Appendix A: Convergence Concepts for Random Variables

A.1 Convergence Definitions

Recall that a real-valued rv $X$ is a function $X : \Omega \rightarrow \mathbb{R}$. Given that a rv is a function, there are many different mechanisms for determining whether a sequence of rvs/functions $(X_n : n \geq 0)$ converges to a limit rv/function $X_\infty$.

Almost Sure Convergence:

We say that $(X_n : n \geq 1)$ converges almost surely to $X_\infty$ if $P(A) = 1$, where

$$A = \{ \omega : X_n(\omega) \rightarrow X_\infty(\omega) \text{ as } n \rightarrow \infty \}$$

and write $X_n \rightarrow X_\infty$ a.s. as $n \rightarrow \infty$ when this convergence holds. This type of convergence is equivalently called: convergence with probability one (written $X_n \rightarrow X_\infty$ w.p. 1 as $n \rightarrow \infty$); convergence almost everywhere (written $X_n \rightarrow X_\infty$ a.e. as $n \rightarrow \infty$); convergence almost certainly (written $X_n \rightarrow X_\infty$ a.c. as $n \rightarrow \infty$).

Remark A.1.1 Note that the event $A$ depends on the infinite-dimensional joint distribution of $(X_n : 1 \leq n \leq \infty)$. As a result, rigorous discussion of probability assignments to events like $A$ needed to await the development of measure-theoretic probability in the early 20th century.

Convergence in $p$’th Mean:

We say that $(X_n : n \geq 1)$ converges in $p$’th mean (for $p > 0$) to $X_\infty$ if $E|X_n|^p < \infty$ for $n \geq 1$ and $\|X_n - X_\infty\|_p \rightarrow 0$ as $n \rightarrow \infty$, where

$$\|Y\|_p = E^{1/p}|Y|^p$$

for $E|Y|^p < \infty$. When this convergence holds, we write $X_n \xrightarrow{L^p} X_\infty$ as $n \rightarrow \infty$.

Convergence in Probability

We say that $(X_n : n \geq 1)$ converges in probability to $X_\infty$ if, for each $\epsilon > 0$,

$$P(|X_n - X_\infty| > \epsilon) \rightarrow 0$$

as $n \rightarrow \infty$, in which case we write $X_n \xrightarrow{P} X_\infty$ as $n \rightarrow \infty$.

Remark A.1.2 Convergence in probability and convergence in $p$’th mean is a statement about the joint distribution of the two rv’s $X_n$ and $X_\infty$ for $n$ large.

Remark A.1.3 Convergence in $p$’th mean implies convergence in probability, since Markov’s inequality implies that

$$P(|X_n - X_\infty| > \epsilon) \leq \frac{E|X_n - X_\infty|^p}{\epsilon^p}.$$
Exercise A.1.1 Prove that almost sure convergence implies convergence in probability.

Weak Convergence

We say that \((X_n : n \geq 1)\) converges weakly to \(X_\infty\) if \(X_\infty\) is a finite-valued rv for which

\[ Ef(X_n) \to Ef(X_\infty) \]

as \(n \to \infty\) for each bounded and continuous function \(f : \mathbb{R} \to \mathbb{R}\), in which case we write

\[ X_n \Rightarrow X_\infty \]

as \(n \to \infty\). Weak convergence is equivalently called “convergence in distribution”.

Remark A.1.4 Weak convergence is a statement about the distribution of \(X_n\) when \(n\) is large.

Remark A.1.5 Weak convergence can be equivalently formulated as: \(X_n \Rightarrow X_\infty\) as \(n \to \infty\) if and only if \(X_\infty\) is a finite-valued rv for which

\[ P(X_n \leq x) \to P(X_\infty \leq x) \]

as \(n \to \infty\) whenever \(x\) is a continuity point of \(P(X_\infty \leq \cdot)\).

Exercise A.1.2 Prove that \(X_n \overset{p}{\to} X_\infty\) as \(n \to \infty\) implies that \(X_n \Rightarrow X_\infty\) as \(n \to \infty\).

Remark A.1.6 Convergence in probability does not imply almost sure convergence. Note that if \(Y = (Y_n : n \geq 0)\) is a nearest neighbor symmetric random walk on \(\mathbb{Z}\), then the recurrence of \(Y\) implies that \(X_n \triangleq I(Y_n = 0) = 1\) infinitely often a.s. so \(X_n \not\to 0\) a.s. as \(n \to \infty\). But \(P(X_n = 0) = P(Y_n \neq 0) \to 1\) as \(n \to \infty\), so \(X_n \overset{p}{\to} 0\) as \(n \to \infty\).

The following diagram makes clear the relationship between these convergence concepts:

```
Almost Sure Convergence
   /               /
  /                /
Convergence in p'th mean
   /                /
  /                /
Convergence in Probability
   /                /
  /                /
Weak Convergence
```

A.2 Basic Facts on Almost Sure Convergence

Fact 1: Suppose that \(X_n \to X_\infty\) a.s. as \(n \to \infty\) and \(Y_n \to Y_\infty\) a.s. as \(n \to \infty\), where \(X_\infty\) and \(Y_\infty\) are finite-valued. If \(f : \mathbb{R}^2 \to \mathbb{R}\) is continuous, then \(h(X_n, Y_n) \to h(X_\infty, Y_\infty)\) a.s. as \(n \to \infty\).
APPENDIX A: CONVERGENCE CONCEPTS FOR RANDOM VARIABLES

Fact 2: Suppose that $X_n \to X_\infty$ a.s. as $n \to \infty$ and $(T_n : n \geq 1)$ is a sequence of $\mathbb{Z}_+\text{-valued rv's}$ for which $T_n \to \infty$ a.s. as $n \to \infty$. Then,

$$X_{T_n} \to X_\infty \text{ a.s.}$$

as $n \to \infty$.

A.3 Basic Facts on Convergence in Probability

Fact 1: Suppose that $X_n \xrightarrow{p} X_\infty$, where $X_\infty$ is finite-valued. If $h : \mathbb{R} \to \mathbb{R}$ is continuous, then $h(X_n) \xrightarrow{p} h(X_\infty)$ as $n \to \infty$.

Remark A.3.1 If $X_n \xrightarrow{p} X_\infty$ and $Y_n \xrightarrow{p} Y_\infty$ as $n \to \infty$ and $((X_n, Y_n) : 1 \leq n \leq \infty)$ are all defined on a common sample space (i.e. are jointly distributed), then $(X_n, Y_n) \xrightarrow{p} (X_\infty, Y_\infty)$ as $n \to \infty$.

Fact 2: Suppose that $X_n \to X_\infty$ a.s. as $n \to \infty$ and $T_n \xrightarrow{p} \infty$ as $n \to \infty$. Then, $X_{T_n} \xrightarrow{p} X_\infty$ as $n \to \infty$.

Remark A.3.2 If $X_n \xrightarrow{p} X_\infty$ and $T_n \xrightarrow{p} \infty$ as $n \to \infty$, it is not always the case that $X_{T_n} \xrightarrow{p} X_\infty$ as $n \to \infty$.

A.4 Basic Facts on Weak Convergence

Fact 1: Suppose that $X_n \Rightarrow X_\infty$ as $n \to \infty$. If $h : \mathbb{R} \to \mathbb{R}$ is continuous, then $h(X_n) \Rightarrow h(X_\infty)$ as $n \to \infty$. In fact, if $P(X_\infty \in D_h) = 0$ (where $D_h = \{x \in \mathbb{R} : h(\cdot) \text{ is discontinuous at } x\}$), then $h(X_n) \Rightarrow h(X_\infty)$ as $n \to \infty$.

Remark A.4.1 The above result is called the continuous mapping principle.

Fact 2: Suppose that $X_n \Rightarrow X_\infty$ as $n \to \infty$ and $Y_n \xrightarrow{p} c$ as $n \to \infty$, where $c$ is deterministic. If $h : \mathbb{R}^2 \to \mathbb{R}$ is such that $P((X_\infty, c) \in D_h) = 0$, then $h(X_n, Y_n) \Rightarrow h(X_\infty, c)$ as $n \to \infty$.

Remark A.4.2 The above fact implies that if $X_n \Rightarrow X_\infty$ and $Y_n \xrightarrow{p} c$, then $X_n + Y_n \Rightarrow X_\infty + c$ and $X_n Y_n \Rightarrow cX_\infty$ as $n \to \infty$.

Remark A.4.3 If $X_n \Rightarrow X_\infty$ and $Y_n \xrightarrow{p} Y_\infty$ as $n \to \infty$, it is not always the case that $h(X_n, Y_n) \Rightarrow h(X_\infty, Y_\infty)$ (even if $h : \mathbb{R}^2 \to \mathbb{R}$ is continuous everywhere).

Remark A.4.4 If $X_n \Rightarrow X_\infty$ as $n \to \infty$ and $T_n \xrightarrow{p} \infty$ as $n \to \infty$, it is not always the case that $X_{T_n} \Rightarrow X_\infty$ as $n \to \infty$. 

3