L-13

① $X_n \to X$ a.s. $\Leftrightarrow P(\{\omega \mid X_n(\omega) \to X(\omega)\}) = 1$

② $X_n \xrightarrow{P} X \Leftrightarrow \lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0 \ \forall \varepsilon > 0$.

Does $E[X_n]$ converge as $n \to \infty$??

Q) Does convergence of $E[X_n]$ tell us anything?

Example

$$X_n = \begin{cases} n & \text{w.p. } 1/n \\ 0 & \text{o.w.} \end{cases}$$

We have $X_n \geq 0$.
The expectation is $E[X_n] = n \cdot \frac{1}{n} + 0 \cdot (1 - \frac{1}{n}) = 1$ for all $n \geq 1$.

However, $P(X_n \neq 0) = 1/n \downarrow 0$.

This implies that $X_n \xrightarrow{P} 0$, since for any $\varepsilon > 0$, for large enough n, So, $P(|X_n - 0| > \varepsilon) = P(X_n = n) = 1/n \to 0$.

So we have $X_n \xrightarrow{P} 0$ but $E[X_n] = 1 \to 1$ as $n \to \infty$.

$$\sum_{n=1}^{\infty} P(X_n \neq 0) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

So if the sequence $(X_n)_{n\geq 1}$ consists of independent random variables, then by the 2nd Borel-Cantelli Lemma, we have $P(X_n \neq 0 \text{ i.o.}) = 1$, which means $P(X_n = n \text{ i.o.}) = 1$.

Definition: We say X_n converges to X in L_1 and write

$$X_n \xrightarrow{L_1} X$$

if $E[|X_n - X|] \to 0$ as $n \to \infty$.

Note: If $X_n \xrightarrow{L_1} X$, then $E[X_n] \to E[X]$ as $n \to \infty$. But need not be otherwise.

The proof follows from the triangle inequality for expectations:

$$|E[X_n] - E[X]| = |E[X_n - X]| \leq E[|X_n - X|]$$

Since $X_n \xrightarrow{L_1} X$, we can conclude $E[|X_n - X|] \to 0$ as $n \to \infty$.

Therefore, $|E[X_n] - E[X]| \to 0$, which implies $\lim_{n\to\infty} E[X_n] = E[X]$ using the above result.

The converse is not true.

For our example, we know $E[X_n] = 1$. Let's see if X_n converges in L_1 to the constant random variable $X = 1$.

Q: Does $X_n \xrightarrow{L_1} X = 1$ hold?

$E[|X_n - 1|] = ??$.
Recall

$$X_n = \begin{cases} n & \text{w.p. } 1/n \\ 0 & \text{o.w.} \end{cases}$$

$$\begin{aligned} E[|X_n - 1|] &= |n-1| \cdot P(X_n = n) + |0-1| \cdot P(X_n = 0) \\ &= (n-1) \cdot \frac{1}{n} + 1 \cdot \left(1 - \frac{1}{n}\right) \\ &= 1 - \frac{1}{n} + 1 - \frac{1}{n} \\ &= 2\left(1 - \frac{1}{n}\right) \to 2 \text{ as } n \to \infty. \end{aligned}$$

Since the limit is not 0, X_n does not converge to $X = 1$ in L_1.

Theorem: $X_n \xrightarrow{L_1} X \Rightarrow X_n \xrightarrow{P} X$.

But not the other way.

Example: The sequence $X_n = \begin{cases} n & \text{w.p. } 1/n \\ 0 & \text{o.w.w.p. } 1 - 1/n \end{cases}$ shows that convergence in probability does not imply convergence in L_1. We showed $X_n \xrightarrow{P} 0$, but $E[|X_n - 0|] = E[X_n] = 1$, which does not go to 0.

Proof of Theorem: For any fixed $\varepsilon > 0$, by Markov's inequality:

$$P(|X_n - X| > \varepsilon) \leq \frac{\mathbb{E}[|X_n - X|]}{\varepsilon} \to 0$$

Since $X_n \xrightarrow{L_1} X$, $E[|X_n - X|] \to 0$. Thus, $P(|X_n - X| > \varepsilon) \to 0$.

Summary of Convergence Modes:

① Almost Sure Convergence: $X_n \to X$ a.s.

② Convergence in Probability: $X_n \xrightarrow{P} X$

③ Convergence in L_p: $X_n \xrightarrow{L_p} X$, $p \geq 1$

Relationships: ① ⇒ ②, ③ ⇒ ②, ② ⇏ ①, ② ⇏ ③.

Definition: Let $p > 1$ and consider the set of random variables X such that $\mathbb{E}[|X|^p] < \infty$. We will write this set as:

$$L_p = \{X : \Omega \to \mathbb{R} \,|\, \mathbb{E}[|X|^p] < \infty\}.$$

This is called the L_p *space*.

On this space, we define a notion of convergence:

$$X_n \xrightarrow{L_p} X \quad \Longleftrightarrow \quad \lim_{n\to\infty} \mathbb{E}\left[|X_n - X|^p\right] = 0,$$

In particular when $p = 1$ (L_1 convergence) and $p = 2$ (L_2 or mean-square convergence). Two special values of p that we will need are $p = 1$ and $p = 2$.

Theorem: For $p \geq 1$, $X_n \xrightarrow{L_p} X \Rightarrow X_n \xrightarrow{P} X$.
Proof: For any $\varepsilon > 0$,

$$\begin{aligned} P(|X_n - X| > \varepsilon) &= P(|X_n - X|^p > \varepsilon^p) \\ &\leq \frac{E[|X_n - X|^p]}{\varepsilon^p} \quad \text{(by Markov's inequality)} \end{aligned}$$

As $n \to \infty$, the numerator $E[|X_n - X|^p] \to 0$ by the definition of L_p convergence. Therefore, $P(|X_n - X| > \varepsilon) \to 0$.

L_2 - Convergence and its relation with WLLN
Recall the Weak Law of Large Numbers (WLLN): If $X_1, X_2, \ldots$ are i.i.d. random variables with mean μ and finite variance σ^2, and $S_n = X_1 + \cdots + X_n$, then $S_n/n \xrightarrow{P} \mu$ as $n \to \infty$.

We can show a stronger result using L_2 convergence. Consider the mean squared error:

$$E\left[\left(\frac{S_n}{n} - \mu\right)^2\right] = \mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{\mathrm{Var}(S_n)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

As $n \to \infty$,

$$E\left[\left(\frac{S_n}{n} - \mu\right)^2\right] = \frac{\sigma^2}{n} \to 0.$$

This is precisely the definition of convergence in L_2.

Therefore, $S_n/n \to \mu$ in L_2.

L-14 Notes on L_p Spaces

L_p Space Definition

Given a probability space $(\Omega, \mathcal{F}, P)$.

Definition: A set $L_p = \{X : \Omega \to \mathbb{R} \,| \mathbb{E}[|X|^p] < \infty\}$ is called the L_p space associated with $(\Omega, \mathcal{F}, P)$.

Definition: A sequence of random variables $\{X_n\}_{n=1}^{\infty}$ from L_p we say " X_n converges to X in L_p", and write as $X_n \xrightarrow{L_p} X$, if $\mathbb{E}[|X_n - X|^p] \to 0$ as $n \to \infty$.

Vector Space

A set $V \neq \phi$ with two operations:

1. **Vector Addition**: $+ : V \times V \to V$, such that $(\underline{x}, \underline{y}) \mapsto \underline{x} + \underline{y}$. For $V = \mathbb{R}^d$, if $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}, \underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{pmatrix}$, then $\underline{x} + \underline{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_d + y_d \end{pmatrix}$.

2. **Scalar Multiplication**: $\cdot : \mathbb{R} \times V \to V$, such that $(c, \underline{v}) \mapsto c \cdot \underline{v}$. For $V = \mathbb{R}^d$, if $c \in \mathbb{R}$, then $c \cdot \underline{x} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_d \end{pmatrix}$.

The pair $(V, +)$ has the following properties:

- Associativity
- Commutativity
- Identity element: $\exists \underline{0} \in V$ s.t. $\underline{v} + \underline{0} = \underline{0} + \underline{v} = \underline{v}$
- Inverse element: $\forall \underline{v} \in V, \exists \underline{u} \in V$ such that $\underline{v} + \underline{u} = \underline{u} + \underline{v} = \underline{0}$. Such a $\underline{u}$ will often be denoted by $(-\underline{v})$. Together these properties make $(V, +)$ a commutative group.

Properties of scalar multiplication:

- **Associativity**: $c_1(c_2 \cdot \underline{v}) = (c_1 c_2)\underline{v}$
- **Distributivity**: $c \cdot (\underline{u} + \underline{v}) = c \cdot \underline{u} + c \cdot \underline{v}$
- **Identity**: $1 \cdot \underline{v} = \underline{v}$, where $1 \in \mathbb{R}$

The zero vector and the inverse vector in $\mathbb{R}^d$ are:

$$\underline{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad -\underline{x} = \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_d \end{pmatrix}$$

Theorem: Let $p \geq 1$. The space L_p is a vector space under the usual addition of random variables and multiplication of a random variable and real number.

Given X, Y are random variables and $c \in \mathbb{R}$:

- Point by point addition: $(X+Y) : \Omega \to \mathbb{R}$ is defined by $\omega \mapsto X(\omega) + Y(\omega)$.
- Scalar multiplication: $(c \cdot X) : \Omega \to \mathbb{R}$ is defined by $\omega \mapsto c \cdot X(\omega)$.

To show that L_p is a vector space, we must show closure under addition and scalar multiplication i.e.

1. If $X, Y \in L_p$, then $X + Y \in L_p$.
2. If $c \in \mathbb{R}$ and $X \in L_p$, then $c \cdot X \in L_p$.

Also, note that $L_p \neq \phi$ because the zero random variable $X \equiv 0$ is in L_p, as $\mathbb{E}[|0|^p] = 0 < \infty$.

If $X, Y \in L_p$, it means $\mathbb{E}[|X|^p] < \infty$ and $\mathbb{E}[|Y|^p] < \infty$. We need to show $\mathbb{E}[|X + Y|^p] < \infty$.

Consider the case when $p = 1$ (L_1):
Given $\mathbb{E}[|X|] < \infty$ & $\mathbb{E}[|Y|] < \infty$. We must show $\mathbb{E}[|X + Y|] < \infty$. The triangle inequality will do: $\mathbb{E}[|X + Y|] \leq \mathbb{E}[|X| + |Y|] = \mathbb{E}[|X|] + \mathbb{E}[|Y|] < \infty$.

Consider the case when $p = 2$ (L_2):
Given $\mathbb{E}[|X|^2] < \infty, \mathbb{E}[|Y|^2] < \infty$. We must show $\mathbb{E}[(X + Y)^2] < \infty$.

Using the Cauchy-Schwarz inequality: $|E[XY]| \leq \sqrt{E[X^2]E[Y^2]} < \infty$.

$$\mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2 + 2XY + Y^2] = \mathbb{E}[X^2] + 2\mathbb{E}[XY] + \mathbb{E}[Y^2] < \infty$$

For general $p \geq 1$:
Consider $a, b \geq 0$. We can establish the inequality $(a+b)^p \leq 2^p(a^p + b^p)$. Since $(a+b) \leq 2\max(a,b)$, we have $(a+b)^p \leq (2\max(a,b))^p = 2^p \max(a^p, b^p) \leq 2^p(a^p + b^p)$.

For general $p \geq 1$, consider

$$X, Y > 0$$

Step 1:

$$(X+Y)^p \leq 2^p \left(X^p + Y^p\right)$$

$$E\left[(X+Y)^p\right] \leq 2^p \left(E[X^p] + E[Y^p]\right) < \infty$$

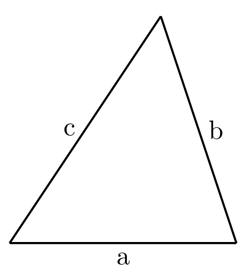
Step 2: General X, Y

$$E\left[|X+Y|^p\right] \leq E\left[|X| + |Y|\right]^p$$

$$\leq E\left[|X|^p\right] + E\left[|Y|^p\right] < \infty$$

Triangle Inequality:

$$|a+b| \leq |a| + |b|$$



Closure under scalar multiplication:
Let $c \in \mathbb{R}$ and $X \in L_p$.

$$\mathbb{E}[|c \cdot X|^p] = \mathbb{E}[|c|^p|X|^p] = |c|^p\mathbb{E}[|X|^p] < \infty$$

This shows $c \cdot X \in L_p$. If

$$E\big[|c \cdot x|^p\big] = E\big[|c|^p|x|^p\big] = |c|^p E\big[|x|^p\big] < \infty \quad \square$$

$$\Omega = \{A\}$$

$$X : \Omega \to \mathbb{R}$$

$$X(A) \in \mathbb{R}$$

Set of r.v's is same as $\mathbb{R}$.

$$L_p = \text{all r.v's on } \Omega.$$

The set of random variables on Ω, and thus the space L_p, can be very large.

- If $\Omega = \{H, T\}$, a random variable $X : \Omega \to \mathbb{R}$ is defined by the pair $(X(H), X(T))$. The set of all such random variables is isomorphic to $\mathbb{R}^2$. So L_p is isomorphic to $\mathbb{R}^2$.
- Ω is 'very large'.

  $X : \Omega \to \mathbb{R}$ plenty of r.v.s.

  L_p may be having many very different r.v.s.

  $f : \Omega \to S$

  S^Ω

L_p Norm

For any $X \in L_p$, we can define a function, often written as $|| \cdot ||_p$:

$$||X||_p = \left(\mathbb{E}[|X|^p]\right)^{1/p}$$

The function $|| \cdot ||_p : L_p \to [0, \infty)$ is defined as $X \mapsto ||X||_p = (\mathbb{E}[|X|^p])^{1/p}$.

In particular:

- For $p = 1$, $||X||_1 = \mathbb{E}[|X|]$
- For $p = 2$, $||X||_2 = \sqrt{\mathbb{E}[X^2]}$

Note: If $\mathbb{E}[X] = 0$, then $||X||_2$ is the standard deviation of X, $SD(X)$.

Properties of a Norm

In particular

For $p = 1$:

$$\|X\|_1 = E[|X|]$$

For $p = 2$:

$$\|X\|_2 = \sqrt{E[X^2]}$$

Note: if $E[X] = 0$ then $\|X\|_2 = SE(X)$.

NORM

1. Non-negative
2. Follows Triangle inequality
3. Non-zero if vector is not a null vector.
4. $\|c \cdot v\| = |c|\|v\|$

1. For $c \in \mathbb{R}$, $X \in L_p$:

$$\begin{aligned} \|c \cdot X\|_p &= \left(E\left[|c \cdot X|^p\right]\right)^{\frac{1}{p}} \\ &= \left(|c|^p E\left[|X|^p\right]\right)^{\frac{1}{p}} \\ &= |c| \cdot \|X\|_p \end{aligned}$$

2. If $\|X\|_p = 0$:

$$\begin{aligned} \|X\|_p = 0 &\implies \left(E\left[|X|^p\right]\right)^{\frac{1}{p}} = 0 \\ &\iff E\left[|X|^p\right] = 0, \text{ a.s.} \\ &\iff X = 0, \text{ a.s.} \end{aligned}$$

The NORM requirement fails. :(

L-14 ends