Math 625
Professor Erhan Bayraktar

Linfeng Li
llinfeng@umich.edu

September 11, 2014

No responding to emails!

1 Week 1

1.1 Tue: 2014-09-02

1.1.1 Basics

Exam (in class, 90 minutes): Tue, Oct 7; Tue, Nov 11; Tue, Dec 9. —— 25%, 30 % and 40 %. 5% attendance.

Textbook: Probability and Statistics (took the class from the writer, available through library). Probability with **Martingale**, the latter is the emphasis.

Exercises: Book has exercises, but not graded homework.

1.1.2 Beginning: measure theory

Sigma-algebra Given a set E (a universal set), $\mathcal{E}$, a nonempty collection of subsets of E, is called a σ-algebra if closed under complements & countable unions.

- The most trivial sigma-algebra $\{\emptyset, E\}$ is called the trivial σ-algebra.

Definition 1.1.1 (σ algebra generated by $\mathcal{C}$). *Given a collection $\mathcal{C}$ of subsets (of E), $\sigma(\mathcal{C})$ will denote the smaller σ-algebra containing $\mathcal{C}$.*

Definition 1.1.2. *σ-algebra generated by open sets is called **Borel σ-algebra**.*

p-system A collection of $\mathcal{C}$ which is closed under (finite) intersection.

"p" for product, could also use π-system. The latter in Greek.

d-system A collection $\mathcal{D}$ is called a d-system if

(i) $E \in \mathcal{D}$,

(ii) $A, B \in \mathcal{D}$, $A \supset B \implies A \setminus B \in \mathcal{D}$

(iii) $(A_n) \subset \mathcal{D}$ and $A_n \uparrow A \implies A \subset \mathcal{D}$.

("d" for Dynkin)

Note: curly characters are for collection of sets.

Proposition 1.1.3. *$\mathcal{E}$ is a σ-algebra if and only if it is a p-system and a d-system.*

Proof. $\Rightarrow$ is trivial.
$\Leftarrow$: Let $\mathcal{E}$ be a collection which is a p-system and a d-system.

1. Closed under complements (to be a sigma algebra). Let $A \in \mathcal{E} \implies E \setminus A \in \mathcal{E}$ by (ii) for property of d-system.
2. Closed under finite unions: $A, B \in \mathcal{E} \implies A \bigcup B = (A^c \cap B^c)^c$ by 1 above and property of p-system.
3. Closed under countable unions: for $(A_n) \subset \mathcal{E}$, $\bigcup_n A_n \in \mathcal{E}$? We construct an increasing sequence of (B_n):
   Let $B_1 = A_1$, $B_2 = A_1 \bigcup A_2 \in \mathcal{E} \ldots \bigcup_n A_n = \bigcup_n B_n$. Then by (iii) for property of d-system, the conclusion follows.

$\square$

Lemma 1.1.4. *For $\mathcal{D}$, a d-system, fix $D \in \mathcal{D}$. Define $\hat{\mathcal{D}} := \{A \in \mathcal{D} : A \cap D \in \mathcal{D}\}$. Then, $\hat{\mathcal{D}}$ is also a d-system.*

Monotone Class Theorem [Very useful tool in showing an arbitrary collection of sets is a σ-algebra]

Theorem 1.1.5. *If a d-system contains a p-system, then it also contains the σ-algebra generated by the p-system.*

Proof. Symbolic expression: $\mathcal{C} \subset \mathcal{D} \implies \sigma(\mathcal{C}) \subset \mathcal{D}$.
Step 1:
Let $\mathcal{C}$ be a p-system. $\mathcal{D}$ is the smallest d-system that contains $\mathcal{C}$. [1]
Enough to show $\mathcal{D} \supset \sigma(\mathcal{C})$.
If fact, we will show $\mathcal{D}$ is a σ-algebra. By proposition, it is enough to show it is a p-system.
Step 2:
Fix $B \in \mathcal{C}$ and let $\mathcal{D}_1 := \{A \in \mathcal{D} : A \bigcap B \in \mathcal{D}\}$. 1) By the lemma 1.1.4, $\mathcal{D}_1$ is a d-system. 2) $\mathcal{C} \subset \mathcal{D}_1$.
1) and 2) $\implies \mathcal{D}_1 = \mathcal{D}$.
Setp 3:
Fix $A \in \mathcal{D}$, let $\mathcal{D}_2 := \{B \in \mathcal{D} : B \bigcap A \in \mathcal{D}\}$.
1) by the lemma, $\mathcal{D}_2$ is a d-system. 2) by **step 2**, $\mathcal{C} \subset D_2$.
1) and 2) $\implies \mathcal{D}_2 = \mathcal{D}$.
Step 1-3 gives that $\mathcal{D}$ is a p-system.
In here, $\mathcal{D} = \sigma(\mathcal{C})$. [But in the theorem, this is not a necessary conclusion.] $\square$

Measurable space $(E, \mathcal{E})$ is a measurable space. [$\mathcal{E}$ is a σ-algebra on E.]

Products of measure spaces $(E, \mathcal{E})$, $(F, \mathcal{F})$. Then $(E \times F, \mathcal{E} light - product \mathcal{F})$ where $\times$ is regular set product; and the light-product is σ (generated by measurable rectangles)

Measurable functions (random variables)

Lemma 1.1.6. *A mapping $f : E \to F$ and (inverse mapping) $f^{-1}(A) := \{x \in E : f(x) \in A\}$. Then, $f^{-1}\emptyset = \emptyset$. $f^{-1}(F) = E$. $f^{-1}(B \setminus C) = f^{-1}(B) \setminus f^{-1}(C)$. $f^{-1}\left(\bigcup_i B_i\right) = \bigcup_i f^{-1}(B_i)$ and $f^{-1}\left(\bigcap B_i\right) = \bigcap_i f^{-1}(B_i)$.*
(set operation passes through the inverse function operation.)

[1](***: little result – the intersections of d-systems is a d-system [to obtain the "smallest"]. Also, the "smallest" matters.)

Definition 1.1.7. *$(E, \mathcal{E})$, $(F, \mathcal{F})$. $f : E \to F$ is "measurable" relative to $\mathcal{E}\&\mathcal{F}$ if $f^{-1}(B) \in \mathcal{E}$, $\forall B \in \mathcal{F}$.*

1.1.3 Thu: 2014-09-04

measurable functions (To "measure" a measurable function: just to integrate it).

Proposition 1.1.8. *A function $f : E \to F$ is meausrbale if and only if for some collection $\mathcal{F}_0$ with $\mathcal{F} = \sigma(\mathcal{F}_0)$, $f^{-1}(B) \in \mathcal{E}$.*

Proof. Necessicty is trivial; (by definitoin)

First collect all the sets s.t. $\mathcal{F}_1 = \left\{B \in \mathcal{F} : f^{-1}(B) \in \mathcal{E}\right\} \supset \mathcal{F}$. We show this by showing that this is a sigma algebra. [through checking the properties of inverse fucntions.] □

Lemma 1.1.9 (Composition of measurable functions are measurbal)**.** *[2.5 Proposition]*

Let $m(\mathcal{E})$ note the collection of measurable functions. Abuse of notation: let $\mathcal{E}$ also note $m(\mathcal{E})$ since the context would be clear.

Theorem 1.1.10. *Proof.* **Step 1** the sup would exist:

...

□

Approximation of measurbale functions

Lemma 1.1.11. *For $r \in \mathbb{R}_+$, $d_n(r) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} 1_{[\frac{k-1}{2^n}, \frac{k}{2^n}]}(r) + n 1_{[n,\infty)}(r)$. THen $d_n(r) \to r$ as $n \to \infty$.*

Theorem 1.1.12. *A positive function is measurable if and only if it is a limit of positive simple functions ($\sum_{i=1}^{n} a_i 1_{A_i}$ for $a_i \in \mathbb{R}$ and $A_i \in \mathcal{E}$).*

Proof. Sufficiency is given by the previous theorem.

Necessity: Let $f_n = d_n \circ f$ where $f_n \uparrow f$. (By construction of $d_n(r)$, f_n is simple measurable function.) □

i

Lemma 1.1.13. *If $h_1, h_2, h_3 \in \mathcal{E}$, $h_1 + h_2, h_1 \dot{h}_2, \lambda h \in \mathcal{E}$ for $\lambda \in \mathbb{R}$.*

Decomposition of positive part and negative part of function f Let $f = f^+ - f^-$ where $f^+ = f \ldots$

Monotone Class Theorem for Functions

Definition 1.1.14. *For $\mathcal{M}$, a collection of functions is a monotone class if*

1. *$1 \in \mathcal{M}$ (1 is the function assigning all elements in E to 1)*
2. *$f, g \in \mathcal{M}_b \implies af + bg \in \mathcal{M}_b$. where $\mathcal{M}_b$ denote bounded functions in the set of functions denoted by $\mathcal{M}$.*
3. *$(f_n) \subset \mathcal{M}_+$, and $f_n \uparrow f \implies f \in \mathcal{M}_+$. where $\mathcal{M}_+$ denote non-negative functions in $\mathcal{M}$.*

Theorem 1.1.15 (Monotone Convergence Theorem)**.** *Let $\mathcal{M}$ be a monotone class. Suppose that for some p-system $\mathcal{C}$, $\mathcal{E} = \sigma(\mathcal{C})$.*

$1_A \in \mathcal{M}, \quad \forall A \in \mathcal{C} \implies \mathcal{M}$ includes all positive measurable functions ($\mathcal{E}_+$) and all bounded measurable functions (E

1_A here is an indicator function

Proof. **Step 1:** we want to show that, for all $1_A \in \mathcal{M}$, $\forall A \in \mathcal{E}$.

Define

use the defection of a monotone class to show that $\mathcal{D}$ is a d-system.

$\mathcal{D}$ being a d-system implies that ...

Step 2 Simple functions are also in $\mathcal{M}$. [find a reason to this.]

Step 3 By the previous theorem on the simple function, we see that for arbitrary $f \in \mathcal{E}_+$, $\exists\,(f_n) \uparrow f$ where f_n is simple measurable functions.

Then, by (3) in definition of monotone class (of functions) , $f \in \mathcal{M}_+$.

Step 4 For $f \in \mathcal{E}_b$, as $f = f^+ - f^-$ by (2) in definition of monotone class, we have $f \in \mathcal{M}_b$. □

Definition 1.1.16. *$X : \Omega \to (E, \mathcal{E})$, $\sigma(X) = X^{-1}\mathcal{E} := \left\{X^{-1}(A) : A \in \mathcal{E}\right\}$ is called the σ-algebra generated by X. Note that X here is a*

Hereby we define a new σ-algebra on Ω.

Proposition 1.1.17. *Let $X : \Omega \to (E, \mathcal{E})$ and another mapping $V : \Omega \to \bar{\mathbb{R}}$ belongs to $\sigma(X)$ if and only if $V = f \circ X$ for some function $f \in \mathcal{E}$.*

Proof. Sufficiency part is trivial;

Necessity: let $\mathcal{M}$ be the collection of all $V f \circ X$ for some $f \in \mathcal{E}$. (This is enough to show if Y is bounded measurable w.r.t. $\sigma(X)$, $\exists$ a bounded measurable function f s.t. $Y = f(X)$)

Step 1 Show that $\mathcal{M}$ is a monotone class:

Step 2 $\mathcal{M}$ includes every indicator function in $\sigma(X)$. The set $H \in \sigma(X)$, $H = X^{-1}(A)$ for $A \in \mathcal{E}$. Then

$$1_H = 1_A \circ X$$

Step 3 Use MCT

□

Measures

Definition 1.1.18. *In a measurable space $(E, \mathcal{E})$, $\mathcal{E}$, for $\mu : \mathcal{E} \to \bar{\mathbb{R}}_+$, if*

(a) $\mu(\emptyset) = 0$

(b) $\mu(\bigcup_n A_n) = \sum_n \mu(A_n)$ *for A_n's being disjoint.*

μ is called a measure. (Note that this measure is infinite as $\mu : \mathcal{E} \to \bar{\mathbb{R}}_+$, with a bar overhead of $\mathbb{R}$.)

Proposition 1.1.19. *For A and B being measurable sets,*

(i) (Monotonicity) $A \subset B \implies \mu(A) \le \mu(B)$; [Implied by finite additivity.]

(ii) (Continuity under increasing limits) $A_n \uparrow A \to \mu(A_n) \uparrow \mu(A)$.

Proof. For $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, ..., $\bigcup_{n=1}^n B_n = \bigcup_{n=1}^n A_n$ [This is finite additivity.] [Could finish the proof by taking limits at both sides.] □

(iii) (Sub-additivity) $\mu(\bigcup_n A_n) \le \sum(A_n)$.

Notation $\mu(E) \leq \infty \implies \mathcal{M}$ is a finite measure. $\mu(E) = 1$ implies that μ is a probability measure.
σ-finite if $\exists$ a partition[2] (E_n) of E s.t. $\sum(E_n) < \infty$.
Σ-finite if $\exists \mu_n$ s.t. $\mu = \sum_n \mu_n$ for $\mu_n(E) < \infty$.
σ-finite $\implies$ Σ-finite.

Theorem 1.1.20. *Let $(E, \mathcal{E})$ be a measurable space and measures μ and ν are $\mu(E) = \nu(E) < \infty$. Moreover, μ and ν agree on $\mathcal{C}$, which is a p-system satisfying $\mathcal{E} = \sigma(\mathcal{C})$, then*

$$\mu = \nu$$

This is why we can specify the Lebesgue measure by only assigning measure to the intervals. The above theorem would generalize the measure.

Again is a consequence of monotone class theorem

2 Week 2

The main goal is to cover mathematical finance topics. But far away from being practical. Be discouraged if you are taking the course for this purpose.

2.1 Tue: quick review of measure theory

(Suggested) Homework: (not to be collected) Suggested homework: 1.9, 1.10, 1.11, 1.12, 1.13, 1.17;
Suggested homework: 2.21, 2.29, 2.31, 2.32;
Suggested homework: 3.11(b), 3.12;

2.1.1 Specification of measures: only about uniqueness.

Theorem 2.1.1 (Specificaion of measures)**.** *For measure: $(E, \mathcal{E})$, let μ and ν be measures with $\mu(E) = \nu(E) < \infty$. If μ and ν agree on $\mathcal{C}$, a p-sysem, $\mathcal{E} = \sigma(\mathcal{C})$,*

$$\implies \nu = \mu \text{ on } \mathcal{E}$$

[Uniqueness is clear. Existence is anothe story.]

Proof. Let $\mathcal{C}$ be the p-system s.t. $\mu(A) = \nu(A)$ for $A \in \mathcal{C}$, let

$$\mathcal{D} = \{A \in \mathcal{E} : \mu(A) = \nu(A)\} =?\mathcal{E}$$

(We want to use monotone class theorem) Knowing that $\mathcal{D} \supset C$ (by statement), we need to show that $\mathcal{D}$ is a d-system, which by Monotone Class Theorme (for sets) would imply that

$$\mathcal{D} \supset \sigma(\supset C) = \mathcal{E}$$

What want to show:

1. $E \in \mathcal{D}$, by the statement;

2. $A, B \in \mathcal{D}$ $A \supset B \implies A \setminus B \in \mathcal{D}$?

$$\mu(A \setminus B) = \mu(A) - \mu(B)$$
$$\nu(A \setminus B) = \nu(A) - \nu(B)$$

So $\mu(A \setminus B) = \nu(A \setminus B)$.

3. $(A_n) \subset \mathcal{D}, A_n \uparrow A$, (by the sequential continuity of measures)

□

[2] Only countable, not necessarily finite. Each element in the partition is disjoint.

2.1.2 Atoms, atomic & diffuse measures

For a measure space (equipped with measures) $(E, \mathcal{E}, \mu)$,

Atom For $x \in E$ with $\mu(\{x\}) > 0$, we call $\{x\}$ an **atom**.

Proposition 2.1.2. *If μ is finite, it can have at most countable many atoms.*

Completeness & Negligible (Null) Sets In measure space $(E, \mathcal{E}, \mu)$, let $B \in \mathcal{E}$ be a non-empty set. If $\mu(B) = 0$, B is negligible. Furthermore, $A \subset E$ is negligible if $A \subset B \in \mathcal{E}$ s.t. $\mu(B) = 0$¿

A measure space if "complete" if every negligible set is measurable.

Though not all measure space is complete, we can **complete** (v.) the space as in the following proposition:

Proposition 2.1.3. *Let $\mathcal{N}$ be the collection of all negligible subsets of E, define $\bar{\mathcal{E}} = \sigma(\mathcal{E} \bigcup \mathcal{N})$*

a) $B \in \bar{\mathcal{E}}$ implise that $B = A \bigcup N, A \in \mathcal{E}, N \in \mathcal{N}$;

b) $\bar{\mu}(A \bigcup N) = \mu(A)$ defines a unique measure on $\bar{\mathcal{E}}$.

Proof. for a) and b) □

¡++¿

2.1.3 Integration

Simple Functions $f = \sum a_i 1_{A_i \in \mathcal{E}}$. So the measure of a simple function is defined as:

$$\mu f = \int f d\mu := \sum_{i=1}^{n} a_i \mu(A_i)$$

This is a monotone operator.

Positive functions For $f \in \mathcal{E}_+$, there exists $f_n = d_n \circ f \uparrow f$. Define

$$\mu f := \lim \mu f_n$$

General Measurable function For $f \in \mathcal{E}$ (measurable functions mapping from E to $\bar{\mathbb{R}}$ (adding some convention):

$$f = f^+ - f^- \implies \mu f = \mu f^+ - \mu f^-$$

Integrable We say $f \in L^1$ (integrable) if μf^+ & $\mu f^- < \infty$.

Remark: properties

1. $f \geq 0 \implies \mu f \geq 0$;
2. $f \leq g$ implies $\mu f \leq \mu g$.

Notation: $\mu(f 1_A) = \int_A f d\mu$.
Remark: (homework) $A, B \in \mathcal{E}$, $A \bigcap B = \emptyset$, $A \bigcup B = C$, then

$$\mu(f 1_A) + \mu(f 1_B) = \mu(f 1_C)$$

2.1.4 Monotone Convergence Theorem

For $(f_n) \subset \mathcal{E}_+$ and $f_n \uparrow f$, $\mu(\lim f_n) = \lim \mu f_n$. [due to this: we have everything well-defined.]

Proof. For $f \geq f_n$, by remark 2) above, $\mu f \geq \mu f_n$. Therefore, $\mu f \geq \lim_n \mu f_n$.
Now we want to show: $\mu f \leq \lim_n \mu f_n$ in three steps.

Step 1 Fix $b \in \mathbb{R}_+$ and $B \in \mathcal{E}$. $f(x) > b$ for all $x \in B$. [$\geq$ case will be discussed later].

$$\{f_n > b\} \uparrow \{f > b\} \implies B_n := B \cap \{f_n > b\} \uparrow B$$

In the limit, by the sequential continuity of measure: $\lim_n \mu(B_n) = \mu(B)$. So we have:

$$f_n 1_B \geq f_n 1_{B_n} \geq b 1_{B_n}$$

(As 1_{B_n} might be 0, so we did not write strict inequality.) By the remark 2), we have:

$$\mu(f_n 1_B) \geq b\mu(B_n) \implies \lim_n \mu(f_n 1_B) \geq b\mu(B)$$

this is the conclusion, but only given that $f(x) > b$.

Further: want to show $\lim_n \mu(f_n 1_B) \geq b\mu(B)$ if $f \geq b$ on B

Case a If $b = 0$, done;

Case b For $b > 0$, have $b_m \uparrow b$. Then if $f \geq b$ on B, we know $f > b_n$.
$\forall m$, apply step 1, we know that $\lim_n \mu(f_n 1_B) \geq b_m \mu(B)$. Take the limit over m in the end will yield the conclusion.

Step 2 We can the simple function g (simple function representation) so that for arbitrary f (I added this arbitrary part)

$$f \geq g = \sum_{i=1}^{m} b_i 1_{B_i}$$

where $\{B_i\}$ is a finite measurable partition. Then

$$f(x) \geq b_i \qquad \text{for } x \in B_i$$

By Step 1, $\lim_n \mu(f_n 1_{B_i}) \geq b_i \mu({}_i B)$ Therefore, we conclude that:

$$\lim_n \mu(f_n) = \lim_n \sum_{i=1}^{n} \mu(f_n 1_{B_i}) \geq \sum b_i \mu(B_i) = \mu g$$

Step 3 We can "decompose" f in the following fashion:

$$\mu f = \lim_k \mu(d_k \circ f)$$

where $d_k \circ f$ is simple function. Apply Step 2: by construction, we have:

$$f \geq d_k \circ f$$

This implies that:

$$\lim_n \mu f_n \geq \mu(d_k \circ f)$$

Lastly, taking k to the limit will yield:

$$\lim_n \mu f_n \geq \mu f$$

This is true by definition of integral: in a way that is tagged to a particular sequence.

$\square$

Proposition

1) If $A \in \mathcal{E}$ is null, $\mu(f1_A) = 0 \; \forall f \in \mathcal{E}$.

2) $f, g \in \mathcal{E}$ and that $f = g$ almost everywhere (a.e.), $\mu f = \mu g$. (By inferring to the decomposition of simple sets, 1) implies 2).

3) If $f \in \mathcal{E}_+$, $\mu f = 0 \implies f = 0$ a.e.

Proof of 3). For $f \in mE_+$ and $\mu f = 0$. Let $N = \{f > 0\}$ and $N_k = \{f > \varepsilon_k \to 0\}$. So we have: $N_k \uparrow N$. $f \geq \varepsilon 1_{N_k} \implies \mu f \geq \varepsilon_k \mu(N_k) \implies \mu(N_k) = 0$ □

2.1.5 Fatou's lemma and Corollary

Fatou's lemma This lemma would be applied quite frequently.

For $(f_n) \subset \mathcal{E}_+ \implies \mu(\liminf f_n) \leq \liminf \mu(f_n)$ (lower semi-continuity of integration). [This is the part that we lose whne taking $\liminf$.]

Proof. Define $g_m = \inf_{m \leq n} f_n$, it is an increasing sequence where $\liminf f_n = \lim_m g_m$. By monotone convergence theorem, we have:

$$\mu(\liminf f_n) = \lim_m \mu g_m$$

Obseve that $g_m \leq f_n$, $\forall n \geq m$. Monotonicity of integration implies that:

$$\mu g_m \leq \mu f_n, \quad n \geq m$$

Then, in the limit, we have:

$$\mu g_m \leq \inf_{n \geq m} \mu f_n$$

□

Lemma: Reverse Fatou For $f_n \leq g$, $g \in L^1$. Then

$$\mu(\limsup f_n) \geq \limsup \mu f_n$$

(uper semi-continuity of integration). Follows from Fatou directly.

Corollary $|f_n| \leq g \in L^1$, and that if (f_n) has a point-wise limit, by the previous two lemmas,

$$\mu(\lim f_n) = \lim_n \mu(f_n)$$

Pointwise holds (as above), and a.e. also holds! This marks the intensity of measure to null sets.

2.1.6 Schefle's Lemma

Let $X_n, X \in L^1_+$ and $X_n \to X$ a.s. (almost surely). Then

$$\mu(|X_n - X|) \to 0 \text{ if and only if } \mu(X_n) \to \mu(X)$$

(need to be a positive sequence (of functions) in $\mathcal{E}$).

Proof. Under the assumption that $X_n, X \in L^1_+$ implies $\mu(|X_n - X|) \to 0$. □

2.2 Thu:

2.2.1 Scheffe's lemma

If $f_n, f \in L^1(E, \mathcal{E}, \mu)^+$ and $f_n \to f$ a.e., then $\mu(|f_n - f|) \to 0 \iff \mu(f_n) \to \mu(f)$.

Proof. (Showing that RHS implies LHS).

Step 1: show that the negative part converge to 0.

$$(f_n - f)^- = \max(f - f_n, 0) \le f$$

As f is integrable ($\in L^1$), by dominated convergence theorem, we have: $\mu\left(f_n\right) - f)^- \to 0$.

Step 2: show that the negative part converge to 0. Convergence + step 1 implies that $\mu((f_n - f)^+) \to 0$ as

$$\mu\left((f_n + f)^+\right) = \mu(f_n - f) + \mu((f_n - f)^-) \to 0$$

The necessary part is trivial (from LHS to RHS).

$\square$

General idea: for $\mu(|g|) = \mu(g^+) + \mu(g^-)$.

2.2.2 Characterization of Integral

If μ is measure, then all integrals have the following properties:

(a) $\mu(0) = 0$

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

(c) For $0 \le f_n \uparrow f$, $\mu(f_n) \to \mu(f)$.

Remember that

Theorem 2.2.1. *See Theorem 4.21 in the book.*

Proof. Complement to the proof in the book.

It remains toa show that for any $f \in \mathcal{E}_+$,

1. For f being simple functions; $f = \sum_{i=1}^n a_i 1_{A_i}$, $a_i \ge 0$, $A_i \in \mathcal{E}$, then:

   $$\mu(f) = \sum a_i \mu(A_i) = \sum a_i L(1_{A_i}) = L(f)$$

   where the last equality follows by (b).

2. In general, for $f \in \mathcal{E}_+$, let $f_n \in \mathcal{E}_+$ s.t. $f_n \uparrow f$ and f_n is simple.

   By the above step, we have: $L(f_n) = \mu(f_n)$. For $\mu(f_n)$, by Monotone Convergence Theorem, we have: $\mu(f_n) \to \mu(f)$.

   On the LHS, by (c), the condition listed above in the thoerem, we have: $L(f_n) \to L(f)$.

   This completes the proof.

$\square$

2.2.3 Image measures

For measure space $(F, \mathcal{F})$ and $(E, \mathcal{E})$, let $h : F \to E$ be a measure w.r.t. $\mathcal{F}$ and $\mathcal{E}$.

Define $\nu \circ h^{-1} : \mathcal{E} \to \bar{\mathbb{R}}_+$ by $\nu \circ h^{-1}(B) = \nu(h^{-1}B)$, $\forall B \in \mathcal{E}$, where ν is a measure on $(F, \mathcal{F})$. "$\nu \circ h^{-1}$" is a whole thing and is called "**image of ν under** h.

Theorem: Change of Variable $\forall f \in \mathcal{E}_+$, we have $(\nu \circ h^{-1})f = \nu(f \circ h)$.

Intuition: this is nothing but a change of variable.

$$\int f(y)\mu(dy) = \int f(h(x))\nu(dx)$$

Proof. We use theorem 1 to prove it. Define $L(f) = \nu(f \circ h)\ \forall f \in \mathcal{E}_+$. We can check that (a), (b), (c) in Theorem 2.2.1 holds. Then by the theorem, $\exists! \mu$ s.t. $\mu(f) = L(f)$, $\forall f \in \mathcal{E}_+$. Then, $\forall B \in \mathcal{E}$,

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

That is, $\mu \equiv \nu \circ h^{-1}$.

Hence, $\forall f \in \mathcal{E}_+$,

$$(\nu \circ h^{-1})f = \mu(f) = L(f) = \nu(f \circ h)$$

□

2.2.4 Randon-Nikodym Theorem

Definition: Let μ, ν be measures on $(E, \mathcal{E})$, ν is absolutely continuous w.r.t. μ (**notation**: $\nu << \mu$). If $\forall A \in \mathcal{E}$, $\nu(A) = 0 \implies \nu(A) = 0$.

Theorem 3 Assume that ν is σ-finite and $\nu << \mu$, then $\exists p \in \mathcal{E}_+$ s.t. $\forall f \in \mathcal{E}_+$,

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

Moreover, p is unique up to almost everywhere equivalence. (if there is a p' that also satisfies the above equation, $p = p'$ a.e.

Notation-wise, we write:

$$\nu(dx) = p(x)\nu(dx)$$

and define $\frac{\nu(dx)}{\mu(dx)} = p(x)$ as the **Randon-N Derivative**.

2.2.5 Transition Kernel

Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be measure spaces.

Definition: k is a transition kernel from $(E, \mathcal{E})$ into $(F, \mathcal{F})$ if [1]

$$K : E \times \mathcal{F} \to \bar{\mathbb{R}}_+$$

(a) For any (fixed) $B \in \mathcal{F}$, the map $X \to k(x, B)$ is $\mathcal{E}$-measurable for $x \in E$.

(b) For nay (fixed) $x \in E$, the map: $B \to K(x, B)$ is a measure on $(F, \mathcal{F})$ [note: $B \in \mathcal{F}$.] where B defines the transition probability.

[1] note that E is the set and $\mathcal{F}$ is the σ-algebra of F, the set.

Theorem 4: Let K be a transition kernel from $(E, \mathcal{E})$ into $(F, \mathcal{F})$, then:

(a) $\forall f \in \mathcal{F}_+$,

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

defines a function Kf that is in $\mathcal{E}_+$.

(b) For any measure μ on $(E, \mathcal{E})$,

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

defines a measure μK on $(F, \mathcal{F})$.

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

Proof. • We show (a) in two steps:

Step 1: if $f = \sum b_i 1_{B_i}$, $B_i \in \mathcal{F}$.

$$\begin{aligned} kf(x) &= \int_F \sum b_i 1_{B_i} K(x, dy) \\ &= \sum b_i \int_F 1_{B_i} K(x, dy) \\ &= \sum b_i K(x, B_i) \in \mathcal{E}_+ \end{aligned}$$

Step 2: In general, for $f \in \mathcal{E}_+$, take f_n being simple functions s.t. $f_n \uparrow f$. By Monotone convergence theorem,

$$Kf(x) = \lim_{n\to\infty} Kf_n(x) \in \mathcal{E}_+$$

- We show (b) and (c) together in two steps:

Step 1 Define $L(f) := \mu(Kf), f \in \mathcal{F}_+$, show that (a), (b) and (c) would hold in Theorem 1.

(a) $L(0) = \mu(K0) = \mu(0) = 0$

(b) For $f, g \in \mathcal{F}_+$ and $a, b \geq 0$, then

$$L(af + bg) = \mu(K(af + bg)) = \mu(aKf + bKg)$$

by linearity of integrals, due to the way in which we had defined Kf. Further, due to linearity of μ:

$$L(af + bg) = a\mu(Kf) + b\mu(kg) = aL(f) + bL(f)$$

(c) For $f_n, f \in \mathcal{F}_+$, $f_n \uparrow f$. Write:

$$L(f) = \mu(Kf) = \mu(K(\lim_n f_n)) = \mu(\lim_n Kf_n)$$

due to MCT;
since K is a transition kernel and $f_n \in \mathcal{F}_+$, we have:

$$L(f) = \lim_n \mu(Kf_n) = \lim_n L(f_n)$$

Then, by Theorem 1, there is a unique measure ν on $(F, \mathcal{F})$ s.t. $\nu(f) = L(f)$, $\forall f \in \mathcal{F}_+$. Let $f \in 1_B$ $B \in \mathcal{F}$, then

$$\nu(B) = L(1_B) = \nu(K1_B) = \mu(K(B))$$

i.e. $\nu = \mu K$ is a measure on $(F, \mathcal{F})$.
Now, $\forall f \in \mathcal{F}_+$, $(\mu K)f = \nu(f) = L(f) = \mu(Kf)$ [by definition in the begining of the proof for this part.

□

Propositoin: Let K be a finite kernel from (E, ε) into $(F, \mathcal{F})$ (i.e. $\forall x \in E$, $k(x, F) < \infty$). The forall positive function $f \in \mathcal{E} \otimes \mathcal{F}$,

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

defines a fucntion Tf in $\mathcal{E}_+$.

Proof. Idea: use Monotone Class Theorem. Define $\mathcal{M} := \{f \in \mathcal{E} \otimes \mathcal{F} : f \text{ is positive or bounded}, Tf \in \mathcal{E}\}$. We show $\mathcal{M}$ is a monotone class. For any rectangle $A \times B \in \mathcal{E} \otimes \mathcal{F}$.

$$T1_{A\times B}(x) = \int_F K(x, dy) 1_{A\times B}(x, y)$$

Upon re-writing the indicator function:

$$T1_{A\times B}(x) = \int_F K(x, dy) 1_A(x) 1_B(y)$$

Since $1_A(x)$ is independent of y, we have:

$$T1_{A\times B}(x) = 1_A(x) K(x, B) \in \mathcal{E}_+$$

By monotone class thoerem, $\mathcal{M}$ includes all positive function $\in \mathcal{E} \otimes \mathcal{F}$.

□

¡++¿