October 13 **October 2018**

Tutorial 12: Random Processes

Lecturer: Utpal Mukherji/ Parimal Parag *Prepared by: Karthik*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

Contents

12.1 Preliminaries: Convergence of Real Sequences

We begin this lecture with some preliminaries on convergence of real sequences. We provide a mathematical definition for what it means for a real sequence $(x_n)_{n\in\mathbb{N}}$ to converge to a real number $x \in \mathbb{R}$. We then proceed to study four forms of convergences of sequences of random variables: almost sure, mean-squared, in probability, and in distribution. Further, we also indicate the relationship between these forms, and provide conditions under which one may be inferred from another. We finally present the Borel-Cantelli lemma and emphasize its role in ascertaining almost sure convergence, given knowledge about convergence in probability. We present examples to motivate the definitions.

We first begin with a formal definition of a real sequence (or a sequence of real numbers).

Definition 12.1.1. A real sequence or a sequence of real numbers is a function $f : \mathbb{N} \to \mathbb{R}$ that associates with each natural number $n \in \mathbb{N}$ a real number $x_n = f(n)$. Such a sequence is represented as $(x_n)_{n\in\mathbb{N}}$ or $(x_n)_{n\geq 1}$, where $x_n \in \mathbb{R}$ for each $n \in \mathbb{N}$.

Examples: $f(n) = \frac{1}{n}, \frac{1}{n^2}, n\log n, n^2, e^{n\log n}, \frac{1}{2^n}$, etc.

We now present a mathematical definition for convergence of a real sequence $(x_n)_{n\in\mathbb{N}}$ to a real number $x \in \mathbb{R}$.

Definition 12.1.2. Let $x \in \mathbb{R}$, and let $(x_n)_{n\in\mathbb{N}}$ be a real sequence. Then, we say that x_n converges to x as $n \to \infty$, denoted as

$$x_n \overset{n\to\infty}{\longrightarrow} x \quad \text{OR} \quad \lim_{n\to\infty} x_n = x,$$

if the following condition holds: given any real number $\epsilon > 0$, there exists a natural number $N = N(\epsilon) \in \mathbb{N}$ (N could possibly depend on the choice of ϵ) such that

$$|x_n - x| \leq \epsilon \quad \text{for all } n \geq N.$$

Thus for instance, $\frac{1}{n} \overset{n\to\infty}{\longrightarrow} 0$, $\left(1 - \frac{1}{n^2}\right) \overset{n\to\infty}{\longrightarrow} 1$, and so on.

Interpretation:
The above definition says that however small the choice of $\epsilon > 0$ is, there is always a stage $N \in \mathbb{N}$ (possibly depending on the choice of ϵ) such that each of the terms $x_N, x_{N+1}, x_{N+2}\cdots$ are all within ϵ distance from x. In other words, the terms of the sequence get closer to x as $n \to \infty$.

The following is an equivalent way of stating the above definition by replacing $\epsilon > 0$ with rational numbers of the form $\frac{1}{m}$, where $m \in \mathbb{N}$.

Definition 12.1.3. (Equivalent definition for convergence of real sequences) Let $x \in \mathbb{R}$, and let $(x_n)_{n\in\mathbb{N}}$ be a real sequence. Then, we say that x_n converges to x as $n \to \infty$, denoted as

$$x_n \overset{n\to\infty}{\longrightarrow} x \quad \text{OR} \quad \lim_{n\to\infty} x_n = x,$$

if the following condition holds: given any $m \in \mathbb{N}$, there exists a natural number $N = N(m) \in \mathbb{N}$ (N could possibly depend on the choice of m) such that

$$|x_n - x| \leq \frac{1}{m} \quad \text{for all } n \geq N.$$

Notice that the above definition is equivalent to the one presented in 12.1.2 since for any $\epsilon > 0$, there always exists $m \in \mathbb{N}$ such that $0 < \frac{1}{m} < \epsilon$. Similarly, given any $m \in \mathbb{N}$, there always exists a real number ϵ such that $0 < \epsilon < \frac{1}{m}$.

12.2 Modes of Convergence of Sequence of Random Variables

With the above definitions for convergence of real sequences in place, we now present some modes of convergence for a sequence of random variables. Since random variables are functions from sample space to $\mathbb{R}$, the following are also notions of convergence of a sequence of functions.

Definition 12.2.1. (Pointwise convergence or sure convergence) Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $(X_n)_{n\in\mathbb{N}}$ be a sequence (or a collection) of $\mathcal{F}$-measurable random variables. Further, let X be an $\mathcal{F}$-measurable random variable. Then, we say that X_n converges to X pointwise, denoted as

$$X_n \overset{n\to\infty}{\longrightarrow} X \quad \text{OR} \quad \lim_{n\to\infty} X_n = X,$$

if the condition

$$X_n(\omega) \overset{n\to\infty}{\longrightarrow} X(\omega) \quad \text{OR} \quad \lim_{n\to\infty} X_n(\omega) = X(\omega)$$

holds <u>for all</u> $\omega \in \Omega$.

Remark 1. Notice that for each $\omega \in \Omega$ and for each $n \in \mathbb{N}$, $X_n(\omega)$ and $X(\omega)$ are real numbers. Hence, the expression

$$X_n(\omega) \overset{n\to\infty}{\longrightarrow} X(\omega) \quad \text{OR} \quad \lim_{n\to\infty} X_n(\omega) = X(\omega)$$

is a statement about convergence of the real sequence $(X_n(\omega))_{n\in\mathbb{N}}$ to $X(\omega)$. Pointwise convergence requires that such a convergence happens for all $\omega \in \Omega$.

Remark 2. A more detailed manner of stating pointwise convergence is as follows: given any $\omega \in \Omega$ and real number $\epsilon > 0$, there exists a natural number $N = N(\omega, \epsilon) \in \mathbb{N}$ (here, N could possibly depend on the choices of ω and ϵ) such that

$$|X_n(\omega) - X(\omega)| \leq \epsilon \quad \text{for all } n \geq N,$$

or alternatively, given any $\omega \in \Omega$ and natural number $m \in \mathbb{N}$, there exists a natural number $N = N(\omega, m) \in \mathbb{N}$ (here, N could possibly depend on the choices of ω and m) such that

$$|X_n(\omega) - X(\omega)| \leq \frac{1}{m} \quad \text{for all } n \geq N.$$

Definition 12.2.2. (Almost sure convergence) Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $(X_n)_{n\in\mathbb{N}}$ and X be $\mathcal{F}$-measurable random variables. Then, we say that X_n converges to X almost surely as $n \to \infty$, denoted as $X_n \overset{a.s.}{\longrightarrow} X$, if the set

$$A := \{X_n \overset{n\to\infty}{\longrightarrow} X\} = \{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}$$

satisfies $P(A) = 1$.

Remark 3. Notice that pointwise convergence of a sequence $(X_n)_{n\in\mathbb{N}}$ of random variables to a random variable X implies almost sure convergence since for pointwise convergence,

$$\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} = \Omega,$$

and $P(\Omega) = 1$.

Before proceeding further, we note an important point: the above definition for almost sure convergence says that

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1,$$

which hides in it the fact that the set $\{\lim_{n\to\infty} X_n = X\} = \{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}$ belongs to the σ-algebra. The justification for this comes from the following expression:

$$\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} = \underbrace{\bigcap_{m=1}^{\infty} \underbrace{\bigcup_{N=1}^{\infty} \underbrace{\bigcap_{n=N}^{\infty} \underbrace{\left\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \le \frac{1}{m}\right\}}_{\in\mathcal{F} \text{ for each } n\ge N}}_{\in\mathcal{F} \text{ for each } N\ge 1}}_{\in\mathcal{F} \text{ for each } m\ge 1}}$$

Definition 12.2.3. (Convergence in mean-squared sense) Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $(X_n)_{n\in\mathbb{N}}$ and X be $\mathcal{F}$-measurable random variables. Then, we say that X_n converges to X in the mean-squared sense as $n \to \infty$, denoted as $X_n \xrightarrow{m.s.} X$, if

$$E[(X_n - X)^2] \to 0 \quad \text{as } n \to \infty.$$

Remark 4. Note that for each $n \in \mathbb{N}$, $(X_n - X)^2$ is a non-negative random variable, and therefore its expectation is a well-defined non-negative real number. The above definition is a statement about convergence of the real sequence $a_n = E[(X_n - X)^2]$ to zero as $n \to \infty$.

Definition 12.2.4. (Convergence in probability) Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $(X_n)_{n\in\mathbb{N}}$ and X be $\mathcal{F}$-measurable random variables. Then, we say that X_n converges to X in probability as $n \to \infty$, denoted as $X_n \xrightarrow{p} X$, if

$$P(|X_n - X| > \epsilon) \to 0 \quad \text{as } n \to \infty \text{ holds for every choice of real number } \epsilon > 0.$$

Remark 5. The above definition is a statement about convergence of the real sequence $b_n = P(|X_n - X| > \epsilon)$ to zero as $n \to \infty$, for every choice of $\epsilon > 0$.

Remark 6. The condition "for every $\epsilon > 0$" is very important in the above definition. If the condition in the definition does not hold for some $\epsilon > 0$, then convergence in probability does not occur; see example 4 of subsection 12.3.1 below.

Remark 7. Notice that the above definition for convergence in probability hides in it the fact that for every choice of real number $\epsilon > 0$, the set $\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \epsilon\} \in \mathcal{F}$. A verification of this fact is left as exercise.

Definition 12.2.5. (Convergence in distribution) Suppose $X_1, X_2, \ldots$ is a collection of random variables (not necessarily on a common probability space). Further, let X be another random variable. Then, X_n we say that converges to X in distribution as $n \to \infty$, denoted as $X_n \xrightarrow{d} X$, if

$$F_{X_n}(x) \xrightarrow{n\to\infty} F_X(x) \quad \text{for all } x \text{ where } F_X(\cdot) \text{ is continuous.}$$

Notice that in the above definition, the convergence of the distribution functions is required to take place only for $x \in \mathbb{R}$ where $F_X(\cdot)$ is continuous. It is not necessary for convergence to take place for all $x \in \mathbb{R}$.

12.3 Implications, Reverse Implications and Examples

In this section, we provide examples for each of the aforementioned modes of convergence. Before we do so, we note the following relationship of implications among the four forms of convergence (see Figure 12.1):

1. Almost sure convergence implies convergence in probability.

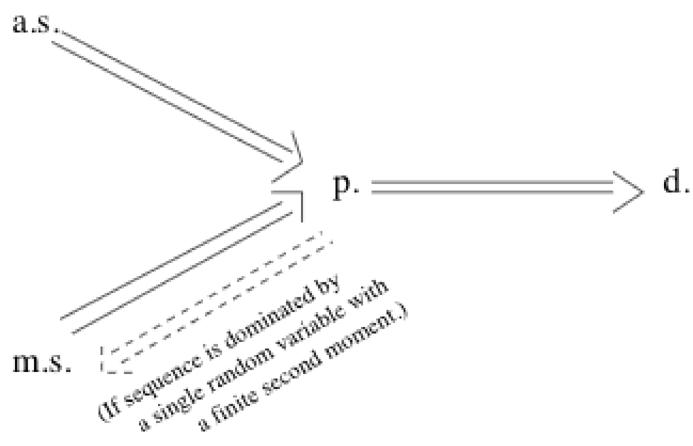

Figure 12.1: Implications among the various modes of convergence. (Picture courtesy: Prof. Bruce Hajek's lecture notes.)

2. Convergence in probability implies convergence in distribution.
3. Convergence in mean-squared sense implies convergence in probability.
4. There is no direct relationship between almost sure convergence and mean squared convergence. There are examples where almost sure convergence holds, but mean-squared convergence does not hold, and vice-versa.

We also note the following reverse implications:

1. For some constant $c \in \mathbb{R}$, if $X_n \xrightarrow{d} c$, then $X_n \xrightarrow{p} c$.
2. If $X_n \xrightarrow{p} X$, and X_n's are uniformly bounded, i.e.,
$$P(|X_n| \leq L) = 1$$
for all $n \in \mathbb{N}$, where L is a constant independent of n, then $X_n \xrightarrow{m.s.} X$.

12.3.1 Examples

1. Consider $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$ and $P = \lambda$, where λ denotes the Lebesgue measure. For each $n \geq 1$, let
$$X_n(\omega) = \begin{cases} 1, & \text{if } \omega \in \left[0, \frac{1}{n}\right), \\ 0, & \text{otherwise.} \end{cases}$$
We immediately infer that
$$P(X_n = 1) = \frac{1}{n} = 1 - P(X_n = 0),$$

which indicates that the sequence $(X_n)_{n\in\mathbb{N}}$ is possibly converging to zero since the probability of X_n taking the value zero is increasing to 1 as $n \to \infty$. Indeed, we have

$$\begin{aligned}
\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = 0\} &= \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \left\{\omega \in \Omega : |X_n(\omega) - 0| \le \frac{1}{m}\right\} \\
&= \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \{\omega \in \Omega : X_n(\omega) = 0\} \\
&= \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \left[\frac{1}{n}, 1\right] \\
&= \bigcap_{m=1}^{\infty} \bigcup_{N=1}^{\infty} \left[\frac{1}{N}, 1\right] \\
&= \bigcap_{m=1}^{\infty} (0, 1] \\
&= (0, 1],
\end{aligned}$$

from which it follows that $P(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = 0\}) = P((0,1]) = 1$, hence implying almost sure convergence of $(X_n)_{n\in\mathbb{N}}$ to zero. This then implies convergence in probability and convergence in distribution of the sequence $(X_n)_{n\in\mathbb{N}}$ to zero. Further we have

$$E[(X_n - 0)^2] = \frac{1}{n} \to 0 \quad \text{as } n \to \infty,$$

hence implying that $(X_n)_{n\in\mathbb{N}}$ converges to 0 in the mean-squared sense.

2. (Moving rectangles)
   The following example is to show that we can have convergence in probability (and hence in distribution) and convergence in mean-squared sense, but not almost surely.
   Let $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$ and $P = \lambda$, where λ denotes the Lebesgue measure. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables defined as follows:

   (a) $X_1 = 1_{[0,1]}$.

   (b) Each of X_2 and X_3 os defined on an interval of size 0.5 as

   $$X_2 = 1_{[0,\frac{1}{2}]}, \quad X_3 = 1_{[\frac{1}{2},1]}.$$

   (c) Each of X_4, X_5, X_6 and X_7 is defined on an interval of size 0.25 as

   $$X_4 = 1_{[0,\frac{1}{4}]}, \quad X_5 = 1_{[\frac{1}{4},\frac{1}{2}]}, \quad X_6 = 1_{[\frac{1}{2},\frac{3}{4}]}, \quad X_7 = 1_{[\frac{3}{4},1]}.$$

   (d) Each of $X_9, \ldots, X_{15}$ is defined on an interval of size 0.125 as

   $$X_8 = 1_{[0,\frac{1}{8}]}, \quad X_9 = 1_{[\frac{1}{8},\frac{2}{8}]}, \quad X_{10} = 1_{[\frac{2}{8},\frac{3}{8}]}, \quad \cdots \quad X_{15} = 1_{[\frac{7}{8},1]},$$

   and so on. Thus, X_n's are "moving rectangles of unit height and width shrinking to zero as $n \to \infty$".

   It follows that for any $\omega \in [0,1]$, there exists $n_0 \in \mathbb{N}$ for which $X_{n_0}(\omega) = 0$, and there exists $m_0 \in \mathbb{N}$ for which $X_{m_0}(\omega) = 1$. Thus, there is no almost sure convergence since for any $\omega \in \Omega$, the corresponding sequence $(X_n(\omega))_{n\in\mathbb{N}}$ will be a string with zeros and ones interleaved.

However, we note that for any $\epsilon > 0$,

$$P(|X_n - 0| > \epsilon) = \begin{cases} 0, & \epsilon \geq 1, \\ \underbrace{P(X_n = 1)}_{\text{width of the interval where } X_n \text{ takes the value 1}}, & 0 < \epsilon < 1, \end{cases}$$

from which we conclude that $P(|X_n - 0| > \epsilon) \to 0$ as $n \to \infty$ for every choice of $\epsilon > 0$ since the width of X_n's shrinks to zero as $n \to \infty$. Since the above is true for any $\epsilon > 0$, it follows that $(X_n)_{n \in \mathbb{N}}$ converges to 0 in probability (and hence in distribution). Furthermore, we have

$$P(|X_n| \leq 1) = 1 \quad \text{for all } n \in \mathbb{N},$$

hence implying convergence in the mean-squared sense.

3. Let $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$, and let P be any probability measure on $\mathbb{R}$. Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of random variables with

$$X_n \sim \mathcal{N}\left(0, \frac{1}{n}\right), \quad n \in \mathbb{N}.$$

Then, intuitively, X_n's seem to converge to 0 since for each $n \in \mathbb{N}$, $E[X_n] = 0$, and the variance of X_n is decreasing to 0 as $n \to \infty$. Indeed, we have

$$E[(X_n - 0)^2] = \frac{1}{n} \to 0 \qquad \text{as } n \to \infty,$$

thus implying that $X_n \xrightarrow{m.s.} 0$. This then implies that $X_n \xrightarrow{p} 0$ and $X_n \xrightarrow{d} 0$.

4. (In distribution but not in probability)
The following example is to demonstrate that a sequence of random variables may converge in distribution but not necessarily in probability.
Let U be a random variable distributed uniformly on $[0,1]$. For each $n \in \mathbb{N}$, define

$$U_n = 1 - U.$$

Then, clearly, U_n and U are identically distributed for each $n \in \mathbb{N}$, and we have

$$F_{U_n}(u) = F_U(u) = \begin{cases} 0, & u \leq 0, \\ u, & 0 < u < 1, \\ 1, & u \geq 1 \end{cases}$$

holding for all $n \in \mathbb{N}$. Since $F_U(\cdot)$ is continuous everywhere, we have

$$F_{U_n}(u) \to F_U(u) \qquad \text{as } n \to \infty \qquad \text{for all } u \in \mathbb{R},$$

thereby implying that $U_n \xrightarrow{d} U$. However, we note that

$$\begin{aligned} P(|U_n - U| > 0.5) &= P(|2U - 1| > 0.5) \\ &= P(U > 0.75) + P(U < 0.25) \\ &= 0.25 + 0.25 \\ &= 0.5 \\ &\nrightarrow 0, \end{aligned}$$

therefore implying that U_n <u>does not</u> converge to U in probability.

12.4 Borel-Cantelli Lemma

In this section, we formally state and prove an important result that, in some cases, helps infer almost sure convergence given information about convergence in probability. This result is known as the Borel-Cantelli lemma, and has two parts to it, as mentioned below:

Lemma 12.4.1. *(Borel-Cantelli Lemma) Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $(A_n)_{n\in\mathbb{N}}$ be a sequence of events in $\mathcal{F}$.*

1. *If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(A_n \text{ i.o.}) = 0$.*
2. *If $\sum_{n=1}^{\infty} P(A_n) = \infty$, and further, A_n's are mutually independent, then $P(A_n \text{ i.o.}) = 1$.*

Interpretation:

The first part of the lemma says: if $\sum_{n=1}^{\infty} P(A_n) < \infty$, then infinitely many of the A_n's cannot occur. In other words, with probability 1, only finitely many of the A_n's occur.

The second part of the lemma says: if $\sum_{n=1}^{\infty} P(A_n) = \infty$, and further, A_n's are mutually independent, then infinitely many of the A_n's must occur with probability 1.

Proof. 1. To prove the first part, we note that

$$\begin{aligned}\{A_n \text{ i.o.}\} &= \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k \\ &\subseteq \bigcup_{k=n}^{\infty} A_k \quad \text{for each } n \in \mathbb{N}.\end{aligned}$$

Therefore, for each $n \in \mathbb{N}$, we have

$$P(A_n \text{ i.o.}) \leq \sum_{k=n}^{\infty} P(A_n),$$

which implies that

$$P(A_n \text{ i.o.}) \leq \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k).$$

Since $\sum_{n=1}^{\infty} P(A_n) < \infty$, it follows that the right hand side of the above equation is equal to 0, thereby implying that $P(A_n \text{ i.o.}) = 0$.

2. To prove the second part, we note that

$$\{A_n \text{ i.o.}\}^c = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k^c.$$

For each $n \in \mathbb{N}$, define

$$B_n := \bigcap_{k=n}^{\infty} A_k^c.$$

Then, it follows that $B_1 \subseteq B_2 \subseteq B_3 \subseteq \cdots$. Therefore, by continuity of probability, we get

$$\begin{aligned}
P((A_n \text{ i.o.})^c) &= \lim_{n\to\infty} P(B_n) \\
&= \lim_{n\to\infty} P\left(\bigcap_{k=n}^{\infty} A_k^c\right) \\
&\overset{(a)}{=} \lim_{n\to\infty} \prod_{k=n}^{\infty} P\left(A_k^c\right) \\
&= \lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P\left(A_k\right)) \\
&\overset{(b)}{\leq} \lim_{n\to\infty} \prod_{k=n}^{\infty} e^{-P(A_k)} \\
&= \lim_{n\to\infty} \exp\left(-\sum_{k=n}^{\infty} P(A_k)\right) \\
&= 0,
\end{aligned}$$

therefore implying that $P(A_n \text{ i.o.}) = 1$. In the above set of equations, (a) follows from the mutual independence of A_n's (and therefore of A_n^c's), (b) follows from the inequality $1 - x \leq e^{-x}$ for $x \in [0,1]$, and the last line follows from the fact that $\sum_{n=1}^{\infty} P(A_n) = \infty$ implies that for each $n \in \mathbb{N}$, the partial sum $\sum_{k=n}^{\infty} P(A_k) = \infty$.

□

We now demonstrate the usefulness of Borel-Cantelli lemma in inferring almost sure convergence from convergence in probability in some cases.

12.4.1 Convergence in Probability to Convergence Almost Surely

Let $(X_n)_{n\in\mathbb{N}}$ and X be random variables on a common probability space $(\Omega, \mathcal{F}, P)$. Suppose that $X_n \xrightarrow{p} X$, i.e., for every choice of $\epsilon > 0$,

$$P(|X_n - X| > \epsilon) \to 0 \quad \text{as } n \to \infty.$$

Define $A_n = \{|X_n - X| > \epsilon\}$. Then, if the condition

$$\sum_{n=1}^{\infty} P(A_n) < \infty \quad \text{for every choice of } \epsilon > 0$$

is satisfied, then by the first part of Borel-Cantelli lemma, we have that for every choice of $\epsilon > 0$, the probability $P(A_n \text{ i.o.}) = 0$. That is, for every choice of $\epsilon > 0$, the event $\{|X_n - X| > \epsilon\}$ does not occur infinitely many times, which is equivalent to saying that for every choice of $\epsilon > 0$, the event $\{|X_n - X| \leq \epsilon\}$ occurs after some stage. It therefore follows that $X_n \xrightarrow{a.s.} X$.

We now present examples to drive the above mentioned point.

12.4.1.1 Examples

1. Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables with the distribution of X_n being

$$P(X_n = 1) = \frac{1}{n^2} = 1 - P(X_n = 0), \quad n \in \mathbb{N}.$$

Then, intuitively, X_n's seem to be converging to 0. Indeed, for every choice of $\epsilon > 0$, we have

$$P(|X_n - 0| > \epsilon) = P(X_n = 1) = \frac{1}{n^2},$$

and thus, by defining $A_n = \{|X_n - 0| > \epsilon\}$, we have

$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

for every choice of $\epsilon > 0$. Thus, by the first part of Borel-Cantelli lemma, we conclude that for every choice of $\epsilon > 0$, we have $P(A_n \text{ i.o.}) = 0$. That is, for every choice of $\epsilon > 0$, the event $\{|X_n - 0| > \epsilon\}$ does not occur infinitely many times. Equivalently, we have that for every choice of $\epsilon > 0$, the event $\{|X_n - 0| \leq \epsilon\}$ occurs for all but finitely many values of n, therefore implying that $X_n \xrightarrow{a.s.} 0$.

2. We now revisit example 3 of subsection 12.3.1. Notice that we did not mention anything about almost sure convergence in example 3. We now show that almost sure convergence indeed holds for this example. Towards this, we first recall that if $Y \sim \mathcal{N}(0, \sigma^2)$, then the moment generating function of Y is given by

$$M_Y(t) = E[e^{tY}] = e^{\frac{t^2\sigma^2}{2}}, \quad t \in \mathbb{R}.$$

Using this, we have

$$M_{X_n}(t) = E[e^{tX_n}] = e^{\frac{t^2}{2n}}, \quad t \in \mathbb{R}.$$

Then, for any choice of $\epsilon > 0$, we have

$$\begin{aligned} P(X_n > \epsilon) &= P(tX_n > t\epsilon) \quad \text{for any } t > 0 \\ &= P(e^{tX_n} > e^{t\epsilon}) \quad \text{for any } t > 0 \\ &\overset{(a)}{\leq} \frac{M_{X_n}(t)}{e^{t\epsilon}} \quad \text{for any } t > 0 \\ &= e^{-(t\epsilon - \log M_{X_n}(t))} \quad \text{for any } t > 0, \end{aligned}$$

where (a) above follows from the application of Markov's inequality to the strictly positive random variable e^{tX_n}. We now notice that the right hand side of the last line above depends on $t > 0$, while the left hand side does not. Thus, we can optimize the right hand side over all $t > 0$ and get the tightest bound. Doing so, we get

$$\begin{aligned} P(X_n > \epsilon) &\leq \min_{t>0} \left\{ e^{-(t\epsilon - \log M_{X_n}(t))} \right\} \\ &= \min_{t>0} e^{-(t\epsilon - \frac{t^2}{2n})} \\ &= e^{-\max\limits_{t>0}(t\epsilon - \frac{t^2}{2n})} \\ &= e^{-\frac{n\epsilon^2}{2}}. \end{aligned}$$

On similar lines, noting that X_n and $-X_n$ have the same characteristic function, we get

$$P(X_n < -\epsilon) = P(-X_n > \epsilon) \leq e^{-\frac{n\epsilon^2}{2}}.$$

Combining the above inequalities, we get

$$\begin{aligned} P(|X_n - 0| > \epsilon) &= P(X_n > \epsilon) + P(X_n < -\epsilon) \\ &\leq 2e^{-\frac{n\epsilon^2}{2}}. \end{aligned}$$

Thus, we notice that

$$\sum_{n=1}^{\infty} P(|X_n - 0| > \epsilon) \leq \sum_{n=1}^{\infty} 2e^{-\frac{n\epsilon^2}{2}} < \infty \quad \text{for every choice of } \epsilon > 0,$$

therefore implying by the first part of Borel-Cantelli lemma that $P(\{|X_n - 0| > \epsilon\} \text{ i.o.}) = 0$ for every choice of $\epsilon > 0$, which means that $X_n \xrightarrow{a.s.} 0$.