Discussion of: Empirical Processes and Applications by Evarist Giné

Jon A. Wellner [1]

University University of Washington

August, 1993

This is a discussion of the *Forum Lectures* by Evarist Giné on the subject of *Empirical Processes and Applications* presented at the European Meeting of Statisticians held in Bath, England, September 13 - 18, 1992.

The discussion includes short sketches of developments in probability theory (Gaussian process theory, weak convergence theory, and probability in Banach spaces), empirical process theory, and applications thereof in statistics. I comment briefly on the formulation of central limit theorems for empirical processes in terms of the presence or absence of a Gaussian hypothesis, expand on Professor Giné's discussion of the bootstrap, and briefly explain my recent results for exchangeably weighted bootstraps obtained jointly with Jens Praestgaard. The discussion closes with some problems.

[1]Research supported in part by National Science Foundation grants DMS-9108409 and DMS-.

AMS 1980 *subject classifications.* Primary: 60F17, 62E20; secondary 60B12.

Key words and phrases. Bootstrap, exchangeability, empirical process.

1. Introduction. First, let me congratulate Professor Giné on his lucid and enthusiastic lectures. He has done a wonderful job of conveying the feel and excitement connected with recent developments in empirical process theory and the application of this theory in statistics. His simple and elegant presentation of the inequalities clearly shows their power for obtaining the basics of the theory.

There has been tremendous progress over the past 15 years on empirical process theory – and in its applications to problems in statistics. As I tried to argue in my recent review article [Wellner (1992)] the time lag between the introduction of problems and their solution using modern empirical process techniques seems to be decreasing rapidly.

This progress in empirical process theory has gone hand in hand with considerable progress in some of the related areas of probability theory. Three general areas in particular are:

- Gaussian process theory
- weak convergence theory
- probability in Banach spaces.

In the area of Gaussian process theory, major developments include: (a) Exponential bounds resulting from the Borell inequalities; (b) Introduction and use of "majorizing measures" to characterize continuity of Gaussian processes; and (c) Systematic development of Gaussian comparison theorems. Major contributions have been made by Dudley, Marcus and Shepp, Fernique, Borell, Pisier, Sudakov, and Talagrand; see Ledoux and Talagrand (1991), chapters 3 and 11 for much of this.

In the area of weak convergence theory, it was recognized early on by Chibisov (1965) that even the classical empirical process is not Borel-measurable in the nonseparable metric space $(D[0,1], \|\cdot\|_\infty)$. Dudley (1966) suggested one solution to deal with this difficulty, and a useful summary of solutions via separable metrizations is given in Billingsley (1968). However the current modern approach via outer measures did not become clear until the work of Hoffmann-Jørgensen (1984) and Dudley (1985).

Finally, developments in the area of probability in Banach spaces has had a profound impact on empirical process theory. Major developments include the Hoffmann-Jørgensen inequalities and the isoperimetric methods exposited in the recent book by Ledoux and Talagrand (1991). Important contributions have been made by Dudley, Giné, Kuelbs, Ledoux, Pisier, Talagrand, and Zinn among many others.

On the empirical process side of the fence, progress has has also been rapid, building on the probability tools, and in turn providing further problems for the theory. Here's a very brief thumbnail sketch of developments in the theory of (general) empirical processes:

- Vapnik and Chervonenkis (1971): Glivenko - Cantelli theorems for sets.

- Dudley (1978): General central limit theorems for empirical processes indexed by sets: VC-classes and sets with smooth boundaries.
- Kolcinskii (1981) and Pollard (1982): Central limit theorems for classes of square integrable functions satisfying uniform entropy conditions.
- Giné and Zinn (1984): Systematic use of Gaussianization and the *multiplier inequality* begins. [We will elaborate on this below.]
- Alexander (1984): Exponential bounds for suprema of empirical processes indexed by classes of sets and by uniformly bounded classes of functions.
- Massart (1986): Rates of convergence for Pollard's uniform entropy central limit theorem; more exponential bounds.
- Ossiander (1987): General central limit theorem for classes functions satisfying an entropy with bracketing condition.
- Talagrand (1987): Study of measurability issues for the Glivenko - Cantelli theorems.
- Dudley (1987): Study of Universal Donsker classes of functions and entropy bounds for convex hulls of polynomial classes.
- Giné and Zinn (1990): Bootstrap CLT for the general empirical process.
- Giné and Zinn (1991); Sheehy and Wellner (1992): Study of Uniform Donsker classes of functions: exponential bounds for such classes and applications to model - based bootstrapping.

This progress in empirical process theory has enabled a large number of new applications in statistics. Statistical problems have, in turn, continue to generate new and challeging problems for empirical process theory. Here is a short list of selected areas of application in statistics, chosen with a view toward potential for further development and application: [I have not given detailed references in an effort to save space; but see Wellner (1992) for a (nearly) complete list.]

- M - estimators: Huber, Pollard, Arcones and Giné
- Infinite-dimensional M - estimators: Gill, van der Vaart, Murphy.
- The Delta-method: Gill, Dudley.
- Rates of convergence: van de Geer, Birgé and Massart, Wong and Shen.
- Smoothing: Pollard, Yukich, Nolan.
- Nonstandard asymptotics: Kim and Pollard, Nolan, Groeneboom, Donoho.

See Wellner (1992) for a more complete review.

2. Donsker theorems: PreGaussianity as hypothesis or conclusion. In considering the Donsker theorems in Professor Giné's lectures, it is useful to distinguish those theorems which involve a *Gaussian hypothesis* in contrast to those which include existence of a tight $P-$Brownian bridge process as one of the *conclusions*. For example, theorems 2 and 3(b) include the *hypothesis* that the class $\mathcal{F}$ is $P-$preGaussian. On the other hand, Pollard's (1982) corollary of theorem 3(b) and Ossiander's (1987) theorem 4(ii) include $P-$preGaussianity of $\mathcal{F}$ as a part of their conclusions. Hypothesizing preGaussianity of $\mathcal{F}$ provides an interesting way of exploring conditions and of obtaining sharp theorems – but with an additional (often difficult) hypothesis to check in order to apply the result. For most applications in statistics, I find the latter type of Donsker theorems, with preGaussianity as a *conclusion*, more convenient and easier to apply.

Of course preGaussianity as a first step in the proof of a Donsker theorem can be extremely illuminating, and can indeed lead to sufficient conditions: one nice example of this is Marcus (1981) in which necessary and sufficient conditions for weak convergence of the empirical characteristic function are found based on preGaussian considerations.

3. Multiplier Inequalities, Empirical Processes, and Bootstrap Empirical Processes In this section our goal is to expand upon Professor Giné's treatment of the multiplier inequality, and to explain its role and consequences in empirical process theory.

The basic multiplier inequality (Proposition 6 of Giné's lectures) was apparently first discovered (independently) by Pisier and Fernique in 1977 or 1978. It first appeared in Giné and Zinn (1984) (also see Giné and Zinn (1986)) where it was used with Gaussian multipliers; see also Giné and Zinn (1986). Alexander (1985), solving a problem posed by Hoffmann-Jørgensen, shows that no "universal multiplier moment" exists: there is no function $\psi : R^2 \rightarrow R$ so that YZ satisfies the central limit theorem, whenever $EYZ = 0$, $E\psi\Big(|Y|, \|Z\|\Big) < \infty$ and Z satisfies the central limit theorem, for independent real and Banach space valued random elements Y and Z. Proposition 6 continues to hold for mean zero multipliers $\{\xi_i\}$ without the assumption of symmetry at the price of a factor of $2\sqrt{2}$ multiplying the right side. For empirical process theory, the processes $Y_i = \delta_{X_i} - P$ usually, but the multipliers ξ_i can be Gaussian, centered Poisson, symmetrized Poisson, centered exponential, centered $Gamma(4,1)$, and so forth.

The following (unconditional) multiplier central limit Theorem is implicit in Giné and Zinn (1984) and is stated explicitly in Giné and Zinn (1986); also see proposition 10.4, page 279 in Ledoux and Talagrand (1991).

THEOREM 3.1. *(Unconditional multiplier CLT). Let $\mathcal{F}$ be a class of measurable functions. Let $\xi_1, \ldots, \xi_n$ be i.i.d. random variables with mean zero, variance $c^2 > 0$ and $\Lambda_{2,1}(\xi_1) < \infty$, independent from $X_1, \ldots, X_n$. Then the following are equivalent:*

A. *$\mathcal{F}$ is Donsker:* $\mathbb{G}_n \Rightarrow \mathbb{G}_P$ *in* $l^\infty(\mathcal{F})$

B. $n^{-1/2}\sum_{i=1}^n \xi_i(\delta_{X_i} - P) \Rightarrow c\,\mathbb{G}_P$ *in* $l^\infty(\mathcal{F})$.

Ledoux & Talagrand (1986) show that the $L_{2,1}$-hypothesis on the multipliers ξ_i cannot be relaxed in the sense that for every ξ with $\Lambda_{2,1}(\xi) = \infty$ there exists a Banach space valued Y that satisfies the central limit theorem, but ξY does not satisfy the central limit theorem. Ledoux and Talagrand (1986) and Ledoux and Talagrand (1991), Proposition 10.4 on page 279, give a different proof of the basic multiplier inequality.

The following almost sure *conditional multiplier central limit theorem* is due to Ledoux, Talagrand (and Zinn) (1988). It apparently originated simply from a desire to better understand the nature of the multiplier CLT. The original proof of Ledoux and Talagrand (1988) used martingale difference methods orginating in Yurinskii (1974); another proof based on isoperimetric methods is given by Ledoux and Talagrand (1991), Theorem 10.14 on page 293.

THEOREM 3.2. *(Almost sure conditional multiplier CLT). Let $\mathcal{F}$ be a class of measurable functions with $\|Pf\|_{\mathcal{F}} < \infty$. Let $\xi_1, \ldots, \xi_n$ be i.i.d. random variables with mean zero, variance $c^2 > 0$ and $\Lambda_{2,1}(\xi_1) < \infty$, independent from $X_1, \ldots, X_n$. Then the following are equivalent:*

A. *$\mathcal{F}$ is Donsker and $P(F^2) < \infty$:* $\mathbb{G}_n \Rightarrow \mathbb{G}_P$ *in* $l^\infty(\mathcal{F})$

B. $n^{-1/2}\sum_{i=1}^n \xi_i(\delta_{X_i(\omega)} - P) \Rightarrow c\,\mathbb{G}_P$ *almost surely in* $l^\infty(\mathcal{F})$.

The following "in probability" *conditional multiplier central limit theorem* is implicit in Giné and Zinn (1990).

THEOREM 3.3. *("In probability" conditional multiplier CLT). Let $\mathcal{F}$ be a class of measurable functions with $\|Pf\|_{\mathcal{F}} < \infty$. Let $\xi_1, \ldots, \xi_n$ be i.i.d. random variables with mean zero, variance $c^2 > 0$ and $\Lambda_{2,1}(\xi_1) < \infty$, independent from $X_1, \ldots, X_n$. Then the following are equivalent:*

A. *$\mathcal{F}$ is Donsker:* $\mathbb{G}_n \Rightarrow \mathbb{G}_P$ *in* $l^\infty(\mathcal{F})$

B. $\mathbb{Z}_n^\omega \equiv n^{-1/2}\sum_{i=1}^n \xi_i(\delta_{X_i(\omega)} - P) \Rightarrow c\,\mathbb{G}_P$ *in probability in* $l^\infty(\mathcal{F})$; *i.e.* $d_{BL^*}(\mathbb{Z}_n^\omega, c\mathbb{G}_P) \to_p 0$ *where* d_{BL^*} *denotes the dual bounded Lipschitz metric.*

In view of these last two theorems, the bootstrap central limit theorems of Giné and Zinn (1990) (theorem 7 of Giné's lectures) seems quite natural when we realize that the bootstrap empirical processes is just a multiplier process with Multinomial weights: Let $\mathbb{P}_n^\omega$ be the empirical measure of the X_i's as above, let

$$\hat{X}_1, \ldots, \hat{X}_n$$

be a "bootstrap sample" from $\mathbb{P}_n^\omega$, and let $\widehat{N}_n \sim Poisson(n)$ be independent of the X_i's and of the $\hat{X}_i$'s. The bootstrap empirical process $\widehat{\mathbb{G}}_n$ is

$$\widehat{\mathbb{G}}_n = \sqrt{n}(\widehat{\mathbb{P}}_n - \mathbb{P}_n^\omega) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n \delta_{\hat{X}_i} - \mathbb{P}_n^\omega\right)$$
$$= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n \hat{M}_{ni}\delta_{X_i(\omega)} - \mathbb{P}_n^\omega\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^n (\hat{M}_{ni} - 1)\,\delta_{X_i(\omega)}$$

where

$$\widehat{\underline{M}}_n \sim Mult_n(n, (\frac{1}{n}, \ldots, \frac{1}{n})) \qquad \text{is independent of the } X_i\text{'s.}$$

We can write

$$\widehat{\underline{M}}_k = \sum_{j=1}^k \underline{1}_j = \sum_{j=1}^k (1_{1j}, \ldots, 1_{nj})$$

where

$$(1_{1j}, \ldots, 1_{nj}) \sim Mult_n\left(1, \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)\right)$$

are i.i.d., $j = 1, \ldots, k$, $k = 1, 2, \ldots$. Note that if we "Poissonize" $\widehat{\underline{M}}_n$ by forming $\widehat{\underline{M}}_{\hat{N}_n}$, the result is:

$$\widehat{\underline{M}}_{\hat{N}_n} \sim (\xi_1, \ldots, \xi_n)$$

where $\xi_1, \ldots, \xi_n$ are i.i.d. Poisson(1). This fact is exploited by Klaassen and Wellner (1992) to give alternative proofs of the Giné and Zinn (1990) bootstrap central limit theorems based on the multiplier CLT's 2.2 and 2.3.

Now we turn to alternative bootstrap methods based on *exchangeable weights* instead of the multinomial weights used in Efron's bootstrap. A key element of the proof is the following multiplier inequality for exchangeable weights given in Praestgaard and Wellner (1993). It shows that the expectation of the norm of an "exchangeably weighted" bootstrap empirical process can be bounded by the expectation of a "randomly permuted" sum – with a random permutation $\underline{R}$ playing the role of the Rademacher random variables in the unconditional multiplier inequality, lemma 2.1.

LEMMA 3.1. *(Exchangeable Multiplier Inequality). Let $W \equiv (W_1, \cdots, W_n)$ be a nonnegative, exchangeable random vector with $\Lambda_{2,1}(W_1) < \infty$, and let R denote a random permutation uniformly distributed on Π_n, the set of permutations of $\{1, \cdots, n\}$. Let $Y_1, \cdots, Y_n$ be random elements of $l^\infty(\mathcal{F})$ so that (W, R) and $(Y_1, \cdots, Y_n)$ are independent (in fact defined on a product probability space). Let $\|\cdot\|$ denote a pseudonorm on $l^\infty(\mathcal{F})$. Then for any $n_0 < n$*

$$E^*\|\frac{1}{\sqrt{n}}\sum_{j=1}^n W_j Y_j\| \le \frac{n_0}{\sqrt{n}} E(\max_{j\le n} W_j)\frac{1}{n}E^*\sum_{j=1}^n \|Y_j\|$$
$$+ \Lambda_{2,1}(W_1) \max_{n_0<k\le n} E^*\|\frac{1}{\sqrt{k}}\sum_{j=n_0+1}^k Y_{R(j)}\|$$

where in the second line the expectation is with respect to both $Y_1, \cdots, Y_n$ and R.

In most of our applications of lemma 2.1 the Y_j's are deterministic.

This version of the multiplier inequality plays a key role, together with Hoeffding's (1963) inequality relating sampling without replacement to sampling with replacement, in the the proof of the following "exchangeably weighted bootstrap central limit theorem". In fact the following theorem, obtained in joint work with Aad van der Vaart, is a slight generalization of the exchangeable bootstrap central limit theorem of Praestgaard and Wellner (1993). It is formulated to include the important case of sampling without replacement. To do this requires introduction of a new norming sequence r_n and hypotheses on the weights $\{W_{ni}\}$ as follows:

A1. $\{W_{ni}\}$ are nonnegative and exchangeable.

A2. $\sum_{i=1}^n W_{ni} = n$.

A3. $\sup_n \|r_n W_{n1}\|_{2,1} < \infty$.

A4. $\lim_{n\to\infty} \sup_{t \geq \epsilon\sqrt{n}} t^2 \Pr(|r_n W_{n1}| > t) = 0$ for every $\epsilon > 0$.

A5. $\frac{r_n^2}{n} \sum_{i=1}^n (W_{ni} - 1)^2 \to_p c^2 > 0$.

The sampling without replacement weights $\{W_{ni}\}$ defined by

$$W_{ni} \equiv \frac{n}{m} \sum_{j=1}^n 1\{R_j = i\}, \qquad i = 1, \ldots, n\,,$$

where $R = (R_1, \ldots, R_n)$ is a random permutation of $\{1, \ldots, n\}$, satisfy these conditions with $r_n = (m/n)/(1 - (m/n))$ and $c = 1$ if $\sup(m/n) < 1$.

In this setting we redefine the bootstrap empirical process to be

$$\widehat{\mathbb{G}}_n^W \equiv r_n \sqrt{n}(\widehat{\mathbb{P}}_n^W - \mathbb{P}_n^\omega) = \frac{r_n}{\sqrt{n}} \sum_{i=1}^n (W_{ni} - 1)\delta_{X_i(\omega)}\,.$$

The following theorem shows that these conditions suffice for conditional weak convergence of the bootstrap process.

THEOREM 3.4. *("Exchangeably weighted" bootstrap CLT). Suppose that $\mathcal{F}$ is a Donsker class, and, for each n, $W = (W_{n1}, \ldots, W_{nn})$ is a vector of weights satisfying conditions A1 - A5 for a sequence r_n with $r_n^2 = o(n)$. Then, under measurability,*

$$\widehat{\mathbb{G}}_n^W \Rightarrow c\,\mathbb{G}_P \qquad \text{in probability in} \qquad l^\infty(\mathcal{F});$$

as $n \to \infty$. If $P^*\|f - Pf\|_{\mathcal{F}}^2 < \infty$ then the convergence is also outer almost surely.

The proof follows the same pattern as the proof of the main result in Praestgaard and Wellner (1993); in fact the main modification needed is in the proof of finite-dimensional convergence. A complete proof will be given in van der Vaart and Wellner (1994).

Here is the corollary for sampling without replacement. Write

$$\tilde{\mathbb{P}}_{m,n} = \frac{1}{m}\sum_{i=1}^{m} \delta_{X_{R_i}}, \qquad \tilde{\mathbb{Q}}_{n-m,n} = \frac{1}{n-m}\sum_{i=m+1}^{n} \delta_{X_{R_i}} .$$

COROLLARY 3.1. *(Bootstrap Without Replacement). Let $\mathcal{F}$ be a Donsker class. If $m \wedge (n-m) \to \infty$, then, subject to measurability,*

$$\sqrt{\frac{nm}{n-m}}\,(\tilde{\mathbb{P}}_{m,n} - \mathbb{P}_n) = \sqrt{\frac{m(n-m)}{n}}\left(\tilde{\mathbb{P}}_{m,n} - \tilde{\mathbb{Q}}_{n-m,n}\right) \Rightarrow \mathbb{G}_P$$

given $X_1, X_2, \ldots$ in probability. Here $\mathbb{G}_P$ is a tight $P-$Brownian bridge process. If in addition $\mathcal{F}$ possesses an envelope function F with $P^*F^2 < \infty$, then the weak convergence also holds given almost every sequence $X_1, X_2, \ldots$.

PROOF OF COROLLARY 2.1. First note that

$$\sqrt{\frac{nm}{n-m}}\,(\tilde{\mathbb{P}}_{m,n} - \mathbb{P}_n) = \frac{r_n}{\sqrt{n}}\sum_{i=1}^{n}(W_{ni}-1)\delta_{X_i}$$

for the bootstrap without replacement weights $\{W_{ni}\}$.

Suppose first that $\limsup_n \lambda_n \equiv \limsup_n (m/n) < 1$. The the result follows from the preceding theorem by checking that the sampling without replacement weights $\{W_{ni}\}$ satisfy A1 - A5 with the choice $r_n^2 \equiv (m/n)/(1-(m/n))$ and with $c = 1$. In fact the hypothesis A5 holds with exact equality.

If $\limsup_n \lambda_n \le 1$, but $\liminf_n \lambda_n > 0$, then we can argue the same way, but using instead the identity

$$-\sqrt{\frac{n(n-m)}{m}}(\tilde{\mathbb{Q}}_{n-m,n} - \mathbb{P}_n) = \sqrt{\frac{m(n-m)}{n}}\left(\tilde{\mathbb{P}}_{m,n} - \tilde{\mathbb{Q}}_{n-m,n}\right);$$

Thus we have the desired conclusion in probability if *either* $\liminf_n \lambda_n > 0$ or $\limsup_n \lambda_n < 1$. But for any given subsequence $\{n'\}$ there exists a further subsequence $\{n''\}$ such that $\lambda_{n''}$ converges to some number in $[0,1]$, and for this subsequence one of the preceding arguments yields the convergence of the process to $\mathbb{G}_P$ along this further subseqence. □

This theorem for "bootstrap sampling without replacement" is closely related to some nice results of Romano and Politis (1992) concerning validity of the "sampling without replacement bootstrap" for a general real-valued statistic T_n with $\tau_n(T_n - \theta) \to_d Z$; their results are focussed on the case $m/n \to 0$. It is also closely related to results for two-sample permutation test obtained in Praestgaard (1992). The main difference is that in Praestgaard (1992) it is important to study the two-sample permutation empirical process under fixed alternatives $P \neq Q$, whereas the above theorem corresponds to the null hypothesis $P = Q$ in Praestgaard's two-sample setting.

The main point to be made here is that the multiplier inequalities, and the multiplier central limit theorems based thereon, are very useful for a variety of statistical problems.

4. Problems. Here are a few selected problems connected with empirical process theory and the application of this theory to statistics. Problem 4 below is from Pyke (1992), to which we refer for further problems in connection with "product processes". Not all of the problems are directly connected to Professor Giné's lectures

1. Suppose that $(S, \mathcal{S})$ is an arbitrary sample space. Is there always a class of functions $\mathcal{F}$ satisfying: (i) $\mathcal{F}$ is $P-$Donsker; (ii) $\mathcal{F}$ is a determining class (i.e. $\int f dP = \int f dQ$ for all $f \in \mathcal{F}$ implies $P = Q$?

2. Is there an analogue of the Hoffmann-Jørgensen inequality for $U-$processes?

3. Is there a $\mathcal{P}-$uniform version of the Giné-Zinn bootstrap theorem?

4. Give conditions on $r = r_n \to \infty$ and $\mathcal{C}_r \subset \mathcal{S}^r$ so that
$$\|\mathbb{P}_n - \mathrm{P}\|^*_{\mathcal{C}_{r_n}} \to_{a.s.} 0\,.$$

5. When does the bootstrap "work" for dependent data? When does the bootstrap "work" for U-processes?

6. In what sense(s) is $\mathbb{P}_n$ optimal as an estimator of P? What are the appropriate extensions of the classical results of Dvoretzky, Kiefer, and Wolfowitz (1956), Kiefer and Wolfowitz (1958)?

7. Suppose that $(X_1, Y_1), \ldots, (X_n, Y_n), \ldots$ are i.i.d. as H on $\mathcal{X} \times \mathcal{Y}$ with empirical measure $\mathbb{H}_n$, and marginal empirical measures $\mathbb{P}_n$ and $\mathbb{Q}_n$ on $\mathcal{X}$ and $\mathcal{Y}$ respectively. What are the natural bootstrap and permutation central limit theorems for the independence empirical process $\sqrt{n}(\mathbb{H}_n - \mathbb{P}_n \cdot \mathbb{Q}_n)$?

ADDITIONAL REFERENCES

Alexander, K. S. (1984). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probability* **12**, 1041 - 1067. Correction. *Ann. Probability* **15**, 428 - 430.

Alexander, K. S. (1985). The nonexistence of a universal multiplier moment for the central limit theorem, in Probability in Banach Spaces V, *Lecture Notes in Math.* **1153**, 15 - 16.

Alexander, K. S. (1987). Rates of growth and sample moduli for weighted empirical processes indexed by sets. *Probab. Th. Rel. Fields* **75**, 379 - 423.

Andersen, N.T. (1985). The central limit theorem for non-separable valued functions. *Zeit. Wahrsch. Verv. Geb.* **70**, 445-455.

Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley, New York.

Chibisov, D. M. (1965). An investigation of the asymptotic power of the tests of fit. *Theor. Probability Appl.* **10**, 421-437.

Dudley, R. M. (1966). Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Ill. J. Math.* **10**, 109-126.

Dudley, R. M. (1990). Nonlinear functionals of empirical measures and the bootstrap. *Probability in Banach Spaces, Progress in Probability 21* **7**, 63-82. Birkhäuser, Boston.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7**, 1-26.

Efron, B. (1982). The Jackknife, the Bootstrap and Other Resampling Plans. *CMBS-NSF Regional Conference Series in Applied Mathematics* **38**. Society for Industrial and Applied Mathematics, Philadelphia.

Giné, E. and Zinn, J. (1992). On Hoffmann-Jørgensen's inequality for $U-$processes. In *Probability in Banach Spaces, 8*, Dudley, Hahn, Kuelbs eds., pp. 80–91, Progress in Probability, 30, Birkhäuser, Boston.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13–30.

Klaassen, C. A. J. and Wellner, J. A. (1992). Kac empirical processes and the bootstrap. In *Probability in Banach Spaces, 8*, Birkhäuser, Boston; edited by R.M. Dudley, M. G. Hahn, and J. Kuelbs.

Kolcinskii, V. I. (1981). On the central limit theorem for empirical measures *Theor. Probability Math. Statist.* **24**, 71 - 82.

Marcus, M. B. (1981). Weak convergence of the empirical characteristic function. *Ann. Probability* **9**, 194 - 201.

Massart, P. (1986). Rates of convergence in the central limit theorem for empirical measures. *Ann. Inst. Henri Poincare* **22**, 381 - 423.

Pisier, G. (1975). Le théorème de la limite centrale et la loi du logarithme itere dans les espaces de Banach. *Seminaire Maurey-Schwartz 1975-1976.* Expose IV, Ecole Polytechnique, Paris.

Pollard, D. (1984). *Convergence in Distribution of Stochastic Processes.* Springer Verlag, New York.

Pollard, D. (1990). Empirical processes: Theory and Applications. *NSF-CMBS Regional Conference Series in Probability and Statistics.* **2**. Institute of Mathematical Statistics, Hayward, California.

Praestgaard, J. (1992) Permutation and bootstrap Kolmogorov-Smirnov tests for the equality of two distributions. Technical Report, Department of Statistics, University of Iowa.

Pyke, R. (1992). Probability in Mathematics and Statistics: A century's predictor of future directions. *Jber.d.D.Math.-Verein, Jubiläumstagung 1990,* 239 - 264.

Romano, J. P. and Politis, D.N. (1992). A general theory for large sample confidence regions based on subsamples under minimal conditions. *Technical Report* **399**, Department of Statistics, Stanford University.

Sheehy, A. and Wellner, J. A. (1988). Uniformity in P of some limit theorems for empirical measures and processes *Technical Report 134*, Department of Statistics, University of Washington.

Van der Vaart, A. W. and Wellner, J. A. (1989). Prohorov and continuous mapping theorems in the Hoffmann-Jørgensen weak convergence theory with application to convolution and asymptotic minimax theorems. *Technical Report No. 157.* Department of Statistics, University of Washington.

Van der Vaart, A. W. and Wellner, J. A. (1993). *Weak Convergence and Empirical Processes* To appear in the IMS Lecture Notes-Monograph Series.

Yurinskii, V.V. (1974). Exponential bounds for large deviations. *Theor. Probability Appl.* **19**, 154-155.

Department of Statistics, GN-22
B313 Padelford Hall
University of Washington
Seattle, Washington 98195