Introduction

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

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

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

If the sequence of random variables $(\boldsymbol{X}_n)$ 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 $(\boldsymbol{X}_n)$ **converges in probability** to zero, expressed $\boldsymbol{X}_n = o_p(1)$, if for $\forall\epsilon > 0$:

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

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

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

i. $\boldsymbol{X}_n = o_p(\alpha_n) \iff \alpha_n^{-1}\boldsymbol{X}_n = o_p(1)$.

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

■

2.2 Properties

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

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 $n \geq \nu_\delta$, $\mathbb{P}(||\boldsymbol{X}_n|| > \epsilon) < \delta$. ■

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

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

♦

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

$$\Big|\Big|\sum_{i=1}^{n} \boldsymbol{X}_i\Big|\Big|_2 \leq \sum_{i=1}^{n} ||\boldsymbol{X}_i|| \leq \sum_{i=1}^{n} \frac{\epsilon}{n} = \epsilon,$$

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

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

By sub-additivity of the probability measure:

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

■

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

♦

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:

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

ii.

$$\mathbb{P}\big(||X_nY_n|| > M_XM_Y\big) \leq \mathbb{P}\big(||X_n|| > M_X \cup ||Y_n|| > M_Y\big)$$
$$\leq \mathbb{P}\big(||X_n|| > M_X\big) + \mathbb{P}\big(||Y_n|| > M_Y\big) \leq \epsilon/2 + \epsilon/2 = \epsilon.$$

■

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. $X_nY_n = o_p(1)$.

♦

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 $\forall(M, \delta)$,

$$\mathbb{P}\big(||X_nY_n|| > M\big) = \mathbb{P}\big(||X_nY_n > \delta \cap ||Y_n|| > M) + \mathbb{P}\big(||X_nY_n|| \leq \delta \cap ||Y_n|| \leq M\big)$$
$$\leq \mathbb{P}\big(||Y_n|| > M\big) + \mathbb{P}\big(||X_n|| > \delta/M\big)$$

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 $X_n = o_p(1)$, $\lim_{n\to\infty} \mathbb{P}(||X_n|| > \delta/M) = 0$. Thus:

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

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

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

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

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

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

□

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

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

By subadditivity:

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

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

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

By subadditivity:

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

iii.

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

■

Proposition 1.2.5.

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

♦

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

$$\lim_{n\to\infty}\mathbb{P}\big(||\boldsymbol{Y}_n|| > \epsilon\big) = 0,$$
$$\lim_{n\to\infty}\mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}|| > \epsilon\big) = 0.$$

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

$$\lim_{n\to\infty}\mathbb{P}\big(|Y_n| > \epsilon\big) = 0,$$
$$\lim_{n\to\infty}\mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}|| > \epsilon\big) = 0.$$

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

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

Proof. By the triangle inequality:

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

■

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

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

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

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. ♦

Proof. i. ($\implies$):

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

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

($\impliedby$):

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

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

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

Since $\mathbb{P}(|X_{nj} - X_j| > M_\epsilon) \leq \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}|| > M_\epsilon\big)$, conclude $X_{nj} = \mathcal{O}_p(1)$.
For $\epsilon > 0$ there $\exists M_\epsilon, \nu_\epsilon$ s.t. when $n \geq \nu_\epsilon$:

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

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

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

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

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

■

Limits of Sets

3.1 Definitions

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

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

$\omega \in \lim_{n\to\infty} B_n$ 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 $\lim_{n\to\infty} C_n$ is defined as their union:

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

$\omega \in \lim_{n\to\infty} C_n$ 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:

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

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:

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

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

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

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, than the **limit** exists and is defined as:

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

■

3.2 Properties

Proposition 1.3.1. For any sequence of sets,

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

♦

Proof. Suppose $\omega \in \liminf_{n\to\infty} A_n$, 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 $\omega \in \limsup_{n\to\infty} A_n$. ■

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

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

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

♦

Proof. Define the sequence of *disjoint* sets $D_1 = C_1$, and $D_k = D_k - D_{k-1}$ for $k \geq 2$. Clearly $C_n = \cup_{k=1}^{n} D_k$. 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:

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

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

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

■

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

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

♣

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

$$\mathbb{P}\left(\lim_{n\to\infty} B_n^c\right) = \lim_{n\to\infty} \mathbb{P}(B_n^c),$$
$$1 - \mathbb{P}\left(\lim_{n\to\infty} B_n^c\right) = 1 - \lim_{n\to\infty} \mathbb{P}(B_n^c).$$

The LHS is:

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

The RHS is:

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

■

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

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{k\geq n} A_k\right) = \mathbb{P}\left(\limsup_{n\to\infty} A_n\right)$$
$$\lim_{n\to\infty} \mathbb{P}\left(\inf_{k\geq n} A_k\right) = \mathbb{P}\left(\liminf_{n\to\infty} A_n\right)$$

♣

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:

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

□

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 $n \to \infty$:

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

Since $A_n \to A$:

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

Conclude that:

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

■

Almost Sure Convergence

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

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

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

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

■

Remark 1.4.1. In the following, the notation:

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

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

4.1 Criteria

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

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

then $\boldsymbol{X}_n \xrightarrow{as} \boldsymbol{X}$. ♦

Proof.

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

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

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

■

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

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

then $\boldsymbol{X}_n\xrightarrow{as}\boldsymbol{X}$. ♦

Proof.

$$\mathbb{P}\left\{\limsup_{n\to\infty}||\boldsymbol{X}_n-\boldsymbol{X}||>\epsilon\right\}\leq\lim_{n\to\infty}\sum_{k=n}^{\infty}\mathbb{P}\big(||\boldsymbol{X}_k-\boldsymbol{X}||>\epsilon\big)$$
$$\leq\frac{1}{\epsilon^p}\lim_{n\to\infty}\sum_{k=n}^{\infty}E||\boldsymbol{X}_n-\boldsymbol{X}||^p=0.$$

■

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 $\boldsymbol{X}_n\xrightarrow{as}\boldsymbol{X}$, then $\lim_{n\to\infty}\boldsymbol{X}_n(\omega)=\boldsymbol{X}(\omega)$, 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:

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

□

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

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

□

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

$$\lim_{n\to\infty}E(\boldsymbol{X}_n)=E(\boldsymbol{X}).$$

□

4.3 Relation to Convergence in Probability

Proposition 1.4.3. Almost sure convergence implies convergence in probability:

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

♦

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

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

■

L^p Convergence

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

$$\lim_{n\to\infty} E\big(||\boldsymbol{X}_n - \boldsymbol{X}||^p\big) = 0.$$

■

Proposition 1.5.1 (**Markov's Inequality**)**.**

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

♦

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

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

■

Corollary 1.5.1. For $p > 0$,

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

♣

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

$$\lim_{n\to\infty} E||\boldsymbol{X}_n||^p = \lim_{n\to\infty} E||\boldsymbol{X}||^p.$$

♦

Proof. By Minkowski's inequality:

$$||\boldsymbol{X}_n||^p \leq ||\boldsymbol{X}_n - \boldsymbol{X}||^p + ||\boldsymbol{X}||^p, \qquad ||\boldsymbol{X}||^p \leq ||\boldsymbol{X} - \boldsymbol{X}_n||^p + ||\boldsymbol{X}_n||^p.$$

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

$$\big|E||\boldsymbol{X}_n||^p - E||\boldsymbol{X}||^p\big| \leq E||\boldsymbol{X}_n - \boldsymbol{X}||^p \to 0.$$

■

5.1 Relation to Convergence in Probability

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

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

♦

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

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

■

Convergence in Probability

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

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

■

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

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

■

6.1 Relation to Almost Sure Convergence

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

Proof. By sub-additivity (1.2.1):

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

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

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

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

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

■

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

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

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

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

$$\sum_{j=1}^{\infty} \mathbb{P}\big(||\boldsymbol{X}_{j+1} - \boldsymbol{X}_j|| > 2^{-j}\big) < \infty,$$

by the Borel-Canteli lemmas:

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

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

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

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

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

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

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

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

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

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

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

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

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

■

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

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

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

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

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

Theorem 1.6.1 (Cauchy Criteria).

i. $(\boldsymbol{X}_n)$ converges in probability if and only if $(\boldsymbol{X}_n)$ is Cauchy in probability.

ii. $\boldsymbol{X}_n \xrightarrow{p} \boldsymbol{X}$ if and only if every subsequence $(\boldsymbol{X}_{n_j})$ contains a further subsequence $(\boldsymbol{X}_{n_{jk}})$ converging almost surely to $\boldsymbol{X}$.

□

6.2 Relation to Expectation

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

$$\lim_{n\to\infty} E(\boldsymbol{X}_n) = E(\boldsymbol{X}).$$

$\square$

Remark 1.6.1. The result is established by showing that every convergent subsequence of $E(\boldsymbol{X}_n)$ converges to the same limit. Suppose $E(\boldsymbol{X}_{n_j})$ is a convergent subsequence. Since $\boldsymbol{X}_{n_j} \xrightarrow{p} \boldsymbol{X}$, there exists a subsequence $\boldsymbol{X}_{n_{jk}}$ converging almost surely to $\boldsymbol{X}$. By the dominated convergence theorem for almost sure convergence, $E(\boldsymbol{X}_{n_j}) \to E(\boldsymbol{X})$. Since this holds for any subsequence, every convergent subsequence of $E(\boldsymbol{X}_n)$ 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 $(\boldsymbol{X}_t)$ is **uniformly integrable** if:

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

$\blacksquare$

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

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

⧫

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

$$\lim_{M\to\infty} E\big\{||\boldsymbol{X}||I(||\boldsymbol{X}|| > M)\big\} = 0.$$

That is, an individual $\boldsymbol{X}$ has finite expectation if and only if it is uniformly integrable. ⧫

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

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

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

$$\lim_{M\to\infty} g_M(\omega) = ||\boldsymbol{X}(\omega)||.$$

By the monotone convergence theorem:

$$\lim_{M\to\infty} E\big\{||\boldsymbol{X}||I(||\boldsymbol{X}|| \leq M)\big\} = E\big\{||\boldsymbol{X}||)\big\}.$$

Since:

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

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

$$\lim_{M\to\infty} E\big\{||\boldsymbol{X}||I(||\boldsymbol{X}|| > M)\big\} = E||\boldsymbol{X}|| - \lim_{M\to\infty} E\big\{||\boldsymbol{X}||I(||\boldsymbol{X}|| \leq M)\big\} = 0.$$

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

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

and:

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

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

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

Thus $\lim_{M\to\infty} E\big\{||\boldsymbol{X}||I(||\boldsymbol{X}|| > M)\big\} \neq 0$. ■

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

i. $\sup_{t\in\mathcal{T}} E||\boldsymbol{X}_t|| < \infty$,

ii. For $\forall\epsilon > 0$ there $\exists\delta > 0$ s.t.

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

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

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

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

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

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

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

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

Observe that by Markov's inequality:

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

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

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

as required. ■

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

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

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

Proof.

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

as $M \to \infty$. ■

Theorem 1.6.3. The following statements are equivalent:

i. $\boldsymbol{X}_n \xrightarrow{L^1} \boldsymbol{X}$,

ii. $(\boldsymbol{X}_n)$ is uniformly integrable, and $\boldsymbol{X}_n \xrightarrow{p} \boldsymbol{X}$.

□

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

$$E\big\{||\boldsymbol{X}||\big\} = E\big\{\liminf_{n_k\to\infty} ||\boldsymbol{X}_{n_k}||\big\} \leq \liminf_{n_k\to\infty} E||\boldsymbol{X}_{n_k}|| \leq \sup_{n\in\mathbb{N}} E||\boldsymbol{X}_n|| < \infty,$$

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:

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

The first term is bounded as:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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:

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

♦

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

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

Similarly:

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

Thus:

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

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

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

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

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

■

Convergence in Distribution

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

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

■

7.1 Relation to Convergence in Probability

Proposition 1.7.1. Convergence in distribution to a constant $\boldsymbol{\alpha}$ implies convergence in probability:

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

♦

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

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

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

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

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

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

■

7.2 Characteristic Functions

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

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

■

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

$$\boldsymbol{X}_n \xrightarrow{d} \boldsymbol{X} \iff \phi_n \to \phi.$$

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

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. $E\{g(X_n)\} \to E\{g(X)\}$ for every bounded, continuous function g.

iii. $E\{\mathcal{L}(X_n)\} \to E\{\mathcal{L}(X)\}$ 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 $\boldsymbol{X}_n \xrightarrow{d} \boldsymbol{X}$, then there exists a sequence of random variables $(\boldsymbol{\xi}_n)$ *defined on a common probability space* such that $\boldsymbol{X}_n \overset{d}{=} \boldsymbol{\xi}_n$, $\boldsymbol{X} \overset{d}{=} \boldsymbol{\xi}$, and $\boldsymbol{\xi}_n \xrightarrow{as} \boldsymbol{\xi}$. □

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

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

$$\lim_{n\to\infty} E(\boldsymbol{X}_n) = E(\boldsymbol{X}).$$

□

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

$$\lim_{n\to\infty} E(\boldsymbol{X}_n) = \lim_{n\to\infty} E(\boldsymbol{\xi}_n) = E(\boldsymbol{\xi}) = E(\boldsymbol{X}).$$

■

Summary of Convergence Relations

- Convergence almost surely implies convergence in probability:

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

- Convergence almost surely may be established by checking:
  - The series $\sum_{n=1}^{\infty} \mathbb{P}(||\boldsymbol{X}_n - \boldsymbol{X}|| > \epsilon)$ is finite for $\forall \epsilon > 0$.
  - The series $\sum_{n=1}^{\infty} E||\boldsymbol{X}_n - \boldsymbol{X}||^p$ is finite for some $p > 0$.

- Convergence in L^p implies convergence in probability:

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

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

- Convergence in probability implies convergence in distribution:

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

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

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

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

$$\lim_{n\to\infty} E(\boldsymbol{X}_n) = E(\boldsymbol{X}).$$