Lecture 7

Instructor: Quan Zhou

For more details about the materials covered in this note, see Chapters 4.1 and 4.2 of Resnick [2] and Chapter 2.1 of Durrett [1].

7.1 Independence

Definition 7.1. Independence for two events/σ-algebras/random variables.

(i) Two events A and B are independent if $\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$.

(ii) Two σ-algebras are independent if for any $A \in \mathcal{F}, B \in \mathcal{G}$, the events A and B are independent.

(iii) Two random variables X and Y are independent if for all $A, B \in \mathcal{B}(\mathbb{R})$, we have $\mathsf{P}(X \in A, Y \in B) = \mathsf{P}(X \in A)\mathsf{P}(Y \in B)$.

Definition 7.2. Mutual independence.

(i) σ-algebras $\mathcal{F}_1, \dots, \mathcal{F}_n$ are independent if whenever $A_i \in \mathcal{F}_i$ for each i, we have $\mathsf{P}(\cap_{i=1}^n A_i) = \prod_{i=1}^n \mathsf{P}(A_i)$.

(ii) Random variables $X_1, \dots, X_n$ are independent if whenever $B_i \in \mathcal{B}(\mathbb{R})$ for each i, we have $\mathsf{P}(\cap_{i=1}^n \{X_i \in \mathcal{B}_i\}) = \prod_{i=1}^n \mathsf{P}(X_i \in \mathcal{B}_i)$.

(iii) Sets (events) $A_1, \dots, A_n$ are independent if whenever $I \subset \{1, 2, \dots, n\}$ we have $\mathsf{P}(\cap_{i \in I} A_i) = \prod_{i \in I} \mathsf{P}(A_i)$.

(iv) Let $\mathcal{A}_1, \dots, \mathcal{A}_n$ be collections of measurable subsets of Ω. We say they are independent if whenever $A_i \in \mathcal{A}_i$ for each i, the events $A_1, \dots, A_n$ are independent.

Theorem 7.1. *If X and Y are independent, then $\sigma(X)$ and $\sigma(Y)$ are independent.*

Proof. If a set $A \in \sigma(X)$, then by definition $A = \{\omega \colon X(\omega) \in C\}$ for some $C \in \mathcal{B}(\mathbb{R})$. Similarly, if $B \in \sigma(Y)$, then $B = \{\omega \colon Y(\omega) \in D\}$ for some $D \in \mathcal{B}(\mathbb{R})$. Hence,

$$\mathsf{P}(A \cap B) = \mathsf{P}(\{\omega \colon X(\omega) \in C, \, Y(\omega) \in D\}) = \mathsf{P}(X(\omega) \in C)\mathsf{P}(Y(\omega) \in D)$$

since X, Y are independent. But the right-hand side is just $\mathsf{P}(A)\mathsf{P}(B)$. $\square$

Theorem 7.2. *If $\mathcal{F}$ and $\mathcal{G}$ are independent, $X \in \mathcal{F}$ and $Y \in \mathcal{G}$, then X and Y are independent.*

Proof. If X is a measurable function with respect to $\mathcal{F}$ and Y is measurable with respect to $\mathcal{G}$, then by definition for any $A, B \in \mathcal{B}(\mathbb{R})$ we have $\{X \in A\} \in \mathcal{F}$ and $\{Y \in B\} \in \mathcal{G}$. Since $\mathcal{F}$ and $\mathcal{G}$ are independent, the two events $\{X \in A\}$ and $\{Y \in B\}$ are independent. □

Example 7.1. Pairwise independence does not imply (mutual) independence. Consider a box containing 4 tickets labeled $112, 121, 211, 222$. Let A_i denote the event that the i-th digit is 1 for $i = 1, 2, 3$. Clearly, $\mathsf{P}(A_1) = \mathsf{P}(A_2) = \mathsf{P}(A_3) = 1/2$. Further, $\mathsf{P}(A_1 \cap A_2) = \mathsf{P}(A_1 \cap A_3) = \mathsf{P}(A_2 \cap A_3) = 1/4$. However, $\mathsf{P}(A_1 \cap A_2 \cap A_3) = 0$.

Example 7.2. Definition 7.2 (iii) may seem complicated but it cannot be simplified. Consider $\Omega = \{1, 2, 3, 4 \ldots, 16\}$ with a uniform probability measure (i.e. probability 1/16 for each outcome) and the following 4 events

$$\begin{aligned} A &= \{1, 2, 4, 5, 6, 9, 10, 16\}, \quad & B &= \{1, 2, 3, 4, 7, 8, 11, 12\}, \\ C &= \{1, 3, 4, 5, 7, 8, 11, 12\}, \quad & D &= \{1, 2, 3, 5, 6, 9, 10, 15\}. \end{aligned}$$

Clearly $\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = \mathsf{P}(D) = 1/2$. One can check that

$$\begin{aligned} &\mathsf{P}(A \cap B \cap C \cap D) = \mathsf{P}(\{1\}) = 1/16, \\ &\mathsf{P}(A \cap B \cap C) = \mathsf{P}(\{1, 4\}) = 1/8, \\ &\mathsf{P}(A \cap B \cap D) = \mathsf{P}(\{1, 2\}) = 1/8, \\ &\mathsf{P}(A \cap C \cap D) = \mathsf{P}(\{1, 5\}) = 1/8, \\ &\mathsf{P}(B \cap C \cap D) = \mathsf{P}(\{1, 3\}) = 1/8. \end{aligned}$$

However, we do not have any pairwise independence: $\mathsf{P}(A \cap B) = \mathsf{P}(A \cap C) = \mathsf{P}(B \cap D) = \mathsf{P}(C \cap D) = 3/16, \mathsf{P}(A \cap D) = 6/16, \mathsf{P}(B \cap C) = 7/16$.

7.2 Properties of independent random variables

Lemma 7.1. *If $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n$ are independent and each $\mathcal{A}_i$ is a π-system, then $\sigma(\mathcal{A}_1), \ldots, \sigma(\mathcal{A}_n)$ are independent.*

Proof. See the textbook. □

Theorem 7.3 (Factorization theorem)**.** *Random variables $X_1, \dots, X_n$ are independent if for all $x_1, \dots, x_n \in (-\infty, \infty]$, we have*

$$\mathsf{P}(X_1 \le x_1, \dots, X_n \le x_n) = \prod_{i=1}^{n} \mathsf{P}(X_i \le x_i).$$

Proof. Let $\mathcal{A}_i = \{\{X_i \le x\} : x \in (-\infty, \infty]\}$ for $i = 1, \dots, n$. It is straightforward to check that $\mathcal{A}_i$ is a π-system. Further, $\sigma(\mathcal{A}_i) = \sigma(X_i)$ by Proposition 3.3. The result then follows from Lemma 7.1. □

Corollary 7.1. *Discrete random variables $X_1, \dots, X_n$ are independent if*

$$\mathsf{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^{n} \mathsf{P}(X_i = x_i),$$

for all possible values of $(x_1, \dots, x_n)$.

Proof. Try it yourself. □

Example 7.3. Let $X_1, X_2, \dots$ be a sequence of i.i.d. continuous random variables with distribution function $F(x)$. X_n is called a record if $X_n > \max\{X_i : i = 1, \dots, n-1\}$. It can be proven that the events $A_n = \{X_n \text{ is a record}\}$ are independent. See Resnick [2, §4.3].

Theorem 7.4. *If $X_1, \dots, X_n$ are independent random variables and X_i has distribution μ_i, then $(X_1, \dots, X_n)$ has distribution $\mu_1 \times \cdots \times \mu_n$.*

Proof. It follows from Dynkin's π-λ theorem and $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R})^n$. □

Theorem 7.5. *Suppose X, Y are independent random variables, and $f, g : \mathbb{R} \to \mathbb{R}$ are measurable functions such that either $f, g \ge 0$, or both $f(X)$ and $g(Y)$ are integrable, then $E[f(X)g(Y)] = E[f(X)]\, E[g(Y)]$.*

Proof. Here we only prove a special case: $E[XY] = E[X]E[Y]$ for non-negative independent random variables X and Y. The proof for the general case is very similar.

Let $(\Omega, \mathcal{F}, \mathsf{P})$ be the underlying probability space. Denote the laws of X and Y by $\mathsf{P}_X = \mathsf{P} \circ X^{-1}$ and $\mathsf{P}_Y = \mathsf{P} \circ Y^{-1}$ respectively. Let $Z = (X, Y)$

and denote the distribution of Z by $\mathsf{P}_Z = \mathsf{P} \circ Z^{-1}$. By the independence assumption, for any Borel sets A, B,

$$\mathsf{P}_Z(A \times B) = \mathsf{P}(Z \in A \times B) = \mathsf{P}_X(A)\mathsf{P}_Y(B).$$

By Dynkin's theorem, this equality holds on the σ-algebra generated by all measurable rectangle sets, which is $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$; that is, $\mathsf{P}_Z = \mathsf{P}_X \times \mathsf{P}_Y$ on $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$. By the change-of-variable formula,

$$\begin{aligned} E[XY] &= \int_\Omega X(\omega)Y(\omega)\mathsf{P}(d\omega) \\ &= \int_{\mathbb{R}^2_+} g(z)\mathsf{P}_Z(dz), \qquad \text{(we define } g(x,y) = xy) \\ &= \int_{\mathbb{R}^2_+} g(z)(\mathsf{P}_X \times \mathsf{P}_Y)(dz) \\ &= \int_{\mathbb{R}_+} y \left\{ \int_{\mathbb{R}_+} x\mathsf{P}_X(dx) \right\} \mathsf{P}_Y(dy) \qquad \text{(by Fubini's theorem)} \\ &= E[X]E[Y] \end{aligned}$$

where in the last step we have used the change-of-variable formula again. □

References

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

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

[3] Jordan M Stoyanov. *Counterexamples in probability.* Courier Corporation, 2013.