A UNIFORM LAW OF LARGE NUMBERS

DANIEL V. TAUSK

In what follows all random objects are assumed to be defined on a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and the probability distribution of a random object Y (i.e., the push-forward of $\mathbb{P}$ by Y) will be denoted by $\mathbb{P}_Y$. The real finite-dimensional vector space V appearing in the statements below will be assumed to be endowed with an arbitrary fixed norm $\|\cdot\|$ and with the corresponding topology. Recall that a topological space is called *separable* if it admits a countable dense subset and *first countable* if every point admits a countable fundamental system of neighborhoods. It holds that every metrizable space is first countable and every compact metrizable space is separable.

Proposition 1 (uniform strong law)**.** *Let $(\mathcal{Y}, \mathcal{B})$ be a measurable space, K be a compact separable and first countable topological space and $g : \mathcal{Y} \times K \to V$ be a map taking values in a real finite-dimensional vector space V. Let $(Y_n)_{n\geq 1}$ be an independent and identically distributed sequence of $\mathcal{Y}$-valued random objects and assume that the following conditions hold:*

(a) *for $\mathbb{P}_{Y_1}$-almost every $y \in \mathcal{Y}$, the map $K \ni \theta \mapsto g(y, \theta) \in V$ is continuous;*

(b) *for all $\theta \in K$, the map $\mathcal{Y} \ni y \mapsto g(y, \theta) \in V$ is measurable;*

(c) *there exists a measurable function $h : \mathcal{Y} \to [0, +\infty[$ such that the expected value $E\big(h(Y_1)\big)$ is finite and such that, for $\mathbb{P}_{Y_1}$-almost every $y \in \mathcal{Y}$, we have $\|g(y, \theta)\| \leq h(y)$ for every $\theta \in K$.*

Under such conditions, we have that for $\mathbb{P}$-almost every $\omega \in \Omega$ the equality

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{i=1}^{n} g\big(Y_i(\omega), \theta\big) = E\big(g(Y_1, \theta)\big)$$

holds for every $\theta \in K$, with the limit being uniform in θ.

Proof. By replacing $\mathcal{Y}$ with a subset of $\mathbb{P}_{Y_1}$-probability 1 and restricting all Y_n to a subset of Ω with $\mathbb{P}$-probability 1, we can assume that conditions (a) and (c) hold with "for $\mathbb{P}_{Y_1}$-almost every $y \in \mathcal{Y}$" replaced with "for every $y \in \mathcal{Y}$". Given a nonempty open subset U of K and $\theta \in K$, we define a map $H_{U,\theta} : \mathcal{Y} \to [0, +\infty[$ by setting

$$H_{U,\theta}(y) = \sup_{\theta' \in U} \|g(y, \theta') - g(y, \theta)\|, \tag{1}$$

Date: May 15th, 2024.

for all $y \in \mathcal{Y}$. Note that $H_{U,\theta} \leq 2h$ and that the function $H_{U,\theta}$ is measurable since, by the continuity of $g(y,\cdot)$, one can take the supremum in (1) over a countable dense subset of U. Given $\theta \in K$ and a decreasing countable fundamental system of open neighborhoods $(U_k)_{k\geq 1}$ of θ, the continuity of $g(y,\cdot)$ implies that $(H_{U_k,\theta})_{k\geq 1}$ converges pointwise to zero and therefore the Dominated Convergence Theorem yields $\lim_{k\to+\infty} E\big(H_{U_k,\theta}(Y_1)\big) = 0$. Moreover, the continuity of $g(y,\cdot)$ and the Dominated Convergence Theorem also yield that the map $K \ni \theta \mapsto E\big(g(Y_1,\theta)\big) \in V$ is continuous.

Now let $\varepsilon > 0$ be fixed. To prove the proposition, it is sufficient to show that, for $\mathbb{P}$-almost every $\omega \in \Omega$, we have

$$\left\| \frac{1}{n}\left[\sum_{i=1}^{n} g\big(Y_i(\omega),\theta\big)\right] - E\big(g(Y_1,\theta)\big)\right\| < \varepsilon,$$

for all $\theta \in K$ and for $n \geq 1$ sufficiently large (keep in mind that one will be able to choose the set of $\mathbb{P}$-probability 1 for which this holds independently of $\varepsilon > 0$ since it is sufficient to consider countably many $\varepsilon > 0$). By the observations in the beginning of the proof, each $\theta \in K$ has an open neighborhood $U(\theta)$ such that

$$E\big(H_{U(\theta),\theta}(Y_1)\big) < \varepsilon \tag{2}$$

and

$$\big\|E\big(g(Y_1,\theta')\big) - E\big(g(Y_1,\theta)\big)\big\| < \varepsilon, \tag{3}$$

for all $\theta' \in U(\theta)$. Since K is compact, we can find a finite subset F of K such that $K = \bigcup_{\theta\in F} U(\theta)$. By the Strong Law of Large Numbers, for $\mathbb{P}$-almost every $\omega \in \Omega$, the conditions

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} H_{U(\theta),\theta}\big(Y_i(\omega)\big) = E\big(H_{U(\theta),\theta}(Y_1)\big), \tag{4}$$

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} g\big(Y_i(\omega),\theta\big) = E\big(g(Y_1,\theta)\big) \tag{5}$$

hold for all $\theta \in F$. Let $\omega \in \Omega$ for which (4) and (5) are satisfied for all $\theta \in F$ be fixed, so that

$$\left\| \frac{1}{n}\left[\sum_{i=1}^{n} H_{U(\theta),\theta}\big(Y_i(\omega)\big)\right] - E\big(H_{U(\theta),\theta}(Y_1)\big)\right\| < \varepsilon, \tag{6}$$

$$\left\| \frac{1}{n}\left[\sum_{i=1}^{n} g\big(Y_i(\omega),\theta\big)\right] - E\big(g(Y_1,\theta)\big)\right\| < \varepsilon, \tag{7}$$

for all $\theta \in F$ and for $n \geq 1$ sufficiently large. Now let $n \geq 1$ be fixed such that (6) and (7) hold for all $\theta \in F$. For arbitrary $\theta' \in K$, we pick $\theta \in F$

with $\theta' \in U(\theta)$ and using (2), (6) and the definition of $H_{U(\theta),\theta}$ we obtain:

$$\left\| \frac{1}{n}\left[\sum_{i=1}^{n} g\big(Y_i(\omega),\theta'\big)\right] - \frac{1}{n}\left[\sum_{i=1}^{n} g\big(Y_i(\omega),\theta\big)\right]\right\| \le \frac{1}{n}\sum_{i=1}^{n} H_{U(\theta),\theta}\big(Y_i(\omega)\big) < E\big(H_{U(\theta),\theta}(Y_1)\big) + \varepsilon < 2\varepsilon. \tag{8}$$

Finally from (3), (7) and (8) we get

$$\left\| \frac{1}{n}\left[\sum_{i=1}^{n} g\big(Y_i(\omega),\theta'\big)\right] - E\big(g(Y_1,\theta')\big)\right\| < 2\varepsilon + \varepsilon + \varepsilon = 4\varepsilon,$$

concluding the proof. □

Corollary 2 (uniform weak law)**.** *Under the assumptions of Proposition 1, we have that the sequence of functions*

$$S_n(\omega) = \sup_{\theta\in K}\left\| \frac{1}{n}\left[\sum_{i=1}^{n} g\big(Y_i(\omega),\theta\big)\right] - E\big(g(Y_1,\theta)\big)\right\|, \quad \omega\in\Omega$$

converges to zero in probability.

Proof. As in the proof of Proposition 1 we obtain (by possibly deleting a subset of $\mathbb{P}$-probability zero from Ω) that the functions S_n are measurable because by continuity one can replace the supremum over K in the definition of S_n with the supremum over a countable dense subset of K. By Proposition 1, the sequence $(S_n)_{n\ge1}$ converges almost surely to zero and hence it also converges in probability to zero. □

Example 3. Assumption (c) in the statement of Proposition 1 is really necessary. Namely, consider the set of positive integers endowed with the discrete topology and let $K = \{1,2,\ldots\}\cup\{\infty\}$ denote its one-point compactification. A map $g : \mathcal{Y}\times K \to V$ satisfying assumptions (a) and (b) in the statement of Proposition 1 is determined by a sequence $(g_m)_{m\ge1}$ of measurable maps $g_m : \mathcal{Y}\to V$ that converges pointwise to a map $g_\infty : \mathcal{Y}\to V$. We set $V=\mathbb{R}$ and we let $\mathcal{Y}=[0,1]$ be endowed with its Borel σ-algebra and g_m be equal to m times the indicator function of the interval $\left]0,\frac{1}{m}\right]$ (so that $g_\infty = 0$). Let $(Y_n)_{n\ge1}$ be an independent sequence of random variables with a uniform distribution in $[0,1]$. We have $E\big(g_m(Y_1)\big)=1$ for all $m\ge1$. Given $n\ge1$ and $\omega\in\Omega$, we can take $m\ge1$ such that $\frac{1}{m}$ is less than every nonzero element of the finite set $\{Y_1(\omega),\ldots,Y_n(\omega)\}$ so that $\sum_{i=1}^{n} g_m\big(Y_i(\omega)\big)=0$. Hence

$$\sup_{m\in K}\left| \frac{1}{n}\left[\sum_{i=1}^{n} g_m\big(Y_i(\omega)\big)\right] - E\big(g_m(Y_1)\big)\right| \ge 1, \tag{9}$$

for all $n\ge1$ and all $\omega\in\Omega$ and we conclude that not even the thesis of the uniform weak law of large numbers (Corollary 2) holds.

In the statement of Proposition 1, if K is not compact we might in some cases be able to find a first countable compactification $\overline{K}$ of K such that $g(y,\cdot)$ admits a continuous extension to $\overline{K}$ for $\mathbb{P}_{Y_1}$-almost every $y \in \mathcal{Y}$. In this case all assumptions of Proposition 1 will be valid for the extension of g to $\mathcal{Y} \times \overline{K}$ and hence the thesis of the proposition will hold. Note that the desired extension of g will always exist if we let $\overline{K}$ be the Stone–Čech compactification of K, but Stone–Čech compactifications are typically not first countable. Below we show that Proposition 1 does not hold in general if K is not compact.

Example 4. Let V, $\mathcal{Y}$ and $(Y_n)_{n\geq 1}$ be as in Example 3 and let K be the set of positive integers endowed with the discrete topology. A map $g : \mathcal{Y} \times K \to V$ satisfying assumptions (a) and (b) in the statement of Proposition 1 is then determined by a sequence $(g_m)_{m\geq 1}$ of measurable maps $g_m : [0,1] \to \mathbb{R}$ and assumption (c) is also satisfied for instance if the sequence $(g_m)_{m\geq 1}$ is uniformly bounded. Since the Banach space $C\big([0,1]\big)$ of continuous real-valued functions on $[0,1]$ endowed with the supremum norm is separable, we can pick g_m such that the set $\big\{g_m : m \geq 1\big\}$ is dense in the unit ball of $C\big([0,1]\big)$. Given $n \geq 1$ and $\omega \in \Omega$, we can for any $\varepsilon > 0$ find a continuous map $g : [0,1] \to [0,1]$ such that $\big|g\big(Y_i(\omega)\big)\big| < \varepsilon$ for all $i = 1,\ldots,n$ and such that $E\big(g(Y_1)\big) > 1 - \varepsilon$. Thus, there exists $m \geq 1$ such that such inequalities hold with $g = g_m$. Hence inequality (9) is satisfied for all $n \geq 1$ and all $\omega \in \Omega$ so that again not even the thesis of the uniform weak law of large numbers holds.

Departamento de Matemática,
Universidade de São Paulo, Brazil
Email address: tausk@ime.usp.br
URL: http://www.ime.usp.br/~tausk