University of Sydney

STAT4528:

Probability and Martingale theory

Professor: Sasha Fish

Notes By: Yining Ding

First Semester Honours Year

Last Updated: April 1, 2022

Contents

0 Introduction

Modern Probability comes with measure
Axiomatic approach to Probability space(devoted by Kolmogorov)
Probability space $(\Omega, \mathcal{F}, \mathbb{P})$

- a space or set Ω consist of points $\omega \in \Omega$
- a class $\mathcal{F}$ of subsets of Ω are called ($\sigma-$algebra)
- A Probability measure on $(\Omega, \mathcal{F})$

Terminology:

- Ω: sample space, all possible outcomes
- ω: sample point/ elementary outcome
- event: An element of $\mathcal{F}$ (Measurable sets)
- $\mathbb{P}$: Measure on measurable space

1 Measuralbe set sigma algebra

$(\Omega, \mathcal{F})$ is called measurable space
any set $A \in \mathcal{F}$ is a measurable set
Def. A family $\mathcal{F}$ of subsets of Ω is said to be $\sigma - algebra$ on Ω if:

A.1 $\Omega \in \mathcal{F}$

A.2 $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

A.3 $A_1, A_2, ... \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

Note: "Sigma" means countable, replace (A.3) with finite the definition becomes field/algebra

1.1 Properties of sigma algebra

, let $\mathcal{F}$ be a $\sigma algebra$

1. $A_n \in \mathcal{F} \Rightarrow \cap_{n=1}^{\infty} A_n \in \mathcal{F}$
2. $\emptyset \in \mathcal{F}$
3. (Finite additivity)$A_1, A_2, ...A_N \in \mathcal{F} \Rightarrow \cap_{n=1}^{N} and \cup_{n=1}^{N} A_n \in \mathcal{F}$
4. $A \subset B, A, B \in \mathcal{F} \Rightarrow B \setminus A \in \mathcal{F}$
5. $A, B \in \mathcal{F} \Rightarrow A \triangle B \in \mathcal{F}$

Exercise 1.1. Show $\mathcal{F}$ is a σ-algebra

1. smallest σ-algebra: $\mathcal{F} = \{\emptyset, \Omega\}$
2. Power set: $\mathcal{F} = 2^{\Omega}$
3. $\Omega = \{1, 2, 3\}, \mathcal{F} = \{\emptyset, \Omega, \{1\}, \{2, 3\}\}$
4. $\Omega = [0, 1], \mathcal{F} = \{A | A \text{ or } A^c \text{ is countable}\}$

Exercise 1.2. Is $\mathcal{F}$ a σ-algebra in the following examples ?

1. $\Omega = [0,1], \mathcal{F} = \{F \subset [0,1] | F \text{ is closed set}\}$
2. $\Omega = [0,1], \mathcal{F} = \{G \subset (0,1) | G \text{ is open set}\}$

Theorem 1.1. *(σ-algebra generated by $\mathcal{G}$)*
For any family $\mathcal{G}$ of subsets of Ω, there exists a unique σ-algebra, denoted by $\sigma(\mathcal{G})$ and called the σ-algebra generated by $\mathcal{G}$ .s.t

1. $\mathcal{G} \subset \sigma(\mathcal{G})$
2. *if $\mathcal{H}$ is a σ-algebra and $\mathcal{G} \subset \mathcal{H}$, then $\sigma(\mathcal{G}) \subset \mathcal{H}$*

That is: $\sigma(\mathcal{G})$ is the smallest σ-algebra on Ω containing $\mathcal{G}$ *Proof.*

Exercise 1.3. Consider following corollary of Definition of σ-algebra generated by $\mathcal{G}$

1. if A is σ-algebra, then $\sigma(A) = A$
2. if $A \subset A' \Rightarrow \sigma(A) \subset \sigma(A')$
3. if $A \subset A' \subset \sigma(A) \Rightarrow \sigma(A') = \sigma(A)$

Def: if $\Omega = \mathbf{R}$ or any other topological space

$$\mathcal{B} = \sigma(\mathcal{G}|\mathcal{G} \text{ is open})$$

is called the Borel sigma field
Lemma: $\mathcal{B} = \sigma(\{(-\infty, x], x \in \mathcal{R}\}$:= A is equivalent definition of Borel set as above

Exercise 1.4. show that the following set are all equivalent in the sense that Borel σ-algebra on $\mathbb{R}$ is generated.

1. closed subsets of $\mathbb{R}$ 2. $\sigma(\{[a,b], a, b \in \mathcal{R}\}$ 3. $\sigma(\{(a,b], a, b \in \mathcal{R}\}$
4. $\sigma(\{(-\infty, b], b \in \mathbb{Q}\}$ 5. $\sigma(\{[a,b], a, b \in \mathbb{Q}\}$ 6. $\sigma(\{(a,b], a, b \in \mathbb{Q}\}$

2 Measure

After defining the measurable set on Ω, next we consider about the measure **Def:** Let $(\Omega, \mathcal{F})$ be a measurable space, A set function $\mu : \mathcal{F} \to [0, +\infty)$ is a measure on $(\Omega, \mathcal{F})$ if:

1. $\mu(\emptyset) = 0$
2. if $A_1, A_2, ... \in \mathcal{F}$ are disjoint then (Countable additivity)

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$$

Then the triple $(\Omega, \mathcal{F}, \mu)$ is called a measure space

Note: if $\mu(\Omega) = 1$, then the measure is called Probability measure denoted by $\mathbb{P}$, the triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space with following properties

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
2. $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(B)$ if $A \subset B$
3. $\mathbb{P}(A) \leq 1$

Def. A measure μ is finite if $\mu(\Omega) < \infty$
A measure μ is σ-finite if $A_1, A_2, ... \in \mathcal{F}$ with $\bigcup\limits_{n=1}^{\infty} A_n = \Omega, \quad \mu(A_n) < \infty, \forall n$

Exercise 2.1. List few measure examples

1. **classical probability measure**: $\Omega = \{1, 2, 3\}, \mathcal{F} = \{\emptyset, \Omega, \{1\}, \{2, 3\}\}, \mathbb{P}(A) = \frac{|A|}{|\Omega|}$
2. **Counting measure: Power set**: $\mathcal{F} = 2^{\Omega} \quad \mu(A) = |A|$ (, this is not probability measure)
3. **Discrete probability measure**: $\Omega = \{\omega_1, \omega_2, ...\}$ countable with $\mathcal{F} = 2^{\Omega}$ be the power set, $\mathbb{P}(A) = \sum\limits_{n:\omega_n \in A} p_n$
   where $p_n \geq 0$with $\sum\limits_{n:\geq 1} p_n = 1$

2.1 properties of measure

let $(\Omega, \mathcal{F}, \mu)$ be a measure space s.t. $A, B, A_1, A_2, ... in\mathcal{F}$ then:

1. if $A_1, A_2, ...A_N \in \mathcal{F}$ are disjoint then (Finite additivity)

$$\mu(\bigcup_{n=1}^{N} A_n) = \sum_{n=1}^{N} \mu(A_n)$$

2. $A \subset B \Rightarrow \mu(A) \leq \mu(B)$
3. $A \subset A_n \bigcup_n \Rightarrow \mu(A) \leq \sum_{n\geq 1} \mu(A_n)$
4. (Inclusion-exclusion formula) if $\mu(\Omega) < \infty$, for $A_1, A_2, ...A_n \in \mathcal{F}$

$$\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i) - \sum_{1\leq i<j\leq n} \mu(A_i \cap A_j) + \sum_{1\leq i<j<k\leq n} \mu(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1}\mu(A_1 \cap \cdots \cap A_n)$$

Continuity

5. if $A_n \uparrow A (A_n \subset A_{n+1}, \bigcup A_n = A)$, then $\mu(A_n) \uparrow \mu(A)$

6. if $A_n \downarrow A (A_{n+1} \subset A_n, \bigcap A_n = A)$, with $\mu(A_1) < \infty$ then $\mu(A_n) \downarrow \mu(A)$

Q: Do we need $\mu(A_1) < \infty$ **?** Hint: Yes

3 Construction of Lebesgue measure on real number

Non trivial question, How one can define "measurable set" on $\mathbf{R}$ that satisfy the condition of measure

Def. A family $\mathcal{F}_0$ of subsets of 2^Ω is said to be algebra/field on if:

A.1 $\Omega \in \mathcal{F}_0$

A.2 $A \in \mathcal{F}_0 \Rightarrow A^c \in \mathcal{F}_0$

A.3 $A_1, A_2, ...A_n \in \mathcal{F}_0 \Rightarrow \bigcup_{k=1}^{n} A_k \in \mathcal{F}_0$ (Finite additivity)

Def: Let $(\Omega, \mathcal{F}_0)$ be a measurable space, A set function $\mu : \mathcal{F}_0 \mapsto [0, +\infty)$ is a (pre)measure on $(\Omega, \mathcal{F}_0)$ if:

1. $\mu(\emptyset) = 0$
2. if $(A_n)_{n=1}^{\infty} \in \mathcal{F}_0$ are disjoint and $\bigcup_{n\geq 1} A_n \in \mathcal{F}_0$ then

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$$

Note: the difference between "real" measure is that we require an infinite sequence. And the problem is does $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}_0$ at all?

Exercise 3.1. Show that such sequence (if $(A_n)_{n=1}^{\infty} \in \mathcal{F}_0$ are disjoint and $\bigcup_{n\geq 1} A_n \in \mathcal{F}_0$) can exist:

Now, let's consider an algebra but not σ-algebra and define a measure (Lebesgue measure) on it, then we will use Caratheodory's extension theorem to establish Lebesgue measure on σ-algebra

Def. $\mathcal{B}_0$: an algebra on $\Omega = (0,1]$ as defined below

$$\mathcal{B}_0 := \{\bigcup_{i=1}^{n}(a_i, b_i] : 0 \le a_i < b_i \le 1, n \in \mathbb{N}\}$$

Def. Legesgue's measure
let $A \in \mathcal{B}_0 \Rightarrow A = \bigcup_{i=1}^{n}(a_i, b_i]$ WLOG, $(a_i, b_i]$ are disjoint sets which is disconnected $(a_i > b_{i-1})$

$$\lambda(A) := \sum_{i=1}^{n}(b_i - a_i)$$

which is the sum of total length of interval
Next, WTS

1. $\mathcal{B}_0$ is an algebra but not σ-algebra
2. λ is a (pre)measure on $\mathcal{B}_0$

Theorem 3.1 (Carathodory's Extension theorem). *if μ is a σ-finite (pre)measure on an algebra $\mathcal{F}_0$, then μ has a unique extension to a measure on $\sigma(\mathcal{F}_0)$*

Corollary 3.1 (The Lebesgue measure on $((0,1), \mathcal{B}(0,1))$ is well defined). *Note:* $\sigma(\mathcal{B}_0) = \mathcal{B}(0,1))$

Construction of the extension:

Def (outer measure). A map $\mu^* : P(X) \mapsto [0, +\infty]$ with following property:

1. $\mu^*(\emptyset) = 0$
2. μ^* is monotonic: $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$
3. countably sub-additive
$$\mu^*(\bigcup_n A_n) \leq \sum_n \mu^*(A_n)$$

Natural question: How to "build" a measure out of outer measure?
Notice: Power set won't work, we have to construct a σ-algebra out of it, sub-additivity won't work, we have to make equality holds.

Def . A set $A \in P(\Omega)$ is called μ^*-measurable if

$$\mu^*(A \cap E) + \mu^*(A^c \cap E) = \mu^*(E) \quad \text{for every set E}$$

Remark. By sub-additivity of μ^* we have

$$\mu^*(E) \leq \mu^*(A \cap E) + \mu^*(A^c \cap E) \quad \text{for every set E}$$

Hence A is μ^*-measurable iff

$$\mu^*(A \cap E) + \mu^*(A^c \cap E) \leq \mu^*(E) \quad \text{for every set E}$$

if ν is a measure on $(\Omega, \mathcal{F})$ with $\mathcal{F}_0 \subset \mathcal{F}$ which agrees with μ on $\mathcal{F}_0$, then for $A \subset \mathcal{F}$ with $A \subset \bigcup_n A_n$, $A_n \subset \mathcal{F}_0$ we have the following:

$$\nu(A) \leq \sum_{n \geq 1} \nu(A_n) = \sum_{n \geq 1} \mu(A_n) \Rightarrow \nu(A) \leq \mu^*(A)$$

Hence μ^* is an upper bound

Now let's put things up, we have $\mathcal{B}_0$: an algebra on $\Omega = (0, 1]$ as defined below

$$\mathcal{B}_0 := \{\bigcup_{i=1}^{n}(a_i, b_i] : 0 \leq a_i < b_i \leq 1, n \in \mathbb{N}\}$$

And Lebesgue measure (premeasure) on $\mathcal{B}_0$.

For each subset $A \in P(\Omega)$, define its outer measure by

$$\mu^*(A) := inf\{\sum_n \mu(A_n) | A_n \in \mathcal{B}_0 \quad \forall n, A \subset \bigcup_n A_n\}$$

We denote by $\mathcal{M}$ a class of all μ^*-measurable sets.

Idea: We are estimating all subsets of power set by the information we have (premeasure on algebra) from outside.
Then we have to make outer measure a real measure, so we have to find a subset of power set to form σ-algebra and establish a measure on it. That is: WTS: $\mathcal{M}$ is sigma algebra that contains $\mathcal{B}_0$ and the "real" measure μ on $\mathcal{M}$ agrees with μ^*, this two will be sufficient

3.1 Existence

Lemma 3.1 (The class $\mathcal{M}$ is a field)**.**

Lemma 3.2 (If $A_1, A_2, ...$ are disjoint $\mathcal{M}$ sets, then for each $E \subset \Omega$)**.**

$$\mu^*(E \cap (\bigcup A_k)) = \sum_k \mu^*(E \cap A_k)$$

Lemma 3.3 (The class $\mathcal{M}$ is a σ-algebra, and μ^* restricted to $\mathcal{M}$ (denote by μ) is countably additive)**.**

Lemma 3.4 ($\mathcal{B}_0 \subset \mathcal{M}$)**.**

Lemma 3.5 ($A \in \mathcal{B}_0$ then $\mu(A) = \mu^*(A)$)**.**

Proof.

3.2 Uniqueness

$\pi - system$ **and** $\lambda - system$

Def ($\pi - system$)**.** For a collection of sets $\mathcal{P}$ if$A, B \in \mathcal{P} \Rightarrow A \cap B \in \mathcal{P}$
(closed under finite intersections)

Def ($\lambda - system$)**.** For a collection of sets λ if

(λ1) $\Omega \in \mathcal{L}$

(λ2) if $A, B \in \mathcal{L}, A \subset B \Rightarrow B \setminus A \in \mathcal{L}$

(λ3) if $A_n \in \mathcal{L}$ and $A_n \uparrow A \Rightarrow A \in \mathcal{L}$

Exercise 3.2. Show that $\mathcal{L}$ is $\lambda - system$ if

(λ1) $\Omega \in \mathcal{L}$

($\lambda 2'$) $A \in \mathcal{L} \Rightarrow A^c \in \mathcal{L}$

($\lambda 3'$) $A_1, A_2, ... \in \mathcal{L}$ $\quad disjoint \Rightarrow \bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{L}$

Note: The definition is different from sigma algebra and

$$(\lambda 1)(\lambda 2)(\lambda 3) \Longleftrightarrow (\lambda 1)(\lambda 2')(\lambda 3')$$

Exercise 3.3. Intersection of $\lambda - system$ is again $\lambda - system$

Exercise 3.4. $\lambda - system$ that is not $\sigma - algebra$

Exercise 3.5 (Important exercise)**.** Show that $\mathcal{F} \subset 2^{\Omega}$ is a σ-algebra iff $\mathcal{F}$ is both $\pi - system$ and $\lambda - system$

Theorem 3.2 (Dynkin's $\pi - \lambda$ Theorem)**.** *if $\mathcal{P}$ is a $\pi -$ system and $\mathcal{L}$ is a $\lambda -$ system, then*

$$\mathcal{P} \subset \mathcal{L} \Rightarrow \sigma(\mathcal{P}) \subset \mathcal{L}$$

Proof.

Theorem 3.3. *Suppose $\mathbb{P}_1$ and $\mathbb{P}_2$ are probability measure (or any other finite measure) on $\sigma(\mathcal{P})$ where $\mathcal{P}$ is a $\pi - system$, if $\mathbb{P}_1$ and $\mathbb{P}_2$ agree on $\mathcal{P}$, then they agree on $\sigma(\mathcal{P})$*

Proof. This can be generalized to any other finite measure

Exercise 3.6. Let $\mathcal{P}$ be $\pi - system$, if μ and ν are measure on $\sigma(\mathcal{P})$ that agree on $\mathcal{P}$ and there is a sequence $A_n \in \mathcal{P}$ with $A_n \uparrow \Omega$ and $\mu(A_n), \nu(A_n) < \infty$ then μ and ν agree on $\sigma(\mathcal{P})$

Note: This complete the uniqueness of Theorem 3.1, for σ-finite measure on an algebra $\mathcal{F}_0$,if there are two extension (measures) on $\sigma(\mathcal{F}_0)$ that agrees on $\mathcal{F}_0$ then they agree on $\sigma(\mathcal{F}_0)$ by Exercise3.6 which makes extension unique. This complete the proof of existence and uniqueness of Caratheodory's extension theorem $\square$

Exercise 3.7. Show that Lebesgue measure is translation invariant by 3.2 Dynkin's theorem

Theorem 3.4 (Stieltjes measure)**.** *Suppose* $F : \mathbb{R} \mapsto \mathbb{R}$ *is a non-decreasing and right continuous* $(F(x) = \lim_{X_n \downarrow x} F(x_n))$*, Then there exists a unique measure* μ *on* $(\mathbb{R}, \mathcal{B})$ *with* $\mu((a, b]) = F(b) - F(a)$

Remark: Choose $F(x) = x$ then we get Lebesgue measure.

Example 3.1. Non-decreasing and right continuous are both important

Completion of a measure space

Def . A measure space $(\Omega, \mathcal{F}, \mu)$ is complete if $A \subset B, B \in \mathcal{F}$ and $\mu(B) = 0 \Rightarrow A \in \mathcal{F}$, (Hence $\mu(A) = 0$ by monotonicity)

Exercise 3.8. if $(\Omega, \mathcal{F}, \mu)$ is complete, then $A \in \mathcal{F}, A \triangle A' \subset B \in \mathcal{F}$ with $\mu(B) = 0 \Rightarrow A' \in \mathcal{F}$ and $\mu(A) = \mu(A')$

Example 3.2. $(\Omega, \mathcal{M}, \mu^*)$ is a complete measure space

Theorem 3.5. *if $(\Omega, \mathcal{F}, \mu)$ is a measure space, then there is a complete measure space $(\Omega, \bar{\mathcal{F}}, \bar{\mu})$ called the completion of $(\Omega, \mathcal{F}, \mu)$ s.t.*

1. $E \in \bar{\mathcal{F}} \iff E = A \cup B$, *where* $A \in \mathcal{F}$ *and* $B \subset N \subset \mathcal{F}, \quad \mu(N) = 0$
2. $\bar{\mu}|_{\mathcal{F}} = \mu$

Note: The completion is unique

4 Sets and Measurability

Terminology:

Probability space $(\Omega, \mathcal{F}, \mathbb{P})$
Ω consists of all possible outcomes ω of an experiment and is called **sample space**, $\omega \in \Omega$ is called a **sample point**.
An element of $\mathcal{F}$ is called an **event**.
We say an **event** $A \in \mathcal{F}$ **occurred** if the outcome or a sample point $\omega \in \Omega$ satisfies $\omega\ inA$

Def . Event $A \in \mathcal{F}$ occurs almost surely (a.s.) if $\mathbb{P}(A) = 1$, equivalently, A occurs a.s. if $\mathbb{P}(A^c) = 0$

The later half definition is important, for example:

For general measure space, we say that S (measurable subset of Ω) holds almost everywhere (a.e.) if

$$\mu(S^c) = 0$$

Example 4.1. $(\mathbb{R}, \mathcal{B}, \lambda), S = \mathbb{R} \setminus \mathbb{Q}, \Rightarrow S^c = \mathbb{Q}$ and $\lambda(S^c) = \lambda(\mathbb{Q}) = 0$

4.1 limsup and liminf sets

Def . let $(A_n)_n$ be a sequence of events, we define the subset $\limsup A_n := \bigcap_{m \geq 1} \bigcup_{n \geq m} A_n$,

$$(B_n = \bigcup_{n \geq m} A_n, \quad B_n \downarrow \limsup A_n)$$

Similarly, $\liminf A_n := \bigcup_{m \geq 1} \bigcap_{n \geq m} A_n$

$$(C_n = \bigcap_{n \geq m} A_n, \quad C_n \uparrow \liminf A_n)$$

Remark.

$$\begin{aligned}
\limsup A_n &:= \bigcap_{m \geq 1} \bigcup_{n \geq m} A_n \\
&= \{\omega | \forall m \geq 1, \exists n = n(\omega) \geq m, \omega \in A_n\} \\
&= \text{The event } A_n \text{ occurs infinitely often (i.o.)} \\
\liminf A_n &:= \bigcup_{m \geq 1} \bigcap_{n \geq m} A_n \\
&= \{\omega | \exists m \geq 1, \forall n \geq m, \omega \in A_n\} \\
&= \text{The event } A_n \text{ eventually occurs}
\end{aligned}$$

Exercise 4.1. Show that

1. $\limsup A_n \in \mathcal{F}$ and $\liminf A_n \in \mathcal{F}$
2. $\liminf A_n \subset \limsup A_n$
3. $\liminf A_n^c = (\limsup A_n)^c$

Remark. If $\{x_n\}$ is a real sequence, then

$$\begin{aligned}\limsup x_n &:= \inf_m \{\sup_{n \geq m} x_n\} \quad \text{The sequence is non-increasing as m increasing} \\ &= \downarrow \lim_m \{\sup n \geq m\} \in [-\infty, +\infty]\end{aligned}$$

$$\begin{aligned}\liminf x_n &:= \sup_m \{\inf_{n \geq m} x_n\} \quad \text{The sequence is non-decreasing as m increasing} \\ &= \uparrow \lim_m \{\inf n \geq m\} \in [-\infty, +\infty]\end{aligned}$$

Theorem 4.1. $\exists \lim X_n \in [-\infty, +\infty] \Longleftrightarrow \liminf x_n = \limsup x_n$

Exercise 4.2. Recall the indicator function $\mathbb{1}_A(\omega)$, Prove that $\forall \omega \in \Omega$

1. $\limsup_n \mathbb{1}_{A_n}(\omega) = \mathbb{1}_{\limsup_n A_n(\omega)}$
2. $\liminf_n \mathbb{1}_{A_n}(\omega) = \mathbb{1}_{\liminf_n A_n(\omega)}$

Lemma 4.1 (The first Borel-Cantelli Lemma). *Let (A_n) be a sequence of events s.t.*

$$\sum_n \mathbb{P}(A_n) < \infty \textit{ then }, \mathbb{P}(\limsup_n A_n) = 0$$

Proof.

Example 4.2. What about the other half of Lemma 4.1 ? Is it true that (A_n) events

$$\sum_n \mathbb{P}(A_n) = \infty \text{ then }, \mathbb{P}(\limsup_n A_n) = 1 \quad \textit{No!}$$

$\Omega = [0,1], \mathcal{F} = \mathcal{B}, \mathbb{P} = \lambda$, let $A_n = \{(0, \frac{1}{n})\} \Rightarrow \mathbb{P}(A_n) = \frac{1}{n}$ and $\sum \mathbb{P}(A_n) = \infty$
But $\limsup_n A_n = \bigcap_{m\geq 1} \bigcup_{n \geq m} A_n = \bigcap_{m \geq 1} A_m = \bigcap_{m\geq 1} (0, \frac{1}{m}) = \{0\} \quad \Rightarrow \mathbb{P}(\limsup_n A_n) = \mathbb{P}(\{0\}) = 0$

Lemma 4.2. *For events $A_n \in \mathcal{F}$,*

1. $\mathbb{P}(\limsup_n A_n) \geq \limsup_n \mathbb{P}(A_n)$
2. $\mathbb{P}(\liminf_n A_n) \leq \liminf_n \mathbb{P}(A_n)$

Proof.

1. $\mathbb{P}(\limsup_n A_n) \geq \limsup_n \mathbb{P}(A_n)$

$$B_m = \bigcup_{n \geq m} A_n \downarrow \bigcap_m B_m = \limsup_n A_n \Rightarrow \mathbb{P}(B_n) \downarrow \mathbb{P}(\limsup_n A_n)$$
$$\mathbb{P}(\limsup_n A_n) = \lim \mathbb{P}(B_m) = \limsup_n \mathbb{P}(B_m) \geq \limsup_n \mathbb{P}(A_m)$$
$$\text{Since } A_m \subset B_m \Rightarrow \mathbb{P}(B_m) \geq \mathbb{P}(A_m)$$

2. $\mathbb{P}(\liminf_n A_n) \leq \liminf_n \mathbb{P}(A_n)$

$$B_m = \bigcap_{n \geq m} A_n \uparrow \bigcup_m B_m = \liminf_n A_n \Rightarrow \mathbb{P}(B_n) \uparrow \mathbb{P}(\liminf_n A_n)$$
$$\mathbb{P}(\liminf_n A_n) = \lim \mathbb{P}(B_m) = \liminf_n \mathbb{P}(B_m) \leq \liminf_n \mathbb{P}(A_m)$$
$$\text{Since } B_m \subset A_m \Rightarrow \mathbb{P}(B_m) \leq \mathbb{P}(A_m)$$

□

Lemma 4.3 (Fatou's Lemma for sets)**.**

$$\int \liminf_n \mathbb{1}_{A_n} d\mathbb{P} \leq \liminf_n \int \mathbb{1}_{A_n} d\mathbb{P}$$

Proof.

$$\mathbb{P}(\liminf_n A_n) = \int \liminf_n \mathbb{1}_{A_n} d\mathbb{P} \overset{(2)}{\leq} \liminf_n \mathbb{P}(A_n) = \liminf_n \int \mathbb{1}_{A_n} d\mathbb{P}$$
$$\Rightarrow \int \liminf_n \mathbb{1}_{A_n} d\mathbb{P} \leq \liminf_n \int \mathbb{1}_{A_n} d\mathbb{P}$$

□

4.2 Measurable functions

Recall that $f : (X, \tau) \mapsto (X', \tau')$ is continuous if $\forall$ open set $G \in \tau'$ we have $f^{-1}(G) \in \tau$

Def . let $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ be measurable spaces, then $f : (\Omega, \mathcal{F}) \mapsto (\Omega', \mathcal{F}')$ is measurable if $\forall A \in \mathcal{F}', f^{-1}(A) \in \mathcal{F}$

Example 4.3. The constant function $f \equiv c$ is always measurable (as a function $f : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$)

$$f^{-1}(A) = \begin{cases} \Omega & c \in A \\ \emptyset & c \notin A \end{cases}, \quad A \in \mathcal{B}$$

Def . A mapping $X : \Omega \mapsto \mathbb{R}$ is a random variable if it is a measurable function of $(\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$

Lemma 4.4. *Let $X : \Omega \mapsto \mathbb{R}$ be a map, and $g := \{X^{-1}(B) | B \in \mathcal{B}\}$, then g is a σ-algebra on Ω and it is the smallest σ-algebra w.r.t which X is a random variable*

Def . g from above lemma is called σ-algebra generated by X, it is denoted by $\sigma(X)$

Proof. g is the smallest σ-algebra

Lemma 4.5. *:*

1. *if* $f : \mathbb{R} \mapsto \mathbb{R}$ *is continuous then* $f : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$ *is measurable*
2. *if* $f : \mathbb{R} \mapsto \mathbb{R}$ *is monotone, then* $f : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$ *is measurable*

Proposition 4.1. *Suppose* X, Y *are RV's and* $f : (\mathbb{R}, \mathcal{B}) \mapsto (\mathbb{R}, \mathcal{B})$ *is measurable, then* $X + Y, XY, f(X)$ *are RV's*

Remark. $f : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$ is measruable $\iff \exists \mathcal{A} \subset \mathcal{F}', \sigma(\mathcal{A}) = \mathcal{F}$ s.t. $\forall A \in \mathcal{A}, f^{-1}(A) \in \mathcal{F}$

Proof. ($X + Y$ is RV:)
let $Z := X + Y$ then $Z^{-1}(-\infty, z) = \{\omega \in \Omega | Z(\omega) < z\} = \{\omega \in \Omega | X(\omega) + Y(\omega) < z\}$

Note: $X + y < z \iff x < z - y \iff \exists r \in \mathbb{Q} : x < r < z - y \iff \exists r \in \mathbb{Q} : x < r, y < z - r$

$$= \{\omega \in \Omega | \exists r \in \mathbb{Q} X(\omega) < r, Y(\omega) < z - r\} = \bigcup_{r \in \mathbb{Q}} \{\omega | X(\omega) < r, Y(\omega) < z - r\}$$
$$= \bigcup_{r \in \mathbb{Q}} \{\omega | X(\omega) < r\} \cap \{\omega | Y(\omega) < z - r\} \in \mathcal{F}$$

$\square$

($Z = f(X)$ is measurable)
For $B \in \mathcal{B}, Z^{-1}(B) = X^{-1}(f^{-1}(B)), f^{-1}(B) \in \mathcal{B}, X^{-1}(f^{-1}(B)) \in \mathcal{F}$

Exercise 4.3. $Z = XY$ is measurable

Why we concern measurable function rather than continuous function?The limit of a sequence of continuous function is generally not continuous, However the limit preserves measurability

In the context of limits, it is convenient to allow RV's to take the value $\pm\infty$

$$X : \Omega \mapsto \bar{\mathbb{R}} := [-\infty, +\infty]$$

is a RV if it is measurable as mapping from $(\Omega, \mathcal{F}) \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}})$,
where $\bar{\mathbb{R}}$ is called the extended real line: unions of intervals $[-\infty, x), (x, y), (y, +\infty]$
$\bar{\mathcal{B}}$ is the sigma algebra generated by open sets in $\bar{\mathbb{R}}$

Proposition 4.2. *If X_n are RV's then so are*

1. $\inf_n X_n$
2. $\sup_n X_n$
3. $\liminf_n X_n$
4. $\limsup_n X_n$
5. $\lim_n X_n \mathbb{1}_{\{\exists \lim_n X_n \in \mathbb{R}\}}$
6. $\lim_n X_n \mathbb{1}_{\{\exists \lim_n X_n \in \bar{\mathbb{R}}\}}$

Proof. :
1. Let $X = \inf_n X_n$ then

$$X^{-1}([-\infty, x)) = \{\omega | \inf_n X_n(\omega) < x\} = \{\exists n | X_n(\omega) < x\}$$
$$= \bigcup_n \{\omega | X_n(\omega) < x\} = \bigcup_n X_n^{-1}([-\infty, x)) \in \mathcal{F}$$

Given X_n are RVs, union of countable sets in $\mathcal{F}$ is again in $\mathcal{F}$
2. **Note:** $\sup_n X_n = -\inf_n(-X_n)$
Given X_n is RVs, so is $-X_n$, by similar reason, $\sup_n X_n$ is also RVs
3.

$$\liminf_n X_n = \sup_m \{\inf_{n \geq m} X_n\} = \sup_m Y_m$$
$$\text{Where } Y_m = \inf_{n \geq m} X_n \text{ This is RV by 1}$$
$$\liminf_n X_n = \sup_m Y_m \text{ This is sup of RV which is RV again by 2}$$

4. **Note:** $\limsup_n X_n = -\liminf_n(-X_n)$
Given X_n is RVs, so is $-X_n$, by similar reason,
5.

$$A = \{\omega | \limsup_n X_n(\omega) < \infty\} \in \mathcal{F} \text{ by 4}$$
$$B = \{\omega | \liminf_n X_n(\omega) > \infty\} \in \mathcal{F} \text{ by 3}$$
$$A \cap B \in \mathcal{F} \Rightarrow \mathbb{1}_{A \cap B} \text{ is RV}$$
$$\text{(Indicator function of measurable set is measurable function)}$$
$$\Rightarrow X := \mathbb{1}_{A \cap B}(\limsup_n X_n - \liminf_n X_n) \text{is well-defined RVs}$$
$$\Rightarrow E := \{\omega | X(\omega) = 0\} \cap A \cap B \in \mathcal{F}$$
$$= \{\omega | \exists \lim_n X_n, \lim_n X_n \in \mathbb{R}\}$$
$$\text{Hence } \lim_n X_n \mathbb{1}_E = \liminf_n X_n \mathbb{1}_E \text{ is RV}$$

6. Everything stays the same as in 5, except we have to include $\pm\infty$ in A and B

5 Random variables and distribution

Def . A mapping $X : \Omega \mapsto \mathbb{R}$ is a random variable if it is a measurable function of $(\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$

5.1 Dsitrbution function

Def . If X is a RV then it induces a Probability measure on $(\mathbb{R}, \mathcal{B})$, called its distribution:

$$\mu(A) = \mathbb{P}(X \in A), \forall A \in \mathcal{B}$$

We typically describe the distribution of X through distribution function (CDF):

$$F_X = \mathbb{P}(X \in (-\infty, t]) \quad t \in \mathbb{R}$$

Proposition 5.1. *F is a CDf iff:*

1. *F is non decreasing*
2. $\lim_{t \to \infty} F(t) = 1, \lim_{t \to -\infty} F(t) = 0,$
3. *F is right continuous,* $\forall a \in R, \lim_{t \downarrow a} F(t) = F(a)$

Example 5.1. Let U has uniform distribution on $[0, 1]$, i.e.

$$F_u(t) = \begin{cases} t, & t \in [0,1] \\ 0, & t < 0 \\ 1, & t > 1 \end{cases}$$

Let F be a function satisfy (1)(2)(3)

Def (quantile function)**.**

$$G(u) := inf\{t | F(t) \geq u\}, \quad u \in (0, 1)$$

G is generalised inverse of F

G is non-decreasing, left-continuous and $\{u | G(u) \leq t\} = \{u | F(t) \geq u\}$
Note G is measurable (Non decreasing function is always measurable), so $Y = G : (0, 1) \mapsto \mathbb{R}$ is a RV with DF F_Y satisfying:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\{u | G(u) \leq y\}) = \mathbb{P}(\{u | F(y) \geq u\}) = F(y)$$

$U \sim U[0, 1], X := Q(U) \sim F$
Proof. $\mathbb{P}(X \leq t) = \mathbb{P}(Q(U) \leq t) = \mathbb{P}(U \leq F(t)) = F(t)$
NB: If F is continuous and $X \sim F$, then also $F(X) \sim U[0, 1]$

5.2 Independency

Recall: Events $A_1, A_2, ..., A_n$ are independent if for any k distinct indices. Note: This could also be defined as indicator function of events which is consistent with next definition

$$\mathbb{P}(A_{i_1} \cap ... \cap A_{i_k}) = \prod_{j=1}^{k} \mathbb{P}(A_{i_j}) \quad 1 \leq i_1 < i_2 < ... < i_k \leq n$$

Def . The RV's, $X_1, ..., X_n$ are independent if for any $B_i \in \mathcal{B}, i = 1, ..., n$

$$\mathbb{P}(X_1 \in B_1, X_2 \in B_2, ..., X_n \in B_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in B_i)$$

Exercise 5.1. Consider following DF

1. Let $X_1, ..., X_n$ be independent RV's, what is the CDF of $Y = max\{X_1, ..., X_n\}$ ($F_Y = f(F_1, ..., F_n)$, where F_i is CDF of X_i)
2. CDF of $Z = min\{X_1, ..., X_n\}$
3. Take $X_i \sim Unif([0, 1])$

1.

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(max\{X_1, ..., X_n\} \leq y) \\ &= \mathbb{P}(X_1 \leq y, X_2 \leq y, ...X_n \leq y) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq y) = \prod_{i=1}^{n} F_i(y) \end{aligned}$$

2. Consider tail distribution.

$$\begin{aligned} 1 - F_Z(z) &= \mathbb{P}(Z \geq z) = \mathbb{P}(min\{X_1, ..., X_n\} \geq z) \\ &= \mathbb{P}(X_1 \geq z, X_2 \geq z, ...X_n \geq z) = \prod_{i=1}^{n} \mathbb{P}(X_i \geq z) = \prod_{i=1}^{n}(1 - F_i(z)) \\ F_Z(z) &= 1 - \prod_{i=1}^{n}(1 - F_i(z)) \end{aligned}$$

3. Take $X_i \sim Unif([0, 1])$,
$F_Y(y) = \prod_{i=1}^{n} F_i(y) = y^n$ if $y \in [0, 1]$
$F_Z(z) = 1 - \prod_{i=1}^{n}(1 - F_i(z)) = 1 - (1 - z)^n$ if $z \in [0, 1]$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space a set $g \subset 2^{\Omega}$ is sub-σ-algebra of $\mathcal{F}$ if g is a σ-algebra and $g \in \mathcal{F}$

Def . The sub-σ-algebra $g_1, ..., g_n$ are independent if whenever $A_i \in g_i, i = 1, 2, ..., n$ we have

$$\mathbb{P}(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} \mathbb{P}(A_i)$$

Exercise 5.2. Show that TFAE

1. If $A_1, ..., A_n$ are independent events then so are $A_1^c, ..., A_n^c$
2. Apply 1 to conclude that $A_1, ..., A_n$ are independent iff $B_1, ..., B_n$ are independent, where for each i, $B_i = A_i$ or $B_i = A_i^c$
3. For $A_1, ..., A_n \in \mathcal{F}$
   (a) A_i are independent events
   (b) The indicator function $\mathbb{1}_{A_i}$ are independent RV's
   (c) $g_i = \{\emptyset, \Omega, A_i, A_i^c\}$, and g_i are independent σ-algebra
   (d) The RVs $X_1, ..., X_n$ are independent iff $g_i = \sigma(X_i)$ are independent
   (e) If $g_1, ..., g_n$ are independent σ-algebra and if X_i is g_i measurable then $X_1, ..., X_n$ are independent RVs

These examples connects independent of RVs with independent of sigma algebras

Extension Let us extend the notion of independence to sets of events $\mathcal{A}_1, ..., \mathcal{A}_n$ where for each i, $\mathcal{A}_i \subset \mathcal{F}$ and $\Omega \in \mathcal{A}_i$

Def . $\mathcal{A}_i$ are independent if for any choice of $A_i \in \mathcal{A}_i, i = 1, ..., n$

$$\mathbb{P}(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} \mathbb{P}(A_i)$$

Exercise 5.3. $\mathcal{A}_1, ..., \mathcal{A}_n$ are independent $\Longleftrightarrow$ whenever $A_i \in \mathcal{A}_i, i = 1, ..., n$, A_i are independent

The following is the key lemma which allows us to check independence only on "generator"

Lemma 5.1. *Suppose that for $i = 1, ..., n$, $\mathcal{A}_i \subset \mathcal{F}$ satisfy*

1. $\Omega \in \mathcal{A}_i$
2. $\mathcal{A}_i$ *is a π-system (closed under intersection)*
3. $\mathcal{A}_1, ..., \mathcal{A}_n$ *are independent*

Then $\sigma(\mathcal{A}_1), \mathcal{A}_2, ..., \mathcal{A}_n$ are independent

Proof. (Using $\pi - \lambda$ theorem of Dynkin's)
For $A_i \in \mathcal{A}_i, i = 2, 3, ..., n$ define the set function μ and $\nu : \sigma(\mathcal{A}_1) \mapsto [0, 1]$ by

$$\mu(A) := \mathbb{P}(A \cap A_2 \cap ... \cap A_n)$$
$$\nu(A) := \mathbb{P}(A)\mathbb{P}(A_2)...\mathbb{P}(A_n)$$

μ and ν are measures on $\sigma(\mathcal{A}_1)$

1. $\mu(\emptyset) = \mathbb{P}(\emptyset) = 0, \nu(\emptyset) = \mathbb{P}(\emptyset)\mathbb{P}(A_2)...\mathbb{P}(A_n) = 0$
2. countable additivity is easy to get since $\mathbb{P}$ is countably additive.

Let $\mathcal{L} = \{A \in \sigma(\mathcal{A}_1) | \mu(A) = \nu(A)$ Then $\mathcal{L}$ is a λ-system.

1. $\Omega \in \mathcal{L}$ since $A_1, ..., A_n$ are independent.
2. $A, B \in \mathcal{L}$ and $A \subset B$
   $\mu(B \setminus A) = \mu(B) - \mu(A) = \nu(B) - \nu(A) = \nu(B \setminus A) \Rightarrow (B \setminus A \in \mathcal{L})$
3. $B_1, B_2, ... \in \mathcal{L}$ and $B_n \uparrow B \Rightarrow C_n := B_n \setminus B_{n-1}, C_0 := \emptyset$
   Note they are disjoint and $\bigcup_n C_n = \bigcup_n B_n$, the $\mathcal{L}$ is closed under set minus $C_n \in \mathcal{L}$
   $$\mu(B) = \mu(\bigcup_n B_n) = \mu(\bigcup_n C_n) = \sum_n \mu(C_n) = \sum_n \nu(C_n) = \nu(\bigcup_n C_n) = \nu(\bigcup_n B_n) = \nu(B) \Rightarrow B \in \mathcal{L}$$

Here $\mathcal{A}_1$ is a π-system s.t. $\mu(A) = \nu(A)$ on $\mathcal{A}_1$ which is subset of $\mathcal{L}$, Then by Dynkin's Theorem 3.2, $\sigma(\mathcal{A}_1) \subset \mathcal{L}$. Also $\mathcal{L} \subset \sigma(\mathcal{A}_1) \Rightarrow \mathcal{L} = \sigma(\mathcal{A}_1)$, Hence $\sigma(\mathcal{A}_1), \mathcal{A}_2, ..., \mathcal{A}_n$ are independent $\square$

Note: we only need $\mathcal{A}_1$ to be π-system, the rest is for induction step in Corollary5.1

Corollary 5.1. *Under the same conditions of the Lemma we have $\sigma(\mathcal{A}_1), \sigma(\mathcal{A}_2), ..., \sigma(\mathcal{A}_n)$ are independent*

Proof. We apply the lemma n times which shifts the indices, the first 2 steps goes like this:
Apply the lemma once, we have $\sigma(\mathcal{A}_1), \mathcal{A}_2, ..., \mathcal{A}_n$ are independent,
then consider the set $\sigma(\mathcal{A}_1), \mathcal{A}_2, ..., \mathcal{A}_n$ instead of $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$,
now put A_2 at the front, now $\sigma(\mathcal{A}_i) \subset \mathcal{F}$ and $\sigma(\mathcal{A}_i)$ is a π-system,
now we have $\sigma(\mathcal{A}_1), \sigma(\mathcal{A}_2), \mathcal{A}_3, ..., \mathcal{A}_n$ are independent, continue the process we will have the lemma □

Corollary 5.2. *The RV's $X_1, ..., X_n$ are independent iff*

$$\mathbb{P}(\bigcap_{i=1}^{n}\{X_i \leq x_i\}) = \prod_{i=1}^{n}\mathbb{P}(X_i \leq x_i), \quad \forall x_i \in \bar{\mathbb{R}}$$

Proof. The set $\mathcal{A}_i = \{\{\omega | X_i(\omega) \leq x\} | x \in \bar{\mathbb{R}}\}$ are π-system, $\mathcal{A}_i \subset \mathcal{F}, \Omega \in \mathcal{A}_i$ and $\sigma(\mathcal{A}_i) = \sigma(X_i) \Rightarrow X_1, ..., X_n$ are independent by exercise (Since $\sigma(\mathcal{A}_i), ..., \sigma(\mathcal{A}_n)$ are independent by Corollary 5.1 □

We may extend the definition of independence to infinite sequence of events, RV's on σ-algebra

Def . Sequence of events $A_1, A_2, ...$ is independent if for every $n \in \mathbb{N}$ events $A_1, ..., A_n$ are independent.

Analogous definition holds for RVs and σ-algebras.

Example 5.2. (Infinite sequence of independent events) Let $(\Omega, \mathcal{F}, \mathbb{P}) = ((0,1], \mathcal{B}((0,1]), \lambda)$ which each $\omega \in (0,1]$ associated its (non-terminating) dyadic expansion

$$\omega = \sum_{n=1}^{\infty} \frac{d_n(\omega)}{2^n} = 0.d_1(\omega)d_2(\omega)...$$

Thus each $d_n(\omega)$ is 0 or 1, and $(d_1(\omega), d_2(\omega), ...)$ is a sequence of binary digits in the expansion of ω
Let $A_n = \{\omega \in (0,1] | d_n(\omega) = 0\}, n = 1, 2, ...$
Then $\mathbb{P}(A_n) = \frac{1}{2}$, Obviously, they are all independent for any number n, explicitly:

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_n) = \frac{1}{2^n} = \mathbb{P}(A_1)\mathbb{P}(A_2)...\mathbb{P}(A_n)$$

Lemma 5.2 (The second Borel-Cantelli lemma)**.** *recall The first Borel-Cantelli lemma 4.1:*
Let (A_n) be a sequence of events s.t.

$$\sum_n \mathbb{P}(A_n) < \infty \Rightarrow \mathbb{P}(\limsup_n A_n) = 0$$

Let A_n be a sequence of independent events s.t.

$$\sum_n \mathbb{P}(A_n) = \infty \Rightarrow \mathbb{P}(\limsup_n A_n) = 1$$

Proof. Let $p_n := \mathbb{P}(A_n)$ then $\mathbb{P}(\bigcap_{n=m}^{N} A_n^c) \xrightarrow{independency} \prod_{n=m}^{N}(1-p_n) \leq \prod_{n=m}^{N} e^{-p_n} = e^{-\sum_{n=m}^{N} p_n} \xrightarrow{N\to\infty} 0$
Note: $1 - x \leq e^{-x}, \forall x \geq 0$
$\mathbb{P}(\bigcap_{n=m}^{N} A_n^c) \xrightarrow{N\to\infty} 0$ Hence,

$$\begin{aligned}
\mathbb{P}((\limsup_n A_n)^c) &= \mathbb{P}((\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n)^c) \\
\text{By De Morgan's law } &= \mathbb{P}((\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n^c)) \leq \sum_{m=1}^{\infty} \mathbb{P}(\bigcap_{n=m}^{\infty} A_n^c) \to 0 \text{ as } \mathbb{P}(\bigcap_{n=m}^{N} A_n^c) \xrightarrow{N\to\infty} 0 \\
&\Rightarrow \mathbb{P}(\limsup_n A_n) = 1 - \mathbb{P}((\limsup_n A_n)^c) = 1
\end{aligned}$$

□

Example 5.3. (Monkey typing Shakespeare) $\{X_n\}$ is i.i.d sequence of RVs with uniform distribution on $\{1, ..., k\}$ where k is the number of keys on the keyboard.
Let $x_1, x_2, ..., x_N$ be the sequence of keys that produces the complete word of W.Shakespeare.
Let $A_n = \bigcap_{i=1}^{N} \{X_{n+i} = x_i\}$, the complete work of W.S. starting with (n+1)-st key strokes
$\mathbb{P}(A_n) = \frac{1}{k^N} \geq 0$
Note A_n's are not independent since they have common keys, so we could make $A_N, A_{2N}, A_{3N}, ...$ and they are indeed independent, since they depend on independent RV's.

$$\sum_{j=1}^{\infty} \mathbb{P}(A_{jN}) = \infty \xRightarrow{BCL2} \mathbb{P}(\limsup_j A_{jN}) = 1 \Rightarrow \mathbb{P}(\limsup_n A_n) = 1$$

If A_n are independent, then for any sequence of A_n, $\mathbb{P}(\limsup_n A_n) \in \{0, 1\}$

Proof. If they converge, we get 0 without using independency(BCL1), if they diverge, we get one with independency (BCL2)

Def . Let $X_1, X_2, ...$ be RVs and define $\tau_n := \sigma(X_{n+1}, X_{n+2}, ...)$ and
1tail σ-algebra of the sequence $\tau = \bigcap_{n=1}^{\infty} \tau_n$

Exercise 5.4. Show that $\tau = \bigcap_{n=m}^{\infty} \tau_n$ for any $m \in \mathbb{N}$

Exercise 5.5. The following events are in τ:

1. $\{\omega | \lim_n X_n(\omega) \text{ exists}\}$
2. $\{\omega | \sum_n^{\infty} X_n(\omega) \text{ converges }\}$
3. $\{\omega | \lim_n \frac{S_n(\omega)}{n} \text{ exists}\}$ where $S_n := \sum_{k=1}^{n} X_k$

The following events are not in τ:

1. $\{S_n > 0\}$, cannot ignore the first n elements
2. $\{\limsup_n S_n > 0\}$
   Similarly, the RV $\limsup_n \frac{S_n}{n}$ is τ-measurable but $\limsup_n S_n$ is not

Theorem 5.1 (Kolmogorov's 0-1 law). *If $X_1, X_2, ...$ are independent RV's then for any $A \in \tau$,*

$$\mathbb{P}(A) \in \{0,1\}$$

Proof. Let $\chi_n := \sigma(X_1, X_2, ..., X_n)$, $\tau_n = \sigma(X_{n+1}, X_{n+2}, ...)$

Lemma 5.3. *χ_n, τ_n are independent*

Proof. Let $\mathcal{A}_1 = \{\{X_i \le x_1 | i \le i \le n\}, x_i \in \bar{\mathbb{R}}\}$, $\mathcal{A}_2 = \{\{X_j \le x_j | n+1 \le j \le n+r\}, r \in \mathbb{N}, x_j \in \bar{\mathbb{R}}\}$, We can easily check for $i = 1,2 : \Omega \in \mathcal{A}_i \in \mathcal{F}, \mathcal{A}_i$ is a π-system. and $\mathcal{A}_1, \mathcal{A}_2$ are independent. This implies $\chi_n = \sigma(A_1)$ and $\tau_n = \sigma(\mathcal{A}_2)$ are independent.

Lemma 5.4. *χ_n and τ are independent.*

Proof. It is enough to note that $\tau \subset \tau_n$,since χ_n and τ_n are independent.

Lemma 5.5. *$\chi_\infty := \sigma(X_1, X_2, ...)$ and τ are independent.*

Proof. Let $\mathcal{A}_1 = \bigcup_{n=1}^{\infty} \chi_n$, We have $\Omega \in \mathcal{A}_1 \subset \mathcal{F}$ and since $\chi_n \uparrow, \mathcal{A}_1$ is a π-system. Since each χ_n is independent of $\tau \Rightarrow \mathcal{A}_1$ is independent of τ
Hence by Lemma $\sigma(\mathcal{A}_1) = \chi_\infty$ and τ are independent.
But $\tau \subset \chi_\infty = \sigma(X_1, X_2, ...)$ so,τ is independent of τ. i.e.

$$\forall A \in \tau, \quad \mathbb{P}(A \cap A) = \mathbb{P}(A)\mathbb{P}(A) \Rightarrow \mathbb{P}(A) \in \{0,1\}$$

Corollary 5.3. *The following are direct results from Thm5.1*

1. *If $A_1, A_2, ...$ are independent events then*

$$\mathbb{P}(\limsup_n A_n) \in \{0,1\}$$

2. *If y is τ-measurable then $\exists c \in \bar{R}$ s.t. $\mathbb{P}(y = c) = 1$*
3. $\mathbb{P}(\lim_n S_n \text{ exists}) \in \{0,1\}$
4. $\mathbb{P}(\lim_n \frac{S_n}{n} = \mu) \in \{0,1\} \forall \mu$ *by* 2

Proof.

6 Lebesgue's integral

Let $(\Omega, \mathcal{F}, \mu)$ be a measurable space, let S^+ denote the set of simple and non-negative $f : \Omega \mapsto \mathbb{R}^+$ defined by

$$f(\omega) = \sum_{k=1}^{n} \alpha_i \mathbb{1}_{A_i}(\omega)$$

where $\alpha_i \in \bar{\mathbb{R}}^+$ and $A_i \in \mathcal{F}$ disjoint (if not, we can make it disjoint)

Exercise 6.1 (f is measurable $(\Omega, \mathcal{F}) \mapsto (\bar{R}, \bar{B})$).

Def . The Lebesgue Integral

1. if $f \in S^+$ then

$$\int f(\omega) d\mu(\omega) = \int_\Omega f d\mu := \sum_{i=1}^{n} \alpha_i \mu(A_i)$$

2. if $f : \Omega \mapsto (\bar{R}^+, \bar{\mathcal{B}})$ is measurable and non-negative then the integral of f w.r.t μ is

$$\int_\Omega f d\mu := \sup_{h \leq f, h \in S^+} \int h d\mu$$

Remark: Since we allow the value of ∞, we define $0 \cdot \infty := 0$

Exercise 6.2. Check that $\int_\Omega f d\mu$ is well defined.
(if we partition Ω in different ways, will we get same value back ?)

Exercise 6.3. Check that for $f \in S^+$ both definition (1,2) agree

Theorem 6.1 (Properties of Lebesgue's integral)**.** *Let $f, g : X \mapsto (\bar{R}^+, \bar{\mathcal{B}})$ be measurable, non-negative functions*

1. *if $f \leq g$ then $\int f d\mu \leq \int g d\mu$*
2. *$c \geq 0 \int cf d\mu = c \int f d\mu$*
3. *if $A, B \in \mathcal{F}$ and $A \subset B$ then*

$$\int_A f d\mu \leq \int_B f d\mu \text{ where } \int_A f d\mu := \int_\Omega f \mathbb{1}_A d\mu$$

4. *$A \in \mathcal{F}$, then $\int_A f d\mu = 0$ iff $f = 0$ a.e on A, i.e. $\mu\{\{f > 0\} \cap A\} = 0$*
5. *if $f, g \geq 0$ then $\int_\Omega (f+g) d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu$*

Proof.

Lemma 6.1 (Approximating function by simple functions)**.** *Let f be measurable non-negative function then $\exists (g_n) \subset S^+$ s.t. $g_n(\omega) \uparrow f(\omega)$ as $n \to \infty$ for any $\omega \in \Omega$*

Proof. Construction:

$$g_n := \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbb{1}_{\{\frac{k-1}{2^n} \leq f \leq \frac{k}{2^n}\}} + n\mathbb{1}_{\{f \geq n\}}$$

Then $g_n \in S^+$, fix $\omega \in \Omega$:

1. if f is bounded, $f(\omega) < n$ then $0 \leq f(\omega) - g_n(\omega) \leq \frac{1}{2^n}$
2. if f is unbounded, $f(\omega) = \infty$(unbounded), then $g_n(\omega) \to \infty$

In both case $g(\omega) \xrightarrow{n\to\infty} f(\omega)$ for any $\omega \in \Omega$

Theorem 6.2 (Monotone Convergence Theorem(MCT))**.** *let f_n be non-decreasing of non-negative measurable functions, and $f_n(\omega) \uparrow f(\omega)$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu$$

Proof. From property (1) of integrals (monotonicity) of integrals:

$$\int_\Omega f_n d\mu \leq \int_\Omega f d\mu \quad \left(\int_\Omega f_n d\mu\right) \text{ is non decreasing}$$

Note: The limits of bounded monotonic sequence exist
It is enough to show

$$\lim_{n\to\infty} \int_\Omega f_n d\mu \leq \int_\Omega f d\mu$$

Let $g \in S^+, g \leq f$. Fix $t \in (0,1)$ Then $A_n := \{\omega | f_n(\omega) \geq tg(\omega)$ is an increasing measurable sets and $\bigcup\limits_n A_n = \Omega$

$$\int_\Omega f_n(\omega) \geq \int_{A_n} f_n(\omega) d\mu \geq t \int_{A_n} g(\omega) d\mu$$

Represent $g = \sum\limits_{k=1}^{N} \alpha_k \mathbb{1}_{E_k}$ then

$$\int_{A_n} g(\omega) d\mu = \int_\Omega \sum_{k=1}^{N} \alpha_k \mathbb{1}_{E_k \cap A_n}(\omega) d\mu = \sum_{k=1}^{N} \alpha_k \mu(E_k \cap A_n) = \sum_{k=1}^{N} \alpha_k \mu(E_k) = \int_\Omega g(\omega) d\mu$$

Since A_n is increasing and we have continuity of measure from below we have $\mu(E_k \cap A_n) \to \mu(E_k)$ as $n \to \infty$ for all k
Recall we have

$$\int_\Omega f_n(\omega) \geq \int_{A_n} f_n(\omega) d\mu \geq t \int_{A_n} g(\omega) d\mu \to t \int_\Omega g(\omega) d\mu \quad \forall t \in (0,1)$$

This is true for all $t \in (0,1)$ Take $t \to 1$

$$\int_\Omega f_n(\omega) \geq \int_\Omega g(\omega) d\mu \quad \forall g \in S^+, g \leq f$$

$$\Rightarrow \int_\Omega f_n(\omega) \geq \sup_{g \leq f, g \in S^+} \int_\Omega g d\mu = \int_\Omega f d\mu$$

□

Example 6.1 (Things might go wrong if f_n is not "non-decreasing"). let $(\Omega, \mathcal{F}, \mathbb{P}) = ((0,1), \mathcal{B}, \lambda)$ and $f_n := n\mathbb{1}_{0,\frac{1}{n}}$.

$$\int_\Omega f_n d\mu = n \cdot \lambda((0, \frac{1}{n})) = 1 \quad n \in \mathbb{N}$$
$$f_n \xrightarrow{n\to\infty} 0 \quad \forall \omega \in (0,1)$$
$$\lim_{n\to\infty} \int_\Omega f_n d\mu = 1 \neq \int_\Omega 0 d\mu = 0$$

For example, for $\omega = \frac{3}{4}, f_1(\omega) = 1, f_2(\omega) = 0$ which is decreasing, does not satisfy MCT conditions.

But this satisfies condition of Fatou's lemma 6.3.
$f_n \to 0, \liminf_n f_n = 0, \int_\Omega f_n d\mu = 1 \to \liminf_n \int_\Omega f_n d\mu = 1$

$$\int_\Omega \liminf_n f_n = 0 d\mu = 0 \leq \liminf_n \int_\Omega f_n d\mu = 1$$

Exercise 6.4. If $f, g \geq 0$ then $\int_\Omega (f+g) d\mu = \int_\Omega f d\mu + \int_\Omega g d\mu$

Theorem 6.3 (Fatou's Lemma)**.** *let f_n be sequence of measurable non negative functions, then*

$$\int_\Omega \liminf_n f_n d\mu \leq \liminf_n \int_\Omega f_n d\mu$$

Proof. let $g_n := \inf_{k\geq n} f_k$ This is non decreasing, and measurable sequence.
$g_n \leq f_n$ and $g_n \uparrow \liminf_n f_n$.

$$\begin{aligned}
&\lim_{n\to\infty} \int_\Omega g_n d\mu = \int_\Omega \liminf_n f_n d\mu \\
&\lim_{n\to\infty} \int_\Omega g_n d\mu = \liminf_n \int_\Omega g_n d\mu \leq \liminf_n \int_\Omega f_n d\mu \text{ by MCT and monotonicity} \\
\Rightarrow &\int_\Omega \liminf_n f_n d\mu \leq \liminf_n \int_\Omega f_n d\mu
\end{aligned}$$

□

Exercise 6.5 (Reverse Fatou's Lemma)**.** let f_n be sequence of measurable non negative functions s.t, $f_n \leq g$ for some measurable g with $\int g d\mu < \infty$ then

$$\int_\Omega \limsup_n f_n d\mu \geq \limsup_n \int_\Omega f_n d\mu$$

Def . For $f : \Omega \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ measurable, we define:

1. $f^+ := \max(f, 0) \geq 0$
2. $f^- := \max(-f, 0) = -\min(f, 0) \geq 0$

observe: $f = f^+ - f^-$ and $f = f^+ + f^-$, f is measurable iff f^+, f^- are both measurable.

Def (Lebesgue integral for measurable functions)**.** If $f : \Omega \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ measurable, we define:

$$\int_\Omega f d\mu := \int_\Omega f^+ d\mu - \int_\Omega f^- d\mu$$

provided that at least one of the integrals on the RHS is finite, otherwise $(\infty - \infty)$ is problematic

Now we have all three definitions of Lebesgue integral.

Def (Integrable)**.** we say f is integrable if

$$\text{both } \int_\Omega f^+ d\mu, \int_\Omega f^- d\mu \text{ are finite } \iff \int_\Omega |f| d\mu < \infty$$

Theorem 6.4 (Properties)**.** *Let f, g are integrable*

1. *$f + g$ is integrable and $\int f + g d\mu = \int f d\mu + \int g d\mu$*
2. *If $c \in \mathbb{R}$ then $\int c f d\mu = c \int f d\mu$*
   (Combing 1 and 2 we have $\int \alpha f + \beta g d\mu = \alpha \int f d\mu + \beta \int g d\mu \quad \forall \alpha, \beta \in \mathbb{R}$)
3. *if $f \leq g$ then, $\int f d\mu \leq \int g d\mu$*
4. *$|\int_\Omega f d\mu| \leq \int_\Omega |f| d\mu$*
5. *If $A_n \in \mathcal{F}$ disjoint, $\bigcup_n A_n = \Omega$ then*

$$\int_\Omega f d\mu = \sum_n \int_{A_n} f d\mu$$

6. *If $\int_A f d\mu = 0 \quad \forall A \in \mathcal{F}$ iff $f = 0$ a.e.*

Corollary 6.1. *If f and g are measurable and $f = g$ a.e. then f is integrable iff g is integrable and*

$$\int f d\mu = \int g d\mu$$

Corollary 6.2 (Fatou's Lemma in more general sense)**.** *let f_n be sequence of measurable functions with $f_n \geq 0$ a.e., then*

$$\int_\Omega \liminf_n f_n d\mu \leq \liminf_n \int_\Omega f_n d\mu$$

Corollary 6.3 (MCT in more general sense)**.** *let f_n be non-decreasing of non-negative measurable functions, and $f_n(\omega) \uparrow f(\omega)$ a.e Then*

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu$$

Proof. Let $g_n = \max(f_n, 0)$, then $g_n \in \mathcal{F}^+$ (max of measurable function is measurable and it is also non negative). Let $N = \{\omega | g_n \neq f_n\}$ observe that:

$$N = \{\omega | g_n \neq f_n\} = \{\omega | \max(f_n, 0) \neq f_n\} = \{\omega | f_n < 0\}$$

$\mu(\{f_n < 0\}) = 0$ since $f_n \geq 0$ a.e., $\mu(N) = 0$. Hence $g_n = f_n$ a.e. and hence

$$\liminf_n g_n = \liminf_n f_n \quad \text{a.e.}$$

By Corollary 6.1 and Fatou's Lemma (Assume we have integrability):

$$\int \liminf_n f_n d\mu = \int \liminf_n g_n d\mu \leq \liminf_n \int g_n d\mu = \liminf_n \int f_n d\mu$$

Riemann Integral

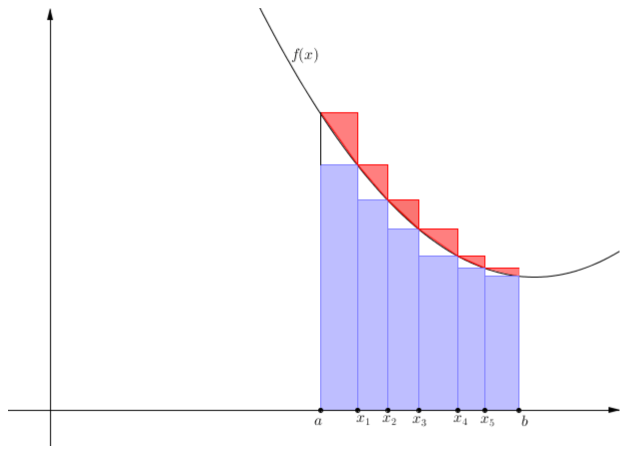

Figure 1: Upper and Lower Riemann sums

Consider function $f : [a, b] \mapsto \mathbb{R}$
We partition $[a, b], \mathcal{P} := \{a_i\}_{i=1}^{k}$ s.t. $a = a_0 < a_1 < ... < a_k = b$
Define:

- $m_i := \inf\{f(x)|x \in [a_{i-1}, a_i]\}$
- $M_i := \sup\{f(x)|x \in [a_{i-1}, a_i]\}$

- Lower sum corresponding to partition $\mathcal{P}$ $\quad L(f, \mathcal{P}) := \sum_{i=1}^{k} m_i(a_i - a_{i-1})$
- Upper sum corresponding to partition $\mathcal{P}$ $\quad U(f, \mathcal{P}) := \sum_{i=1}^{k} M_i(a_i - a_{i-1})$

- Lower Riemann integral $L(f) := \sup\{L(f, \mathcal{P}) : \mathcal{P} \in \tilde{\mathcal{P}}\}$
- Upper Riemann integral $U(f) := \inf\{U(f, \mathcal{P}) : \mathcal{P} \in \tilde{\mathcal{P}}\}$
  $\tilde{\mathcal{P}}$ be the set of all possible partitions of $[a, b]$

Def (Riemann Integrable)**.** $f : [a, b] \mapsto \mathbb{R}$ is Riemann integrable if $L(f) = U(f)$. In that case we define the integral of f by

$$\int_a^b f(x)dx = I = U(f) = L(f)$$

Connections between Riemann and Lebesgue integral

Suppose we have a Riemann integrable function f, WTS: This function is also Lebesgue integrable, first it has to be measurable
Suppose $f : [a,b] \mapsto \mathbb{R}^+$ is Riemann integrable with $\int_a^b f(x)dx = I$ Then there exists sequence $l_n, u_n \in S^+$ s.t.

1. $l_n \uparrow l \leq f$ and $u_n \downarrow u \geq f$
2. For $u_n, l_n \in S^+$ the Lebesgue integral is defined equal to Riemann integrals $\int l_n d\lambda = \int_a^b l_n dx$.

$$L(f, \mathcal{P}_n) = \int_a^b l_n dx \uparrow L \Rightarrow \int l_n d\lambda \uparrow L$$

   Similarly, $\int u_n d\lambda \downarrow U$

Hence l and u are both measurable (limit of measurable functions are measurable) with $l \leq u$ and

$$\int l_n d\lambda \leq \int l d\lambda \leq \int u d\lambda \leq \int u_n d\lambda$$

And by MCT, we have

$$\int l_n d\lambda \to \int l d\lambda, \int u_n d\lambda \to \int u d\lambda$$
$$\int l_n d\lambda \to I, \int u_n d\lambda \to I$$

We have, by additivity:

$$\int l d\lambda + \int (u - l) d\lambda = \int u d\lambda$$
$$\int (u - l) d\lambda = 0$$

We also know that $u - l$ is positive and measurable (difference of limits of measurable function), by property 4 we have $u = l$ a.e. (*)

$$\begin{aligned} & l_n \leq f \leq u_n \text{ for any n} \\ & l_n \uparrow l, u_n \downarrow u \\ \Rightarrow & l \leq f \leq u \\ & u = f = l \text{ on } \{u = l\} (a.e.) \qquad (*) \end{aligned}$$

Denote $\tilde{f} = f \mathbb{1}_{\{u=l\}} = u I_{\{u=l\}}$ which is measurable (product of two measurable functions) and $f = \tilde{f} + f \mathbb{1}_{\{u \neq l\}}$

Exercise 6.6. Show that with $\mathcal{L}$ =Lebesgue (completion of Borel σ-algebra w.r.t λ)

$$f([a,b], \mathcal{L}, \lambda) \mapsto (\mathbb{R}, \mathcal{B})$$

is measurable.

Corollary 6.4. *f is Lebesgue $([a,b], \mathcal{L}, \lambda)$ integrable and $\int f d\lambda = I$*

Comments: This allows us to compute Lebesgue integral $\int f d\lambda$ using The Fundamental Theorem of Calculus for Riemann integrable function.

Example 6.2. $(\Omega, \mathcal{F}, \mu) = ([0,1], \mathcal{B}, \lambda)$ Let $f_n(x) = (n+1)x^n$ then

$$\int f_n d\lambda = \int_0^1 f_n(x)dx = x^{n+1}\Big|_0^1 = 1$$

7 Convergence

7.1 L^p spaces (Lebesgue spaces)

The spaces of all integrable functions is defined by L^1 or $L_1(\Omega, \mathcal{F}, \mu)$, i.e, $f \in L^1$ iff f is measurable and $\int |f| d\mu < \infty$
L^1 is a vector space as

- $|\alpha f + \beta g| \leq |\alpha||f| + |\beta||g|$
- Integral is a linear function on L^1 by Theorem 6.4 (1, 2)
  - $f + g$ is integrable and $\int f + g d\mu = \int f d\mu + \int g d\mu$
  - If $c \in \mathbb{R}$ then $\int c f d\mu = c \int f d\mu$

Def . For $p \geq 1$, L^p or $L^p(\Omega, \mathcal{F}, \mu)$ is the space of all measurable functions f s.t.

$$\int |f|^p d\mu < \infty \qquad \text{(i.e. } |f|^p \in L^1\text{)}$$

The space L^p is equipped with p-norm given by $\|f\|_p = \left(\int_\Omega |f|^p \, d\mu\right)^{\frac{1}{p}}$
$\|\cdot\|_p$ satisfies:

1. $\|f\|_p = 0 \iff f = 0$ a.e.
2. $\|cf\|_p = |c| \|f\|_p, \quad c \in \mathbb{R}$
3. $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ (Minkowski's inequality)
   Proof of Minkowski's inequality is at section inequality
4. $\|f + g\|_p^p \leq 2^{p-1}(\|f\|_p^p + \|g\|_p^p)$

Proof. 4
Notice $h(t) = t^p \quad (p \geq 1)$ is convex for $t \geq 0$ For $x, y \geq 0$ we have

$$h(\frac{1}{2}x + \frac{1}{2}y) \leq \frac{1}{2}(h(x) + h(y))$$
$$\frac{(x+y)^p}{2^p} \leq \frac{1}{2}(x^p + y^p)$$

WLOG, we assume $f, g \geq 0$ This will make it larger on the LHS but RHS stays the same

$$\frac{(|f(\omega) + g(\omega)|)^p}{2^p} \leq \frac{1}{2}(|f(\omega)|^p + |g(\omega)|^p) \qquad \forall \omega \in \Omega$$

Now integrate over Ω w.r.t μ

$$\|f + g\|_p^p \leq 2^{p-1}(\|f\|_p^p + \|g\|_p^p)$$

□

$(V, \|\cdot\|)$ is normed vector space over a field $(\mathbb{R})$ if

1. V is a vector space over field
2. $\|\cdot\|$ is a norm on V over the field
   (a) $\|v\| \geq 0$ and $\|v\| \iff V = 0_v \forall v \in V$
   (b) $\|\alpha v\| = |\alpha| \|v\|, \quad \forall \alpha \in \mathbb{R}, v \in V$
   (c) $\|u + v\| \leq \|u\| + \|v\|, \quad \forall u, v \in V$

We denote by $L^p(\Omega, \mathcal{F}, \mu)$ the space:

$$L^p = \{f : \Omega \mapsto \mathbb{R} | \text{measurable and } \|f\|_p = \left(\int_\Omega |f|^p \, d\mu \right)^{\frac{1}{p}} < \infty\}$$

quotient the space by equivalence relation $\sim$, that is:

$$f, g \in L^p, f \sim g \iff f - g = 0 \text{ a.e. } \mu$$

Because without this, $\|f\|_p = 0 \;\not\Longrightarrow\; f = 0$, say for Lebesgue measure on $[0, 1]$, 1 for rational and 0 for irrational.

Theorem 7.1 (Banach space)**.** $L^p, p \geq 1$ *is complete normed vector space (Banach space) i.e. every Cauchy sequence* $\{f_n\} \subset L^p$ *has a limit in* L^p

Cauchy sequence:$\forall \epsilon > 0, \exists N > 0, \forall n, m \geq N, \|f_n - f_m\| \leq \epsilon$
L^2 is even more special, since the $\|\cdot\|_2$ is generated by the inner product:

$$f, g \in L^2 : \langle f, g \rangle := \int f g d\mu$$

is well define by Holder's inequality 7.7 $\int |fg| \, d\mu \leq \|f\|_2 \|g\|_2 < \infty$
L^2 is **Hilbert space**, i.e. it is a complete normed space with the norm generated by the inner product

Theorem 7.2 (Dominated Convergence Theorem)**.** *Let f_n be sequence of measurable functions, if $\exists$ integrable g s.t. $|f_n| \leq g$ a.e. $\forall n \in \mathbb{N}$, then each f_n is integrable. If $f_n \to f$ a.e. and f is measurable, then f is integrable*

$$\lim_n \int_\Omega |f_n - f| \, d\mu = 0 \quad \textit{and} \quad \lim_n \int_\Omega f_n d\mu = \int_\Omega f d\mu$$

Proof. Since $|f_n| \leq g$ a.e. Define $A_n := \{\omega | |f_n| \geq g\}, B_n := \{\omega | \lim_n f_n \neq f\}$
By definition we have $\mu(A_n) = \mu(B_n) = 0 \quad \forall n \in \mathbb{N}$ also, $\mu(\bigcup_{n\geq 1}(A_n \cup B_n)) = 0$
Hence, on $(\bigcup_{n\geq 1}(A_n \cup B_n))^c$ We have desired property that $|f| \leq g$ a.e.
$\int |f| \, d\mu \leq \int g d\mu < \infty$, f is integrable and measurable, also we have $|f_n - f| \leq |f_n| + |f| \leq 2g$ a.e.
Apply Fatou's Lemma we have:

$$\begin{aligned}
\int \liminf_n (2g - |f_n - f|) d\mu &\leq \liminf_n \int 2g - |f_n - f| \, d\mu \\
&= \int 2g + \liminf_n \int -|f_n - f| \, d\mu \\
&= \int 2g - \limsup_n \int |f_n - f| \, d\mu
\end{aligned}$$

$$\begin{gathered}
LHS = \int \liminf_n (2g - |f_n - f|) d\mu = \int 2g \\
\limsup_n \int |f_n - f| \, d\mu \leq \int 2g - \int 2g = 0 \\
0 \leq \liminf_n \int |f_n - f| \, d\mu \leq \limsup_n \int |f_n - f| \, d\mu \leq 0 \\
\liminf_n = \limsup_n, \text{ Limit exists and } \lim_n \int |f_n - f|) d\mu = 0 \\
0 \leq \left| \int (f_n - f)) d\mu \right| \leq \int |f_n - f|) d\mu \xrightarrow{n\to\infty} \\
\lim_{n\to\infty} \int f_n d\mu = \int f d\mu
\end{gathered}$$

7.2 Modes of convergence

Def . Sequence of measurable functions f_n converges to a measurable function f:

1. Almost everywhere
$$\mu(\{\omega| \lim_n f_n(\omega) = f(\omega)\}^c) = 0$$
This is denoted by $f_n \xrightarrow{\text{a.e.}} f$
2. In measure
$$\forall \epsilon > 0 \quad \mu(\{\omega| |f_n(\omega) - f(\omega)| > \epsilon\}) \xrightarrow{n\to\infty} 0$$
This is denoted by $f_n \xrightarrow{\mu} f$
3. In L^p $\quad (1 \leq p \leq \infty)$ if $f_n \in L^p$
$$\int |f_n - f|^p \, d\mu \xrightarrow{n\to\infty} 0$$
This is denoted by $f_n \xrightarrow{L^p} f$

Lemma 7.1. *if $f_n \xrightarrow{a.e.} f$ and $\mu(\Omega) < \infty$ then $f_n \xrightarrow{\mu} f$*

Lemma 7.2. *Let f_n be measurable functions and $f_n \xrightarrow{L^1} f$ Then $f_n \xrightarrow{\mu} f$*

Lemma 7.3 (Scheffe's lemma)**.** *Suppose $f_n, f \in L^1, f_n \xrightarrow{a.e.} f$ Then*
$$f_n \xrightarrow{L^1} f \iff \int |f_n| \, d\mu \to \int |f| \, d\mu$$

Proof. of Lemma 7.2 Fix $\epsilon > 0, |f_n - f| \in \mathcal{F}^+$, Then by Markov inequality 7.5
$$\mu(|f_n - f| \geq \epsilon) \leq \frac{1}{\epsilon} \int |f_n - f| \, d\mu \xrightarrow{n\to\infty} 0$$

Example 7.1. Difference between modes of convergence.(Counterexample)

1. $f_n \xrightarrow{\text{a.e.}} f \;\not\Longrightarrow\; f_n \xrightarrow{L^1} f$
   Let $f_n := n\mathbb{1}_{(0,\frac{1}{n})}$ on $((0,1),\mathcal{B},\lambda)$.

$$f_n(\omega) \to 0 \quad \forall \omega \in \Omega$$
$$\int |f_n - f|\, d\mu = \int f_n d\mu = n\lambda((1,\frac{1}{n})) = 1 \neq 0$$

   Note: we have $f_n \xrightarrow{\mu} f$
   Let $\epsilon > 0 \quad \lambda(\left|f_n - 0\right| > \epsilon) = \lambda(n\mathbb{1}_{(0,\frac{1}{n})} > \epsilon) \leq \frac{1}{n} \xrightarrow{n\to\infty} 0$

2. $f_n \xrightarrow{L^1} f + f_n \xrightarrow{\mu} f \;\not\Longrightarrow\; f_n \xrightarrow{\text{a.e.}} f$
   Define $f : f_1 = \mathbb{1}_{(0,\frac{1}{2})}, f_2 = \mathbb{1}_{(\frac{1}{2},1)}, f_3 = \mathbb{1}_{(0,\frac{1}{4})}, f_4 = \mathbb{1}_{(\frac{1}{4},\frac{1}{2})} \cdots$

   $\int |f_n - f|\, d\mu = \int f_n d\mu = \lambda(\text{support } f_n) =\xrightarrow{n\to\infty} 0$
   $\lambda(\left|f_n\right| > \epsilon) \leq \lambda(\text{support } f_n) \xrightarrow{n\to\infty} 0$
   $\forall \omega \in \Omega, f_n(\omega) = (0,0,1,...,0,0,1,0,..)$ $\quad$ $\exists\infty$ many n's the sequence does not converge for all omega

3. $f_n \xrightarrow{\text{a.e.}} f \;\not\Longrightarrow\; f_n \xrightarrow{\mu} f$
   Let $f_n := \mathbb{1}_{[n,\infty)}$ on $([0,\infty),\mathcal{B},\lambda)$.

$$\forall \omega \in [0,\infty), f_n \xrightarrow{\text{a.e.}} 0$$
$$\lambda(\left| f_n(\omega) \geq \frac{1}{2} \right|) = \lambda([n,\infty)) = \infty \neq 0$$

Theorem 7.3. *Suppose that $f_n \xrightarrow{\mu} f$, Then $\exists$ a subsequences $\{n_k\}$ s.t. $f_{n_k} \xrightarrow{a.e.} f$*

Remark. Since $f_n \xrightarrow{L^1} f \implies f_n \xrightarrow{\mu} f$, this claim also holds if we have L^1 convergence.

Proof. For any $k \geq 1 \quad \exists n_k \geq 1$ s.t.

$$\mu(|f_n - f| > \frac{1}{2^k}) < \frac{1}{2^k} \qquad \forall n \geq n_k$$

WLOG, we could assume $n_k \uparrow$
Let $A_k := \{\omega | |f_{n_k}(\omega) - f(\omega)| > \frac{1}{2^k}\} \qquad \sum_k \mu(A_k) \leq \sum_k \frac{1}{2^k} \leq \infty$
by Borel-cantelli's Lemma 4.1 we have $\mu(\limsup_k A_k) = 0$
That is, A_k occurs infinitely often have zero measure, that is for a.e. $\omega \in \Omega$, omega is inside of finitely many of $A_k's$
For a.e $\omega \in \Omega \quad \exists K(\omega) \geq 1$ s.t $\omega \notin A_k \forall k \geq K(\omega)$ we have $|f_{n_k}(\omega) - f(\omega)| \leq \frac{1}{2^k} \xrightarrow{k \to \infty} 0$ □

Theorem 7.4. *If $f_n \xrightarrow{\mu} f$ and $|f_n| \leq g$ a.e with $g \in L^1$ then $f_n \xrightarrow{L^1} f$ (Stronger than DCT which requires a.e convergence for $\mu(\Omega) < \infty$)*

Proof. Suppose $f_n \nrightarrow f$ in L^1 Then $\exists \epsilon > 0$ and $n_k \uparrow \infty$ s.t

$$\int |f_{n_k} - f| \, d\mu > \epsilon, \qquad \forall k \qquad (*)$$

But $f_{n_k} \xrightarrow{\mu} f$ By previous Thm, $\exists m_k \subset n_k$ s.t $f_{m_k} \xrightarrow{\text{a.e.}} f$
By DCT we have $\lim_n \int |f_{m_k} - f| = 0, f_{m_k} \xrightarrow{L^1} f$ which contradicts (*) □

7.3 Inequalities

Theorem 7.5 (Markov inequality). *For $f \in \mathcal{F}^+$ and $\alpha > 0$*

$$\mu\{\omega | f(\omega) \geq \alpha\} \leq \frac{1}{\alpha} \int f d\mu$$

The next inequality will be used in proving Holder's inequality

Theorem 7.6 (Young's inequality). *for $\omega, z > 0$*

$$\omega z \leq \frac{1}{p}\omega^p + \frac{1}{q}z^q$$

Theorem 7.7 (Holder's inequality). *Let f, g be measurable functions, $p, q \in (1, \infty)$ with $\frac{1}{q} + \frac{1}{p} = 1$ Then*

$$\int |fg| \, d\mu \leq \left(\int |f|^p \, d\mu \right)^{\frac{1}{p}} \left(\int |g|^q \, d\mu \right)^{\frac{1}{q}} = \|f\|_p \|g\|_q$$

Theorem 7.8 (Minkowski's inequality(triangle inequality for p-norms)). *For measurable functions f, g and $p \geq 1$*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

proof of Markov inequality:

Proof.

$$f \geq f\mathbb{1}_{\{f\geq\alpha\}} \geq \alpha\mathbb{1}_{\{f\geq\alpha\}} \geq 0$$

Integrate both side also we have monotonicity

$$\int \alpha\mathbb{1}_{\{f\geq\alpha\}} = \alpha\mu\{\omega|f(\omega) \leq \alpha\} \leq \int f d\mu \implies \mu\{\omega|f(\omega) \leq \alpha\} \leq \frac{1}{\alpha}\int f d\mu$$

□

Proof of Minkowski's inequality:

Proof. WLOG $f, g \geq 0$ For $p = 1, |f+g| \leq |f| + |g|$, Apply monotonicity of Lebesgue measure, we have desired relation.

If $p > 1$, let $q \in (1, +\infty)$ s.t. $\frac{1}{q} = 1 - \frac{1}{p} \implies q(p-1) = p$

we will utilise

$$(f+g)^p = f(f+g)^{p-1} + g(f+g)^{p-1} \tag{7.1}$$

Apply Holder's inequality on RHS of above equation

$$\int f(f+g)^{p-1}d\mu \leq \left(\int |f|^p \, d\mu\right)^{\frac{1}{p}} \left(\int |f+g|^{(p-1)q} \, d\mu\right)^{\frac{1}{p}\frac{p}{q}} = \|f\|_p \|f+g\|_p^{\frac{p}{q}}$$
$$\int g(f+g)^{p-1}d\mu \leq \|g\|_p \|f+g\|_p^{\frac{p}{q}}$$

Integrate equation 7.1 we have

$$\begin{aligned}\int (f+g)^p d\mu &\leq \int f(f+g)^{p-1}d\mu + \int g(f+g)^{p-1}d\mu \\ &= \|f\|_p \|f+g\|_p^{\frac{p}{q}} + \|g\|_p \|f+g\|_p^{\frac{p}{q}} \\ &= \|f+g\|_p^{\frac{p}{q}} (\|f\|_p + \|g\|_p)\end{aligned}$$

LHS = $\int (f+g)^p d\mu = \|f+g\|_p^p$ divided by $\|f+g\|_p^{\frac{p}{q}}$ both sides we have

$$\|f+g\|_p^{p(1-\frac{1}{q})} = \|f+g\| \leq \|f\|_p + \|g\|_p$$

□

8 Expectation

Recall that a RV is a measurable function $X : (\Omega, \mathcal{F}) \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}})$

Def (Expectation)**.** The expectation (expected value/ mean) of RV X defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is

$$\mathbb{E}(X) := \int_{\Omega} X d\mathbb{P}$$

Provided the integral is well-defined.
Note: $\mathbb{E}(X)$ is well-defined iff $\min(\mathbb{E}(X^+), \mathbb{E}(X^-)) < \infty$

Recall that for a RV X defined on $(\Omega, \mathcal{F}, \mathbb{P})$, we define μ_X the probability measure induced by X on $(\bar{R}, \bar{\mathcal{B}})$ (push forward of $\mathbb{P}$ under map X)

$$\mu_X(B) := \mathbb{P}(\{\omega | X(\omega) \in B\}) \quad \forall B \in \bar{\mathcal{B}}$$

Def . A RV X on $(\Omega, \mathcal{F}, \mathbb{P})$ is **discrete** if $\exists \{x_k\} \subset \bar{\mathbb{R}}$ be countable, $x_k \neq x_l$ for $k \neq l$ s.t.

$$\mathbb{P}(X = x_k) = \mu_X(\{x_k\}) > 0$$
$$\sum_k \mathbb{P}(X = x_k) = \sum_k \mu_X(\{x_k\}) = 1$$

we denote by $p(x) := \mu_X(x)$ the probability mass function of the RV X. The set $\{x | p(x) > 0\}$ is called the support of the pmf of μ_X

Proposition 8.1. *For a discrete RV X with support* $\{x_k\}$

1. $\mathbb{E}(X)$ *is defined* $\iff \min\{\sum\limits_{k:x_k>0} x_k p(x_k), -\sum\limits_{k:x_k<0} x_k p(x_k)\} < \infty$
2. *If* $\mathbb{E}(X)$ *is defined then*
$$\mathbb{E}(X) = \sum_k x_k p(x_k)$$
3. $X \in L^1 \iff \mathbb{E}(|X|) < \infty \iff \sum\limits_k |x_k| \, p(x_k) < \infty$
4. *For any measurable function,* $h : \bar{\mathbb{R}} \mapsto \bar{\mathbb{R}}$ *the RV* $Y = h(X)$ *is discrete and 1-3 hold with* Y *and* $h(x_k)$

Example 8.1. If $h(x) = \sin(x)$ then $\mathbb{E}(\sin(X)) = \sum\limits_x \sin(x_k) p(x_k)$ if expectation for X is well-defined

Proof. First ignore the zero measure part: Let $A_k := X^{-1}(x_k) = \{\omega | X(\omega) = x_k\}$ and $X' := \sum\limits_k x_k \mathbb{1}_{A_k}$
Then X' is a discrete RV and $X = X'$ a.s. Assume first $X \geq 0$ a.e. $\implies$ $X' \geq 0$ a.e. and we have the integral are equal:

$$\mathbb{E}(X) = \mathbb{E}(X') = \int \sum_k x_k \mathbb{1}_{A_k} d\mathbb{P}$$

Let $X_n := \sum\limits_{k=1}^{n} x_k \mathbb{1}_{A_k}$, $X_n \uparrow X'$ and by MCT:

$$\mathbb{E}(X) = \mathbb{E}(X') = \int X' d\mathbb{P} = \int \lim_n X_n d\mathbb{P} = \lim_n \int X_n d\mathbb{P} = \lim_n \sum_{k=1}^{n} x_k \mathbb{P}(A_k) = \sum_k x_k \mathbb{P}(A_k)$$
$$\mathbb{E}(X) = \sum_k x_k \mathbb{P}(A_k)$$

□

Def . AC and pmf
The probability measure induced by RV X, μ_X is **absolutely continuous (AC)** w.r.t Lebesgue's measure λ, if $\mu_X(\mathbb{R}) = 1$ and if $\exists f : (\mathbb{R}, \mathcal{B}) \mapsto ([0, \infty), \mathcal{B})$ measurable s.t.

$$\forall B \in \mathcal{B}, \mathbb{P}(X \in B) = \mu_X(B) = \int_B f d\lambda = \int_{\mathbb{R}} f \mathbb{1}_B d\lambda$$

f is called the **density** of induced measure μ_X

Theorem 8.1 (Change of variables formula)**.** *Let X be a RV on $(\Omega, \mathcal{F}, \mathbb{P})$ and induced measure μ_X be distribution of X*

1. *suppose that $h : (\bar{\mathbb{R}}, \bar{\mathcal{B}}) \mapsto (\bar{R}^+, \bar{\mathcal{B}}^+)$ is measurable,let $Y := h(X)$ Then Y is non-negative RV with*

$$\mathbb{E}(Y) = \int Y \, d\mathbb{P} = \int h(X) \, d\mathbb{P} = \int_{\bar{\mathbb{R}}} h \, d\mu_X$$

2. *If $h : (\bar{\mathbb{R}}, \bar{\mathcal{B}}) \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ is measurable then for $Y = h(X)$*
   (a) *$\mathbb{E}(Y) = \int Y \, d\mathbb{P}$ is defined iff $\int h \, d\mu_X$ is defined and then*

$$\mathbb{E}(Y) = \int h \, d\mu_X$$

   (b) *$Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ iff $h \in L^1(\bar{\mathbb{R}}, \bar{\mathcal{B}}, \mu_X)$*

   Remark. Part 1 and 2 holds for any RVs (discrete/ continuous/ singular ...)

3. *If X has a density f ,then*

$$\mathbb{E}(Y) = \int_{\mathbb{R}} hf \, d\lambda$$

   (Lebesgue integral) with $\mathbb{E}(Y)$ is defined iff $\int_{\mathbb{R}} hf \, d\lambda$ exists

4. *If X has density f and $h : \mathbb{R} \mapsto \mathbb{R}^+$ is measurable s.t. $g(x) := h(x) \cdot f(x)$ is Riemann integrable on any finite interval, then*

$$\mathbb{E}(Y) = \int_{\mathbb{R}} fh \, d\lambda = \int_{-\infty}^{\infty} h(x) f(x) \, dx$$

5. *Condition on h in 4 and we generalise it to the case $\mathbb{R} \mapsto \mathbb{R}$*
   $\mathbb{E}(Y)$ is well-defined if $\min\{\int_{-\infty}^{\infty} h^+ f(x) \, dx, \int_{-\infty}^{\infty} h^- f(x) \, dx\} < \infty$ and then we have:

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} h(x) f(x) \, dx$$

Proof.

Corollary 8.1. $X \in L^1$ *iff* $\int_{-\infty}^{\infty} |x| f(x)\, dx < \infty$

Corollary 8.2. *If X has a density f s.t. $g(x) = xf(x)$ is Riemann integrable on any finite interval, then $E(X)$ is defined iff*

$$\int_{-\infty}^{0} xf(x)\, dx > -\infty \quad or \quad \int_{0}^{\infty} xf(x)\, dx < \infty$$

and if defined then

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)\, dx$$

Remark. This definition does not means the expectation is finite, only existence can be guaranteed

Proposition 8.2. *Let X_1, X_2 be RVs on $(\Omega, \mathcal{F}, \mathbb{P})$*

1. *If X_n are uniformly bounded ($|X_n| \le c < \infty \quad a.s$). and $X_n \to X$ a.s. then $X_n \to X$ in L^1 Bounded Convergence Theorem*
2. *If $X_i \in L^1$ then $\mathbb{E}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathbb{E}(X_i)$*
3. *If $X \ge 0$ a.s. then $\mathbb{E}(X) \ge 0$. Also if $X_i \ge 0$ a.s. $\mathbb{E}(\sum_{i=1}^{\infty} X_i) = \sum_{i=1}^{\infty} \mathbb{E}(X_i)$, this is always well defined*
4. *If $\mathbb{E}(X) < \infty$ then $\mathbb{P}(X < \infty) = 1$*
5. *If $X_n \ge 0$ a.s. and $\sum_{n=1}^{\infty} \mathbb{E}(X_n) < \infty$ then by 4, $\sum_{n=1}^{\infty} X_n < \infty$ a.s. and in particular, $X_n \xrightarrow{a.s.} 0$*

Proposition 8.3. *Suppose X is a measurable function defined on $(\Omega, \mathcal{F}, \mu)$ with $\mu(\Omega) < \infty$ Then*

1. *if $X \in L^p$ for $p \geq 1$ then $X \in L^r$ for $r \in [1, p]$ (i.e. L^p spaces are nested and it is not necessarily the case for infinite measure)*
2. *If $\mu = \mathbb{P}$, then $\|X\|_r \leq \|X\|_p$, for $p \geq r \geq 1$, that is existence of higher moment always imply existence of lower moment.*

Proof. If $p > 1$ and $r \in [1, p)$, define $p' = \frac{p}{r} > 1$ then $\exists q' > 1 \quad \frac{1}{p'} + \frac{1}{q'} = 1$ Conjugate. and then we have (By Holder's inequality 7.7):

$$\begin{aligned} \|X\|_r^r = \int |X|^r \cdot 1 \, d\mu &\leq \left(\int |X|^{rp'} \, d\mu \right)^{\frac{1}{p'}} \left(\int |1|^{q'} \, d\mu \right)^{\frac{1}{q'}} \\ &= \left(\int |X|^p \, d\mu \right)^{\frac{r}{p}} \mu(\Omega)^{\frac{1}{q'}} \\ &= \|X\|_p^r \, \mu(\Omega)^{\frac{1}{q'}} \\ \implies \|X\|_r &\leq \|X\|_p \, \mu(\Omega)^{\frac{1}{rq'}} \leq \infty \end{aligned}$$

□

convex function and Jensen's inequality

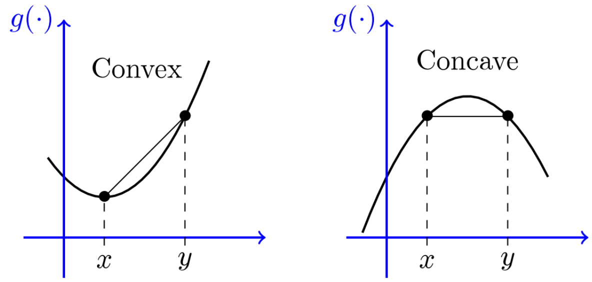

Figure 2: Convex and concave function

Def (Convex function). A function $\phi : I \mapsto \mathbb{R}$ is convex on an open interval I if $\forall x, y \in I$ and parameterise by $t \in (0,1)$:

$$\phi(tx + (1-t)y) \leq t\phi(x) + (1-t)\phi(y), \quad \forall t \in (0,1)$$

Lemma 8.1 (Slope inequality). *If $a < b < c \in I$ then we have (Notation: $S([a,b]) := \frac{\phi(b)-\phi(a)}{b-a}$)*

$$\frac{\phi(b)-\phi(a)}{b-a} \leq \frac{\phi(c)-\phi(a)}{c-a} \leq \frac{\phi(c)-\phi(b)}{c-b}$$
$$S([a,b]) \leq S([a,c]) \leq S([b,c])$$

Consider slope of the triangle formed by a,b,c:

Corollary 8.3. *Given convex function ϕ, For any $x \in I$ left and right derivatives $D_{-\phi}$ and $D_{+\phi}$ exist and satisfy:*

$$D_{-\phi} := \lim_{h \to 0^+} \frac{\phi(x) - \phi(x-h)}{h} \leq \lim_{h \to 0^+} \frac{\phi(x+h) - \phi(x)}{h} =: D_{+\phi}$$

Corollary 8.4. *For any $x_0 \in I$ and $m \in [D_{-\phi}(x_0), D_{+\phi}(x_0)]$, we have:*

$$m(x - x_0) + \phi(x_0) \leq \phi(x)$$

$m(x - x_0) + \phi(x_0)$ is supporting line at x_0 of ϕ

Proof. of Lemma and Corollary on previous page

Theorem 8.2 (Jensen's inequality). *If $\phi : I \mapsto \mathbb{R}$ is convex function on an open interval $I \subset \mathbb{R}$, and if X is a RV in L^1 with $\mathbb{P}(X \in I) = 1$, then $\mathbb{E}(\phi(X))$ is well-defined (possibly equal to ∞)*

$$\mathbb{E}(\phi(X)) \geq \phi(\mathbb{E}(X))$$

Proof. Since $\mathbb{P}(X \in I) = 1$ and $X \in L^1$, then $\mathbb{E}(X)$ exists and belongs to I. Let $x_0 = \mathbb{E}(X) \implies \forall y \in I$ we have

$$\begin{aligned}
&m(y - \mathbb{E}(X)) + \phi(\mathbb{E}(X)) \leq \phi(y) \\
&\implies m(X - \mathbb{E}(X)) + \phi(\mathbb{E}(X)) \leq \phi(X) \quad a.s.(*) \\
&\implies \left|\phi^-(X)\right| \leq K|X| + c \quad \phi(X) \text{ is bounded below since } X \in L^1 \\
&\implies \mathbb{E}(\phi(X)) \quad \text{is well defined} \\
Integrate(*) \\
&\mathbb{E}(m(X - \mathbb{E}(X)) + \phi(\mathbb{E}(X))) \leq \mathbb{E}(\phi(X)) \\
&\implies \mathbb{E}(m(X - \mathbb{E}(X)) + \phi(\mathbb{E}(X))) = \phi(\mathbb{E}(X)) \leq \mathbb{E}(\phi(X))
\end{aligned}$$

□

Remark. Taking $\phi(x) = x^2$ be convex, then we have

$$\mathbb{E}(X^2) \geq \mathbb{E}(X)^2 \iff \|X\|_1 \leq \|X\|_2$$

Exercise 8.1. Using Jensen's inequality to prove $\|X\|_r \leq \|X\|_p$, $r \in [1, p]$

Variance and Covariance

If $X \in L^1$ then $m := \mathbb{E}(X)$ is well-defined and finite. Since we have $(X-m)^2 = X^2 - 2mX + m^2$ then $(X-m)^2 \in L^1 \iff \mathbb{E}(X^2) < \infty \iff X^2 \in L^1 \iff X \in L^2$ Now:

Def (Variance of X).

$$V(X) := \mathbb{E}((X-m)^2) < \infty \iff X \in L^2$$

Theorem 8.3 (Chebyshev's Inequality). *For $X \in L^2$ with $m = \mathbb{E}(X)$*

$$\mathbb{P}(|X-m| \geq \epsilon) \leq \frac{V(X)}{\epsilon^2}$$

Proof. Define $Y = (X-m)^2 \in L^1$, By Markov inequality 7.5

$$\mathbb{P}(Y \geq \epsilon) \leq \frac{\mathbb{E}(Y)}{\epsilon^2} = \frac{V(X)}{\epsilon^2}$$

□

Def . For $X, Y \in L^1$ with $X \cdot Y \in L^1$ we define

$$Cov(X,Y) := \mathbb{E}((X-m_X)(Y-m_Y))$$

$m_X = \mathbb{E}(X), m_Y = \mathbb{E}(Y)$

Notice: $Cov(X,Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X)\mathbb{E}(Y)$

Def (Non-correlated). $\mathbb{E}(X \cdot Y) = \mathbb{E}(X)\mathbb{E}(Y) \iff Cov(X,Y) = 0$

Lemma 8.2. *If $X_1, ..., X_n \in L^2$ then $X_k X_l \in L^1 \quad \forall k, l$ and*

$$V(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n} V(X_k) + \sum_{k \neq l} Cov(X_k, X_l)$$

Proof. If $k \neq l$:

$$\|X_k X_l\|_1 \leq \|X_k\|_2 \|X_l\|_2 < \infty$$

By Holder's inequality 7.7,The rest are just binomial expansion

Generally $X, Y \in L^1 \not\Longrightarrow XY \in L^1$ but this is the true if we add independency:

Lemma 8.3. *If X and Y are independent, $X, Y \in L^1$ then $X \cdot Y \in L^1$ and*

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

Corollary 8.5. *Relation between independency and correlation*

1. $Cov(X,Y) = 0$ *provided X, Y independent*
2. *If $X, Y \in L^2$ and independent*

$$V(X+Y) = V(X) + V(Y)$$

Proof. (a)The claim holds true if $X = \mathbb{1}_A$ and $Y = \mathbb{1}_B$, where $A, B \in \mathcal{F}$ and A, B are independent events.
(b)$X = \sum\limits_k \alpha_k \mathbb{1}_{A_k}, Y = \sum\limits_l \beta_l \mathbb{1}_{B_l}$ where $A_k, B_l \in \mathcal{F}, \alpha_k, \beta_l \geq 0$ note this is finite sum.
(c) $X, Y \geq 0$, Let X_n, Y_b be the setting as we did for Lebesgue integral, then we have two increasing and non negative sequence of RVs, by MCT

$$\mathbb{E}(X \cdot Y) \overset{MCT}{=} \lim_n \mathbb{E}(X_n Y_n) \overset{(b)}{=} \lim_n \mathbb{E}(X_n)\mathbb{E}(Y_n) = \lim_n \mathbb{E}(X_n) \lim_n \mathbb{E}(Y_n) \overset{MCT}{=} \mathbb{E}(X)\mathbb{E}(Y)$$

Also we have

$$\mathbb{E}(X \cdot Y) = \mathbb{E}(X)\mathbb{E}(Y) < \infty \implies X \cdot Y \in L^1$$

(d) Use the fact $X = X^+ - X^-, Y = Y^+ - Y^-$ and they are also independent. and $XY = X^+Y^+ + X^-Y^- - X^+Y^- - X^-Y^+$ and linearity of expectation. $\square$

SLLN

Def (A version of The strong law of large number)**.** Suppose X_i are i.i.d RVs with $\|X_i\|_4 \leq c < \infty$, let $S_n := \sum_{i=1}^{n} X_i$. Then

$$\frac{S_n}{n} \xrightarrow{a.s.} m := \mathbb{E}(X_i)$$

Proof. Step1: We may assume $m = \mathbb{E}(X_1) = 0$, otherwise, introduce

$$X_i' := X_i - m, S_n' := \sum_{i=1}^{n} X_i' \implies \left\|X_i'\right\|_4 \leq c + |m|\,, S_n' = S_n - mn, \mathbb{E}(X_i') = 0$$

. This is WLOG.
step2: Assuming $m = 0$,

$$S_n^4 = (\sum_{i=1}^{n} X_i)^4 = \sum_{i=1}^{n} X_i^4 + \sum_{i<j} \binom{4}{2} X_i^2 X_j^2 + \text{...The rest are all have terms with odd power}$$

Every term is integrable given $\|X_i\|_4 < \infty \implies S_n^4$ is integrable and take expectation.

$$\begin{aligned}
\mathbb{E}(S_n^4) &= n\mathbb{E}(X^4) + 3n(n-1)\mathbb{E}(X^2)E(X^2) \\
&= n\|X\|_4^4 + 3n(n-1)\|X\|_2^4 \\
&\leq nc^4 + 3n(n-1)c^4 \leq 3n^2c^4 \qquad (*) \\
\mathbb{E}((\frac{S_n}{n})^4) &\leq 3c^4/n^2
\end{aligned}$$

Hence from (*):

$$\mathbb{E}(\sum_{n\geq 1} (\frac{S_n}{n})^4) = \sum_{n\geq 1} \mathbb{E}((\frac{S_n}{n})^4) \leq \sum_{n\geq 1} \frac{3n^2c^4}{n^4} < \infty \implies (\frac{S_n}{n})^4 \xrightarrow{a.s.} 0 \implies \frac{S_n}{n} \xrightarrow{a.s.} 0$$

□

9 Product measure

Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be measurable spaces, and let $\Omega := \Omega_1 \times \Omega_2$ be their Cartesian product $(\Omega = \{(\omega_1, \omega_2) | \omega_i \in \Omega_i\})$ Let $\tau := \{A_1 \times A_2 | A_i \in \mathcal{F}_i\}$.The τ is a π-system (intersection of rectangle is again rectangle) and we define:

$$\mathcal{F} := \sigma(\tau)$$

$\mathcal{F}$ is called the product σ-algebra, and denoted $\mathcal{F}_1 \times \mathcal{F}_2$

Remark. $\mathcal{F}_1 \times \mathcal{F}_2$ is Not Cartesian product.

Actually, τ is the Cartesian product.

Exercise 9.1. Show the following results.

1. Verify τ is a π-system
2. Define $\rho_i : \Omega \mapsto \Omega_i$ as $\rho_i(\omega_1, \omega_2) = \omega_i$ The i-th coordinate map, Show that ρ_i is $\mathcal{F}$ measurable.
3. Show that $\mathcal{F} = \sigma(\rho_1, \rho_2)$ or equivalently $\mathcal{F} = \sigma(\{B_1 \times \Omega_2, \Omega_1 \times B_2 | B_i \in \mathcal{F}_i\})$

Remark. Both ways to define $\mathcal{F}$ work for a countable product of measurable spaces.

This is important for stochastic process X indexed by time.

Main issue If $f : (\Omega, \mathcal{F}) \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ measurable, does it imply that

$$\forall \omega_1 \in \Omega_1, f(\omega_1, \cdot) : (\Omega_2, \mathcal{F}_2) \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}})$$
$$\forall \omega_2 \in \Omega_2, f(\cdot, \omega_2) : (\Omega_1, \mathcal{F}_1) \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}})$$

are measurable ? Yes

Theorem 9.1 (The Monotone Class Theorem). *Let $\mathcal{M}$(monotone class) be a class of bounded (not necessarily measurable) functions $\Omega \mapsto \mathbb{R}$ s.t.*

1. *$\mathcal{M}$ is a vector space over $\mathbb{R}$*
2. *$\mathbb{1}_\Omega \in \mathcal{M}$*
3. *If $(f_n) \in \mathcal{M}$ satisfy: $0 \le f_n \le c < \infty$ and $f_n \uparrow$, then we have $\lim_{n\to\infty} f_n \in \mathcal{M}$*

If $\mathcal{P} \subset \Omega$ is a π-system, and $\forall A \in \mathcal{P} : \mathbb{1}_A \in \mathcal{M}$ then $\mathcal{M}$ contains every bounded $\sigma(\mathcal{P})$ measurable function

Proof. Let $\mathcal{L} := \{A \subset \Omega | \mathbb{1}_A \in \mathcal{M}\}$

1. $\mathbb{1}_\Omega \in \mathcal{M} \implies \Omega \in \mathcal{L}$
2. If $A, B \in \mathcal{L}, A \subset B$ we have $\mathbb{1}_A, \mathbb{1}_B \in \mathcal{M}$ also $\mathbb{1}_{B\setminus A} = \mathbb{1}_B - \mathbb{1}_A \in \mathcal{M}$ since $\mathcal{M}$ is a vector space, Hence $\mathbb{1}_{B\setminus A} \in \mathcal{M} \implies B \setminus A \in \mathcal{L}$
3. Let $A_n \in \mathcal{L}$ and $A_n \uparrow A$, Obviously $\mathbb{1}_{A_n} \leq 1 < \infty$ and $\mathbb{1}_{A_n} \uparrow$ then we have

$$\lim_{n\to\infty} \mathbb{1}_{A_n} = \mathbb{1}_{\lim\limits_{n\to\infty} A_n} \in \mathcal{M} \implies A \in \mathcal{L}$$

Hence, $\mathcal{L}$ is a λ-system. Since $\mathcal{P} \subset \mathcal{L}$ By Dynkin's theorem 3.2 $\sigma(\mathcal{P}) \subset \mathcal{L}$
But $\mathcal{M}$ is a vector space $\implies$ any bounded simple functions over $\sigma(\mathcal{P})$ is in $\mathcal{M}$.
For any f which is $\sigma(\mathcal{P})$ measurable and bounded, $\exists M < \infty, 0 \leq f \leq M < \infty$ and non-negative, Define

$$f_n := \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbb{1}_{\{\frac{k-1}{2^n} \leq f \leq \frac{k}{2^n}\}} + n\mathbb{1}_{\{f \geq n\}}$$

$f_n \in \mathcal{M}$ as it is a simple bounded $\sigma(\mathcal{P})$-measurable function, also $0 \leq f_n \leq M$ and $f_n \uparrow f$, so $f \in \mathcal{M}$ by property 3 of $\mathcal{M}$
Finally, any $\sigma(\mathcal{P})$-measurable function with $|f| \leq M < \infty$ can be written as $f = f^+ - f^-$ and $f^+, f^- \in \mathcal{M}$ since they are measurable, bounded, non-negative. Again, $f = f^+ - f^-$ since $\mathcal{M}$ is a vector space.

Remark (Boundness of $f \in \mathcal{M}$)**.** This makes the proof less technical, but this is not necessary.

Proposition 9.1. *Let* $f : (\Omega, \mathcal{F}) \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ *be measurable, then*

$$(*) \begin{cases} \forall \omega_1 \in \Omega_1, f(\omega_1, \cdot) : (\Omega_2, \mathcal{F}_2) \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}}) \\ \forall \omega_2 \in \Omega_2, f(\cdot, \omega_2) : (\Omega_1, \mathcal{F}_1) \mapsto (\bar{\mathbb{R}}, \bar{\mathcal{B}}) \end{cases}$$

are measurable

Proof. Let $\mathcal{M} := \{f : \Omega \mapsto \mathbb{R}, \text{f is bounded and (*) holds}\}$Then

1. $\mathcal{M}$ is a vector space over $\mathbb{R}$
   If $f, g \in \mathcal{M}, a, b \in \mathbb{R}$ then $af + bg \in \mathcal{M}$ since linear combinations of bounded function is bounded same for measurability.
2. $\mathbb{1}_\Omega \in \mathcal{M}$ obviously.
3. If $(f_n) \in \mathcal{M}$ satisfy $0 < (f_n) \leq c < \infty$ and $f_n \uparrow$, then $\lim\limits_n f_n \in \mathcal{M}$ (limit exist because bounded convergent sequence, and limit of measurable functions are measurable)

Denote $\tau := \{A_1 \times A_2 | A_i \in \mathcal{F}_i\}$, $\mathbb{1}_A \in \mathcal{M}$ since $\forall A \in \tau, \mathbb{1}_A \omega_1, \omega_2 = \mathbb{1}_{A_1}(\omega_1) \cdot \mathbb{1}_{A_2}(\omega_2)$ which is clearly bounded, and if we restrict one coordinate, it it either 0 or indicator function of the other coordinate which is measurable.
But τ is a π-system, therefore by Monotone Class theorem 9.1:

$$\{\text{all bounded } \mathcal{F}\text{-measurable functions} \subset \mathcal{M}\}$$

Note $\sigma(\tau) = \mathcal{F}$

Recall:

Theorem 9.2 (Fubini–Tonelli theorem for Riemann integral)**.** *If $f : \mathbb{R}^2 \mapsto \mathbb{R}^+$ is continuous, then with $I = [a, b], J = [c, d]$ where $a < b, c < d$*

$$\iint_{I\times J} f(x,y)\,dx\,dy = \int_I \int_J f(x,y)\,dy\,dx = \int_J \int_I f(x,y)\,dx\,dy$$

LHS: double integral, RHS: iterated integral.

The same holds for $f : \mathbb{R}^2 \mapsto \mathbb{R}$ continuous provided

$$\iint |f|\,dx\,dy < \infty$$

We have analogous result in the product space for Lebesgue's integral, but we need to define the analogue of double integral (Measure on the product space).
For Riemann integral, the double integral is defined in terms of Riemann sums w.r.t a 2-d lattice of rectangles. The contribution of each rectangle is

$$f(x_i, y_i)\left|R_{ij}\right| \qquad (x_i, y_i) \in R_{ij}, \quad \left|R_{ij}\right| = \text{Area of } R_{ij}$$

The Lebesgue analogue would be to integrate on $\Omega_1 \times \Omega_2$ with the "product measure"

$$\mu(A_1 \cdot A_2) = \mu(A_1) \cdot \mu(A_2) \text{where } A_i \in \mathcal{F}_i$$

Remark. This definition is defined on the π-system which is τ, but we have to extend it to the whole product σ-algebra

We can define this product measure starting with the "rectangle" $A_1 \cdot A_2$ above then extending it using the Carathéodory's machinery.But since we have no idea what result measure we want, we could pursue the second method: However, Fubini's theorem suggest an alternative approach (easier to calculate the measure of a complicated $A \in \mathcal{F}$) using the iterated integrals.

Setup: Assume μ_i are finite measure no $(\Omega_i, \mathcal{F}_i), i = 1, 2$. Let $f : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$ be bounded and measurable and $\mu_i(\Omega_i) < \infty$

Def . The functions

$$I_1(\omega_1, f) := \int_{\Omega_2} f(\omega_1, \omega_2)\, d\mu_2(\omega_2)$$
$$I_2(\omega_2, f) := \int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)$$

are well defined from $\Omega_i \mapsto \mathbb{R}$ follow the proposition 9.1

Lemma 9.1. *If $f : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$ is bounded and measurable then*

$$(**) \begin{cases} (i) & \text{For } i = 1, 2 : I_i(\cdot, f) : (\Omega_i, \mathcal{F}_i) \mapsto (\mathbb{R}, \mathcal{B}) \text{ is bounded and measurable} \\ (ii) & \text{Fubini: } \int_{\Omega_1} I_1(\cdot, f)\, d\mu_1 = \int_{\Omega_2} I_2(\cdot, f)\, d\mu_2 \end{cases}$$

Proof. Let $\mathcal{M}$ be the class of bounded measurable functions for (**) holds. Then

Exercise 9.2. Preparation work for Monotone Class theorem

1. $\mathcal{M}$ is a vector space over $\mathbb{R}$
2. $\mathbb{1}_\Omega \in \mathcal{M}$
3. If $(f_n) \in \mathcal{M}$ satisfy $0 < f_n \leq c < \infty$ and $f_n \uparrow$ then $\lim_n f_n \in \mathcal{M}$
4. $\mathcal{M}$ contains $\mathbb{1}_A$ for $A = A_1 \cdot A_2 \in \tau$

By Monotone Class theorem 9.1: Every bounded $\mathcal{F}$-measurable function is inside $\mathcal{M}$

Remark. We can extend above result to more general case (σ-finite)

1. By replacing (i) in the above lemma by
   (i')For $i = 1, 2 : I_i(\cdot, f) : (\Omega_i, \mathcal{F}_i) \mapsto (\mathbb{R}, \mathcal{B})$ is measurable
   The lemma can be extended to σ-finite measures μ_i and $f \geq 0$, so (i')(ii) holds
2. If follows that if μ_i are σ-finite and f is measurable, then $I_i(\cdot, f^+)$ and $I_i(\cdot, f^-)$ are measurable and if for $A_i := \{\omega \in \Omega_i | I_i(\cdot, f^+) = \infty, I_i(\cdot, f^-) = \infty\}, \mu_i(A_i) = 0$ then $I_i(\cdot, f)$ is well defined on $\Omega_i \setminus \mathcal{A}_i$ and $I_i(\cdot, f)\mathbb{1}_{A_i^c}$ is measurable, and on A_i^c we have

$$I_i(\cdot, f) = I_i(\cdot, f^+) - I_i(\cdot, f^-)$$

product measure

Given $(\Omega_i, \mathcal{F}_i, \mu_i)$ with μ_i finite measures for $i = 1, 2$ we define $\mu : \mathcal{F} \mapsto \mathbb{R}^+$ as

$$\bigoplus \quad \mu(A) := \int_{\Omega_1} I_1(\cdot, \mathbb{1}_A)\, d\mu_1 = \int_{\Omega_2} I_2(\cdot, \mathbb{1}_A)\, d\mu_2$$

Proposition 9.2. *After we define the product measure*

1. *μ is a measure on $(\Omega, \mathcal{F})$ called the product measure*
2. *$\mu(A_1 \cdot A_2) = \mu_1(A_1) \cdot \mu_2(A_2), \forall A_i \in \mathcal{F}_i$*
3. *If ν is another measure on $(\Omega, \mathcal{F})$ with $\nu(A_1 \cdot A_2) = \mu(A_1 \cdot A_2) = \mu_1(A_1) \cdot \mu_2(A_2), \forall A_i \in \mathcal{F}_i$ then $\mu = \nu$*

Proof. Exercise :)

Remark. The identity $\bigoplus$ as well as the notion of the product measure can be extended to σ-finite measures

Corollary 9.1 (Lebesgue measure on $\mathbb{R}^2$)**.** *There exists a unique measure λ_2 on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ s.t. for any intervals $I_i : (a_i, b_i]$ we have*

$$\lambda_2(I_1 \cdot I_2) = \lambda(I_1)\lambda(I_2) = (b_1 - a_1)(b_2 - a_2)$$

Theorem 9.3. *Let μ be the product measure on $(\Omega, \mathcal{F})$ of $(\Omega_i, \mathcal{F}_i, \mu_i)$ with μ_i are σ-finite measures for $i = 1, 2$*

1. *If f is non-negative measurable functions of $(\Omega, \mathcal{F})$ then*

$$\bigotimes \quad \int f\, d\mu := \int_{\Omega_1} I_1(\cdot, f)\, d\mu_1 = \int_{\Omega_2} I_2(\cdot, f)\, d\mu_2$$

2. *$f \in L^1(\Omega) \iff \mathbb{1}_1(\cdot, |f|) \in L^1(\Omega_1) \iff \mathbb{1}_2(\cdot, |f|) \in L^1(\Omega_2)$ and then $\bigotimes$ holds*

Proof. Exercise, again :)