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.

KIEFER, J. (1961). 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.

KIEFER, J. and WOLFOWITZ, J. (1959). Asymptotic minimax character of the sample d.f. for vector chance variables. *Ann. Math. Statist.* **30** 463–489.

MASSART, P. (1983). Vitesse de convergence dans le théorème de la limite centrale pour le processus empirique. Thesis, Université de Paris-Sud, Centre d'Orsay.

DEPARTMENT OF MATHEMATICS, GN-50
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195

PROFESSOR R. M. DUDLEY (*Massachusetts Institute of Technology*). For central limit theorems (CLT) on empirical measures uniformly over classes of sets, there had been, roughly, one criterion (asymptotic equicontinuity) and at least three disparate sufficient conditions: 1) an integral condition on metric entropy with inclusion, 2) the Vapnik–Červonenkis combinatorial condition with some measurability, and 3) to have a sequence of sets converging fast enough to the empty set and/or whole space. Now Giné and Zinn have illuminated much more of the surrounding landscape, providing additional criteria, and further sufficient conditions which are at least close to necessary.

Limit theorems for empirical measures uniformly over classes of functions are equivalent to limit theorems in general Banach spaces (Dudley, 1984) (the former are not only special cases of the latter). For example, E. Mourier's classical strong law of large numbers in separable Banach spaces can be subsumed by an extension of the Blum–DeHardt strong law for empiricals (Dudley, 1984, Sec. 6.1), although the already short proof is not thereby shortened.

There have also been disparate conditions for the CLT in separable Banach spaces: the Jain–Marcus (1975) theorem and results depending on the geometry of the space, notably Hoffman–Jørgensen and Pisier's (1976) result that the classical conditions $EX_1 = 0$, $E\|X_1\|^2 < \infty$ imply the CLT in Type 2 spaces. See also Giné (1981). Zinn (1977) deduced the Jain–Marcus theorem from the Type 2 theorem, but one can still ask for a more general result including both, of the kind in the paper under discussion. Giné and Zinn deduce, from their CLT (8.11) for unbounded classes $\mathscr{F}$, the theorem of Pollard (1982). This, in turn, implies (Dudley, 1984, Sec. 11.3) the Jain–Marcus (1975) theorem. What about the Type 2 theorem?

For metric entropy of functions with bracketing, as far as I know, there is a substantial gap between the conditions proved sufficient for the CLT and counterexamples (Yukich, 1982). We also do not know whether the entropy is being taken in the "right" metric for sharp results.

Relations between the geometry of Banach spaces and combinatorics of Vapnik-Červonenkis classes, not necessarily by way of probability, are also being studied, e.g., Milman (1982) and Pajor (1983).

Pisier (1984) has shown the following. Let $\mathscr{M}(X, \mathscr{A})$ be the set of all finite signed measures on a measurable space $(X, \mathscr{A})$, with total variation norm $\|\cdot\|$. For $\mathscr{C} \subset \mathscr{A}$ let $\|\mu\|_{\mathscr{C}} := \sup_{A\in\mathscr{C}} |\mu(A)|$, $\mu \in \mathscr{M}(X, \mathscr{A})$. Then the identity map from $\mathscr{M}(X, \mathscr{A})$ to itself is a type 2 operator from $\|\cdot\|$ to $\|\cdot\|_{\mathscr{C}}$, if and only if $\mathscr{C}$ is a Vapnik-Červonenkis class.

REFERENCES

DUDLEY, R. M. (1984). A course on empirical processes. Ecole d'été de probabilités de St.-Flour, 1982. *Lecture Notes in Math.* Springer, New York. To appear.

GINÉ, E. (1981). Central limit theorems in Banach spaces: a survey, Probability in Banach Spaces III (Medford, 1980). *Lecture Notes in Math.* **860** 138–152.

HOFFMANN-JØRGENSEN, J., and G. PISIER (1976). The law of large numbers and the central limit theorem in Banach spaces. *Ann. Probab.* **4** 587–599.

JAIN, N. C., and M. MARCUS (1975). Central limit theorems for $C(S)$-valued random variables. *J. Funct. Anal.* **19** 216–231.

MILMAN, V. D. (1982). Some remarks about embeddings of ℓ_1^k in finite-dimensional spaces. *Israel J. Math.* **43** 129–138.

PAJOR, A. (1983). Plongement de ℓ_1^n dans les espaces de Banach complexes. *C. R. Acad. Sci. Paris Sér. I* **296** 741–743.

PISIER, G. (1984). Remarques sur les classes de Vapnik–Červonenkis. Preprint.

POLLARD, D. (1982). A central limit theorem for empirical processes. *J. Austral. Math. Soc. Ser. A.* **33** 235–248.

YUKICH, J. (1982). Convergence of empirical probability measures. Ph.D. thesis, Massachusetts Institute of Technology.

ZINN, J. (1977). A note on the central limit theorem in Banach spaces. *Ann. Probab.* **5** 283–286.

DEPARTMENT OF MATHEMATICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139

PROFESSOR PETER GAENSSLER (*University of Munich*). This is a masterpiece in the spirit of Dudley's recent contributions to the asymptotic theory of empirical measures on arbitrary sample spaces.

The present paper incorporates, in addition, some fundamental and—as the main results (especially Theorem 3.2 and Theorem 5.7) show—powerful ideas and results from the theory of Gaussian and sub-Gaussian processes into the present framework which is mainly concerned with central limit theorems (CLT's) for empirical processes indexed by classes $\mathscr{C}$ of sets or, more generally, classes $\mathscr{F}$ of functions defined on the sample spaces under consideration. Here, by CLT it is meant an invariance principle in probability (as explained in the Introduction) which implies a functional CLT in a certain sense of weak convergence of laws of the corresponding processes as presented in the monograph (Gaenssler, 1984). So, the representation of Dudley's (1978) work in the revisited sense of weak convergence of laws in nonseparable metric spaces (named as $\mathscr{L}_b$-convergence in Gaenssler, 1984) may (and this was one of the main intentions) serve to round off and to link additionally the present paper with classical results whose starting point was Donsker's famous functional CLT for the uniform empirical process as presented in Billingsley's (1968) book.

Since the appearance of Billingsley's (1968) book one knows about the significant impact it had on certain important fields of nonparametric statistics (based on Kolmogorov–Smirnov- or Cramér–von Mises-type statistics). Concerning the present and now really most advanced shape of the *probabilistic* theory on empirical processes on arbitrary sample spaces the question naturally arises about the possible effects on statistical inference based on spatial data.

There is an interesting paper by Thomas W. Sager in the *Annals of Statistics* (1982) on "Nonparametric Maximum Likelihood Estimation of Spatial Patterns" where general Glivenko–Cantelli-type theorems are used in proving consistency results. Further results on asymptotic normality should be possible within the context of the present paper.

On the other hand, when looking into the theory of functional CLT's for empirical $\mathscr{C}$-processes together with the concept of $\mathscr{L}_b$-convergence in the sense of Gaenssler (1984) one realizes that the metric spaces, on which the laws (of the empirical $\mathscr{C}$-processes and the limiting Gaussian processes) can be defined in an appropriate way, depend on the underlying distribution P of the random observations. But in spite of this one can show (i.e. in the case when P is unknown) that the sufficient entropy—and measurability—conditions of Gaenssler, 1984 based on a dominating measure λ (for a class of possible P's) imply functional CLT's for the empirical $\mathscr{C}$-processes based on random observations with distribution P admitting a bounded density w.r.t. λ.

Finally, the results of the present paper on LLN's (Glivenko–Cantelli-type results) in the spirit of the Vapnik–Červonenkis theory are very substantial as well.

It would be interesting to compare these results with recent contributions by Michel Talagrand (personal communication) prepared for the *Annals of Mathematics* where a general (in terms of nonasymptotic conditions) description of the GC-classes (i.e. classes $\mathscr{C}$ of sets or classes $\mathscr{F}$ of functions on which the empirical processes $P_n(f), f \in \mathscr{F}$, converge uniformly to $P(f), f \in \mathscr{F}$) is given.

REFERENCES

GAENSSLER, P. (1984). Empirical processes. Vol. 3. *IMS Lecture Notes Monograph Series.* Ed. by Shanti S. Gupta.

MATHEMATISCHES INSTITUT
UNIVERSITÄT MÜNCHEN
THERESIENSTRASSE 39
D-8000 MÜNCHEN 2
WEST GERMANY

PROFESSOR WALTER PHILIPP (*University of Illinois*). 1. Although the main theme of this beautiful paper is the central limit theorem and the strong law of large numbers, it is perhaps nevertheless worth mentioning that all the results of this paper dealing with the central limit theorem also imply under a sharp moment condition an almost sure invariance principle with rate $o((n \log \log n)^{1/2})$ and as a consequence a functional as well as a compact law of the iterated logarithm. This can be viewed as an extension of the well-known fact that if a sequence of independent identically distributed random variables with values in a separable Banach space satisfies the central limit theorem and a moment condition then the law of the iterated logarithm and even an almost sure invariance principle with error term $o((n \log \log n)^{1/2})$ holds (see Philipp, 1978 and the remarks before Theorem 1.2 of Dudley and Philipp, 1983.) For instance in the case of Theorems 3.1 and 3.2 of the paper under discussion the formal argument goes like this: Since the functions $f \in \mathscr{F}$ are uniformly bounded the moment condition in Dudley and Philipp, 1983, Theorem 1.3 holds with $F = 1$. The proofs of Theorems 3.1 and 3.2 show that conditions (a) and (b) of Theorem 2.12 hold, and thus the above claim follows from Dudley and Philipp, 1983, Theorem 1.3, (1.19) with e_p replaced by ρ_p. (The proof as given on pages 513–514 does not depend on the particular choice of the pseudometric.)

2. For similar theorems frequently the question is raised whether a generalization to weakly dependent random variables is possible. Undoubtedly this question will also be asked in connection with the present paper. Except for a more stringent moment condition, Theorem 1.1 of Dudley and Philipp (1983), and thus the sufficiency part of Theorem 2.12 continue to hold for stationary sequences of mixing random elements. (See Philipp, 1982.) (The necessity part also holds by the proof of Dudley, 1984, Theorem 4.1.1.) But in order to establish condition (b) of Theorem 2.12 and thus applying Theorem 2.12 to special classes of functions, such as indicators of convex sets, a proof based on the chaining argument, as given in Dudley, 1979, Theorem 5.1, requires sharp exponential bounds à la Bernstein. But such bounds are not yet available for mixing random variables. Only much weaker bounds (Philipp, 1982, Theorem 4) have been proved thus far. These bounds permit the metric entropy with inclusion to grow only logarithmically instead of polynomially, as it is permitted to grow for independent random variables. Since the proofs of the paper under discussion also require Bernstein's inequality, (see (2.18) and the proof of Theorem 3.1) there does not appear to be much hope for a generalization of many of these results to mixing random variables.

The only thing I can say about the nonstationary case is: Good luck!

3. This paper deals with empirical processes for classes of functions indexed by (linear) sets $\subset \mathbb{N}$. In a forthcoming paper (Morrow and Philipp, 1984) it is shown that the results of the paper under discussion on the central limit theorem and on the law of the iterated logarithm mentioned in Dudley (1979) continue to hold for empirical processes for classes of functions indexed by classes $\mathscr{A}$ of subsets in $\mathbb{N}^q$ satisfying an entropy condition and having not too much mass on the boundary. Although the following example is perhaps a curiosity, it is a good illustration of what I am writing about since it requires the least amount of notation: Let $\{\xi_j, j \in \mathbb{N}^2\}$ be independent identically distributed random variables with values in $\mathbb{R}^2$ and having uniform distribution over $[0, 1]^2$. Let $\mathscr{C}$ denote the class of convex sets contained in $[0, 1]^2$. Then with probability 1

$$H_n := \sup\{|\textstyle\sum_{j\in nC} (1_D(\xi_j) - \lambda(D) - Y_j(1_D))| : C, D \in \mathscr{C}\}$$
$$= o(n(\log\log n)^{1/2})$$

and

$$n^{-1}H_n \to 0 \quad \text{in} \quad L^2(S, \mathscr{S}, Pr).$$

Here λ denotes Lebesgue measure and $Y_j(1_D)$ are Gaussian random variables. Thanks to the papers by Dudley and to the paper under discussion, the theory of empirical processes indexed by linear sets is now very well developed. I believe it is important to develop the theory of empirical processes for classes of functions indexed by more general sets to the same level by establishing sharp conditions on the classes $\mathscr{A}$ mentioned above.

REFERENCES

DUDLEY, R. M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* **6** 899–929. Correction (1979) ibid. **7** 909–911.

DUDLEY, R. M. (1984). A course on empirical processes. *Lecture Notes in Math.* Springer, New York. To appear.

DUDLEY, R. M. and PHILIPP, W. (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. *Z. Wahrsch. verw. Gebiete* **62** 509–552.

MORROW, G. J. and PHILIPP, W. (1984). Invariance principles for partial sum processes and empirical processes indexed by sets. Unpublished manuscript.

PHILIPP, W. (1978). Almost sure invariance principles for sums of B-valued random variables, Probability in Banach spaces II (Proc. Conf. Oberwolfach.) *Lecture Notes in Math.* **709** 171–193. Springer, Berlin (1979).

PHILIPP, W. (1982). Invariance principles for sums of mixing random elements and the multivariate empirical process. Colloquia Soc. János Bolyai **36**. (Proc. Conf. Limit Theorems in Probability Theory and Statistics, Veszprém, Hungary).

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801

PROFESSOR DAVID POLLARD (*Yale University*). The proof of Theorem 3.2 is beautiful. The reduction to the mixture of Gaussian processes $G_{P_n} = n^{-1/2} \sum_{i=1}^{n} g_i f(X_i)$ will be a much-copied technique in future. It suggests that there may be some way of coupling Gaussian processes G_λ and G_μ: perhaps G_μ can be written as G_λ plus an independent, centered Gaussian process if μ is bigger than λ in some appropriate sense. That would give

$$\mathbb{P}\{\sup_{\mathscr{F}'_\delta} G_\lambda(f) > t\} \leq 2\mathbb{P}\{\sup_{\mathscr{F}'_\delta} G_\mu(f) > t\} \quad \text{for all} \quad t.$$

The details concerning the bounding of $\Pr(A^c)$ could then be interpreted as a proof that P_n concentrates increasingly in the set of λ where this inequality holds, with $\mu = 4P$.

For the chaining argument, as in Theorem 3.1, the sub-Gaussian property enters through the exponential bound on the tail probabilities of empirical processes. The Bernstein inequality provides such a bound only when the variance contribution stays bigger than $\varepsilon n^{-1/2}$; that is why the chaining has to stop when $\rho^2(f, g)$ gets down near $\varepsilon n^{-1/2}$. The Hoeffding inequality gives the sub-Gaussian tails for all pairs of functions; that is why the chaining underlying (8.20) can continue forever.

The fancy definition of $\delta_T^{(2)}(f, g)$ in (8.22) turns out to be unnecessary: Pollard (1984, Section VII.4).

The inequality (5.3) can be interpreted as an assertion that

$$\sup(P_n f^2)^{1/2} \leq 8 \sup(Pf^2)^{1/2} \quad \text{with high probability}$$

if the covering numbers $N_{n,2}$ do not increase too rapidly. The proof can be rewritten as an application of the symmetrization inequality (2.1) to the L^2 seminorms: Pollard (1984, Section II.6).

REFERENCE

POLLARD, D. (1984). *Convergence of Stochastic Processes.* Springer-Verlag, Berlin. To appear.

DEPARTMENT OF STATISTICS
YALE UNIVERSITY
BOX 2179 YALE STATION
NEW HAVEN, CONNECTICUT 06520

PROFESSOR RONALD PYKE (*University of Washington*). The authors are to be commended for their comprehensive unification and extension of recent developments for empirical processes. The thoroughness of their exposition makes discussion difficult.

As one that was involved with the formulation and first proof of the weak convergence result for set-indexed uniform empirical processes, I have watched with considerable interest the rapid development of the subject during the past decade. This paper together with some lecture notes and monographs about to be published shows the excellent state of the literature. There should now be a thrust toward the application of these theoretical advances, as encouraged in my 1978 IMS Special Invited Lecture on "Empirical Processes and Inference". For example, Kolmogorov-type statistics of the form $D_n(C) := \| P_n - P \|_C$ which have been defined for many years, can now be considered for tests of hypotheses (cf. Pyke, 1984). The asymptotic distributions of many such statistics are now known in theory. Practically, however, two important problems remain; (i) accurate simulations are needed to estimate the limiting (and so why not the exact?) distributions, and (ii) computational procedures are needed to facilitate the evaluation of the statistic. In the latter case some interesting theoretical problems involving random sets arise. For example, envisage an interactive work station where the data $X_1, X_2, \ldots, X_n$, is entered into a smart graphics computer which can compute $P_n(A) - P(A)$ for any specified $A \in C$. Assume for simplicity that C is the class of convex sets. If an operator is able to manipulate these sets on a screen, then after a finite number of sets have been sampled, say $C_0 \subset C$, the computer will have calculated $D_n(C_0)$. Here C_0 is a random, data dependent and finite subfamily of C. When the problem is more clearly defined, will it be possible to give probability bounds for $D_n(C) - D_n(C_0)$ in terms of $card(C_0)$? By contrast, if C_0 is a fixed pre-specified family (such as a δ-net for C) the above paper reviews some of the bounds on this difference that are available in terms of the entropy, $card(C_0)$. For applications of course V-C classes should be large enough.

A second problem that brings random sets $A \in C_0$ into the discussion of empirical processes is that of finding a set, $\hat{A}_n$ say, of minimal area among all sets with a specified coverage, say $P_n(A) \geq 1/2$. How does the weak convergence of $\hat{A}_n$ to $\hat{A}$, the minimal area member of C for which $Z(\hat{A}) \geq 1/2$, follow best from the results of this paper?

Empirical processes form a special case of random set functions that arise in many real life situations. Data often is of the form (X, M) where X is a location and M is a measurement taken at X. Such data determine in the obvious way a random signed measure

$$S(A) = \textstyle\sum_{X_i \in A} M_i, \quad A \in C. \tag{1}$$

If the locations are random but the masses are constant, one is led to empirical processes. If on the other hand, the locations are fixed (such as at equally spaced lattice points of a cube) but the masses are random, one obtains the other important special case of partial-sum processes (cf. Pyke, 1984, Bass and Pyke, 1984, and references therein). The ties between partial-sum and empirical processes are well known, particularly in the one-dimensional case. However, it is possible to develop this relationship much further, and this work is in progress.

By introducing $Y_i = (X_i, M_i)$ and $f(x, m) = m1_A(x)$, (1) can be written as $\sum f(Y_i)$ which is in the form central to this paper provided the Y_i's are independent and the X_i's are bounded. Extensions to the unbounded case are considered in a forthcoming dissertation by Mina Ossiander. (Notice that the "unbounded case" referred to in this paper is the case of non-uniformly bounded classes of *bounded* functions.) The need for a CLT in the unbounded case has also arisen in Ossiander and Pyke (1984) where a set-indexed context for Lévy's Brownian process is given.

REFERENCES

OSSIANDER M. and PYKE, R. (1984). Lévy's Brownian motion as a set-indexed process and related central limit theorems. Tech. Rpt. No. 42, Univ. of Washington.

BASS, R. F. 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.

PYKE, R. (1984). Asymptotic results for empirical and partial-sum processes: A review. *Can. J. Statist.* To appear.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195

PROFESSOR WINFRIED STUTE (*Universität Giessen*). The paper under discussion is a further development of the theory of invariance principles for empirical processes as initiated by Professor Dudley. The motivation for Dudley's work seems to have been:

(1) the need for extending the theory of weak convergence (appropriately modified) so as to apply it to distributional results for stochastic processes indexed by quite arbitrary parameter sets.
(2) the observation that tightness, the crucial part in proving weak convergence, is closely connected with the realizability of stochastic processes in certain function sub-spaces, e.g. in classes of continuous functions.

In working out this program it turns out at an already early stage that, when applying some standard exponential bounds, one has to have control over the size of the parameter set. This leads to the notion of metric entropy as introduced below (2.19). Theorem 3.1 stresses the importance of the role played by the metric entropy for $\mathscr{F}$ to be a Donsker class. (3.1) is the familiar boundedness condition on the size of $\mathscr{F}$. The same condition also occurs when proving $\mathscr{F}$ to be pre-Gaussian. Theorem 3.2, which is a major result of the paper, is a real gain on Theorem 3.1. Under no condition on the metric entropy, it reveals the way how the properties of being pre-Gaussian and Donsker are connected.

As mentioned by Professors Giné and Zinn, the results of this paper do not yield a direct approach to the invariance principle for the empirical characteristic function process, i.e. for $\mathscr{F} = \{f_t\}$, with $f_t(x) = \exp[itx]$, $x \in \mathbb{R}$. This is amazing, especially since $\mathscr{F}$ is a uniformly bounded family of nice functions of any desirable degree of smoothness. That the general approach does not lead to any result seems to be due to the fact that even for bounded $\mathscr{F}$'s the invariance principle may be valid only under moment assumptions on P. Such conditions do not have any simple counterpart in the abstract set-up of Theorems 3.1 and 3.2. Similar problems occur when considering $\mathscr{F} = \{f: |f| \le 1, f$ is Lipschitz of order $1\}$. This class appears when studying the dual-bounded Lipschitz distance between P_n and P. Moreover, Huber (1981), page 40, also points out the curse of dimensionality (of the underlying sample space) when dealing with this $\mathscr{F}$.

The paper Stute (1983) has been written mainly to offer an alternative concept of weak convergence, within which such $\mathscr{F}$'s can be easily handled under simple conditions only involving the tails of P.

REFERENCES

HUBER, P. J. (1981). *Robust Statistics.* Wiley. New York.

STUTE, W. (1983). Empirical processes indexed by smooth functions. *Stoch. Process. Appl.* **14** 55–66.

MATHEMATISCHES INSTITUT
UNIVERSITÄT GIESSEN
ARNDTSTRAßE 2
D-6300 GIESSEN
WEST GERMANY