Lecture 4

1 Conditional probability spaces

Definition 1.1 (Conditional probability space). *Let* $(\Omega, \mathbb{P})$ *be a probability space and let* $B \subseteq \Omega$ *be an event for which* $\mathbb{P}(B) > 0$. *Define the* conditional probability space given B *to be the probability space* $(\Omega, \mathbb{P}(\cdot \mid B))$, *where*

$$\mathbb{P}(\omega \mid B) = \begin{cases} 0 & \text{if } \omega \notin B \\ \frac{\mathbb{P}(\omega)}{\mathbb{P}(B)} & \text{if } \omega \in B \end{cases}$$

We first need to verify that the conditional probability space given B is indeed a probability space for all events B where $\mathbb{P}(B) > 0$. That is, it suffices to show that Ω is finite or countably infinite, that $\mathbb{P}(\omega \mid B) \geq 0$ for every $\omega \in \Omega$, and that

$$\sum_{\omega \in \Omega} \mathbb{P}(\omega \mid B) = 1.$$

Since $(\Omega, \mathbb{P})$ is a probability space, Ω is finite or countably infinite, and, by its definition, $\mathbb{P}(\omega \mid B) \geq 0$ for every $\omega \in \Omega$. Moreover,

$$\begin{aligned} \sum_{\omega \in \Omega} \mathbb{P}(\omega \mid B) &= \sum_{\omega \in B} \mathbb{P}(\omega \mid B) + \sum_{\omega \in \Omega \setminus B} \mathbb{P}(\omega \mid B) = \sum_{\omega \in B} \frac{\mathbb{P}(\omega)}{\mathbb{P}(B)} + \sum_{\omega \in \Omega \setminus B} 0 \\ &= \frac{1}{\mathbb{P}(B)} \cdot \sum_{\omega \in B} \mathbb{P}(\omega) = \frac{1}{\mathbb{P}(B)} \cdot \mathbb{P}(B) = 1. \end{aligned}$$

Therefore this is indeed a probability space. In particular, all previous results on probability spaces apply here as well. For example

$$\mathbb{P}(A_1 \cup A_2 \mid B) = \mathbb{P}(A_1 \mid B) + \mathbb{P}(A_2 \mid B) - \mathbb{P}(A_1 \cap A_2 \mid B).$$

Example 1: A coin with probability of 1/3 for heads is tossed twice. It is known that in at least one of the tosses the coin came out heads. What is the probability that the result of the first coin toss was heads?

Solution: Define the probability space $(\Omega, \mathbb{P})$ as follows

$$\begin{aligned} &\Omega = \{(h,h), (h,t), (t,h), (t,t)\}, \\ &\mathbb{P}((h,h)) = \frac{1}{9},\ \mathbb{P}((h,t)) = \frac{2}{9}, \\ &\mathbb{P}((t,h)) = \frac{2}{9},\ \mathbb{P}((t,t)) = \frac{4}{9}. \end{aligned}$$

Let B be the event that at least one of the coin tosses came out heads, then

$$B = \{(h,h),(h,t),(t,h)\}.$$

We would like to condition on the event B and for that we first need to verify that its probability is positive. Indeed, it holds that

$$\mathbb{P}(B) = \mathbb{P}((h,h)) + \mathbb{P}((h,t)) + \mathbb{P}((t,h)) = \frac{1}{9} + \frac{2}{9} + \frac{2}{9} = \frac{5}{9} > 0.$$

Now, the conditional probability space given B is defined to be $(\Omega, \mathbb{P}(\cdot \mid B))$, where

$$\begin{aligned}\mathbb{P}((h,h) \mid B) &= \frac{1}{5},\ \mathbb{P}((h,t) \mid B) = \frac{2}{5},\\ \mathbb{P}((t,h) \mid B) &= \frac{2}{5},\ \mathbb{P}((t,t) \mid B) = 0.\end{aligned}$$

Let A be the event that the first coin toss came out heads. Then

$$\mathbb{P}(A \mid B) = \mathbb{P}((h,h) \mid B) + \mathbb{P}((h,t) \mid B) = \frac{1}{5} + \frac{2}{5} = \frac{3}{5}.$$

In general, since the *conditional probability space* is a probability space itself, given an event A we can calculate its *conditional probability* just as we would in any other probability space. Namely, for every event A it holds that

$$\mathbb{P}(A \mid B) = \sum_{\omega \in A} \mathbb{P}(\omega \mid B).$$

Next, we state and prove several useful lemmata regarding conditional probability.

Lemma 1.2. *If A and B are events in a probability space $(\Omega, \mathbb{P})$ such that $\mathbb{P}(B) > 0$, then*

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

Proof. It holds that

$$\begin{aligned}\mathbb{P}(A \mid B) &= \sum_{\omega \in A} \mathbb{P}(\omega \mid B)\\ &= \sum_{\omega \in A \cap B} \mathbb{P}(\omega \mid B) + \sum_{\omega \in A \setminus B} \mathbb{P}(\omega \mid B)\\ &= \sum_{\omega \in A \cap B} \frac{\mathbb{P}(\omega)}{\mathbb{P}(B)}\\ &= \frac{1}{\mathbb{P}(B)} \cdot \sum_{\omega \in A \cap B} \mathbb{P}(\omega)\\ &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.\end{aligned}$$

□

Lemma 1.3. *[Rule of Multiplication] Let* $A_1, \ldots, A_n$ *be events in a probability space* $(\Omega, \mathbb{P})$ *such that* $\mathbb{P}\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$*, then*

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 \mid A_1) \cdot \mathbb{P}(A_3 \mid A_1 \cap A_2) \cdot \ldots \cdot \mathbb{P}\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right).$$

Proof. It follows by monotonicity and by our assumption that $\mathbb{P}\left(\bigcap_{i=1}^{k} A_i\right) \geq \mathbb{P}\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$ for every $1 \leq k \leq n-1$. Hence, it follows by Lemma 1.2 that

$$\begin{aligned}
&\mathbb{P}(A_1) \cdot \mathbb{P}(A_2 \mid A_1) \cdot \mathbb{P}(A_3 \mid A_1 \cap A_2) \cdot \ldots \cdot \mathbb{P}\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right) \\
&= \mathbb{P}(A_1) \cdot \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)} \cdot \frac{\mathbb{P}(A_1 \cap A_2 \cap A_3)}{\mathbb{P}(A_1 \cap A_2)} \cdot \ldots \cdot \frac{\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right)}{\mathbb{P}\left(\bigcap_{i=1}^{n-1} A_i\right)} \\
&= \mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right).
\end{aligned}$$

□

Lemma 1.4 (Law of Total Probability)**.** *If* A *and* B *are events in a probability space* $(\Omega, \mathbb{P})$ *such that* $0 < \mathbb{P}(B) < 1$*, then*

$$\mathbb{P}(A) = \mathbb{P}(B) \cdot \mathbb{P}(A \mid B) + \mathbb{P}(B^c) \cdot \mathbb{P}(A \mid B^c).$$

Proof. Observe that

$$A = (A \cap B) \cup (A \cap B^c),$$

and the events $A \cap B$ and $A \cap B^c$ are disjoint. Therefore

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(B) \cdot \mathbb{P}(A \mid B) + \mathbb{P}(B^c) \cdot \mathbb{P}(A \mid B^c),$$

where the second equality holds by the Rule of Multiplication (i.e., Lemma 1.3). □

Lemma 1.5 (Bayes' Rule)**.** *If* A *and* B *are events in a probability space* $(\Omega, \mathbb{P})$ *such that* $\mathbb{P}(A) > 0$ *and* $\mathbb{P}(B) > 0$*, then*

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B) \cdot \mathbb{P}(A \mid B)}{\mathbb{P}(A)}.$$

Proof. It holds that

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(B) \cdot \mathbb{P}(A \mid B)}{\mathbb{P}(A)},$$

where both equalities hold by Lemma 1.2. □

Example 2: A fraction $0 < p < 1$ of the population are sick with some disease, and there is a procedure used to diagnose whether or not a given person has this disease. When a healthy person is going through the procedure, he/she will be diagnosed with the disease with probability q. When a sick person is going through the procedure, he/she will be diagnosed with the disease with probability r. A person is chosen from the population uniformly at random.

1. What is the probability that the chosen person is sick?
2. What is the probability that the chosen person is sick and the procedure will diagnose him/her as having the disease?
3. What is the probability that the procedure will diagnose the chosen person as having the disease?
4. Given that the procedure diagnosed the chosen person as having the disease, what is the probability that he/she are in fact sick?

Solution: Let B be the event that the chosen person is sick and let A by the event that he/she was diagnosed with the disease. Then

$$\mathbb{P}(B) = p,\ \mathbb{P}(A \mid B) = r,\ \mathbb{P}(A \mid B^c) = q.$$

1. $\mathbb{P}(B) = p$ as was stated previously.
2. By the Multiplication rule we have
$$\mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A \mid B) = p \cdot r.$$
3. By the Law of total probability we have
$$\mathbb{P}(A) = \mathbb{P}(B) \cdot \mathbb{P}(A \mid B) + \mathbb{P}(B^c) \cdot \mathbb{P}(A \mid B^c) = p \cdot r + (1-p) \cdot q.$$
4. By Bayes' rule we have
$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B) \cdot \mathbb{P}(A \mid B)}{\mathbb{P}(A)} = \frac{p \cdot r}{p \cdot r + (1-p) \cdot q}.$$

2 Independent Events

The following definition comprises a convenient way of testing whether two events are independent of one another.

Definition 2.1. *[Independent Events] Events A and B in some probability space $(\Omega, \mathbb{P})$ are said to be* independent *if* $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

Example 3: A fair die is rolled. Let A be the event that the outcome is even, let B be the event that the outcome is at most 4, and let C be the event that the outcome is a multiple of 3. Which pairs of these 3 events are independent?
Solution:

$$\begin{array}{lll} \mathbb{P}(A) = \frac{1}{2} & \mathbb{P}(B) = \frac{2}{3} & \mathbb{P}(C) = \frac{1}{3} \\ \mathbb{P}(A \cap B) = \frac{1}{3} & \mathbb{P}(A \cap C) = \frac{1}{6} & \mathbb{P}(B \cap C) = \frac{1}{6} \\ \mathbb{P}(A) \cdot \mathbb{P}(B) = \frac{1}{3} & \mathbb{P}(A) \cdot \mathbb{P}(C) = \frac{1}{6} & \mathbb{P}(B) \cdot \mathbb{P}(C) = \frac{2}{9} \end{array}$$

Therefore A and B are independent, and A and C are independent. On the other hand, B and C are not independent.

The following claim provides an equivalent definition of independence. It is better motivated, but less practical.

Claim 2.2 (Independent Events – Equivalent Definition)**.** *Let A and B be two events in some probability space $(\Omega, \mathbb{P})$ such that $\mathbb{P}(B) > 0$. Then A and B are independent if and only if*

$$\mathbb{P}(A) = \mathbb{P}(A \mid B).$$

Proof. Since

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)},$$

it follows that

$$\mathbb{P}(A) = \mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \iff \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

□

Claim 2.3. *Let A and B be two events in some probability space $(\Omega, \mathbb{P})$. Then A and B are independent if and only if the following four equations hold*

1. $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$
2. $\mathbb{P}(A \cap B^c) = \mathbb{P}(A) \cdot \mathbb{P}(B^c)$
3. $\mathbb{P}(A^c \cap B) = \mathbb{P}(A^c) \cdot \mathbb{P}(B)$
4. $\mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c) \cdot \mathbb{P}(B^c)$

The proof of this claim is left as an exercise.