LAW OF LARGE NUMBERS WITH ESTIMATED PARAMETER

DANIEL V. TAUSK

ABSTRACT. We prove a simple version of the law of large numbers in which the independent and identically distributed random variables are dependent on an unknown parameter that is replaced by a consistent estimate.

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 statement below will be assumed to be endowed with an arbitrary fixed norm $\|\cdot\|$ and with the corresponding topology. Topological spaces are assumed to be endowed with their respective Borel σ-algebras.

Proposition 1. *Let $(\mathcal{Y}, \mathcal{B})$ be a measurable space, Θ be a topological space, $\theta_0 \in \Theta$ be a point, $g : \mathcal{Y} \times \Theta \to V$ be a map taking values in a real finite-dimensional vector space V, $(Y_n)_{n\geq 1}$ be an independent and identically distributed sequence of $\mathcal{Y}$-valued random objects and $(\hat{\theta}_n)_{n\geq 1}$ be a sequence of Θ-valued random objects. Assume that the following conditions hold:*

(a) *the map g is measurable with respect to the product σ-algebra on $\mathcal{Y} \times \Theta$;*

(b) *the point θ_0 has a countable fundamental system of neighborhoods in the space Θ;*

(c) *for every U belonging to some fundamental system of neighborhoods of θ_0 in Θ, the map*

$$G_U : \mathcal{Y} \ni y \longmapsto \sup_{\theta \in U} \big\|g(y, \theta) - g(y, \theta_0)\big\| \in [0, +\infty]$$

is measurable with respect to the completion of the probability measure $\mathbb{P}_{Y_1} : \mathcal{B} \to [0, 1]$;

(d) *the sequence $(\hat{\theta}_n)_{n\geq 1}$ converges to θ_0 in probability, i.e., for every Borel neighborhood U of θ_0 in Θ we have $\lim_{n\to+\infty} \mathbb{P}(\hat{\theta}_n \in U) = 1$;*

(e) *for $\mathbb{P}_{Y_1}$-almost every $y \in \mathcal{Y}$, the map $\Theta \ni \theta \mapsto g(y, \theta) \in V$ is continuous at the point θ_0;*

Date: September 4th, 2024.

(f) *there exists a neighborhood U_0 of θ_0 in Θ and 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 $\|g(y, \theta)\| \le h(y)$ for every $\theta \in U_0$.*

Under such conditions, we have that

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} g(Y_i, \hat{\theta}_n) = E\big(g(Y_1, \theta_0)\big) \tag{1}$$

in probability.

The technical measurability condition (c) in the statement of Proposition 1 would be hard to verify in practice, but luckily it is automatically satisfied under very mild conditions, as we will show later (Proposition 3 and Corollary 5).

Proof of Proposition 1. Note first that by replacing the σ-algebra $\mathcal{B}$ with the domain of the completion of $\mathbb{P}_{Y_1}$ and the probability measure $\mathbb{P}$ with its completion, the random objects Y_n remain independent with common distribution given by the completion of $\mathbb{P}_{Y_1}$, so that we can assume without loss of generality that the probability measure $\mathbb{P}_{Y_1}$ is complete.

It follows from conditions (b) and (c) that there exists a countable decreasing fundamental system of neighborhoods $(U_k)_{k\ge1}$ of θ_0 in Θ such that G_{U_k} is measurable for all $k \ge 1$. By (e), the sequence $(G_{U_k})_{k\ge1}$ converges $\mathbb{P}_{Y_1}$-almost surely to zero and by (f) we have $G_{U_k} \le 2h$ $\mathbb{P}_{Y_1}$-almost surely for k sufficiently large, so that the Dominated Convergence Theorem yields that:

$$\lim_{k\to+\infty} E\big(G_{U_k}(Y_1)\big) = 0. \tag{2}$$

Let $\varepsilon > 0$ be given. By (2) there exists a neighborhood U of θ_0 in Θ such that G_U is measurable and $E\big(G_U(Y_1)\big) < \varepsilon$. By the weak law of large numbers, we have

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} G_U(Y_i) = E\big(G_U(Y_1)\big) \tag{3}$$

in probability and therefore:

$$\lim_{n\to+\infty} \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} G_U(Y_i) < \varepsilon\right) = 1. \tag{4}$$

For $\theta \in U$ we have

$$\left\|\frac{1}{n}\sum_{i=1}^{n} g(Y_i, \theta) - \frac{1}{n}\sum_{i=1}^{n} g(Y_i, \theta_0)\right\| \le \frac{1}{n}\sum_{i=1}^{n} G_U(Y_i)$$

and therefore:

$$\left[\left\|\frac{1}{n}\sum_{i=1}^{n} g(Y_i,\hat{\theta}_n) - \frac{1}{n}\sum_{i=1}^{n} g(Y_i,\theta_0)\right\| \geq \varepsilon\right] \subset \left[\frac{1}{n}\sum_{i=1}^{n} G_U(Y_i) \geq \varepsilon\right] \cup [\hat{\theta}_n \notin U'],$$

where U' is some Borel neighborhood of θ_0 contained in U. From (4) and (d) we then obtain that

$$\lim_{n\to+\infty}\left[\frac{1}{n}\sum_{i=1}^{n} g(Y_i,\hat{\theta}_n) - \frac{1}{n}\sum_{i=1}^{n} g(Y_i,\theta_0)\right] = 0$$

in probability (condition (a) is required to ensure that $g(Y_i,\hat{\theta}_n)$ is measurable). The conclusion now follows from the fact that the weak law of large numbers yields that

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n} g(Y_i,\theta_0) = E\big(g(Y_1,\theta_0)\big)$$

in probability. □

Remark 2. Proposition 1 admits a "strong law" version i.e., if we replace condition (d) in the statement of the proposition by the condition that $(\hat{\theta}_n)_{n\geq 1}$ converges almost surely to θ_0 then the thesis can be replaced with the statement that the limit (1) holds almost surely as well. Namely, by the strong law of large numbers, $\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n} g(Y_i,\theta_0) = E\big(g(Y_1,\theta_0)\big)$ almost surely and therefore it is sufficient to show that for every $\varepsilon > 0$ there exists a set of $\mathbb{P}$-probability 1 such that for ω in such set we have

$$\left\|\frac{1}{n}\sum_{i=1}^{n} g\big(Y_i(\omega),\hat{\theta}_n(\omega)\big) - \frac{1}{n}\sum_{i=1}^{n} g\big(Y_i(\omega),\theta_0\big)\right\| < \varepsilon \tag{5}$$

for n sufficiently large (keep in mind that one is then able to choose such set of $\mathbb{P}$-probability 1 independently of $\varepsilon > 0$ since it is sufficient to consider countably many $\varepsilon > 0$). To prove the latter statement, pick a neighborhood U of θ_0 such that G_U is measurable and $E\big(G_U(Y_1)\big) < \varepsilon$, as in the proof of Proposition 1. The strong law of large numbers then yields that the limit (3) holds almost surely and for those $\omega \in \Omega$ at which both $\lim_{n\to+\infty}\hat{\theta}_n = \theta_0$ and (3) hold we have that (5) holds for n sufficiently large.

Now we establish some sufficient conditions for the validity of condition (c) in the statement of Proposition 1.

Proposition 3. *Condition (c) in the statement of Proposition 1 follows from condition (a) if we assume that θ_0 admits a separable neighborhood in Θ and that there exists a neighborhood U of θ_0 in Θ such that the map $g(y,\cdot)$ is continuous on U for $\mathbb{P}_{Y_1}$-almost every $y \in \mathcal{Y}$.*

Proof. If U' is an open neighborhood of θ_0 that is contained both in U and in a separable neighborhood of θ_0 then U' is itself separable and

$$G_{U'}(y) = \sup_{\theta \in D} \big\| g(y, \theta) - g(y, \theta_0) \big\|$$

for $\mathbb{P}_{Y_1}$-almost every $y \in \mathcal{Y}$, where D is a countable dense subset of U'. This implies that $G_{U'}$ is measurable with respect to the completion of $\mathbb{P}_{Y_1}$. □

We recall that a measurable space is called *standard Borel* if it is isomorphic to a Borel subset of a Polish space (i.e., a complete separable metric space) endowed with its Borel σ-algebra. In the statement below we denote by $\overline{\mathbb{R}} = [-\infty, +\infty]$ the extended real line.

Lemma 4. *If $\mathcal{Y}$ and Θ are standard Borel spaces and $f : \mathcal{Y} \times \Theta \to \overline{\mathbb{R}}$ is a measurable map with respect to the product σ-algebra on $\mathcal{Y} \times \Theta$ then the map*

$$F : \mathcal{Y} \ni y \longmapsto \sup_{\theta \in \Theta} f(y, \theta) \in \overline{\mathbb{R}}$$

is measurable with respect to the completion of any finite measure on $\mathcal{Y}$.

Proof. Without loss of generality, we assume that $\mathcal{Y}$ and Θ are Borel subsets of Polish spaces endowed with their respective Borel σ-algebras. For every $c \in \mathbb{R}$, the set $[F > c]$ is equal to the image under the first projection of the Borel set $[f > c]$ and it is therefore analytic. The conclusion follows from the fact that an analytic subset of a Polish space is measurable with respect to the completion of any finite measure defined on its Borel σ-algebra ([1, Theorem 4.10.12]). □

Corollary 5. *Condition (c) in the statement of Proposition 1 follows from condition (a) if the spaces $\mathcal{Y}$ and Θ are standard Borel.* □

References

[1] S. M. Srivastava, A Course on Borel Sets, Springer, 1998.

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