Lecture 3

Instructor: Quan Zhou

For more details about the materials covered in this note, see Chapters 3.1 and 3.2 of Resnick [3] and Chapter 1.3 of Durrett [2].

3.1 Inverse maps

Definition 3.1. Let Ω, Λ be two sets and consider a function $f \colon \Omega \to \Lambda$. For $A \subset \Lambda$, the inverse image of A under f is

$$f^{-1}(A) = \{\omega \in \Omega : \ f(\omega) \in A\}.$$

Example 3.1. A simple function means a function with a finite range (finite number of possible values). For a real-valued simple function (i.e. $\Lambda = \mathbb{R}$), we may denote the range by $\{a_1, \dots, a_k\}$, where a_i's are distinct real numbers. Define $A_i = f^{-1}(\{a_i\})$. Then, $\{A_i \colon i = 1, \dots, k\}$ partitions Ω. ("Partition" means $\cup_{i=1}^k A_i = \Omega$ and A_i's are disjoint.) Further, the function can be expressed by $f = \sum_{i=1}^k a_i \mathbb{1}_{A_i}$.

Proposition 3.1. *f^{-1} preserves complementation, unions and intersections; that is, $f^{-1}(A^c) = (f^{-1}(A))^c$, $f^{-1}(\cup_{t \in T} A_t) = \cup_{t \in T} f^{-1}(A_t)$ and $f^{-1}(\cap_{t \in T} A_t) = \cap_{t \in T} f^{-1}(A_t)$.*

Proof. Try it yourself. □

Lemma 3.1. *Let $\mathcal{G}$ be a σ-algebra on Λ. Then, $f^{-1}(\mathcal{G}) = \{f^{-1}(A) : \ A \in \mathcal{G}\}$ is a σ-algebra on Ω.*

Proof. We only need to verify the three postulates. (i) Since $\Lambda \in \mathcal{G}$, we have $\Omega = f^{-1}(\Lambda) \in f^{-1}(\mathcal{G})$. (ii) If $f^{-1}(A) \in f^{-1}(\mathcal{G})$, so is $(f^{-1}(A))^c = f^{-1}(A^c)$ by Proposition 3.1. (iii) If $f^{-1}(A_i) \in f^{-1}(\mathcal{G})$ for $i = 1, 2, \dots,$ we have $\cup_i f^{-1}(A_i) = f^{-1}(\cup_i A_i) \in f^{-1}(\mathcal{G})$ since $\cup_i A_i \in \mathcal{G}$ and f^{-1} preserves unions by Proposition 3.1. □

Remark 3.1. Sometimes we also use the notation $\sigma(f) = f^{-1}(\mathcal{G})$, and we say that $\sigma(f)$ is the σ-algebra generated by f. Of course, when $\sigma(f)$ is used, it is assumed that $\mathcal{G}$ is clear from context; for example, when $\Lambda = \mathbb{R}$, the notation $\sigma(f)$ means $f^{-1}(\mathcal{B}(\mathbb{R}))$.

Theorem 3.1. *If $\mathcal{A} \subset \mathcal{P}(\Lambda)$ (i.e. $\mathcal{A}$ is a collection of subsets of Λ), then $f^{-1}(\sigma(\mathcal{A})) = \sigma(f^{-1}(\mathcal{A}))$.*

Proof. First, by Lemma 3.1 $f^{-1}(\sigma(\mathcal{A}))$ is a σ-algebra and thus $f^{-1}(\sigma(\mathcal{A})) \supset \sigma(f^{-1}(\mathcal{A}))$. Second, define $\mathcal{C} = \{B \subset \Lambda : f^{-1}(B) \in \sigma(f^{-1}(A))\}$ and show that $\mathcal{C}$ is also a σ-algebra. Clearly, $\mathcal{A} \subset \mathcal{C}$ and thus $\sigma(\mathcal{A}) \subset \mathcal{C}$. It follows that the other direction also holds, i.e. $f^{-1}(\sigma(\mathcal{A})) \subset \sigma(f^{-1}(\mathcal{A}))$, which concludes the proof. □

3.2 Measurable functions and random variables

Definition 3.2. Let $(\Omega, \mathcal{F})$ and $(\Lambda, \mathcal{G})$ be two measurable spaces and $f\colon \Omega \to \Lambda$ be a function. We say f is a measurable function if $f^{-1}(\mathcal{G}) \subset \mathcal{F}$ and we write $f\colon (\Omega, \mathcal{F}) \to (\Lambda, \mathcal{G})$. When Ω and Λ are clear from text and we only want to emphasize the σ-algebra, we may write $f \in \mathcal{F}/\mathcal{G}$.

If $(\Lambda, \mathcal{G}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, we say f is Borel measurable or a Borel function and often simply write $f \in \mathcal{F}$.

Definition 3.3. In probability theory, a real valued Borel function is called a random variable for $d = 1$ and a random vector for $d > 1$ and is often denoted by $X, Y, \ldots$.

Example 3.2. Consider a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Let $X = \mathbb{1}_A$ for some $A \in \mathcal{F}$. Then X is a random variable and $\sigma(X) = \{\emptyset, \Omega, A, A^c\}$.

Proposition 3.2 (Test for measurability)**.** *Consider measurable spaces $(\Omega, \mathcal{F})$, $(\Lambda, \mathcal{G})$ and function $f\colon \Omega \to \Lambda$. If $f^{-1}(\mathcal{A}) \subset \mathcal{F}$ for some $\mathcal{A}$ that generates $\mathcal{G}$, then f is measurable.*

Proof. If $f^{-1}(\mathcal{A}) \subset \mathcal{F}$, we have $\sigma(f^{-1}(\mathcal{A})) \subset \mathcal{F}$ by the minimality of the generated σ-algebra. Then apply Theorem 3.1. □

Corollary 3.1. *The real valued function $X\colon \Omega \to \mathbb{R}$ is a random variable iff $X^{-1}((-\infty, b]) \in \mathcal{F}$ for any $b \in \mathbb{R}$.*

Proof. Try it yourself. □

Proposition 3.3. *Let X be a random variable. If $\mathcal{A}$ generates $\mathcal{B}(\mathbb{R})$, then $\sigma(X) = \sigma(\{X^{-1}(A) : A \in \mathcal{A}\})$.*

Proof. Try it yourself. □

Proposition 3.4 (Composition). *Let $f : (\Omega_1, \mathcal{B}_1) \to (\Omega_2, \mathcal{B}_2)$ and $g : (\Omega_2, \mathcal{B}_2) \to (\Omega_3, \mathcal{B}_3)$ where $(\Omega_i, \mathcal{B}_i)$ $(i = 1, 2, 3)$ are measurable spaces. Define the composition $g \circ f : \Omega_1 \to \Omega_3$ by $g \circ f(\omega_1) = g(f(\omega_1))$ for $\omega_1 \in \Omega_1$. Then $g \circ f \in \mathcal{B}_1/\mathcal{B}_3$.*

Proof. Try it yourself. □

Proposition 3.5 (Converse to Proposition 3.4). *Let $f : (\Omega_1, \mathcal{B}_1) \to (\Omega_2, \mathcal{B}_2)$ and $h : \Omega_1 \to \mathbb{R}$. Then, $h \in \sigma(f)/\mathcal{B}(\mathbb{R})$ if and only if there exists some $g : (\Omega_2, \mathcal{B}_2) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $h = g \circ f$.*

Proof. Try it yourself after Lecture 4. □

Proposition 3.6. *Let $f : \mathbb{R}^m \to \mathbb{R}^d$ be a continuous function. Then $f \in \mathcal{B}(\mathbb{R}^m)/\mathcal{B}(\mathbb{R}^d)$.*

Proof. It follows from the definition of Borel σ-algebra and the fact that $f^{-1}(A)$ is open if $A \subset \mathbb{R}^d$ is open and f is continuous. □

Lemma 3.2. *$X = (X_1, \ldots, X_n)$ is a random vector iff X_i is a random variable for every i.*

Proof. The proof relies on the fact that $\mathcal{B}(\mathbb{R}^n)$ is generated by the collection of all the rectangles in $\mathbb{R}^n$. See Durrett [2, Theorem 1.3.5]. □

Theorem 3.2. *If $X_1, \ldots, X_n$ are random variables and $f : (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, then $f(X_1, \ldots, X_n)$ is a random variable and $X = (X_1, \ldots, X_n)$ is a random vector.*

Proof. By Proposition 3.4, if $(X_1, \ldots, X_n)$ is a measurable function which maps from $(\Omega, \mathcal{F})$ to $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, then $f(X_1, \ldots, X_n)$ is a random variable. In other words, we need to show $(X_1, \ldots, X_n)$ is a random vector. But this follows from Lemma 3.2. □

Theorem 3.3. *If $X_1, X_2, \ldots,$ are random variables, then $\inf_n X_n$, $\sup_n X_n$, $\liminf_n X_n$, $\limsup_n X_n$ are measurable. Note that they take value in the measurable space $(\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$.*

Proof. Observe that $\{\inf_n X_n < x\} = \cup_n \{X_n < x\}$ and $\{\sup_n X_n > x\} = \cup_n \{X_n > x\}$. Then use the property that a σ-algebra is closed under countable unions/intersections. □

Proposition 3.7. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $(\Lambda, \mathcal{G})$ be a measurable space and $f : (\Omega, \mathcal{F}) \to (\Lambda, \mathcal{G})$. Define a function on $\mathcal{G}$, denoted by $\mu \circ f^{-1}$ (or $f_{\#}\mu$), as $(\mu \circ f^{-1})(A) = \mu(f^{-1}(A))$ for any $A \in \mathcal{G}$. Then $\mu \circ f^{-1}$ is a measure on $(\Lambda, \mathcal{G})$. It is called the push-forward measure of μ or the measure induced by f.*

Proof. Try it yourself. □

Definition 3.4. For a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and a random variable X defined on it, $\mathsf{P} \circ X^{-1}$ is called the distribution or the law of X.

Example 3.3. Consider tossing two dice, which corresponds to the sample space $\Omega = \{(i, j) : 1 \leq i, j \leq 6\}$. Let $\Lambda = \{2, 3, \ldots, 12\}$. Define $X : \Omega \to \Lambda$ by $X((i, j)) = i + j$. Then $X^{-1}(\{2, 3\}) = \{(1, 1), (1, 2), (2, 1)\}$. The distribution of X is given by the push-forward measure $\mathsf{P} \circ X^{-1}$ where P denotes the probability measure on Ω. Hence, $\mathsf{P} \circ X^{-1}(\{2, 3\}) = 3/36$.

References

[1] Dennis D. Cox. *The Theory of Statistics and Its Applications*. Unpublished.

[2] Rick Durrett. *Probability: Theory and Examples*, volume 49. Cambridge university press, 2019.

[3] Sidney Resnick. *A Probability Path*. Springer, 2019.