A REMARK ON SLUTSKY'S THEOREM

Freddy Delbaen

Departement für Mathematik, ETH Zürich

1. Introduction and Notation.

In Theorem 1 of the paper by [BEKSY] a generalisation of a theorem of Slutsky is used. In this note I will present a necessary and sufficient condition that assures that whenever X_n is a sequence of random variables that converges in probability to some random variable X, then for each Borel function f we also have that $f(X_n)$ tends to $f(X)$ in probability. The abstract way of formulating the result has the advantage that it shows how to decompose the problem. The key result is the Dunford-Pettis characterisation of relatively weakly compact subsets of the space L^1. Because of this immediate relationship I believe that the result is known. However I could not find a reference.
In the sequel $(\Omega, \mathcal{A}, \mathbb{P})$ is a fixed probability space and $(E, \mathcal{E})$ is a measurable space. The sequence $(X_n)_{n\geq 1}$ denotes a sequence of measurable functions of Ω into E. Also X denotes a measurable function of Ω into E. The distributions (image measures) of X_n, resp. X are denoted by μ_n, resp. μ.
The symbol $\mathcal{H}$, a subset of the space of measurable functions from E into $\mathbb{R}$, denotes the set which consists of those functions g such that $g(X_n)$ tends to $g(X)$ in probability. It is clear that $\mathcal{H}$ satifies some stability properties. First of all it is clear that $\mathcal{H}$ is a vector space stable for multiplication, i.e. an algebra. Also if $\phi\colon \mathbb{R} \to \mathbb{R}$ is continuous and $f \in \mathcal{H}$, then $\phi(f) \in \mathcal{H}$. It follows that for each $m \geq 0$ and $f \in \mathcal{H}$, the truncation f^m of f is also in $\mathcal{H}$, f^m is defined as $f^m(x) = f(x)$ if $|f(x)| \leq m$, $f^m(x) = m$ if $f(x) > m$ and $f^m(x) = -m$ if $f(x) < -m$. Conversely if all the truncations f^m are in $\mathcal{H}$, then also $f \in \mathcal{H}$. It is also obvious that $\mathcal{H}$ is closed for uniform convergence.
Let $M(E, \mathcal{E})$, M for short, be the space of all signed measures defined on the space $(E, \mathcal{E})$. A subset K of $M(E, \mathcal{E})$ is said to be relatively weakly compact if it is relatively weakly compact for the weak topology (i.e. $\sigma(M, M^*)$) on M. The Dunford-Pettis theorem states that K is relatively weakly compact if and only if there is a probability measure $\lambda \in M$ such that every element $\nu \in K$ is

1991 *Mathematics Subject Classification.* 28A20, 28A33 60B10, 60B12.
Key words and phrases. Slutsky's Theorem, weak compactness.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

absolutely continuous with respect to λ and such that the set $\left\{ \frac{d\nu}{d\lambda} \mid \nu \in K \right\}$ of Radon-Nikodym derivatives, is uniformly integrable in $L^1(\lambda)$. For information on weak compactness and related topics I refer to [G], last chapter.

Theorem 1. *Let us assume that the set $(\mu_n)_{n\geq 1}$ of distributions of X_n is relatively weakly compact. If $(f_k)_{k\geq 1}$ is a sequence of functions in $\mathcal{H}$ that converges pointwise to a function f then also $f \in \mathcal{H}$, i.e. $\mathcal{H}$ is stable for taking pointwise convergent limits.*

Proof. Let $K = \{\mu_n \mid n \geq 1\} \cup \{\mu\}$. Clearly K is relatively weakly compact. Because of the stability properties of $\mathcal{H}$, we may and do assume that the sequence f_k is uniformly bounded, e.g. for each k, we have $|f_k| \leq 1$. Since the measures in K have uniformly integrable RN derivatives, we immediately obtain that $\sup_{\nu\in K} \int_E |f - f_k| \, d\nu$ tends to zero. For given $\epsilon > 0$ we now take k_0 big enough to assure that $\sup_{\nu\in K} \int_E |f - f_{k_0}| \, d\nu < \epsilon$. Now we take n_0 so that for $n \geq n_0$, $\int_\Omega |f_{k_0}(X_n) - f_{k_0}(X)| \, d\mathbb{P} < \epsilon$. For $n \geq n_0$ we then have $\int_\Omega |f(X_n) - f(X)| \, d\mathbb{P} < 3\epsilon$. This reasoning shows that $f(X_n)$ tends to $f(X)$ in $L^1(\mathbb{P})$ and hence in probability. □

By a standard argument on monotone classes we can now deduce the next theorem, which I give without proof.

Theorem 2. *If the set $(\mu_n)_{n\geq 1}$ of distributions of X_n is relatively weakly compact and if $\mathcal{H} \supset \mathcal{G}$, then $\mathcal{H}$ contains all measurable functions with respect to the sigma algebra $\mathcal{B}$, generated by $\mathcal{G}$.*

In the paper by [BEKSY], the functions X_n take values in a separable metric space S and X_n tend to X in probability. Since in their case, all the X_n have the same distribution, it immediately follows that for every Borel function h on S, we have that $h \circ X_n$ tend to $h \circ X$ in probability. More precisely we have the following.

Theorem 3. *Let S be a metric space and suppose that the sequence of S-valued random variables X_n converges to X in probability. In order that for each Borel measurable function f, the sequence $f(X_n)$ converges to $f(X)$ in probability, it is necessary and sufficient that the sequence of distributions $(\mu_n)_{n\geq 1}$ is relatively weakly compact.*

Proof. The sufficency is dealt with in Theorem 1 and 2 above. The necessity of the weak compactness condition is rather trivial. Suppose that the sequence of distributions, $(\mu_n)_{n\geq 1}$, is not weakly compact. Then there is a bounded measurable function such that $\int g \, d\mu_n$ does not converge to $\int g \, d\mu$. It follows that $g(X_n)$ cannot converge to $g(X)$ in probability. □

If in the previous theorem we replace convergence in probability by convergence almost surely, then the statement is wrong. To see this we will give a counterexample. We start with the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ equipped with the usual normalised Lebesgue measure m. Let O be an open subset of $\mathbb{T}$ such that $m(O) < 1/2$ and

such that O is dense in $\mathbb{T}$. I will construct a sequence X_n, defined on some probability space, such that X_n converges to a random variable X almost surely. All the variables will be distributed uniformly on $\mathbb{T}$, i.e. $\mu_n = m$ for all n. However it will turn out that the almost sure covergence of $\mathbf{1}_O(X_n)$ to $\mathbf{1}_O(X)$ is false. The construction goes as follows. For each $\delta > 0$ and $x \in \mathbb{T}$, we put $g(\delta, x) = \frac{1}{2\delta} m(O \cap I_\delta^x)$, where I_δ^x is the symmetric interval around x with length 2δ. Since O is dense we obtain that $g(\delta, x) > 0$ for all $x \in \mathbb{T}$ and all $\delta > 0$. It is now easy to find integers $(k_l)_{l \geq 1}$ such that for almost every $x \in \mathbb{T}$ we have that

$$\sum_{l \geq 1} k_l \, g\left(x, \frac{1}{l+1}\right) = \infty.$$

To construct the variables X_n, we need a sequence of independent variables $(V_n)_{n \geq 1}$, uniformly distributed on $[-1, 1]$. The variable X is taken to be independent of the sequence V_n and to have a distribution equal to m. Let us put $K_0 = 0$ and $K_{l+1} = K_l + k_{l+1}$. For each n, $K_l < n \leq K_{l+1}$, we define $X_n = X + \frac{1}{l+1} V_n$. The distributions of the X_n are easily seen to be equal to m. Since for almost every $x \in \mathbb{T}$ we have that

$$\sum_n \mathbb{P}[X_n \in O \mid X = x] = \sum_{l \geq 1} k_l \, g\left(x, \frac{1}{l+1}\right) = \infty,$$

it follows from independence and the Borel Cantelli lemma that for almost every $\omega \in X^{-1}(O^c)$, $X_n(\omega) \in O$ infinitely often. The construction of the counterexample is therefore complete.

Acknowledgement. The author thanks the participants in the Mathematical Finance workshop in Oberwolfach, September 97 for discussions on this topic.

References

[BEKSY] M.T. Barlow, M.Émery, F.B. Knight, S. Song, M. Yor, *Autour d'un théorème de Tsirelson sur des filtrations browniennes et non browniennes*, Séminaire de Probabilités XXXII (1998).

[G] A. Grothendieck, *EVT*, Publicação da Sociedade de Matemática e S. Paulo (1954).

Eidgenössische Technische Hochschule Zürich, CH-8092 Zürich, Switzerland
E-mail address: delbaen@math.ethz.ch

AN ADDENDUM TO A REMARK ON SLUTSKY'S THEOREM

Freddy Delbaen

Departement für Mathematik, ETH Zürich

In [D], I gave a counter-example to the following statement. If X_n is a sequence of measurable functions taking values in a Polish space E, and converging almost surely to a measurable function X, then for every Borel function h defined on E, $h(X_n)$ converges a.s. to $h(X)$. In the case of convergence in probability, the statement holds provided the image measures (or distributions) form a relatively weakly compact sequence (Slutsky's theorem). After the paper was printed, I discovered in the paper by Dellacherie, Feyel and Mokobodzki, [DFM], that the counterexample was already known. In fact there the authors show:

Theorem. *If $(X_n)_{n\geq 1}$ is a sequence of measurable mappings and if X is a measurable function, all defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in the Polish Space $(E, \mathcal{T})$, then the following are equivalent*

(1) *for each real valued Borel function $h\colon E \to \mathbb{R}$ we have that $h(X_n)$ tends to $h(X)$ almost surely*
(2) *X_n tends to X almost surely in a stationary way, i.e. for almost every $\omega \in \Omega$ there is n_0 (depending on ω) such that for all $n \geq n_0$ we have $X_n(\omega) = X(\omega)$*

The aim of this addendum is not only to give credit to Dellacherie, Feyel and Mokobodzki but also to give some extra background information. As an example we will see that the Slutsky result can be stated in a different way. To fix notation, let $(E, \mathcal{T}, \mathcal{E})$ be a Polish space equipped with its topology $\mathcal{T}$ and its Borel structure $\mathcal{E}$. Let μ_n be a sequence of probability measures on $(E, \mathcal{E})$. We recall that the sequence μ_n tends weak* to the probability μ if for every $\mathcal{T}$-continuous bounded function f on E we have $\int f\, d\mu_n$ tends to $\int f\, d\mu$. We say that μ_n tends to μ weakly if for every Borel function f on E we have that $\int f\, d\mu_n$ tends to $\int f\, d\mu$. In this case we have that μ_n tends to μ weakly in the sense of the topology $\sigma(\mathcal{M}, \mathcal{M}^*)$. By a result of Grothendieck, to have weak convergence, it is sufficient to ask that for every open set $O \in \mathcal{T}$ we have that

1991 *Mathematics Subject Classification.* 28A20, 28A33 60B10, 60B12.
Key words and phrases. Slutsky's Theorem, weak compactness.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

$\mu_n(O)$ tends to $\mu(O)$. But for Polish spaces there are many topologies that give the same Borel structure, in fact every finer topology $\mathcal{T}'$, which is still Polish, gives, by Blackwell's theorem, the same Borel sets $\mathcal{E}$ on E. The relation between the weak* and the weak convergence becomes clearer thanks to the following result, see [S] page 91–93.

Theorem. *If A_n is a sequence of Borel sets in E, then there exists a finer topology $\mathcal{T}'$ on E, still Polish and such that each A_n is an open-closed set in $\mathcal{T}'$.*

Corollary. *If f_n is a sequence of Borel functions $f_n: E \to \mathbb{R}$, then there is a finer, still Polish, topology $\mathcal{T}'$ on E such that each f_n is continuous.*

This result gives us the following theorem.

Theorem. *For a sequence of probability measures on $(E, \mathcal{E})$, the following two properties are equivalent*

(1) *the sequence μ_n converges weakly to μ,*
(2) *for every Polish topology $\mathcal{T}'$, finer than $\mathcal{T}$, the sequence μ_n converges weak* to μ.*

The result of Dellacherie, Feyel and Mokobodzki can now be rephrased as

Theorem. *For a sequence of measurable mappings defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in a Polish space $(E, \mathcal{T})$, the following are equivalent*

(1) *the sequence X_n tends to X is a stationary way (as above)*
(2) *for every Polish topology $\mathcal{T}'$ finer than $\mathcal{T}$, we have that X_n tends to X almost surely.*

References

[D] F. Delbaen, *A remark on Slutsky's Theorem*, Séminaire de Probabilités XXXII, Lecture Notes in Mathematics **1686,** (1998), Spinger, 313–315.

[DFM] C. Dellacherie, D. Feyel, G. Mokobodzki, *Integrales de capacités fortement sous-additives*, Séminaire de Probabilités XVI, Lecture Notes in Mathematics **920** (1982), Spinger, 8–28.

[S] S. M. Srivastava, *A course on Borel sets* (1998), Springer.

Eidgenössische Technische Hochschule Zürich, CH-8092 Zürich, Switzerland
E-mail address: delbaen@math.ethz.ch