Le Gall: Measure Theory, Probability, and Stochastic Processes Companion Notes

Taken by: Mark Zhu

August 2025

Contents

Part I
Measure Theory

1 Measurable Spaces

`Def 1.2` $\mathcal{C}$ is a collection of subsets of E.

`p. 4` "Clearly, closed subsets of E are Borel sets."
Let $U \subset E$, and U is open $\implies U \in \mathcal{B}(E)$. Since $\mathcal{B}(E)$ is a σ-field, $E \setminus U \in \mathcal{B}(E)$, which is closed.

`Lem 1.5` To prove $\mathcal{B}(E) \otimes \mathcal{B}(F) \subset \mathcal{B}(E \times F)$, consider the following class

$$\mathcal{C}_B = \{A \in \mathcal{B}(E) : A \times B \in \mathcal{B}(E \times F)\}$$

Fix $B \subset F$ be open. $\mathcal{B}(E \times F) = \sigma(\mathcal{O})$ where $\mathcal{O} = \{\text{all open subsets of } E \times F\}$, and $E \times F = \{(e, f) : e \in E, f \in F\}$. If $A \subset E$ is also open, then $A \times B \in \mathcal{B}(E \times F), A \in \mathcal{C}_B, \mathcal{B}(E) \subset \mathcal{C}_B$. Clearly $\mathcal{C}_B \subset \mathcal{B}(E)$, so $\mathcal{C}_B = \mathcal{B}(E)$.

`p. 7` In the proof of property (4), discuss why

$$\mu\left(B_1 \setminus \bigcap_{n\in\mathbb{N}} B_n\right) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right)$$

Derive $B_1 \setminus \bigcap_{n\in\mathbb{N}} B_n = B_1 \cap (\bigcap_{n\in\mathbb{N}} B_n)^c = B_1 \cap (\bigcup_{n\in\mathbb{N}} B_n^c) = \bigcup_{n\in\mathbb{N}}(B_1 \cap B_n^c) = \bigcup_{n\in\mathbb{N}} A_n$.

`p. 9` Discuss the equivalence of limsup and liminf for sequences of numbers $\{a_n\}$ and of sets $\{A_n\}$. Recall

$$\limsup_{n\to\infty}\{a_n\} = \lim_{n\to\infty} \sup_{k\geq n}\{a_k\}, \quad \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

Define an indicator function

$$a_n(x) = \begin{cases} 1 & x \in A_n \\ 0 & x \notin A_n \end{cases}$$

$\sup_{k\geq n}\{a_k(x)\} = 1$ if $\exists k \geq n$ such that $x \in A_k$. $\sup_{k\geq n}\{a_k(x)\} = 0$ otherwise. Take limits, then $\lim_{n\to\infty} \sup_{k\geq n}\{a_k\} = 1$ if $\forall n \geq 1, \exists k \geq n$ such that $x \in A_k$.
Note that $\forall n \geq 1 \iff \bigcap_{n\geq 1}$; $\exists k \geq n \iff \bigcup_{k\geq n}$, so

$$\limsup_{n\to\infty}\{a_n(x)\} = \begin{cases} 1 & x \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \\ 0 & \text{otherwise} \end{cases}$$

For liminf, it's more natural to consider the inverse statement. $\inf_{k\geq n}\{a_k(x)\} = 0$ if $\exists k \geq n$ such that $x \notin A_k$, so $a_k(x) = 0$. $\lim_{n\to\infty} \inf_{k\geq n}\{a_k\} = 0$ if $\forall n \geq 1, \exists k \geq n$ such that $x \notin A_k$. That means if $\exists n \geq 1, \forall k \geq n, x \in A_k$, then $\lim_{n\to\infty} \inf_{k\geq n}\{a_k\} = 1$. Thus

$$\liminf_{n\to\infty}\{a_n(x)\} = \begin{cases} 1 & x \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k \\ 0 & \text{otherwise} \end{cases}$$

Conclusion: the gap between real numbers and sets is closed by *indicator functions*.

`Lem 1.7` Because $\bigcap_{k=n}^{\infty} A_k \subset A_k$, for all $k \geq n$,

$$\mu\left(\bigcap_{k=n}^{\infty} A_k\right) \leq \mu(A_k)$$

So $\mu\left(\bigcap_{k=n}^{\infty} A_k\right) \leq \inf_{k\geq n} \mu(A_k)$. Since $\bigcap_{k=n}^{\infty} A_k$ is increasing with respect to n, use property (3) in p. 6, and get the desired result.

`Prop 1.10` To let f measurable, two conditions should suffice

1. $\exists$ generator $\mathcal{C} \subset \mathcal{B}$, such that $\sigma(\mathcal{C}) = \mathcal{B}$
2. f is measurable on $\mathcal{C}$

By Def 1.8 and the second condition, $\forall B \in \mathcal{C}, f^{-1}(B) \in \mathcal{A}$. So $B \in \mathcal{G}, \mathcal{C} \subset \mathcal{G}$. Now show that $\mathcal{G}$ is a σ-field.

- $F \in \mathcal{B}, f^{-1}(F) = E \in \mathcal{A} \implies F \in \mathcal{G}$
- Let $B \in \mathcal{G}$, so $f^{-1}(B) \in \mathcal{A}$. Since $\mathcal{A}$ is a σ-field, $\mathcal{A} \ni E \backslash f^{-1}(B) = f^{-1}(F \backslash B)$, so $B^c = F \backslash B \in \mathcal{G}$
- If $B_n \in \mathcal{G}$, then $f^{-1}(B_n) \in \mathcal{A}, f^{-1}(\bigcup_{n\geq 1} B_n) = \bigcup_{n\geq 1} f^{-1}(B_n) \in \mathcal{A}$. So $\bigcup_{n\geq 1} B_n \in \mathcal{G}$

So $\mathcal{G}$ is a σ-field. Since $\mathcal{C} \subset \mathcal{G}$, $\mathcal{B} = \sigma(\mathcal{C}) \subset \mathcal{G}$. By construction of $\mathcal{G}$, $\mathcal{G} \subset \mathcal{B}$. Thus $\mathcal{G} = \mathcal{B} \implies \forall B \in \mathcal{B}, f^{-1}(B) \in \mathcal{A}$, so f is measurable.

`Lem 1.15`

1. "The set A of all $x \in E$ such that $f_n(x)$ converges in $\mathbb{R}$ as $n \to \infty$ is measurable" where 'converges' means $\limsup_n f_n = \liminf_n f_n$.
2. $A = G^{-1}(\Delta)$ is followed by Δ is a measurable set, G is a measurable function, so $G^{-1}(\Delta) \in \mathcal{A}$ under measurable space $(E, \mathcal{A})$.
3. Then move on to the second statement that $h : E \to \mathbb{R}$ is measurable. If $\forall F \subset \mathbb{R}, h^{-1}(F) \in \mathcal{A}$, i.e. $h^{-1}(F)$ is a measurable set, then h is a measurable function.
   (Case 1) If $0 \notin F$, then h collapses into $h(x) = \lim_n f_n(x), \forall x \in A$. Then $h^{-1}(F)$ is given by
   $$h^{-1}(F) = A \cap \{x \in E : \limsup f_n(x) \in F\}$$
   and both sets on the right are measurable.
   (Case 2) If $0 \in F$, then $0 \notin F^c$.
   $$x \in (h^{-1}(F))^c \iff h(x) \notin F \iff h(x) \in F^c \iff x \in h^{-1}(F^c)$$
   which means $(h^{-1}(F))^c = h^{-1}(F^c)$. From Case 1 we know $h^{-1}(F^c)$ is a measurable set, so does $(h^{-1}(F))^c$. Because the collection of measurable sets is a σ-field (by Def 1.1), $h^{-1}(F)$ is also measurable.

`p. 13` "Any σ-field is also a monotone class. Conversely, a monotone class $\mathcal{M}$ is a σ-field if and only if it is closed under finite intersections."
Question: How 'closure under finite intersections' helps $\mathcal{M}$ becoming a σ-field?
For property (ii), set $B = E$, then $A^c = E \setminus A \in \mathcal{M}$. For property (iii), given $\mathcal{M}$ is closed under finite intersections, let $A, B \in \mathcal{M}, A^c \cap B^c \in \mathcal{M} \implies A \cup B = (A^c \cap B^c)^c$. So it's also closed under finite union. Define
$$B_n = \bigcup_{k=1}^{n} A_k$$
clearly $B_n \subset B_{n+1}$. Apply (iii),
$$\mathcal{M} \ni \bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} \bigcup_{k=1}^{n} A_k = \bigcup_{k\in\mathbb{N}} A_k$$
without the restriction of $A_n \subset A_{n+1}$.

`Thm 1.18` Explanation about the proof

1. When verifying $\mathcal{M}_A$ is a monotone class, in the complement property
   $$\mathcal{M}(\mathcal{C}) \ni (A \cap B') \setminus (A \cap B) = A \cap (B' \setminus B)$$
   is based on the construction of $\mathcal{M}_A$ and nothing else.

2. Before the last step, we reached the result

$$\forall A \in \mathcal{C}, \forall B \in \mathcal{M}(\mathcal{C}), A \cap B \in \mathcal{M}(\mathcal{C})$$

But to show $\mathcal{M}(\mathcal{C})$ is closed under finite intersections, we need

$$\forall A, B \in \mathcal{M}(\mathcal{C}), A \cap B \in \mathcal{M}(\mathcal{C})$$

Previously, we fixed $A \in \mathcal{C}$ to reach the property $\mathcal{C} \subset \mathcal{M}_A$ and everything follows. So we cannot simply replace $A \in \mathcal{C}$ by $A \in \mathcal{M}(\mathcal{C})$. Here's what we do:

Fix $B \in \mathcal{C}$. Now A can be an arbitrary set in $\mathcal{M}(\mathcal{C})$. Use the conclusion above, $A \cap B \in \mathcal{M}(\mathcal{C})$, so $B \in \mathcal{M}_A$. That is, $\forall B \in \mathcal{C}, B \in \mathcal{M}_A$, thus $\mathcal{C} \subset \mathcal{M}_A, \forall A \in \mathcal{M}(\mathcal{C})$. Since $\mathcal{M}_A$ is a monotone class, $\mathcal{M}(\mathcal{C}) \subset \mathcal{M}_A$, $\mathcal{M}(\mathcal{C})$ is closed under finite intersection.

$\boxed{\texttt{Cor 1.19}}$ In (1), $\mu(E) = v(E)$ ensures that $\mathcal{G}$ satisfies its first property for being a monotone class. In (2), μ_n, v_n are respective restrictions of μ, v to E_n, i.e.

$$\mu_n(A) = \mu(A \cap E_n), \quad \forall A \in \mathcal{A}$$

$\mu(E_n) = v(E_n)$ allows us to apply (1) to $\mu_n, v_n, \implies \mu_n = v_n$. Then

$$\mu(A) = \mu(A \cap E) = \mu\left(A \cap \bigcup_{n \in \mathbb{N}} E_n\right) = \mu\left(\bigcup_{n \in \mathbb{N}} (A \cap E_n)\right) = \lim_{n \to \infty} \uparrow \mu(A \cap E_n)$$

by property (3) in p. 6, where $A \cap E_n$ is increasing.

2 Integration of Measurable Functions

`Prop 2.3` Rigorously prove property (2). Let $h \in \mathcal{E}_+$ such that $h \leq f$. Then if $h(x) > 0, f(x) \geq h(x) > 0$, so

$$\{x \in E : h(x) > 0\} \subset \{x \in E : f(x) > 0\}$$

Given $\mu(\{x \in E : f(x) > 0\}) = 0$, so $\mu(\{x \in E : h(x) > 0\}) = 0$. The integral of h would be

$$\int h \mathrm{d}\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i) = 0$$

since either $\alpha_i = 0$ or $\alpha_i > 0$ but $\mu(A_i) = 0$. Because $h \leq f$ always holds,

$$\int f \mathrm{d}\mu = \sup_{h \leq f} \int h \mathrm{d}\mu = 0$$

`Thm 2.4`

1. Why is f measurable?

   $f_n \to f \implies f = \limsup f_n = \liminf f_n$, then by Prop 1.14.

2. Why $\int f \mathrm{d}\mu \geq \lim_n \int f_n \mathrm{d}\mu$?

   Since f_n is increasing, $f_n \leq \lim_n f_n = f$. Take integrals, $\int f \mathrm{d}\mu \geq \int f_n \mathrm{d}\mu$, then take limit $n \to \infty$ on the RHS.

3. Why $f_n \to f, a < 1$ leads to $E_n \uparrow E$?

   We want to prove (1) $E_n \subset E_{n+1}, \forall n$, (2) $\bigcup_{n=1}^{\infty} E_n = E$.

   (1) Take $x \in E_n, ah(x) \leq f_n(x) \leq f_{n+1}(x)$, so $x \in E_{n+1}$.

   (2) Let $x \in E, h(x) \leq f(x), 0 \leq a < 1$.

   (a) $h(x) = 0$. Then $ah(x) = 0 \leq f(x)$, so $x \in E_n, \forall n$.

   (b) $h(x) > 0$. Then $ah(x) < h(x) \leq f(x)$. Since $f_n(x) \uparrow f(x), \exists N \in \mathbb{N}$, such that for $n \geq N, ah(x) < f_n(x)$. Thus $\forall n \geq N, x \in E_n$.

   We conclude that, $\forall x \in E, \exists N \in \mathbb{N}$, such that $\forall n \geq N, x \in E_n$, so

   $$\bigcup_{n \in \mathbb{N}} E_n = E$$

   Combined with $E_n \subset E_{n+1}, E_n \uparrow E$.

4. Given $f_n \geq a\mathbf{1}_{E_n} h$, then

   $$\int a\mathbf{1}_{E_n} h \mathrm{d}\mu = \int a\mathbf{1}_{E_n} \left(\sum_{i=1}^{k} \alpha_i \mathbf{1}_{A_i} \right) \mathrm{d}\mu = \int \left(a \sum_{i=1}^{k} \alpha_i \mathbf{1}_{E_n \cap A_i} \right) \mathrm{d}\mu = a \sum_{i=1}^{k} \alpha_i \mu(E_n \cap A_i)$$

5. In the last step

   $$\int h \mathrm{d}\mu \leq \lim_{n \to \infty} \uparrow \int f_n \mathrm{d}\mu$$

   take $\sup_{h \in \mathcal{E}_+, h \leq f}$ on both sides.

6. Conclusion:

   The most insightful part of this proof is to construct a simple function $h = \sum_{i=1}^{k} \alpha_i \mathbf{1}_{A_i}$ and restrict $h \leq f$, then use Def 2.2.

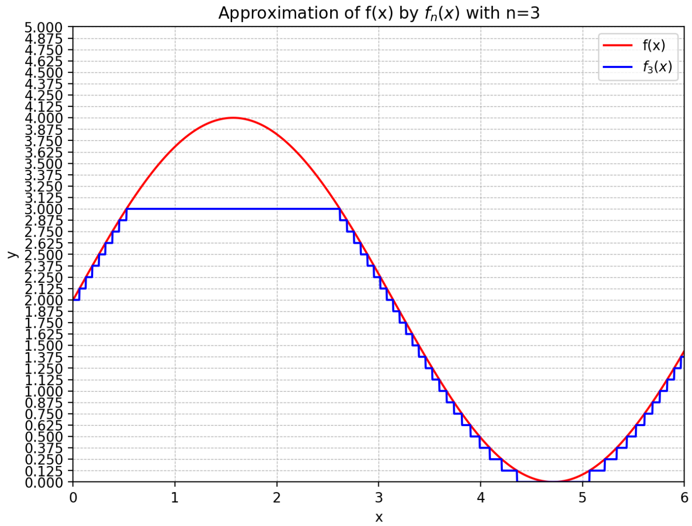

Figure 1: Example for Prop 2.5(1)

`Prop 2.5` The author used constructive proof for (1). Recall uniform convergence: $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ such that $\forall n \geq N, x \in E, |f_n(x) - f(x)| < \varepsilon$. In uniform convergence, N cannot depend on x. The constructed f_n satisfies $0 \leq f - f_n \leq 2^{-n}$ where 2^{-n} does not depend on x.
Example: $f(x) = 2\sin(x) + 2, n = 3$, see Fig 1.

`p. 23` Counting measure is simply summing up, analogously, `np.sum(x)` in numpy.

1. $f_n(k) = f(n)\mathbf{1}_{\{n\}}(k)$, which means only when $k = n$ does $\mathbf{1}_{\{n\}}(k) = 1$, thus $\sum_n f_n(k) = \sum_n f(n)\mathbf{1}_{\{n\}}(k) = f(k)$ (fix k and treat n as the variable). So $\sum_n f_n = f$. By Prop 2.5(3),

$$\int f \mathrm{d}\mu = \int \left(\sum_{n \in \mathbb{N}} f_n \right) \mathrm{d}\mu = \sum_{n \in \mathbb{N}} \int f_n \mathrm{d}\mu$$

Since f_n is a simple function,

$$\int f_n \mathrm{d}\mu = \int f(n)\mathbf{1}_{\{n\}}(k) \mathrm{d}\mu(k) = \sum_{k \in \mathbb{N}} f(n)\mathbf{1}_{\{n\}}(k) = f(n)$$

We conclude that

$$\int f \mathrm{d}\mu = \sum_{n \in \mathbb{N}} f(n).$$

2. $f_n(k) = a_{n,k}$, remember that k is the variable.

$$\int \left(\sum_{n \in \mathbb{N}} f_n \right) \mathrm{d}\mu = \int \left(\sum_{n \in \mathbb{N}} a_{n,k} \right) \mathrm{d}\mu = \sum_{k \in \mathbb{N}} \left(\sum_{n \in \mathbb{N}} a_{n,k} \right)$$

$$\sum_{n \in \mathbb{N}} \int f_n \mathrm{d}\mu = \sum_{n \in \mathbb{N}} \int a_{n,k} \mathrm{d}\mu = \sum_{n \in \mathbb{N}} \left(\sum_{k \in \mathbb{N}} a_{n,k} \right)$$

Then use Prop 2.5(3).

`Cor 2.6` Detailed proof for equation (2.1).

(Step 1) For indicator function $f = \mathbf{1}_A$

$$\int \mathbf{1}_A \mathrm{d}v = v(A) = \int \mathbf{1}_A g \mathrm{d}\mu$$

by definition of integral (see Def 2.1) and $v(A)$.

(Step 2) Extend to simple function using Prop 2.5(2),$f_n = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$,

$$\int f_n \mathrm{d}v = \int \left(\sum_{i=1}^n \alpha_i \mathbf{1}_{A_i} \right) \mathrm{d}v = \sum_{i=1}^n \alpha_i \int \mathbf{1}_A \mathrm{d}v$$

By Step 1,

$$= \sum_{i=1}^n \alpha_i \int \mathbf{1}_A g \mathrm{d}\mu = \int f_n g \mathrm{d}\mu$$

(Step 3) Extend to all nonnegative measurable function using Prop 2.5(1). $\exists \{f_n\}_n$ such that $f_n \uparrow f$. By Thm 2.4,

$$\int f \mathrm{d}v = \lim_{n\to\infty} \uparrow \int f_n \mathrm{d}v = \lim_{n\to\infty} \uparrow \int f_n g \mathrm{d}\mu = \int f g \mathrm{d}\mu$$

`Prop 2.7` (2) Logic: write ∞ as $\forall n \geq 1, \exists f(x) > n$. Call $\{f(x) > n\}$ an event, denoted by A_n. $\forall n \geq 1$ in set theory language is $\bigcap_{n\geq 1}$.
Question: Does the reverse hold? Answer: No! Counterexample: $f(x) = 1/x, x \in (0, 1]$, and let μ be Lebesgue measure. $f < \infty$ a.e., but $\int_0^1 \frac{1}{x} \mathrm{d}x = \infty$.
(3) To show the set such that $f(x)$ lies in $(0, +\infty)$ is measure zero, find a sequence of sets such that

$$(0, +\infty) = \bigcup_{n=1}^{\infty} [1/n, +\infty)$$

`Prop 2.9` The proof appears extremely similar to Cor 2.6.

(Step 1) Let $h = \mathbf{1}_B$, then $h = 1$ only when $\varphi(x) \in B, x \in \varphi^{-1}(B)$.

$$\int_E \mathbf{1}_B(\varphi(x))\mu(\mathrm{d}x) = \int_E \mathbf{1}_{\varphi^{-1}(B)}(x)\mu(\mathrm{d}x) = \mu(\varphi^{-1}(B)) = v(B) = \int_F \mathbf{1}_B(y)v(\mathrm{d}y)$$

(Step 2) If h_n is a simple function, $h_n = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$.

$$\int_E h_n(\varphi(x))\mu(\mathrm{d}x) = \sum_{i=1}^n \alpha_i \int_E \mathbf{1}_{A_i}(\varphi(x))\mu(\mathrm{d}x) = \sum_{i=1}^n \alpha_i v(A_i) = \int_F h_n(y)v(\mathrm{d}y)$$

(Step 3) For all nonnegative measurable function h, contruct $h_n \uparrow h$ using the method in Prop 2.5(1). By Thm 2.4 (MCT),

$$\int_E h(\varphi(x))\mu(\mathrm{d}x) = \lim_{n\to\infty} \uparrow \int_E h_n(\varphi(x))\mu(\mathrm{d}x) = \lim_{n\to\infty} \uparrow \int_F h_n(y)v(\mathrm{d}y) = \int_F h(y)v(\mathrm{d}y)$$

`p. 34` "Notice that (iii)'⇒(iii) thanks to the mean value theorem."
$\exists \xi \in (u_0, u)$ or $\in (u, u_0)$, such that

$$|f(u,x) - f(u_0,x)| \leq \left| \frac{\partial f}{\partial u}(\xi, x) \right| \cdot |u - u_0| \leq g(x) \cdot |u - u_0|$$

3 Construction of Measures

Def 3.1

Can σ-subadditive $\Rightarrow$ σ-additive? No. Counterexample: define

$$\mu(A) = \begin{cases} 1 & A \neq \varnothing \\ 0 & A = \varnothing \end{cases}$$

Then $\forall \{A_k\}_k$,

$$\mu\left(\bigcup_{k\geq 1} A_k\right) = \mathbf{1}_{\{\exists k: A_k \neq \varnothing\}} \leq \sum_{k\geq 1} \mathbf{1}_{\{A_k \neq \varnothing\}} = \sum_{k\geq 1} \mu(A_k)$$

so σ-subadditive holds. Take $A_k = \{x_k\}$, all disjoint, but

$$\mu\left(\bigcup_{k\geq 1} A_k\right) = \mu(\{x_1, \cdots, x_n\}) = 1 \neq \infty \impliedby \sum_{k\geq 1} \mu(A_k) = \sum_{k\geq 1} 1$$

Thm 3.3

(i) How does the author come up with equation (3.1) to prove (i)? We want to show $\bigcup_{k\in\mathbb{N}} B_k \in \mathcal{M}$, then show

$$\mu^*(A) \geq \underbrace{\mu^*\left(A \cap \bigcup_{k\in\mathbb{N}} B_k\right)}_{(1)} + \underbrace{\mu^*\left(A \cap \left(\bigcup_{k\in\mathbb{N}} B_k\right)^c\right)}_{(2)}$$

Apply De Morgan's Law on (2). For (1), rewrite it as

$$\mu^*\left(A \cap \bigcup_{k\in\mathbb{N}} B_k\right) = \mu^*\left(\bigcup_{k\in\mathbb{N}} (A \cap B_k)\right)$$

and by σ-subadditive of μ^*,

$$\leq \sum_{k\in\mathbb{N}} \mu^*(A \cap B_k)$$

So the minimal requirement is

$$\mu^*(A) = \sum_{k\in\mathbb{N}} \mu^*(A \cap B_k) + \mu^*\left(A \cap \bigcap_{k\in\mathbb{N}} B_k^c\right)$$

Plug in $k = 1$ and check, $\mu^*(A) = \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c)$, meaning the equation is very likely to hold, then we can safely explore the proof.

(ii) $\mathcal{M}(\mu^*) \subset \mathcal{P}(E)$, restriction of μ^* to $\mathcal{M}$ simply changes the domain of μ^* from $\mathcal{P}(E)$ to $\mathcal{M}$, maintaining the value unchanged. So for $\mu := \mu^*|_{\mathcal{M}}$,

$$\mu : \mathcal{M} \to [0, +\infty], \quad \mu(B) = \mu^*(B), \forall B \in \mathcal{M}$$

p. 44

"The infimum is over all countable covers of A by open intervals $(a_i, b_i), i \in \mathbb{N}$ (it is trivial that such covers exist)." This is not trivial, it's *Heine-Borel Theorem*.

p. 46

1. Summing up the two inequalities, the RHS gives,

$$\begin{aligned}[(b_i \wedge \alpha) - (a_i \wedge \alpha)] + [(b_i \vee \alpha) - (a_i \vee \alpha)] &= [(b_i \wedge \alpha) + (b_i \vee \alpha)] - [(a_i \wedge \alpha) + (a_i \vee \alpha)] \\ &= (b_i + \alpha) - (a_i + \alpha) = b_i - a_i\end{aligned}$$

2. By Lem 3 in the lecture note MIT18.S190 Lec4, that a metric space X being sequentially compact implies that X is *totally bounded* ($\forall \varepsilon > 0, \exists x_1, \cdots, x_k \in X$ with finite k such that $\{B_\varepsilon(x_i)\}_k$ is an open cover of X). It follows $\exists N, [a, b] \subset \bigcup_{i=1}^N (a_i, b_i)$.

`Prop 3.6` $\mathcal{N}$ shows the situation when a set $A \notin \mathcal{A}$ but it's small enough to be negligible. Another name for $\bar{A}$ is *completion*, and this proposition says there exists a unique extended measure that (1) gives the same value for all sets in $\mathcal{A}$, and (2) assigns zero measure to all sets in $\mathcal{N}$.

1. Verify that $\mathcal{B}$ is a σ-field

   (a) Let $B = B' = E \in \mathcal{A}, \mu(E \setminus E) = 0 \implies E \in \mathcal{B}$.

   (b) If $A \in \mathcal{B}$, then $\exists B, B' \in \mathcal{A}, B \subset A \subset B', \mu(B' \setminus B) = 0$. $B^c \supset A^c \supset B'^c, \mu(B^c \setminus B'^c) = \mu(B^c \cap B') = \mu(B' \setminus B) = 0$. So $A^c \in \mathcal{B}$.

   (c) If $A_n \in \mathcal{B}, \forall n \geq 1$, then $\exists B_n, B'_n \in \mathcal{A}, B_n \subset A_n \subset B'_n, \mu(B'_n \setminus B_n) = 0$. Let $B := \bigcup_{n\geq 1} B_n \subset \bigcup_{n\geq 1} A_n \subset \bigcup_{n\geq 1} B'_n =: B'$,

$$\mu(B' \setminus B) = \mu\left(\bigcup_{n\geq 1} B'_n \cap \bigcap_{n\geq 1} B_n^c\right) = \mu\left(\bigcup_{n\geq 1}(B'_n \cap \bigcap_{n\geq 1} B_n^c)\right)$$

   By σ-subadditive,

$$\leq \sum_{n\geq 1} \mu(B'_n \cap \bigcap_{n\geq 1} B_n^c) \leq \sum_{n\geq 1} \mu(B'_n \cap B_n^c) = \mu(B'_n \setminus B_n) = 0$$

   So $\bigcup_{n\geq 1} A \in \mathcal{B}$.

`Thm 3.8` (1) To prove the first statement, we used the pushforward of measure. $\sigma_x(\lambda)$ is a pushforward that has property $\lambda(x + A) = \lambda(A)$, we want to show $\sigma_x(\lambda) = \lambda$. Use Cor 1.19, and check if the conditions of this corollary hold. First, a $\mathcal{C} \subset \mathcal{A}$ such that $\sigma(\mathcal{C}) = \mathcal{A}$, we find open box $\mathcal{C} = (-K, K)^d$. Second, $\forall A \in \mathcal{C}, \sigma_x(\lambda)(A) = \lambda(A)$, apparently the volume of a box does not change by shifting. So apply Cor 1.19 and get $\sigma_x(\lambda) = \lambda$.
(2) Since $\lambda([0,1)^d) = 1$, cut $[0,1)^d$ into n^d boxes, since we have already set $c = \mu\left([0,1)^d\right)$,

$$\mu\left([0,1)^d\right) = n^d \cdot \mu\left(\left[0, \tfrac{1}{n}\right)^d\right) \Rightarrow \mu\left(\left[0, \tfrac{1}{n}\right)^d\right) = \frac{c}{n^d}$$

`Prop 3.9` $F = C \setminus U$ is a compact set because C is compact, U is open, then $C \setminus U$ is a closed subset of a compact set, which is also compact. (See Lem 20 of MIT18.S190 Lec3).
Why $\lambda(F) \geq \lambda(C) - \lambda(U)$? Cut C into two disjoint sets, $\lambda(C) = \lambda(C \cap U) + \lambda(C \setminus U) \leq \lambda(U) + \lambda(C \setminus U)$.

`Thm 3.12` (ii) statement: Given a function $F : \mathbb{R} \to \mathbb{R}_+$ with some properties (increasing, bounded, right-continuous, and $F(-\infty) = 0$), then $\exists!$ measure μ such that $F(x) = \mu((-\infty, x])$. Construct

$$\mu^*(A) = \inf\left\{\sum_{i\in\mathbb{N}}(F(b_i) - F(a_i)) : A \subset \bigcup_{i\in\mathbb{N}}(a_i, b_i]\right\}$$

(Step 1) $\mu^*(A)$ is an outer measure, rf. Thm 3.4(i). ("increasing", "bounded")

(Step 2) $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}(\mu^*)$, so that restricting μ^* on $\mathcal{B}(\mathbb{R})$ gives a measure μ, rf. Thm 3.4(ii).

(Step 3) We still need to show $F(x) = \mu((-\infty, x])$. It suffices to check $F(b) - F(a) = \mu((a, b])$ ("$F(-\infty) = 0$"). Let $\varepsilon > 0$. Since $\mathbb{R}$ is compact, exists *finite* collection of open sets that covers $[a + \varepsilon, b]$.

$$F(b) - F(a + \varepsilon) \leq \sum_{i=1}^{\infty}(F(y_i) - F(x_i)) + \varepsilon$$

As $\varepsilon \to 0, F(a + \varepsilon) \to F(a)$ ("right-continuous"). Thus $F(b) - F(a) \leq \mu((a, b])$. The reverse $\geq$ is by construction of μ^*.

4 L^p Spaces

`Thm 4.5` See the proof that l^p is a Banach space here.

$$l^p : \|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \quad , \quad L^p : \|f\|_p = \left(\int |f|^p \,\mathrm{d}\mu\right)^{1/p}$$

In ℓ^p, the input is a sequence (or vector) $x = \{x_i\}_{i\geq 1}$. In L^p, the input is an equivalence class of measurable functions $[f]$.

`Def 4.10` Absolute continuity $v \ll \mu$ is an analogy from real analysis. $\Delta x \to 0 \implies \Delta f \to 0$; $\mu(A) = 0 \implies v(A) = 0$. f is controlled by x and v is controlled by μ. See Fig 2.

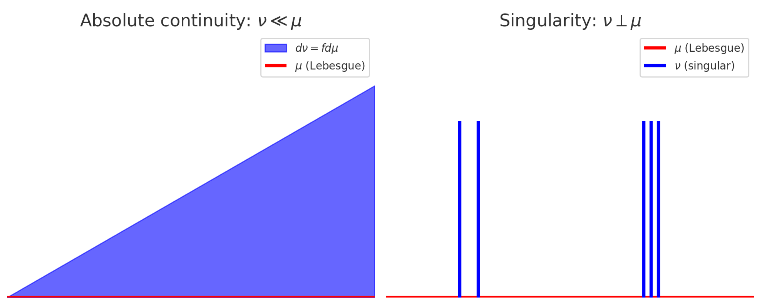

Figure 2: Absolute continuous and Singular

`Thm 4.11` Companion for Radon-Nikodym theorem.

1. (Step 1). (1.a) $L^2(\mu) \subset L^1(\mu)$. If $f \in L^2(\mu)$ and μ is a finite measure ($\mu(E) < \infty$), then set $g \equiv 1$, by Cauchy-Schwarz inequality

$$\int |f| \,\mathrm{d}\mu \leq \left(\int |f|^2 \,\mathrm{d}\mu\right)^{1/2} \cdot (\mu(E))^{1/2} < \infty$$

So $f \in L^1(\mu)$, thus $L^2(\mu) \subset L^1(\mu)$.

(1.b) Because we assume $v \leq \mu$, if $\mu(A) = 0$, then $v(A) \leq \mu(A) = 0$, thus $v \ll \mu$. Therefore $f = \tilde{f}$, μ-a.e. $\Rightarrow f = \tilde{f}$, ν-a.e.

(1.c) Let $C = v(E)^{1/2}$, because $|\Phi(f)| \leq C\|f\|_{L^2(\mu)}$, $\Phi(f)$ is a bounded linear map. Therefore it is continuous, see Proposition 1.3 in MIT18.102. So now we can apply Riesz.

(1.d) Show $g \geq 0$. For every $\varepsilon > 0$,

$$\mu(\{x \in E : g(x) \leq \varepsilon\}) \geq v(\{x \in E : g(x) \leq \varepsilon\}) = \int_{\{x\in E: g(x)\leq\varepsilon\}} g \,\mathrm{d}\mu \geq \varepsilon \cdot \mu(\{x \in E : g(x) \leq \varepsilon\})$$

So $\mu(\{x \in E : g(x) \leq \varepsilon\}) = 0$, take a decreasing sequence of $\varepsilon \to 0$, then $\mu(\{x \in E : g(x) < 0\}) = 0$. $g \geq 0$ μ a.e..

(1.e) The result of Step 1: Given a measure μ, then there exists a $v \leq \mu$ that is the measure of density g with respect to μ, i.e. $v = g \cdot \mu$ (see Cor 2.6),

$$\int f \,\mathrm{d}v = \int fg \,\mathrm{d}\mu$$

2. (Step 2). (2.a) We need f to be a bounded measurable function for

$$\int f \ \mathrm{d}v = \int fh \ \mathrm{d}\mu + \int fh \ \mathrm{d}v \implies \int f(1-h) \ \mathrm{d}v = \int fh \ \mathrm{d}\mu$$

To avoid $\infty - \infty$ in rearrangement, we need to show $\int f \, \mathrm{d}v, \int fh \, \mathrm{d}\mu$ and $\int fh \, \mathrm{d}v$ are all finite. By $0 \le h \le 1$, only need to show $\int f \, \mathrm{d}v, \int f \, \mathrm{d}\mu$ is finite. So if f is bounded,

$$\int |f| \, \mathrm{d}v \le \sup |f| \cdot \int \, \mathrm{d}v = \|f\|_\infty \cdot v(E) < \infty$$

similar for μ.

(2.b) Then set $f_n := f' \wedge n, f_n \uparrow f'$, so f_n is bounded while f' is any nonnegative measurable function.

$$\int f'(1-h) \, \mathrm{d}v = \lim_{n\to\infty} \int f_n(1-h) \, \mathrm{d}v, \quad \int f'h \, \mathrm{d}\mu = \lim_{n\to\infty} \int f_n h \, \mathrm{d}\mu$$

(2.c) $N = \{x \in E : h(x) = 1\}$ is the founded N such that $\mu(N) = 0$, $v_s(N^c) = v(N^c \cap N) = 0$, so $v_s \perp \mu$.

(2.d) $\forall A \in \mathcal{A}$, if $\mu(A) = 0$, since $g = \mathbf{1}_{N^c} \frac{h}{1-h}$, then $v_a(A) = \int_A \mathrm{d}v = \int_A g \, \mathrm{d}\mu = \frac{h}{1-h}\mu(A \cap N^c) \le \frac{h}{1-h}\mu(A) = 0$. So $v_a \ll \mu$.

3. (Remark). Perfect counterexample. A counting measure μ on $([0,1], \mathcal{B}([0,1]))$ is not σ-finite because $[0,1]$ is uncountable. We need to find a sequence $\{A_n\}_n$ that $[0,1] = \bigcup_{n\ge1} A_n, \mu(A_n) < \infty$ to make μ σ-finite. But a countable union of finite sets is at most countable. So μ cannot be σ-finite.

4. (Conclusion). Given a measure μ, consider a new measure v. Lebesgue decomposition says $v = v_a + v_s$. Radon-Nikodym theorem tells us there $\exists! g$, such that $v_a = g \cdot \mu$. In the proof, we also determined what exactly is v_a, v_s. So the process would be, first find a set N where $\mu(N) = 0$. Then

$$v_s = \mathbf{1}_N \cdot v, \quad v_a = \mathbf{1}_{N^c} \frac{\mathrm{d}v}{\mathrm{d}\mu} \cdot \mu$$

p. 80 Example (2). Consider measurable space $([0,1], \mathcal{F}_n)$, denote f_n as the Radon-Nikodym derivative of v with respect to λ. $\mathcal{F}_n = \sigma\left(I_i^{(n)}; i \in \{1, 2, \cdots, 2^n\}\right)$ where

$$I_i^{(n)} := \left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)$$

Claim: $f_n(x)$ is constant on the atoms of $\mathcal{F}_n$.

Proof. Recall Def 1.8, consider $((0,1], \mathcal{F}_n), (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, let $f_n : (0,1] \to \mathbb{R}$, f_n is measurable, so

$$\forall A \in \mathcal{B}(\mathbb{R}), f_n^{-1}(A) \in \mathcal{F}_n$$

If $\exists a, b \in I_i^{(n)}, f_n(a) \ne f_n(b)$, then $f^{-1}(B_\epsilon(a)) \subset I_i^{(n)}$ and $f^{-1}(B_\epsilon(a)) \notin \mathcal{F}_n$. $\square$

For $x \in I_i^{(n)}$, $f_n(x) = c_i$. Since $v = f_n \cdot \lambda$ by Radon-Nikodym theorem,

$$v(I_i^{(n)}) = \int_{I_i^{(n)}} f_n(x)\lambda(\mathrm{d}x) = c_i \lambda(I_i^{(n)}) \implies c_i = \frac{v(I_i^{(n)})}{2^{-n}}$$

since the Lebesgue measure $\lambda(I_i^{(n)}) = 2^{-n}$. So we conclude that

$$f_n(x) = \sum_{i=1}^{2^n} \mathbf{1}_{I_i^{(n)}} \frac{v(I_i^{(n)})}{2^{-n}}(x)$$

5 Product Measures

`Prop 5.1` (i) The goal is to show $\mathcal{C} = \mathcal{A} \otimes \mathcal{B}$. So $\mathcal{C}$ should contain all measurable rectangles like $\mathcal{A} \otimes \mathcal{B}$, and should be a σ-field.

1. $\forall A \in \mathcal{A}, B \in \mathcal{B}$, let $C = A \times B \in \mathcal{C}$, then

$$C_x = \{y \in F : (x, y) \in A \times B\} = \begin{cases} B & \text{if } x \in A \\ \varnothing & \text{if } x \notin A \end{cases}$$

2. $\mathcal{C}$ is a σ-field.

   (a) Let $C = E \times F, x \in E, C_x = F \in \mathcal{B}$, so $E \times F \in \mathcal{C}$

   (b) If $C \in \mathcal{C}$, then $C_x \in \mathcal{B}$.

$$(C^c)_x = \{y \in F : (x, y) \in (A \times B)^c\} = (C_x)^c \in \mathcal{B}$$

   since $\mathcal{B}$ is a σ-field and is closed under complements.

   (c) If $C_n \in \mathcal{C}$, then $(C_n)_x \in \mathcal{B}$.

$$\left(\bigcup_{n \geq 1} C_n\right)_x = \{y \in F : (x, y) \in \bigcup_{n \geq 1} C_n\} = \bigcup_{n \geq 1} \{y \in F : (x, y) \in C_n\} = \bigcup_{n \geq 1} (C_n)_x \in \mathcal{B}$$

   since $\mathcal{B}$ is a σ-field.

(ii) The text already derived $f_x^{-1}(D) = (f^{-1}(D))_x$. Because $D \in \mathcal{G}$ and f is measurable, $f^{-1}(D) \in \mathcal{A} \otimes \mathcal{B}$ by Def 1.8. By (i), $(f^{-1}(D))_x \in \mathcal{B}$, so $f_x^{-1}(D) \in \mathcal{B}$, which means f_x is $\mathcal{B}$-measurable by Def 1.8. (ii) is the main result of this proposition, it tells us if $f : E \times F \to G$, then $f_x(y), f^y(x)$ are measurable on the corresponding σ-field of y and x.

`Thm 5.2` A trick to show uniqueness is using σ-finite together with Cor 1.19. Another trick to show $x \mapsto v(C_x)$ is $\mathcal{A}$-measurable for all sets is constructing a class

$$\mathcal{G} = \{C \in \mathcal{A} \otimes \mathcal{B} : x \mapsto \nu(C_x) \text{ is } \mathcal{A}\text{-measurable}\}$$

and show $\mathcal{G}$ equals the original σ-field $\mathcal{A} \otimes \mathcal{B}$. The structure is always: first suppose finite measures, then extend to σ-finite by restricting the measures on partitions of underlying set E.
In this proof, we used Cor 1.19 for uniqueness. As for existence, assume v is finite, define

$$m(C) := \int_E v(C_x) \mu(\mathrm{d}x).$$

Consider the measurable spaces $(E, \mathcal{A}, \mu)$ and $(F, \mathcal{B}, v)$. Verify $v(C_x)$ is valid, that is $C_x \in \mathcal{B}$ (Prop 5.1). Then verify $x \mapsto \nu(C_x)$ is $\mathcal{A}$-measurable, so we can legally take integral (Thm 1.18). Now release v from finite to σ-finite. Then check $m(C)$ is actually a measure (Def 1.6). Because

$$v(C_x) = \begin{cases} v(B) & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}, \implies v(C_x) = \mathbf{1}_A(x) v(B)$$

the property in (i) holds,

$$m(C) = v(B) \int_E \mathbf{1}_A(x) \mu(\mathrm{d}x) = v(B) \mu(A)$$

`Thm 5.3` We had a hard time proving Thm 5.2, which says

$$\mu \otimes v(C) = \int_E v(C_x) \mu(\mathrm{d}x) = \int_F \mu(C^y) v(\mathrm{d}y).$$

This is a special case of Thm 5.3 with $f = \mathbf{1}_C$,

$$\int_{E\times F} f \,\mathrm{d}\mu \otimes v = \int_E \left(\int_F f(x,y) v(\mathrm{d}y) \right) \mu(\mathrm{d}x) = \int_F \left(\int_E f(x,y) \mu(\mathrm{d}x) \right) v(\mathrm{d}y)$$

The extension from Thm 5.2 to 5.3 follows the route: indicator function $\to$ simple function $\to$ approximation (Prop 2.5(i)). Thm 5.3 applies only to nonnegative f, Thm 5.4 removed this condition.

$\boxed{\texttt{p. 95}}$ Write $\varphi(s,t) = \mathbf{1}_{\{s\leq t\}} f(t) g(s)$, which is on space $([a,b], \mathcal{B}([a,b]), \lambda) \times ([a,b], \mathcal{B}([a,b]), \lambda)$. Consider $h(s,t) := |f(t)g(s)|$ is nonnegative and measurable, apply Thm 5.3,

$$\begin{aligned}\int_{[a,b]^2} h(s,t)\,\mathrm{d}(\lambda\otimes\lambda) &= \int_a^b \left(\int_a^b |f(t)||g(s)|\lambda(\mathrm{d}s) \right) \lambda(\mathrm{d}t) \\ &= \int_a^b |f(t)| \left(\int_a^b |g(s)| \lambda(\mathrm{d}s) \right) \lambda(\mathrm{d}t) = \left(\int_a^b |g(s)|\,\mathrm{d}s \right) \left(\int_a^b |f(t)|\,\mathrm{d}t \right)\end{aligned}$$

Notation $\lambda(\mathrm{d}x) \iff \mathrm{d}x$. Since f, g are integrable, $\varphi(s,t) \in \mathcal{L}^1(\lambda\otimes\lambda)$. So

$$\int_a^b \left(\int_a^b \varphi(s,t)\mathrm{d}s \right) \mathrm{d}t = \int_a^b \left(\int_a^b \varphi(s,t)\mathrm{d}t \right) \mathrm{d}s$$

$\boxed{\texttt{p. 100}}$ Derive $I_n = n/(n+1) \cdot I_{n-2}$. By symmetry, $I_n = 2\int_0^1 (1-x^2)^{n/2}\,\mathrm{d}x$. Let $x = \sin\theta, \mathrm{d}x = \cos\theta\,\mathrm{d}\theta$, $x \in [0,1] \implies \theta \in [0, \pi/2]$.

$$I_n = 2\int_0^1 (1-x^2)^{n/2}\,\mathrm{d}x = 2\int_0^{\pi/2} \cos^n\theta \cdot \cos\theta \;\mathrm{d}\theta = 2\int_0^{\pi/2} \cos^{n+1}\theta \;\mathrm{d}\theta$$

Integrate by part, let $u = \cos^n\theta, \mathrm{d}v = \cos\theta\,\mathrm{d}\theta$. This result in

$$\int_0^{\pi/2} \cos^{n+1}\theta \;\mathrm{d}\theta = n\int_0^{\pi/2} (1-\cos^2\theta)\cos^{n-1}\theta \;\mathrm{d}\theta \implies I_n = \frac{n}{n+1} I_{n-2}$$

7 Change of Variables

Prop 7.1

(1) Given $A \in \mathcal{B}(\mathbb{R}^d)$, it is not guaranteed that $f(A) \in \mathcal{B}(\mathbb{R}^d)$, even though f is continuous. It will always be Lebesgue measurable, but not necessarily Borel measurable.
Claim: All continuous functions are measurable.

Proof. Consider $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. $f : \mathbb{R}^d \to \mathbb{R}^d$ is continuous $\iff$ $\forall$ open set $U \subset \mathbb{R}^d$, $f^{-1}(U)$ is open in $\mathbb{R}^d$ (see Lemma 30 in MIT18.S190 Lec2). So if f is continuous, then $\forall U \in \mathcal{B}(\mathbb{R}^d), f^{-1}(U) \in \mathcal{B}(\mathbb{R}^d)$. This is exactly the definition of f being Boral measurable, Def 1.8. $\square$

Since M is invertible, $g := f^{-1}$ is also continuous. Then $g^{-1}(A)$ is continuous, thus measurable. So $f(A) = (f^{-1})^{-1}(A)$ is measurable.
(2) For special cases P is orthonormal matrix and S is symmetric positive definite, we have $1 = |\det(P)|, c = |\det(S)|$.

Lemma (Polar Decomposition)**.** *Any invertible matrix M can be decomposed as $M = PS$ where P is an orthonormal matrix and S is symmetric positive definite (spd).*

Proof. The intuition is rotate (P) and stretch out (S). If M is invertible, then $M^T M$ is spd for sure. Let $S^2 = M^T M$. Set $P = MS^{-1}$, now show that P is orthonormal,

$$P^T P = (MS^{-1})^T MS^{-1} = (S^{-1})^T M^T MS^{-1} = S^{-1}S^2S^{-1} = I$$

So $M = PS$. $\square$

Since M is invertible, apply polar decomposition, $M = PS$ and $|\det(M)| = |\det(P)| \cdot |\det(S)| = c$.

Thm 7.2

Let $f = \mathbf{1}_A$, then $f(\varphi(u)) = \mathbf{1}_A(\varphi(u)) = 1$ only if $\varphi(u) \in A, u \in \varphi^{-1}(A)$, otherwise $f(\varphi(u)) = 0$. So $\int_U f(\varphi(u))|J_\varphi(u)|\,\mathrm{d}u = \int_{\varphi^{-1}(A)} |J_\varphi(u)|\,\mathrm{d}u$. Note that $A \subset D, B := \varphi^{-1}(A) \subset U, \varphi(B) = A$. Replace A by $\varphi(B)$,

$$\lambda_d(\varphi(B)) = \int_{\varphi^{-1}(\varphi(B))} |J_\varphi(u)|\,\mathrm{d}u = \int_B |J_\varphi(u)|\,\mathrm{d}u$$

The remaining proof is way too complicated, so skip it. Being able to use polar coordinates technique would be enough.

p. 128

Recall equation (5.6),

$$\gamma_{2k} = \frac{\pi^k}{k!}, \quad \gamma_{2k+1} = \frac{\pi^k}{(k+\frac{1}{2})(k-\frac{1}{2})\cdots\frac{3}{2}\frac{1}{2}}$$

and gamma function

$$\Gamma(n) = (n-1)!, \quad \Gamma(\frac{1}{2}+n) = (n-\frac{1}{2})(n-\frac{3}{2})\cdots\frac{3}{2}\frac{1}{2}\sqrt{\pi}$$

Consider $d = 2k, k \in \mathbb{N}$ and $d = 2k+1, k \in \mathbb{N}$ respectively, If $d = 2k, k \in \mathbb{N}$, then

$$\gamma_{2k} = \frac{\pi^k}{\Gamma(k+1)} = \frac{\pi^{d/2}}{\Gamma(d/2+1)}, \quad \gamma_{2k+1} = \frac{\pi^k \cdot \sqrt{\pi}}{\Gamma(\frac{1}{2}+k+1)} = \frac{\pi^{d/2}}{\Gamma(d/2+1)}$$

It generates a unified form of volume for high-dimensional balls.

Part II
Probability Theory

8 Foundations of Probability Theory

`Def 8.2` ω is hidden in formulas of probability theory.

$$\mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega \in \Omega : \omega \in X^{-1}(B)\}) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}) =: \mathbb{P}(X \in B)$$

`Prop 8.4` Review Fubini theorem, the only condition is $f \in \mathcal{L}^1(\mu \otimes v)$ and μ, v are σ-finite. In this case, X is defined on $(\Omega, \mathcal{A}, \mathbb{P})$.

$$\{X \geq x\} = X^{-1}([x, +\infty))$$

X is a measurable function, $[x, +\infty) \subset \mathbb{R}$ is a Borel set, therefore $\{x \leq X\} \in \mathcal{A}$ is a measurable subset in Ω. $\mathbf{1}_{\{x \leq X\}} : \Omega \to \{0, 1\}$, check $\forall B \in \mathcal{P}(\{0, 1\})$, so $B \in \mathcal{B}(\mathbb{R})$,

$$\mathbf{1}^{-1}_{\{x \leq X\}}(B) = \{\omega \in \Omega : \mathbf{1}_{\{x \leq X\}}(\omega) \in B\} \in \mathcal{A}$$

so $\mathbf{1}_{\{x \leq X\}}$ is a Lebesgue measurable function. $\mathbf{1}_{\{x \leq X\}} \in \mathcal{L}^1(\mathbb{P} \otimes \lambda)$, Fubini theorem is applicable.

`Prop 8.5` Recall Prop 2.9,

$$\int_E h(\varphi(x))\mu(\mathrm{d}x) = \int_F h(y)v(\mathrm{d}y)$$

here $\mu = \mathbb{P}, v = \mathbb{P}_X, \varphi = X$,

$$\mathbb{E}[f(X)] = \int_\Omega f(X(\omega))\mathbb{P}(\mathrm{d}\omega) = \int_E f(x)\mathbb{P}_X(\mathrm{d}x)$$

`p. 142` In the discussion after Prop 8.5, the text

> Since the indicator function of an open rectangle is the increasing limit of a sequence of continuous functions with compact support, the probability measure μ is also determined by the values of $\int \varphi(x)\mu(\mathrm{d}x)$ when φ varies in the space $C_c\left(\mathbb{R}^d\right)$ of all continuous functions with compact support from $\mathbb{R}^d$ into $\mathbb{R}$.

means any indicator function of a rectangle $\mathbf{1}_R$ can be approximated from below by a sequence $\varphi_n \in C_c(\mathbb{R}^d)$. See a construction of Lemma 2.2 in MIT18.102. So

$$\mathbf{1}_R(x) = \lim_{n\to\infty} \varphi_n(x) \quad \text{with each } \varphi_n \in C_c(\mathbb{R}^d),\ \varphi_n \leq \mathbf{1}_R$$

By the *monotone convergence theorem*,

$$\mu(R) = \int \mathbf{1}_R(x)\,\mu(dx) = \lim_{n\to\infty} \int \varphi_n(x)\,\mu(dx)$$

`Prop 8.9` $\mathbf{1}_{A_i} = \mathbf{1}_{B_i} \circ X$.

$$1_{A_i}(\omega) = \begin{cases} 1 & \text{if } \omega \in A_i, \\ 0 & \text{otherwise,} \end{cases} = \begin{cases} 1 & \text{if } X(\omega) \in B_i, \\ 0 & \text{otherwise,} \end{cases} = 1_{B_i}(X(\omega)) = (1_{B_i} \circ X)(\omega).$$

`p. 153` Show Markov inequality. Let $A_a = \{\omega \in \Omega : X(\omega) \geq a\}, X \geq a\mathbf{1}_{A_a}$.

$$\int X\,\mathrm{d}\mathbb{P} \geq a \int \mathbf{1}_{A_a}\,\mathrm{d}\mathbb{P} = a\mathbb{P}(A_a), \implies \mathbb{P}(A_a) \leq \frac{1}{a}\int \mathbf{1}_{A_a}\,\mathrm{d}\mathbb{P} = \frac{1}{a}\mathbb{E}[X]$$

`Prop 8.13` Given $\alpha_0 = \mathbb{E}[X]$ in equation (8.5), then

$$Z = \mathbb{E}[X] + \sum_{j=1}^{n} \alpha_j (Y_j - \mathbb{E}[Y_j])$$

To let equation (8.6) holds, plug in Z,

$$\mathbb{E}[(X - \mathbb{E}[X])(Y_k - \mathbb{E}[Y_k])] - \sum_{j=1}^{n} \alpha_j \mathbb{E}[(Y_j - \mathbb{E}[Y_j])(Y_k - \mathbb{E}[Y_k])] = 0$$

This is exactly $\mathrm{cov}(X, Y_k) = \sum_{j=1}^{n} \alpha_j \mathrm{cov}(Y_j, Y_k)$.

`Def 8.14` A probability measure is absolutely continuous with respect to Lebesgue measure if and only if it has a density function $\varphi(x)$. Then applies Radon-Nikodym theorem, $\mathbb{P}_X(\mathrm{d}x) = \varphi(x)\lambda(\mathrm{d}x)$. Recall *Fourier Transform* in p. 32,

$$\hat{\varphi}(\xi) = \int e^{i\xi x} \varphi(x) \lambda(\mathrm{d}x) = \int e^{i\xi x} \mathbb{P}_X(\mathrm{d}x) =: \Phi(\xi)$$

`Lem 8.15` (1) Parity argument: $e^{i\xi X} = \cos(\xi X) + i \sin(\xi X)$.
(2) To use Thm 2.13, let $h(u, x) = e^{-x^2/2} \cos(ux)$, find a function $g(x)$ such that

$$|h(u, x) - h(\xi, x)| \leq g(x) \cdot |u - \xi|$$

By mean value theorem, $\exists c \in (u, \xi)$ (or (ξ, u)), such that $|\cos(ux) - \cos(\xi x)| = |-x \sin(cx)||u - \xi|$.

$$|e^{-x^2/2}(\cos(ux) - \cos(\xi x))| \leq e^{-x^2/2} |\cos(ux) - \cos(\xi x)| \leq e^{-x^2/2} |x| \cdot |u - \xi| = g(x) \cdot |u - \xi|$$

where $g(x) = |x| \cdot e^{-x^2/2} \in \mathcal{L}^1_+$. So condition (ii) satisfies.
(3) If $Z \sim \mathcal{N}(0, 1)$, then $\Phi_Z(\xi) = \exp(-\xi^2/2)$. For $X = m + \sigma Z \sim \mathcal{N}(m, \sigma^2)$,

$$\Phi_X(\xi) = \mathbb{E}[\exp(i\xi X)] = \mathbb{E}[\exp(i\xi(m + \sigma Z))] = e^{i\xi m} \cdot \Phi_Z(\sigma\xi) = e^{i\xi m} \cdot e^{-\sigma^2 \xi^2/2}$$

`Thm 8.16` (i) To establish (1), use Lem 8.15 to find that $X \sim \mathcal{N}(0, 1/\sigma)$, and use Def 8.14 to write

$$\exp(-\frac{x^2}{2\sigma^2}) = \Phi_X(x) = \int_{\mathbb{R}} e^{ix \cdot \xi} g_{1/\sigma}(\xi) \, \mathrm{d}\xi$$

9 Independence

p. 170 Given X_j is in $(E_j, \mathcal{E}_j)$. The following expressions are equivalent:

1. $\mathbb{P}(A_j)$: A_j is an element in $\sigma(X_j)$
2. $\mathbb{P}(\{X_j \in F_j\})$: F_j is an element in $\mathcal{E}_j$

In (1), $\sigma(X_j) = \{X_j^{-1}(F) : F \in \mathcal{E}_j\}$. An element in $\sigma(X_j)$ means for some $F_j \in \mathcal{E}_j$, $A_i = \{X_j^{-1}(F_j)\} = \{\omega \in \Omega : X_j(\omega) \in F_j\}$. This is exactly $\{X_j \in F_j\}$ in (2).