# Probability and Stochastic Processes Course Notes

Chia-Hao Chang, chchangkh@utexas.edu

Fall 2018

## 1 Convergence of Random Variables

In this chapter, we are going to discuss to the following problem: what does it mean by

$$\lim_{n\to\infty} X_n \to X$$

Namely, what do we mean by that a sequence of random variables $(X_n)$ **converges** to a random variable $X$ ?

### 1.1 Almost Sure Convergence

**Definition 1** (Almost Sure Convergence)**.** Given a probability triplet $(\Omega, \mathcal{F}, \mathbb{P})$ and a sequence of random variables $\{X_1, X_2, X_3, \cdots, X_n, \cdots\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$. We say a the sequence of random variables $(X_n)$ **converges almost surely** (a.s.) to $X^*$ if and only if

$$\mathbb{P}(\{\omega : \lim_{n\to\infty} X_n(\omega) = X^*(\omega)\}) = 1$$

or if we define $\mathcal{A}$

$$\mathcal{A} = \{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X^*(\omega)\}$$

$$X_n \to X^* \quad a.s. \quad \text{iff} \quad \mathbb{P}(\mathcal{A}) = 1$$

That is, $(X_n)$ converges to $X^*$ **almost everywhere** except on a subset of $\mathbb{P}$**-measure** $-$ **0**.

The almost sure convergence is defined through inspecting "one-by-one" the elements in $\Omega$ and see if it converges finally. If it does, we collect it into $\mathcal{A}$ and calculate the probability of $\mathcal{A}$ to see if the convergence is **almost everywhere**.

### 1.2 Convergence in Probability

**Definition 2** (Convergence in Probability)**.** Given a probability triplet $(\Omega, \mathcal{F}, \mathbb{P})$ and a sequence of random variables $\{X_1, X_2, X_3, \cdots, X_n, \cdots\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$. We say that the sequence of random variables $(X_n)$ **converges in probability** to $X^*$ if

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}(\{|X_n(\omega) - X^*(\omega)| > \varepsilon\}) = 0$$

More precisely,

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}(\{\omega \in \Omega : |X_n(\omega) - X^*(\omega)| > \varepsilon\}) = 0$$

In fact, this is equivalent saying that **probability of error** should be arbitrarily small as $n$ grows up no matter what tolerance interval is given. To see this, let's define an "event of error." Given a tolerance interval $\varepsilon > 0$, the "event of error" $B_n^{(\varepsilon)}$ is defined as

$$B_n^{(\varepsilon)} = \{\omega : |X_n(\omega) - X^*(\omega)| > \varepsilon\}$$

The probability of error at the $n^{th}$ random variable $P_e^{(n)}(\varepsilon)$ is

$$P_e^{(n)}(\varepsilon) = \mathbb{P}(B_n^{(\varepsilon)})$$

We may view $(P_n^{(e)}(\varepsilon))$ as a sequence of "probability of error". i.e.

$$P_e^{(1)}(\varepsilon), P_e^{(2)}(\varepsilon), \cdots, P_e^{(n)}(\varepsilon), \cdots$$

Then, $X_n \to X^*$ in probability if and only if

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} P_e^{(n)}(\varepsilon) = 0$$

Convergence in probability is defined through the sequence of "probability of error" under a given tolerance (i.e. $\varepsilon$). That is, we look at every "cross section" of the random variable $X_n$ and calculate the probability of error. If such probability decreases to 0 as $n \to \infty$ for every $\varepsilon > 0$. We say the sequence $(X_n) \to X^*$ **in probability**.

## 1.3 Convergence in Mean Square

**Definition 3** (Convergence in Mean Square)**.** Given a probability triplet $(\Omega, \mathcal{F}, \mathbb{P})$ and a sequence of random variables $\{X_1, X_2, \cdots X_n, \cdots\}$ on it, we say the sequence $(X_n)$ converges to $X^*$ in **Mean Square** if and only if

$$\lim_{n\to\infty} \mathbb{E}[|X_n - X^*|^2] = 0$$

The mean square convergence is analogous to convergence in probability in the sense they are said to converge if and only if the probability of error decreases to 0 when $n \to \infty$.

If we define the probability of error at random variable $X_n$, $P_e^{(n)}$, to be

$$P_e^{(n)} = \mathbb{E}[|X_n - X^*|^2]$$

We say $X_n$ converges to $X^*$ in **mean square** if $P_e^{(n)} \to 0$ as $n \to \infty$.

## 1.4 Point-wise Convergence

**Definition 4** (Point-wise Convergence)**.** Given a probability triplet $(\Omega, \mathcal{F}, \mathbb{P})$ and a sequence of random variables $\{X_1, X_2, \cdots, X_n, \cdots\}$ on it. We say $(X_n)$ converges to $X^*$ $\mathbf{point-wise}$ if $X_n$ converges to $X^*$ on every point of $\Omega$.

$$X_n \to X^* \;\; \text{point} - \text{wise} \quad \Longleftrightarrow \forall \omega \in \Omega, \quad X_n(\omega) \to X^*(\omega)$$

One might want to compare the definition of point-wise convergence in random variables with the definition of point-wise convergence in real analysis. In fact, these two definition coincides with each other as long as $X_n(\omega)$ are regarded as sequence of functions $X_n : \Omega \to \mathbb{R}$, where $\mathbb{R}$ is a metric space with standard measure $d(x, y) = |x - y|$. Different from real analysis, the $\mathbf{point-wise}$ convergence is the "strongest" convergence condition among all types of convergence.

### 1.5 Convergence in Distribution

**Definition 5** (Convergence in Distribution)**.** Given a sequence of random variables $\{X_1, X_2, \cdots, X_n, \cdots\}$ (possibly not in the same probability space). We say $X_n$ converges to $X^*$ **in distribution** if $\lim_{n\to\infty} F_{X_n}(x) = F(x)$ for $x$ being a point of continuity.

$$X_n \xrightarrow{\text{d}} X^* \quad \iff \forall x : \text{continuous}, \quad \lim_{n\to\infty} F_{X_n} = F_X(x)$$

## 2 Limit of Events and Borel-Cantelli Lemma

Let $(A_n) = \{A_1, A_2, \cdots, A_n, \cdots\}$ be a sequence of events. We define the upper limit and the lower limit of the sequence $(A_n)$ in the following sense.

**Definition 6** (Limit Supremum and Infimum)**.** The limit supremum $\limsup_{n\to\infty} A_n$ is defined as

$$\limsup_{n\to\infty} A_n = \bigcap_{n\geq 1}^{\infty} \bigcup_{k\geq n}^{\infty} A_k$$

On the other hand, the limit infimum $\liminf_{n\to\infty} A_n$ is defined as

$$\liminf_{n\to\infty} A_n = \bigcup_{n\geq 1}^{\infty} \bigcap_{k\geq n}^{\infty} A_k$$

Intuitively, if $x$ is in the limit supremum of $A_n$'s, $x \in \limsup_{n\to\infty} A_n$, this means

$$x \in \Lambda_n = \bigcup_{k\geq n}^{\infty} A_k, \quad \forall n \in \mathbb{N}$$

This means that no matter how far $N$ we choose, we can always find an index $n > N$ s.t. $x$ lies in $A_n$, implying that $x$ occurs **infinitely many times** or **infintely often** in the collection $A_n$'s (otherwise, if $x$ lies in $A_n$'s for finitely many times, then there exists $n_0 \in \mathbb{N}$ s.t. $x \notin A_n$ whenever $n > n_0$ and $x \notin \bigcup_{k\geq n} A_k$ for $n > n_0$). This means the elements in $\limsup_{n\to\infty} A_n$ are those occurs infinitely many times in $A_n$'s.

Similarly, if $y$ lies in the limit infimum of $A_n$'s, then $y$ lies in at least one the sets $\bigcap_{k\geq n^*} A_k$ for some $n^* \in \mathbb{N}$.

$$y \in \bigcap_{k\geq n^*}^{\infty} A_k$$

implying that $y \in A_k$, $\forall k \geq n^*$. This means that for $y$ to be in $\liminf_{n\to\infty} A_n$, $y$ must be in all $A_n$'s, with finitely many exceptions. Namely, $y$ **eventually stays in** $A_n$'s **forever**.

To rephrase the above explanations,

$$\limsup_{n\to\infty} A_n = \{x : x \in A_n \text{ occurs infinitely often}\}$$
$$\liminf_{n\to\infty} A_n = \{x : x \in A_n \text{ occurs all except finitely often}\}$$

**Remark 1.** *The following are equivalent*

*(a)* $A = \limsup_{n\to\infty} A_n$

*(b) event $A$ occurs* **infinitely often** *(i.o.)*

*(c) the event $A$ "never leaves forever."*

**Remark 2.** *The following are equivalent*

*(a)* $B = \limsup_{n\to\infty} A_n$

*(b) event $B$ occurs* **all except finitely often** *(a.e.f.o.)*

*(c) the event $A$ "eventually stays forever."*

Having these definition in mind, we proceed to the famous **Borel − Cnatelli Lemma** in measure theory.

**Theorem 1** (Borel-Cantelli Lemma)**.** *Given a probability triplet $(\Omega, \mathcal{F}, \mathbb{P})$ and a sequence of events $\{A_1, A_2, \cdots, A_n, \cdots\}$ over it. If*

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

*then*

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = 0$$

*where* $\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$

Sometimes the Borel-Cantelli Lemma is also called the **first Borel − Cantelli Lemma**.

The Borel-Cantelli Lemma states that, if we have a sequence of events and we would like to determine whether "$A_n$'s occurs infinitely often", we can instead evaluate the infinite series sum of the probability of $A_n$'s.

**Corollary 1.** *Let $(\Omega, \mathcal{F})$ be a measurable space and $\mu$ be a measure. $\{A_1, A_2, \cdots, A_n, \cdots\}$ be an infinite collection of subsets of $\Omega$ such that $A_j \in \mathcal{F}$, $\forall j \in \mathbb{N}$. If $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, then $\mu(\limsup_{n\to\infty} A_n) = 0$.*

**Theorem 2** (Second Borel-Cantelli Lemma)**.** *Suppose $\{B_1, B_2, \cdots, B_n, \cdots\}$ are* **independent** *Borel sets (i.e. events). If $\sum_{n=1}^{\infty} \mathbb{P}(B_n) = \infty$, then $\mathbb{P}(\limsup_{n\to\infty} B_n) = 1$.*

*Proof.* The proof utilizes the property that $1 - x \le e^{-x}$, for $x \ge 0$. The rest follows. □

## 2.1 Application of Borel-Cantelli Lemma to Almost Sure Convergence

Let's define $A_n^{\varepsilon}$ as follows,

$$A_n^{\varepsilon} = \{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}$$

If we can show that

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n^{\varepsilon}) < \infty$$

By Borel-Cantelli Lemma, we have

$$\mathbb{P}(\limsup_{n\to\infty} A_n^\varepsilon) = 0$$

That is, $A_n$'s can not happen **infinitely often**, implying that there exists some $n_0 \in \mathbb{N}$, s.t. $A_n$ does not occur whenever $n > n_0$. If this holds for arbitrary $\varepsilon > 0$, then we have $X_n \to X$ almost surely.
It is noteworthy that the opposite is not true. That is, even though $\sum_{n=1}^{\infty} \mathbb{P}(A_n^\varepsilon) = \infty$, $\mathbb{P}(\limsup_{n\to\infty} A_n^\varepsilon)$ might still be 0, implying almost sure convergence of $X_n \to X$.

# 3 Convergence and Cauchy Criteria

**Proposition 1.** *(a) If $X_n \to X$ a.s, then $X_n \to X$ in probability.*

*(b) If $X_n \to X$ m.s., then $X_n \to X$ in probability.*

*(c) If $\mathbb{P}(|X_n| \le Y) = 1$, $\forall n$ for some fixed random variable $Y$ with $\mathbb{E}[|Y|^2] \le \infty$ and if $X_n \to X$ in probability, then $X_n \to X$ m.s.*

*(d) If $X_n \to X$ in probability, then $X_n \to X$ in distribution.*

*(e) Suppose $X_n \to X$ in probability/ mean square/ almost sure sense and $X_n \to Y$ in probability/ mean square/ almost sure sense. Then $\mathbb{P}(\{X = Y\}) = 1$. That is, the measure of $\{X \neq Y\}$ is 0. i.e. the limit of a sequence of random variables (if exists) is* **unique**. *Thus, we say $X$ (or equivalently $Y$) is "the" limit of the sequence.*

*(f) Suppose $X_n \to X$ in distribution and $X_n \to Y$ in distribution. Then $X$ and $Y$ have the same distribution ($F_x(c) = F_Y(c), \forall c$).*

**Theorem 3** (Monotone Convergence Theorem for Random Variables)**.** *Let $\{X_1, X_2, \cdots, X_n, \cdots\}$ be a sequence of random variables. If $\mathbb{P}(\{X_n \le X_{n+1}\}) = 1$, $\forall n$ and if $\exists$ a fixed random variable $Y$ s.t. $\mathbb{P}(\{X_n \le Y\}) = 1$, $\forall n$, then the sequence $(X_n)$ converges* **almost surely**.

**Proposition 2** (Cauchy Criteria for Random Variables)**.** *Let $(X_n)$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

*(a) $(X_n)$ converges* **almost surely** *(to some random variable $X$)* **if and only if**

$$\mathbb{P}(\{\omega \in \Omega : \lim_{n,m\to\infty} |X_n(\omega) - X_m(\omega)| = 0\}) = 1$$

*or, equivalently*

$$\mathbb{P}(\{\omega \in \Omega : \forall \epsilon > 0, \exists N \in \mathbb{N}, \text{ s.t. } |X_n(\omega) - X_m(\omega)| < \epsilon \text{ whenever } n, m > N\}) = 1$$

*(b) $(X_n)$ converges (to some random variable) in* **mean square if and only if** *$(X_n)$ is a Cauchy sequence in mean square sense. i.e.*

$$\mathbb{E}[X_n^2] < \infty, \forall n$$

*and*

$$\lim_{n,m\to\infty} \mathbb{E}[|X_n - X_m|^2] = 0$$

*(c)* $(X_n)$ *converges* **in probability** *(to some random variable)* **if and only if**

$$\forall \varepsilon > 0, \quad \lim_{n,m\to\infty} \mathbb{P}(\{|X_n - X_m| > \varepsilon\}) = 0$$

**Proposition 3.** *The following statements are equivalent*

1. *The sequence* $X_n \to X$ *in distribution.*
2. $\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)]$ *for* **any bounded continuous** *function* $f$*.*
3. $\lim_{n\to\infty} \phi_{X_n}(u) = \phi_X(u)$

*where* $\phi_X(u) = \mathbb{E}[e^{juX}]$ *is the characteristic function of the random variable* $X$

# Appendix I: Useful Lemmas

**Lemma 1** (Triangular Inequality in $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$)**.** *Let* $X$ *and* $Y$ *be random variables on* $(\Omega, \mathcal{F}, \mathbb{P})$*. Define the* $\mathcal{L}^2$ **norm** $||X||_2$ *as*

$$||X||_2 \coloneqq (\mathbb{E}[|X|^2])^{1/2}$$

*The* $\mathcal{L}^2$ *triangular inequality is*

$$||X + Y||_2 \leq ||X||_2 + ||Y||_2$$

*or equivalently,*

$$\mathbb{E}[|X + Y|^2]^{1/2} \leq \mathbb{E}[|X|^2]^{1/2} + \mathbb{E}[|Y|^2]^{1/2}$$

In fact, the $p$-norm $||\cdot||_p$ of a random variable $X$ is defined as

$$||X||_p \coloneqq (\mathbb{E}[|X|^p])^{1/p}$$

**Lemma 2** (Probability of Infinite Intersection of Events)**.** *Let* $\{A_1, A_2, \cdots, A_n, \cdots\}$ *be a collection of events on* $(\Omega, \mathcal{F}, \mathbb{P})$*. s.t.*

$$\mathbb{P}(A_n) = 1, \quad \forall n \in \mathbb{N}$$

*Then, the* **infinite union** *of the events has probability* $1$*. i.e.*

$$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = 1$$

*Also, the probability of* **infinite intersection** *of events has probability* $1$*. i.e.*

$$\mathbb{P}(\bigcap_{n=1}^{\infty} A_n) = 1$$

# Appendix: Limit Supremum and Infimum of Sets

Let $\{A_1, A_2, \cdots, A_n, \cdots\}$ be a sequence of sets on $\Omega$. The supremum of $(A_n)$ $\sup_{k\geq n} A_k$ is defined as

$$\sup_{k\geq n} A_k = \bigcup_{k=n}^{\infty} A_k$$

which is the union of $A_j$'s from $n$. On the other hand, the infimum of $(A_n)$, $\inf_{k\geq n} A_k$ is

$$\inf_{k\geq n} A_k = \bigcap_{k=n}^{\infty} A_k$$

which is the intersection of $A_k$'s from $n$. As an analogy to the sequence on $\mathbb{R}$, one might think of $\sup_{k\geq n} A_k$ as the **smallest** set covering $\{A_n, A_{n+1}, A_{n+2}, \cdots\}$. On the other hand, $\inf_{k\geq n} A_k$ can be thought of as the **greatest** set that is contained in $\{A_n, A_{n+1}, A_{n+2}, \cdots\}$. i.e. "Least Upper Bound" and "Greatest Lower Bound" for "sets."

**Definition 7** (Limit Supremum)**.** The **Limit Supremum** of $(A_n)$ $\limsup_{n\to\infty} A_n$ is defined as

$$\limsup_{n\to\infty} A_n \coloneqq \inf_{n\geq 1} \sup_{k\geq n} A_k = \lim_{n\to\infty} \sup_{k\geq n} A_k$$

On the other hand, the limit infimum is defined similarly

**Definition 8** (Limit Infimum)**.** The **Limit Infimum** of $(A_n)$ $\liminf_{n\to\infty} A_n$ is defined as

$$\liminf_{n\to\infty} A_n \coloneqq \sup_{n\geq 1} \inf_{k\geq n} A_k = \lim_{n\to\infty} \inf_{k\geq n} A_k$$

The definition of the upper limit and the lower limit of the sets is rather similar to that of a sequence in $\mathbb{R}$. Applying the definition of supremum and infimum of sets, we have

$$\limsup_{n\to\infty} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

and

$$\liminf_{n\to\infty} = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

Recall that for a sequence $(x_n : n \geq 1)$ in $\mathbb{R}$, the limit of $(x_n)$ exists if and only if its upper limit and lower limit exists and equal. Similarly,

**Definition 9** (Limit of Sets)**.** The limit of a sequence of sets $\lim_{n\to\infty} A_n$ exists **if and only if**

$$\limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n \coloneqq \lim_{n\to\infty} A_n$$

### 3.1 Monotone Sequence

**Definition 10** (Monotonicity)**.** A sequence of sets $(A_n)$ is said to be **monotonically increasing** if $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots \subseteq A_n \subseteq \cdots$. If $(A_n)$ is monotonically increasing we write $A_n \uparrow$. On the other hand, if $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots \supseteq A_n \supseteq \cdots$, we $(A_n)$ is **monotonically decreasing** and denote it as $A_n \downarrow$.

**Proposition 4** (Limit of Monotone Sequence)**.** *1. If $A_n \uparrow$, then $\lim_{n\to\infty} A_n$ exists and $\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$.*

*2. If $A_n \downarrow$, then $\lim_{n\to\infty} A_n$ exists and $\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$*

That is, if a sequence of sets is monotonic, then its *limit* always exists.

**Proposition 5** (Properties)**.** *The following are some properties regarding the limit of sets. [1]*

*(a)* $\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$.

# References

[1] S. Resnick: A Probability Path, *Birkhauser*, 1999.