The Annals of Probability
1984, Vol. 12, No. 4, 990–998

DISCUSSION OF THE PAPER OF PROFESSORS GINÉ AND ZINN

PROFESSOR KENNETH S. ALEXANDER (*University of Washington*). The most important idea in this paper, perhaps, is the comparison to a Gaussian process of a process (here an empirical process) involving summation of random variables, conditionally on the values of those variables. This technique will no doubt find broader application; in fact, Ronald Pyke and I are already attempting to use it to obtain central limit theorems for partial sum processes. These are processes of the form

$$Z_n(A) = n^{-d/2} \sum_{j \in \mathbb{Z}_+^d \cap nA} X_j, \quad A \in \mathbf{A}$$

where $\mathbf{A}$ is a collection of subsets of $[0, 1]^d$, $\mathbb{Z}_+$ denotes the nonnegative integers, nA is $\{nx: x \in A\}$, and $\{X_j: j \in \mathbb{Z}_+^d\}$ is an array of i.i.d. random variables. We hope to reduce the moment condition required on X_j for the CLT in Bass and Pyke (1984) toward the minimal condition $EX_j^2 < \infty$, under metric entropy conditions on $\mathbf{A}$. Z_n can be represented as a weighted sum of processes each of which is qualitatively like an empirical process, and this sum may be compared conditionally to a weighted sum of Gaussian processes.

The "square root trick" (Lemma 5.2 in the paper) gives a convenient and ingenious method of bounding

$$\Pr{}^*[\sup_{f,g \in \mathscr{F}, \rho^2(f,g) \le \varepsilon/n^{1/2}} |\nu_n(f-g)| > \tau\varepsilon]$$

in (3.2) in the paper (see Remark 5.3). The bound is not sharp, however—a factor of 2 is lost, for example, when (2.3) of Lemma 2.7 is used. For the CLT sharpness is of course not needed, but it becomes important for other asymptotic results, including laws of the iterated logarithm (Alexander, 1984, Kiefer, 1961) and minimax properties of the empirical distribution function as an estimator of the true d.f. (Dvoretzky, Kiefer, and Wolfowitz, 1956, Kiefer and Wolfowitz, 1959). For sharp bounds, rather than randomize through use of Rademacher variables ε_i and the square root trick, the idea is to take $N > n$ i.i.d. variables $X_1, \cdots, X_N$, then randomly select n of the N and construct P_n from these; $P_n - P_N$ is then studied. This technique is used in Alexander (1984), Devroye (1982), and Massart (1983).

Finally, it would be of interest to know whether (5.1), (5.14), and (5.15) are actually equivalent statements.

REFERENCES

ALEXANDER, K. S. (1984). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* **12** 1041–1067.

BASS, R. and PYKE, R. (1984). Functional law of the iterated logarithm and uniform central limit theorem for partial-sum processes indexed by sets. *Ann. Probab.* **12** 13–34.

DEVROYE, L. (1982). Bounds for the uniform deviation of empirical measures. *J. Multivariate Anal.* **12** 72–79.

DVORETZKY, A., KIEFER, J., and WOLFOWITZ, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27** 642–669.