

Lecture-03: Independence

1 Law of Total Probability

Exercise 1.1 (Countably infinite coin tosses). Consider a sequence of coin tosses, such that the sample space is $\Omega = \{H, T\}^{\mathbb{N}}$. For set of outcomes $E_n \triangleq \{\omega \in \Omega : \omega_n = H\}$, we consider an event space generated by $\mathcal{F} \triangleq \sigma(\{E_n : n \in \mathbb{N}\})$. Let \mathcal{F}_n be the event space generated by the first n coin tosses, i.e. $\mathcal{F}_n \triangleq \sigma(\{E_i : i \in [n]\})$. Let A_n be the set of outcomes corresponding to at least one head in first n outcomes $A_n \triangleq \{\omega \in \Omega : \omega_i = H \text{ for some } i \in [n]\} = \cup_{i=1}^n E_i \in \mathcal{F}$, and B_n be the set of outcomes corresponding to first head at the n th outcome $B_n \triangleq \{\omega \in \Omega : \omega_1 = \dots = \omega_{n-1} = T, \omega_n = H\} = \cap_{i=1}^{n-1} E_i^c \cap E_n \in \mathcal{F}$.

1. Show that $\mathcal{F} = \sigma(\{\mathcal{F}_n : n \in \mathbb{N}\})$.
2. Show that $\sigma(\{A_n : n \in \mathbb{N}\}) \subseteq \mathcal{F}$ and $\sigma(\{B_n : n \in \mathbb{N}\}) \subseteq \mathcal{F}$.

Theorem 1.2 (Law of total probability). For a probability space (Ω, \mathcal{F}, P) , consider a sequence of events $B \in \mathcal{F}^{\mathbb{N}}$ that partitions the sample space Ω , i.e. $B_m \cap B_n = \emptyset$ for all $m \neq n$, and $\cup_{n \in \mathbb{N}} B_n = \Omega$. Then, for any event $A \in \mathcal{F}$, we have

$$P(A) = \sum_{n \in \mathbb{N}} P(A \cap B_n).$$

Proof. We can expand any event $A \in \mathcal{F}$ in terms of any partition B of the sample space Ω as

$$A = A \cap \Omega = A \cap (\cup_{n \in \mathbb{N}} B_n) = \cup_{n \in \mathbb{N}} (A \cap B_n).$$

From the mutual disjointness of the events $B \in \mathcal{F}^{\mathbb{N}}$, it follows that the sequence $(A \cap B_n \in \mathcal{F} : n \in \mathbb{N})$ is mutually disjoint. The result follows from the countable additivity of probability of disjoint events. \square

Example 1.3 (Countably infinite coin tosses). Consider the sample space $\Omega = \{H, T\}^{\mathbb{N}}$ and event space \mathcal{F} generated by sequence $E \in \mathcal{F}^{\mathbb{N}}$ defined in Exercise 1.1. We observe that any event $A \in \mathcal{F}_n$ can be written as

$$A = \cup_{\omega \in A} \{\omega\} = \cup_{\omega \in A} \cap_{i=1}^n (\{\omega \in E_i\} \cup \{\omega \notin E_i\}).$$

2 Independence

Definition 2.1 (Independence of events). For a probability space (Ω, \mathcal{F}, P) , a family of events $A \in \mathcal{F}^I$ is said to be independent, if for any finite set $F \subseteq I$, we have

$$P(\cap_{i \in F} A_i) = \prod_{i \in F} P(A_i).$$

Remark 1. The certain event Ω and the impossible event \emptyset are always independent to every event $A \in \mathcal{F}$.

Example 2.2 (Two coin tosses). Consider two coin tosses, such that the sample space is $\Omega = \{HH, HT, TH, TT\}$, and the event space is $\mathcal{F} = \mathcal{P}(\Omega)$. It suffices to define a probability function $P : \mathcal{F} \rightarrow [0, 1]$ on the sample space. We define one such probability function P , such that

$$P(\{HH\}) = P(\{HT\}) = P(\{TH\}) = P(\{TT\}) = \frac{1}{4}.$$

Let event $E_1 \triangleq \{HH, HT\}$ and $E_2 \triangleq \{HH, TH\}$ correspond to getting a head on the first or the second toss respectively.

From the defined probability function, we obtain the probability of getting a tail on the first or the second toss is $\frac{1}{2}$, and identical to the probability of getting a head on the first or the second toss. That is, $P(E_1) = P(E_2) = \frac{1}{2}$ and the intersecting event $E_1 \cap E_2 = \{HH\}$ with the probability $P(E_1 \cap E_2) = \frac{1}{4}$. That is, for events $E_1, E_2 \in \mathcal{F}$, we have

$$P(E_1 \cap E_2) = P(E_1)P(E_2).$$

That is, events E_1 and E_2 are independent.

Example 2.3 (Countably infinite coin tosses). Consider the outcome space $\Omega = \{H, T\}^{\mathbb{N}}$ and event space \mathcal{F} generated by the sequence E defined in Exercise 1.1. We define a probability function $P : \mathcal{F} \rightarrow [0, 1]$ by $P(\cap_{i \in F} E_i) = p^{|F|}$ for any finite subset $F \subseteq \mathbb{N}$. By definition, $E \in \mathcal{F}^{\mathbb{N}}$ is a sequence of independent events. Consider $A, B \in \mathcal{F}^{\mathbb{N}}$, where $A_n \triangleq \cup_{i=1}^n E_i$ and $B_n \triangleq \cap_{i=1}^{n-1} E_i^c \cap E_n \in \mathcal{F}$ for all $n \in \mathbb{N}$. It follows that $P(A_n) = 1 - (1 - p)^n$ and $P(B_n) = p(1 - p)^{n-1}$ for $n \in \mathbb{N}$.

For any $\omega \in \Omega$, we can define the number of heads in first n trials by $k_n(\omega) \triangleq \sum_{i=1}^n \mathbb{1}_{\{H\}}(\omega_i) = \sum_{i=1}^n \mathbb{1}_{\{\omega \in E_i\}}$. For any general event $A \in \mathcal{F}_n = \sigma(\{E_i : i \in [n]\})$, we can write

$$P(A) = \sum_{\omega \in A} \prod_{i=1}^n \left[P\{\omega \in E_i\} + P\{\omega \in E_i^c\} \right] = \sum_{\omega \in A} p^{k_n(\omega)} (1 - p)^{n - k_n(\omega)}.$$

Example 2.4 (Counter example). Consider a probability space (Ω, \mathcal{F}, P) and the events $A_1, A_2, A_3 \in \mathcal{F}$. The condition $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$ is not sufficient to guarantee independence of the three events. In particular, we see that if

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3), \quad P(A_1 \cap A_2 \cap A_3^c) \neq P(A_1)P(A_2)P(A_3^c),$$

then $P(A_1 \cap A_2) = P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3^c) \neq P(A_1)P(A_2)$.

Definition 2.5. A family of collections of events $(\mathcal{A}_i \subseteq \mathcal{F} : i \in I)$ is called independent, if for any finite set $F \subseteq I$ and $A_i \in \mathcal{A}_i$ for all $i \in F$, we have

$$P(\cap_{i \in F} A_i) = \prod_{i \in F} P(A_i).$$

3 Conditional Probability

Consider N trials of a random experiment over an outcome space Ω and an event space \mathcal{F} . Let $\omega_n \in \Omega$ denote the outcome of the experiment of the n th trial. Consider two events $A, B \in \mathcal{F}$ and denote the number

of times event A and event B occurs by $N(A)$ and $N(B)$ respectively. We denote the number of times both events A and B occurred by $N(A \cap B)$. Then, we can write these numbers in terms of indicator functions as

$$N(A) = \sum_{n=1}^N \mathbb{1}_{\{\omega_n \in A\}}, \quad N(B) = \sum_{n=1}^N \mathbb{1}_{\{\omega_n \in B\}}, \quad N(A \cap B) = \sum_{n=1}^N \mathbb{1}_{\{\omega_n \in A \cap B\}}.$$

We denote the relative frequency of events $A, B, A \cap B$ in N trials by $\frac{N(A)}{N}, \frac{N(B)}{N}, \frac{N(A \cap B)}{N}$ respectively. We can find the relative frequency of events A , on the trials where B occurred as

$$\frac{\frac{N(A \cap B)}{N}}{\frac{N(B)}{N}} = \frac{N(A \cap B)}{N(B)}.$$

Inspired by the relative frequency, we define the conditional probability function conditioned on events.

Definition 3.1. Fix an event $B \in \mathcal{F}$ such that $P(B) > 0$, we can define the conditional probability $P(\cdot|B) : \mathcal{F} \rightarrow [0,1]$ of any event $A \in \mathcal{F}$ conditioned on the event B as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Lemma 3.2 (Conditional probability). For any event $B \in \mathcal{F}$ such that $P(B) > 0$, the conditional probability $P(\cdot|B) : \mathcal{F} \rightarrow [0,1]$ is a probability measure on space (Ω, \mathcal{F}) .

Proof. We will show that the conditional probability satisfies all three axioms of a probability measure.

Non-negativity: For all events $A \in \mathcal{F}$, we have $P(A|B) \geq 0$ since $P(A \cap B) \geq 0$.

σ -additivity: For an infinite sequence of mutually disjoint events $(A_i \in \mathcal{F} : i \in \mathbb{N})$ such that $A_i \cap A_j = \emptyset$ for all $i \neq j$, we have $P(\cup_{i \in \mathbb{N}} A_i | B) = \sum_{i \in \mathbb{N}} P(A_i | B)$. This follows from disjointness of the sequence $(A_i \cap B \in \mathcal{F} : i \in \mathbb{N})$.

Certainty: Since $\Omega \cap B = B$, we have $P(\Omega|B) = 1$.

□

Remark 2. For two independent events $A, B \in \mathcal{F}$ such that $P(A \cap B) > 0$, we have $P(A|B) = P(A)$ and $P(B|A) = P(B)$. If either $P(A) = 0$ or $P(B) = 0$, then $P(A \cap B) = 0$.

Remark 3. For any partition B of the sample space Ω , if $P(B_n) > 0$ for all $n \in \mathbb{N}$, then from the law of total probability and the definition of conditional probability, we have

$$P(A) = \sum_{n \in \mathbb{N}} P(A|B_n)P(B_n).$$

4 Conditional Independence

Definition 4.1 (Conditional independence of events). For a probability space (Ω, \mathcal{F}, P) , a family of events $A \in \mathcal{F}^I$ is said to be conditionally independent given an event $C \in \mathcal{F}$ such that $P(C) > 0$, if for any finite set $F \subseteq I$, we have

$$P(\cap_{i \in F} A_i | C) = \prod_{i \in F} P(A_i | C).$$

Remark 4. Let $C \in \mathcal{F}$ be an event such that $P(C) > 0$. Two events $A, B \in \mathcal{F}$ are said to be conditionally independent given event C , if

$$P(A \cap B | C) = P(A | C)P(B | C).$$

If the event $C = \Omega$, it implies that A, B are independent events.

Remark 5. Two events may be independent, but not conditionally independent and vice versa.

Example 4.2. Consider two independent events $A, B \in \mathcal{F}$ such that $P(A \cap B) > 0$ and $P(A \cup B) < 1$. Then the events A and B are not conditionally independent given $A \cup B$. To see this, we observe that

$$P(A \cap B | A \cup B) = \frac{P((A \cap B) \cap (A \cup B))}{P(A \cup B)} = \frac{P(A \cap B)}{P(A \cup B)} = \frac{P(A)P(B)}{P(A \cup B)} = P(A | A \cup B)P(B).$$

We further observe that $P(B | A \cup B) = \frac{P(B)}{P(A \cup B)} \neq P(B)$ and hence $P(A \cap B | A \cup B) \neq P(A | A \cup B)P(B | A \cup B)$.

Example 4.3. Consider two non-independent events $A, B \in \mathcal{F}$ such that $P(A) > 0$. Then the events A and B are conditionally independent given A . To see this, we observe that

$$P(A \cap B | A) = \frac{P(A \cap B)}{P(A)} = P(B | A)P(A | A).$$