Introduction

Remark 1.1.1. This document considers stochastic order notation, limits of sets, and the different modes of convergence for random variables: almost sure, in L^p, in probability, and in distribution. 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 a *common* probability space $(\Omega, \mathscr{F}, P)$. ♦

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:

$$P\big(||\boldsymbol{X}_n||_2 > M_\epsilon\big) < \epsilon.$$

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 $\forall \epsilon > 0$:

$$\lim_{n \to \infty} \mathbb{P}(||\boldsymbol{X}_n||_2 > \epsilon) = 0.$$

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 $P(||\boldsymbol{X}_n||_2 > \epsilon) < \delta$. ■

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 $n \geq \nu_\delta$, $\mathbb{P}(||\boldsymbol{X}_n||_2 > \epsilon) < \delta$. ■

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

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

♦

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

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

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

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

By sub-additivity of the probability measure:

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

■

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 $n \geq \nu_X, \mathbb{P}(||X_n||_2 > M_X) < \epsilon/2$ and when $n \geq \nu_Y, \mathbb{P}(||Y_n||_2 > M_Y) < \epsilon/2$. Set $M = \max(M_X, M_Y)$, then:

$$\mathbb{P}\big(||X_n + Y_n||_2 > 2M\big) \leq \mathbb{P}\big(||X_n||_2 > M\big) + \mathbb{P}\big(||Y_n||_2 > M\big) \leq \epsilon/2 + \epsilon/2 = \epsilon.$$

ii.

$$\mathbb{P}\big(||X_nY_n||_2 > M_XM_Y\big) \leq \mathbb{P}\big(||X_n||_2 > M_X \cup ||Y_n||_2 > M_Y\big)$$
$$\leq \mathbb{P}\big(||X_n||_2 > M_X\big) + \mathbb{P}\big(||Y_n||_2 > M_Y\big) \leq \epsilon/2 + \epsilon/2 = \epsilon.$$

■

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 $\forall(M, \delta)$,

$$\mathbb{P}\big(||X_nY_n||_2 > M\big) = \mathbb{P}\big(||X_nY_n > \delta \cap ||Y_n||_2 > M) + \mathbb{P}\big(||X_nY_n||_2 \leq \delta \cap ||Y_n||_2 \leq M\big)$$
$$\leq \mathbb{P}\big(||Y_n||_2 > M\big) + \mathbb{P}\big(||X_n||_2 > \delta/M\big)$$

Since $Y_n = \mathcal{O}_p(1)$, for any $\epsilon > 0$ there $\exists M_\epsilon$ s.t. when $n \geq \nu_\epsilon$, $\mathbb{P}(||Y_n||_2 > M_\epsilon) < \epsilon$. Moreover, since $X_n = o_p(1)$, $\lim_{n\to\infty}\mathbb{P}(||X_n||_2 > \delta/M) = 0$. Thus:

$$\lim_{n\to\infty}\mathbb{P}\big(||X_nY_n||_2 > M_\epsilon\big) \leq \epsilon + \lim_{n\to\infty}\mathbb{P}\big(||X_n||_2 > \delta/M\big) = \epsilon.$$

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. $||\boldsymbol{X}_n||_2^r = o_p(\alpha_n^r)$ where $r > 0$.

□

Proof. i. If $||\boldsymbol{X}_n + \boldsymbol{Y}_n||_2/\max(\alpha_n, \beta_n) > \epsilon$, then either:

$$\frac{||\boldsymbol{X}_n||_2}{\alpha_n} > \frac{\epsilon}{2} \vee \frac{||\boldsymbol{Y}_n||_2}{\beta_n} > \frac{\epsilon}{2}.$$

By subadditivity:

$$\lim_{n\to\infty}\mathbb{P}\left\{\frac{||\boldsymbol{X}_n + \boldsymbol{Y}_n||_2}{\max(\alpha_n, \beta_n)} > \epsilon\right\} \leq \lim_{n\to\infty}\mathbb{P}\left(\alpha_n^{-1}||\boldsymbol{X}_n||_2 > \frac{\epsilon}{2}\right) + \mathbb{P}\left(\beta_n^{-1}||\boldsymbol{Y}_n||_2 > \frac{\epsilon}{2}\right) = 0.$$

ii. If $||\boldsymbol{X}_n\boldsymbol{Y}_n||_2/(\alpha_n\beta_n) > \epsilon$, then either:

$$\left\{\alpha_n^{-1}||\boldsymbol{X}_n||_2 \leq 1,\ \beta_n^{-1}||\boldsymbol{Y}_n||_2 > \epsilon\right\} \bigcup \left\{\alpha_n^{-1}||\boldsymbol{X}_n||_2 > 1, (\alpha_n\beta_n)^{-1}||\boldsymbol{X}_n\boldsymbol{Y}_n||_2 > \epsilon\right\}.$$

By subadditivity:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{||\boldsymbol{X}_n\boldsymbol{Y}_n||_2}{\alpha_n\beta_n} > \epsilon\right) \leq \lim_{n\to\infty} \mathbb{P}\left(\beta_n^{-1}||\boldsymbol{Y}_n||_2 > \epsilon\right) + \mathbb{P}\left(\alpha_n^{-1}||\boldsymbol{X}_n||_2 > 1\right) = 0.$$

iii.

$$\lim_{n\to\infty} \mathbb{P}\left(\alpha_n^{-r}||\boldsymbol{X}_n||_2^r > \epsilon\right) = \lim_{n\to\infty} \mathbb{P}\left(\alpha_n^{-1}||\boldsymbol{X}_n||_2 > \epsilon^{1/r}\right) = 0.$$

■

Proposition 1.2.5.

$$\boldsymbol{X}_n - \boldsymbol{X} = o_p(1) \iff ||\boldsymbol{X}_n - \boldsymbol{X}||_2 = o_p(1).$$

♦

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

$$\lim_{n\to\infty} \mathbb{P}\big(||\boldsymbol{Y}_n||_2 > \epsilon\big) = 0,$$
$$\lim_{n\to\infty} \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\big) = 0.$$

Now let $Y_n = ||\boldsymbol{X}_n - \boldsymbol{X}||_2$, then $Y_n = o_p(1)$ if and only if for $\forall\epsilon > 0$:

$$\lim_{n\to\infty} \mathbb{P}\big(|Y_n| > \epsilon\big) = 0,$$
$$\lim_{n\to\infty} \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\big) = 0.$$

Thus, the statements $\boldsymbol{X}_n - \boldsymbol{X} = o_p(1)$ and $||\boldsymbol{X}_n - \boldsymbol{X}||_2 = o_p(1)$ 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:

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

■

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. ($\implies$):

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

$$\lim_{n\to\infty} \mathbb{P}(|X_{nj} - X_j| > \epsilon) \leq \lim_{n\to\infty} \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\big) = 0.$$

($\impliedby$):

$$\lim_{n\to\infty} \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\big) \leq \lim_{n\to\infty} \sum_{j=1}^{J} \mathbb{P}\big(|X_{nj} - X_j| > \epsilon/J\big) = 0.$$

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

$$\mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > M_\epsilon\big) < \epsilon.$$

Since $\mathbb{P}(|X_{nj} - X_j| > M_\epsilon) \leq \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > M_\epsilon\big)$, conclude $X_{nj} = \mathcal{O}_p(1)$.
For $\epsilon > 0$ there $\exists M_\epsilon, \nu_\epsilon$ s.t. when $n \geq \nu_\epsilon$:

$$\mathbb{P}(|X_{nj} - X_j| > M_\epsilon) \leq \mathbb{P}\left(\sqrt{\sum_{j=1}^{k}(X_{nj} - X_j)^2} > M_\epsilon\right) = \mathbb{P}\left(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > M_\epsilon\right) < \epsilon.$$

($\impliedby$) 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 $n \geq \nu$:

$$\mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > JM\big) \leq \sum_{j=1}^{J} \mathbb{P}\left(|X_{nj} - X_j| > M\right) \leq J \cdot \frac{\epsilon}{J} = \epsilon.$$

■

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 $\forall\epsilon > 0$:

$$\mathbb{P}\left\{\omega : \limsup_{n\to\infty} ||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\right\} = 0.$$

■

Remark 1.4.1. In the following, the notation:

$$\left\{\limsup_{n\to\infty} ||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\right\},$$

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

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

$$\begin{aligned}
0 &= \mathbb{P}\left\{\limsup_{n\to\infty} ||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\right\} \\
&= \lim_{n\to\infty} \mathbb{P}\left\{\sup_{k\ge n} ||\boldsymbol{X}_k - \boldsymbol{X}||_2 > \epsilon\right\} \\
&\ge \lim_{n\to\infty} \mathbb{P}\big\{||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\big\}.
\end{aligned}$$

■

Proposition 1.4.2. If for $\forall\epsilon > 0$:

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

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

Proof.

$$\mathbb{P}\left\{\limsup_{n\to\infty} ||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\right\} = \lim_{n\to\infty} \mathbb{P}\left\{\sup_{k\geq n} ||\boldsymbol{X}_k - \boldsymbol{X}||_2 > \epsilon\right\}$$
$$= \lim_{n\to\infty} \mathbb{P}\left\{\bigcup_{k\geq n} ||\boldsymbol{X}_k - \boldsymbol{X}||_2\right\} \leq \lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}\big(||\boldsymbol{X}_k - \boldsymbol{X}||_2 > \epsilon\big).$$

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

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}\big(||\boldsymbol{X}_k - \boldsymbol{X}||_2 > \epsilon\big) = 0.$$

■

Proposition 1.4.3. If $\exists(p > 0)$ such that:

$$\sum_{n=1}^{\infty} E||\boldsymbol{X}_n - \boldsymbol{X}||_2^p < \infty,$$

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

Proof.

$$\mathbb{P}\left\{\limsup_{n\to\infty} ||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\right\} \leq \lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}\big(||\boldsymbol{X}_k - \boldsymbol{X}||_2 > \epsilon\big)$$
$$\leq \frac{1}{\epsilon^p} \lim_{n\to\infty} \sum_{k=n}^{\infty} E||\boldsymbol{X}_n - \boldsymbol{X}||_2^p = 0.$$

■

L^p Convergence

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

$$\lim_{n\to\infty} E\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2^p\big) = 0.$$

■

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

$$\mathbb{P}\big(||\boldsymbol{X}||_2 \geq t\big) \leq \frac{E||\boldsymbol{X}||_2}{t}.$$

♦

Proof. Let $Y = ||\boldsymbol{X}||_2$,

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

■

Corollary 1.5.1. For $p > 0$,

$$\mathbb{P}\big(||\boldsymbol{X}||_2 \geq t\big) \leq \frac{E||\boldsymbol{X}||_2^p}{t^p}.$$

♣

Proposition 1.5.2. 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 $\forall \epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\big) \leq \frac{1}{t^p} \lim_{n\to\infty} E||\boldsymbol{X}_n - \boldsymbol{X}||_2^p = 0.$$

■

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 $\forall \epsilon > 0$:

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

■

Proposition 1.6.1. 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:

$$\begin{aligned}\mathbb{P}(\boldsymbol{X}_n \leq \boldsymbol{t}) &= \mathbb{P}\Big(\big\{\boldsymbol{X}_n \leq \boldsymbol{t}, ||\boldsymbol{X}_n - \boldsymbol{X}||_2 < \epsilon\boldsymbol{1}\big\} \cup \big\{\boldsymbol{X}_n \leq \boldsymbol{t}, ||\boldsymbol{X}_n - \boldsymbol{X}|| + 2 \geq \epsilon\boldsymbol{1}\big\}\Big) \\ &\leq \mathbb{P}\big(\boldsymbol{X} \leq \boldsymbol{t} + \epsilon\boldsymbol{1}\big) + \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\big).\end{aligned}$$

Similarly:

$$\mathbb{P}(\boldsymbol{X} \leq \boldsymbol{t} - \epsilon\mathbf{1}) \leq \mathbb{P}\big(\boldsymbol{X}_n \leq \boldsymbol{t}\big) + \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\big).$$

Thus:

$$\mathbb{P}(\boldsymbol{X} \leq \boldsymbol{t} - \epsilon\mathbf{1}) - \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\big) \leq \mathbb{P}(\boldsymbol{X}_n \leq \boldsymbol{t}) \leq \mathbb{P}\big(\boldsymbol{X} \leq \boldsymbol{t} + \epsilon\mathbf{1}\big) + \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{X}||_2 > \epsilon\big).$$

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

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

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

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

$\blacksquare$

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:

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

$\blacksquare$

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

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

$\blacklozenge$

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 $||\boldsymbol{X}_n - \boldsymbol{\alpha}||_2 < \epsilon$ is:

$$\mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{\alpha}||_2 < \epsilon\big) = \mathbb{P}\big(\boldsymbol{\alpha} - \epsilon\mathbf{1} \leq \boldsymbol{X}_n \leq \boldsymbol{\alpha} + \epsilon\mathbf{1}\big) = F_n(\boldsymbol{\alpha} + \epsilon\mathbf{1}) - F_n(\boldsymbol{\alpha} - \epsilon\mathbf{1}).$$

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

$$\lim_{n\to\infty} \mathbb{P}\big(||\boldsymbol{X}_n - \boldsymbol{\alpha}||_2 < \epsilon\big) = I(\boldsymbol{\alpha} + \epsilon\mathbf{1} > \boldsymbol{\alpha}) - I(\boldsymbol{\alpha} - \epsilon\mathbf{1} > \boldsymbol{\alpha}) = 1.$$

$\blacksquare$

Theorem 1.7.1 (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 $\boldsymbol{X}_n \stackrel{d}{=} \boldsymbol{\xi}_n$ and $\boldsymbol{\xi}_n \xrightarrow{as} \boldsymbol{\xi}$. $\square$

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 $\sum_{n=1}^{\infty} E||\boldsymbol{X}_n - \boldsymbol{X}||_2^p$ 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 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:

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