Lecture-03: Independence

1 Independence

Definition 1.1 (Independence of events). For a probability space $(\Omega, \mathcal{F}, P)$, a family of events $(A_i \in \mathcal{F} : i \in I)$ 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 1.2 (Two coin tosses). Consider two coin tosses, such that the sample space is $\Omega = \{HH, HT, TH, TT\}$, and the event space is $\mathcal{F} = 2^{\Omega}$. It suffices to define a probability function $P : \mathcal{F} \to [0,1]$ 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 $A_1 \triangleq \{HH, HT\}$ and $B_2 \triangleq \{HH, TH\}$ 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, $P(A_1) = P(A_2) = \frac{1}{2}$ and the intersecting event $A_1 \cap A_2 = \{HH\}$ with the probability $P(A_1 \cap A_2) = \frac{1}{4}$. That is, for events $A_1, A_2 \in \mathcal{F}$, we have

$$P(A_1 \cap A_2) = P(A_1)P(A_2).$$

That is, events A_1 and A_2 are independent.

Example 1.3 (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 $\mathcal{F} \triangleq \sigma(\{E_n : n \in \mathbb{N}\})$. We define a probability function $P : \mathcal{F} \to [0,1]$ by $P(\cap_{i \in F} E_i) = p^{|F|}$ for any finite subset $F \subseteq \mathbb{N}$. By definition, $(E_n : n \in \mathbb{N})$ is a sequence of independent events.

We observe that 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 set of outcomes corresponding to first head at the nth outcome

$$B_n \triangleq \{\omega \in \Omega : \omega_1 = \cdots = \omega_{n-1} = T, \omega = H\} = \cap_{i=1}^{n-1} E_i^c \cap E_n \in \mathcal{F}.$$

In particular this implies that $\sigma(\{A_n : n \in \mathbb{N}\}) \subseteq \mathcal{F}$ and $\sigma(\{B_n : n \in \mathbb{N}\}) \subseteq \mathcal{F}$. We can show that $P(A_n) = 1 - (1-p)^n$ and $P(B_n) = p(1-p)^{n-1}$ for $n \in \mathbb{N}$.

Let $\mathcal{F}_n$ be the event space generate by the first n coin tosses, i.e. $\mathcal{F}_n \triangleq \sigma(\{E_i : i \in [n]\})$. Then, we can show that $\mathcal{F} = \sigma(\{\mathcal{F}_n : n \in \mathbb{N}\})$. For any $\omega \in \Omega$, we can define the number of heads in first n trials by $k_n \triangleq \sum_{i=1}^n \mathbb{1}_{\{\omega_i = H\}}$. Then, we observe that any event $A \in \mathcal{F}_n$ can be written as union of $\cap_{i=1}^n C_i$ where $C_i = E_i$ or E_i^c. That is, we can specify the first n outcomes for each $\omega \in A$. Since the probability $P(\cap_{i=1}^n C_i) = \prod_{i=1}^n P(C_i)$, we have

$$P(A) = \sum_{\omega \in A} p^{k_n(\omega)}(1-p)^{n-k_n(\omega)}.$$

Example 1.4 (Counter example). Consider a probability space $(\Omega, \mathcal{F}, P)$ 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

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

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 1.5. 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

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

2 Law of Total Probability

Theorem 2.1 (Law of total probability). *For a probability space $(\Omega, \mathcal{F}, P)$, consider a sequence of events $B = (B_n \in \mathcal{F} : n \in \mathbb{N})$ 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 $(B_n \in \mathcal{F} : n \in \mathbb{N})$, 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. □

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 nth 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

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

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 $P(\cdot|B) : \mathcal{F} \to [0,1]$ 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 $P(\cdot|B) : \mathcal{F} \to [0,1]$ is a probability measure on space $(\Omega, \mathcal{F})$.*

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

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

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 $(\Omega, \mathcal{F}, P)$, a family of events $(A_i \in \mathcal{F} : i \in I)$ 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

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