Lecture 5

0.1 Independence of More Than Two Events

Definition 0.1. *Three events A, B, and C in some probability space $(\Omega, \mathbb{P})$ are said to be* independent *if the following four equations hold:*

a. $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

b. $\mathbb{P}(A \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(C)$

c. $\mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C)$

d. $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$

Remark 0.2. *For each of the four equations in Definition 0.1, one can construct a probability space and define three events such that this equation does not hold but the other three do. This is left as an exercise.*

Below is another definition of the independence of three events; we will later show that it is equivalent to Definition 0.1.

Definition 0.3. *Three events A, B, and C in some probability space $(\Omega, \mathbb{P})$ are said to be* independent *if the following eight equations hold:*

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

Before proving the equivalence of these two definitions, we extend both of them to n events.

Definition 0.4. *[Independent Events – First definition] Events $A_1, \ldots, A_n$ are said to be* independent *if for every $\mathcal{I} \subseteq \{1, \ldots, n\}$ of size at least 2 it holds that*

$$\mathbb{P}\left(\bigcap_{i \in \mathcal{I}} A_i\right) = \prod_{i \in \mathcal{I}} \mathbb{P}(A_i)$$

Definition 0.5. *[Independent Events – Second definition] Events $A_1, \ldots, A_n$ are said to be* independent *if*

$$\mathbb{P}\left(\bigcap_{i=1}^{n} B_i\right) = \prod_{i=1}^{n} \mathbb{P}\left(B_i\right),$$

for all choices of $B_i \in \{A_i, A_i^c\}$.

Theorem 0.6. *Definition 0.4 and Definition 0.5 are equivalent.*

Proof. We will prove the assertion of the theorem for $n = 3$, while the general case follows similar lines. We first show that Definition 0.5 implies Definition 0.4. Equation d in the first definition is the same as Equation 1 in the second definition, hence there is no need to prove it. We next prove Equation a. Observe that

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

and this union is of disjoint events. Hence

$$\begin{aligned}
\mathbb{P}(A \cap B) &= \mathbb{P}(A \cap B \cap C) + \mathbb{P}(A \cap B \cap C^c) \\
&= \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C) + \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C^c) \\
&= \mathbb{P}(A) \cdot \mathbb{P}(B) \left(\mathbb{P}(C) + \mathbb{P}(C^c)\right) \\
&= \mathbb{P}(A) \cdot \mathbb{P}(B).
\end{aligned}$$

By symmetry we get that Equations b and c hold as well.

We now prove that Definition 0.4 implies Definition 0.5. As in the proof of the opposite implication, Equation d in the first definition is the same as Equation 1 in the second definition, and so there is no need to prove Equation 1. We next prove Equation 2. Using Equations a and d we obtain

$$\begin{aligned}
\mathbb{P}(A \cap B \cap C^c) + \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C) &= \mathbb{P}(A \cap B \cap C^c) + \mathbb{P}(A \cap B \cap C) \\
&= \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathbb{P}(A \cap B \cap C^c) &= \mathbb{P}(A) \cdot \mathbb{P}(B) - \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C) \\
&= \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot (1 - \mathbb{P}(C)) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C^c).
\end{aligned}$$

Equations 3 and 4 can be derived similarly. We next prove Equation 5. By Equation a the events A and B are independent according to the definition for two events. Hence, by a homework exercise, the events A and B^c are also independent. Combining this with Equation 3, which we already proved, and we get that

$$\begin{aligned}
\mathbb{P}(A \cap B^c \cap C^c) + \mathbb{P}(A) \cdot \mathbb{P}(B^c) \cdot \mathbb{P}(C) &= \mathbb{P}(A \cap B^c \cap C^c) + \mathbb{P}(A \cap B^c \cap C) \\
&= \mathbb{P}(A \cap B^c) = \mathbb{P}(A) \cdot \mathbb{P}(B^c).
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathbb{P}(A \cap B^c \cap C^c) &= \mathbb{P}(A) \cdot \mathbb{P}(B^c) - \mathbb{P}(A) \cdot \mathbb{P}(B^c) \cdot \mathbb{P}(C) \\
&= \mathbb{P}(A) \cdot \mathbb{P}(B^c) \cdot (1 - \mathbb{P}(C)) = \mathbb{P}(A) \cdot \mathbb{P}(B^c) \cdot \mathbb{P}(C^c).
\end{aligned}$$

Equations 6, 7 and 8 can be derived similarly. □

0.2 Conditional Independence

Let A, B, and C be events in some probability space $(\Omega, \mathbb{P})$ such that $\mathbb{P}(C) > 0$. One could ask whether A and B are independent in the conditional space $(\Omega, \mathbb{P}(\cdot \mid C))$, namely, if

$$\mathbb{P}(A \cap B \mid C) = \mathbb{P}(A \mid C) \cdot \mathbb{P}(B \mid C) \tag{1}$$

holds. We will present two examples, demonstrating that independence in the conditional probability space does not necessarily imply, nor is it implied by independence in the original probability space.

Example 1:

1. Toss a fair coin twice, the two tosses being independent. Let A be the event that the outcome of the first toss is heads, let B be the event that the outcome of the second toss is heads, and let C be the event that the outcome of exactly one of the two tosses is heads. Then
$$\mathbb{P}(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A) \cdot \mathbb{P}(B),$$
that is, A and B are independent in $(\Omega, \mathbb{P})$.

   On the other hand, denoting heads by 1 we have
$$\mathbb{P}(A \mid C) = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(\{(1,0)\})}{\mathbb{P}(\{(1,0),(0,1)\})} = \frac{1/4}{1/2} = \frac{1}{2}$$
and, similarly,
$$\mathbb{P}(B \mid C) = \frac{1}{2}.$$
On the other hand
$$\mathbb{P}(A \cap B \mid C) = 0.$$
We conclude that A and B are not independent in $(\Omega, \mathbb{P}(\cdot \mid C))$.

2. Toss a fair coin three times, all coin tosses being mutually independent. Let A be the event that the outcome of the first two tosses is heads, let B be the event that the outcome of the last two tosses is heads, and let C be the event that the outcome of the second toss is heads. Then
$$\mathbb{P}(A \cap B) = \frac{1}{8} \neq \frac{1}{4} \cdot \frac{1}{4} = \mathbb{P}(A) \cdot \mathbb{P}(B),$$
that is, A and B are not independent in $(\Omega, \mathbb{P})$.

   On the other hand,
$$\mathbb{P}(A \mid C) = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(A)}{\mathbb{P}(C)} = \frac{1/4}{1/2} = \frac{1}{2}$$

and, similarly,

$$\mathbb{P}\left(B \mid C\right) = \frac{1}{2}.$$

Moreover

$$\mathbb{P}\left(A \cap B \mid C\right) = \frac{\mathbb{P}\left(A \cap B \cap C\right)}{\mathbb{P}\left(C\right)} = \frac{1/8}{1/2} = \mathbb{P}\left(A \mid C\right) \cdot \mathbb{P}\left(B \mid C\right).$$

We conclude that A and B are independent in $(\Omega, \mathbb{P}\left(\cdot \mid C\right))$.

1 Random Variables

Definition 1.1 (Random Variable)**.** *Let* $(\Omega, \mathbb{P})$ *be a probability space and let* S *be a set. A* random variable *over the set* S *is a function* $X : \Omega \to S$*. Whenever* S *is not explicitly specified, we assume that* $S = \mathbb{R}$*.*

Example 2:

1. A fair die is rolled twice, the two rolls being independent. Let X be the sum of the two results. That is

$$\begin{aligned}
&\Omega = \{1,2,3,4,5,6\}^2 = \{(i,j) : 1 \le i,j \le 6\},\\
&\mathbb{P}\left((i,j)\right) = \frac{1}{36} \text{ for every } (i,j) \in \Omega,\\
&X : \Omega \to \mathbb{R},\\
&X\left((i,j)\right) = i + j.
\end{aligned}$$

2. A coin with probability 1/3 for heads is tossed 4 times, all coin tosses being mutually independent. Let Y_1 be the number of heads in the first two tosses, let Y_2 be the number of heads in the second and third tosses, and let Y_3 be the number of heads in the last two tosses. In the following we denote heads by 1 and tails by 0. Then

$$\begin{aligned}
&\Omega = \{0,1\}^4,\\
&\mathbb{P}\left((\omega_1,\omega_2,\omega_3,\omega_4)\right) = \left(\frac{1}{3}\right)^{\omega_1+\omega_2+\omega_3+\omega_4} \cdot \left(\frac{2}{3}\right)^{4-(\omega_1+\omega_2+\omega_3+\omega_4)} \quad \text{for every } (\omega_1,\omega_2,\omega_3,\omega_4) \in \Omega,\\
&Y_1, Y_2, Y_3 : \Omega \to \mathbb{R},\\
&Y_1\left((\omega_1,\omega_2,\omega_3,\omega_4)\right) = \omega_1 + \omega_2,\\
&Y_2\left((\omega_1,\omega_2,\omega_3,\omega_4)\right) = \omega_2 + \omega_3,\\
&Y_3\left((\omega_1,\omega_2,\omega_3,\omega_4)\right) = \omega_3 + \omega_4.
\end{aligned}$$

3. Let $(\Omega, \mathbb{P})$ be a probability space, and let $A \subseteq \Omega$ be an event. Then the *indicator of* A, denoted by 1_A, is defined as follows

$$1_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$