

Berti, Patrizia; Pratelli, Luca; Rigo, Pietro

Working Paper

A Skorohod Representation Theorem for Uniform Distance

Quaderni di Dipartimento, No. 109

Provided in Cooperation with:

University of Pavia, Department of Economics and Quantitative Methods (EPMQ)

Suggested Citation: Berti, Patrizia; Pratelli, Luca; Rigo, Pietro (2010) : A Skorohod Representation Theorem for Uniform Distance, Quaderni di Dipartimento, No. 109, Università degli Studi di Pavia, Dipartimento di Economia Politica e Metodi Quantitativi (EPMQ), Pavia

This Version is available at:

<https://hdl.handle.net/10419/95306>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Quaderni di Dipartimento

**A Skorohod Representation Theorem
for Uniform Distance**

Patrizia Berti
(Università di Modena e Reggio Emilia)

Luca Pratelli
(Accademia Navale di Livorno)

Pietro Rigo
(Università di Pavia)

109 (01-10)

Dipartimento di economia politica
e metodi quantitativi
Università degli studi di Pavia
Via San Felice, 5
I-27100 Pavia

Gennaio 2010

A SKOROHOD REPRESENTATION THEOREM FOR UNIFORM DISTANCE

PATRIZIA BERTI, LUCA PRATELLI, AND PIETRO RIGO

ABSTRACT. Let μ_n be a probability measure on the Borel σ -field on $D[0, 1]$ with respect to Skorohod distance, $n \geq 0$. Necessary and sufficient conditions for the following statement are provided. On some probability space, there are $D[0, 1]$ -valued random variables X_n such that $X_n \sim \mu_n$ for all $n \geq 0$ and $\|X_n - X_0\| \rightarrow 0$ in probability, where $\|\cdot\|$ is the sup-norm. Such conditions do not require μ_0 separable under $\|\cdot\|$. Applications to exchangeable empirical processes and to pure jump processes are given as well.

1. INTRODUCTION

Let D be the set of real cadlag functions on $[0, 1]$ and

$$\|x\| = \sup_t |x(t)|, \quad u(x, y) = \|x - y\|, \quad x, y \in D.$$

Also, let d be Skorohod distance and $\mathcal{B}_d, \mathcal{B}_u$ the Borel σ -fields on D with respect to (w.r.t.) d and u , respectively.

In real problems, one usually starts with a sequence $(\mu_n : n \geq 0)$ of probabilities on \mathcal{B}_d . If $\mu_n \rightarrow \mu_0$ weakly (under d), Skorohod representation theorem yields $d(X_n, X_0) \xrightarrow{a.s.} 0$ for some D -valued random variables X_n such that $X_n \sim \mu_n$ for all $n \geq 0$. However, X_n can fail to approximate X_0 uniformly. A trivial example is $\mu_n = \delta_{x_n}$, where $(x_n) \subset D$ is any sequence such that $x_n \rightarrow x_0$ according to d but not according to u .

Lack of uniform convergence is sometimes a trouble. Thus, given a sequence $(\mu_n : n \geq 0)$ of laws on \mathcal{B}_d , it is useful to have conditions for:

- (1) On some probability space (Ω, \mathcal{A}, P) , there are random variables $X_n : \Omega \rightarrow D$ such that $X_n \sim \mu_n$ for all $n \geq 0$ and $\|X_n - X_0\| \xrightarrow{P} 0$.

Convergence in probability cannot be strengthened into a.s. convergence in condition (1). In fact, it may be that (1) holds, and yet there are not D -valued random variables Y_n such that $Y_n \sim \mu_n$ for all n and $\|Y_n - Y_0\| \xrightarrow{a.s.} 0$; see Example 7.

This paper is concerned with (1). The main result is Theorem 4, which states that (1) holds if and only if

$$(2) \quad \limsup_n \sup_{f \in L} |\mu_n(f) - \mu_0(f)| = 0,$$

2000 *Mathematics Subject Classification.* 60B10, 60A05, 60A10.

Key words and phrases. Cadlag function – Exchangeable empirical process – Separable probability measure – Skorohod representation theorem – Uniform distance – Weak convergence of probability measures.

where L is the set of functions $f : D \rightarrow \mathbb{R}$ satisfying

$$\sigma(f) \subset \mathcal{B}_d, \quad -1 \leq f \leq 1, \quad |f(x) - f(y)| \leq \|x - y\| \text{ for all } x, y \in D.$$

Theorem 4 can be commented as follows. Say that a probability μ , defined on \mathcal{B}_d or \mathcal{B}_u , is *u-separable* in case $\mu(A) = 1$ for some *u-separable* $A \in \mathcal{B}_d$. Suppose μ_0 is *u-separable* and define $\mu_0^*(H) = \mu_0(A \cap H)$ for $H \in \mathcal{B}_u$, where $A \in \mathcal{B}_d$ is *u-separable* and $\mu_0(A) = 1$. Since μ_n is defined only on \mathcal{B}_d for $n \geq 1$, we adopt Hoffmann-Jørgensen's definition of convergence in distribution for non measurable random elements; see e.g. [7] and [9]. Let I_0 be the identity map on $(D, \mathcal{B}_u, \mu_0^*)$ and I_n the identity map on $(D, \mathcal{B}_d, \mu_n)$, $n \geq 1$. Further, let D be regarded as a metric space under u . Then, since μ_0^* is *u-separable*, one obtains:

- (i) Condition (1) holds (with $\|X_n - X_0\| \xrightarrow{a.s.} 0$) provided $I_n \rightarrow I_0$ in distribution;
- (ii) $I_n \rightarrow I_0$ in distribution if and only if $\lim_n \sup_{f \in L} |\mu_n(f) - \mu_0(f)| = 0$.

Both (i) and (ii) are known facts; see Theorems 1.7.2, 1.10.3 and 1.12.1 of [9].

The spirit of Theorem 4, thus, is that one can dispense with *u-separability* of μ_0 to get (1). This can look surprising, as separability of the limit law is crucial in Skorohod representation theorem; see [5]. However, $X_n \sim \mu_n$ is asked only on \mathcal{B}_d and not on \mathcal{B}_u . Indeed, X_n can even fail to be measurable w.r.t. \mathcal{B}_u .

Non *u-separable* laws on \mathcal{B}_d are quite usual. A cadlag process Z , with jumps at random time points, has typically a non *u-separable* distribution on \mathcal{B}_d . One example is $Z(t) = B_{M(t)}$, where B is a standard Brownian bridge, M an independent random distribution function and the jump-points of M have a non discrete distribution. Such a Z is the limit in distribution, under d , of certain exchangeable empirical processes; see [1] and [3].

In applications, unless μ_0 is *u-separable*, checking condition (2) is usually difficult. In this sense, Theorem 4 can be viewed as a "negative" result, as it states that condition (1) is quite hard to reach. This is partly true. However, there are also meaningful situations where (2) can be proved with a reasonable effort. Two examples are exchangeable empirical processes, which motivated Theorem 4, and a certain class of jump processes. Both are discussed in Section 4.

Our proof of Theorem 4 is admittedly long and it is confined in a final appendix. Some preliminary results, of possible independent interest, are needed. We mention Proposition 2 and Lemma 13 in particular.

A last remark is that Theorem 4 is still valid if D is replaced by $D([0, 1], \mathcal{X})$, the space of cadlag functions from $[0, 1]$ into a separable Banach space \mathcal{X} .

2. A PRELIMINARY RESULT

Let (Ω, \mathcal{A}, P) be a probability space. The outer and inner measures are

$$P^*(H) = \inf\{P(A) : H \subset A \in \mathcal{A}\}, \quad P_*(H) = 1 - P^*(H^c), \quad H \subset \Omega.$$

Given a metric space (S, ρ) and maps $X_n : \Omega \rightarrow S$, $n \geq 0$, say that X_n *converges to X_0 in (outer) probability*, written $X_n \xrightarrow{P} X_0$, in case

$$\lim_n P^*(\rho(X_n, X_0) > \epsilon) = 0 \quad \text{for all } \epsilon > 0.$$

In the sequel, d_{TV} denotes *total variation distance* between two probabilities defined on the same σ -field.

Proposition 1. *Let (F, \mathcal{F}) be a measurable space and μ_n a probability on (F, \mathcal{F}) , $n \geq 0$. Then, on some probability space (Ω, \mathcal{A}, P) , there are measurable maps $X_n : (\Omega, \mathcal{A}) \rightarrow (F, \mathcal{F})$ such that*

$$P_*(X_n \neq X_0) = P^*(X_n \neq X_0) = d_{TV}(\mu_n, \mu_0) \text{ and } X_n \sim \mu_n \text{ for all } n \geq 0.$$

Proposition 1 is well known, even if in a slightly different form; see Theorem 2.1 of [8]. A proof of the present version is in Section 3 of [5].

Next proposition is fundamental for proving our main result. Among other things, it can be viewed as an improvement of Proposition 1.

Proposition 2. *Let λ_n be a probability on $(F \times G, \mathcal{F} \otimes \mathcal{G})$, $n \geq 0$, where (F, \mathcal{F}) is a measurable space and (G, \mathcal{G}) a Polish space equipped with its Borel σ -field. The following conditions are equivalent:*

(a) *There are a probability space (Ω, \mathcal{A}, P) and measurable maps $(Y_n, Z_n) : (\Omega, \mathcal{A}) \rightarrow (F \times G, \mathcal{F} \otimes \mathcal{G})$ such that*

$$(Y_n, Z_n) \sim \lambda_n \text{ for all } n \geq 0, \quad P^*(Y_n \neq Y_0) \rightarrow 0, \quad Z_n \xrightarrow{P} Z_0;$$

(b) *For each bounded Lipschitz function $f : G \rightarrow \mathbb{R}$,*

$$\lim_n \sup_{A \in \mathcal{F}} \left| \int I_A(y) f(z) \lambda_n(dy, dz) - \int I_A(y) f(z) \lambda_0(dy, dz) \right| = 0.$$

To prove Proposition 2, we first recall a result of Blackwell and Dubins [6].

Theorem 3. *Let G be a Polish space, \mathcal{M} the collection of Borel probabilities on G , and m the Lebesgue measure on $(0, 1)$. There is a Borel measurable map*

$$\Phi : \mathcal{M} \times (0, 1) \rightarrow G$$

such that, for every $\nu \in \mathcal{M}$,

- (i) $\Phi(\nu, \cdot) \sim \nu$ under m ;
- (ii) *There is a Borel set $A_\nu \subset (0, 1)$ such that $m(A_\nu) = 1$ and*

$$\Phi(\nu_n, t) \rightarrow \Phi(\nu, t) \quad \text{whenever } t \in A_\nu, \nu_n \in \mathcal{M} \text{ and } \nu_n \rightarrow \nu \text{ weakly.}$$

We also need to recall disintegrations. Let λ be a probability on $(F \times G, \mathcal{F} \otimes \mathcal{G})$, where (F, \mathcal{F}) and (G, \mathcal{G}) are arbitrary measurable spaces. In this paper, λ is said to be *disintegrable* if there is a collection $\alpha = \{\alpha(y) : y \in F\}$ such that:

- $\alpha(y)$ is a probability on \mathcal{G} for $y \in F$;
- $y \mapsto \alpha(y)(C)$ is \mathcal{F} -measurable for $C \in \mathcal{G}$;
- $\lambda(A \times C) = \int_A \alpha(y)(C) \mu(dy)$ for $A \in \mathcal{F}$ and $C \in \mathcal{G}$, where $\mu(\cdot) = \lambda(\cdot \times G)$.

Such an α is called a *disintegration* for λ . For λ to admit a disintegration, it suffices that G is a Borel subset of a Polish space and \mathcal{G} the Borel σ -field on G .

Proof of Proposition 2. "(a) \Rightarrow (b)". Under (a), for each $A \in \mathcal{F}$ and bounded Lipschitz $f : G \rightarrow \mathbb{R}$, one obtains

$$\begin{aligned} \left| \int I_A(y) f(z) \lambda_n(dy, dz) - \int I_A(y) f(z) \lambda_0(dy, dz) \right| &= \left| E_P \{ I_A(Y_n) f(Z_n) - I_A(Y_0) f(Z_0) \} \right| \\ &\leq E_P \left| f(Z_n) (I_A(Y_n) - I_A(Y_0)) \right| + E_P \left| I_A(Y_0) (f(Z_n) - f(Z_0)) \right| \\ &\leq \sup |f| P^*(Y_n \neq Y_0) + E_P \left| f(Z_n) - f(Z_0) \right| \rightarrow 0. \end{aligned}$$

"(b) \Rightarrow (a)". Let $\mu_n(A) = \lambda_n(A \times G)$, $A \in \mathcal{F}$. By (b), $d_{TV}(\mu_n, \mu_0) \rightarrow 0$. Hence, by Proposition 1, on a probability space (Θ, \mathcal{E}, Q) there are measurable maps $h_n : (\Theta, \mathcal{E}) \rightarrow (F, \mathcal{F})$ satisfying $h_n \sim \mu_n$ for all n and $Q^*(h_n \neq h_0) \rightarrow 0$. Let

$$\Omega = \Theta \times (0, 1), \quad \mathcal{A} = \mathcal{E} \otimes \mathcal{B}_{(0,1)}, \quad P = Q \times m,$$

where $\mathcal{B}_{(0,1)}$ is the Borel σ -field on $(0, 1)$ and m the Lebesgue measure.

Since G is Polish, each λ_n admits a disintegration $\alpha_n = \{\alpha_n(y) : y \in F\}$. By Theorem 3, there is a map $\Phi : \mathcal{M} \times (0, 1) \rightarrow G$ satisfying conditions (i)-(ii). Let

$$Y_n(\theta, t) = h_n(\theta) \quad \text{and} \quad Z_n(\theta, t) = \Phi\{\alpha_n(h_n(\theta)), t\}, \quad (\theta, t) \in \Theta \times (0, 1).$$

For fixed θ , condition (i) yields $Z_n(\theta, \cdot) = \Phi\{\alpha_n(h_n(\theta)), \cdot\} \sim \alpha_n(h_n(\theta))$ under m . Since α_n is a disintegration for λ_n , for all $A \in \mathcal{F}$ and $C \in \mathcal{G}$ one has

$$\begin{aligned} P(Y_n \in A, Z_n \in C) &= \int_{\Theta} I_A(h_n(\theta)) m\{t : Z_n(\theta, t) \in C\} Q(d\theta) \\ &= \int_{\{h_n \in A\}} \alpha_n(h_n(\theta))(C) Q(d\theta) = \int_A \alpha_n(y)(C) \mu_n(dy) = \lambda_n(A \times C). \end{aligned}$$

Also, $P^*(Y_n \neq Y_0) = Q^*(h_n \neq h_0) \rightarrow 0$ by Lemma 1.2.5 of [9].

Finally, we prove $Z_n \xrightarrow{P} Z_0$. Write $\alpha_n(y)(f) = \int f(z) \alpha_n(y)(dz)$ for all $y \in F$ and $f \in L_G$, where L_G is the set of Lipschitz functions $f : G \rightarrow [-1, 1]$. Since $Q^*(h_n \neq h_0) \rightarrow 0$, there are $A_n \in \mathcal{F}$ such that $Q(A_n^c) \rightarrow 0$ and $h_n = h_0$ on A_n . Given $f \in L_G$,

$$\begin{aligned} E_Q \left| \alpha_n(h_n)(f) - \alpha_0(h_0)(f) \right| - 2Q(A_n^c) &\leq E_Q \left\{ I_{A_n} \left| \alpha_n(h_0)(f) - \alpha_0(h_0)(f) \right| \right\} \\ &\leq E_Q \left| \alpha_n(h_0)(f) - \alpha_0(h_0)(f) \right| = \int \left| \alpha_n(y)(f) - \alpha_0(y)(f) \right| \mu_0(dy). \end{aligned}$$

Using condition (b), it is not hard to see that $\int |\alpha_n(y)(f) - \alpha_0(y)(f)| \mu_0(dy) \rightarrow 0$. Therefore, $\alpha_n(h_n)(f) \xrightarrow{Q} \alpha_0(h_0)(f)$ for each $f \in L_G$, and this is equivalent to

each subsequence (n') contains a further subsequence (n'')

such that $\alpha_{n''}(h_{n''}(\theta)) \rightarrow \alpha_0(h_0(\theta))$ weakly for Q -almost all θ ;

see Remark 2.3 and Corollary 2.4 of [2]. Thus, by property (ii) of Φ , each subsequence (n') contains a further subsequence (n'') such that $Z_{n''} \xrightarrow{a.s.} Z_0$. That is, $Z_n \xrightarrow{P} Z_0$ and this concludes the proof. \square

3. EXISTENCE OF CADLAG PROCESSES, WITH GIVEN DISTRIBUTIONS ON THE SKOROHOD BOREL σ -FIELD, CONVERGING UNIFORMLY IN PROBABILITY

As in Section 1, \mathcal{B}_d and \mathcal{B}_u are the Borel σ -fields on D w.r.t. d and u . Also, L is the class of functions $f : D \rightarrow [-1, 1]$ which are measurable w.r.t. \mathcal{B}_d and Lipschitz w.r.t. u with Lipschitz constant 1. We recall that, for $x, y \in D$, the Skorohod distance $d(x, y)$ is the infimum of those $\epsilon > 0$ such that

$$\|x - y \circ \gamma\| \leq \epsilon \quad \text{and} \quad \sup_{s \neq t} \left| \log \frac{\gamma(s) - \gamma(t)}{s - t} \right| \leq \epsilon$$

for some strictly increasing homeomorphism $\gamma : [0, 1] \rightarrow [0, 1]$. The metric space (D, d) is separable and complete.

We write $\mu(f) = \int f d\mu$ whenever μ is a probability on a σ -field and f a real bounded function, measurable w.r.t. such a σ -field.

Motivations for the next result have been given in Section 1.

Theorem 4. *Let μ_n be a probability measure on \mathcal{B}_d , $n \geq 0$. Then, conditions (1) and (2) are equivalent, that is,*

$$\limsup_n \sup_{f \in L} |\mu_n(f) - \mu_0(f)| = 0$$

if and only if there are a probability space (Ω, \mathcal{A}, P) and measurable maps $X_n : (\Omega, \mathcal{A}) \rightarrow (D, \mathcal{B}_d)$ such that $X_n \sim \mu_n$ for each $n \geq 0$ and $\|X_n - X_0\| \xrightarrow{P} 0$.

The proof of Theorem 4 is given in the Appendix. Here, we state a corollary and an open problem and we make two examples.

In applications, the μ_n are often probability distributions of random variables, all defined on some probability space $(\Omega_0, \mathcal{A}_0, P_0)$. In the spirit of [4], a (minor) question is whether condition (1) holds with the X_n defined on $(\Omega_0, \mathcal{A}_0, P_0)$ as well.

Corollary 5. *Let $(\Omega_0, \mathcal{A}_0, P_0)$ be a probability space and $Z_n : (\Omega_0, \mathcal{A}_0) \rightarrow (D, \mathcal{B}_d)$ a measurable map, $n \geq 1$. Suppose $\lim_n \sup_{f \in L} |E_{P_0}\{f(Z_n)\} - \mu_0(f)| = 0$ for some probability measure μ_0 on \mathcal{B}_d . If P_0 is nonatomic, there are measurable maps $X_n : (\Omega_0, \mathcal{A}_0) \rightarrow (D, \mathcal{B}_d)$, $n \geq 0$, such that*

$$X_0 \sim \mu_0, \quad X_n \sim Z_n \text{ for each } n \geq 1, \quad \|X_n - X_0\| \xrightarrow{P_0} 0.$$

Also, P_0 is nonatomic if $\mu_0\{x\} = 0$ for all $x \in D$, or if $P_0(Z_n = x) = 0$ for some $n \geq 1$ and all $x \in D$.

Proof. Since (D, d) is separable, P_0 is nonatomic if $P_0(Z_n = x) = 0$ for some $n \geq 1$ and all $x \in D$. By Corollary 5.4 of [4], $(\Omega_0, \mathcal{A}_0, P_0)$ supports a D -valued random variable Z_0 with $Z_0 \sim \mu_0$. Hence, P_0 is nonatomic even if $\mu_0\{x\} = 0$ for all $x \in D$. Next, by Theorem 4, on a probability space (Ω, \mathcal{A}, P) there are D -valued random variables Y_n such that $Y_0 \sim \mu_0$, $Y_n \sim Z_n$ for $n \geq 1$ and $\|Y_n - Y_0\| \xrightarrow{P} 0$. Let $(D^\infty, \mathcal{B}_d^\infty)$ be the countable product of (D, \mathcal{B}_d) and

$$\nu(A) = P((Y_0, Y_1, \dots) \in A), \quad A \in \mathcal{B}_d^\infty.$$

Then, ν is a Borel probability on a Polish space. Thus, if P_0 is nonatomic, $(\Omega_0, \mathcal{A}_0, P_0)$ supports a D^∞ -valued random variable $X = (X_0, X_1, \dots)$ with $X \sim \nu$; see e.g. Theorem 3.1 of [4]. Since $(X_0, X_1, \dots) \sim (Y_0, Y_1, \dots)$, this concludes the proof. \square

Let (S, ρ) be a metric space such that $(x, y) \mapsto \rho(x, y)$ is measurable w.r.t. $\mathcal{E} \otimes \mathcal{E}$, where \mathcal{E} is the ball σ -field on S . This is actually true in case $(S, \rho) = (D, u)$ and it is very useful to prove Theorem 4. Thus, a question is whether (D, u) can be replaced by (S, ρ) in Theorem 4. Precisely, let $(\mu_n : n \geq 0)$ be a sequence of laws on \mathcal{E} and L_S the class of functions $f : S \rightarrow [-1, 1]$ such that $\sigma(f) \subset \mathcal{E}$ and $|f(x) - f(y)| \leq \rho(x, y)$ for all $x, y \in S$. Then,

Conjecture: $\lim_n \sup_{f \in L_S} |\mu_n(f) - \mu_0(f)| = 0$ if and only if $\rho(X_n, X_0) \rightarrow 0$ in probability for some S -valued random variables X_n such that $X_n \sim \mu_n$ for all n .

We finally give two examples. The first shows that condition (2) cannot be weakened into $\mu_n(f) \rightarrow \mu_0(f)$ for each fixed $f \in L$.

Example 6. For each $n \geq 0$, let $h_n : (0, 1) \rightarrow [0, \infty)$ be a Borel function such that $\int_0^1 h_n(t) dt = 1$. Suppose that $h_n \rightarrow h_0$ in $\sigma(L_1, L_\infty)$ but not in L_1 under Lebesgue measure m on $(0, 1)$, that is,

$$(3) \quad \limsup_n \int_0^1 |h_n(t) - h_0(t)| dt > 0, \\ \lim_n \int_0^1 h_n(t) g(t) dt = \int_0^1 h_0(t) g(t) dt \quad \text{for all bounded Borel functions } g.$$

Take a sequence $(T_n : n \geq 0)$ of $(0, 1)$ -valued random variables, on a probability space (Θ, \mathcal{E}, Q) , such that each T_n has density h_n w.r.t. m . Define

$$Z_n = I_{[T_n, 1]} \quad \text{and} \quad \mu_n(A) = Q(Z_n \in A) \quad \text{for } A \in \mathcal{B}_d.$$

Then $Z_n = \phi(T_n)$, with $\phi : (0, 1) \rightarrow D$ given by $\phi(t) = I_{[t, 1]}$, $t \in (0, 1)$. Hence, for fixed $f \in L$, one obtains

$$\mu_n(f) = E_Q\{f \circ \phi(T_n)\} = \int_0^1 h_n(t) f \circ \phi(t) dt \longrightarrow \int_0^1 h_0(t) f \circ \phi(t) dt = \mu_0(f).$$

Suppose now that $X_n \sim \mu_n$ for all $n \geq 0$, where the X_n are D -valued random variables on some probability space (Ω, \mathcal{A}, P) . Since

$$P\{\omega : X_n(\omega)(t) \in \{0, 1\} \text{ for all } t\} = Q\{\theta : Z_n(\theta)(t) \in \{0, 1\} \text{ for all } t\} = 1,$$

it follows that

$$P(\|X_n - X_0\| > \frac{1}{2}) = P(X_n \neq X_0) \geq d_{TV}(\mu_n, \mu_0) = \frac{1}{2} \int_0^1 |h_n(t) - h_0(t)| dt.$$

Therefore, X_n fails to converge to X_0 in probability.

A slight change in Example 6 shows that convergence in probability cannot be strengthened into a.s. convergence in condition (1). Precisely, it may be that (1) holds, and yet there are not D -valued random variables Y_n satisfying $Y_n \sim \mu_n$ for all n and $\|Y_n - Y_0\| \xrightarrow{a.s.} 0$.

Example 7. In the notation of Example 6, instead of (3) assume

$$\lim_n \int_0^1 |h_n(t) - h_0(t)| dt = 0 \quad \text{and} \quad m(\liminf_n h_n < h_0) > 0$$

where m is Lebesgue measure on $(0, 1)$. Since

$$d_{TV}(\mu_n, \mu_0) = \frac{1}{2} \int_0^1 |h_n(t) - h_0(t)| dt \longrightarrow 0,$$

condition (1) trivially holds by Proposition 1. Suppose now that $Y_n \sim \mu_n$ for all $n \geq 0$, where the Y_n are D -valued random variables on a probability space (Ω, \mathcal{A}, P) . As $m(\liminf_n h_n < h_0) > 0$, Theorem 3.1 of [8] yields $P(Y_n = Y_0 \text{ ultimately}) < 1$. On the other hand, since $P(Y_n(t) \in \{0, 1\} \text{ for all } t) = 1$, one obtains

$$P(\|Y_n - Y_0\| \longrightarrow 0) = P(Y_n = Y_0 \text{ ultimately}) < 1.$$

4. APPLICATIONS

Condition (2) is not always hard to be checked, even if μ_0 is not u -separable. We illustrate this fact by two examples. To this end, we first note that conditions (1)-(2) are preserved under certain mixtures.

Corollary 8. *Let G be the set of distribution functions on $[0, 1]$ and \mathcal{G} the σ -field on G generated by the maps $g \mapsto g(t)$, $0 \leq t \leq 1$. Let π be a probability on \mathcal{G} and μ_n and λ_n probabilities on \mathcal{B}_d . Then, condition (1) holds provided*

$$\sup_{f \in L} |\lambda_n(f) - \lambda_0(f)| \longrightarrow 0 \quad \text{and}$$

$$\mu_n(A) = \int \lambda_n\{x : x \circ g \in A\} \pi(dg) \quad \text{for all } n \geq 0 \text{ and } A \in \mathcal{B}_d.$$

Proof. By Theorem 4, there are a probability space (Θ, \mathcal{E}, Q) and measurable maps $Z_n : (\Theta, \mathcal{E}) \longrightarrow (D, \mathcal{B}_d)$ such that $Z_n \sim \lambda_n$ for all n and $\|Z_n - Z_0\| \xrightarrow{Q} 0$. Define $\Omega = \Theta \times G$, $\mathcal{A} = \mathcal{E} \otimes \mathcal{G}$, $P = Q \times \pi$, and $X_n(\theta, g) = Z_n(\theta) \circ g$ for all $(\theta, g) \in \Theta \times G$. It is routine to check that $X_n \sim \mu_n$ for all n and $\|X_n - X_0\| \xrightarrow{P} 0$. \square

Example 9. (Exchangeable empirical processes). Let $(\xi_n : n \geq 1)$ be a sequence of $[0, 1]$ -valued random variables on the probability space $(\Omega_0, \mathcal{A}_0, P_0)$. Suppose (ξ_n) exchangeable and define

$$F(t) = E_{P_0}(I_{\{\xi_1 \leq t\}} \mid \tau)$$

where τ is the tail σ -field of (ξ_n) . Take F to be regular, i.e., each F -path is a distribution function. Then, the n -th empirical process can be defined as

$$Z_n(t) = \frac{\sum_{i=1}^n \{I_{\{\xi_i \leq t\}} - F(t)\}}{\sqrt{n}}, \quad 0 \leq t \leq 1, n \geq 1.$$

Since $Z_n : (\Omega_0, \mathcal{A}_0) \rightarrow (D, \mathcal{B}_d)$ is measurable, one can define $\mu_n(\cdot) = P_0(Z_n \in \cdot)$. Also, let μ_0 be the probability distribution of

$$Z_0(t) = B_M(t)$$

where B is a standard Brownian bridge on $[0, 1]$ and M an independent copy of F (with B and M defined on some probability space). Then, $\mu_n \rightarrow \mu_0$ weakly (under d) but μ_0 can fail to admit any extension to \mathcal{B}_u ; see [3] and Example 11 of [1]. Thus, Z_n can fail to converge in distribution, under u , according to Hoffmann-Jørgensen's definition. However, Corollaries 5 and 8 grant that:

On $(\Omega_0, \mathcal{A}_0, P_0)$, there are measurable maps $X_n : (\Omega_0, \mathcal{A}_0) \rightarrow (D, \mathcal{B}_d)$ such that $X_n \sim Z_n$ for each $n \geq 0$ and $\|X_n - X_0\| \xrightarrow{P_0} 0$.

Define in fact $B_n(t) = n^{-1/2} \sum_{i=1}^n \{I_{\{u_i \leq t\}} - t\}$, where u_1, u_2, \dots are i.i.d. random variables (on some probability space) with uniform distribution on $[0, 1]$. Then, $B_n \rightarrow B$ in distribution, under u , according to Hoffmann-Jørgensen's definition. Let λ_n and λ_0 be the probability distributions of B_n and B , respectively. Since λ_0 is u -separable, $\sup_{f \in L} |\lambda_n(f) - \lambda_0(f)| \longrightarrow 0$ (see Section 1). Thus, the first condition of Corollary 8 holds. The second condition follows from de Finetti's representation theorem, by letting $\pi(A) = P_0(F \in A)$ for $A \in \mathcal{G}$. Hence, condition (1) holds.

It remains to see that the X_n can be defined on $(\Omega_0, \mathcal{A}_0, P_0)$. To this end, it can be assumed $\mathcal{A}_0 = \sigma(\xi_1, \xi_2, \dots)$. If P_0 is nonatomic, it suffices to apply Corollary 5. Suppose P_0 has an atom A . Since $\mathcal{A}_0 = \sigma(\xi_1, \xi_2, \dots)$, up to P_0 -null sets, A is of the

form $A = \{\xi_n = t_n \text{ for all } n \geq 1\}$ for some constants t_n . Let $\sigma = (\sigma_1, \sigma_2, \dots)$ be a permutation of $1, 2, \dots$ and $A_\sigma = \{\xi_n = t_{\sigma_n} \text{ for all } n \geq 1\}$. By exchangeability,

$$P_0(A_\sigma) = P_0(A) > 0 \quad \text{for all permutations } \sigma,$$

and this implies $t_n = t_1$ for all $n \geq 1$. Let H be the union of all P_0 -atoms. Up to P_0 -null sets, one obtains

$$H \subset \{\xi_n = \xi_1 \text{ for all } n \geq 1\} \subset \{Z_n = 0 \text{ for all } n \geq 1\}.$$

If $P_0(H) = 1$, thus, it suffices to let $X_n = 0$ for all $n \geq 0$. If $0 < P_0(H) < 1$, since $P_0(\cdot \mid H^c)$ is nonatomic and (ξ_n) is still exchangeable under $P_0(\cdot \mid H^c)$, it is not hard to define the X_n on $(\Omega_0, \mathcal{A}_0, P_0)$ in such a way that $X_n \sim Z_n$ for all $n \geq 0$ and $\|X_n - X_0\| \xrightarrow{P_0} 0$.

Example 10. (Pure jump processes). For each $n \geq 0$, let

$$C_n = (C_{n,j} : j \geq 1) \quad \text{and} \quad Y_n = (Y_{n,j} : j \geq 1)$$

be sequences of real random variables, defined on the probability space $(\Omega_0, \mathcal{A}_0, P_0)$, such that

$$0 \leq Y_{n,j} \leq 1 \quad \text{and} \quad \sum_{j=1}^{\infty} |C_{n,j}| < \infty.$$

Define

$$Z_n(t) = \sum_{j=1}^{\infty} C_{n,j} I_{\{Y_{n,j} \leq t\}}, \quad 0 \leq t \leq 1, n \geq 0.$$

Since $Z_n : (\Omega_0, \mathcal{A}_0) \rightarrow (D, \mathcal{B}_d)$ is measurable, one can define $\mu_n(\cdot) = P_0(Z_n \in \cdot)$. Then, condition (1) holds provided

C_n is independent of Y_n for every $n \geq 0$,

$$\sum_{j=1}^{\infty} |C_{n,j} - C_{0,j}| \xrightarrow{P_0} 0 \quad \text{and} \quad d_{TV}(\nu_{n,k}, \nu_{0,k}) \rightarrow 0 \text{ for all } k \geq 1,$$

where $\nu_{n,k}$ denotes the probability distribution of $(Y_{n,1}, \dots, Y_{n,k})$.

For instance, $\nu_{n,k} = \nu_{0,k}$ for all n and k in case $Y_{n,j} = V_{n+j}$ with V_1, V_2, \dots a stationary sequence. Also, independence between C_n and Y_n can be replaced by

$$\sigma(C_{n,j}) \subset \sigma(Y_{n,1}, \dots, Y_{n,j}) \quad \text{for all } n \geq 0 \text{ and } j \geq 1.$$

To prove (1), define $Z_{n,k}(t) = \sum_{j=1}^k C_{n,j} I_{\{Y_{n,j} \leq t\}}$. For each $f \in L$,

$$\begin{aligned} |\mu_n(f) - \mu_0(f)| &\leq |Ef(Z_n) - Ef(Z_{n,k})| + |Ef(Z_{n,k}) - Ef(Z_{0,k})| + |Ef(Z_{0,k}) - Ef(Z_0)| \\ &\leq E\{2 \wedge \|Z_n - Z_{n,k}\|\} + |Ef(Z_{n,k}) - Ef(Z_{0,k})| + E\{2 \wedge \|Z_0 - Z_{0,k}\|\} \\ &\leq E\{2 \wedge \sum_{j>k} |C_{n,j}|\} + |Ef(Z_{n,k}) - Ef(Z_{0,k})| + E\{2 \wedge \sum_{j>k} |C_{0,j}|\} \end{aligned}$$

where $E(\cdot) = E_{P_0}(\cdot)$. Given $\epsilon > 0$, take $k \geq 1$ such that $E\{2 \wedge \sum_{j>k} |C_{0,j}|\} < \epsilon$. Then,

$$\limsup_n \sup_{f \in L} |\mu_n(f) - \mu_0(f)| < 2\epsilon + \limsup_n \sup_{f \in L} |Ef(Z_{n,k}) - Ef(Z_{0,k})|.$$

It remains to show that $\sup_{f \in L} |Ef(Z_{n,k}) - Ef(Z_{0,k})| \rightarrow 0$. Since C_n is independent of Y_n , up to changing $(\Omega_0, \mathcal{A}_0, P_0)$ with some other probability space, it can be assumed

$$P_0(Y_{n,j} \neq Y_{0,j} \text{ for some } j \leq k) = d_{TV}(\nu_{n,k}, \nu_{0,k});$$

see Proposition 1. The same is true if $\sigma(C_{n,j}) \subset \sigma(Y_{n,1}, \dots, Y_{n,j})$ for all n and j . Then, letting $A_{n,k} = \{Y_{n,j} = Y_{0,j} \text{ for all } j \leq k\}$, one obtains

$$\begin{aligned} \sup_{f \in L} |Ef(Z_{n,k}) - Ef(Z_{0,k})| &\leq E\{I_{A_{n,k}} 2 \wedge \|Z_{n,k} - Z_{0,k}\|\} + 2P_0(A_{n,k}^c) \\ &\leq E\{2 \wedge \sum_{j=1}^{\infty} |C_{n,j} - C_{0,j}|\} + 2d_{TV}(\nu_{n,k}, \nu_{0,k}) \rightarrow 0. \end{aligned}$$

Thus, condition (2) holds, and an application of Theorem 4 concludes the proof.

APPENDIX

Three preliminary lemmas are needed to prove Theorem 4. The first is part of the folklore about Skorohod distance, and we state it without a proof. Let $\Delta x(t) = x(t) - x(t-)$ denote the jump of $x \in D$ at $t \in (0, 1]$.

Lemma 11. *Fix $\epsilon > 0$ and $x_n \in D$, $n \geq 0$. Then, $\limsup_n \|x_n - x_0\| \leq \epsilon$ whenever $d(x_n, x_0) \rightarrow 0$ and*

$$|\Delta x_n(t)| > \epsilon \quad \text{for all large } n \text{ and } t \in (0, 1) \text{ such that } |\Delta x_0(t)| > \epsilon.$$

The second lemma is a consequence of Remark 6 of [5], but we give a sketch of its proof as it is basic for Theorem 4. Let μ, ν be laws on \mathcal{B}_d and $\mathcal{F}(\mu, \nu)$ the class of probabilities λ on $\mathcal{B}_d \otimes \mathcal{B}_d$ such that $\lambda(\cdot \times D) = \mu(\cdot)$ and $\lambda(D \times \cdot) = \nu(\cdot)$. Since the map $(x, y) \mapsto \|x - y\|$ is measurable w.r.t. $\mathcal{B}_d \otimes \mathcal{B}_d$, one can define

$$W_u(\mu, \nu) = \inf_{\lambda \in \mathcal{F}(\mu, \nu)} \int 1 \wedge \|x - y\| \lambda(dx, dy).$$

Lemma 12. *For a sequence $(\mu_n : n \geq 0)$ of probabilities on \mathcal{B}_d , condition (1) holds if and only if $W_u(\mu_0, \mu_n) \rightarrow 0$.*

Proof. The "only if" part is trivial. Suppose $W_u(\mu_0, \mu_n) \rightarrow 0$. Let $\Omega = D^\infty$, $\mathcal{A} = \mathcal{B}_d^\infty$ and $X_n : D^\infty \rightarrow D$ the n -th canonical projection, $n \geq 0$. Take $\lambda_n \in \mathcal{F}(\mu_0, \mu_n)$ such that $\int 1 \wedge \|x - y\| \lambda_n(dx, dy) < \frac{1}{n} + W_u(\mu_0, \mu_n)$. Since (D, d) is Polish, λ_n admits a disintegration $\alpha_n = \{\alpha_n(x) : x \in D\}$ (see Section 2). By Ionescu-Tulcea theorem, there is a unique probability P on \mathcal{B}_d^∞ such that $X_0 \sim \mu_0$ and

$$\beta_n(x_0, x_1, \dots, x_{n-1})(A) = \alpha_n(x_0)(A), \quad (x_0, x_1, \dots, x_{n-1}) \in D^n, A \in \mathcal{B}_d,$$

is a regular version of the conditional distribution of X_n given $(X_0, X_1, \dots, X_{n-1})$ for all $n \geq 1$. Under such P , one obtains $(X_0, X_n) \sim \lambda_n$ (so that $X_n \sim \mu_n$) and

$$\epsilon P(\|X_0 - X_n\| > \epsilon) \leq E_P\{1 \wedge \|X_0 - X_n\|\} < \frac{1}{n} + W_u(\mu_0, \mu_n) \rightarrow 0 \quad \text{for all } \epsilon \in (0, 1).$$

□

The third lemma needs some more effort. Let $\phi_0(x, \epsilon) = 0$ and

$$\phi_{n+1}(x, \epsilon) = \inf\{t : \phi_n(x, \epsilon) < t \leq 1, |\Delta x(t)| > \epsilon\}$$

where $n \geq 0$, $\epsilon > 0$, $x \in D$ and $\inf \emptyset := 1$. The map $x \mapsto \phi_n(x, \epsilon)$ is universally measurable w.r.t. \mathcal{B}_d for all n and ϵ .

Lemma 13. *Let \mathcal{F}_k be the Borel σ -field on \mathbb{R}^k and $I \subset (0, 1)$ a dense subset. For a sequence $(\mu_n : n \geq 0)$ of probabilities on \mathcal{B}_d , condition (1) holds provided*

$$\sup_{A \in \mathcal{F}_k} \left| \int f(x) I_A(\phi_1(x, \epsilon), \dots, \phi_k(x, \epsilon)) \mu_n(dx) - \int f(x) I_A(\phi_1(x, \epsilon), \dots, \phi_k(x, \epsilon)) \mu_0(dx) \right| \longrightarrow 0$$

for each $k \geq 1$, $\epsilon \in I$ and function $f : D \rightarrow [-1, 1]$ such that $|f(x) - f(y)| \leq d(x, y)$ for all $x, y \in D$.

Proof. Fix $\epsilon \in I$ and write $\phi_n(x)$ instead of $\phi_n(x, \epsilon)$. As each ϕ_n is universally measurable w.r.t. \mathcal{B}_d , there is a set $T \in \mathcal{B}_d$ such that

$$\mu_n(T) = 1 \quad \text{and} \quad I_T \phi_n \text{ is } \mathcal{B}_d\text{-measurable for all } n \geq 0.$$

Thus, ϕ_n can be assumed \mathcal{B}_d -measurable for all n . Let k be such that

$$\mu_0\{x : \phi_r(x) \neq 1 \text{ for some } r > k\} < \epsilon.$$

For such a k , define $\phi(x) = (\phi_1(x), \dots, \phi_k(x))$, $x \in D$, and

$$\lambda_n(A) = \mu_n\{x : (\phi(x), x) \in A\}, \quad A \in \mathcal{F}_k \otimes \mathcal{B}_d.$$

Since (D, d) is Polish, Proposition 2 applies to such λ_n with $(F, \mathcal{F}) = (\mathbb{R}^k, \mathcal{F}_k)$ and $(G, \mathcal{G}) = (D, \mathcal{B}_d)$. Condition (b) holds by the assumption of the Lemma. Thus, by Proposition 2, on a probability space (Ω, \mathcal{A}, P) there are measurable maps $(Y_n, Z_n) : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^k \times D, \mathcal{F}_k \otimes \mathcal{B}_d)$ satisfying

$$(Y_n, Z_n) \sim \lambda_n \text{ for all } n \geq 0, \quad P(Y_n \neq Y_0) \longrightarrow 0, \quad d(Z_n, Z_0) \xrightarrow{P} 0.$$

Since $P(Y_n = \phi(Z_n)) = \lambda_n\{(\phi(x), x) : x \in D\} = 1$, one also obtains

$$(4) \quad \lim_n P(\phi(Z_n) = \phi(Z_0)) = 1.$$

Next, by (4) and $d(Z_n, Z_0) \xrightarrow{P} 0$, there is a subsequence (n_j) such that

$$\limsup_n P(\|Z_n - Z_0\| > \epsilon) = \lim_j P(\|Z_{n_j} - Z_0\| > \epsilon),$$

$$d(Z_{n_j}, Z_0) \xrightarrow{a.s.} 0, \quad P(\phi(Z_{n_j}) = \phi(Z_0) \text{ for all } j) > 1 - \epsilon.$$

Define $U = \limsup_j \|Z_{n_j} - Z_0\|$ and

$$H = \{\phi_r(Z_0) = 1 \text{ for all } r > k\} \cap \{\phi(Z_{n_j}) = \phi(Z_0) \text{ for all } j\} \cap \{d(Z_{n_j}, Z_0) \longrightarrow 0\}.$$

For each $\omega \in H$, Lemma 11 applies to $Z_0(\omega)$ and $Z_{n_j}(\omega)$, so that $U(\omega) \leq \epsilon$. Further,

$$\begin{aligned} P(H^c) &\leq P(\phi_r(Z_0) \neq 1 \text{ for some } r > k) + P(\phi(Z_{n_j}) \neq \phi(Z_0) \text{ for some } j) \\ &< \mu_0\{x : \phi_r(x) \neq 1 \text{ for some } r > k\} + \epsilon < 2\epsilon. \end{aligned}$$

Since $U \leq \epsilon$ on H ,

$$\begin{aligned} \limsup_n P(\|Z_n - Z_0\| > \epsilon) &= \lim_j P(\|Z_{n_j} - Z_0\| > \epsilon) \leq P(U \geq \epsilon) \\ &\leq P(U = \epsilon) + P(H^c) < P(U = \epsilon) + 2\epsilon. \end{aligned}$$

On noting that $E_P\{1 \wedge \|Z_0 - Z_n\|\} \leq \epsilon + P(\|Z_n - Z_0\| > \epsilon)$, one obtains

$$\limsup_n W_u(\mu_0, \mu_n) \leq \limsup_n E_P\{1 \wedge \|Z_0 - Z_n\|\} < P(U = \epsilon) + 3\epsilon.$$

Since I is dense in $(0, 1)$, then $P(U = \epsilon) + 3\epsilon$ can be made arbitrarily small for a suitable $\epsilon \in I$. Thus, $\limsup_n W_u(\mu_0, \mu_n) = 0$. An application of Lemma 12 concludes the proof. \square

We are now ready for the last attack to Theorem 4.

Proof of Theorem 4. "(1) \Rightarrow (2)". Just note that

$$\begin{aligned} |\mu_n(f) - \mu_0(f)| &= |E_P\{f(X_n)\} - E_P\{f(X_0)\}| \leq E_P|f(X_n) - f(X_0)| \\ &\leq E_P\{2 \wedge \|X_n - X_0\|\} \longrightarrow 0, \quad \text{for each } f \in L, \text{ under (1).} \end{aligned}$$

"(2) \Rightarrow (1)". Let $B_\epsilon = \{x : |\Delta x(t)| = \epsilon \text{ for some } t \in (0, 1]\}$. Then, B_ϵ is universally measurable w.r.t. \mathcal{B}_d and $\mu_0(B_\epsilon) > 0$ for at most countably many $\epsilon > 0$. Hence, $I = \{\epsilon \in (0, 1) : \mu_0(B_\epsilon) = 0\}$ is dense in $(0, 1)$.

Fix $\epsilon \in I$, $k \geq 1$, and a function $f : D \rightarrow [-1, 1]$ such that $|f(x) - f(y)| \leq d(x, y)$ for all $x, y \in D$. By Lemma 13, for condition (1) to be true, it is enough that

$$(5) \quad \lim_n \sup_{A \in \mathcal{F}_k} |\mu_n\{f I_A(\phi)\} - \mu_0\{f I_A(\phi)\}| = 0$$

where $\phi(x) = (\phi_1(x), \dots, \phi_k(x))$, $x \in D$, and $\phi_j(x) = \phi_j(x, \epsilon)$ for all j .

In order to prove (5), given $b \in (0, \frac{\epsilon}{2})$, define

$$F_b = \{x : |\Delta x(t)| \notin (\epsilon - 2b, \epsilon + 2b) \text{ for all } t \in (0, 1]\}, \quad G_b = \{x : d(x, F_b) \geq \frac{b}{2}\}.$$

Then,

$$(i) \ G_b^c \subset F_{b/2}; \quad (ii) \ \phi(x) = \phi(y) \text{ whenever } x, y \in F_b \text{ and } \|x - y\| < b.$$

Statement (ii) is straightforward. To check (i), fix $x \notin G_b$ and take $y \in F_b$ with $d(x, y) < b/2$. Let $\gamma : [0, 1] \rightarrow [0, 1]$ be a strictly increasing homeomorphism such that $\|x - y \circ \gamma\| < b/2$. For all $t \in (0, 1]$,

$$|\Delta x(t)| \leq |\Delta y \circ \gamma(t)| + 2\|x - y \circ \gamma\| < |\Delta y(\gamma(t))| + b.$$

Similarly, $|\Delta x(t)| > |\Delta y(\gamma(t))| - b$. Since $y \in F_b$, it follows that $x \in F_{b/2}$.

Next, define

$$\psi_b(x) = \frac{d(x, G_b)}{d(x, F_b) + d(x, G_b)}, \quad x \in D.$$

Then, $\psi_b = 0$ on G_b and ψ_b is Lipschitz w.r.t. d with Lipschitz constant $2/b$. Hence, ψ_b is Lipschitz w.r.t. u with Lipschitz constant $2/b$ (since $d \leq u$). Basing on (i)-(ii) and such properties of ψ_b , it is not hard to check that $\psi_b I_A(\phi)$ is Lipschitz w.r.t. u , with Lipschitz constant $2/b$, for every $A \in \mathcal{F}_k$. In turn, since $d \leq u$ and f is Lipschitz w.r.t. d with Lipschitz constant 1,

$$f_A = f \psi_b I_A(\phi), \quad A \in \mathcal{F}_k,$$

is Lipschitz w.r.t. u with Lipschitz constant $(1 + 2/b)$. Moreover,

$$|\mu_n\{f I_A(\phi)\} - \mu_n(f_A)| \leq \mu_n|f I_A(\phi)(1 - \psi_b)| \leq \mu_n(1 - \psi_b).$$

On noting that $(1 + 2/b)^{-1} f_A \in L$ for every $A \in \mathcal{F}_k$, condition (2) yields

$$\begin{aligned} &\limsup_n \sup_{A \in \mathcal{F}_k} |\mu_n\{f I_A(\phi)\} - \mu_0\{f I_A(\phi)\}| \\ &\leq \limsup_n \{\mu_n(1 - \psi_b) + \sup_{A \in \mathcal{F}_k} |\mu_n(f_A) - \mu_0(f_A)| + \mu_0(1 - \psi_b)\} \\ &= 2\mu_0(1 - \psi_b) \leq 2\mu_0(F_b^c). \end{aligned}$$

Since $\epsilon \in I$ and $\bigcap_{b>0} F_b^c = \{x : |\Delta x(t)| = \epsilon \text{ for some } t\} = B_\epsilon$, one obtains

$$\limsup_n \sup_{A \in \mathcal{F}_k} |\mu_n\{f I_A(\phi)\} - \mu_0\{f I_A(\phi)\}| \leq 2 \lim_{b \rightarrow 0} \mu_0(F_b^c) = 2 \mu_0(B_\epsilon) = 0.$$

Therefore, condition (5) holds and this concludes the proof. \square

REFERENCES

- [1] Berti P., Rigo P. (2004) Convergence in distribution of non measurable random elements, *Ann. Probab.*, 32, 365-379.
- [2] Berti P., Pratelli L., Rigo P. (2006) Almost sure weak convergence of random probability measures, *Stochastics*, 78, 91-97.
- [3] Berti P., Pratelli L., Rigo P. (2006) Asymptotic behaviour of the empirical process for exchangeable data, *Stoch. Proc. Appl.*, 116, 337-344.
- [4] Berti P., Pratelli L., Rigo P. (2007) Skorohod representation on a given probability space, *Prob. Theo. Rel. Fields*, 137, 277-288.
- [5] Berti P., Pratelli L., Rigo P. (2009) Skorohod representation theorem via disintegrations, *Submitted*, currently available at: http://economia.unipv.it/pagp/pagine_personali/prigo/skorsankhya.pdf
- [6] Blackwell D., Dubins L.E. (1983) An extension of Skorohod's almost sure representation theorem, *Proc. Amer. Math. Soc.*, 89, 691-692.
- [7] Dudley R.M. (1999) *Uniform central limit theorems*, Cambridge University Press.
- [8] Sethuraman J. (2002) Some extensions of the Skorohod representation theorem, *Sankhya*, 64, 884-893.
- [9] van der Vaart A., Wellner J.A. (1996) *Weak convergence and empirical processes*, Springer.

PATRIZIA BERTI, DIPARTIMENTO DI MATEMATICA PURA ED APPLICATA "G. VITALI", UNIVERSITA' DI MODENA E REGGIO-EMILIA, VIA CAMPI 213/B, 41100 MODENA, ITALY
E-mail address: berti.patrizia@unimore.it

LUCA PRATELLI, ACCADEMIA NAVALE, VIALE ITALIA 72, 57100 LIVORNO, ITALY
E-mail address: pratell@mail.dm.unipi.it

PIETRO RIGO (CORRESPONDING AUTHOR), DIPARTIMENTO DI ECONOMIA POLITICA E METODI QUANTITATIVI, UNIVERSITA' DI PAVIA, VIA S. FELICE 5, 27100 PAVIA, ITALY
E-mail address: prigo@eco.unipv.it