Pairwise Independence vs. Mutual Independence

Recall that we defined mutual independence last class as follows: Events $A_1, \ldots, A_n$ are \textit{mutually independent} if for every subset $I \subseteq \{1, \ldots, n\}$,

$$P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i).$$

Note that we need this property to hold for \textit{every} subset $I$. We now consider two examples where the independence holds for certain subsets of events but not for all subsets.

**Example:** Consider two tosses of an unbiased coin. Define events:

- $A$ = outcome of first flip is H.
- $B$ = outcome of second flip is H.
- $C$ = outcomes of the two flips are the same.

Since the two tosses are independent, events $A$ and $B$ are independent. Now, consider events $A$ and $C$:

$$P(C|A) = \frac{P(A,C)}{P(A)}$$

Now, if $A$ and $C$ both happen, that implies that both tosses were heads, thus, $P(A,C) = 1/4$. Since the coin is unbiased, $P(A) = 1/2$. Thus, $P(C|A) = 1/2$. Now, event $C$ occurs if either both tosses are heads or both are tails, i.e.,

$$P(C) = P(A,B) + P(A^c,B^c) = 1/4 + 1/4 = 1/2$$

Thus, $P(C|A) = P(C)$ and hence events $C$ and $A$ are independent. Similarly, we can show that events $C$ and $B$ are independent. Thus, the events $A$, $B$ and $C$ are pairwise independent.

Now we show that these three events are not mutually independent. Intuitively, if we know whether or not events $A$ and $B$ happened, we know whether $C$ happened. Mathematically,

$$P(C|A,B) = \frac{P(A,B,C)}{P(A,B)}$$

Now, $P(A,B,C)$ is the probability that both tosses are heads and equal, which is the same as the probability of both tosses being heads, which is just $P(A,B)$. Thus, $P(C|A,B) = 1$ which is not the same as $P(C)$, showing that the events are not mutually independent. Hence, pairwise independence does not imply mutual independence.

**Example:** We now provide another example, where in contrast to the previous example, the events are not pairwise independent but the condition is satisfied when $I$ includes all events. Roll two fair dice independently, and define the following events:
A: first die is 1,2,3.
B: first die is 2,3,6.
C: sum of outcomes is 9.

We have
\[ P(A \cap B \cap C) = P(\{(3,6)\}) = \frac{1}{36} \]
\[ P(A) = P(B) = \frac{1}{2} \]
\[ P(C) = P(\{(3,6), (6,3), (4,5), (5,4)\}) = \frac{4}{36} = \frac{1}{9}. \]

Hence, we have
\[ P(A \cap B \cap C) = P(A)P(B)P(C), \]
but note that the events A and B are not pairwise independent, i.e.
\[ P(A \cap B) = \frac{1}{3} \neq \frac{1}{2} \cdot \frac{1}{2} = P(A)P(B). \]

Independent random variables

In the previous section we have defined the notion of independent events. But it seems natural to extend the notion to random variables as well. When should we call two random variables independent?

Suppose \( X_1 \) and \( X_2 \) are the results of the first and second roll of a die respectively. Then it is reasonable to model the events \( X_1 = 3 \) and \( X_2 = 5 \) to be independent. But so should the events \( X_1 = 6 \) and \( X_2 = 6 \). In fact, the events \( X_1 = a \) and \( X_2 = b \) should be independent for every \( a, b \in \{1, \ldots, 6\} \). This example leads to a natural definition for two random variables.

**Definition 4.1 (independence):** Two random variables \( X \) and \( Y \) are independent if the events \( X = a \) and \( Y = b \) are independent for every value \( a \) that \( X \) can take on and for every value \( b \) that \( Y \) can take on.

The above definition generalizes to any finite set of random variables:

**Definition 4.2 (mutual independence):** Random variables \( X_1, \ldots, X_n \) are mutually independent if the events \( X_1 = a_1, X_2 = a_2, \ldots, X_n = a_n \) are mutually independent for every value \( a_i \) that \( X_i \) can take on, \( i = 1, \ldots, n \).

Note that independence of random variables drastically reduces the number of parameters needed to specify a probability model. Indeed, in a general probability model with \( n \) random variables each taking on \( m \) possible values, one needs to specific \( m^n - 1 \) numbers to determine the underlying probabilities. But if the random variables are independent, then one needs only \( n(m - 1) \) numbers!

**Examples**

1. **Coin tosses.** In the previous lecture, we studied the experiment where a fair coin was flipped three times. As before, let \( A \) be the event that all three tosses are heads. Then \( A = A_1 \cap A_2 \cap A_3 \), where \( A_i \) is the event that the \( i \)th toss comes up heads.
It seems reasonable that the tosses should remain mutually independent, even if the coin is biased, since no coin toss is affected by any of the other tosses. If the coin is biased with heads probability \( p \), this independence assumption implies

\[
P(A) = P(A_1) \times P(A_2) \times P(A_3) = p^3.
\]

As another example, let \( B \) denote the event that the first coin toss comes up tails and the next two coin tosses come up heads. Then \( B = A_1^c \cap A_2 \cap A_3 \), and these events are independent, so

\[
P(B) = P(A_1^c) \times P(A_2) \times P(A_3) = p^2(1 - p),
\]

since \( P(A_1^c) = 1 - P(A_1) = 1 - p \). More generally, the probability of any sequence of \( n \) tosses containing \( r \) heads and \( n - r \) tails is \( p^r(1 - p)^{n-r} \). The notion of independence is the key concept that enables us to assign probabilities to these outcomes.

2. **Balls and bins.** Suppose we toss \( m \) (labelled) balls at random into \( n \) bins independently. The event that the first bin is empty is same as the event that each of the \( m \) balls go into a bin other than the first one. The probability that a particular ball goes into a bin other than the first is \( \frac{n - 1}{n} \). Thus, by independence, the probability that the first bin is empty is \( \left( \frac{n - 1}{n} \right)^m \).

Now suppose that \( m = 3 \) and \( n = 3 \). We already know that the probability the first bin is empty is \( \left( \frac{2}{3} \right)^3 = \frac{8}{27} \). What is the probability of this event given that the second bin is empty? Call these events \( A, B \) respectively. To compute \( P(A|B) \) we need to figure out \( P(A \cap B) \). But \( A \cap B \) is the event that the first and second bins are both empty, i.e., all three balls fall in the third bin. So \( P(A \cap B) = \frac{1}{3} \times \frac{1}{3} = \frac{1}{27} \) (why? Hint: use independence again). Therefore,

\[
P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/27}{8/27} = \frac{1}{8}.
\]

Not surprisingly, \( \frac{1}{8} \) is quite a bit less than \( \frac{8}{27} \); knowing that bin 2 is empty makes it significantly less likely that bin 1 will also be empty.

3. **Biased coins.** Suppose we have two biased coins, with probabilities \( p \) and \( q \) of obtaining heads. We consider the experiment where we generate a sequence of length 3 by each time randomly picking one of the two coins and then flipping it. Let \( A_i \) be the event that the \( i \)th flip is a heads. Then the events \( A_1, A_2 \) and \( A_3 \) are independent. However, if we select a coin initially, and then flip that same coin 3 times, the events \( A_1, A_2 \) and \( A_3 \) are not independent anymore.