A REMARK ON THE WEAK CONVERGENCE OF PROCESSES IN THE SKOROHOD TOPOLOGY

by

Jean Jacod and Philip Protter
Université de Paris VI — Purdue University

Technical Report #92-06

Department of Statistics
Purdue University

January 1992

A REMARK ON THE WEAK CONVERGENCE OF PROCESSES IN THE SKOROHOD TOPOLOGY

by

Jean Jacod
Laboratoire de Probabilités
Université de Paris VI
4, Place Jussieu-Tour 56
75230 Paris Cedex 05
FRANCE

and

Philip Protter[1]
Departments of Mathematics and Statistics
Purdue University
West Lafayette, IN 47907-1395
USA

Abstract

For processes converging weakly in the Skorohod topology, it is shown that there exist random time change processes such that the time changed processes converge locally uniformly to the limit, and that the time changed processes can be taken to be *jointly measurable*, and even adapted to an appropriate filtration.

[1]Supported in part by NSF grant #DMS-9103454

1. INTRODUCTION

A wide class of stochastic processes of interest have paths which are right continuous with left limits ("càdlàg" is the French acronym). A natural way to study their convergence is by using the Skorohod topology, which by using changes of time allows one to "move" the jumps of the approximants to the times of the jumps of the limit process. The description of the Skorohod topology is natural and intuitive, but when one applies it to weak convergence of stochastic processes, its appeal breaks down: typically one says that one needs

$$\lim_{n\to\infty} E\{F(X_n)\} = E\{F(X)\},$$

for every bounded, Skorohod continuous function F. But what is a Skorohod continuous function? What are examples of functions that are continuous in the sup norm that are not also continuous for the Skorohod topology? In this note we attempt to remedy this problem by showing that if $X_n = (X_n(t))_{t\geq 0}$ are processes converging weakly to a process X in the Skorohod topology, then there is a random sequence of time changes $\Lambda_n(t,\omega)$ such that $X_n \circ \Lambda_n$ converges locally uniformly to X. The key feature here is that we can take the time change processes Λ_n *jointly measurable* in (t,ω). We also show that the time change processes Λ_n can be taken adapted to an appropriate filtration. It is perhaps surprising that this "obvious" result does not exist in the literature; while the first half is mostly measure theory and the proof we give for the first statement is short, nevertheless it uses the very deep "section theorem" of P. A. Meyer and thus it is ultimately a non-trivial theorem.

Let L denote the class of functions which are increasing and bijective fom $\mathbf{R}_+$ to $\mathbf{R}_+$. Let $\mathbf{D} = \mathbf{D}(\mathbf{R}_+, E)$ denote the space of càdlàg functions mapping $\mathbf{R}_+$ into E, where E is a given Polish space with a distance δ. A sequence $x_n = (x_n(t))_{t\geq 0}$ in $\mathbf{D}$ converges in the *Skorohod topology* to $x \in \mathbf{D}$ if and only if there exist $\lambda_n \in L$ converging uniformly to the identity such that $x_n \circ \lambda_n$ converge locally uniformly to x. This Skorohod convergence can be described using a distance. For example a compatible distance d is given by, for $x, y \in \mathbf{D}$:

$$(1.1)\qquad \begin{cases} d(x,y) &= \inf_{\lambda\in L} d_\lambda(x,y), \qquad \text{where} \\ d_\lambda(x,y) &= \sup_{t\geq 0}|\lambda(t) - t\,| + \sum_{p\in\mathbf{N}} 2^{-p}\min\{1, \sup_{t\leq p} \delta(x\circ\lambda(t), y(t))\}. \end{cases}$$

By "convergence in $\mathbf{D}$*", we will always mean convergence in the Skorohod topology.*

We now turn to stochastic processes. Let X_n, X be E-valued stochastic processes, and suppose that one of the following assumptions holds:

$$(1.2)\qquad X_n \quad \text{converges to } X \text{ a.s. in } \mathbf{D};$$

$$(1.3)\qquad X_n \quad \text{converges to } X \text{ weakly in } \mathbf{D}.$$

Note that for weak convergence ((1.3)), the processes X_n and X can be defined on different probability spaces: $(\Omega_n, \mathcal{F}_n, P_n)$ and $(\Omega, \mathcal{F}, P)$, respectively.

It is now natural to ask if there exist stochastic processes Λ_n which are changes of time such that $d_{\Lambda_n}(X_n, X)$ converges to 0 in an appropriate sense.

For example assume that (1.2) holds. Then for each ω such that $X_n(\omega)$ converges to $X(\omega)$, there exists $\Lambda_n(\omega) \in L$ such that $d_{\Lambda_n(\omega)}(X_n(\omega), X(\omega))$ converges to 0; but it is not *a priori* clear that one can choose measurable versions of the Λ_n that would therefore be true stochastic processes. And it is even less clear that one can choose measurable Λ_n that have good non-anticipating properties. If for example the processes X_n and X are adapted to an underlying filtration $\mathsf{F} = (\mathcal{F}_t)_{t\geq 0}$, one cannot hope to have the Λ_n also adapted to F; but since the Λ_n converge to $\Lambda(t) \equiv t$, one can perhaps choose the Λ_n such that they are adapted to filtrations $\mathsf{F}^n = (\mathcal{F}_{t+\gamma_n})_{t\geq 0}$, where γ_n is a sequence decreasing to 0. In section three we construct such Λ_n.

For background information on weak convergence one can consult any of Billingsley [**1**], Ethier-Kurtz [**3**], Jacod-Shiryaev [**4**], or Pollard [**5**].

2. TIME CHANGES AS STOCHASTIC PROCESSES

In this section we show that if stochastic processes $X_n = (X_n(t))_{t\geq 0}$ converge to X in the Skorohod topology, then there exist time change stochastic processes Λ_n such that $X_n \circ \Lambda_n$ converges to X locally uniformly. (We use the notation established in section one).

(2.1) THEOREM. *Suppose stochastic processes X_n converge a.s. in the Skorohod topology in D to X. Then there exists a sequence $(\Lambda_n)_{n\geq 1}$ of (measurable) processes with paths in L such that* $\lim_{n\to\infty} d_{\Lambda_n}(X_n, X) = 0$ a.s.

Proof: Let

$$U_n = \{(\omega, \lambda) \in \Omega \times L : d_\lambda(X_n(\omega), X(\omega)) \leq d(X_n(\omega), X(\omega)) + 2^{-n}\}.$$

One easily checks that $U_n \in \mathcal{F} \otimes \mathcal{L}$, where $\mathcal{L}$ is the Borel σ-field of L (under the uniform topology). Moreover the projection $\pi(U_n) = \{\omega : \exists \lambda \in L \text{ with } (\omega, \lambda) \in U_n\}$ is equal to all of Ω. By the measurable section theorem (see, e.g., [2, p.18, Theorem T37]), there exists a random variable Λ_n with values in $(L, \mathcal{L})$ such that $P(\omega : (\omega, \Lambda_n(\omega)) \in U_n) = 1$. Since $d(x,y) = \inf_{\lambda \in L} d_\lambda(x,y)$, we have $d_{\Lambda_n}(X_n, X) \leq d(X_n, X) + 2^{-n}$ a.s., whence the result. ∎

(2.2) Remark: Because we have used the section theorem, we cannot omit the "almost surely" in the conclusion of Theorem (2.1), even if $X_n(\omega)$ converges in the Skorohod topology to $X(\omega)$ for all $\omega \in \Omega$.

(2.3) THEOREM. *Let X_n be stochastic processes such that X_n converges weakly in* $\mathbf{D}$ *to X. Then there exists an auxiliary space $(\Omega', \mathcal{F}', P')$, and processes X'_n, X' defined on this space such that $\mathcal{L}(X'_n) = \mathcal{L}(X_n)$, $\mathcal{L}(X') = \mathcal{L}(X)$, and also processes Λ'_n with paths in L such that $\lim_{n\to\infty} d_{\Lambda'_n}(X'_n, X') = 0$ a.s. (dP').*

Proof: $\mathcal{L}(X') = \mathcal{L}(X)$ means that the two processes X' and X have the same distribution. Applying the Skorohod representation theorem, we can find a space $(\Omega', \mathcal{F}', P')$ on which there exist X'_n, X' with $\mathcal{L}(X'_n) = \mathcal{L}(X_n)$ and $\mathcal{L}(X') = \mathcal{L}(X)$, and also X'_n converges a.s. to X' (dP'). Thus we need only to apply Theorem (2.1). ∎

(2.4) THEOREM. *Let $(X_n)_{n\geq 1}$, X be stochastic processes such that X_n converges weakly in* $\mathbf{D}$ *to X. Then for each n one can construct an extension $(\Omega'', \mathcal{F}''_n, P''_n)$ of $(\Omega_n, \mathcal{F}_n, P_n)$ on which there exist a process $Y_n = (Y_n(t))_{t\geq 0}$ with $\mathcal{L}(Y_n) = \mathcal{L}(X)$, and a process Λ''_n with paths in L, such that $d_{\Lambda''_n}(X_n, Y_n)$ converges in distribution to zero.*

Proof: One can easily construct an extension $(\Omega''_n, \mathcal{F}''_n, P''_n)$ of $(\Omega_n, \mathcal{F}_n, P_n)$ and two processes (Y_n, Λ''_n) with paths respectively in $\mathbf{D}$ and L, such that, with the notation of Theorem (2.3),

$$\mathcal{L}(X_n, Y_n, \Lambda''_n) = \mathcal{L}(X'_n, X', \Lambda'_n).$$

In particular $\mathcal{L}(Y_n) = \mathcal{L}(X)$, and $\mathcal{L}(d_{\Lambda''_n}(X_n, Y_n)) = \mathcal{L}(d_{\Lambda'_n}(X'_n, X'))$. However by Theorem (2.3), $d_{\Lambda'_n}(X'_n, X')$ converges a.s. to 0, and we have the result. ∎

3. TIME CHANGES AS ADAPTED STOCHASTIC PROCESSES

In Section 2 we showed that if X_n converges to X in the Skorohod topology, then there exist random time changes Λ_n that *are stochastic processes* such that $X_n \circ \Lambda_n$ converges to X in the appropriate sense. We did not discuss however whether or not the processes Λ_n could be chosen to be adapted to an appropriate underlying filtration of σ-algebras. If for example the processes X_n, X are defined on the same space, and letting (for $t \geq 0$), $\mathcal{F}_n(t) = \sigma(X_n(s), X(s); s \leq t)$ denote the natural filtration, one cannot hope to have $\Lambda_n(t) \in \mathcal{F}_n(t)$, each $t \geq 0$, in general. (After all, the processes Λ_n are "changes of time"!). However since $\Lambda_n(t)$ converges uniformly to $\Lambda(t) = t$, we will show that it is possible to choose the Λ_n such that they are adapted to $\mathcal{G}_n(t) = \mathcal{F}_n(t + \gamma_n)$, $t \geq 0$, where γ_n is a sequence of constants decreasing to 0 as n tends to ∞.

(3.1) THEOREM. *Suppose stochastic processes X_n converge a.s. to X in* $\mathbf{D}$*, and let $\mathcal{F}_n(t) = \sigma(X_n(s), X(s), s \leq t)$. Then there exist a sequence of constants γ_n decreasing to zero, and processes Λ_n with paths in L such that Λ_n is adapted to $\mathcal{G}_n$, where $\mathcal{G}_n(t) = \mathcal{F}_n(t + \gamma_n)$, and moreover $\lim_{n\to\infty} d_{\Lambda_n}(X_n, X) = 0$ a.s.*

Proof: Choose and fix an integer p, and $\alpha > 0$, $\beta > 0$. Let

$$N(\alpha) = \{\omega : \delta(X(t), X(t-)) = \alpha, \text{ for at least one } t > 0\}.$$

Let T^i (respectively T_n^i), for $i \geq 1$, be the successive jump times of X (resp. X_n) where the size of the jump exceeds α. Let $T^0 = T_n^0 = 0$. Let

$$\begin{aligned} I(n) &= I(n,\alpha,\beta,p) \equiv \inf\{i : T^i > p \text{ or } |T^i - T_n^i| > \beta\} \\ S_n^i &= (T^i \wedge T_n^i - \beta)^+ \\ \hat{S}_n^i &= T^i \vee T_n^i + \beta. \end{aligned}$$

We now define a process $\Lambda_n = \Lambda_n(\alpha,\beta,p)$ with paths in L as follows: Λ_n is piecewise linear, with discontinuities of its derivative at the following times only: the times T_n^i for $1 \leq i < I(n)$; times $\hat{S}_n^{i-1}$ and S_n^i if $\hat{S}_n^{i-1} \leq S_n^i$ and $1 \leq i < I(n)$; and $\hat{S}_n^{I(n)-1}$. Therefore it suffices to give the values of Λ_n at these points, to know $\Lambda_n(T_n^i) = T^i$ for $1 \leq i < I(n)$, and $\Lambda_n(\hat{S}_n^{i-1}) = \hat{S}_n^{i-1}$ and $\Lambda_n(S_n^i) = S_n^i$ if $\hat{S}_n^{i-1} \leq S_n^i$ for $1 \leq i < I(n)$, and finally to set $\Lambda_n(t) = t$ if $t \geq \hat{S}_n^{I(n)-1}$.

Since T^i and T_n^i are $(\mathcal{F}_n(t))_{t\geq 0}$ stopping times, we clearly have

$$\Lambda_n(\alpha,\beta,p;t) \text{ is } \mathcal{F}_n(t+2\beta) - \text{measurable,} \tag{3.2}$$

whereas by construction we also have

$$\sup_{t\geq 0}|\Lambda_n(\alpha,\beta,p;t) - t| \leq \sup_i(|T_n^i - T^i|; 1 \leq i \leq I(n)) \leq \beta). \tag{3.3}$$

Last, we note that if

$$\begin{aligned} V_n(\alpha,\beta,p) = \sup_{i\geq 1 \text{ and } T^{i-1}\leq p} \Big\{ & \delta(X(T^i), X_n(T_n^i)) \\ & + \sup_{s,t:|s-t|\leq\beta; T^{i-1}\leq s<T^i\wedge(p+1); T_n^{i-1}\leq t<T_n^i} \delta(X(s), X_n(t)) \Big\}, \end{aligned}$$

(with the convention $\delta(X(T^i), X_n(T_n^i)) = 0$ if $T^i = \infty$ or $T_n^i = \infty$) we deduce from (3.3) and the distance (1.1) that

$$d_{\Lambda_n(\alpha,\beta,p)}(X, X_n) \leq \beta + 2^{-(p-1)} + V_n(\alpha,\beta,p) \text{ on } \{T^{I(n)} > p\}. \tag{3.4}$$

Next we define a "β-modulus of continuity" of X on the interval $[0, p+1]$, without the jumps of size greater than α, as follows:

$$c(\alpha,\beta,p) = \sup_{i\geq 1} \sup\{\delta(X(t), X(s)) : |t-s| \leq \beta, s \leq p+1, T^{i-1} \leq s < t \leq T^i\}.$$

Then $c(\alpha, \beta, p)$ above is a random variable, and moreover

(3.5) $$\limsup_{\beta \to 0} c(\alpha, \beta, p) \leq \alpha.$$

Using the assumption of a.s. convergence of $X_n \to X$ and classical properties of the Skorohod topology (cf., eg, any of [1], [3], [4], or [5]), we have outside a null set:

(3.6) $T_n^i \to T^i$ for all $i \geq 1$,

(3.7) $\limsup_{n \to \infty} V_n(\alpha, \beta, p) \leq c(\alpha, \beta, p)$,

and (3.6) yields, again outside a null set:

(3.8) $T^{I(n)} > p$ for all n large enough (depending on ω).

Now, for each p let $\alpha_p \in (0, 2^{-p}]$ with $P(N(\alpha_p)) = 0$ (such α_p's always exist). By (3.5) there is β_p in $(0, 2^{-p}]$ with $\beta_{p+1} \leq \beta_p$ and

$$P(c(\alpha_p, \beta_p, p) > 2^{-(p-1)}) \leq 2^{-p}.$$

Then (3.7) and (3.8) yield a strictly increasing sequence of integers n_p such that

(3.9) $P\{V_n(\alpha_p, \beta_p, p) \leq 2^{-(p-2)}$ and $T^{I(n)} > p$ for all $n \geq n_p\} \geq 1 - 2^{-(p-1)}$.

Therefore, if $A_p = \{d_{\Lambda_n(\alpha_p, \beta_p, p)}(X, X_n) \leq 2^{-(p-3)}$ for all $n \geq n_p\}$ we deduce from (3.4) and (3.9) that $P(A_p) \geq 1 - 2^{-(p-1)}$. Then the Borel-Cantelli Lemma gives

(3.10) $$P(A) = 1, \text{ where } A = \limsup_{p \to \infty} A_p.$$

Our final step is to set $q_n = sup(p : n_p < n)$ and $\Lambda_n = \Lambda_n(\alpha_{q_n}, \beta_{q_n}, q_n)$ and $\gamma_n = 2\beta_{q_n}$. By (3.2) we have that Λ_n is adapted to the filtration $\mathcal{G}_n(t) = \mathcal{F}_n(t + \gamma_n)$, $t \geq 0$. The sequence q_n increases to $+\infty$, so γ_n decreases to 0. Finally, since $n \geq n_{q_n}$, we have $d_{\Lambda_n}(X, X_n) \leq 2^{-(q_n - 3)}$ on A_{q_n}, hence $\lim_{n \to \infty} d_{\Lambda_n}(X, X_n) = 0$ on the set A of (3.10): that is, $d_{\Lambda_n}(X, X_n) \to 0$ a.s., and the theorem is proved. ∎

(3.11) THEOREM. *Suppose stochastic processes X_n converge weakly in* $\mathbf{D}$ *to a process* X. *Then for each n there exists $\gamma_n > 0$ with $\lim_{n\to\infty} \gamma_n = 0$, and an extension $(\Omega''_n, \mathcal{F}''_n, P''_n)$ of $(\Omega_n, \mathcal{F}_n, P_n)$ with a process Y_n on Ω''_n such that $\mathcal{L}(Y_n) = \mathcal{L}(X)$, and there exists a process Λ''_n with paths in L such that Λ''_n is adapted to $(\mathcal{G}''_n(t))_{t\geq 0} = (\mathcal{F}''_n(t+\gamma_n))_{t\geq 0}$, where $\mathcal{F}''_n(t) = \sigma\{X_n(s), Y_n(s); s \leq t\}$, and such that $\lim_{n\to\infty} d_{\Lambda''_n}(X_n, Y_n) = 0$, with convergence in distribution.*

Proof: As in the proof of Theorem (2.3), by the Skorohod representation theorem we can find a space $(\Omega', \mathcal{F}', P')$ on which there exist X'_n X' with $\mathcal{L}(X'_n) = \mathcal{L}(X_n)$, and $\mathcal{L}(X') = \mathcal{L}(X)$, and $\lim_{n\to\infty} X'_n = X'$ a.s. (dP'). Therefore by Theorem (3.1) there exists a sequence of constants γ_n decreasing to 0, and processes Λ'_n adapted to $\mathcal{G}'_n(t) = \mathcal{F}'_n(t+\gamma_n), t \geq 0$ such that $\lim_{n\to\infty} d_{\Lambda'_n}(X'_n, X') = 0$ a.s.

Next, as in the proof of Theorem (2.4) one can construct an extension $(\Omega'', \mathcal{F}''_n, P''_n)$ of $(\Omega_n, \mathcal{F}_n, P_n)$ such that there exists a process Y_n with $\mathcal{L}(Y_n) = \mathcal{L}(X)$, and a process Λ''_n with paths in L such that $\lim_{n\to\infty} d_{\Lambda''_n}(X_n, Y_n) = 0$, with convergence in distribution. However since $\mathcal{L}(X_n, Y_n, \Lambda''_n) = \mathcal{L}(X'_n, X', \Lambda'_n)$, we have $\mathcal{L}(Y_n) = \mathcal{L}(X)$, and Λ''_n is adapted to $(\mathcal{F}''(t+\gamma_n))_{t\geq 0}$, where $\mathcal{F}''_n(t) = \sigma\{X_n(s), Y_n(s); s \leq t\}$. Since $\lim_{n\to\infty} d_{\Lambda''_n}(X_n, Y_n) = 0$, the theorem is proved. ∎

REFERENCES

[1] P. BILLINGSLEY, "Convergence of Probability Measures," Wiley, New York, 1968.

[2] C. DELLACHERIE, "Capacités et Processes Stochastiques," Springer-Verlag, Berlin Heidelberg New York, 1972.

[3] S. ETHIER AND T. G. KURTZ, "Markov Processes: Characterization and Convergence," Wiley, New York, 1986.

[4] J. JACOD AND A. N. SHIRYAEV, "Limit Theorems for Stochastic Processes," Springer Verlag, Berlin Heidelberg, New York, 1987.

[5] D. POLLARD, "Convergence of Stochastic Processes," Springer Verlag, Berlin Heidelberg New York, 1984.