A TOPOLOGICAL VERSION OF SLUTSKY'S THEOREM

PAUL RÉSSEL

ABSTRACT. For weak convergence of probability measures on a product of two topological spaces the convergence of the marginals is certainly necessary. If however the marginals on one of the factor spaces converge to a one-point measure, the condition becomes sufficient, too. This generalizes a well-known result of Slutsky.

One of the most frequently applied theorems in Mathematical Statistics is the so-called "Slutsky's theorem". Roughly stated this theorem says that if a sequence of random variables converges in distribution to a certain limit law, then so does a slightly disturbed sequence. More precisely: let \( X_1, X_2, \ldots \) be a sequence of real valued random variables converging in distribution to \( \mu \). Let \( Y_1, Y_2, \ldots \) be another sequence of random variables defined on the same probability space, but not necessarily independent of the \( X_j \). Then if \( Y_n \) converges in probability to zero, the "disturbed" sequence \( X_n + Y_n \) converges in distribution to \( \mu \). Similarly, if \( Y_n \) converges to one then \( \mu \) is the limit law of \( X_n Y_n \), cf. [3].

More recently, in connection with infinite dimensional random variables (resp. stochastic processes), a generalisation of Slutsky's theorem was obtained for random variables with values in metric spaces [1, Chapter 1]. But not all range spaces of random variables are in a natural way metrizable and we are presenting now a rather general version of Slutsky's theorem which hopefully will cover all situations where a result like this may be needed.

Let \( E \) be an arbitrary topological space (not necessarily Hausdorff). The Borel-\( \sigma \)-field of \( E \) (i.e. the \( \sigma \)-field generated by the topology) is denoted \( \mathcal{B}(E) \). A probability measure \( \mu \) on \( \mathcal{B}(E) \) is called \( \tau \)-smooth if and only if \( \mu(G) = \sup \mu(G) \) for every net of open subsets \( \{G_\alpha\} \) of \( E \) filtering up to \( G \). This condition can be regarded as a kind of minimal compatibility of the measure with the underlying topology. If \( E \) has a countable base then of course every measure is \( \tau \)-smooth. If \( E \) is completely metrizable or locally compact, then \( \tau \)-smoothness is equivalent with being a Radon measure, cf. [4, p. 16].

Let \( M_1^\tau(E) \) denote the space of all \( \tau \)-smooth probability measures on some topological space \( E \). Then the weak topology on \( M_1^\tau(E) \) is by definition the coarsest topology such that for each bounded lower semicontinuous functions \( f: E \to \mathbb{R} \) the associated function \( \mu \mapsto \int f \, d\mu \) on \( M_1^\tau(E) \) is also lower semicontinuous. The famous Portmanteau theorem (cf. [4, Theorem 8.1]) states a lot of equivalent conditions for the weak convergence of nets of probability measures from which we only need, that \( \mu_\alpha \to \mu \) weakly if and only if \( \liminf \mu_\alpha(G) \geq \mu(G) \) for each open set \( G \subseteq E \).

Let \( E \) and \( F \) be two topological spaces, then always \( \mathcal{B}(E) \otimes \mathcal{B}(F) \subseteq \mathcal{B}(E \times F) \). The usual product of two measures \( \mu \) on \( \mathcal{B}(E) \) and \( \nu \) on \( \mathcal{B}(F) \) is only defined on
$B(E) \otimes B(F)$, but if $\mu$ and $\nu$ are both $\tau$-smooth, then there is a unique $\tau$-smooth extension of $\mu \otimes \nu$ to the larger $\sigma$-field $B(E \times F)$, denoted $\mu \otimes \nu$; cf. [2, Theorem 1].

Now we are able to state our result:

**THEOREM.** Let $E$ and $F$ be two (not necessarily Hausdorff) spaces. Let $\{\rho_\lambda\}$ be a net of $\tau$-smooth probability measures on $E \times F$ with marginals $\{\mu_\lambda\}$ on $E$ and $\{\nu_\lambda\}$ on $F$. Assume that $\{\mu_\lambda\}$ converges to some $\tau$-smooth measure $\mu$ and that $\{\nu_\lambda\}$ converges to a one-point measure $\epsilon_y$, where $y \in F$. Then $\{\rho_\lambda\}$ converges to $\mu \otimes \epsilon_y$.

**PROOF.** Let $G \subseteq E \times F$ be open, then there are open sets $U_\lambda \subseteq E$, $V_\lambda \subseteq F$, $\lambda \in \Lambda$ such that $G = \bigcup_{\lambda \in \Lambda} (U_\lambda \times V_\lambda)$. The measure $\rho := \mu \otimes \epsilon_y$ being $\tau$-smooth, we can find, given $\epsilon > 0$, finitely many $\lambda_1, \ldots, \lambda_n \in \Lambda$ such that

$$\rho \left( \bigcup_{j=1}^n (U_{\lambda_j} \times V_{\lambda_j}) \right) > \rho(G) - \epsilon.$$

Put $G_0 := \bigcup_{j=1}^n (U_{\lambda_j} \times V_{\lambda_j})$ and $I := \{ j \leq n : y \in V_{\lambda_j} \}$. If $I = \emptyset$, then $\rho(G_0) = 0$ and certainly $\liminf \rho_\lambda(G_0) \geq \rho(G_0)$. Suppose now that $I \neq \emptyset$. Let $A := \bigcup_{j \in I} U_{\lambda_j}$ and $B := \bigcap_{j \in I} V_{\lambda_j}$; then $G_1 := A \times B \subseteq G_0$ and $\rho(G_1) = \rho(G_0)$. Using the fact that

$$\limsup \nu_\lambda(B^c) \leq \epsilon_y(B^c) = 0$$

we get

$$\liminf \rho_\lambda(G_0) \geq \liminf \rho_\lambda(G_1) \geq \liminf (\mu_\lambda(A) - \nu_\lambda(B^c)) \geq \mu(A) = \rho(G_0) > \rho(G) - \epsilon.$$

Hence

$$\liminf \rho_\lambda(G) \geq \rho(G)$$

and the above-mentioned $\liminf$-condition shows that $\rho = \lim \rho_\lambda$. □

We now show how easily the classical theorem of Slutsky mentioned in the introduction (and a lot of similar results) are obtained from our theorem.

Suppose that $X_n, Y_n$, $n = 1, 2, \ldots$, are real random variables such that $P X_n$ (the distribution of $X_n$) converges to $\mu$, $Y_n$ converging in probability (and therefore in distribution, too) to zero. Let $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ be defined by $\varphi(x, y) := x + y$. Then

$$P X_n + Y_n \rightarrow P \varphi(X_n, Y_n) = (P X_n, Y_n) \varphi$$

and from our theorem

$$P(X_n, Y_n) \rightarrow \mu \otimes \epsilon_0.$$

The continuity of $\varphi$ now implies

$$P X_n + Y_n \rightarrow (\mu \otimes \epsilon_0) \varphi = \mu \epsilon_0 = \mu.$$

If instead $Y_n$ converges to one, we have to replace $\varphi$ by the continuous function $\psi(x, y) := xy$ to get

$$P X_n Y_n \rightarrow \mu.$$

The extension of Theorem 1 to the product of more than two, but finitely many spaces is more or less obvious. The following lemma shows that the result also extends to countable products.
Lemma. Let $E_1, E_2, \ldots$ be a countable sequence of topological spaces whose product is denoted $E$. Let $\{\rho^n\}$ be a net of $\tau$-smooth probability measures on $E$ such that $\{\rho^n_i\}$, the net of marginals in the $i$th coordinate, converges to a one-point measure $\epsilon_{x_i}$ for each $i \geq 1$. Then $\{\rho^n\}$ converges to $\epsilon_x$, where $x := (x_1, x_2, \ldots)$.

Proof. Let $\pi^n: E \to \prod_{i=1}^n E_i$ be the natural projection. By induction we conclude from Theorem 1 that $\{\rho^{n,r}_n\}$ converges to $\epsilon(x_1, \ldots, x_n)$ for all $n \in \mathbb{N}$. If $G \subseteq E$ is open, then $G$ has the form

$$G = \bigcup_{n=1}^{\infty} \pi^{-1}_n(G_n)$$

where $G_n \subseteq \prod_{i=1}^n E_i$ is open for all $n$.

We have to show $\lim\inf \rho_n(G) \geq \epsilon_x(G)$ and this is obvious if $x \notin G$. But if $x \in G$ then $x \in \pi^{-1}_n(G_n)$ for some $n$ and hence

$$\lim\rho_n(G) \geq \lim\inf \rho_n(\pi^{-1}_n(G_n))$$

$$= \lim\inf \rho^{n,r}_n(G_n)$$

$$\geq \epsilon(x_1, \ldots, x_n)(G_n) = 1. \quad \square$$

References


Institut für Mathematische Statistik der Westfälischen Wilhelms-Universität Münster, Roxeler Strasse 64, D-4400 Münster, Federal Republic of Germany