

1 Review of non-random convergence concepts

By convergence, we mean the limiting behavior of a sequence of numbers (A sequence is just a list of numbers, ordered according to the integers.)

Exs: (i) $X_1 = 10, X_2 = 5.6, x_3 = 7, \dots$ (no pattern); (ii) $X_n = \log(n), n = 1, 2, 3, \dots$; (iii) $X_n = 2$; (iv) $X_n = \frac{1}{n}$; etc.

We say that a sequence converges if the limit $\lim_{n \rightarrow \infty} X_n$ exists and is finite (if the limit is $+\infty$ or $-\infty$, then we say the sequence diverges).

Recall: the statement $\lim_{n \rightarrow \infty} X_n = a$ means that for all $\epsilon > 0$ there exists some integer $n^*(\epsilon)$ such that, for all $n > n^*(\epsilon)$, $|X_n - a| < \epsilon$.

Visually: as n gets large, the “tail” of the sequence X_n lies in an ϵ -neighborhood of a , for all $\epsilon > 0$.



When the limit doesn't exist, then we use:

- $\limsup_{n \rightarrow \infty} X_n \equiv \lim_{n \rightarrow \infty} (\sup_{m \geq n} X_m)$: least upper bound. (The sequence $\sup_{m \geq n} X_m$ is the maximum attained by the “tail” of the X sequence, excluding the first $n - 1$ elements.)
- $\liminf_{n \rightarrow \infty} X_n \equiv \lim_{n \rightarrow \infty} (\inf_{m \geq n} X_m)$: greatest lower bound
- The limsup and liminf always exists. If the limit exists, then $\lim = \limsup = \liminf$.
- Consider: $X_n = (-1)^n$. The limit doesn't exist, but limsup is 1 and liminf is -1.
- Consider $X_n = (-1)^n \cdot 1/n$:

n	X_n	$\sup_{m \geq n} X_m$	$\inf_{m \geq n} X_m$
1	-1	$\frac{1}{2}$	-1
2	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{3}$
3	$-\frac{1}{3}$	$\frac{1}{4}$	$-\frac{1}{3}$
4	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{5}$
5	$-\frac{1}{5}$	$\frac{1}{6}$	$-\frac{1}{5}$
6	$\frac{1}{6}$	$\frac{1}{6}$	$-\frac{1}{7}$
\vdots			
∞	0	0	0



2 Big-O and little-o notation

(Taken from White, *Asymptotic Theory for Econometricians*, chap. II)

- The sequence $\{b_n\}$ is *at most of order* n^λ , denoted $O(n^\lambda)$, if and only if for *some* real number Δ , $0 < \Delta < \infty$, there exists a finite integer N such that for all $n \geq N$, $|n^{-\lambda}b_n| < \Delta$.
- The sequence $\{b_n\}$ is of *order smaller than* n^λ , denoted $o(n^\lambda)$, if and only if for *every* real number $\delta > 0$, there exists a finite integer $N(\delta)$ such that for all $n \geq N(\delta)$, $|n^{-\lambda}b_n| < \delta$.

Remarks:

- $O(n^\lambda)$: $n^{-\lambda}b_n$ is eventually bounded. $O(1)$: b_n eventually bounded.
- $o(n^\lambda)$: $\lim_{n \rightarrow \infty} n^{-\lambda}b_n = 0$. $o(1)$: $b_n \rightarrow 0$.
- $b_n = o(n^\lambda) \implies b_n = O(n^\lambda)$
- $b_n = O(n^\lambda) \implies b_n = o(n^{\lambda+\delta})$ for $\delta > 0$

Let a_n and b_n be scalars. Then:

- If $a_n = O(n^\lambda)$ and $b_n = O(n^\mu)$, then $(a_nb_n) = O(n^{\lambda+\mu})$ and $(a_n + b_n) = O(n^{\max(\lambda, \mu)})$.

- If $a_n = o(n^\lambda)$ and $b_n = o(n^\mu)$, then $(a_n b_n) = o(n^{\lambda+\mu})$ and $(a_n + b_n) = o(n^{\max(\lambda, \mu)})$.
- If $a_n = O(n^\lambda)$ and $b_n = o(n^\mu)$, then $(a_n b_n) = o(n^{\lambda+\mu})$ and $(a_n + b_n) = O(n^{\max(\lambda, \mu)})$.

For asymptotic theory, you will see a lot of negative exponents on n . As examples:

$$\sqrt{n}b_n \text{ eventually bounded} \Leftrightarrow b_n = O(n^{-\frac{1}{2}}) = O\left(\frac{1}{n^{\frac{1}{2}}}\right)$$

$$na_n \text{ eventually bounded} \Leftrightarrow a_n = O(n^{-1}) = O\left(\frac{1}{n}\right).$$

Note that $-\frac{1}{2} > -1$. Then we have:

- $a_n = o\left(\frac{1}{n^{\frac{1}{2}}}\right)$
- $a_n = o(b_n) \Leftrightarrow \frac{a_n}{b_n} = o(1)$
- $a_n + b_n = O(n^{-\frac{1}{2}}) = O\left(\frac{1}{n^{\frac{1}{2}}}\right)$, &etc.

3 Convergence concepts for random sequences

Earlier we talked about convergence ideas for non-random (i.e., deterministic) sequences. Basically, we say that a sequence Z_n converges to some limit point Z^* iff, for any $\epsilon > 0$, we can find some integer $n^*(\epsilon)$ such that, for all $n > n^*(\epsilon)$,

$$|Z_n - Z^*| < \epsilon.$$

We also write $\lim_{n \rightarrow \infty} Z_n = Z^*$. Graphically, the “tail” of the sequence must lie in an arbitrarily small neighborhood of the limit point Z^* .

For random sequences, how do we define convergence? By way of an example, we illustrate two important convergence concepts: **convergence in probability** and **convergence almost surely**.

- Consider this class as a population. What is μ , the population mean of hours slept last night?
- Consider estimating μ by randomly sampling your classmates, and using the sample average. Sample 10 of your classmates:

n	value X_n	$\bar{X}_n \equiv \frac{1}{n} \sum_{i=1}^n X_i$
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

- Now intuitively, the sequence \bar{X}_n “converges” to μ . But in what precise sense?
- Most obvious and natural thought is that *every* sequence compiled by a student in the class should converge (as $n \rightarrow 25$) to μ . This is what is meant by “almost sure” convergence: all possible realizations of the random sequence should converge.

Formally: let $Z_1, Z_2, Z_3, \dots, Z_n, \dots$ and Z^* be defined on the same probability space (Ω, \mathbb{B}, P) . Then for each $\omega \in \Omega$, let $Z_1(\omega), Z_2(\omega), \dots, Z_n(\omega)$ and $Z^*(\omega)$ denote the corresponding values for the random variables. Almost-sure convergence is the statement that:

$$P(\omega : \lim_{n \rightarrow \infty} Z_n(\omega) = Z^*(\omega)) = 1$$

which is equivalent to:

$$\forall \epsilon > 0 : P(\omega : \exists n^*(\epsilon, \omega), \forall n > n^*, |Z_n(\omega) - Z^*(\omega)| < \epsilon) = 1$$

The set of ω 's such that the tails of the sequence $Z_1(\omega), Z_2(\omega), \dots, Z_n(\omega), \dots$ lie in an arbitrarily small neighborhood of $Z^*(\omega)$ for n large enough has probability 1.

When $Z^n = \bar{X}_n$, and $Z^* = \mu$, then the result is the *strong law of large numbers*.

- On the other hand, let's consider the sequences compiled by everyone in the class:

n	$\#(\bar{X}_n^i - \mu > 3)$	$\#(\bar{X}_n^i - \mu > 1)$	$\#(\bar{X}_n^i - \mu > 0.5)$
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

You expect that as $n \rightarrow 25$, the percentage of the sequences for which \bar{X}_n lies outside some neighborhood of μ should go down. This is formalized in the concept of convergence in probability.

Formally: as before, let $Z_1, Z_2, Z_3, \dots, Z_n, \dots$ and Z^* be defined on the same probability space (Ω, \mathbb{B}, P) . Then for each $\omega \in \Omega$, let $Z_1(\omega), Z_2(\omega), \dots, Z_n(\omega)$ and $Z^*(\omega)$ denote the corresponding values for the random variables. Convergence in probability is the statement that:

$$\forall \epsilon, \delta > 0, \exists n^*(\delta; \epsilon) : \forall n > n^*(\delta; \epsilon), P(\omega : |Z_n(\omega) - Z^*(\omega)| < \epsilon) > 1 - \delta.$$

More succinctly, for all $\epsilon > 0$, the sequence $P_n^\epsilon \equiv \text{Prob}(|Z_n - Z^*| < \epsilon)$ is a non-random sequence (recall the last problem set!). The idea of convergence in probability is that the sequence P_n^ϵ converges to 1, or $\lim_{n \rightarrow \infty} P_n^\epsilon = 1$, for all $\epsilon > 0$.

When $Z^n = \bar{X}_n$, and $Z^* = \mu$, then the result is the *weak law of large numbers*. This is usually straightforward to prove, as we showed last class, using some Chebyshev's inequality, or some other probability inequality.

- Convergence in probability is weaker than (implied by) almost-sure convergence. Only requires that “a decreasing percentage” of realized sequences lie outside an arbitrarily small neighborhood of the limit point Z^* for n large enough. Almost-sure convergence requires that *none* of the realized sequences lie outside an arbitrarily small neighborhood of the limit point Z^* .

Theorem: $Z_n \xrightarrow{a.s.} Z^* \implies Z_n \xrightarrow{P} Z^*$.

3.1 Examples

- Prob space: $([0, 1], \mathbb{B}_{[0,1]}, \mu)$
- Consider: for $\omega \in [0, 1]$

$$Z_n(\omega) = \begin{cases} 1 & \text{if } n \text{ odd and } \omega \in [0, \frac{1}{n}] \\ 1 & \text{if } n \text{ even and } \omega \in [1 - \frac{1}{n}, 1] \\ \omega & \text{otherwise} \end{cases}$$

$$Z^*(\omega) = \omega$$

- Example: if $\omega = \frac{1}{4}$, then $Z_1 = 1, Z_2 = \frac{1}{4}, Z_3 = 1, Z_4 = Z_5 = Z_6 = \dots = \frac{1}{4}$.
- Does $Z_n \xrightarrow{p} Z^*$? **YES**. For given ϵ small enough, and given n , the probability that $|Z_n(\omega) - Z^*(\omega)| > \epsilon$ is $\frac{1}{n}$. Therefore, for given δ , choose n^* such that $\frac{1}{n^*} < \delta$.
- Does $Z_n \xrightarrow{as} Z^*$? **YES!** We have $Z_n(\omega) \rightarrow Z^*(\omega) = \omega$, except for the measure-zero set $\{0\}$.

More formally, for all $\epsilon > 0$ and small, and all $\omega \in (0, 1]$, we have that, for $n > \frac{1}{\omega}$, $|Z_n(\omega) - Z^*(\omega)| < \epsilon$. (Actually it is stronger: we have $Z_n(\omega) = Z^*(\omega) = \omega$ for $n > \frac{1}{\omega}$.)

- What if we consider

$$Z_n(\omega) = \begin{cases} 1 & \text{if } \omega \in [0, \frac{1}{n}] \\ \omega & \text{otherwise} \end{cases}$$

Does $Z_n \xrightarrow{p} Z^*$? **YES**. Using similar argument as before.

Does $Z_n \xrightarrow{as} Z^*$? **YES**. Same argument as before.

- What about

$$Z_n(\omega) = \begin{cases} 1 & \text{if } \omega \in [0, 0.00001 + \frac{1}{n}] \\ \omega & \text{otherwise} \end{cases}$$

Does $Z_n \xrightarrow{p} Z^*$? **NO**. Here $P(\omega : |Z_n(\omega) - Z^*(\omega)| > \epsilon) \rightarrow 0.00001$, so that for any $\delta < 0.00001$, we cannot find a n^* .

- Here is an example which does not converge almost surely, but converges in probability.

Again, consider the probability space $([0, 1], \mathbb{B}_{[0,1]}, \mu)$. Define the random variables Z_1, Z_2, \dots as:

$$\begin{aligned}
 Z_1 &= \mathbb{1}_{[0,1]} \\
 Z_2 &= \mathbb{1}_{[0, \frac{1}{2}]} \\
 Z_3 &= \mathbb{1}_{[\frac{1}{2}, 1]} \\
 Z_4 &= \mathbb{1}_{[0, \frac{1}{3}]} \\
 Z_5 &= \mathbb{1}_{[\frac{1}{3}, \frac{2}{3}]} \\
 Z_6 &= \mathbb{1}_{[\frac{2}{3}, 1]} \\
 Z_7 &= \mathbb{1}_{[0, \frac{1}{4}]} \\
 Z_8 &= \mathbb{1}_{[\frac{1}{4}, \frac{2}{4}]} \\
 Z_9 &= \mathbb{1}_{[\frac{2}{4}, \frac{3}{4}]} \\
 &\quad \dots \\
 Z^* &= 0
 \end{aligned} \tag{1}$$

We have $Z_n \xrightarrow{p} Z^*$: for every $\epsilon, \delta > 0$, we have $P(|Z_n - Z^*| < \epsilon) > 1 - \delta$ for $n > \max(\frac{1}{\epsilon}, \frac{1}{\delta})$.

We *do not* have $Z_n \xrightarrow{as} Z^*$: for all $\omega \in [0, 1]$, $\limsup Z_n(\omega) = 1$ and $\liminf Z_n(\omega) = 0$.