Acta Sci. Math., 49 (1985), 397—400

On tightness of random sequences

WINFRIED STUTE

Let $(\xi_n)_n$ be a sequence of random elements in a complete separable metric space (X, d), defined on some probability space $(\Omega, \mathscr{A}, \mathbf{P})$. In many situations, particularly in statistical large sample theory, it is required to show that the laws $\mathscr{L}(\xi_n)$, $n \geqq 1$, converge weakly to some specified (Borel) measure μ. For this a general device is to guarantee that $\mathscr{L}(\xi_n)$, $n \geqq 1$, has at least one cluster point and, in a second step, that there is at most one of such points. While uniqueness may be shown by applying general methods for identifying weak limits (cf. BILLINGSLEY [1]), the existence part usually takes account of Prohorov's theorem. Accordingly, it remains to prove that ξ_n, $n \geqq 1$, is uniformly tight:

(1) for given $\varrho > 0$ there exists some compact subset K_ϱ of X such that $\mathbf{P}(\xi_n \notin K_\varrho) \leqq \varrho$ for all $n \geqq 1$.

Apart from stochastic arguments, to find such a K_ϱ, one has to characterize the (relatively) compact subsets of X. This might cause some difficulties due to the fact that such a description needs a far reaching investigation of the topology induced by d. In many cases, however, there exists a (closed) subspace X_0 of X such that

(2) the ξ_n's, as $n \to \infty$, concentrate more and more on X_0, so that a possible limit distribution is supported by X_0.

(3) the relative topology induced on X_0 admits a simpler characterization of compactness.

An important example we have in mind is the space $X = D[0, 1]$ of right-continuous functions on $[0, 1]$ with left-hand limits, endowed with the Skorohod topology (cf. BILLINGSLEY [1]). The class of processes with paths in D contains appropriate versions of partial sum, empirical and quantile processes. In each case the limit process may be chosen so as to have continuous paths, i.e. we may take $X_0 = C[0, 1]$, the space of continuous functions on $[0, 1]$. As a matter of fact the Skorohod topology on C coincides with the topology of uniform convergence.

Received August 29, 1983.

Thus a characterization of compactness in X_0 is obtained from the classical Arzela—Ascoli Theorem. Identification of the limit of course relies on the convergence of the finite dimensional distributions.

In this paper a simple method for proving tightness is proposed which is based on appropriate X_0-valued transformations $T_\varepsilon(\xi_n)$, $\varepsilon>0$, of ξ_n, $n\geqq 1$.

Proposition 1. *Assume that, for each* $\varepsilon>0$, T_ε: $X\to X_0$ *is a measurable transformation such that*

(4) $T_\varepsilon(\xi_n)$, $n\geqq 1$, *is tight in* (X_0, d) *for each* $\varepsilon>0$,

(5) $\lim_{\varepsilon\to 0}\limsup_{n\to\infty} \mathbf{P}(d(T_\varepsilon(\xi_n), \xi_n)\geqq\eta)=0$ *for all* $\eta>0$.

Then ξ_n, $n\geqq 1$, *is tight in* (X, d), *and each cluster point* μ *of* $\mathscr{L}(\xi_n)$, $n\geqq 1$, *satisfies* $\mu(X_0)=1$.

Proof. Fix some $\eta>0$. By (4) we have, given $\varepsilon>0$,

$$\mathbf{P}(T_\varepsilon(\xi_n)\notin M^\eta) \leqq \eta \quad \text{for all} \quad n \geqq 1$$

for some finite $M=M(\eta, \varepsilon)\subset X_0$, where $M^\eta=\{x\in X: d(x, M)<\eta\}$ is the open η-neighborhood of M in X. For small enough $\varepsilon>0$ (5) implies

$$\mathbf{P}(\xi_n\notin M^{2\eta}) \leqq 2\eta \quad \text{for all} \quad n \geqq n_0(\eta).$$

Since $\xi_1, ..., \xi_{n_0-1}$ are tight in (X, d), we may find some finite $M_0(\eta)\equiv M_0\supset M$ in X such that

$$\mathbf{P}(\xi_n\notin M_0^{2\eta}) \leqq 2\eta \quad \text{for all} \quad n \geqq 1.$$

For K_ϱ we may then take the closure of the set $\bigcap_{k\geqq 1} M_0^{\varrho 2^{-k}}$. To show that each cluster point μ is supported by X_0, assume w.l.o.g. that $\mathscr{L}(\xi_n)\to\mu$ weakly. Since X_0 is closed, $X_0^\eta\downarrow X_0$ as $\eta\downarrow 0$. Hence it remains to prove $\mu(X_0^\eta)=1$. As is well known, the set of η's for which X_0^η has a μ-null boundary forms a dense set in $(0, \infty)$. Hence it suffices to consider only such η's. In this case

$$\mu(X_0^\eta) = \lim_{n\to\infty} \mathbf{P}(\xi_n\in X_0^\eta).$$

That the right-hand side equals one now easily follows from (5) and the fact that $T_\varepsilon(\xi_n)\in X_0$ for all $\varepsilon>0$.

Let us show the usefulness of our approach by giving a straightforward proof of the following important result (cf. Billingsley [1], Theorem 15.5).

Proposition 2. *Let* ξ_n, $n\geqq 1$, *be a random sequence in* $D[0, 1]$ *such that*

(6) *for each* $\varrho>0$ *there exists some finite* $a>0$ *such that*

$$\mathbf{P}(|\xi_n(0)| \geqq a) \leqq \varrho \quad \textit{for all} \quad n \geqq 1.$$

(7) *for all* $\eta, \varrho>0$ *there exists some* $0<\delta<1$ *such that for all* $n \geqq n_0(\eta, \varrho)$

$$\mathbf{P}\Big(\sup_{|t-s|\leqq\delta} |\xi_n(t)-\xi_n(s)| \geqq \eta\Big) \leqq \varrho.$$

Then ξ_n, $n\geqq 1$, *is tight in* $(D[0, 1], d)$, *and each cluster point* μ *satisfies* $\mu(C[0, 1])=1$.

Proof. For $f\in D[0, 1]$, put $f(t)=f(1)$ for $t>1$ and $f(t)=f(0)$ for $t<0$. Let K be a smooth nonnegative kernel function on the real line, integrating to one and vanishing outside some bounded interval. Put

$$Tf(t) \equiv \tilde{f}(t) = \int f(x)K(t-x)\,dx = \int f(t-y)K(y)\,dy, \quad 0 \leqq t \leqq 1.$$

Obviously, $\tilde{f}\in C[0, 1]$. If $\sup_{|t-s|\leqq\delta} |f(t)-f(s)| < \eta$, we have for $|t-s|\leqq\delta$:

$$|\tilde{f}(t)-\tilde{f}(s)| \leqq \int |f(t-y)-f(s-y)|K(y)\,dy < \eta \int K(y)\,dy = \eta,$$

i.e. $\sup_{|t-s|\leqq\delta} |\tilde{f}(t)-\tilde{f}(s)|<\eta$ whenever $\sup_{|t-s|\leqq\delta} |f(t)-f(s)|<\eta$. Furthermore, if $|f(0)|<a$ and $\sup_{|t-s|\leqq\delta} |f(t)-f(s)|<\eta$, we obtain

$$\|f\| \equiv \sup_{0\leqq s\leqq 1} |f(s)| < a+\eta/\delta \equiv b < \infty$$

and thus $|\tilde{f}(0)|\leqq\|f\|<b$. It follows from (6) and (7) and the Arzela—Ascoli Theorem that $T(\xi_n)$, $n\geqq 1$, is tight in $C[0, 1]$.

Now, we may let K depend on ε in such a way that the degree of smoothing decreases as $\varepsilon\to 0$. To be specific, let

$$K(x) = K_\varepsilon(x) = \varepsilon^{-1}K_0(x/\varepsilon),$$

where K_0 is a preassigned probability kernel vanishing outside some finite interval, say $[-1, 1]$. Define

$$T_\varepsilon(f)(t) = \varepsilon^{-1}\int f(x)K_0\big((t-x)/\varepsilon\big)\,dx.$$

We already know that $T_\varepsilon(\xi_n)$, $n\geqq 1$, is tight in $C[0, 1]$ for each $\varepsilon>0$. Furthermore,

$$\tilde{f}(t)-f(t) = \int_{\mathrm{supp}(K)} [f(t-y)-f(t)]K(y)\,dy,$$

whence

$$\sup_{0\leqq t\leqq 1} |\tilde{f}(t)-f(t)| \leqq \sup_{\substack{0\leqq t\leqq 1\\ y\in\mathrm{supp}(K)}} |f(t-y)-f(t)|.$$

For $K=K_\varepsilon$, we have $\operatorname{supp}(K)\subset[-\varepsilon, \varepsilon]$ and thus $\sup_{0\leqq t\leqq 1} |\tilde{f}(t)-f(t)|<\eta$ whenever $\sup_{|t-s|\leqq\varepsilon} |f(t)-f(s)|<\eta$. Observe that $d(\tilde{f},f)\leqq \sup_{0\leqq t\leqq 1} |\tilde{f}(t)-f(t)|$ and conclude that for $\varepsilon\leqq\delta$

$$\mathbf{P}\big(d(T_\varepsilon(\xi_n), \xi_n) \geqq \eta\big) \leqq \varrho, \quad n \geqq n_0(\eta, \varrho).$$

This shows (5) and completes the proof of the proposition.

References

[1] P. Billingsley, *Convergence of probability measures*, Wiley (New York, 1968).

MATHEMATISCHES INSTITUT DER
JUSTUS-LIEBIG-UNIVERSITÄT
ARNDTSTRASSE 2
6300 GIESSEN, FRG