Measures and Random Variables

2.1 Dynkin's $\pi - \lambda$ theorem

A collection of subsets is called a π-system if it is closed under finite intersection. A collection of subsets $\mathcal{D}$ is called a λ-system if it satisfies

i. $\Omega \in \mathcal{D}$
ii. $\mathcal{D}$ is closed under complement, and
iii. $\mathcal{D}$ is closed under countable union of disjoint sets.

So, a λ-system is almost a σ-field except not closed under general countable union. The following will be useful to characterize σ-fields.

Theorem 2.1 (Dynkin's $\pi - \lambda$ theorem)**.** *Let $\mathcal{C}$ be a π-system, $\mathcal{D}$ be a λ-system, and $\mathcal{C} \subseteq \mathcal{D}$. Then, $\sigma(\mathcal{C}) \subseteq \mathcal{D}$.*

The original definition of λ-system is different, but equivalent. We choose this definition as it is typically easy to verify.

2.2 Measures on Real Line

Definition 2.2. *A **Stieltjes measure function** is a function $F : \mathbb{R} \longrightarrow \mathbb{R}$ if*

(i) *F is non-decreasing,*

(ii) *F is right continuous, i.e., if $x_n \downarrow x$, $F(x_n) \to F(x)$.*

We can define $\mu_F((a, b]) = F(b) - F(a)$.

Application: Take a nondecreasing and right-continuous function F on $\mathbb{R}$. Let S be the semi-algebra $\{(a, b] \mid -\infty \leqslant a \leqslant b \leqslant \infty\}$ and $\mu((a, b]) = F(b) - F(a), -\infty < a < b < \infty$. When $a = -\infty$ or $b = \infty$, we take $F(\infty) = \lim_{x\uparrow\infty} F(x)$, $F(-\infty) = \lim_{x\downarrow-\infty} F(x)$.

i) $\overline{\mu}((a, b]) = F(b) - F(a) \geqslant 0$,

ii) If $(a, b] = \cup_{i=1}^{n}(a_i, b_i]$, order a_i in increasing order. We claim that $a_i = a, b_i = a_i, \cdots, b_n = b$. Then $\mu((a, b]) = F(b) - F(a) = \sum_{i=1}^{n}(F(b_i) - F(a_i)) = \sum_{i=1}^{n} \mu((a_i, b_i])$.

iii) If $(a, b] \subseteq \cup_{i=1}^{\infty}(a_i, b_i]$, take $\varepsilon > 0$. Using right-continuity of F, change $b_i \to b_i + \delta_i$ where $\delta_i > 0, F(b_i + \delta_i) - F(b_i) \leqslant \frac{\varepsilon}{2^i}$, and change $a \to a + \delta$ where $\delta > 0, F(a + \delta) - F(a) \leqslant \varepsilon$. Then $[a + \delta, b] \subseteq \cup_{i=1}^{\infty}(a_i, b_i + \delta_i)$. Extract a finite subcover $[a + \delta, b] \subseteq \cup_{i=1}^{n}(a_{j_i}, b_{j_i} + \delta_{j_i})$. Show that

$$F(b) - F(a + \delta) \quad \leqslant \textstyle\sum_{i=1}^{n} \left(F(b_{j_i} + \delta_{j_i}) - F(a_{j_i})\right) \leqslant \sum_{i=1}^{\infty} \left(F(b_i + \delta_i) - F(a_i)\right).$$

This implies $F(b) - F(a) \leqslant 2\varepsilon + \sum_{i=1}^{\infty}(F(b_i) - F(a_i))$. Since ε is arbitrary, we are done.

As an example, let $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$ and μ be a probability measure. Define

$$F(x) = \mu((-\infty, x]),\ x \in \mathbb{R}.$$

Check that F is non-decreasing and right-continuous, $\mu((a,b]) = F(b) - F(a)$ and $\mu(\mathbb{R}) = F(\infty) - F(-\infty)$.

Example 2.3. *Let* $\Omega = \mathbb{R}^2$ *and* $\mathscr{S} = \{(a,b] \times (c,d] \mid -\infty \leqslant a \leqslant b \leqslant \infty, -\infty \leqslant c \leqslant d \leqslant \infty\}$. *Show that* $\mathscr{S}$ *is a semi-algebra. Find the properties needed for a function* $F : \mathbb{R}^2 \to (-\infty, \infty)$ *to define a measure* μ_F *on* $\mathcal{B}^2$ *such that* $\mu_F((a,b] \times (c,d]) = F(b,d) - F(a,d) - F(b,c) + F(a,c)$ *for all* $a \leqslant b, c \leqslant d$.

Given a measure space $(\Omega, \mathcal{F})$ and a measure μ on it, we define the completion of $\mathcal{F}$ w.r.t. μ as

$$\bar{\mathcal{F}} := \{A \cup B \mid A \in \mathcal{F}, B \subseteq N \in \mathcal{F} \text{ such that } \mu(N) = 0\}.$$

We also extend μ to $\bar{\mu}$ on $\bar{\mathcal{F}}$ by $\bar{\mu}(A \cup B) = \mu(A)$ where $B \subseteq N \in \mathcal{F}$ with $\mu(N) = 0$.

Example 2.4. *Show that* $\bar{\mathcal{F}}$ *is a* σ*-field. Moreover,* $\bar{\mu}$ *is well-defined and is a measure on* $(\Omega, \bar{\mathcal{F}})$.

2.3 Random Variables

Definition 2.5. *Given two measure spaces* $(\Omega, \mathcal{F})$ *and* $(S, \mathscr{S})$*, a function* $f : \Omega \to \mathscr{S}$ *is measurable if* $f^{-1}(A) \in \mathcal{F}$ *for any* $A \in \mathscr{S}$.

Definition 2.6. *A random variable is a measurable function from* $(\Omega, \mathcal{F})$ *to* $(\mathbb{R}, \mathcal{B})$*, where* $\mathcal{B}$ *is the Borel* σ*-field of* $\mathbb{R}$*. A random vector is a measurable function from* $(\Omega, \mathcal{F})$ *to* $(\mathbb{R}^d, \mathcal{B}^d)$*, where* $\mathcal{B}^d$ *is the Borel* σ*-field of* $\mathbb{R}^d$.

Notation: We will use X, Y, Z as r.v.s on $(\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$.

Proposition 2.7. (i) *Composition of two measurable functions is measurable.*

(ii) *Let* $\mathcal{C}$ *be a collection of sets from* $\mathscr{S}$ *and* $\sigma(\mathcal{C}) = \mathscr{S}$*. If* $f^{-1}(A) \in \mathcal{F}$ *for any* $A \in \mathcal{C}$*, then* f *is measurable.*
(Hint: $\{A \mid f^{-1}(A) \in \mathcal{F}\}$ *is a* σ*-field.)*

Example 2.8. *Take* $\mathcal{B} = \sigma(\{(-\infty, a] \mid a \in \mathbb{R}\})$*. For* $X : \Omega \to \mathbb{R}$*, it is enough to check* $X^{-1}((-\infty, a]) = \{\omega : X(\omega \leqslant a)\} = \{X \leqslant a\} \in \mathcal{F}$.

Notation: $\{X \in A\} = X^{-1}(A) = \{\omega : X(\omega) \in A)\}$.

Example 2.9. $X : \Omega \to \mathbb{R}^d, \mathcal{B}^d = \sigma(\prod_{i=1}^{d}(-\infty, a_i] : (a_1, a_2, ..., a_d) \in \mathbb{R}^d)$.

Definition 2.10. (i) *If* $f : (\Omega, \mathcal{F}) \to (S, \mathscr{S})$ *is measurable, then* $f^{-1}(\mathscr{S}) = \{f^{-1}(A) : A \in \mathscr{S}\}$ *is also a* σ*-field. This is called the* σ*-field generated by* f.

(ii) *If also,* $(\Omega, \mathcal{F}, \mu)$ *is a measurable space, then we can define* $\gamma = \mu \circ X^{-1}$ *a measure on* $(S, \mathscr{S})$*. Then we have* $\gamma(A) = \mu(X^{-1}(A)) = \mu(x \in A), A \in \mathscr{S}$.

In particular, For random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B})$, $\mu(A) := \mathbb{P}(x \in A)$ where $A \in \mathcal{B}$ is a probability on $(\mathbb{R}, \mathcal{B})$.

2.4 Distributions

Definition 2.11. *The distribution function of X is defined as $F(x) = \mathbb{P}(X \leqslant x)$ for $x \in \mathbb{R}$. Then we have $F(a) - F(b) = \mathbb{P}(X \in (b, a])$.*

Proposition 2.12. (i) *F is non-decreasing.*

(ii) *F is right-continuous.*

(iii) *$F(x) \downarrow 0$ as $x \downarrow -\infty, F(x) \uparrow 1$ as $x \uparrow \infty$.*

(iv) *$\lim_{x_n \to x-} F(x_n)$ exists and $\lim_{x_n \to x-} F(x_n) = \mathbb{P}(X < x)$.*

Lemma 2.13. *Given any F satisfying (i)-(iii), there exists a r.v. X defined on $((0,1), \mathcal{B}, \mathbb{P} = \lambda)$ with distribution function F.*

Proof. Define $X : ((0,1), \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ as

$$X(\omega) = \sup\{y : F(y) < \omega\}.$$

We need to check that i) X is measurable and ii) X has distribution function F.

We claim that, $\{\omega : X(\omega) \leqslant a\} = \{\omega : \omega \leqslant F(a)\}$. Indeed, If $\omega \leqslant F(a)$, then for any y which satisfies $F(y) < \omega$ also satisfies $y < a$ and thus, $X(\omega) \leqslant a$. On the other hand, if $\omega > F(a)$, then $\exists\, \varepsilon > 0$ s.t. $\omega > F(a + \varepsilon)$ by right continuity of F. Thus, by definition of X, $X(\omega) \geqslant a + \varepsilon > a$.

In fact, X behaves like the inverse function of F. We define $F^{-1}(x) := \sup\{y : F(y) < x\}, x \in (0,1)$.

Besides, $\mathbb{P}\{\omega : X(\omega) \leqslant a)\} = \mathbb{P}\{\omega : \omega \leqslant F(a)\} = F(a)$ since $\omega \in (0,1)$. So F is distribution function of X. ■

Remark: If U is a uniform distributed r.v. on $(0,1)$, *i.e.*,

$$\mathbb{P}(U \leqslant x) = \begin{cases} x & 0 < x < 1 \\ 0 & x \leqslant 0 \\ 1 & x > 1 \end{cases}$$

then, $X = F^{-1}(U)$.

2.5 Properties of Random Variables

Definition 2.14. (i) *Two random variables X, Y are equal in distribution if X and Y has the same distribution function.*

(ii) *Two random variables X, Y are equal a.s. if they are defined on the same measurable space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbb{P}(X \neq Y) = 0$.*

Example 2.15. (i) ***Uniform**$(0,1)$ distribution with $F(a) = \int_{-\infty}^{a} f(x)dx$ and*

$$f(x) = \begin{cases} 0 & x \notin (0,1) \\ 1 & x \in (0,1) \end{cases}.$$

(ii) ***Exponential***(λ) *with distribution function*

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geqslant 0 \\ 0 & x < 0 \end{cases} \quad \textit{with density } f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leqslant 0 \end{cases}$$

(iii) *The distribution for the measure* $\mu\{i\} = p_i, i = 1, 2, 3, \dots$ *where* $\sum_{i=1}^{\infty} p_i = 1$ *is a step function.*

(iv) *Check that Cantor function gives a distribution function supported on the Cantor set.*

Proposition 2.16. (i) *Any continuous function is measurable.*

(ii) *If* f *is a continuous function and* $(X_1, \dots, X_n)$ *is a random vector, then* $f(X_1, X_2, \dots, X_n)$ *is measurable.*

(iii) $X_i : (\Omega_i, \mathcal{F}_i) \to (\mathbb{R}, \mathcal{B}), i = 1, \dots, n$ *are r.v.'s, then* $(X_1, \dots, X_n) : (\Omega_1 \times \cdots \times \Omega_n, \mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_n) \to (\mathbb{R}^n, \mathcal{B}^n)$ *is also measurable where we write* $\mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_n = \sigma(\mathcal{F}_1 \times \cdots \times \mathcal{F}_n)$.

(iv) *If* X_n *is a sequence of r.v.'s, then* $\inf_{n \geqslant 1} X_n, \sup_{n \geqslant 1} X_n, \liminf X_n, \limsup X_n$ *are r.v.'s.*

Proof. Proof of (i) follows from the fact that for any open set A, $f^{-1}(A)$ is open and hence measurable when f is continuous. Part (ii) follows from composition of two measurable functions are measurable. For (iii), note that $\{(X_1, ..X_n) \in \prod_{i=1}^n (-\infty, a_i]\} = \prod_{i=1}^n \{X_i \leqslant a_i\}$. Finally, (iv) follows from the fact that

$$\{\inf_{n \geqslant 1} X_n \geqslant a\} = \cap_{n \geqslant 1} \{X_n \geqslant a\}, \{\sup_{n \geqslant 1} X_n \leqslant a\} = \cap_{n \geqslant 1} \{X_n \leqslant a\}$$

and $\mathcal{B} = \sigma([a, \infty) : a \in \mathbb{R}) = \sigma((-\infty, a] : a \in \mathbb{R})$. Finally note that,

$$\liminf X_n = \sup_{n \geqslant 1} \inf_{m \geqslant n} X_m, \quad \limsup X_n = \inf_{n \geqslant 1} \sup_{m \geqslant n} X_m.$$

■

Definition 2.17. (i) *A sequence of r.v.'s* (X_n) *is said to converge a.s. if* $\mathbb{P}(\lim X_n$ *exists*$) = 1$.

(ii) *We say that* (X_n) *converges in probability to* X *if* $\lim_{n \to \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$ *for any* $\varepsilon > 0$.

(iii) *We say that* X_n *converges in distribution to* X *if* $F_n \to F$ *pointwise at all continuity points of* F.

Example 2.18. ***Indicator random variable.***

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \textit{if } \omega \in A \\ 0 & \textit{o.w.} \end{cases}.$$