KAC EMPIRICAL PROCESSES AND THE BOOTSTRAP

Chris A. J. Klaassen and Jon A. Wellner*

0. Introduction

The general empirical process indexed by a subset $\mathcal{F}$ of $L_2(P)$ converges weakly to a P-Brownian bridge process if and only if the corresponding Poissonized, or Kac, empirical process converges to a P-Brownian motion. In this expository paper we explore the role of Poissonization in bootstrap sampling, and give a partial explanation of the relationship between the bootstrap limit theorems of Giné and Zinn (1990) and the multiplier central limit theorems of Ledoux and Talagrand (1988). Giné and Zinn (1990) used two important tools in their proofs of consistency of Efron's bootstrap of the general empirical process: a conditional almost sure bound of Ledoux, Talagrand, and Zinn (lemma 5 of Ledoux and Talagrand (1988), lemma 2.3 of Giné and Zinn (1990)), and a Poissonization inequality of Le Cam (1970). They also relied heavily on symmetrization by Rademacher random variables. The proof we present here of the Giné and Zinn theorem still relies upon the two main tools used by Giné and Zinn (1990), but in a somewhat different way. The major differences are: (i) We avoid symmetrization with Rademacher random variables, and work instead with centered random elements throughout. (ii) We appeal directly to the almost sure multiplier central limit theorem of Ledoux and Talagrand (1988), (1991) rather than to the key tool used in its proof, the conditional almost sure bound. (iii) Poissonization is used differently. In analogy to the Poissonized empirical process, we compare the Poissonized bootstrap empirical process directly with the bootstrap empirical process. The key Poissonization techniques date back to Le Cam (1970) and Araujo and Giné (1980). We conclude with a random sample

* **Research supported in part by:** NSF Grants DMS-8723011 and DMS-8505550

size central limit theorem for general empirical processes.

1. Empirical Processes and Poisson Processes

Let $(A, \mathbf{A}, P)$ be a probability space, and let $X, X_1, \ldots, X_n, \ldots$ be iid P. Let $Y_1, \ldots, Y_n$ be iid Poisson(1), so that $N_n \equiv Y_1 + \ldots + Y_n \sim Poisson(n)$. Let $N_\lambda \sim Poisson(\lambda)$. Suppose that $\mathcal{F}$ is a collection of real-valued measurable functions on A.

We want to consider the processes

$$n\mathbb{P}_n \equiv \sum_{i=1}^{n} \delta_{X_i} \equiv \text{ the empirical point process on } A,$$

$$\mathbb{X}_n \equiv \sqrt{n}(\mathbb{P}_n - P) \equiv \text{ the empirical process},$$

$$\mathbb{Z}_n \equiv \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{N_n} \delta_{X_i} - P\right) \equiv \text{ the Poissonized, or Kac, empirical process},$$

and relations between them as elements of $l^\infty(\mathcal{F})$, the space of all bounded real valued functions on $\mathcal{F}$. Here

$$\mathbb{P}_n(f) \equiv \frac{1}{n}\sum_{i=1}^{n} f(X_i) \quad \text{and} \quad P(f) \equiv \int f\,dP, \quad f \in \mathcal{F}.$$

Let F be defined by $F(x) = (\sup_{f\in\mathcal{F}} |f(x)|)^*$, $x \in A$; F is the (smallest measurable) envelope function of $\mathcal{F}$. Here, if $h : A \to R$ is an arbitrary function, then h^* denotes the least measurable function dominating h; see e.g. Dudley (1984) or (1985). We denote the norm on $l^\infty(\mathcal{F})$ by $\|\cdot\|_\mathcal{F}$, i.e. for $b \in l^\infty(\mathcal{F})$, $\|b\|_\mathcal{F} = \sup_{f\in\mathcal{F}} |b(f)|$. Since it is a norm, $\|\cdot\|_\mathcal{F}$ and also $\|\cdot\|^*_\mathcal{F}$ is convex on $l^\infty(\mathcal{F})$. Note that $P \in l^\infty(\mathcal{F})$ if and only if $\|P\|_\mathcal{F} < \infty$, $\mathbb{P}_n \in l^\infty(\mathcal{F})$ for $n = 1, 2, \ldots$ if and only if $F(X_1) < \infty$ a.s., and $\mathbb{X}_n \in l^\infty(\mathcal{F})$ for $n = 1, 2, \ldots$ if and only if $\tilde{F}(X_1) = (\sup_{f\in\mathcal{F}} |f(X_1) - Pf|)^* < \infty$ a.s. Also note that $\mathbb{Z}_n \in l^\infty(\mathcal{F})$ a.s. if $F(X_1) < \infty$ a.s. and $\|P\|_\mathcal{F} < \infty$.

Two centered Gaussian processes in $l^\infty(\mathcal{F})$ which arise in the following are:

$$\mathbb{Z} \equiv \text{ a } P\text{-Brownian motion process } (Z_P)$$

with

$$Cov(\mathbb{Z}(f), \mathbb{Z}(g)) = P(fg), \quad f, g \in \mathcal{F};$$

and

$$\mathbb{X} \equiv \text{ a P - Brownian bridge process } (G_P)$$

with

$$Cov(\mathbb{X}(f), \mathbb{X}(g)) = P(fg) - P(f)P(g), \quad f, g \in \mathcal{F}.$$

Let ρ_P and e_P denote the natural Gaussian pseudometrics on $\mathcal{F}$ corresponding to $\mathbb{X}$ and $\mathbb{Z}$ respectively: for $f, g \in \mathcal{F}$

$$\rho_P^2(f,g) \equiv Var_P(f(X) - g(X)), \quad e_P^2(f,g) \equiv E_P(f(X) - g(X))^2.$$

Note that since $\mathbb{Z}$ is P-Brownian motion, then

$$\mathbb{Z} - \mathbb{Z}(1)P \qquad \text{is a} \quad \text{P - Brownian bridge process } G_P,$$

and, on the other hand, if the P-Brownian bridge process $\mathbb{X}$ and a standard normal random variable Z are independent, then

$$\mathbb{X} + ZP \quad \text{is a} \quad P\text{-Brownian motion process } Z_P.$$

Simple computations show that, for $f, g \in L_2(P)$,

$$E\,\mathbb{X}_n(f) = 0, \quad Cov(\mathbb{X}_n(f), \mathbb{X}_n(g)) = Cov(\mathbb{X}(f), \mathbb{X}(g)),$$

$$E\,\mathbb{Z}_n(f) = 0, \quad Cov(\mathbb{Z}_n(f), \mathbb{Z}_n(g)) = Cov(\mathbb{Z}(f), \mathbb{Z}(g)).$$

Two basic identities, which follow by simple algebra, are:

$$\mathbb{Z}_n = \sqrt{\frac{N_n}{n}}\,\mathbb{X}_{N_n} + \sqrt{n}\left(\frac{N_n}{n} - 1\right)P \tag{1}$$

(this is a general version of equation (8.4.8), page 339, Shorack and Wellner (1986)): or, since $\mathbb{Z}_n(1) = \sqrt{n}(n^{-1}N_n - 1)$, on $[N_n > 0]$,

$$\mathbb{X}_{N_n} = \frac{1}{\sqrt{N_n/n}}(\mathbb{Z}_n - \mathbb{Z}_n(1)P). \tag{2}$$

Yet another identity which follows by simple algebra is

$$\mathbb{X}_n - (\mathbb{Z}_n - \mathbb{Z}_n(1)P) = \left\{ \frac{1}{\sqrt{n}} \sum_{i=N_n+1}^{n} (\delta_{X_i} - P) \right\} 1_{[n>N_n]} - \left\{ \frac{1}{\sqrt{n}} \sum_{i=n+1}^{N_n} (\delta_{X_i} - P) \right\} 1_{[n<N_n]}. \quad (3)$$

Because $N_n - n = O_p(\sqrt{n})$, each of the two terms on the right side of (3) involves just $O(\sqrt{n})$ terms, so we expect them to converge to zero in probability, and hence convergence of $\mathbb{X}_n$ will be equivalent to convergence of $\mathbb{Z}_n$.

We say that $\mathcal{F}$ is *Z_P-preGaussian* if $\mathbb{Z}$ can be chosen so that $\mathbb{Z} \in C(\mathcal{F}, e_P)$ a.s. where $C(\mathcal{F}, e_P)$ is the collection of e_P-uniformly continuous functions in $l^\infty(\mathcal{F})$. (By convention all the functions $x \in l^\infty(\mathcal{F})$ are bounded: $\|x\|_{\mathcal{F}} \equiv \sup_{f\in\mathcal{F}} |x(f)| < \infty$.) For short this will be written as $\mathcal{F} \in PG(Z_P)$. Similarly, $\mathcal{F}$ is *G_P-preGaussian* if $\mathbb{X}$ can be chosen so that $\mathbb{X} \in C(\mathcal{F}, \rho_P)$ a.s. where $C(\mathcal{F}, \rho_P)$ is the collection of ρ_P-uniformly continuous functions in $l^\infty(\mathcal{F})$, and in this case we write $\mathcal{F} \in PG(G_P)$.

Definition 1. Suppose that $\mathcal{F} \subset L_2(P)$. We say that $\mathcal{F}$ is *P-Kac* if $\mathcal{F} \in PG(Z_P)$ and $\mathbb{Z}_n \Rightarrow \mathbb{Z} \sim Z_P$ in $l^\infty(\mathcal{F})$.

The weak convergence $\Rightarrow$ in this definition is in the sense of Hoffmann-Jørgensen (1984); see e.g. Dudley (1985) or Van der Vaart and Wellner (1990). As usual, $\mathcal{F} \subset L_2(P)$ is called *P-Donsker* if $\mathcal{F} \in PG(G_P)$ and $\mathbb{X}_n \Rightarrow \mathbb{X}$ in $l^\infty(\mathcal{F})$ (i.e. the CLT holds for $\mathbb{X}_n$ in $l^\infty(\mathcal{F})$), and we denote this by $\mathcal{F} \in CLT(P)$.

The following theorem extends the classical result for empirical processes of Kac (1949).

Theorem 1. Suppose that $\mathcal{F} \subset L_2(P)$ and $\|P\|_{\mathcal{F}} < \infty$. Then $\mathcal{F}$ is P-Donsker if and only if $\mathcal{F}$ is P-Kac.

We will repeatedly use the inequality

$$\|E \cdot \|_{\mathcal{F}}^* \le E\| \cdot \|_{\mathcal{F}}^* \quad (4)$$

which follows from the convexity of $\|\cdot\|_{\mathcal{F}}^*$. In particular, if X and Y are independent random elements of $l^\infty(\mathcal{F})$ with $EY = 0$, then

$$E\|X\|_{\mathcal{F}}^* = E\|X + EY\|_{\mathcal{F}}^* \le E\|X + Y\|_{\mathcal{F}}^* . \quad (5)$$

Proof of theorem 1. Let $\epsilon > 0$ (small) and $M \geq 1$ (large). Then, by (3) and (5)

$$Pr^*(\|\mathbb{X}_n - (\mathbb{Z}_n - \mathbb{Z}_n(1)P)\|_{\mathcal{F}} > \epsilon)$$

$$\leq Pr(|N_n - n| > M\sqrt{n}\,) + \sum_{k=-[M\sqrt{n}]}^{[M\sqrt{n}]} P(N_n = n+k)\, Pr^*\left(\|n^{-1/2}\sum_{i=1}^{|k|}(\delta_{X_i} - P)\|_{\mathcal{F}} > \epsilon\right)$$

$$\leq \frac{1}{M^2} + \sum_{k=-[M\sqrt{n}]}^{[M\sqrt{n}]} P(N_n = n+k)\frac{1}{\epsilon}E\|n^{-1/2}\sum_{i=1}^{[M\sqrt{n}]}(\delta_{X_i} - P)\|_{\mathcal{F}}^* \qquad (a)$$

$$\leq \frac{1}{M^2} + \frac{M}{\epsilon}E\|\frac{1}{[M\sqrt{n}]}\sum_{i=1}^{[M\sqrt{n}]}(\delta_{X_i} - P)\|_{\mathcal{F}}^*.$$

Therefore, *if $\mathcal{F}$ is P-Glivenko - Cantelli*, we have

$$\|\mathbb{X}_n - (\mathbb{Z}_n - \mathbb{Z}_n(1)P)\|_{\mathcal{F}} \to_{Pr^*} 0 \qquad (b)$$

Taking $\mathcal{F} \in L_2(P)$ to be a finite class, we already deduce from this that the finite dimensional distributions of $\mathbb{X}_n + \mathbb{Z}_n(1)P$ and of $\mathbb{Z}_n$ have the same limits. In particular

$$\mathbb{Z}_n \to_{f.d.} \mathbb{Z} \sim Z_P.$$

If $\mathcal{F}$ is P- Donsker, then it is *a fortiori* P-Glivenko - Cantelli, so that (b) holds, which yields that $\mathcal{F}$ is P-Kac (the process $\mathbb{X}_n + \mathbb{Z}_n(1)P$ converges in law since $\mathbb{Z}_n(1)P$ converges by the CLT in R and is independent of $\mathbb{X}_n$). If $\mathcal{F}$ is P-Kac, then it is P-Glivenko-Cantelli by Le Cam's lemma (Araujo and Giné (1980), theorem 3.4.8, where truncated expectations can be replaced by expectations if they exist); we give a statement and proof below. Therefore $\mathcal{F}$ is P-Donsker again by (b). The uniform integrability needed to guarantee convergence of the expectations in these arguments follows from e.g. Andersen (1985), proposition 3.7, page 449. □

Lemma 1. (Le Cam's Poissonization inequality)

$$(1-e^{-1})E^*\|\sum_{i=1}^{n}(\delta_{X_i} - P)\|_{\mathcal{F}} \leq E^*\|\sum_{i=1}^{N_n}(\delta_{X_i} - P)\|_{\mathcal{F}}.$$

Equivalently

$$(1-e^{-1})E^*\|\mathbb{X}_n\|_{\mathcal{F}} \le E^*\|\mathbb{Z}_n - \mathbb{Z}_n(1)P\|_{\mathcal{F}}.$$

Proof. Write $Z_i = \delta_{X_i} - P$. For $i = 1, \ldots, n$ let Z_{ij}, $j = 1, 2, \ldots$ be iid copies of Z_i, let $Y_1, \ldots, Y_n$ be iid Poisson(1) random variables, and set $N_n = Y_1 + \ldots + Y_n$. Then, by (4) and (5)

$$\begin{aligned}(1-e^{-1})E^*\|\sum_{i=1}^n Z_i\|_{\mathcal{F}} &= E\|E(\sum_{i=1}^n (Y_i \wedge 1)Z_i | Z_1, \ldots, Z_n)\|_{\mathcal{F}}^* \\ &\le E\|\sum_{i=1}^n (Y_i \wedge 1)Z_i\|_{\mathcal{F}}^* \\ &\le E(E(\|\sum_{i=1}^n \sum_{j=1}^{Y_i} Z_{ij}\|_{\mathcal{F}}^* | Y_1, \ldots, Y_n)) \\ &= E^*\|\sum_{i=1}^{N_n} Z_i\|_{\mathcal{F}}.\end{aligned}$$

This proof is almost exactly the same as the proof of lemma 2.1 of Giné and Zinn (1990). Note that the result continues to hold if the X_i's are independent but not identically distributed. □

We conclude this section with a by - product of the proof of theorem 1.

Proposition 1. If $\mathcal{F} \subset L_2(P)$ and $\mathcal{F}$ is P-Donsker, then

$$n^{1/4}\|\mathbb{X}_{N_n} - \mathbb{X}_n\|_{\mathcal{F}}^* = O_p(1).$$

Proof. Replacing ϵ in (a) of the proof of theorem 1 by $Kn^{-1/4}$ (K large), we see that the left hand side of (3) is $O_p(n^{-1/4})$. Together with (2) this yields the claim. Note that the hypothesis $\|P\|_{\mathcal{F}} < \infty$ is not used here. □

Remarks: The basic method of Poissonization was used by Kolmogorov (1933) and Donsker (1952) in their classic papers; see e.g. Shorack and Wellner (1986), chapter 8. Kac (or Poissonized) empirical processes have also played an important role in obtaining rates of convergence for limit theorems and invariance principles; see e.g. Massart (1989), in particular his lemma 2, and Dudley (1984), section 8.3, where Poissonization is used to study rates of convergence for classes of sets too large to satisfy the CLT. Durst and Dudley (1981) have a partial result in the direction of theorem 1 for classes of sets. In the classical one dimensional case $A = R$ there is a substantial literature concerned with finite sample and asymptotic distributions of statistics connected with the Kac empirical process; see e.g. M. Csörgő and Alvo (1970), M. Csörgő (1972), S. Csörgő (1981), Suzuki (1972), and the references therein.

Part of theorem 1 is contained in theorem 3.4.9 of Araujo and Giné (1980); see also their exercise 2, page 122, and note that their X_{nj} are assumed to be symmetric. Evarist Giné has shown us a proof of theorem 1 via symmetrization and desymmetrization.

Pyke (1968) and S. Csörgő (1974) have related results which go in a somewhat different direction: they study (special cases of) $\mathbb{X}_{\nu_n}$ where ν_n is random and possibly *dependent* on the X_i's; see section 3.

2. The bootstrap empirical process and Poissonization

Now let $\mathbb{P}_n^\omega$ be the empirical measure of the X_i's as above, let

$$X_1^\#, \ldots, X_n^\# \tag{1}$$

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

$$\begin{aligned}\mathbb{X}_n^\# &= \sqrt{n}(\mathbb{P}_n^\# - \mathbb{P}_n^\omega) \\ &= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n \delta_{X_i^\#} - \mathbb{P}_n^\omega\right) \\ &= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n M_{ni}^\# \delta_{X_i(\omega)} - \mathbb{P}_n^\omega\right) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n (M_{ni}^\# - 1)\,\delta_{X_i(\omega)}\end{aligned} \tag{2}$$

where

$$\underline{M}_n^\# \sim Mult_n(n, (\frac{1}{n}, \ldots, \frac{1}{n})) \text{ is independent of the } X_i\text{'s.} \tag{3}$$

We can write

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

where

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

are iid, $j = 1,\ldots,k$, $k = 1,2,\ldots$.

Note that if we "Poissonize" $\underline{M}_n^\#$ by forming $\underline{M}_{N_n^\#}^\#$, the result is:

$$\underline{M}_{N_n^\#}^\# \sim (Y_1,\ldots,Y_n) \tag{4}$$

where $Y_1,\ldots,Y_n$ are iid Poisson(1). Thus we consider the "Poissonized bootstrap empirical process" $\mathbb{Z}_n^\#$ defined by

$$\begin{aligned}\mathbb{Z}_n^\# &= \sqrt{n}\left\{\frac{1}{n}\sum_{i=1}^n M_{N_n^\#,i}^\# \delta_{X_i(\omega)} - \mathbb{P}_n^\omega\right\} \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n (M_{N_n^\#,i}^\# - 1)\delta_{X_i(\omega)} \\ &\sim \frac{1}{\sqrt{n}}\sum_{i=1}^n (Y_i - 1)\delta_{X_i(\omega)}.\end{aligned} \tag{5}$$

For our main results we need a measurability assumption which insures that $\|\mathbb{X}_n\|_{\mathcal{F}'(\delta,\rho_P)}$ is completion measurable under P and Fubini's theorem can be applied to $\|\sum_{i=1}^n Y_i\delta_{X_i}\|_{\mathcal{F}'(\delta,\rho_P)}$ where the Y_i are iid real-valued mean zero rv's independent of the X_i's. In the terminology of Giné and Zinn (1984, 1986, 1990), we require $\mathcal{F}$ to be nearly linearly deviation measurable for P, or $\mathcal{F} \in NLDM(P)$ for short, and that both $\mathcal{F}^2$ and $\mathcal{F}'^2$ be nearly linearly supremum measurable, or $NLSM(P)$. When all of these hold, we say $\mathcal{F} \in M(P)$. It is known that $\mathcal{F} \in M(P)$ if $\mathcal{F}$ is countable, or if the empirical processes $\mathbb{X}_n$ are stochastically separable, or if $\mathcal{F}$ is image admissible Suslin (see Giné and Zinn (1990), pages 853, 854).

Theorem 2. Suppose that $\mathcal{F} \in M(P)$ and $\|P\|_{\mathcal{F}} < \infty$. Then $\mathcal{F} \in CLT(P)$ and $P(F^2) < \infty$ if and only if $\mathbb{Z}_n^\# \Rightarrow \mathbb{Z}^\# \sim Z_P$ a.s. P^∞ in $l^\infty(\mathcal{F})$.

Proof. The essence of this is the Ledoux and Talagrand (1988) theorem 4, page 34 (due to Ledoux, Talagrand, and Zinn) extended slightly to allow for asymmetric multipliers $Y_i - 1$. Write

$$\mathbb{Z}_n^\# = \frac{1}{\sqrt{n}}\sum_{i=1}^n (Y_i - 1)\delta_{X_i(\omega)} \tag{a}$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_i - 1)(\delta_{X_i(\omega)} - P) + \sqrt{n}(\bar{Y}_n - 1)P$$
$$\equiv \mathbb{Y}_n^{\#} + \sqrt{n}(\bar{Y}_n - 1)P = \mathbb{Y}_n^{\#} + \mathbb{Z}_n^{\#}(1)P,$$

so that

$$\mathbb{Y}_n^{\#} = \mathbb{Z}_n^{\#} - \mathbb{Z}_n^{\#}(1)P. \tag{b}$$

Since the Y_i's are iid Poisson(1), they satisfy $\int_0^\infty \sqrt{P(|Y_i| > t)}\, dt < \infty$. By theorem 3.1 of Praestgaard (1990) or Ledoux and Talagrand (1991) theorem 10.14, page 293 (this has a history beginning with Giné and Zinn (1984), lemma 2.9, and continuing with Ledoux and Talagrand (1986), Giné and Zinn (1986) pages 64 and 65, and Ledoux and Talagrand (1988), theorem 4, page 34 and the discussion on page 35) we know that $\mathcal{F} \in CLT(P)$ and $P(F^2) < \infty$ if and only if

$$\mathbb{Y}_n^{\#} \Rightarrow \mathbb{Y}^{\#} \sim G_P \quad \text{a.s.} \quad P^\infty \quad \text{in} \quad l^\infty(\mathcal{F}). \tag{c}$$

But it is trivially true that

$$\sqrt{n}(\bar{Y}_n - 1)P \Rightarrow ZP \quad \text{in} \quad l^\infty(\mathcal{F}), \tag{d}$$

where $Z \sim N(0,1)$ is independent of $\mathbb{Y}^{\#}$. The conclusion follows from (a) through (d). □

To connect $\mathbb{X}_n^{\#}$ and $\mathbb{Z}_n^{\#}$, note that

$$\mathbb{X}_n^{\#} - (\mathbb{Z}_n^{\#} - \mathbb{Z}_n^{\#}(1)P) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (M_{ni}^{\#} - M_{N_n^{\#},i}^{\#})(\delta_{X_i(\omega)} - P) \equiv R_n(\omega), \tag{6}$$

an identity paralleling (1.3). Here the coupling of $\mathbb{X}_n^{\#}$ and $\mathbb{Z}_n^{\#}$ is via the first $n \wedge N_n^{\#}$ $\underline{1}_j$'s they have in common. To obtain suggestive identities paralleling (1.1) and (1.2), we define

$$\mathbb{X}_{n,k}^{\#} = \sqrt{k}(\mathbb{P}_k^{\#} - \mathbb{P}_n)$$
$$= \sqrt{k}\left\{\frac{1}{k}\sum_{j=1}^{k}\left(\sum_{i=1}^{n} 1_{ij}\delta_{X_i(\omega)}\right) - \mathbb{P}_n\right\}$$
$$= \frac{1}{\sqrt{k}} \sum_{i=1}^{n} (M_{ki}^{\#} - \frac{k}{n})\delta_{X_i(\omega)}$$

the bootstrap empirical process for a bootstrap sample of size k from $\mathbb{P}_n^\omega$. Thus $\mathbb{X}_{n,n}^\# = \mathbb{X}_n^\#$, and the identities paralleling (1.1) and (1.2) are:

$$\mathbb{Z}_n^\# = \sqrt{\frac{N_n^\#}{n}}\mathbb{X}_{n,N_n^\#}^\# + \sqrt{n}(n^{-1}N_n^\# - 1)\mathbb{P}_n \tag{7}$$

and, since $\mathbb{Z}_n^\#(1) = \sqrt{n}(n^{-1}N_n^\# - 1)$,

$$\mathbb{X}_{n,N_n^\#}^\# = \frac{1}{\sqrt{N_n^\#/n}}(\mathbb{Z}_n^\# - \mathbb{Z}_n^\#(1)\mathbb{P}_n). \tag{8}$$

Although we will not use these identities directly in our proof, they provide helpful insight and intuition.

Theorem 3. Suppose that $\mathcal{F} \in M(P)$ and $\|P\|_\mathcal{F} < \infty$. Then the following statements are equivalent:

A. $\mathcal{F} \in CLT(P)$ and $P(F^2) < \infty$.

B. $\mathbb{Z}_n^\# \Rightarrow \mathbb{Z}^\# \sim Z_P$ a.s. P^∞ in $l^\infty(\mathcal{F})$.

C. $\mathbb{X}_n^\# \Rightarrow \mathbb{X}^\# \sim G_P$ a.s. P^∞ in $l^\infty(\mathcal{F})$.

Proof. This *follows* from the bootstrap CLT of Giné and Zinn (1990): they prove the equivalence of A and C (even without the hypothesis $\|P\|_\mathcal{F} < \infty$), while the equivalence of A and B is just implied by theorem 2 (essentially Ledoux and Talagrand (1988)). The proofs of Giné and Zinn use symmetrization by Rademacher random variables. Thus, if we write sA, sB, and sC for the corresponding symmetrized parts of theorem 3, and B_u for B unconditionally (integrating over the X_i's in (5), $\mathbb{Z}_n^\# \Rightarrow \mathbb{Z}^\# \sim Z_P$ in $l^\infty(\mathcal{F})$), the Giné and Zinn proof is organized as follows:

$$\text{sA} \Rightarrow \text{sB}_\text{u} \Rightarrow \text{sC} \Rightarrow \text{sA}.$$

Here the first implication follows from lemma 2.9 of Giné and Zinn (1984); the second implication follows from the Ledoux, Talagrand (and Zinn) (1988) almost sure multiplier CLT and proposition 2.2 of Giné and Zinn (1990); and the last implication comes from proposition 2.2 again. They then show that the "s's" can be removed from sA $\Rightarrow$ sC. (Giné and Zinn (1990) do not have sB or sB_u in the statement of their main result, but only in their lemma 2.1 and proposition 2.2.)

An interesting question is: can we use the Poissonization argument to give an alternative proof of the Giné - Zinn bootstrap theorem, with most of the "hard work" contained in the Ledoux and Talagrand (and Zinn) (1988) theorem? In other words, is there a proof of theorem 3 organized in the following way:

$$\mathrm{A} \underset{\text{LT \& thm 2}}{\Leftrightarrow} \mathrm{B} \Leftrightarrow \mathrm{C} \quad ?$$

We now complete this program by showing that B and C are equivalent. In view of (6), it suffices to show that both B and C imply that the right hand side of (6) converges to zero a.s. P^∞ and in the sense of $\Rightarrow$.

Suppose we can show both B and C imply $\mathcal{F} \in GC(P)$; i.e. the Glivenko - Cantelli theorem holds in $l^\infty(\mathcal{F})$:

$$\|\mathbb{P}_n - P\|_{\mathcal{F}}^* \to_{a.s.} 0 \quad \text{as} \quad n \to \infty. \tag{a}$$

Note that conditional on $N_n^\# = k$,

$$\{|M_{N_n^\#,i}^\# - M_{n,i}^\#|\}_{1\le i\le n} \stackrel{d}{=} \{M_{|k-n|,i}^\#\}_{1\le i\le n}.$$

Furthermore, $|N_n^\# - n| = O_p(\sqrt{n})$. Thus we need to consider $M_{m_n}^\#$ with $m_n = O(\sqrt{n})$.

Let the random variable B_n have a binomial distribution with parameters m_n and $1/n$. Then

$$\begin{aligned} &Pr(B_n \le 2) \\ &= (1-\frac{1}{n})^{m_n-2}\Big\{(1-\frac{1}{n})^2 + \frac{m_n}{n}(1-\frac{1}{n}) + \frac{m_n(m_n-1)}{2n^2}\Big\} \\ &= \Big\{1 - \frac{m_n-2}{n} + \frac{m_n^2}{2n^2} + O(n^{-3/2})\Big\}\Big\{1 + \frac{m_n-2}{n} + \frac{m_n^2}{2n^2} + O(n^{-3/2})\Big\} \\ &= 1 - O(n^{-3/2}), \end{aligned}$$

and consequently

$$Pr(\max_{1\le i\le n} M_{m_n,i}^\# > 2) \le n\,Pr(B_n > 2) = O(n^{-1/2}). \tag{b}$$

Thus with probability $> 1-\epsilon$ for $n \geq N_\epsilon$, the maximum of the absolute values of the differences of the components of the multinomials in (6) is 2 or less. This suggests rewriting $R_n(\omega)$ of (6) as follows: let $I_j \equiv \{i : M^{\#}_{|N^{\#}_n - n|, i} \geq j\}$, and $m_n^{(j)} \equiv \#(I_j)$, the cardinality of I_j. Then, for $n \geq N_\epsilon$, with probability exceeding $1-\epsilon$ by (b), we may write

$$R_n(\omega) = sign(n - N^{\#}_n)\frac{1}{\sqrt{n}}\sum_{i=1}^{n} M^{\#}_{|N_n - n|, i}(\delta_{X_i(\omega)} - P)$$

$$= sign(n - N^{\#}_n)\sum_{j=1}^{2}\frac{m_n^{(j)}}{\sqrt{n}}\left\{\frac{1}{m_n^{(j)}}\sum_{i\in I_j}(\delta_{X_i(\omega)} - P)\right\} \qquad (c)$$

where $|N^{\#}_n - n| \geq m_n^{(1)} \geq m_n^{(2)}$. Since $N^{\#}_n - n = O_p(\sqrt{n})$, it follows that $m_n^{(j)}/\sqrt{n} = O_p(1)$ on the right side in (c) while the terms inside the brackets converge to zero for P^∞ a.e. ω by (a) since (conditional on $m_n^{(j)}$) it has the same convergence properties as $\mathbb{P}_{m_n^{(j)}} - P$. We conclude that $R_n(\omega)$ converges in probability to zero for P^∞-a.e. ω if (a) holds. The following lemma (similar to lemma 3.2 of Pollard (1981)) makes this more precise.

Lemma. Let $U_n = (U_{n1}, \ldots, U_{nn})$ be a random vector in $\{0,1\}^n$, defined on $([0,1], \mathbf{B}, \lambda)$, and independent of $X_1, X_2, \ldots$ with an exchangeable (permutation invariant) distribution, and $\sum_{i=1}^n U_{ni} > 0$ a.s. If (a) holds and $\sum_{i=1}^n U_{ni} \to_p \infty$, then

$$\|\sum_{i=1}^{n} U_{ni}(\delta_{X_i(\omega)} - P)/\sum_{i=1}^{n} U_{ni}\|^*_{\mathcal{F}} \to_p 0$$

for a.e. ω as $n \to \infty$.

Proof. Write $S_n \equiv \sum_{i=1}^n U_{ni}(\delta_{X_i} - P)/\sum_{i=1}^n U_{ni}$, and let $\mathbf{A}_n$ be the σ-field generated by $S_n(f), S_{n+1}(f), \ldots, f \in \mathcal{F}$. Fix $f_0 \in \mathcal{F}$. Since by symmetry

$$E(S_n(f_0)|\mathbf{A}_{n+1}) = E(E(S_n(f_0)|U_n, \mathbf{A}_{n+1})|\mathbf{A}_{n+1})$$

$$= E(\sum_{i=1}^{n} U_{ni}E(f_0(X_i) - P(f_0)|U_n, \mathbf{A}_{n+1})/\sum_{i=1}^{n} U_{ni}|\mathbf{A}_{n+1})$$

$$= E(f_0(X_1) - P(f_0)|\mathbf{A}_{n+1})$$

Thus with probability $> 1-\epsilon$ for $n \geq N_\epsilon$, the maximum of the absolute values of the differences of the components of the multinomials in (6) is 2 or less. This suggests rewriting $R_n(\omega)$ of (6) as follows: let $I_j \equiv \{i : M^{\#}_{|N_n^{\#}-n|,i} \geq j\}$, and $m_n^{(j)} \equiv \#(I_j)$, the cardinality of I_j. Then, for $n \geq N_\epsilon$, with probability exceeding $1-\epsilon$ by (b), we may write

$$R_n(\omega) = sign(n - N_n^{\#})\frac{1}{\sqrt{n}}\sum_{i=1}^{n} M^{\#}_{|N_n-n|,i}(\delta_{X_i(\omega)} - P)$$

$$= sign(n - N_n^{\#})\sum_{j=1}^{2}\frac{m_n^{(j)}}{\sqrt{n}}\left\{\frac{1}{m_n^{(j)}}\sum_{i\in I_j}(\delta_{X_i(\omega)} - P)\right\} \qquad (c)$$

where $|N_n^{\#} - n| \geq m_n^{(1)} \geq m_n^{(2)}$. Since $N_n^{\#} - n = O_p(\sqrt{n})$, it follows that $m_n^{(j)}/\sqrt{n} = O_p(1)$ on the right side in (c) while the terms inside the brackets converge to zero for P^∞ a.e. ω by (a) since (conditional on $m_n^{(j)}$) it has the same convergence properties as $\mathbb{P}_{m_n^{(j)}} - P$. We conclude that $R_n(\omega)$ converges in probability to zero for P^∞-a.e. ω if (a) holds. The following lemma (similar to lemma 3.2 of Pollard (1981)) makes this more precise.

Lemma. Let $U_n = (U_{n1}, \ldots, U_{nn})$ be a random vector in $\{0,1\}^n$, defined on $([0,1], \mathbf{B}, \lambda)$, and independent of $X_1, X_2, \ldots$ with an exchangeable (permutation invariant) distribution, and $\sum_{i=1}^n U_{ni} > 0$ a.s. If (a) holds and $\sum_{i=1}^n U_{ni} \to_p \infty$, then

$$\|\sum_{i=1}^{n} U_{ni}(\delta_{X_i(\omega)} - P)/\sum_{i=1}^{n} U_{ni}\|^*_{\mathcal{F}} \to_p 0$$

for a.e. ω as $n \to \infty$.

Proof. Write $S_n \equiv \sum_{i=1}^n U_{ni}(\delta_{X_i} - P)/\sum_{i=1}^n U_{ni}$, and let $\mathbf{A}_n$ be the σ-field generated by $S_n(f), S_{n+1}(f), \ldots, f \in \mathcal{F}$. Fix $f_0 \in \mathcal{F}$. Since by symmetry

$$\begin{aligned} E(S_n(f_0)|\mathbf{A}_{n+1}) &= E(E(S_n(f_0)|U_n, \mathbf{A}_{n+1})|\mathbf{A}_{n+1}) \\ &= E(\sum_{i=1}^{n} U_{ni}E(f_0(X_i) - P(f_0)|U_n, \mathbf{A}_{n+1})/\sum_{i=1}^{n} U_{ni}|\mathbf{A}_{n+1}) \\ &= E(f_0(X_1) - P(f_0)|\mathbf{A}_{n+1}) \end{aligned}$$

and, similarly,

$$\begin{aligned} S_{n+1}(f_0) &= E(S_{n+1}(f_0)|\mathbf{A}_{n+1}) \\ &= E(E(S_{n+1}(f_0)|U_{n+1}, \mathbf{A}_{n+1})|\mathbf{A}_{n+1}) \\ &= E(\sum_{i=1}^{n} U_{n+1,i} E(f_0(X_i) - P(f_0)|U_{n+1}, \mathbf{A}_{n+1}) / \sum_{i=1}^{n+1} U_{n+1}, i|\mathbf{A}_{n+1}) \\ &= E(f_0(X_1) - P(f_0)|\mathbf{A}_{n+1}) \end{aligned}$$

hold, $\{S_n(f_0), \mathbf{A}_n\}$ is a reversed martingale. In view of

$$E(\|S_n\|_{\mathcal{F}}^*|\mathbf{A}_{n+1}) \geq \{\sup_{f\in\mathcal{F}} |E(S_n(f)|\mathbf{A}_{n+1})|\}^* = \|S_n\|_{\mathcal{F}}^*,$$

the sequence $\{\|S_n\|_{\mathcal{F}}^*, \mathbf{A}_n\}$ is a reversed submartingale bounded from below by 0. Consequently, there exists a nonnegative random variable S with

$$\|S_n\|_{\mathcal{F}}^* \to S \quad \text{a.s.} \quad P^\infty \times \lambda.$$

Write $N \equiv \sum_1^n U_{ni}$ and note that

$$\begin{aligned} Pr(\|S_n\|_{\mathcal{F}}^* > \epsilon) &= E\, Pr(\|S_n\|_{\mathcal{F}}^* > \epsilon | U_n) \\ &= E\, Pr(\|\mathbb{P}_N - P\|_{\mathcal{F}}^* > \epsilon | U_n) \to 0 \end{aligned}$$

by (a) and dominated convergence. Consequently S must be degenerate at 0. □

Applying this lemma to the summands in (c), we conclude that $R_n(\omega)$ converges in probability to 0 a.s. P^∞ if (a) holds.

Assume B. Since B implies A by theorem 2, (a) follows immediately since $\mathcal{F} \in CLT(P)$ implies that $\mathcal{F}$ is P-Glivenko-Cantelli in probability and together with $P(F) \leq (P(F^2))^{1/2} < \infty$ this implies $\mathcal{F}$ is P-Glivenko-Cantelli a.s. (e.g. by the proof of lemma 3.2 of Pollard (1981)).

Finally, suppose that C holds. To prove (a), we first show:

Claim 1. C implies $P(F) < \infty$ and $P(F^2) < \infty$.

Proof. Giné and Zinn (1990), page 858, show that C implies $P(F^2) < \infty$ using a result on the CLT in Banach space (Araujo and Giné (1980), theorem 3.5.4). A direct proof of $P(F) < \infty$ (which is all we need here) is also possible (and takes about a page), but will be omitted. □

Claim 2. C implies $EV_n \to 0$.

Proof. Let $\epsilon > 0$. Then

$$V_n \le \frac{1}{n}\sum_{i=1}^{n}(M_{ni}^{\#}+1)F(X_i) \equiv W_n$$

where $EW_n = 2P(F) < \infty$ by claim 1. Since $E(M_{ni}^{\#}+1) = 2$, there exists a $K = K(\epsilon)$ with

$$\max_{1\le i\le n} E(M_{ni}^{\#}+1)1_{[M_{ni}^{\#}\ge K]} \le \epsilon.$$

Then we have, with $\underline{X} \equiv (X_1, \ldots, X_n)$,

$$\begin{aligned} E(V_n|\underline{X}) &\le \epsilon + E(W_n 1_{[V_n>\epsilon]}|\underline{X}) \\ &= \epsilon + \frac{1}{n}\sum_{i=1}^{n} F(X_i)E\{(M_{ni}^{\#}+1)(1_{[M_{ni}^{\#}\le K-1]} + 1_{[M_{ni}^{\#}\ge K]})1_{[V_n>\epsilon]}\} \\ &\le \epsilon + \frac{1}{n}\sum_{i=1}^{n} F(X_i)\{K\,Pr(V_n > \epsilon|\underline{X}) + \epsilon\}, \end{aligned}$$

and hence, by symmetry considerations,

$$\begin{aligned} EV_n &\le \epsilon + \epsilon P(F) + K\,E\{F(X_1)\,Pr(V_n > \epsilon|\underline{X})\} \\ &\to \epsilon(1+P(F)) \quad \text{as} \quad n\to\infty \end{aligned}$$

by (e), claim 1, and dominated convergence. Since $\epsilon > 0$ is arbitrary, this yields the claim. □

Claim 3. $(1-1/n)^n E\|\mathbb{P}_n - P\|_{\mathcal{F}}^* \le EV_n$.

Proof. This follows from two applications of Jensen's inequality:

Since $M_{ni}^{\#} \sim$Binomial$(n, 1/n)$,

$$\begin{aligned} (1-\frac{1}{n})^n E\|\mathbb{P}_n - P\|_{\mathcal{F}}^* &= E\|n^{-1}\sum_{i=1}^{n} E(1_{[M_{ni}^{\#}=0]})(\delta_{X_i} - P)\|_{\mathcal{F}}^* \\ &\le E\|n^{-1}\sum_{i=1}^{n} 1_{[M_{ni}^{\#}=0]}(\delta_{X_i} - P)\|_{\mathcal{F}}^* \\ &\le E\|n^{-1}\sum_{i=1}^{n}(M_{ni}^{\#}-1)(\delta_{X_i} - P)\|_{\mathcal{F}}^* = E\,V_n \end{aligned}$$

where the last inequality follows from $E\|U\| \le E\|U+V\|$ if U, V are independent with $EV = 0$ (applied conditionally). □

Combining claims 2 and 3 yields $E\|\mathbb{P}_n - P\|_{\mathcal{F}}^* \to 0$ as $n \to \infty$. Since $P(F) < \infty$, this implies (a): see e.g. the proof of lemma 3.2 of Pollard (1981). □

Remarks: Praestgaard (1991) has extended the A implies C part of theorem 3 to a very large class of exchangeable weights $\{W_{ni}\}$ in place of the multinomial weights $\{M_{ni}\}$ of Efron's bootstrap used in theorem 3. Wellner (1992) has shown the equivalence of A - C in theorem 3 with convergence of the sequential bootstrap empirical process $\mathbb{K}_n$ defined on $[0,1] \times \mathcal{F}$ by

$$\mathbb{K}_n(t,f) = n^{-1/2} \sum_{i=1}^{n} (M^{\#}_{[nt],i} - [nt]/n)(f(X_i(\omega)) - Pf), \quad 0 \le t \le 1, \quad f \in \mathcal{F}.$$

3. The empirical process with random sample size

Now consider the empirical measure $\mathbb{P}_n$ and the empirical process $\mathbb{X}_n$ based on X_i's iid P on $(A, \mathbf{A})$ as in section 1. As explained by Pyke (1968), it is often the case in practice that the sample size available to the statistician is random, and perhaps dependent on the X_i's. Suppose that $\{N_n : n \ge 1\}$ is a positive, integer - valued stochastic process satisfying

$$n^{-1} N_n \to_p \nu \quad \text{as} \quad n \to \infty \tag{1}$$

where ν is a positive random variable, i.e.

$$Pr(\nu > 0) = 1. \tag{2}$$

The following theorem says that the randomness of N_n does not upset convergence of the empirical process $\mathbb{X}_n$ when the sample size n is replaced by N_n as long as (1) and (2) hold.

Theorem 4. Suppose that $\mathcal{F} \subset L_2(P)$ is P-Donsker so that $\mathbb{X}_n \Rightarrow \mathbb{X} \sim G_P$. If $\{N_n\}$ and ν satisfy (1) and (2), then

$$\mathbb{X}_{N_n} \Rightarrow \mathbb{X} \sim G_P \quad \text{in} \quad l^\infty(\mathcal{F}) \quad \text{as} \quad n \to \infty .$$

Of course this result remains valid if (1) is replaced by

$$c_n N_n \to_p \nu, \quad c_n \downarrow 0 \quad \text{as} \quad n \to \infty. \tag{3}$$

Proof. It suffices to show that

$$\mathbb{Y}_n \equiv \mathbb{X}_{N_n} \to_{f.d.} G_P \quad \text{as} \quad n \to \infty, \tag{a}$$

and, for every $\epsilon > 0$, that

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} Pr_P^*\{\|\mathbb{Y}_n\|_{\mathcal{F}'(\delta,\rho_P)} > \epsilon\} = 0. \tag{b}$$

The finite-dimensional convergence in (a) follows by use of the Cramér-Wold device and the random sample size central limit theorem of Blum, Hanson, and Rosenblatt (1963); this is almost exactly as in Csörgő (1974), pages 20 - 22.

It remains to prove (b). Let $\gamma > 0$. Choose $0 < a < b < \infty$ so that

$$Pr(a < \nu \le b) > 1 - \gamma. \tag{c}$$

With $0 < \eta$ we have

$$\lim_{n \to \infty} Pr_P(|n^{-1} N_n - \nu| > \eta) = 0 \tag{d}$$

by assumption (1). Hence with the choice $0 < \eta < a$ the left side of (b) is bounded by

$$\gamma + \lim_{\delta \downarrow 0} \limsup_{n \to \infty} Pr_P^*\{\max_{n(a-\eta) \le k \le n(b+\eta)} \|\mathbb{X}_k\|_{\mathcal{F}'(\delta,\rho_P)} > \epsilon\}$$

$$\le \gamma + \lim_{\delta \downarrow 0} \limsup_{n \to \infty} Pr_P^*\{\max_{k \le [n(b+\eta)]} \| \frac{\sum_{i=1}^k (\delta_{X_i} - P)}{\sqrt{n(b+\eta)}} \|_{\mathcal{F}'(\delta,\rho_P)} > \epsilon \frac{\sqrt{n(a-\eta)}}{\sqrt{n(b+\eta)}}\}$$

$$\le \gamma + \lim_{\delta \downarrow 0} \limsup_{n \to \infty} 2\, Pr_P^*\{\|\mathbb{X}_{m_n}\|_{\mathcal{F}'(\delta,\rho_P)} > \frac{\epsilon}{2} \frac{\sqrt{a-\eta}}{\sqrt{b+\eta}}\} \tag{e}$$

with $m_n \equiv [n(b+\eta)]$ by Ottaviani's inequality

(see e.g. Dudley (1984), inequality 3.2.7 with $\|\| \equiv \|\|_{\mathcal{F}'(\delta,\rho_P)}$),

provided

$$\sup_{k \le m_n} Pr_P^*(\|\mathbb{X}_k\|_{\mathcal{F}'(\delta,\rho_P)} > \frac{\epsilon\sqrt{n(a-\eta)}}{2\sqrt{k}}) \le \frac{1}{2}. \tag{f}$$

By theorem 4.1.1 of Dudley (1984), $\mathcal{F} \in CLT(P)$ implies that

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} Pr_P^*\{\|\mathbb{X}_n\|_{\mathcal{F}'(\delta,\rho_P)} > \zeta\} = 0 \tag{g}$$

for every $\zeta > 0$. Consequently (f) is valid for every $\epsilon > 0$ and for n sufficiently large, and moreover the right side of (e) equals γ. Since $\gamma > 0$ is arbitrary, (b) holds. □

Remarks: Pyke (1968) proved theorem 4 in the one-dimensional distribution function case: $(A, \mathbf{A}, P) = ([0,1], \mathbf{B}, Lebesgue)$, and $\mathcal{F} = \{1_{[0,t]} : 0 \leq t \leq 1\}$ under the assumption (1) with $P(\nu = 1) = 1$. Billingsley (1968) and S. Csörgő (1974), following Blum, Hanson, Rosenblatt (1963), showed that Pyke's one-dimensional result continues to hold under (1) and (2), allowing a general positive limit random variable ν. These results were extended by Wichura (1968) and Fernandez (1970); the latter includes the "desirable generalization" mentioned by Pyke (1968) in his closing remark. Note that in view of the results of Dudley (1985) sections 6 and 7, theorem 4 also contains these results as well as many others.

Acknowledgments: We owe thanks to Evarist Giné, David Mason, Ronald Pyke, and Joel Zinn for helpful discussions. A referee suggested the present proof of theorem 1.

References

Andersen, N. T. (1985). The central limit theorem for non-separable valued functions. *Z. Wahrsch. verw. Gebiete,* **70**, 445–455.

Araujo, A. and Giné, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables,* Wiley, New York.

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

Blum, J. R., D. L. Hanson, and J. I. Rosenblatt (1963). On the central limit theorem for the sum of a random number of independent random variables. *Z. Wahrsch. verw. Geb.,* **1**, 389–393.

Csörgő , M. (1972). Distribution results for distance functions based on the modified empirical distribution function of M. Kac. *Ann. Inst. Statist. Math.,* **24**, 101 - 110.

Csörgő , M. and Alvo, M. (1970). Distribution results and power functions for Kac statistics. *Ann. Inst. Statist. Math.,* **22**, 257 - 260.

Csörgő , S. (1974). On weak convergence of the empirical process with random sample size. *Acta Sci. Math. Szeged,* **36**, 17–25.

Csörgő , S. (1981). Strong approximation of empirical Kac processes. *Ann. Inst. Statist. Math.,* **33**, 417 - 423.

Donsker, M. D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.,* **23**, 277–281.

Dudley, R. M. (1984). A course on empirical processes; École d'Été de Probabilités de St. Flour. *Lecture Notes in Math.,* **1097**, 2–142, Springer Verlag, New York.

Dudley, R. M. (1985). An extended Wichura theorem, definitions of Donsker class, and weighted empirical distributions. *Lecture Notes in Math.*, **1153**, 141–178, Springer - Verlag, New York.

Durst, M. and R. M. Dudley (1981). Empirical processes, Vapnik-Chervonenkis classes, and Poisson processes. *Prob. Math. Statist. (Wroclaw)*, **1**, 109-115.

Fernandez, P. J. (1970). A weak convergence theorem for random sums of independent random variables. *Ann. Math. Statist.*, **41**, 710–712.

Giné, E. and J. Zinn (1984). Some limit theorems for empirical processes. *Ann. Probability*, **12**, 929–989.

Giné, E. and J. Zinn (1986). Lectures on the central limit theorem for empirical processes. *Lect. Notes in Math.*, **1221**, 50–113, Springer, Berlin.

Giné, E. and J. Zinn (1990). Bootstrapping general empirical measures. *Ann. Probability*, **18**, 851–869.

Giné, E. and J. Zinn (1991). Gaussian characterization of uniform Donsker classes of functions, *Ann. Probability* , **19**, 758 - 782.

Hoffmann-Jørgensen, J. (1984). *Stochastic Processes on Polish Spaces.* Unpublished manuscript.

Kac, M. (1949). On deviations between theoretical and empirical distributions. *Proc. Natl. Acad. Sci.*, **35**, 252–257.

Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giorn. Ist. Ital. Attuari*, **4**, 83–91.

Le Cam, L. (1970). Remarques sur le théorème limite central dans les espaces localement convexes. *Les Probabilités sur les Structures Algébriques*, C.N.R.S., Paris, 233 - 249.

Ledoux, M. and M. Talagrand (1986). Conditions d'integrabilité pour les multiplicateurs dans le TLC Banachique. *Ann. Probability*, **14**, 916–921.

Ledoux, M. and M. Talagrand (1988). Un critère sur les petites boules dans le théorème limite central. *Probab. Th. Rel Fields*, **77**, 29–47.

Ledoux, M. and M. Talagrand (1991). Probability in Banach Spaces. Springer-Verlag, New York.

Massart, P. (1989). Strong approximation for multivariate empirical and related processes, via KMT constructions. *Ann. Probability*, **17**, 266–291.

Pollard, D. (1981). Limit theorems for empirical processes. *Z. Wahrsch. verw. Geb.*, **57**, 181–195.

Praestgaard, J. (1990). *Bootstrap with general weights and multiplier central limit theorems.* Technical Report **195**, Department of Statistics, University of Washington.

Praestgaard, J. (1991). *General-Weights Bootstrap of the Empirical Process.* Unpublished Ph.D. dissertation, Department of Statistics, University of Washington.

Pyke, R. (1968). The weak convergence of the empirical process with random sample size. *Proc. Cambridge Phil. Soc.*, **64**, 155-160.

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

Suzuki, G. (1972). Distributions of Kac-statistics. *Ann. Inst. Statist. Math.*, **24**, 415 - 421.

Van der Vaart, A. W. and J. A. Wellner (1990). *Prohorov and continuous mappings theorems in the Hoffmann-Jørgensen weak convergence theory with applications to convolution and asymptotic minimax theorems.* Technical Report **157**, Department of Statistics, University of Washington, Seattle.

Wellner, J. A. (1992). Bootstrap limit theorems: a partial survey. To appear in *Nonparametric Statistics and Related Topics,* A.K.Md.E. Saleh, editor. Elsevier, Amsterdam.

Wichura, M. J. (1968). *On the weak convergence of non-Borel probabilities on a metric space.* Ph.D. dissertation, Columbia University.

Chris A. J. Klaassen
Department of Mathematics
University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
The Netherlands

Jon A. Wellner
Department of Statistics, GN-22
University of Washington
Seattle, Washington 98195
USA