Notes of Probability and Stochastics

Xie Zejian Zhang Songxin

2021-01-29

Contents

0.1 Notations

$\mathbb{R}$	$(-\infty, \infty)$
$\overline{\mathbb{R}}$	$[-\infty, \infty]$
$\mathbb{R}_+$	$[0, \infty)$
$\overline{A}$	Closure of set A
A°	Interior of set A
$(x_n) \subset A$	A sequence taking value in A
2^A	The power set of A
$\mathcal{A}$	A collection of subsets in A, i.e., $\mathcal{A} \subset 2^A$
$\ker \mathcal{A}$	$\bigcap_{A \in \mathcal{A}} A$
$x_n \nearrow x$	(x_n) is increasing and converges to x.
$\sigma(\mathcal{A})$	σ-algebra generated by $\mathcal{A}$.
$\mathcal{A}_+$	Nonnegative function in $\mathcal{A}$
$\mu \ll \nu$	μ is absolutely continuous w.r.t. ν.
s.t.	such that
w.r.t.	with respect to
r.v.	random variable

Chapter 1

Measure and integrations

1.1 Measurable space

1.1.1 σ algebra

Definition 1.1. A nonempty system of subset of Ω is an algebra on Ω if it's

1. Closed under complement: $A^c \in \mathcal{A} \iff A \in \mathcal{A}$
2. Closed under finite union: $\cup_i A_i \in \mathcal{A} \impliedby \forall i, A_i \in \mathcal{A}$

it's an σ algebra on Ω if it's also closed under countable union.

Remark. $\mathcal{A}$ is an algebra auto implies $\varnothing \in \mathcal{A}$ and $\Omega \in \mathcal{A}$. So $\{\varnothing, \Omega\}$ is the minimum algebra on Ω and thus minimum σ algebra while the discrete algebra 2^Ω is maximum.

Let $\left\{\mathcal{A}_\gamma : \gamma \in \Gamma\right\}$ is a collection of σ algebra, then we have $\mathcal{A} = \cap_{\gamma\in\Gamma}\mathcal{A}_\gamma$ is also a σ algebra. Hence

Definition 1.2. The smallest σ algebra as intersection of all σ algebras contains $\mathcal{A}$, that called the σ algebra **generated** by $\mathcal{A}$ and denoted by $\sigma(\mathcal{A})$.

The smallest $\sigma-$algebra generated by the system of all open sets in a topological space (Ω, τ) is called **Borel σ algebra** on Ω and denoted by $\mathcal{B}(\Omega)$, its elements are called **Borel sets**.

1.1.2 π, λ, m systems

Definition 1.3. A collection of subsets $\mathcal{A}$ is called.

- **m-system** if closed under monotone series, that is if $(A_n) \subset \mathcal{A}$ and $A_n \nearrow A$, then $A \in \mathcal{A}$.
- $\pi-$**system** is closed under finite intersection
- $\lambda-$ **system** if
  1. $\Omega \in \mathcal{A}$
  2. closed under proper difference:
  $$A - B \in \mathcal{A} \impliedby B \subset A \text{ and } B, A \in \mathcal{A}$$
  3. is an m-system.

Theorem 1.1. *Let $\mathcal{A}$ be a collection of subsets of Ω iff it's both a π system and λ system.*

Proof. For $\Rightarrow$, check:

1. $\Omega \in \mathcal{A}$
2. $A - B = A \cap B^c \in \mathcal{A}$
3. is an m-system

For the converse:

1. $A^c = \Omega - A \in \mathcal{A}$
2. $A \cup B = (A^c \cap B^c)^c \in \mathcal{A}$
3. hence $\mathcal{A}$ is an algebra and $\mathcal{A}$ is a m-system.

Similarly, for m, π, λ -system, they also has a minimum system generated by some collection $\mathcal{C}$.

□

Lemma 1.1. *Let $\mathcal{A}$ be an algebra, then*

1. $m(\mathcal{A}) = \sigma(\mathcal{A})$
2. *if $\mathcal{B}$ is an m class and $\mathcal{A} \subset \mathcal{B}$, then $\sigma(\mathcal{A}) = m(\mathcal{A}) \subset \mathcal{B}$*

Similarly, let $\mathcal{A}$ be a π class, then $\lambda(\mathcal{A}) = \sigma(\mathcal{A})$

Then we have **Monotone class theorem**:

Theorem 1.2. $\forall \mathcal{A} \subset \mathcal{B} \subset \mathcal{P}(\Omega), s.t.:$

1. *If $\mathcal{A}$ is a π-class, $\mathcal{B}$ is a $\lambda-$class, then $\sigma(\mathcal{A}) \subset \mathcal{B}$*
2. *If $\mathcal{A}$ is an algebra, $\mathcal{B}$ is a $m-$class, then $\sigma(\mathcal{A}) \subset \mathcal{B}$*

1.1.3 Graphical illustration of different classes

$$\begin{array}{ccccc}
\sigma \text{ algebra} & \Leftrightarrow & \lambda \text{ system} & + & \pi \text{ system} \\
\Downarrow & & \Downarrow & & \\
\text{algebra} & & m \text{ system} & & \\
\Downarrow & & & & \\
\text{semialgebra} & & & & \\
\Downarrow & & & & \\
\pi \text{ system} & & & &
\end{array}$$

1.1.4 Measurable spaces

Definition 1.4 (Measurable Space)**.** Pair $(\Omega, \mathcal{A})$ where $\mathcal{A}$ is a σ-Algebra on Ω.

Definition 1.5 (Products of measurable spaces)**.** Let $(E, \mathcal{E}), (F, \mathcal{F})$ be two measurable spaces. For $A \subset E, B \subset F$, $A \times B$ is the set of all pairs (x, y) : $x \in A, y \in B$. Note that $\mathcal{E} \times \mathcal{F}$ is also a σ-Algebra with all $A \times B$ where $A \in \mathcal{E}, B \in \mathcal{F}$ which is called *the product σ-Algebra*.

1.2 Measurable function

1.2.1 Mappings

Let $f : \Omega_1 \to \Omega_2$ be a mapping, $\forall B \subset \Omega_2$ and $\mathcal{G} \subset \mathcal{P}(\Omega_2)$, the **inverse image** of

- B is $f^{-1}(B) = \{\omega : \omega \in \Omega_1, f(\omega) \in B\} := \{f \in B\}$
- $\mathcal{G}$ is $f^{-1}(\mathcal{G}) = \{f^{-1}(B) : B \in \mathcal{G}\}$

There is some properties:

1. $f^{-1}(\Omega_2) = \Omega_1, f^{-1}(\varnothing) = \varnothing$
2. $f^{-1}(B^c) = \left[f^{-1}(B)\right]^c$
3. $$\begin{array}{l} f^{-1}\left(\cup_{\gamma\in\Gamma} B_\gamma\right) = \cup_{\gamma\in\Gamma} f^{-1}\left(B_\gamma\right) \text{ for } B_\gamma \subset \Omega_2, \gamma \in \Gamma \\ f^{-1}\left(\cap_{\gamma\in\Gamma} B_\gamma\right) = \cap_{\gamma\in\Gamma} f^{-1}\left(B_\gamma\right) \text{ for } B_\gamma \subset \Omega_2, \gamma \in \Gamma \end{array}$$

   where Γ may not countable.
4. $f^{-1}(B_1 - B_2) = f^{-1}(B_1) - f^{-1}(B_2)\, \forall B_1, B_2 \subset \Omega_2$
5. $B_1 \subset B_2 \subset \Omega_2 \implies f^{-1}(B_1) \subset f^{-1}(B_2)$

6. If $\mathcal{B}$ is a $\sigma-$algebra, $f^{-1}(\mathcal{B})$ is also a $\sigma-$algebra. It's easy to check $f^{-1}(\mathcal{B})$ is closed under complement and countable union. (From properties 2 and 3)
7. If $\mathcal{C}$ is nonempty, $f^{-1}(\sigma(\mathcal{C})) = \sigma\left(f^{-1}(\mathcal{C})\right)$

Remarks f^{-1} preserves all the set operations on Ω.

1.2.2 Measurable functions

Definition 1.6. For two measurable spaces $(\Omega_1, \mathcal{A})$, $(\Omega_1, \mathcal{B})$, $f : \Omega_1 \to \Omega_2$ is a **measurable mapping** if $f^{-1}(\mathcal{B}) \subset \mathcal{A}$, where

$$f^{-1}(\mathcal{B}) = \{f^{-1}(B) : B \in \mathcal{B}\}$$

It is a **measurable function** if $(\Omega_2, \mathcal{B}) = (\mathcal{R}^n, \mathcal{B}(\mathcal{R}^n))$, moreover, a **Borel function** if $(\Omega_1, \mathcal{A}) = (\mathcal{R}^m, \mathcal{B}(\mathcal{R}^m))$

Remark. If $\mathcal{B} = \sigma(\mathcal{C})$, the definition can be reduced to $f^{-1}(\mathcal{C}) \subset \mathcal{A}$ since

$$f^{-1}(\mathcal{B}) = f^{-1}(\sigma(\mathcal{C})) = \sigma\left(f^{-1}(\mathcal{C})\right) \subset \sigma(\mathcal{A}) = \mathcal{A}$$

Lemma 1.2. *Suppose $f : \mathcal{E} \to \mathcal{F}$ and $g : \mathcal{F} \to \mathcal{G}$ are measurable, then so is $f \circ g$.*

Proof. The same as how we proved composition of continuous function is continuous.

□

1.3 Random Variable

A r.v. X is a measurable function from $(\Omega_1, \mathcal{A})$ to $(\mathcal{R}, \mathcal{B})$. It denoted by X is $\mathcal{A}$-measurable or $X \in \mathcal{A}$

(Another definition): A r.v. X is a measurable mapping from $(\Omega, \mathcal{A}, P)$ to $(\mathcal{R}, \mathcal{B})$ such that

$$P(|X| = \infty) = P(\{\omega : |X(\omega)| = \infty\}) = 0$$

Lemma 1.3. *X is a r.v. from $(\Omega, \mathcal{A})$ to $(\mathbb{R}, \mathcal{B})$*

$$\begin{aligned} &\iff X \le x = X^{-1}([-\infty], x) \in \mathcal{A} \quad \forall x \in \mathbb{R} \\ &\iff X \le x = X^{-1}([-\infty], x) \in \mathcal{A} \quad \forall x \in D \end{aligned}$$

where D is a dense subset of $\mathbb{R}$, e.g. $\mathbb{Q}$. $\{X \le x\}$ above can be replaced by

$$\{X \le x\}, \quad \{X \ge x\}, \quad \{X < x\}, \quad \{X > x\}, \quad \{x < X < y\}$$

1.3.1 Construction of random variables

Lemma 1.4. $\mathbf{X} = (X_1, \cdots, X_n)$ *is a random vectors if* X_k *is a r.v.* $\forall k$ *iff* $\mathbf{X}$ *is a measurable function from* $(\Omega, \mathcal{A})$ *to* $(\mathcal{R}^n, \mathcal{B}(\mathcal{R}^n))$.

Proof. Note that

$$\{\mathbf{X} \in \prod I_n\} = \bigcap \{X_n \in I_n\} \in \mathcal{A}$$

where $I_k = (a_k, b_k], -\infty \leq a_k \leq b_k \leq \infty$ and follows from $\sigma(\{\prod I_n\}) = \mathcal{B}(\mathcal{R}^n)$. For the other direction, note

$$\{X_k \leq t\} = \{\mathbf{X} \in \prod_{i<k} \mathbb{R} \times \{-\infty, t\} \times \prod_{i>k} \mathbb{R}\} \in \mathcal{A}$$

$\square$

Recall lemma 1.2 we have:

Theorem 1.3. $\forall$ *random* n *vectors* $X = (X_{1:n})$ *and Borel function* f *from* $\mathcal{R}^n \to \mathcal{R}^m$, *then* $f(X)$ *is a random* m *vectors.*

Remark. Note that continuous function are borel measurable since continuity leads to inverse image of an open set is still open. So if $X_{1:n}$ are r.v.'s, so are $\sum X_n$, $\sin(x)$, e^X, $\text{Poly}(X), \cdots$. That implies:

$\forall X, Y \in \mathcal{A}$, so are $aX + bY, X \vee Y = \max\{X, Y\}, X \wedge Y = \min\{X, Y\}, X^2, XY, X/Y, X^+ = \max(x, 0), X^- = -\min(x, 0), |X| = X^+ + X^-$

1.3.2 Limiting opts

Let (X_n) are r.v. on $(\Omega, \mathcal{A})$, then $\sup_{n\to\infty} X_n, \inf_{n\to\infty} X_n, \limsup_{n\to\infty} X_n, \liminf_{n\to\infty} X_n$ are r.v.'s. Moreover, if it exists, $\lim_{n\to\infty} X_n$ is r.v..

Proof. First two follows from, $\forall t \in \mathbb{R}$:

$$\{\sup_{n\to\infty} X_n \leq t\} = \bigcap_{n=1}^{\infty} \{X_n \leq t\} \in \mathcal{A}$$

$$\{\inf_{n\to\infty} X_n \geq t\} = \bigcap_{n=1}^{\infty} \{X_n \geq t\} \in \mathcal{A}$$

and the last two follows from $\limsup_{n\to\inf} = \inf_{k\to\infty} \sup_{m\geq k} X_m$ and $\liminf_{n\to\inf} = \sup_{k\to\infty} \inf_{m\geq k} X_m$.

$\square$

That implies

Lemma 1.5. *If $S = \sum_1^\infty X_n$ exists everywhere, then S is a r.v.*

Proof. Note $\sum_1^\infty X = \lim_{n\to\infty} \sum_n X_n$ is a r.v. □

If $X = \lim_{n\to\infty} X_n$ holds **almost** everywhere, i.e., define Ω_0 is the set of all ω, such that $\lim_n X_n(\omega)$ exists, then $P(\Omega_0) = 1$, we say that X_n converges a.s. and write:

$$X_n \to X \quad a.s.$$

For a measurable function f, we may modify it at a null set into f' and it remain measurable since for any open set G, $f'^{-1}(G)$ differ $f^{-1}(G)$ a at most null set, by the completion of Lebesgue measure space, $f'-1(G)$ is measurable and thus f^{-1} measurable. Hence, for $f_n \to f$ a.s., we may ignore a null set and then $f_n \to f$ everywhere and thus f measurable.

1.3.3 Approximations of r.v. by simple r.v.'s

Definition 1.7. If $A \in \mathcal{A}$, the indicator function $\mathbf{1}_A$ is a r.v. If $\Omega = \sum_1^n A_i$, where $A_i \in \mathcal{A}$, then $X = \sum_1^n a_i \mathbf{1}_{A_i}$ is a r.v. and called **simple r.v.**

Any r.v. can be approximated by simple ones:

Theorem 1.4. *$\forall X \in \mathcal{A}$, $\exists 0 \le X_1 \le X_2 \cdots X_n$ s.t. $X_n(\omega) \nearrow X(\omega)$ everywhere.*

Proof. Suppose

$$X_n(\omega) = \sup\{\frac{j}{2^n} : j \in \mathbb{Z}, \frac{j}{2^n} \le \min(X(\omega), 2^n)\}$$

One can check X_n is simple r.v. and $X_n(\omega) \nearrow X(\omega)$ for all $\omega \in \Omega$. □

1.3.4 σ algebra generated by r.v.

Let $\{X_\lambda, \lambda \in \Lambda\}$ is r.v.s on $(\Omega, \mathcal{A})$. Define

$$\sigma\left(X_\lambda, \lambda \in \Lambda\right) := \sigma\left(X_\lambda \in B, B \in \mathcal{B}, \lambda \in \Lambda\right) = \sigma\left(X_\lambda^{-1}(\mathcal{B}), \lambda \in \Lambda\right) = \sigma\left(\cup_{\lambda\in\Lambda} X_\lambda^{-1}(\mathcal{B})\right)$$

which is called σ algebra generated by $\{X_\lambda, \lambda \in \Lambda\}$, where Λ is a index set which can be uncountable.

For $\Lambda = \mathbb{N}^+$:

1.
$$\sigma\left(X_i\right) = \sigma\left(X_i^{-1}(\mathcal{B})\right) = X_i^{-1}(\mathcal{B}) = \{X_i \in \mathcal{B}\}$$
$$\sigma\left(X_1, \ldots, X_n\right) = \sigma\left(\cup_{i=1}^{n} X_i^{-1}(\mathcal{B})\right) = \sigma\left(\cup_{i=1}^{n} \sigma\left(X_i\right)\right)$$

2.
$$\sigma\left(X_1\right) \subset \sigma\left(X_1, X_2\right) \subset \ldots\ldots \subset \sigma\left(X_1, \ldots, X_n\right)$$
$$\sigma\left(X_1, X_2, \ldots..\right) \supset \sigma\left(X_2, X_3, \ldots..\right) \supset \ldots\ldots \supset \sigma\left(X_n, X_{n+1}, \ldots..\right)$$

3. $\bigcap_1^{\infty} \sigma(X_n, X_{n+1}, \cdots)$ is the tail σ algebra of $X_{1:}$

If $A_{1:n}$ are not mutually exclusive to each other, then we have

$$|\sigma(A_{1:n})| = 2^{2^n}$$

Which follows from for a partition $A_{1:n}$,

$$\sigma(A_1, \cdots, A_n) = \{\bigcup_{i \in J} A_i\}$$

where J is any subset of $\mathbb{N} \leq n$ and $A_0 = \varnothing$. Hence for discrete r.v. Y, $\sigma(Y)$ can be generated from $A_i = \{Y = y_i\}$ forall y_i's. For continuous case, it's generated by all intervals.

1.3.5 Monotone classes of function

Definition 1.8 (monotone class)**.** $\mathcal{M}$ is called a monotone class if: - $1 \in \mathcal{M}$ - $f, g \in \mathcal{M}_b$ and $a, b \in \mathbb{R} \implies af + bg \in \mathcal{M}$ - $(f_n) \subset M_+, f_n \uparrow f \implies f \in \mathcal{M}$

where $\mathcal{M}_+$ is a subcollection consisting of positive functions in $\mathcal{M}$, and $\mathcal{M}_b$ for the bounded function in $\mathcal{M}$.

Theorem 1.5 (Monotone class theorem for functions)**.** *Let $\mathcal{M}$ be a monotone class of functions on E. Suppose for some p-system $\mathcal{C}$ generating $\mathcal{E}$ and $1_A \in \mathcal{M}$ for every $A \in \mathcal{C}$. Then $\mathcal{M}$ includes all positive $\mathcal{E}$-measurable functions and all bounded $\mathcal{E}$-measurable functions.*

Proof. First we need to show that $1_A \in \mathcal{M}$ for every $A \in \mathcal{E}$. Let $\mathcal{D} = \{A \in \mathcal{E} : 1_A \in \mathcal{M}\}$. Now we check that $\mathcal{D}$ is a d-system: - $1_E = 1$, so $E \in \mathcal{D}$. - $B \subset A,\ A, B \in \mathcal{D}$. $1_{A-B} = 1_A - 1_B \in \mathcal{D}$ - $(A_n) \subset \mathcal{D}$ and $A_n \uparrow A$, then $1_{A_n} \uparrow 1_A$, so $1_A \in \mathcal{M}$, then $A \in \mathcal{D}$

By assumption, $\mathcal{C} \subset \mathcal{D}$, and $\sigma(\mathcal{C})$ is the smallest d-system by the theorem above, so $\mathcal{E} \subset \mathcal{D}$, so $1_A \in \mathcal{M}$ for every $A \in \mathcal{E}$.
As $1_A \in \mathcal{M}$ for every $A \in \mathcal{E}$, we can easily prove that all of the positive simple function is generated by the linear combination of 1_A s. And all positive $\mathcal{E}$-measurable functions is generated by a sequence of positive simple functions.

Then for general bounded $\mathcal{E}$-measurable function f, using $f = f^+ - f^-$ where $f^+, f^- \in \mathcal{M}$.

□

Definition 1.9. Let $(E, \mathcal{E}), (F, \mathcal{F})$ be two measurable spaces and f is a bijection $E \to F$. Then f is said to be a isomorphism of $(E, \mathcal{E})$ and $(F, \mathcal{F})$ if f is $\mathcal{E}$-measurable and f^{-1} is $\mathcal{F}$-measurable. These two spaces are called isomorphic if there exists an isomorphisms between them.

Definition 1.10. A measurable space $(\Omega, \mathcal{A})$ is said to be *standard* if it there exist an embedding $f : (\Omega, \mathcal{A}) \hookrightarrow (\mathbb{R}, \mathcal{B})$.

Remark. Clearly, $([0,1], \mathcal{B}([0,1]))$, $(\mathbb{N} \leq n, 2^{N \leq n})$ and $(\mathbb{N}, 2^{\mathbb{N}})$ are all standard. In fact, every standard measurable space is isomorphic to one of them.

1.4 Measure

Let Ω be a space and $\mathcal{A}$ a class, then function $\mu : \mathcal{A} \to R = [-\infty, \infty]$ is a **set function**.

It's

- 1. **finite** if $\forall A \in \mathcal{A}, \quad |\mu(A)| < \infty$
  2. σ**-finite** if $\exists A_n \subset \mathcal{A}, \quad s.t. \quad \cup_{i=1}^{\infty} A_i = \Omega \quad \forall n \quad |\mu(A_n)| < \infty$
  3. **s finite** if there exist countable finite (μ_n) s.t. $\mu = \sum_n \mu_n$.
- 1. **additive** $\iff \mu\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n \mu\left(A_i\right)$
  2. $\sigma-$**additive**$\iff \mu\left(\sum_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \mu\left(A_i\right)$

Remark. Finite implies σ finite and σ finite implies Σ finite.

μ is a **measure** on $\mathcal{A}$ if

1. $\forall A \in \mathcal{A} : \mu(A) \geq 0$
2. It's σ additive.

the triplet $(\Omega, \mathcal{A}, \mu)$ is a **measure space** when μ is a measure and $(\Omega, \mathcal{A})$ is a measurable space. Whose sets are called **measurable sets** or $\mathcal{A}$**-measurable**.A measure space is a **probability space** if $P(\Omega) = 1$.

Assume that $A_{1:n} \in \mathcal{A}$ and $A \in \mathcal{A}$ and μ is a measure.

1. μ is continues from above, if $A_n \searrow A \implies \mu(A_n) \to \mu(A)$
2. μ is continues from below, if $A_n \nearrow A \implies \mu(A_n) \to \mu(A)$
3. μ is continues at A, if $A_n \to A \implies \mu(A_n) \to \mu(A)$

$\forall$ Measure μ is continues from below and may not continues from above. It will be continues from above if $\exists m < \infty, \mu(A_m) < \infty$.So finite measure μ are always continues.

1.4.1 Properties of measure

1.4.1.1 Semialgebras

Let μ be a nonnegative additive set function on a semialgebra $\mathcal{A}$. $\forall A, B \in \mathcal{A}$ and $\{A_n, B_n, n \geq 1\} \in \mathcal{A}$

1. (**Monotonicity**): $A \subset B \implies \mu(A) \leq \mu(B)$
2. (**σ-subadditivity**):
    1. $\sum_1^\infty A_n \subset A, \quad \implies \quad \sum_1^\infty \mu(A_n) \leq \mu(A)$
    2. Moreover, if μ is a measure, then

$$B \subset \sum_{n=1}^{\infty} B_n \implies \mu(B) \leq \sum_{n=1}^{\infty} \mu(B_n)$$

We can assert a nonnegative set function μ is a measure by:

1. μ is additive
2. μ is σ subadditive on $\mathcal{S}$

1.4.1.2 Algebras

Algebras enjoy all the properties of semialgebra, also, we have

Theorem 1.6 (σ subadditivity)**.** *Let μ be a measure on an algebra $\mathcal{A}$*

$$A \subset \cup_1^\infty A_n \implies \mu(A) \leq \sum_1^\infty \mu(A_n)$$

Proof. Note $A = A \cap (\cup A_n) = \cup(A \cap A_n)$, hence we can write A as union in $\mathcal{A}$ by take $B_n = A \cap A_n \in \mathcal{A}$.

$$A = \cup_1^\infty B_n$$

and then we can take $C_n = B_n - \cup_1^{n-1} B_i \in \mathcal{A}$ to write A as disjoint union:

$$A = \sum C_n$$

Then

$$\mu(\mathcal{A}) = \mu(\sum C_n) = \sum \mu(C_n) \leq \sum \mu(B_n) \leq \sum \mu(A_n)$$

as $C_n \subset B_n \subset A_n$.

$\square$

1.4.1.3 σ algebras

Let μ be a measure on an σ algebra $\mathcal{A}$

1. Monotonicity
2. Boole's inequality(Countable Sub-Additivity)

$$\mu\left(\cup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu\left(A_i\right)$$

3. Continuity from below
4. Continuity from above if μ is finite in A_i.

The sense of **4** follows from suppose $A_i \searrow A$, then $A_1 - A_i \nearrow A_1 - A$, then

$$\mu(A_1) - \mu(A) = \mu(A_1 - A) = \lim \mu(A_1 - A_i) = \mu(A_1) - \lim(A_i)$$

where $\mu(A_1)$ cannot be cancelled if $\mu(A_i) = \infty$.

Definition 1.11. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, and $N \subset \Omega$

1. N is a μ **null set** iff $\exists B \in \mathcal{A} \quad s.t. \quad \mu(B) = 0, \quad N \subset B$
2. This measure space is a **complete measure** space if $\forall$ μ null space N, $N \in \mathcal{A}$

Theorem 1.7. *Given any measure space* $(\Omega, \mathcal{A}, \mu)$ *, there exist a complete measure space* $(\Omega, \bar{\mathcal{A}}, \bar{\mu})$*, such that* $\mathcal{A} \subset \bar{\mathcal{A}}$ *and* $\bar{\mu}$ *is an extension of* μ*. This space is called completion of* $(\Omega, \mathcal{A}, \mu)$*.*

Proof. Take

$$\bar{\mathcal{A}} = \{A \cup N : A \in \mathcal{A}\}$$
$$\bar{\mathcal{B}} = \{A \Delta N : A \in \mathcal{A}\}$$

$\bar{\mathcal{A}} = \bar{\mathcal{B}}$ since $A \cup N = (A - B)\Delta(B \cap (A \cup N))$ and $A \Delta N = (A - B) \cup (B \cap (A \Delta N))$.

Then we can show that $\bar{\mathcal{A}}$ is a σ algebra. Let $\Omega_i = A_i \cup N_i \in \bar{\mathcal{A}}$, then

$$\bigcup_1^{\infty} \Omega_i = \bigcup_1^{\infty} A_i \cup \bigcup_1^{\infty} N_i$$

and note $\bigcup_1^{\infty} A_i \in \mathcal{A}$ and $\mu(\bigcup_1^{\infty} N_i) \leq \mu(\cup_1^{\infty} B_i) \leq \cup_1^{\infty} \mu(B_i) = 0$. Thus $\bar{\mathcal{A}}$ is closed by countable union. As for complements, note $\Omega^c = A^c \cap N^c = (A^c \cap N^c \cap B^c) \cup (A^c \cap N^c \cap B) = (A^c \cap B^c) \cup (A^c \cap N^c \cap B) \in \bar{\mathcal{A}}$.

Finally we define a measure $\bar{\mu}$ on $\bar{\mathcal{A}}$ by

$$\bar{\mu}(A \cup N) = \mu(A)$$

We should prove it's well defined. Suppose $A_1 \cup N_1 = A_2 \cup N_2 \in \bar{\mathcal{A}}$, note $A \Delta B \Delta C = A \Delta (B \Delta C)$ and $A \Delta B = B \Delta A$.

$$\begin{aligned}(A_1 \Delta A_2) \Delta (N_1 \Delta N_2) &= (A_1 \Delta A_2 \Delta N_1) \Delta N_2 \\ &= (A_1 \Delta N_1) \Delta (A_2 \Delta N_2) \\ &= \varnothing\end{aligned}$$

Hence $A_1 \Delta A_2 = N_1 \Delta N_2$, note $N_1 \Delta N_2 \subset N_1 \cup N_2 \subset B_1 \cup B_2$, hence $\mu(A_1 \Delta A_2) = 0$ and thus $\mu(A_1 - A_2) = \mu(A_2 - A_1) = 0$. Therefore

$$\mu(A_1) = \mu(A_1 - A_2) + \mu(A_1 \cap A_2) = \mu(A_1 \cap A_2)$$
$$\mu(A_2) = \mu(A_2 - A_1) + \mu(A_1 \cap A_2) = \mu(A_1 \cap A_2)$$

$\bar{\mu}$ is do well defined. μ^* is auto σ additive since so is μ and is easy to check that all μ^* null set is μ null set.

□

1.4.2 Specification of measures

Theorem 1.8. *Let $(\Omega, \mathcal{A})$ be a measurable space. Let μ, ν be measures on it with $\mu(\Omega) = \nu(\Omega) < \infty$. If μ, ν agree on a π system generating $\mathcal{A}$, then μ, ν are identical. >*

Proof. Let $\mathcal{C}$ be the π system generating $\mathcal{A}$ and $\mu(A) = \nu(A)$ for every $A \in \mathcal{C}$. Consider $\mathcal{D} = \{A \in \mathcal{A} : \mu(A) = \nu(A)\}$ which satisfies $\mathcal{C} \subset \mathcal{D} \subset \Omega$. Then we need to prove that $\mathcal{D}$ is a λ system:

- $\Omega \in \mathcal{D}$ by the assumption.
- Let $A, B \in \mathcal{D}$ and $B \subset A$. Then $\mu(A-B) = \mu(A) - \mu(B) = \nu(A) - \nu(B) = \nu(A - B)$, so $A - B \in \mathcal{D}$
- Let $(A_n) \uparrow A$ and $(A_n) \subset \mathcal{D}$, then $\mu(A_n) \uparrow \mu(A)$, $\nu(A_n) \uparrow \nu(A)$, since $\mu(A_n) = \nu(A_n)$ for every n, so $\mu(A) = \nu(A)$.

So $\mathcal{D}$ is a d-system. It follows that $\sigma(\mathcal{C}) = \mathcal{A} \subset \mathcal{D}$.

□

As consequence, we have

Corollary 1.1. *Suppose μ and ν are probability measures on space on $(\overline{\mathbb{R}}, \mathcal{B})$ then $\mu = \nu$ iff $\mu[-\infty, r] = \nu[-\infty, r], \forall r \in \mathbb{R}$.*

Proof. Note $\{[-\infty, r] : r \in \mathbb{R}\}$ is a π system and generates $\mathcal{B}$.

□

1.4.3 Atomic and diffuse measure

Definition 1.12. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space where $\mathcal{A}$ contains all the singletons: $\{x\} \in \mathcal{A}$ for every $x \in \Omega$ (it's true for all the standard measure).

A point x is said to be an **atom** if $\mu(\{x\}) > 0$, the measure is said to be **diffuse** if it has no atoms. It is said to be **purely atomic** if the set D of its atoms is countable and $\mu(\Omega - D) = 0$.

Lemma 1.6. *A s-finite measure has at most countable many atoms.*

Proof. It suffices to show that when μ is finite. Suppose $A_n = \{x : \mu\{x\} > \frac{1}{n}\}$ and A consists all atoms, then the claim follows from $A_n \nearrow A$ and $|A_n| \leq n\mu(\Omega)$ as $A = \bigcup_n A_n$.

□

Theorem 1.9. *Let μ be a s-finite measure on $(\Omega, \mathcal{A})$. Then $\mu = \nu + \lambda$ where λ is a diffuse measure and ν is purely atomic.*

Proof. Let D be set of all atoms and define

$$\lambda(A) = \mu(A - D)$$
$$\nu(A) = \mu(A \cap D)$$

for all $A \in \mathcal{A}$. Clearly, $\lambda + \nu = \mu$. Then

- λ is diffuse as $\lambda\{x\} = 0$ for all $x \in D$ and if $\lambda\{x\} > 0$, it must be $x \in D$.
- ν is purely atomic as $D_\nu = D$ clearly and $\nu(\Omega - D) = \mu(\varnothing) = 0$.

□

1.5 Integration

Let f be Borel measurable on $(\Omega, \mathcal{A}, \mu)$. The **integral** of f w.r.t μ is denoted by

$$\int f(\omega)\mu(d\omega) = \int f d\mu = \int f$$

1. If $f = \sum_1^n a_i \mathbf{1}_{A_i}$ with $a_i \geq 0$,

$$\int f d\mu = \sum_1^n a_i \mu\left(A_i\right)$$

2. If $f \geq 0$,define

$$\int f d\mu = \lim_n \int f_n d\mu$$

   where f_n are simple functions and $f_n \nearrow f$.

3. For any f, we have $f = f^+ - f^-$, define

$$\int f d\mu := \int f^+ d\mu - \int f^- d\mu$$

4. f is said to be integrable w.r.t. μ if $\int |f| d\mu < \infty$. We denote all integrable functions by L^1.

Proposition 1.1. ***(Integral over sets)***

$$\int_A f d\mu = \int f \mathbf{1}_A d\mu = \int f(\omega) \mathbf{1}_A(\omega) \mu(d\omega)$$

*(**Absolute integrability**). $\int f$ is finite iff $\int |f|$ is finite.*

(Linearity) *If $f, g, a, b \geq 0$ or $f, g \in L^1$*

$$\int (af + bg) = a \int f + b \int g$$

(σ additivity over sets) If $A = \sum_{i=1}^\infty A_i$, then

$$\int_A f = \sum_{i=1}^\infty \int_{A_i} f$$

(Positivity) *If $f \geq 0$ a.s., then $\int f \geq 0$*

(Monotonicity) *If $f_1 \leq f \leq f_2$ a.s., then $\int f_1 \leq \int f \leq \int f_2$*

(Mean value theorem) *If $a \leq f \leq b$ a.s., then*

$$a\mu(A) \leq \int_A f \leq b\mu(A)$$

***(Modulus inequality)**: $|\int f| \leq \int |f|$*

1.5.1 Monotone Convergence Theorem

Theorem 1.10 (Monotone Convergence Theorem). *Suppose nonnegative $f_n \nearrow f$ a.e., then $\int f_n d\mu \nearrow \int f d\mu$.*

Theorem 1.11. *We may ignore a null set then $f_n \nearrow f$ and their integration still equal. Note $\int f_n d\mu \leq \int f d\mu$, $\int f_n d\mu$ must converges to some $L \leq \int f$. Then we show $L \geq \int f$.*

Let $s = \sum a_i \chi_{E_i}$ be simple function and $s \leq f$. Let $A_n = \{x : f_n(x) \geq cs(x)\}$ where $c \in (0, 1)$, then $A_n \nearrow X$. For each n

$$\begin{aligned}\int f_n \geq \int_{A_n} f_n \geq c \int_{A_n} s \\ = c \int_{A_n} \sum a_i \chi_{E_i} \\ = c \sum a_i \mu(E_i \cap A_n) \\ \nearrow c \int s\end{aligned}$$

hence $L \geq c \int s \implies L \geq \int s \implies L = \lim L \geq \lim \int s_n = \int f$.

Lemma 1.7 (Fatou's Lemma). *If $f_n \geq 0$ a.e. then*

$$\int \left(\liminf_n f_n\right) \leq \liminf_n \int f_n$$

Proof. Suppose $g_n = \inf_{i \geq n} f_i$ and recall that $\lim g_n = \liminf f_n$. Clearly $g_n \leq f_i \forall i \geq n$ hence

$$\int g_n \leq \inf_{i \geq n} \int f_i$$

Take limit both side and note g_n is increasing:

$$\lim \int g_n = \int \lim g_n = \int \liminf f_n \leq \liminf \int f_n$$

□

Theorem 1.12 (Dominated Convergence Theorem). *Suppose $f_n(x) \to f(x) \forall x$, and there exists a nonnegative integrable g s.t. $|f_n(x)| \leq g(x)$(then we get $f_n \in L^1$ immediately), then*

$$\lim \int f_n d\mu = \int f d\mu$$

Proof. Since $f_n + g \geq 0$

$$\int f + \int g = \int f + g \leq \liminf \int f_n + g = \liminf \int f_n + \int g$$

thus $\int f \leq \liminf_{n\to\infty} \int f_n$. Similarly, we can get $\int f \leq \liminf_{n\to\infty} \int f_n$ from $g - f_n \geq 0$.

□

Theorem 1.13 (Tonelli's Throrem)**.** *If* $\sum_1^\infty \int |f_n| < \infty$*, then*

$$\int \left(\sum_{i=1}^{\infty} f_n \right) = \sum_{i=1}^{\infty} \int f_n$$

Proof. Let $g_k = \sum_1^k |f_n|, g = \sum_1^\infty |f_n|, h_k = \sum_1^k f_n, h = \sum_1^\infty f_n$. Then $g_k \nearrow g$, by MCT, we have

$$\int g = \lim \int g_k = \lim \sum_1^k \int |f_n| = \sum_1^\infty \int |f_n| < \infty$$

Hence we may let g dominate h_k and get

$$\int h = \lim \int h_k = \sum_1^\infty \int f_n$$

□

1.5.2 Criteria for zero a.e.

Lemma 1.8 (Markov inequality)**.** *Let* $A = \{x \in \Omega : f(x) \geq M\}$*, then*

$$\mu(A) \leq \frac{\int f}{M}$$

Proof.

$$\mu(A) = \int \chi_A = \int_A \chi_A \leq \int_A \frac{f}{M} \leq \int_X \frac{f}{M} = \frac{\int f}{M}$$

□

Suppose f is measurable and non-negative and $\int f d\mu = 0$. Then $f = 0$ a.e.

Proof. By lemma 1.8 and define $A_M = \{x \in \Omega : f(x) \geq M\}$. Consequently, $\mu(A_M) = 0$ for all $M > 0$, note $A_{\frac{1}{n}} \nearrow A_0$:

$$A_0 = \bigcup_1^\infty A_{\frac{1}{n}} \implies \mu(A_0) = \sum 0 = 0$$

Hence $f = 0$ a.e.

□

Lemma 1.9. *Suppose f is integrable and $\int_A f = 0$ for all measurable A. Then $f = 0$ a.e.*

Proof. Suppose $A_n = \{x \in \Omega : f(x) \geq \frac{1}{n}\}$, then

$$0 = \int_{A_n} f \geq \frac{\mu(A_n)}{n} \Rightarrow \mu(A_n) = 0$$

thus $\mu\{x \in \Omega : f(x) > 0\} = 0$. Similarly, $\mu\{x \in \Omega : f(x) < 0\} = 0$ and the claim follows.

□

Theorem 1.14. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is integrable and $\int_a^x f = 0$ for all x, then $f = 0$ a.e.*

Proof. For any interval $I = [c, d]$,

$$\int_i f = \int_a^d f - \int_a^c f = 0$$

Then the integral is 0 for finite disjoint union of intervals from additivity. Note open sets G can be written as countable union of disjoint open intervals $G = \sum_1^\infty I_i = \lim \sum I_n \implies$

$$\int_G f = \int f\chi_G = \int f \sum_1^\infty \chi_{I_i} = \int \lim f \sum \chi_{I_i} = \lim \int f \sum \chi_{I_i} = 0$$

If $G_n \searrow H$, then

$$\int_H f = \int f\chi_H = \int \lim f\chi_{G_n} = \lim \int f\chi_{G_n} = \lim \int_{G_n} f = 0$$

where we apply DMT twice and take dominated function $g = |f|$.

Finally, for any borel measurable set E, there is $G_\delta \supset E$ and $m(G_\delta - E) = 0$, then

$$\int_E f = \int f\chi_E = \int f\chi_{G_\delta} = \int_{G_\delta} f = 0$$

□

1.5.3 Characterization of the integral

Theorem 1.15. *Let $(\Omega, \mathcal{A})$ be a measurable space and $L : \mathcal{A} \to \overline{\mathbb{R}}_+$, then there is a unique measure μ on $(\Omega, \mathcal{A})$ s.t. $L(f) = \int f$ for every $f \in \mathcal{A}_+$ iff:*

- $f = 0 \implies L(f) = 0$
- $f, g \in \mathcal{A}_+$ *and* $a, b \in \mathbb{R}_+ \implies L(af + bg) = aL(f) + bL(g)$
- $(f_n) \subset \mathcal{A}_+$ *and* $f_n \nearrow f \implies L(f_n) \nearrow L(f)$

Proof. $\Rightarrow$ follows from the definition and properties of integral. For $\Leftarrow$, let there is a function L satisfies above conditions and give a μ and let $\mu(A) = L(1_A)$, then use those conditions we can prove that μ is a measure a $(\Omega, \mathcal{A})$.

□

1.6 Transforms and Indefinite integral

Definition 1.13 (Image measure)**.** Let $(F, \mathcal{F})$ and $(E, \mathcal{E})$ be measurable spaces. Let ν be a measure on $(F, \mathcal{F})$ and let $h : F \to E$ be measurable relative to $\mathcal{F}$ and $\mathcal{E}$, then define a mapping $\nu \circ h^{-1}(B) = \nu(h^{-1}B),\ B \in \mathcal{E}$. Then $\nu \circ h^{-1}$ is a measure on $(E, \mathcal{E})$, which is called the **image** of ν under h.

Remark. Image inherit finite and s-finite, but not σ-finite.

Theorem 1.16. *For every $f \in \mathcal{E}$, we have $(\nu \circ h^{-1})f = \nu(f \circ h)$.*

Proof. We only need to show that $\mathcal{E}_+$ case and then the general case follows easily.

Let $L : \mathcal{E}_+ \to \overline{\mathbb{R}}_+$ by letting $L(f) = \nu(f \circ h)$. Then as the property of $\nu(f \circ h)$, f satisfies the properties of integral characterization theorem. Then, $L(f) = \mu f$ for some unique measure μ on $(E, \mathcal{E})$. And note $\mu = \nu \circ h$

$$\mu(B) = L(\mathbf{1}_B) = \nu(\mathbf{1}_B \circ h) = \nu(h^{-1}B)$$

□

1.6.1 Images of the Lebesgue measure

By theorem 1.16, we have a convenient tool for creating new measure from old and, we may integral a measure using Lebesgue measure:

Theorem 1.17. *Let $(\Omega, \mathcal{A}, \mu)$ be a standard measure space where μ is s-finite and $b = \mu(\Omega)$. Then there exists a measurable mapping $h : ([0, b), \mathcal{B}_{[0,b]}) \to (\Omega, \mathcal{A})$ s.t. $\mu = \lambda \circ h^{-1}$, where λ is the Lebesgue measure on $[0, b)$.*

Proof. See 5.15 and 5.16 on page 34 in *Probability and Stochastic.*

□

1.6.2 Indefinite integrals

Definition 1.14. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $p \in \mathcal{A}_+$. Define:

$$\nu(A) = \mu(p\mathbf{1}_A) = \int_A p d\mu$$

then ν is a measure on $(\Omega, \mathcal{A})$ and called **indefinite integral** of p w.r.t. μ.

Remark. ν is a measure follows from MCT.

Theorem 1.18. *For any $f \in \mathcal{A}_+$, $\nu f = \mu(pf)$.*

Proof. Let $L(f) = \mu(pf)$. Check L :

- $f = 0 \implies L(f) = 0$
- Give $f, g \in \mathcal{E}_+$ and $a, b \in \mathbb{R}_+ \implies L(af + bg) = \mu(p(af + bg))$ and based on the arithmetic rules on $\mathbb{R}$ and the linearity of μ, $L(af + bg) = aL(f) + bL(g)$
- Give $(f_n) \subset \mathcal{E}_+$ and $f_n \nearrow f$, $L(f_n) = \mu(pf_n)$ and as $f_n \nearrow f$, $pf_n \nearrow pf$ so $\lim L(f_n) = \lim \mu(pf_n)$. According to the monotone converging theorem, $\lim \mu(pf_n) = \mu(pf) = L(f)$

So there exists $\hat{\mu}$ s.t. $\mu(pf) = \hat{\mu} f$ and that force $\hat{\mu} = \nu$ as

$$\hat{\mu}(A) = L(\mathbf{1}_A) = \mu(p\mathbf{1}_A) = \nu(A)$$

□

Remark. Writing above result in an explicit notation:

$$\int_E f d\nu = \int_E p f d\mu$$

that is:

$$d\nu = p d\mu$$

and it's precisely Fundamental theorem of calculus. Thus we can say:

- ν is the indefinite integral of p w.r.t. μ or
- p is the density of ν w.r.t. μ.

1.6.3 Radon-Nikodym theorem

Definition 1.15 (absolutely continuous)**.** Let ν and μ be measures on a measurable space $(\Omega, \mathcal{A})$. Then ν is said to be **absolutely continuous** w.r.t. μ if for any set $A \in \mathcal{E}$, $\mu(A) = 0 \implies \nu(A) = 0$ and denoted by $\nu \ll \mu$.

Clearly, if ν is the indefinite integral of some $p \in \mathcal{A}_+$ w.r.t. μ, then it's absolutely continuous w.r.t. μ. And the follows shows that the converse is true.

Theorem 1.19 (Radon-Nikodym Theorem)**.** *Suppose that μ is σ-finite and ν is absolutely continuous w.r.t. μ. Then there exists unique(up to a.e.) $p \in \mathcal{A}_+$ s.t. for every $f \in \mathcal{A}_+$:*

$$\int_\Omega f d\nu = \int_\Omega p f d\mu$$

1.7 Kernels and Product spaces

Definition 1.16 (transition kernel)**.** Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measurable spaces. Let $K : E \times \mathcal{F} \to \overline{\mathbb{R}}_+$. Then, K is called a transition kernel from $(E, \mathcal{E})$ into $(F, \mathcal{F})$ if:

- the mapping $x \mapsto K(x, B)$ is $\mathcal{E}$-measurable for every set $B \in \mathcal{F}$
- the mapping $B \mapsto K(x, B)$ is a measure on $(F, \mathcal{F})$ for every $x \in E$

Example 1.1. If ν is a finite measure on $(F, \mathcal{F})$, and k is a positive function on $E \times F$ that is $\mathcal{E} \times \mathcal{F}$-measurable, then

$$K(x, B) = \int_B k(x, y) d\nu$$

when fix $x \in E$, $K(x, B) = \nu(k(x, y)\mathbf{1}_B) = \mu(B)$ for some μ which is the measure on $(F, \mathcal{F})$.
when fix $B \in \mathcal{F}$, $K(x, B) = f(x) =$

Theorem 1.20. *Let K be a transition kernel from $(E, \mathcal{E})$ into $(F, \mathcal{F})$. Then*

$$Kf(x) = \int_F K(x, dy) f(y), \qquad x \in E,$$

defines a function $Kf \in \mathcal{E}_+$ for every $f \in \mathcal{F}_+$.

$$\mu K(B) = \int_E \mu(dx) K(x, B), \qquad B \in \mathcal{F},$$

defines a measure μK on $(F, \mathcal{F})$ for each measure μ on $(E, \mathcal{E})$. and

$$(\mu K)f = \mu(Kf) = \int_E \mu(dx) \int_F K(x, dy) f(y)$$

for every measure μ on $(E, \mathcal{E})$ and function f in $\mathcal{F}_+$.

Proof. □

Theorem 1.21. *Let $f \in \mathcal{E} \times \mathcal{F}$, then $x \mapsto f(x, y) \in \mathcal{E}$ for each $y \in F$ and $y \mapsto f(x, y) \in \mathcal{F}$ for each $x \in E$.*

Theorem 1.22. *For every $f \in \mathcal{E}_+$, we have $(\nu \circ h^{-1})f = \nu(f \circ h)$.*

Proof. Let $L : \mathcal{E}_+ \to \overline{\mathbb{R}}_+$ by letting $L(f) = \nu(f \circ h)$. Then as the property of $\nu(f \circ h)$, f satisfies the properties of integral characterization theorem. Then, $L(f) = \mu f$ for some unique measure μ on $(E, \mathcal{E})$. And

$$\mu(B) = L(1_B) = \nu(1_B \circ h) = \nu(h^{-1}B), \qquad B \in \mathcal{E}$$

□

Definition 1.17 (Standard measurable space)**.** Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measurable spaces let f be a bijection from E onto F, then f is called an isomorphism of $(E, \mathcal{E})$ and $(F, \mathcal{F})$ iff f is $\mathcal{E}$-measurable and f^{-1} is $\mathcal{F}$-measurable.
A measurable space $(E, \mathcal{E})$ is said to be a standard measurable space if it is isomorphic to $(F, \mathcal{B}_F)$ for some Borel set F of $\mathbb{R}$.

Definition 1.18. Let μ be a measure on $(E, \mathcal{E})$. It is said to be finite if $\mu(E) < \infty$, then $\mu(A) < \infty$ for every $A \subset E$.
It is called a probability measure if $\mu(E) = 1$.
It is said to be σ-finite if there exists a measurable partition (E_n) of E s.t. for each n, $\mu(E_n) < \infty$.
It is said to be Σ-finite if there exists a sequence of finite measures (μ_n) s.t. $\mu = \sum_n \mu_n$.

Theorem 1.23. *If a measure is σ-finite then it must be Σ-finite.*

Proof. Assume that there is a measure μ on $(E, \mathcal{E})$ is σ-finite, where there is a $(E_n) \subset \mathcal{E}$ s.t. $\bigcup_n E_n = E$. Then give (μ_n) s.t. for each n,

$$\begin{cases} \mu_n(E_n) = \mu(E_n) \\ \mu_n(E - E_n) = 0 \end{cases}$$

then $\mu_n(E) = \mu_n(E_n) = \mu(E_n)$, so a σ-finite measure μ is a Σ-fnite measure.

□

Let $(E, \mathcal{E}, \mu)$ be a measure space. Let p be a positive $\mathcal{E}$-measurable function. Define:

$$\nu(A) = \mu(p1_A) = \int_A \mu(dx)p(x), \qquad A \in \mathcal{E}$$

Theorem 1.24. *ν is a measure on $(E, \mathcal{E})$.*

Theorem 1.25. *Fix a $p \in \mathcal{E}_+$, and give a measure μ on $(E, \mathcal{E})$, then there must be a measure ν satisfying $\nu f = \mu(pf)$ for every $f \in \mathcal{E}_+$.*

Proof. Let $L(f) = \mu(pf)$. Check L : - $f = 0 \implies L(f) = 0$ - Give $f, g \in \mathcal{E}_+$ and $a, b \in \mathbb{R}_+ \implies L(af + bg) = \mu(p(af + bg))$ and based on the arithmetic rules on $\mathbb{R}$ and the linearity of μ, $L(af + bg) = aL(f) + bL(g)$ - Give $(f_n) \subset \mathcal{E}_+$ and $f_n \nearrow f$, $L(f_n) = \mu(pf_n)$ and as $f_n \nearrow f$, $pf_n \nearrow pf$ so $\lim L(f_n) = \lim \mu(pf_n)$. According to the monotone converging theorem, $\lim \mu(pf_n) = \mu(pf) = L(f)$

So there exists a ν satisfying $\nu f = L(f) = \mu(pf)$.

□

Written above result in an explicit notation:

$$\int_E \nu(dx)f(x) = \int_E \mu(dx)p(x)f(x)$$

then we can change the $\nu(dx)$ with $\mu(dx)p(x)$ where $x \in E$, $p \in \mathcal{E}_+$.

Definition 1.19 (absolutely continuous)**.** Let ν and μ be measures on a measurable space $(E, \mathcal{E})$. Then ν is said to be absolutely continuous respect to μ if for any set $A \in \mathcal{E}$, $\mu(A) = 0 \implies \nu(A) = 0$.

Theorem 1.26. *If ν satisfies the conditions $\nu f = \mu(pf)$ for every $f \in \mathcal{E}_+$, then ν is absolutely continuous with respect to μ.*

Proof. Let $\mu(A) = 0$, then we need to prove that $\mu(p\mathbf{1}_A) = 0$: - Let p is a simple function with a cononical form on E with $p = \sum_{i=1}^n a_i \mathbf{1}_{B_i}$ and as (B_i) is a partition, $(B_i \cap A)$ is a partition of A, noted as (A_i). Then

$$\mu(p\mathbf{1}_A) = \int_A pd\mu = \sum_{i=1}^n a_i \mu(A_i)$$

As $\mu(A) = 0$ and $A = \bigcup_{i=1}^n A_i$, $\mu(A) = \sum_{i=1}^n \mu(A_i)$, so $\mu(A_i) = 0$ for every i. So $\mu(p\mathbf{1}_A) = 0$. - Let $p \in \mathcal{E}_+$, give (p_n) s.t. $p_n \nearrow p$ where $p_n = d_n \circ p$. Then $p_n \mathbf{1}_A \nearrow p\mathbf{1}_A$ and for every n, $\mu(p_n \mathbf{1}_A)$ holds. So:

$$\mu(p\mathbf{1}_A) = \mu(\lim p_n \mathbf{1}_A) = \lim \mu(p_n \mathbf{1}_A) = 0$$

according to the monotone converging theorem, so ν is absolutely continuous respect to μ

□

Theorem 1.27 (Radon Nikodym Theorm)**.** *Suppose that μ is σ-finite and ν is absolutely continuous with repect to μ. Then there exists a positive $\mathcal{E}$-measurable function p s.t. $\nu f = \mu(pf)$ on E for every $f \in \mathcal{E}_+$.*
Moreover, p is unique above, if there is a $q \in \mathcal{E}_+$, then $q = p$ a.e.

Notice that Randon Nikodym theorem is the converse of the theorem before it. The proof of the Randon Nikodym theorem is not able to show here now.

Definition 1.20 (transition kernal)**.** Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measurable spaces. Let K be a mapping from $E \times \mathcal{F}$ into $\overline{\mathbb{R}}_+$. Then, K is called a transition kernal from $(E, \mathcal{E})$ into $(F, \mathcal{F})$ if: - the mapping $x \mapsto K(x, B)$ is $\mathcal{E}$-measurable for every set $B \in \mathcal{F}$ - the mapping $B \mapsto K(x, B)$ is a measure on $(F, \mathcal{F})$ for every $x \in E$

For example, if ν is a finite measure on $(F, \mathcal{F})$, and k is a positive function on $E \times F$ that is $\mathcal{E} \times \mathcal{F}$-measurable, then

$$K(x, B) = \int_B \nu(dy) k(x, y), \qquad x \in E, B \in \mathcal{F}$$

when fix $x \in E$, $K(x, B) = \nu(k(x, y)\mathbf{1}_B) = \mu(B)$ for some μ which is the measure on $(F, \mathcal{F})$.
when fix $B \in \mathcal{F}$, $K(x, B) = f(x) =$

Theorem 1.28. *Let K be a transition kernel from $(E, \mathcal{E})$ into $(F, \mathcal{F})$. Then*

$$Kf(x) = \int_F K(x, dy) f(y), \qquad x \in E,$$

defines a function $Kf \in \mathcal{E}_+$ for every $f \in \mathcal{F}_+$.

$$\mu K(B) = \int_E \mu(dx) K(x, B), \qquad B \in \mathcal{F},$$

defines a measure μK on $(F, \mathcal{F})$ for each measure μ on $(E, \mathcal{E})$. and

$$(\mu K) f = \mu(Kf) = \int_E \mu(dx) \int_F K(x, dy) f(y)$$

for every measure μ on $(E, \mathcal{E})$ and function f in $\mathcal{F}_+$.

Proof. $\square$

Theorem 1.29. *Let $f \in \mathcal{E} \times \mathcal{F}$, then $x \mapsto f(x, y) \in \mathcal{E}$ for each $y \in F$ and $y \mapsto f(x, y) \in \mathcal{F}$ for each $x \in E$.*