \, \| \hat{\mu}_n - \mu_n \| \geqslant \lambda) \\ & \leqslant C \sum_{R \in \mathscr{P}_n} \exp(-A\lambda^2 n^{2\delta} \, l_n^{2d} \, m_n^d \, \psi(B\lambda n^{\delta-1/2} l_n^d \, m_n^d)) \\ & \leqslant Cn^{\delta} \exp(-A\lambda^2 n^{\delta} \, \psi(B\lambda n^{\delta-1/2})) \\ & \leqslant Cn^{\delta} \exp(-A\lambda^2 n^{\delta}), \quad \lambda \geqslant 0. \qquad \ldots \text{(3.14)} \end{aligned}$$

It is obvious that

$$C \sum_{n=1}^{\infty} n^{\delta} \exp(-A\lambda^2 n^{\delta}) < \infty \quad \forall \, \lambda > 0, \qquad \ldots \text{(3.15)}$$

and hence (3.10) follows. Q.E.D.

In order to obtain strong convergence properties for $\hat{m}_n - m$, additional smoothness conditions on μ and f will be needed to insure the convergence of $\|\mu_n - \mu\|$ and almost sure convergence of $\|\hat{f}_n - f\|$. Since these considerations are well-known and, moreover, not based on properties of U_n^* they will be omitted.

Acknowledgement. Part of this work was done while the second author was visiting the Indian Statistical Institute, Calcutta. He wishes to express his gratitude for the hospitality and the stimulating environment.

REFERENCES

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

Bhattacharya, P. K. (1974): Convergence of sample paths of normalized sums of induced order statistics. *Ann. Statist.*, **2**, 1034-1039.

Cox, D. D. and Mason, D. M. (1984): Personal communication.

Marcus, M. B. and Zinn, J. (1984): The bounded law of the iterated logarithm for the weighted empirical distribution process in the non-i.i.d. case. *Ann. Probability*, **12**, 335-360.

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

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

Paper received : January, 1985.

Revised : March, 1985.