

Introduction

Remark 1.1.1. This document considers stochastic order notation, limits of sets, different modes of stochastic: almost sure, in L^p , in probability, and in distribution; and uniform integrability. Throughout, assume that (\mathbf{X}_n) is a sequence of scalar or vector-valued random variables with candidate limit \mathbf{X} , and that the (\mathbf{X}_n) and \mathbf{X} are defined on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$. ◆

Order Notation

2.1 Definitions

Definition 1.2.1. Let α_n and β_n denote sequences of real numbers, then $\alpha_n = \mathcal{O}(\beta_n)$ if there exist a bound $M \in \mathbb{R}^+$ and a threshold $\nu \in \mathbb{N}$ s.t. for $n \geq \nu$: $|\alpha_n| \leq M|\beta_n|$. ■

Definition 1.2.2. A sequence of random variables (\mathbf{X}_n) is **bounded in probability**, expressed $\mathbf{X}_n = \mathcal{O}_p(1)$, if for $\forall \epsilon > 0$ there $\exists(M_\epsilon, \nu_\epsilon)$ s.t. $n \geq \nu_\epsilon$ implies:

$$P(\|\mathbf{X}_n\| > M_\epsilon) < \epsilon.$$

If the sequence of random variables (\mathbf{X}_n) is *bounded in probability*, then the corresponding sequence (F_n) of probabilities measures is described as *uniformly tight*. ■

Definition 1.2.3. Let α_n and β_n denote sequences of real numbers, then $\alpha_n = o(\beta_n)$ if for $\forall \epsilon > 0$ there $\exists(\nu_\epsilon)$ s.t. for $n \geq \nu_\epsilon$: $|\alpha_n| \leq \epsilon|\beta_n|$. ■

Definition 1.2.4. A sequence of random variables (\mathbf{X}_n) **converges in probability** to zero, expressed $\mathbf{X}_n = o_p(1)$, if for $\forall \epsilon > 0$:

$$\lim_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_n\| > \epsilon) = 0.$$

Convergence of the sequence (\mathbf{X}_n) in probability to zero requires that for $\forall \epsilon, \delta > 0$ there exists ν_δ s.t. when $n \geq \nu_\delta$ the probability $P(\|\mathbf{X}_n\| > \epsilon) < \delta$. ■

Definition 1.2.5. Suppose (\mathbf{X}_n) is a sequence of random variables, and that $(\alpha_n) \in \mathbb{R}^+$ is a sequence of positive constants.

- i. $\mathbf{X}_n = o_p(\alpha_n) \iff \alpha_n^{-1} \mathbf{X}_n = o_p(1)$.
 - ii. $\mathbf{X}_n = \mathcal{O}_p(\alpha_n) \iff \alpha_n^{-1} \mathbf{X}_n = \mathcal{O}_p(1)$.
-

2.2 Properties

Proposition 1.2.1. If \mathbf{X}_n converges in probability to zero, then \mathbf{X}_n is bounded in probability: $\mathbf{X}_n = o_p(1) \implies \mathbf{X}_n = \mathcal{O}_p(1)$. \blacklozenge

Proof. Fix $\epsilon > 0$, then by the definition of convergence in probability, for $\forall \delta > 0$ there $\exists \nu_\delta \in \mathbb{N}$ s.t. when $n \geq \nu_\delta$, $\mathbb{P}(\|\mathbf{X}_n\| > \epsilon) < \delta$. \blacksquare

Proposition 1.2.2 (Sub-additivity). Suppose $\{\mathbf{X}_i\}$ is a finite collection of random variables, not necessarily independent nor identically distributed. Then:

$$\mathbb{P}\left(\left\|\sum_{i=1}^n \mathbf{X}_i\right\|_2 > \epsilon\right) \leq \sum_{i=1}^n \mathbb{P}(\|\mathbf{X}_i\| > \epsilon/n) \quad (1.2.1)$$

\blacklozenge

Proof. If $\left\|\sum_{i=1}^n \mathbf{X}_i\right\|_2 > \epsilon$, then at least one $\|\mathbf{X}_i\| > \epsilon/n$, for suppose not, then:

$$\left\|\sum_{i=1}^n \mathbf{X}_i\right\|_2 \leq \sum_{i=1}^n \|\mathbf{X}_i\| \leq \sum_{i=1}^n \frac{\epsilon}{n} = \epsilon,$$

which leads to a contradiction. Expressed In terms of events:

$$\left\{\omega \in \Omega : \left\|\sum_{i=1}^n \mathbf{X}_i\right\|_2 > \epsilon\right\} \subset \bigcup_{i=1}^n \{\omega \in \Omega : \|\mathbf{X}_i\| > \epsilon/n\}.$$

By sub-additivity of the probability measure:

$$\mathbb{P}\left(\left\|\sum_{i=1}^n \mathbf{X}_i\right\|_2 > \epsilon\right) \leq \mathbb{P}\left(\bigcup_{i=1}^n \{\omega \in \Omega : \|\mathbf{X}_i\| > \epsilon/n\}\right) \leq \sum_{i=1}^n \mathbb{P}(\|\mathbf{X}_i\| > \epsilon/n).$$

\blacksquare

Proposition 1.2.3. If $X_n = \mathcal{O}_p(1)$ and $Y_n = \mathcal{O}_p(1)$, then:

i. $X_n + Y_n = \mathcal{O}_p(1)$.

ii. $X_n Y_n = \mathcal{O}_p(1)$.

\blacklozenge

Proof. i. Since $X_n = \mathcal{O}_p(1)$ and $Y_n = \mathcal{O}_p(1)$. Fix $\epsilon > 0$, there $\exists(M_X, \nu_X)$ and $\exists(M_Y, \nu_Y)$ s.t. when $n \geq \nu_X$, $\mathbb{P}(\|X_n\| > M_X) < \epsilon/2$ and when $n \geq \nu_Y$, $\mathbb{P}(\|Y_n\| > M_Y) < \epsilon/2$. Set $M = \max(M_X, M_Y)$, then:

$$\mathbb{P}(\|X_n + Y_n\| > 2M) \leq \mathbb{P}(\|X_n\| > M) + \mathbb{P}(\|Y_n\| > M) \leq \epsilon/2 + \epsilon/2 = \epsilon.$$

ii.

$$\begin{aligned}\mathbb{P}(\|X_n Y_n\| > M_X M_Y) &\leq \mathbb{P}(\|X_n\| > M_X \cup \|Y_n\| > M_Y) \\ &\leq \mathbb{P}(\|X_n\| > M_X) + \mathbb{P}(\|Y_n\| > M_Y) \leq \epsilon/2 + \epsilon/2 = \epsilon.\end{aligned}$$

■

Proposition 1.2.4. If $X_n = o_p(1)$ and $Y_n = \mathcal{O}_p(1)$, then:

i. $X_n + Y_n = \mathcal{O}_p(1)$.

ii. $X_n Y_n = o_p(1)$.

◆

Proof. i. Since $X_n = o_p(1) \implies X_n = \mathcal{O}_p(1)$, (i.) follows from the last proposition.

ii. Fix $\epsilon > 0$. For $\forall(M, \delta)$,

$$\begin{aligned}\mathbb{P}(\|X_n Y_n\| > M) &= \mathbb{P}(\|X_n Y_n\| > \delta \cap \|Y_n\| > M) + \mathbb{P}(\|X_n Y_n\| \leq \delta \cap \|Y_n\| \leq M) \\ &\leq \mathbb{P}(\|Y_n\| > M) + \mathbb{P}(\|X_n\| > \delta/M)\end{aligned}$$

Since $Y_n = \mathcal{O}_p(1)$, for any $\epsilon > 0$ there $\exists M_\epsilon$ s.t. when $n \geq \nu_\epsilon$, $\mathbb{P}(\|Y_n\| > M_\epsilon) < \epsilon$. Moreover, since $X_n = o_p(1)$, $\lim_{n \rightarrow \infty} \mathbb{P}(\|X_n\| > \delta/M) = 0$. Thus:

$$\lim_{n \rightarrow \infty} \mathbb{P}(\|X_n Y_n\| > M_\epsilon) \leq \epsilon + \lim_{n \rightarrow \infty} \mathbb{P}(\|X_n\| > \delta/M) = \epsilon.$$

Since $\epsilon > 0$ was arbitrary, conclude $X_n Y_n = o_p(1)$.

■

Theorem 1.2.1. Suppose $\mathbf{X}_n = o_p(\alpha_n)$ and $\mathbf{Y}_n = o_p(\beta_n)$, then:

i. $\mathbf{X}_n + \mathbf{Y}_n = o_p\{\max(\alpha_n, \beta_n)\}$.

ii. $\mathbf{X}_n \mathbf{Y}_n = o_p(\alpha_n \beta_n)$.

iii. $\|\mathbf{X}_n\|^r = o_p(\alpha_n^r)$ where $r > 0$.

□

Proof. i. If $\|\mathbf{X}_n + \mathbf{Y}_n\| / \max(\alpha_n, \beta_n) > \epsilon$, then either:

$$\frac{\|\mathbf{X}_n\|}{\alpha_n} > \frac{\epsilon}{2} \vee \frac{\|\mathbf{Y}_n\|}{\beta_n} > \frac{\epsilon}{2}.$$

By subadditivity:

$$\lim_{n \rightarrow \infty} \mathbb{P}\left\{\frac{\|\mathbf{X}_n + \mathbf{Y}_n\|}{\max(\alpha_n, \beta_n)} > \epsilon\right\} \leq \lim_{n \rightarrow \infty} \mathbb{P}\left(\alpha_n^{-1} \|\mathbf{X}_n\| > \frac{\epsilon}{2}\right) + \mathbb{P}\left(\beta_n^{-1} \|\mathbf{Y}_n\| > \frac{\epsilon}{2}\right) = 0.$$

ii. If $\|\mathbf{X}_n \mathbf{Y}_n\|/(\alpha_n \beta_n) > \epsilon$, then either:

$$\{\alpha_n^{-1} \|\mathbf{X}_n\| \leq 1, \beta_n^{-1} \|\mathbf{Y}_n\| > \epsilon\} \cup \{\alpha_n^{-1} \|\mathbf{X}_n\| > 1, (\alpha_n \beta_n)^{-1} \|\mathbf{X}_n \mathbf{Y}_n\| > \epsilon\}.$$

By subadditivity:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\|\mathbf{X}_n \mathbf{Y}_n\|}{\alpha_n \beta_n} > \epsilon \right) \leq \lim_{n \rightarrow \infty} \mathbb{P} (\beta_n^{-1} \|\mathbf{Y}_n\| > \epsilon) + \mathbb{P} (\alpha_n^{-1} \|\mathbf{X}_n\| > 1) = 0.$$

iii.

$$\lim_{n \rightarrow \infty} \mathbb{P} (\alpha_n^{-r} \|\mathbf{X}_n\|^r > \epsilon) = \lim_{n \rightarrow \infty} \mathbb{P} (\alpha_n^{-1} \|\mathbf{X}_n\| > \epsilon^{1/r}) = 0.$$

■

Proposition 1.2.5.

$$\mathbf{X}_n - \mathbf{X} = o_p(1) \iff \|\mathbf{X}_n - \mathbf{X}\| = o_p(1).$$

◆

Proof. Let $\mathbf{Y}_n = \mathbf{X}_n - \mathbf{X}$, then by definition $\mathbf{Y}_n = o_p(1)$ if and only if for $\forall \epsilon > 0$:

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{Y}_n\| > \epsilon) &= 0, \\ \lim_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) &= 0. \end{aligned}$$

Now let $Y_n = \|\mathbf{X}_n - \mathbf{X}\|$, then $Y_n = o_p(1)$ if and only if for $\forall \epsilon > 0$:

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}(|Y_n| > \epsilon) &= 0, \\ \lim_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) &= 0. \end{aligned}$$

Thus, the statements $\mathbf{X}_n - \mathbf{X} = o_p(1)$ and $\|\mathbf{X}_n - \mathbf{X}\| = o_p(1)$ are identical. ■

Proposition 1.2.6. Suppose (\mathbf{X}_n) and (\mathbf{Y}_n) are sequences of random variables. If 1. $\mathbf{X}_n - \mathbf{Y}_n = o_p(1)$ and 2. $\mathbf{Y}_n - \mathbf{Y} = o_p(1)$, then $\mathbf{X}_n - \mathbf{Y} = o_p(1)$. ◆

Proof. By the triangle inequality:

$$\|\mathbf{X}_n - \mathbf{Y}\| \leq \|\mathbf{X}_n - \mathbf{Y}_n\|_2 + \|\mathbf{Y}_n - \mathbf{Y}\| = o_p(1)$$

■

Proposition 1.2.7. Suppose (\mathbf{X}_n) is a sequence of J dimension random variables, and $(\alpha_n) \in \mathbb{R}^+$ is a sequence of positive constants, then:

i. $\mathbf{X}_n = o_p(1) \iff X_{nj} = o_p(1)$ for $j \in \{1, \dots, J\}$.

ii. $\mathbf{X}_n = \mathcal{O}_p(1) \iff X_{nj} = \mathcal{O}_p(1)$ for $j \in \{1, \dots, J\}$.

That is, a sequence of random variables converges in probability to zero, or is bounded in probability, if and only if the components convergence in probability to zero, or are bounded in probability. \blacklozenge

Proof. i. (\implies):

$$|X_{nj} - X_j| \leq \sqrt{\sum_{j=1}^k (X_{nj} - X_j)^2},$$

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_{nj} - X_j| > \epsilon) \leq \lim_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) = 0.$$

(\impliedby):

$$\lim_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) \leq \lim_{n \rightarrow \infty} \sum_{j=1}^J \mathbb{P}(|X_{nj} - X_j| > \epsilon/J) = 0.$$

ii. (\implies) Since $\mathbf{X}_n = \mathcal{O}_p(1)$, for $\forall \epsilon > 0$ there $\exists (M_\epsilon, \nu_\epsilon)$ s.t. when $n \geq \nu_\epsilon$:

$$\mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > M_\epsilon) < \epsilon.$$

Since $\mathbb{P}(|X_{nj} - X_j| > M_\epsilon) \leq \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > M_\epsilon)$, conclude $X_{nj} = \mathcal{O}_p(1)$.

For $\epsilon > 0$ there $\exists M_\epsilon, \nu_\epsilon$ s.t. when $n \geq \nu_\epsilon$:

$$\mathbb{P}(|X_{nj} - X_j| > M_\epsilon) \leq \mathbb{P}\left(\sqrt{\sum_{j=1}^k (X_{nj} - X_j)^2} > M_\epsilon\right) = \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > M_\epsilon) < \epsilon.$$

(\impliedby) Fix $\epsilon > 0$. For each $j \in \{1, \dots, J\}$ there $\exists M_j, \nu_j$ s.t. when $n \geq \nu_j$:

$$\mathbb{P}(|X_{nj} - X_j| > M_j) < \frac{\epsilon}{J}.$$

Set $M = \max_j M_j$ and $\nu = \max_j \nu_j$. Now when $n \geq \nu$:

$$\mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > JM) \leq \sum_{j=1}^J \mathbb{P}(|X_{nj} - X_j| > M) \leq J \cdot \frac{\epsilon}{J} = \epsilon.$$

\blacksquare

Limits of Sets

3.1 Definitions

Definition 1.3.1. Suppose (B_n) is a *decreasing* sequence of measurable sets. The limit $\lim_{n \rightarrow \infty} B_n$ is defined as their intersection:

$$\lim_{n \rightarrow \infty} B_n = \bigcap_{n=1}^{\infty} B_n.$$

$\omega \in \lim_{n \rightarrow \infty} B_n$ if for $\forall n \in \mathbb{N}$, $\omega \in B_n$. ■

Definition 1.3.2. Suppose (C_n) is an *increasing* sequence of measurable sets. The limit $\lim_{n \rightarrow \infty} C_n$ is defined as their union:

$$\lim_{n \rightarrow \infty} C_n = \bigcup_{n=1}^{\infty} C_n.$$

$\omega \in \lim_{n \rightarrow \infty} C_n$ if there exists an $n \in \mathbb{N}$ s.t. $\omega \in C_n$. ■

Definition 1.3.3. Define the *supremum* of sequence of sets as:

$$\sup_{k \geq n} A_k = \bigcup_{k \geq n} A_k.$$

The **limit supremum** of a sequence (A_n) of sets is:

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

Observe that $B_n = \sup_{k \geq n} A_k$ is a *decreasing* sequence of sets, since each consecutive B_n is the union of fewer A_n . ■

Definition 1.3.4. Define the *infimum* of a sequence of sets as:

$$\inf_{k \geq n} A_k = \bigcap_{k \geq n} A_k.$$

The **limit infimum** of a sequence (A_n) of sets is:

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

Observe that $C_n = \inf_{k \geq n} A_k$ is an *increasing* sequence of sets, since each consecutive C_n is the intersection of fewer A_n . ■

Definition 1.3.5. The limit supremum and limit infimum of a sequence of sets (A_n) always exist. If these two sets are equal, then the **limit** exists and is defined as:

$$\lim_{n \rightarrow \infty} A_n \equiv \liminf_{n \rightarrow \infty} A_n = \limsup_{n \rightarrow \infty} A_n.$$

■

3.2 Properties

Proposition 1.3.1. For any sequence of sets,

$$\liminf_{n \rightarrow \infty} A_n \subseteq \limsup_{n \rightarrow \infty} A_n.$$

◆

Proof. Suppose $\omega \in \liminf_{n \rightarrow \infty} A_n$, then there exists an $n \in \mathbb{N}$ s.t. $\omega \in A_k$ for $\forall k \geq n$. That is, ω belongs to every A_k far enough out in the sequence. Thus, for any $n \in \mathbb{N}$, there exists $k \geq n$ s.t. $\omega \in A_k$. Conclude that $\omega \in \limsup_{n \rightarrow \infty} A_n$. ■

Remark 1.3.1. Since the limit infimum is always a subset of the limit supremum, to prove the limit $\lim_{n \rightarrow \infty} A_n$ exists, it suffices to prove that the limit supremum is a subset of the limit infimum. ◆

Proposition 1.3.2. Suppose $C_n \rightarrow C$ is an increasing sequence of measurable sets on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then:

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} C_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(C_n).$$

◆

Proof. Define the sequence of *disjoint* sets $D_1 = C_1$, and $D_k = C_k - C_{k-1}$ for $k \geq 2$. Clearly $C_n = \bigcup_{k=1}^n D_k$. By finite additivity of the probability measure:

$$\mathbb{P}(C_n) = \mathbb{P}\left(\bigcup_{k=1}^n D_k\right) = \sum_{k=1}^n \mathbb{P}(D_k).$$

Note too that since the C_n are increasing:

$$\lim_{n \rightarrow \infty} C_n = \bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^n D_k = \bigcup_{k=1}^{\infty} D_k.$$

By σ -additivity of the probability measure \mathbb{P} :

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} C_n\right) = \mathbb{P}\left(\bigcup_{k=1}^{\infty} D_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(D_k) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{P}(D_k) = \lim_{n \rightarrow \infty} \mathbb{P}(C_n).$$

■

Corollary 1.3.1. Suppose $B_n \rightarrow B$ is a decreasing sequence of measurable sets, then:

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} B_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(B_n).$$

♣

Proof. Since (B_n) is decreasing, the sequence of complements (B_n^c) is necessarily increasing, thus:

$$\begin{aligned}\mathbb{P}\left(\lim_{n \rightarrow \infty} B_n^c\right) &= \lim_{n \rightarrow \infty} \mathbb{P}(B_n^c), \\ 1 - \mathbb{P}\left(\lim_{n \rightarrow \infty} B_n\right) &= 1 - \lim_{n \rightarrow \infty} \mathbb{P}(B_n^c).\end{aligned}$$

The LHS is:

$$1 - \mathbb{P}\left(\lim_{n \rightarrow \infty} B_n\right) = P\left(\bigcup_{n=1}^{\infty} \bigcap_{k \geq n} B_k^c\right)^c = P\left(\bigcap_{n=1}^{\infty} \bigcup_{k \geq n} B_k\right) = \mathbb{P}\left(\lim_{n \rightarrow \infty} B_n\right).$$

The RHS is:

$$1 - \lim_{n \rightarrow \infty} \mathbb{P}(B_n^c) = \lim_{n \rightarrow \infty} \{1 - \mathbb{P}(B_n^c)\} = \lim_{n \rightarrow \infty} \mathbb{P}(B_n).$$

■

Corollary 1.3.2. Let (A_n) denote a sequence of sets, then:

$$\begin{aligned}\lim_{n \rightarrow \infty} \mathbb{P}\left(\sup_{k \geq n} A_k\right) &= \mathbb{P}\left(\limsup_{n \rightarrow \infty} A_n\right) \\ \lim_{n \rightarrow \infty} \mathbb{P}\left(\inf_{k \geq n} A_k\right) &= \mathbb{P}\left(\liminf_{n \rightarrow \infty} A_n\right)\end{aligned}$$

♣

Proof. The conclusion follows because $B_n = \sup_{k \geq n} A_k$ is decreasing and $C_n = \inf_{k \geq n} A_k$ is increasing. ■

Theorem 1.3.1 (Continuity). If (A_n) is a sequence of sets converging to A , then:

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}\left(\lim_{n \rightarrow \infty} A_n\right). \quad (1.3.2)$$

□

Proof. Since $\inf_{k \geq n} A_k \subseteq A_n \subseteq \sup_{k \geq n} A_k$:

$$\mathbb{P}\left(\inf_{k \geq n} A_k\right) \leq \mathbb{P}(A_n) \leq \mathbb{P}\left(\sup_{k \geq n} A_k\right).$$

Taking the limit as $n \rightarrow \infty$:

$$\begin{aligned}\lim_{n \rightarrow \infty} \mathbb{P}\left(\inf_{k \geq n} A_k\right) &\leq \lim_{n \rightarrow \infty} \mathbb{P}(A_n) \leq \lim_{n \rightarrow \infty} \mathbb{P}\left(\sup_{k \geq n} A_k\right), \\ \mathbb{P}\left(\liminf_{n \rightarrow \infty} A_n\right) &\leq \lim_{n \rightarrow \infty} \mathbb{P}(A_n) \leq \mathbb{P}\left(\limsup_{n \rightarrow \infty} A_n\right).\end{aligned}$$

Since $A_n \rightarrow A$:

$$\lim_{n \rightarrow \infty} A_n = \liminf_{n \rightarrow \infty} A_n = \limsup_{n \rightarrow \infty} A_n.$$

Conclude that:

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} A_n\right) \leq \lim_{n \rightarrow \infty} \mathbb{P}(A_n) \leq \mathbb{P}\left(\lim_{n \rightarrow \infty} A_n\right).$$

■

Almost Sure Convergence

Definition 1.4.1. A sequence of random variables (\mathbf{X}_n) **converges almost surely** to \mathbf{X} , expressed $\mathbf{X}_n \xrightarrow{as} \mathbf{X}$, if:

$$\mathbb{P}\left\{\omega : \lim_{n \rightarrow \infty} \mathbf{X}_n(\omega) = \mathbf{X}(\omega)\right\} = 1.$$

Equivalently, for $\forall \epsilon > 0$:

$$\mathbb{P}\left\{\omega : \limsup_{n \rightarrow \infty} \|\mathbf{X}_n - \mathbf{X}\| > \epsilon\right\} = 0.$$

■

Remark 1.4.1. In the following, the notation:

$$\left\{\limsup_{n \rightarrow \infty} \|\mathbf{X}_n - \mathbf{X}\| > \epsilon\right\},$$

will implicitly refer to the set of $\omega \in \Omega$ where the condition $\{\cdot\}$ holds.

◆

4.1 Criteria

Proposition 1.4.1. If for $\forall \epsilon > 0$:

$$\sum_{n=1}^{\infty} \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) < \infty,$$

then $\mathbf{X}_n \xrightarrow{as} \mathbf{X}$.

◆

Proof.

$$\begin{aligned} \mathbb{P}\left\{\limsup_{n \rightarrow \infty} \|\mathbf{X}_n - \mathbf{X}\| > \epsilon\right\} &= \lim_{n \rightarrow \infty} \mathbb{P}\left\{\sup_{k \geq n} \|\mathbf{X}_k - \mathbf{X}\| > \epsilon\right\} \\ &= \lim_{n \rightarrow \infty} \mathbb{P}\left\{\bigcup_{k \geq n} \|\mathbf{X}_k - \mathbf{X}\| > \epsilon\right\} \leq \lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} \mathbb{P}(\|\mathbf{X}_k - \mathbf{X}\| > \epsilon). \end{aligned}$$

Since the series $\sum_{n=1}^{\infty} \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\|_2 > \epsilon)$ converges,

$$\lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} \mathbb{P}(\|\mathbf{X}_k - \mathbf{X}\| > \epsilon) = 0.$$

■

Proposition 1.4.2. If $\exists(p > 0)$ such that:

$$\sum_{n=1}^{\infty} E\|\mathbf{X}_n - \mathbf{X}\|^p < \infty,$$

then $\mathbf{X}_n \xrightarrow{as} \mathbf{X}$.

◆

Proof.

$$\begin{aligned} \mathbb{P}\left\{\limsup_{n \rightarrow \infty} \|\mathbf{X}_n - \mathbf{X}\| > \epsilon\right\} &\leq \lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} \mathbb{P}(\|\mathbf{X}_k - \mathbf{X}\| > \epsilon) \\ &\leq \frac{1}{\epsilon^p} \lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} E\|\mathbf{X}_k - \mathbf{X}\|^p = 0. \end{aligned}$$

■

4.2 Relation to Expectation

Remark 1.4.2. The following theorems carry over from the analogous results for sequences of real valued functions since if $\mathbf{X}_n \xrightarrow{as} \mathbf{X}$, then $\lim_{n \rightarrow \infty} \mathbf{X}_n(\omega) = \mathbf{X}(\omega)$, except possibly on a set of measured zero.

◆

Theorem 1.4.1 (Monotone Convergence). If (X_n) is a non-negative, increasing sequence of scalar random variables with $X_n \xrightarrow{as} X$, then:

$$\lim_{n \rightarrow \infty} E(X_n) = E(X).$$

□

Theorem 1.4.2 (Fatou's). If (\mathbf{X}_n) is a non-negative sequence of random variables, with $\mathbf{X}_n \xrightarrow{as} \mathbf{X}$, then:

$$E(\mathbf{X}) = E\left(\liminf_{n \rightarrow \infty} \mathbf{X}_n\right) \leq \liminf_{n \rightarrow \infty} E(\mathbf{X}_n).$$

□

Theorem 1.4.3 (Dominated Convergence). If $\mathbf{X}_n \xrightarrow{as} \mathbf{X}$ and $\|\mathbf{X}_n\| \leq Y$ with $E(Y) < \infty$, then:

$$\lim_{n \rightarrow \infty} E(\mathbf{X}_n) = E(\mathbf{X}).$$

□

4.3 Relation to Convergence in Probability

Proposition 1.4.3. Almost sure convergence implies convergence in probability:

$$\mathbf{X}_n \xrightarrow{as} \mathbf{X} \implies \mathbf{X}_n \xrightarrow{p} \mathbf{X}.$$

◆

Proof. Suppose $\mathbf{X}_n \xrightarrow{as} \mathbf{X}$, then for $\forall \epsilon > 0$:

$$\begin{aligned} 0 &= \mathbb{P} \left\{ \limsup_{n \rightarrow \infty} \|\mathbf{X}_n - \mathbf{X}\| > \epsilon \right\} \\ &= \lim_{n \rightarrow \infty} \mathbb{P} \left\{ \sup_{k \geq n} \|\mathbf{X}_k - \mathbf{X}\| > \epsilon \right\} \\ &\geq \lim_{n \rightarrow \infty} \mathbb{P} \{ \|\mathbf{X}_n - \mathbf{X}\| > \epsilon \}. \end{aligned}$$

■

L^p Convergence

Definition 1.5.1. A sequence of random variables (\mathbf{X}_n) **converges in L^p** , expressed $\mathbf{X}_n \xrightarrow{L^p} \mathbf{X}$ if:

$$\lim_{n \rightarrow \infty} E(\|\mathbf{X}_n - \mathbf{X}\|^p) = 0.$$

■

Proposition 1.5.1 (Markov's Inequality).

$$\mathbb{P}(\|\mathbf{X}\| \geq t) \leq \frac{E\|\mathbf{X}\|}{t}.$$

◆

Proof. Let $Y = \|\mathbf{X}\|$,

$$\mathbb{P}(Y \geq t) = E\{I(Y \geq t)\} = E\{I(Y/t \geq 1)\} \leq E(Y/t).$$

■

Corollary 1.5.1. For $p > 0$,

$$\mathbb{P}(\|\mathbf{X}\| \geq t) \leq \frac{E\|\mathbf{X}\|^p}{t^p}.$$

♣

Proposition 1.5.2. If $\mathbf{X}_n \xrightarrow{L^p} \mathbf{X}$ then:

$$\lim_{n \rightarrow \infty} E\|\mathbf{X}_n\|^p = \lim_{n \rightarrow \infty} E\|\mathbf{X}\|^p.$$

◆

Proof. By Minkowski's inequality:

$$\|\mathbf{X}_n\|^p \leq \|\mathbf{X}_n - \mathbf{X}\|^p + \|\mathbf{X}\|^p, \quad \|\mathbf{X}\|^p \leq \|\mathbf{X} - \mathbf{X}_n\|^p + \|\mathbf{X}_n\|^p.$$

Since $E\|\mathbf{X}_n\|^p - E\|\mathbf{X}\|^p \leq E\|\mathbf{X}_n - \mathbf{X}\|^p$ and $E\|\mathbf{X}\|^p - E\|\mathbf{X}_n\|^p \leq E\|\mathbf{X}_n - \mathbf{X}\|^p$:

$$|E\|\mathbf{X}_n\|^p - E\|\mathbf{X}\|^p| \leq E\|\mathbf{X}_n - \mathbf{X}\|^p \rightarrow 0.$$

■

5.1 Relation to Convergence in Probability

Proposition 1.5.3. For $p > 1$, convergence in L^p implies convergence in probability:

$$\mathbf{X}_n \xrightarrow{L^p} \mathbf{X} \implies \mathbf{X} \xrightarrow{p} \mathbf{X}.$$

◆

Proof. For $\forall \epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) \leq \frac{1}{\epsilon^p} \lim_{n \rightarrow \infty} E\|\mathbf{X}_n - \mathbf{X}\|^p = 0.$$

■

Convergence in Probability

Definition 1.6.1. A sequence of random variables (\mathbf{X}_n) **converges in probability** to \mathbf{X} , expressed $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$, if for $\forall \epsilon > 0$:

$$\lim_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) = 0.$$

■

Definition 1.6.2. A sequence of random variables (\mathbf{X}_n) is **Cauchy in probability** if $\forall(\epsilon, \delta) > 0$ there exists $\nu_{\epsilon, \delta}$ such that when $n, m \geq \nu$:

$$\mathbb{P}(\|\mathbf{X}_n - \mathbf{X}_m\| > \epsilon) < \delta.$$

■

6.1 Relation to Almost Sure Convergence

Proposition 1.6.1. If $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$, then (\mathbf{X}_n) is Cauchy in probability. ◆

Proof. By sub-additivity (1.2.1):

$$\mathbb{P}(\|\mathbf{X}_n - \mathbf{X}_m\| > \epsilon) \leq \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon/2) + \mathbb{P}(\|\mathbf{X} - \mathbf{X}_m\| > \epsilon/2).$$

Since $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$, there $\exists \nu_{\epsilon, \delta}$ such that when $n \geq \nu$:

$$\mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon/2) < \delta/2$$

Now for $n, m \geq \nu$:

$$\mathbb{P}(\|\mathbf{X}_n - \mathbf{X}_m\| > \epsilon) < \delta/2 + \delta/2 = \delta.$$

■

Proposition 1.6.2. If (\mathbf{X}_n) is Cauchy in probability, then (\mathbf{X}_n) contains a subsequence (\mathbf{X}_{n_j}) that converges almost surely. ◆

Proof. Let $n_1 = 1$, and define n_j by:

$$n_j = \inf_{n \in \mathbb{N}} \{n > n_{j-1} : \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}_m\| > 2^{-j}) < 2^{-j} \text{ for } \forall (n, m) \geq n\}.$$

Thus, n_j is the smallest n exceeding n_{j-1} such that the probability of the distance between \mathbf{X}_n and \mathbf{X}_m exceeding $\epsilon = 2^{-j}$ is less than $\delta = 2^{-j}$. Now since:

$$\sum_{j=1}^{\infty} \mathbb{P}(\|\mathbf{X}_{n_{j+1}} - \mathbf{X}_{n_j}\| > 2^{-j}) < \infty,$$

by the Borel-Canteli lemmas:

$$\mathbb{P} \left\{ \limsup_{j \rightarrow \infty} (\|\mathbf{X}_{n_{j+1}} - \mathbf{X}_{n_j}\| > 2^{-j}) \right\} = 0$$

almost everywhere on Ω . Hence for ω in a set of probability one:

$$\alpha_j(\omega) = \|\mathbf{X}_{n_{j+1}}(\omega) - \mathbf{X}_{n_j}(\omega)\| \leq 2^{-j}.$$

Now $\{\alpha_j(\omega)\}$ is a Cauchy sequence in \mathbb{R} , and by completeness, $\alpha_j(\omega)$ converges to some limit $\alpha(\omega)$. Thus, $\lim_{j \rightarrow \infty} \mathbf{X}_{n_j}(\omega)$ exists almost everywhere on Ω . ■

Proposition 1.6.3. If (\mathbf{X}_n) is Cauchy in probability, then \mathbf{X}_n converges in probability to some limit \mathbf{X} . ◆

Proof. Since (\mathbf{X}_n) is Cauchy in probability, there exists a subsequence (\mathbf{X}_{n_j}) converging almost surely to some limit \mathbf{X} . Now:

$$\mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) \leq \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}_{n_j}\| > \epsilon/2) + \mathbb{P}(\|\mathbf{X}_{n_j} - \mathbf{X}\| > \epsilon/2).$$

Since (\mathbf{X}_n) is Cauchy, there exists ν_1 such that when $n, n_j \geq \nu_1$:

$$\mathbb{P}(\|\mathbf{X}_n - \mathbf{X}_{n_j}\| > \epsilon/2) < \delta/2.$$

Since $\mathbf{X}_{n_j} \xrightarrow{as} \mathbf{X}$, which implies $\mathbf{X}_{n_j} \xrightarrow{p} \mathbf{X}$, there exists ν_2 such that when $n_j \geq \nu_2$:

$$\mathbb{P}(\|\mathbf{X}_{n_j} - \mathbf{X}\| > \epsilon/2) < \delta/2.$$

Let $\nu = \max(\nu_1, \nu_2)$, then for $n, n_j \geq \nu$:

$$\mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) < \delta.$$

■

Proposition 1.6.4. $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$ if and only if every subsequence (\mathbf{X}_{n_j}) itself contains a subsequence $(\mathbf{X}_{n_{j_k}})$ converging almost surely to \mathbf{X} . ♦

Proof. (\implies) If $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$, then any subsequence (\mathbf{X}_{n_j}) converges in probability to the same limit, and the subsequence (\mathbf{X}_{n_j}) contains a further subsequence converging almost surely.

(\impliedby) Suppose for contradiction that every subsequence contains an almost surely converging subsequence, but that \mathbf{X}_n does not converge in probability. Since it is not the case that $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$, for $\delta > 0$ and $\epsilon > 0$ there exist infinitely many k such that:

$$\mathbb{P}(\|\mathbf{X}_k - \mathbf{X}\| > \epsilon) \geq \delta.$$

Let \mathbf{X}_{n_j} denote a subsequence along such k , then (\mathbf{X}_{n_j}) can itself contain no subsequence converging almost surely to \mathbf{X} . ■

Theorem 1.6.1 (Cauchy Criteria).

- i. (\mathbf{X}_n) converges in probability if and only if (\mathbf{X}_n) is Cauchy in probability.
- ii. $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$ if and only if every subsequence (\mathbf{X}_{n_j}) contains a further subsequence $(\mathbf{X}_{n_{j_k}})$ converging almost surely to \mathbf{X} .

□

6.2 Relation to Expectation

Theorem 1.6.2 (Dominated Convergence). If $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$ and $\|\mathbf{X}_n\| \leq Y$ with $E(Y) < \infty$, then:

$$\lim_{n \rightarrow \infty} E(\mathbf{X}_n) = E(\mathbf{X}).$$

□

Remark 1.6.1. The result is established by showing that every convergent subsequence of $E(\mathbf{X}_n)$ converges to the same limit. Suppose $E(\mathbf{X}_{n_j})$ is a convergent subsequence. Since $\mathbf{X}_{n_j} \xrightarrow{p} \mathbf{X}$, there exists a subsequence $\mathbf{X}_{n_{j_k}}$ converging almost surely to \mathbf{X} . By the dominated convergence theorem for almost sure convergence, $E(\mathbf{X}_{n_{j_k}}) \rightarrow E(\mathbf{X})$. Since this holds for any subsequence, every convergent subsequence of $E(\mathbf{X}_n)$ approaches the same limit. See Resnick (2014), corollary 6.3.2. ♦

6.3 Uniform Integrability

Definition 1.6.3. Let \mathcal{T} denote an index set. The collection of random variables (\mathbf{X}_t) is **uniformly integrable** if:

$$\lim_{M \rightarrow \infty} \sup_{t \in \mathcal{T}} E\{\|\mathbf{X}_t\| I(\|\mathbf{X}_t\| > M)\} = 0.$$

■

Remark 1.6.2. If (\mathbf{X}_t) is uniformly integrable, then for any $\epsilon > 0$, there exists M_ϵ such that for $m \geq M_\epsilon$:

$$\sup_{t \in \mathcal{T}} E\{\|\mathbf{X}_t\| I(\|\mathbf{X}_t\| > m)\} < \epsilon.$$

♦

Proposition 1.6.5. A random variable \mathbf{X} has finite expectation $E\|\mathbf{X}\| < \infty \iff$

$$\lim_{M \rightarrow \infty} E\{\|\mathbf{X}\| I(\|\mathbf{X}\| > M)\} = 0.$$

That is, an individual \mathbf{X} has finite expectation if and only if it is uniformly integrable. ♦

Proof. (\implies) Consider the sequence of functions:

$$g_M(\omega) = \|\mathbf{X}(\omega)\| I(\|\mathbf{X}(\omega)\| \leq M)$$

for $M \rightarrow \infty$. This sequence is non-negative and increasing in M , with:

$$\lim_{M \rightarrow \infty} g_M(\omega) = \|\mathbf{X}(\omega)\|.$$

By the monotone convergence theorem:

$$\lim_{M \rightarrow \infty} E\{\|\mathbf{X}\|I(\|\mathbf{X}\| \leq M)\} = E\{\|\mathbf{X}\|\}.$$

Since:

$$E\{\|\mathbf{X}\|I(\|\mathbf{X}\| > M)\} = E\|\mathbf{X}\| - E\{\|\mathbf{X}\|I(\|\mathbf{X}\| \leq M)\},$$

upon taking the limit as $M \rightarrow \infty$:

$$\lim_{M \rightarrow \infty} E\{\|\mathbf{X}\|I(\|\mathbf{X}\| > M)\} = E\|\mathbf{X}\| - \lim_{M \rightarrow \infty} E\{\|\mathbf{X}\|I(\|\mathbf{X}\| \leq M)\} = 0.$$

(\Leftarrow) Suppose on the contrary that $E\|\mathbf{X}\| = \infty$. Since:

$$E\{\|\mathbf{X}\|I(\|\mathbf{X}\| \leq M)\} \leq M\mathbb{P}(\|\mathbf{X}\| \leq M) \leq M,$$

and:

$$E\|\mathbf{X}\| = E\{\|\mathbf{X}\|I(\|\mathbf{X}\| \leq M)\} + E\{\|\mathbf{X}\|I(\|\mathbf{X}\| > M)\},$$

if $E\|\mathbf{X}\| = \infty$, then for any M :

$$E\{\|\mathbf{X}\|I(\|\mathbf{X}\| > M)\} = \infty.$$

Thus $\lim_{M \rightarrow \infty} E\{\|\mathbf{X}\|I(\|\mathbf{X}\| > M)\} \neq 0$. ■

Proposition 1.6.6. If (\mathbf{X}_t) is uniformly integrable, if and only if:

- i. $\sup_{t \in \mathcal{T}} E\|\mathbf{X}_t\| < \infty$,
- ii. For $\forall \epsilon > 0$ there $\exists \delta > 0$ s.t.

$$\mathbb{P}(A) < \delta \implies \sup_{t \in \mathcal{T}} E\{\|\mathbf{X}_t\|I_A\} < \epsilon.$$

Importantly, $\sup_{t \in \mathcal{T}} E\|\mathbf{X}_t\| < \infty$ is necessary but not sufficient for a sequence (\mathbf{X}_t) to achieve uniform integrability. ◆

Proof. (\implies) Suppose (\mathbf{X}_t) is uniformly integrable. Fix $\epsilon > 0$. Choose M_ϵ such that $E\{\|\mathbf{X}\|I(\|\mathbf{X}\| > M)\} < \epsilon$. Now for $\forall t$:

$$E\|\mathbf{X}_t\| = E\{\|\mathbf{X}_t\|I(\|\mathbf{X}_t\| \leq M)\} + E\{\|\mathbf{X}_t\|I(\|\mathbf{X}_t\| > M)\} \leq M + \epsilon.$$

Thus $\sup_{t \in \mathcal{T}} E\|\mathbf{X}_t\| < \infty$. Now select M_ϵ such that $E\{\|\mathbf{X}\|I(\|\mathbf{X}\| > M)\} < \epsilon/2$ and set $\delta = \epsilon/(2M_\epsilon)$. If $\mathbb{P}(A) < \delta$, then:

$$\begin{aligned} E\{\|\mathbf{X}_t\|I_A\} &= E\{\|\mathbf{X}_t\|I(A \cap \|\mathbf{X}_t\| \leq M)\} + E\{\|\mathbf{X}_t\|I(A \cap \|\mathbf{X}_t\| > M)\} \\ &\leq M\mathbb{P}(A) + E\{\|\mathbf{X}_t\|I(\|\mathbf{X}_t\| > M)\} \leq M \frac{\epsilon}{2M} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

(\Leftarrow) By condition (i.), there exists K such that $\sup_{t \in \mathcal{T}} E\|\mathbf{X}_t\| < K$. By condition (ii.), for $\epsilon > 0$, there exists $\delta > 0$ s.t.

$$\mathbb{P}(A) < \delta \implies \sup_{t \in \mathcal{T}} E\{\|\mathbf{X}_t\| I_A\} < \epsilon.$$

Observe that by Markov's inequality:

$$\mathbb{P}(\|\mathbf{X}_i\| > M) \leq \frac{E(\|\mathbf{X}_t\|)}{M} \leq \frac{K}{M}.$$

Choose M such that $K/M < \delta$, then $\mathbb{P}(\|\mathbf{X}_i\| > M) < \delta$, which implies that:

$$\sup_{t \in \mathcal{T}} E\{\|\mathbf{X}_t\| I(\|\mathbf{X}_i\| > M)\} < \epsilon,$$

as required. ■

Proposition 1.6.7. If there exists $\delta > 0$ such that:

$$\sup_{t \in \mathcal{T}} E\|\mathbf{X}_t\|^{1+\delta} < \infty,$$

then (\mathbf{X}_t) is uniformly integrable. ♦

Proof.

$$\begin{aligned} E\{\|\mathbf{X}_t\| I(\|\mathbf{X}_t\| > M)\} &= E\{\|\mathbf{X}_t\| \cdot 1 \cdot I(\|\mathbf{X}_t\|/M > 1)\} \\ &\leq E\left\{\|\mathbf{X}_t\| \cdot \frac{\|\mathbf{X}_t\|^\delta}{M^\delta} \cdot I(\|\mathbf{X}_t\|/M > 1)\right\} \\ &\leq \frac{1}{M^\delta} E\{\|\mathbf{X}_t\|^{1+\delta}\} \rightarrow 0 \end{aligned}$$

as $M \rightarrow \infty$. ■

Theorem 1.6.3. The following statements are equivalent:

- i. $\mathbf{X}_n \xrightarrow{L^1} \mathbf{X}$,
- ii. (\mathbf{X}_n) is uniformly integrable, and $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$.

□

Proof. (\Leftarrow) Since $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$, there exists a subsequence (\mathbf{X}_{n_k}) converging almost surely to \mathbf{X} . By Fatou's lemma:

$$E\{\|\mathbf{X}\|\} = E\left\{\liminf_{n_k \rightarrow \infty} \|\mathbf{X}_{n_k}\|\right\} \leq \liminf_{n_k \rightarrow \infty} E\|\mathbf{X}_{n_k}\| \leq \sup_{n \in \mathbb{N}} E\|\mathbf{X}_n\| < \infty,$$

where the last inequality follows from uniform integrability. Define the bounded, continuous function:

$$\psi_M(t) = \begin{cases} -M, & t \leq -M, \\ t, & t \in (-M, M), \\ M, & t \geq M. \end{cases}$$

Consider:

$$E\|\mathbf{X}_n - \mathbf{X}\| \leq E\|\mathbf{X}_n - \psi_M(\mathbf{X}_n)\| + E\|\psi_M(\mathbf{X}_n) - \psi_M(\mathbf{X})\| + E\|\psi_M(\mathbf{X}) - \mathbf{X}\|.$$

The first term is bounded as:

$$E\|\mathbf{X}_n - \psi_M(\mathbf{X}_n)\| \leq E\{\|\mathbf{X}_n\|I(\|\mathbf{X}_n\| > M)\}.$$

By uniform integrability, there exists M_1 such that for $m \geq M_1$:

$$\sup_{n \in \mathbb{N}} E\{\|\mathbf{X}_n\|I(\|\mathbf{X}_n\| > m)\} < \frac{\epsilon}{3}.$$

Since $\psi_M(t)$ is a bounded, continuous function, and $\mathbf{X}_n \xrightarrow{d} \mathbf{X}$, by the portmanteau theorem there exists ν such that for $n \geq \nu$:

$$E\|\psi_M(\mathbf{X}_n) - \psi_M(\mathbf{X})\| < \frac{\epsilon}{3}.$$

Finally, since $E\|\mathbf{X}\|$ is bounded, there exists M_2 such that for $m \geq M_2$:

$$E\|\psi_M(\mathbf{X}) - \mathbf{X}\| \leq E\{\|\mathbf{X}\|I(\|\mathbf{X}\| > M)\} < \frac{\epsilon}{3}.$$

Take $n \geq \nu$ and $m \geq \max(M_1, M_2)$, then:

$$E\|\mathbf{X}_n - \mathbf{X}\| < \epsilon.$$

(\implies) Since $\mathbf{X}_n \xrightarrow{L^1} \mathbf{X}$, the sequence of real numbers $E\|\mathbf{X}_n\|$ converges to $E\|\mathbf{X}\| < \infty$. Since convergent sequences in \mathbb{R} are bounded:

$$\sup_{n \in \mathbb{N}} E\|\mathbf{X}_n\| < \infty.$$

Now since $E\|\mathbf{X}_n - \mathbf{X}\| \rightarrow 0$, there exists ν_ϵ such that for $n \geq \nu$:

$$E\|\mathbf{X}_n - \mathbf{X}\| < \epsilon/2.$$

Let $K_1 = E\|\mathbf{X}\|$ and $K_2 = \max_{n < \nu} E\|\mathbf{X}_n\|$. Set $K = \max(K_1, K_2)$. Choose δ_1 such that if $\mathbb{P}(A) < \delta_1$, then:

$$E\{\|\mathbf{X}\|I_A\} \leq K\mathbb{P}(A) < \frac{\epsilon}{2}.$$

Choose δ_2 such that if $\mathbb{P}(A) < \delta_1$, then:

$$\max_{n < \nu} E\{\|\mathbf{X}_n - \mathbf{X}\|_{I_A}\} \leq K\mathbb{P}(A) < \frac{\epsilon}{2}$$

Set $\delta = \min(\delta_1, \delta_2)$, then for $\mathbb{P}(A) < \delta$:

$$\sup_{n \in \mathbb{N}} E\{\|\mathbf{X}_n\|_{I_A}\} \leq \sup_{n \in \mathbb{N}} E\{\|\mathbf{X}_n - \mathbf{X}\|_{I_A}\} + E\{\|\mathbf{X}\|_{I_A}\} < \frac{\epsilon}{2} + \frac{\epsilon}{2} < \epsilon.$$

The inequality holds for $n < \nu$ by choice of δ , and for $n \geq \nu$ by convergence in L^1 . ■

6.4 Relation to Convergence in Distribution

Proposition 1.6.8. Convergence in probability implies convergence in distribution:

$$\mathbf{X}_n \xrightarrow{p} \mathbf{X} \implies \mathbf{X}_n \xrightarrow{d} \mathbf{X}.$$

◆

Proof. Suppose $\mathbf{X}_n \xrightarrow{p} \mathbf{X}$ and that \mathbf{t} is a continuity point of F_X , then:

$$\begin{aligned} \mathbb{P}(\mathbf{X}_n \leq \mathbf{t}) &= \mathbb{P}\left(\{\mathbf{X}_n \leq \mathbf{t}, \|\mathbf{X}_n - \mathbf{X}\| < \epsilon\mathbf{1}\} \cup \{\mathbf{X}_n \leq \mathbf{t}, \|\mathbf{X}_n - \mathbf{X}\| + 2 \geq \epsilon\mathbf{1}\}\right) \\ &\leq \mathbb{P}(\mathbf{X} \leq \mathbf{t} + \epsilon\mathbf{1}) + \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon). \end{aligned}$$

Similarly:

$$\mathbb{P}(\mathbf{X} \leq \mathbf{t} - \epsilon\mathbf{1}) \leq \mathbb{P}(\mathbf{X}_n \leq \mathbf{t}) + \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon).$$

Thus:

$$\mathbb{P}(\mathbf{X} \leq \mathbf{t} - \epsilon\mathbf{1}) - \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon) \leq \mathbb{P}(\mathbf{X}_n \leq \mathbf{t}) \leq \mathbb{P}(\mathbf{X} \leq \mathbf{t} + \epsilon\mathbf{1}) + \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon).$$

Taking the limit as $n \rightarrow \infty$:

$$\mathbb{P}(\mathbf{X} \leq \mathbf{t} - \epsilon\mathbf{1}) \leq \lim_{n \rightarrow \infty} \mathbb{P}(\mathbf{X}_n \leq \mathbf{t}) \leq \mathbb{P}(\mathbf{X} \leq \mathbf{t} + \epsilon\mathbf{1}).$$

Since \mathbf{t} is a continuity point of F_X as $\epsilon \rightarrow 0$:

$$\mathbb{P}(\mathbf{X} \leq \mathbf{t}) \leq \lim_{n \rightarrow \infty} \mathbb{P}(\mathbf{X}_n \leq \mathbf{t}) \leq \mathbb{P}(\mathbf{X} \leq \mathbf{t}).$$

■

Convergence in Distribution

Definition 1.7.1. A sequence of random variables (\mathbf{X}_n) **converges in distribution** if the sequence of distribution functions (F_n) converges *pointwise* to the distribution function F_X of \mathbf{X} on the set $\mathcal{C}(F_X)$ of continuity points of F_X :

$$\lim_{n \rightarrow \infty} F_n(\mathbf{t}) = F_X(\mathbf{t}) \text{ for } \mathbf{t} \in \mathcal{C}(F_X)$$

■

7.1 Relation to Convergence in Probability

Proposition 1.7.1. Convergence in distribution to a constant α implies convergence in probability:

$$\mathbf{X}_n \xrightarrow{d} \alpha \implies \mathbf{X}_n \xrightarrow{p} \alpha.$$

◆

Proof. If $\mathbf{X}_n \xrightarrow{d} \alpha$, then:

$$\lim_{n \rightarrow \infty} \mathbb{P}(\mathbf{X}_n \leq t) = I(t \geq \alpha).$$

Now, probability that $\|\mathbf{X}_n - \alpha\| < \epsilon$ is:

$$\mathbb{P}(\|\mathbf{X}_n - \alpha\| < \epsilon) = \mathbb{P}(\alpha - \epsilon \mathbf{1} \leq \mathbf{X}_n \leq \alpha + \epsilon \mathbf{1}) = F_n(\alpha + \epsilon \mathbf{1}) - F_n(\alpha - \epsilon \mathbf{1}).$$

Taking the limit as $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_n - \alpha\| < \epsilon) = I(\alpha + \epsilon \mathbf{1} > \alpha) - I(\alpha - \epsilon \mathbf{1} > \alpha) = 1.$$

■

7.2 Characteristic Functions

Definition 1.7.2. The **characteristic function** $\phi : \mathbb{R} \rightarrow \mathbb{C}$ of a random variable \mathbf{X} with distribution F is defined by:

$$\phi(t) = E(e^{itX}) = \int e^{itx} dF(x).$$

■

Theorem 1.7.1 (Uniqueness). The characteristic function $\phi(t)$ of a random variable uniquely determines its distribution. □

Remark 1.7.1. See Resnick (2014), theorem 9.5.1. ◆

Theorem 1.7.2 (Levy Continuity). Suppose (\mathbf{X}_n) is a sequence of random variables, and $\phi_n(t)$ the corresponding sequence of characteristic functions. \mathbf{X}_n converges in distribution to a random variable \mathbf{X} with characteristic function $\phi(t)$ if and only if $\phi_n(t) \rightarrow \phi(t)$:

$$\mathbf{X}_n \xrightarrow{d} \mathbf{X} \iff \phi_n \rightarrow \phi.$$

Moreover, if (ϕ_n) converges *pointwise* to ϕ_∞ , and ϕ_∞ is continuous at zero, then ϕ_∞ is the characteristic function of some random variable \mathbf{X} and $\mathbf{X}_n \xrightarrow{d} \mathbf{X}$. □

Remark 1.7.2. See Resnick (2014), theorem 9.5.2; van der Vaart (1998), theorem 2.13. ◆

7.3 Additional Properties

Theorem 1.7.3 (Portmanteau). The following statements are equivalent:

- i. $X_n \xrightarrow{d} X$,
- ii. $E\{g(X_n)\} \rightarrow E\{g(X)\}$ for every bounded, continuous function g .
- iii. $E\{\mathcal{L}(X_n)\} \rightarrow E\{\mathcal{L}(X)\}$ for every bounded, Lipschitz function \mathcal{L} .

□

Remark 1.7.3. For proof, see van der Vaart (1998), lemma 2.2. ♦

Theorem 1.7.4 (Skorokhod Representation). If $\mathbf{X}_n \xrightarrow{d} \mathbf{X}$, then there exists a sequence of random variables (ξ_n) defined on a common probability space such that $\mathbf{X}_n \stackrel{d}{=} \xi_n$, $\mathbf{X} \stackrel{d}{=} \xi$, and $\xi_n \xrightarrow{as} \xi$. □

Remark 1.7.4. The proof is by construction. See Billingsley (1995), theorem 25.6. ♦

Theorem 1.7.5 (Dominated Convergence). If $\mathbf{X}_n \xrightarrow{d} \mathbf{X}$ and $\|\mathbf{X}_n\| \leq Y$ with $E(Y) < \infty$, then:

$$\lim_{n \rightarrow \infty} E(\mathbf{X}_n) = E(\mathbf{X}).$$

□

Proof. By the Skorokhod representation theorem, there exist a sequence of random variables (ξ_n) and ξ such that $\xi_n \stackrel{d}{=} \mathbf{X}_n$, $\xi \stackrel{d}{=} \mathbf{X}$, and $\xi_n \xrightarrow{as} \xi$. By the dominated convergence theorem for almost sure convergence, $E(\xi_n) \rightarrow E(\xi)$. Thus:

$$\lim_{n \rightarrow \infty} E(\mathbf{X}_n) = \lim_{n \rightarrow \infty} E(\xi_n) = E(\xi) = E(\mathbf{X}).$$

■

Summary of Convergence Relations

- Convergence almost surely implies convergence in probability:

$$\mathbf{X}_n \xrightarrow{as} \mathbf{X} \implies \mathbf{X}_n \xrightarrow{p} \mathbf{X}.$$

- Convergence almost surely may be established by checking:

- The series $\sum_{n=1}^{\infty} \mathbb{P}(\|\mathbf{X}_n - \mathbf{X}\| > \epsilon)$ is finite for $\forall \epsilon > 0$.
- The series $\sum_{n=1}^{\infty} E\|\mathbf{X}_n - \mathbf{X}\|^p$ is finite for some $p > 0$.

- Convergence in L^p implies convergence in probability:

$$\mathbf{X}_n \xrightarrow{L^p} \mathbf{X} \implies \mathbf{X}_n \xrightarrow{p} \mathbf{X}.$$

- Convergence in probability and uniform integrability implies convergence in L^1 .

- Convergence in probability implies convergence in distribution:

$$\mathbf{X}_n \xrightarrow{p} \mathbf{X} \implies \mathbf{X}_n \xrightarrow{d} \mathbf{X}.$$

- Convergence in distribution to a constant $\boldsymbol{\alpha}$ implies convergence in probability to that constant:

$$\mathbf{X}_n \xrightarrow{d} \boldsymbol{\alpha} \implies \mathbf{X}_n \xrightarrow{p} \boldsymbol{\alpha}.$$

- (Dominated convergence) Suppose $\mathbf{X}_n \rightarrow \mathbf{X}$ almost surely, in probability, or in distribution. If $\|\mathbf{X}_n\| \leq Y$ and $E(Y) < \infty$, then:

$$\lim_{n \rightarrow \infty} E(\mathbf{X}_n) = E(\mathbf{X}).$$