Lecture-03: Independence

1 Conditional Probability

Consider N trials of a random experiment over outcome space Ω and the event space $\mathcal{F}$. Let $X_n \in \Omega$ 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

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

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. 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 1.1 (Conditional probability). *For any event $B \in \mathcal{F}$ such that $P(B) \geqslant 0$, 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 $P(\Omega|B) = 1$.

□

2 Law of Total Probability

Theorem 2.1 (Law of total probability). *Consider a countable partition $B = (B_n : n \in \mathbb{N})$ of 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 disjointness of the sets (B_n), it follows that $(A \cap B_n \in \mathcal{F} : n \in \mathbb{N})$ is a sequence of disjoint sets. The result follows from the countable additivity of probability of disjoint sets. ☐

Remark 1. 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

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

3 Independence

Definition 3.1 (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 2. The certain event Ω and the impossible event $\emptyset$ are always independent to every event $A \in \mathcal{F}$.

Example 3.2 (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 $A = \{HH, HT\}$ and $B = \{HH, TH\}$ 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, $P(A) = P(B) = \frac{1}{2}$ and the intersecting event $A \cap B = \{HH\}$ with the probability $P(A \cap B) = \frac{1}{4}$. That is, for events $A, B \in \mathcal{F}$, we have

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

That is, events A and B are independent.

Remark 3. 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 $P(A \cap B) = 0$.

Example 3.3 (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 $P(A_1 \cap A_2) = P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3^c) \neq P(A_1)P(A_2)$.

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 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

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

Example 4.3. 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

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