Chapter 2

Convergence concepts

In the following, $\{X_n\}_{n\in I\!N}$ will be a sequence of r.v.s defined over the same probability space $(\Omega, \mathcal{A}, P)$.

Definition 2.1 (Almost sure convergence)
A sequence $\{X_n\}_{n\in I\!N}$ is said to converge almost surely (a.s.) to the r.v. X, represented by $X_n \stackrel{a.s.}{\to} X$, if and only if

$$P\left(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}\right) = 1.$$

Convergence "almost surely" is also called "almost everywhere" (a.e.), "with probability 1" or "strong" convergence.

Remark 2.1 Let us define the set

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

Definition 2.1 has sense only if $\mathcal{C} \in \mathcal{A}$.

Definition 2.2 (Almost sure equality)
We say that two r.v.s X and Y are almost surely equal ($X = Y$ a.s.) iff $P(X \neq Y) = 0$.

Theorem 2.1 (Uniqueness of limiting r.v.)

$$X_n \stackrel{a.s.}{\to} X \Rightarrow \left[X_n \stackrel{a.s.}{\to} Y \Leftrightarrow X = Y \text{ a.s.}\right]$$

Theorem 2.2 (Another characterization of a.s convergence)

$X_n \stackrel{a.s.}{\to} X \Leftrightarrow \lim_{n\to\infty} P\left(\{\omega \in \Omega : |X_m(\omega) - X(\omega)| \leq \epsilon, \forall m \geq n\}\right) = 1.$

Example 2.1 Consider the probability space associated with tossing a coin, $(\Omega, \mathcal{A}, P)$ with $\Omega = \{H, T\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ and $P(\{H\}) = P(\{T\}) = 1/2$. consider also the sequence of r.v.s $\{X_n\}_{n\in I\!N}$, defined by

$$X_n(\{H\}) = 1/n, \quad X_n(\{T\}) = -1/n, \quad n \in I\!N.$$

Observe that $X_n \stackrel{a.s.}{\to} X \equiv 0$.

Definition 2.3 (Convergence in probability)
A sequence $\{X_n\}_{n\in I\!N}$ is said to converge in probability to the r.v. X, represented by $X_n \stackrel{P}{\to} X$, if and only if

$$\forall \epsilon > 0, \quad \lim_{n\to\infty} P\left(\{\omega \in \Omega : \ |X_n(\omega) - X(\omega)| \geq \epsilon\}\right) = 0.$$

Theorem 2.3 (Uniqueness of limiting r.v.)

$$X_n \stackrel{P}{\to} X \Rightarrow \left[X_n \stackrel{P}{\to} Y \Leftrightarrow X = Y \text{ a.s.}\right]$$

Theorem 2.4 (a.s convergence implies convergence in prob.)

$$X_n \stackrel{a.s.}{\to} X \Rightarrow X_n \stackrel{P}{\to} X.$$

Example 2.2 (Counterexample: reciprocal of Theorem 2.4)
We are going to build a sequence of r.v.s that converge in probability to a r.v. X but does not converge a.s. Consider the probability space $(\Omega, \mathcal{A}, P)$, with $\Omega = (0, 1]$, $\mathcal{A} = \mathcal{B}_{(0,1]}$ and $P = U(0, 1]$.

Consider the sequence of r.v.s:

$$\begin{array}{lll} X_1 = \mathbf{1}_{(0,1]}, & X_2 = \mathbf{1}_{(0,1/2]}, & X_3 = \mathbf{1}_{(1/2,1]} \\ X_4 = \mathbf{1}_{(0,1/4]}, & X_5 = \mathbf{1}_{(1/4,1/2]}, & X_6 = \mathbf{1}_{(1/2,3/4]} \\ X_7 = \mathbf{1}_{(3/4,1]}, & X_8 = \mathbf{1}_{(0,1/8]}, & X_9 = \mathbf{1}_{(1/8,1/4]}, \ldots \end{array}$$

Observe that $X_n \stackrel{P}{\to} X$ but $X_n(\omega)$ does not converge for any $\omega \in (0,1]$, so that $X_n \stackrel{a.s}{\nrightarrow} X$.

Theorem 2.5 (Partial converse of Th. 2.4)

$X_n \stackrel{a.s.}{\to} X \Rightarrow$ There is a subsequence of $\{X_n\}_{n\in I\!N}$ that converges a.s to X.

Definition 2.4 (Convergence in r-th mean)
For any real number $r > 0$, we say that $\{X_n\}_{n\in I\!N}$ converges in r-th mean to a r.v. X ($X_n \stackrel{r}{\to} X$), iff

$$E\left(|X_n - X|^r\right) \underset{n\to\infty}{\longrightarrow} 0.$$

In particular, for $r = 2$ we say that $\{X_n\}_{n\in I\!N}$ converges to X in quadratic mean, and we represent it by $X_n \stackrel{q.m.}{\to} X$.

Theorem 2.6 (Uniqueness of limiting r.v.)

$$X_n \stackrel{q.m.}{\to} X \Rightarrow \left[X_n \stackrel{q.m.}{\to} Y \Leftrightarrow X = Y \text{ a.s.}\right]$$

Theorem 2.7 (Convergence in q.m. implies convergence in prob.)

$$X_n \stackrel{q.m.}{\to} X \Rightarrow X_n \stackrel{P}{\to} X.$$

Example 2.3 (Counterexample: reciprocal of Theorem 2.7)
Consider the situation of Example 2.2. We have seen that $X_n \stackrel{a.s.}{\nrightarrow} X$ and it is easy to see that $X_n \stackrel{q.m.}{\to} X \equiv 0$.

Example 2.4 (Counterexample: reciprocal of Theorem 2.7)
Consider the probability space $(\Omega, \mathcal{A}, P)$ with $\Omega = (0,1)$, $\mathcal{A} = \mathcal{B}_{(0,1)}$ and $P = U(0,1)$. Check that $X_n \stackrel{a.s.}{\to} X$ but $X_n \stackrel{q.m.}{\not\to} X \equiv 0$ for the following sequences of r.v.s:

(a) $X_n(\Omega) = 0$ if $\omega \in (0, 1-1/n)$ and $X_n(\Omega) = n$ otherwise.

(b) $X_n(\Omega) = 2^n$ if $\omega \in [0, 1/n)$ and $X_n(\Omega) = 0$ otherwise.

Theorem 2.8 (Fatou-Lebesgue's Lemma)
If $X_n \stackrel{a.s.}{\to} X$ and $X_n \geq Y$, $\forall n \in I\!N$ and for a r.v. Y with $E|Y| < \infty$, then

$$\liminf_{n\to\infty} E(X_n) \geq E(X).$$

Example 2.5 (Fatou-Lebesgue's Lemma)
Consider the sequence of r.v.s of Example 2.4 (b). Check the Fatou-Lebesgue's Theorem.

Theorem 2.9 (Monotone Convergence Theorem for r.v.s)
If $0 \leq X_1 \leq X_2 \leq \cdots$ and $X_n \stackrel{a.s.}{\to} X$, then

$$E(X_n) \underset{n\to\infty}{\to} E(X).$$

Theorem 2.10 (Dominated Convergence Theorem for r.v.s)
If $X_n \stackrel{a.s.}{\to} X$ and $|X_n| \leq Y$ for a r.v. Y with $E|Y| < \infty$, then

$$E(X_n) \underset{n\to\infty}{\to} E(X).$$

Definition 2.5 (Convergence in law/distribution)
Let $\{F_n\}_{n\in I\!N}$ be a sequence of d.f.s defined over $(I\!R, \mathcal{B})$, We say

that $\{F_n\}_{n\in I\!N}$ converges in law/distribution to a d.f. F ($F_n \stackrel{d}{\to} F$) iff

$$\lim_{n\to\infty} F_n(x) = F(x), \quad \forall x \in \mathcal{C}(F),$$

where $\mathcal{C}(F)$ is the set of points where F is continuous.

Definition 2.6 (Convergence in law of r.v.s)
A sequence $\{X_n\}_{n\in I\!N}$ of r.v.s with respective d.f.s $\{F_n\}_{n\in I\!N}$ is said to converge in law/distribution to a r.v. X with d.f. F ($X_n \stackrel{d}{\to} X$) iff $F_n \stackrel{d}{\to} F$.

Proposition 2.1 (Uniqueness of limiting r.v.)
If $F_n \stackrel{d}{\to} F$ and $F_n \stackrel{d}{\to} F^*$, then $F = F^*$.

Remark 2.2 (Knowledge of the d.f. on a dense set sufficient)
The proof of Proposition 2.1 implies that F is completely determined when it is given for a dense subset of $I\!R$.

Theorem 2.11 (Convergence in prob. implies convergence in law)

$$X_n \stackrel{P}{\to} X \Rightarrow X_n \stackrel{d}{\to} X.$$

Example 2.6 (Convergence in law)
Consider the sequence of degenerated r.v.s given by $P(X_n = 1/n) = 1$, $n = 1, 2, \ldots$. Consider also the degenerated r.v. at zero, $P(X = 0) = 1$. Check that $X_n \stackrel{d}{\to} X$.

Example 2.7 (Counterexample: Convergence in law does not imply convergence in prob.)
Consider the probability space $(\Omega, \mathcal{A}, P)$ with $\Omega = (0, 1)$, $\mathcal{A} =$

$\mathcal{B}_{(0,1)}$ and $P = U(0,1)$. Consider the sequence of r.v.s defined over $(\Omega, \mathcal{A}, P)$ by

$$X_{2n-1}(\omega) = \begin{cases} 1, & \text{if } \omega \in (0,1/2) \\ 0, & \text{if } \omega \in [1/2,1) \end{cases}, \quad X_{2n}(\omega) = \begin{cases} 0, & \text{if } \omega \in (0,1/2) \\ 1, & \text{if } \omega \in [1/2,1) \end{cases}$$

Check that $X_n \overset{P}{\nrightarrow} X$ for any r.v. X but $X_n \overset{d}{\rightarrow} X$, where $X \overset{d}{=} \text{Bern}(1/2)$.

Theorem 2.12 (Convergence in law to a constant implies convergence in prob.)

$$X_n \overset{d}{\nrightarrow} c \in \mathbb{R} \Rightarrow X_n \overset{P}{\rightarrow} c.$$

The following example shows that in the definition of convergence in law as the pointwise limit at only the continuity points is convenient, we cannot define it at all points.

Example 2.8 (Conv. in law does not imply convergence in prob.)
Consider the sequence of i.i.d. r.v.s $\{X_n\}_{n\in\mathbb{N}}$, with $X_n \overset{d}{=} N(0,1)$, $\forall n \in \mathbb{N}$.

Theorem 2.13 (Borel-Cantelli's Lemma)
Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of sets from Ω and let $(\Omega, \mathcal{A}, P)$ be a probability space. It holds that:

(a) If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(\{\omega \in \Omega : \omega \in A_n \text{ i.o.}\}) = 0$;

(b) Reciprocally, if $\{A_n\}_{n=1}^{\infty}$ are independent and $\sum_{n=1}^{\infty} P(A_n) = \infty$, then $P(\{\omega \in \Omega : \omega \in A_n \text{ i.o.}\}) = 1$.

Example 2.9 (Borel-Cantelli's Lemma)
Let $\{X_n\}_{n\in I\!N}$ be a sequence of r.v.s with $X_n \stackrel{d}{=} \text{Bern}(p_n)$ with $p_n = 1/n^2$, $\forall n \in I\!N$. Obtain the probability of $X_n = 1$ i.o.

Remark 2.3 (Another characterization of the Borel-Cantelli's Lemma)
The Borel-Cantelli's Lemma is useful in problems related to the a.s. convergence. In fact, another characterization of the a.s. convergence is the following:

$$P(|X_n - X| > \epsilon \text{ i.o.}) = 0, \quad \forall \epsilon > 0.$$

Lemma 2.1 (Boundedness of measurable functions)
Let X be a r.v. defined from a probability space $(\Omega, \mathcal{A}, P)$ on $(I\!R, \mathcal{B})$. Let g be a bounded measurable function from $(I\!R, \mathcal{B})$ on $(I\!R, \mathcal{B})$, with $P(X \in \mathcal{C}(g)) = 1$. Then, it holds that $\forall \epsilon > 0$, there exist two continuous and bounded functions f and g such that $f \leq g \leq h$ pointwise and $E\left[h(X) - f(X)\right]$.

Definition 2.7 (Function that vanishes out of a compact set)
Let g be a real function defined on $I\!R$. We say that g vanishes out of a compact set if and only if there is a compact set $C \subset I\!R^d$ such that $g(X) = 0$, $\forall x \notin C$.

Theorem 2.14 (Relation between conv. in law and conv. of expectations)
The following conditions are equivalent:

(a) $X_n \stackrel{d}{\to} X$;

(b) $E\left[g(X_n)\right] \to E\left[g(X)\right]$, for each continuous function g that vanishes out of a compact set;

(c) $E[g(X_n)] \to E[g(X)]$, for each continuous and bounded function g;

(d) $E[g(X_n)] \to E[g(X)]$, for all bounded and measurable function g with $P(X \in \mathcal{C}(g)) = 1$, where $\mathcal{C}(g)$ is the continuity set of g.

Implications (a)$\Rightarrow$ (b), (a)$\Rightarrow$ (c) and (a)$\Rightarrow$ (d) are known as Helly-Bray's Theorem.

Example 2.10 (Boundedness is necessary in (c) and (d))
Consider the function $g(x) = x$ and the sequence of r.v.s defined by $P(X_n = n) = 1/n$ and $P(X_n = 0) = 1 - 1/n$. Check that $X_n \xrightarrow{d} X$, with $X \equiv 0$ but $E[g(X_n)] \nrightarrow E[g(X)]$. Why does this happen?

Example 2.11 (Continuity is necessary)
Consider the function

$$g(x) = \begin{cases} 0, & x = 0 \\ 1, & x > 0 \end{cases}.$$

Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of r.v.s with $P(X_n = 1/n) = 1$, $\forall n \in \mathbb{N}$. See that $X_n \xrightarrow{d} X$, with $X \equiv 0$ but $E[g(X_n)] \nrightarrow E[g(X)]$. Why does this happen?

Sometimes it is easier to calculate the characteristic function $\varphi_n(t)$ than the d.f. $F_n(t)$ if a sequence of r.v.s $\{X_n\}_{n\in\mathbb{N}}$. Then, the convergence in law of $\{X_n\}_{n\in\mathbb{N}}$ will be studied through the pointwise convergence of $\varphi_n(t)$ by using the following theorem.

Theorem 2.15 (Continuity Theorem)

$$X_n \xrightarrow{d} X \leftrightarrow \varphi_{X_n}(t) \underset{n\to\infty}{\to} \varphi_X(t), \quad \forall t \in \mathbb{R}.$$

Problems

1. Consider the sequence of r.v.s given by $P(X_n = n) = 1/n$ and $P(X_n = 0) = 1 - 1/n$, $n = 1, 2, \ldots$. Consider also the degenerated r.v. at zero, $P(X = 0) = 1$. Check that $X_n \xrightarrow{d} X$.

2. Consider the sequence of d.f.s defined as

$$F_n(x) = \begin{cases} 0, & x < 0 \\ nx, & 0 \leq x < 1/n, \\ 1, & x \geq 1/n \end{cases}$$

   Prove that the pointwise limit function is not a d.f. but $F_n \underset{n\to\infty}{\rightarrow} F$, where F is the d.f. given by

$$F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

3. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of r.v.s with $X_n \stackrel{d}{=} \text{Beta}(1/n, 1/n)$ and the r.v. $X \stackrel{d}{=} \text{Bin}(1, 1/2)$. Prove that $X_n \xrightarrow{d} X$. Is that true also if $X_n \stackrel{d}{=} \text{Beta}(\alpha/n, \beta/n)$?

4. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of r.v.s, where X_n is uniformly distributed in the set $\{1/n, 2/n, 3/n, \ldots, 1\}$. Prove that $X_n \xrightarrow{d} X$, where $X \stackrel{d}{=} U(0, 1)$. Does $X_n \xrightarrow{P} X$?

5. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of r.v.s and X another r.v. With the help of Hölder's inequality, prove that:

   (a) If $0 < r < s$ and $E|X_n - X|^s < \infty$, then $E|X_n - X|^r < \infty$.

   (b) If $0 < r < s$ and $X_n \xrightarrow{s} X$, then $X_n \xrightarrow{r} X$.

6. Give and example of a sequence of r.v.s such that

$$E|X_n| \underset{n\to\infty}{\to} 0 \quad \text{and} \quad E(X_n)^2 \underset{n\to\infty}{\to} 0.$$

7. Let μ be a constant. Show that

$$X_n \overset{q.m.}{\to} \mu \leftrightarrow E(X_n) \underset{n\to\infty}{\to} \mu \quad \text{and} \quad V(X_n) \underset{n\to\infty}{\to} 0.$$