

Different Notions of Convergence For Measurable Functions

Jacob Denson*

November 10, 2022

Let's discuss the notions of convergence one can have for random variables, and for measurable functions. Random variables are just measurable functions on a measure space with total measure one, so for every notion of convergence we obtain for measurable functions, we will obtain an analogous definition for random variables, which is the same definition, but written in a more probabilistic language. Let's begin by listing out the main types of convergence one can have on a measure space Ω , equipped with a measure μ :

- A sequence of measurable functions $\{f_n : \Omega \rightarrow \mathbb{R}\}$ *converges in measure* to a function $f : \Omega \rightarrow \mathbb{R}$ if, for any $\varepsilon > 0$,

$$\limsup_{n \rightarrow \infty} \mu\left(\{x \in \Omega : |f_n(x) - f(x)| \geq \varepsilon\}\right) = 0.$$

Analogously, for a sequence of random variables $\{X_n : \Omega \rightarrow \mathbb{R}\}$, we say these random variables *converge in probability* to a random variable $X : \Omega \rightarrow \mathbb{R}$ if, for any $\varepsilon > 0$,

$$\limsup_{n \rightarrow \infty} \mathbb{P}\left(|X_n - X| \geq \varepsilon\right) = 0.$$

- A sequence of measurable functions $\{f_n : \Omega \rightarrow \mathbb{R}\}$ *converges almost everywhere* to a measurable function $f : \Omega \rightarrow \mathbb{R}$ if, there exists a measurable set $\Omega_0 \subset \Omega$ with $\mu(\Omega_0^c) = 0$, such that for any $x \in \Omega_0$, $\lim_{n \rightarrow \infty} f_n(x) = f(x)$. Equivalently,

$$\mu\left(\{x \in \Omega : \limsup_{n \rightarrow \infty} |f_n(x) - f(x)| \neq 0\}\right) = 0.$$

Analogously, a sequence of random variables $\{X_n : \Omega \rightarrow \mathbb{R}\}$ *converges almost surely* to a random variable $X : \Omega \rightarrow \mathbb{R}$ if there exists a measurable set $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that for any $x \in \Omega_0$, $\lim_{n \rightarrow \infty} f_n(x) = f(x)$. Equivalently,

$$\mathbb{P}\left(\limsup_{n \rightarrow \infty} |X_n - X| \neq 0\right) = 0.$$

*University of Madison Wisconsin, Madison, WI, jcdenson@wisc.edu

- A sequence of measurable functions $\{f_n : \Omega \rightarrow \mathbb{R}\}$ *converges in L^p* to a function $f : \Omega \rightarrow \mathbb{R}$ if

$$\limsup_{n \rightarrow \infty} \int |f_n(x) - f(x)|^p dx = 0.$$

In order for this definition to make sense, one normally assumes that each function in the family $\{f_n\}$, and the limiting function f , lies in $L^p(\Omega)$, the space of all functions $g : \Omega \rightarrow \mathbb{R}$ such that

$$\int |g(x)|^p dx < \infty.$$

A sequence of random variables $\{X_n : \Omega \rightarrow \mathbb{R}\}$ *converges in L^p* to a random variable $X : \Omega \rightarrow \mathbb{R}$ if

$$\limsup_{n \rightarrow \infty} \mathbb{E} |X_n - X|^p = 0.$$

In order for this definition to make sense, one normally assumes that each random variable in the family $\{X_n\}$, and the limiting variable X , lies in $L^p(\Omega)$, the space of all random variables Y such that $\mathbb{E} |Y|^p < \infty$.

The notions of convergence in measure and convergence almost everywhere are very closely related, by virtue of the fact that they look at quantities associated with how fast a function is converging pointwise. Convergence in measure tells us that for any $\varepsilon > 0$ and $\delta > 0$, if n is suitably large, then all points $x \in \Omega$ outside a set of measure δ will satisfy $|f_n(x) - f(x)| \leq \varepsilon$. Convergence almost everywhere tells us that for any point x outside a set of measure zero, for any $\varepsilon > 0$, if n is taken suitably large, then $|f_n(x) - f(x)| \leq \varepsilon$.

These definitions are *almost* exactly the same, but the problem is in the order of the quantifiers, which means that, in general, neither condition implies the other condition. In other words, there exists sequences of functions converging in measure, but not converging almost surely, and there also exists sequences of functions converging almost surely, but not converging in measure. For convergence in measure, the value n selected is allowed to depend on ε and δ , but not the point x (except that we may throw away a set of ‘bad points’ that has measure at most δ). For convergence almost everywhere, the value n is allowed to depend on ε and x (but we cannot throw away a set of ‘bad points’ with measure δ as in convergence in measure).

1 Convergence in Measure Implies Convergence Almost Everywhere

One way convergence in measure can imply convergence almost everywhere is if one has a more quantified estimate on the rate of convergence result and applying the *Borel-Cantelli Lemma*. For instance, suppose we can justify that, for some sequence $\{f_n\}$,

$$\mu\left(\{x \in \Omega : |f_n(x) - f(x)| \geq 1/n\}\right) \leq 1/2^n.$$

If we define $E_n = \{x \in \Omega : |f_n(x) - f(x)| \geq 1/n\}$, then we find that

$$\sum_{n=1}^{\infty} \mu(E_n) \leq \sum_{n=1}^{\infty} 1/2^n < \infty.$$

Thus the Borel Cantelli Lemma implies that the complement of the set

$$E^* = \bigcup_{n_0=1}^{\infty} \bigcap_{n=n_0}^{\infty} E_n = \limsup_{n \rightarrow \infty} E_n$$

is a set of measure zero. But

$$E^* = \{x \in \Omega : \limsup_{n \rightarrow \infty} n \cdot |f_n(x) - f(x)| \leq 1\}$$

and so if $x \in E^*$, then $f_n(x)$ converges to $f(x)$. More explicitly, if $x \in E^*$, then there exists n_0 such that $x \in E_n$ for all $n \geq n_0$, which means that

$$|f_n(x) - f(x)| \leq 1/n.$$

for all such n . This implies that $f_n(x) \rightarrow f(x)$ for all $x \in E^*$, and thus the sequence $\{f_n\}$ converges to f almost everywhere.

We now use another secret trick: *For any sequence converging qualitatively, by taking a clever subsequence we can often introduce a more quantitative convergence.* Thus if $\{f_n\}$ is a sequence converging in measure to a function f , then for any n , we take $\varepsilon = 1/n$. We can then choose an integer k_n such that

$$\mu\left(\{x \in \Omega : |f_{k_n}(x) - f(x)| \geq 1/n\}\right) \leq 1/2^n.$$

The Borel-Cantelli method in the last paragraph can then be applied to the subsequence $\{f_{k_n}\}$. Thus we conclude every sequence converging in measure has a *subsequence* converging almost everywhere.

On the other hand, there are sequences converging in probability but not almost everywhere. The standard example is the *typewriter sequence*, given for $f_n : [0, 1] \rightarrow \{0, 1\}$ defined to be the indicator function of the set

$$E_n = \{[x] : \log n \leq x \leq \log(n+1)\},$$

where $[x]$ denotes the decimal part of x . Then E_n has Lebesgue measure at most $\log(n+1) - \log(n) \lesssim 1/n$. This implies $\{f_n\}$ converges to zero in measure since for $\varepsilon < 1$,

$$\mu\left(\{x \in [0, 1] : |f_n(x)| \geq \varepsilon\}\right) = \mu(E_n) \leq 1/n,$$

and so

$$\lim_{n \rightarrow \infty} \mu\left(\{x \in [0, 1] : |f_n(x)| \geq \varepsilon\}\right) = \lim_{n \rightarrow \infty} \mu(E_n) \lesssim \lim_{n \rightarrow \infty} 1/n = 0.$$

On the other hand, for any $x \in [0, 1]$, x is contained in infinitely many of the sets E_n and so

$$\limsup |f_n(x)| = 1.$$

Thus f_n does not converge pointwise to zero for any particular value.

2 Convergence Almost Everywhere Implies Convergence in Measure

What does convergence almost everywhere imply about convergence in L^p ? Let $\{f_n\}$ be a sequence of measurable functions converging almost surely to a function f , i.e. that if

$$E^* = \{x \in \Omega : \limsup_{n \rightarrow \infty} |f_n(x) - f(x)| = 0\},$$

then $\mu((E^*)^c) = 0$. We claim that *in a finite measure space*, convergence almost everywhere implies convergence in measure. We must show that for any $m > 0$,

$$\lim_{n \rightarrow \infty} \mu\left(\{x \in \Omega : |f_n(x) - f(x)| \geq 1/m\}\right) = 0.$$

The key trick here is to introduce *monotonicity* into the problem. Let

$$E_{n,m} = \{x \in \Omega : |f_n(x) - f(x)| \geq 1/m\}.$$

If it was true that $E_{1,m} \supset E_{2,m} \supset \dots$, then the monotone convergence theorem would imply that (provided that we are in a *finite* measure space) that

$$\limsup_{n \rightarrow \infty} \mu(E_{n,m}) = \mu\left(\bigcap_{n=1}^{\infty} E_{n,m}\right).$$

We have

$$\bigcap_{n=1}^{\infty} E_{n,m} = \left\{x \in \Omega : \sup_{n \rightarrow \infty} |f_n(x) - f(x)| \geq 1/m\right\} \subset (E^*)^c,$$

and so we would conclude

$$\limsup_{n \rightarrow \infty} \mu(E_{n,m}) = \mu\left(\bigcap_{n=1}^{\infty} E_{n,m}\right) = 0,$$

which would complete the argument. Unfortunately, $\{E_{n,m}\}$ are not monotone, but we can *replace* them with bigger sets that *are* monotone. Define

$$F_{n,m} = \left\{x \in \Omega : \sup_{n' \geq n} |f_{n'}(x) - f(x)| \geq 1/m\right\},$$

then $E_{n,m} \subset F_{n,m}$ for all n and m , and so

$$\limsup_{n \rightarrow \infty} \mu(E_{n,m}) \leq \limsup_{n \rightarrow \infty} \mu(F_{n,m}).$$

It thus suffices to show the right hand side is zero for all m . But the sets $\{F_{n,m}\}$ *are* monotone decreasing, and so

$$\limsup_{n \rightarrow \infty} \mu(F_{n,m}) = \mu\left(\bigcap_{n=1}^{\infty} F_{n,m}\right).$$

We still have

$$\bigcap_{n=1}^{\infty} F_{n,m} \subset (E^*)^c,$$

and so the rest of the proof from before still gives us that

$$\limsup_{n \rightarrow \infty} \mu(F_{n,m}) = \mu\left(\bigcap_{n=1}^{\infty} F_{n,m}\right) = 0.$$

Thus we have shown that in a finite measure space, convergence almost everywhere implies convergence in measure. In particular, convergence almost surely implies convergence in probability. One way to remember this is through the *weak* and *strong* law of large numbers, since the *strong* law implies the *weak* law, the weak law is about convergence in probability, and the strong law is about almost everywhere convergence (so convergence almost everywhere implies convergence in probability).

On an infinite measure space, however, convergence almost surely does not imply convergence almost everywhere. For instance, on the measure space $\Omega = \{0, 1, \dots\}$ equipped with the counting measure, we can define

$$f_n(x) = \mathbf{I}(x = n).$$

Then $f_n(x) \rightarrow 0$ for any $x \in \Omega$, but for $\varepsilon < 1$,

$$\mu\left(\{x \in \Omega : |f_n(x)| \geq \varepsilon\}\right) = \mu(\{n\}) = 1.$$

Thus $\{f_n\}$ does not converge to zero in measure. Similar examples exist for any infinite measure space, by taking a sequence of functions which are the indicator functions of disjoint sets of measure one.

3 Convergence in L^p

Note that convergence almost everywhere and convergence in measure tell us most points are close to converging when n is large, *but they tell us nothing about the points that are not close to converging, except that this set is small*. We can get more control over these non-converging points by introducing the L^p norms, which measure convergence ‘on average’. For $1 \leq p < \infty$, convergence in L^p is equivalent to

$$\limsup_{n \rightarrow \infty} \int |f_n(x) - f(x)|^p = 0.$$

This implies *most points* x have $|f_n(x) - f(x)|$ small, i.e. Chebyshev’s inequality implies that for any $\varepsilon > 0$, if n is suitably large, for all $t \geq 0$,

$$\mu\left(\{x \in \Omega : |f_n(x) - f(x)| \geq t\}\right) \leq \varepsilon/t^p.$$

Again the order of quantifiers has changed. But in this case the order of quantifiers here is stronger than convergence in measure, so that L^p convergence implies convergence in

measure. But it does not imply convergence almost surely for $p < \infty$ (the typewriter sequence converges in L^p to zero for all $1 \leq p < \infty$). Convergence in L^∞ , on the other hand, *does* imply convergence almost surely, again by looking at quantifiers.

Finally, we discuss the relation between convergence in various different L^p spaces. On a finite measure space, L^p convergence for larger p implies L^p convergence for lower L^p (by *Hölder's inequality*). The situation is reversed for ‘discrete’ measure spaces like the integers, convergence in L^p here for lower p implies convergence in L^p for higher L^p . For general measure spaces, L^p convergences are disjoint from one another. Examples to show this follow by taking

$$f_n(x) = H_n \mathbf{I}(x \in E_n)$$

for some measurable set E_n with $|E_n| = W_n$ for some number W_n . One can check that

$$\int |f_n(x)|^p = H_n^p W_n.$$

To show L^{p_1} convergence does not imply L^{p_2} convergence, it suffices to choose sequences $\{H_n\}$ and $\{W_n\}$ such that $H_n^{p_1} W_n \rightarrow 0$, but $H_n^{p_2} W_n$ does not converge to zero. If $p_1 < p_2$, then one will have to choose $W_n \rightarrow 0$ (which is why this argument doesn’t work for ‘discrete spaces’), and if $p_1 > p_2$, then one will have to choose $W_n \rightarrow \infty$ (which is why this argument doesn’t work for finite measure spaces).