1 Dependent and Independent Events

Let A and B be events. We say that A is independent of B if $P(A|B) = P(A)$. That is, the marginal probability of A is the same as the conditional probability of A, given B. This means that the probability of A occurring is not affected by B occurring. It turns out that, in this case, B is independent of A as well. So, we just say that A and B are independent.

We say that A depends on B if $P(A|B) \neq P(A)$. That is, the marginal probability of A is not the same as the conditional probability of A, given B. This means that the probability of A occurring is affected by B occurring. It turns out that, in this case, B depends on A as well. So, we just say that A and B are dependent.

Consider these events from the card draw: A = drawing a king, B = drawing a spade, C = drawing a face card.

Events A and B are independent. If you know that you have drawn a spade, this does not change the likelihood that you have actually drawn a king. Formally, the marginal probability of drawing a king is $P(A) = 4/52$. The conditional probability that your card is a king, given that it a spade, is $P(A|B) = 1/13$, which is the same as $4/52$.

Events A and C are dependent. If you know that you have drawn a face card, it is much more likely that you have actually drawn a king than it would be ordinarily. Formally, the marginal probability of drawing a king is $P(A) = 4/52$. The conditional probability that your card is a king, given that it is a face card, is $P(A|C) = 4/12$, which is much larger than $4/52$.

Recall the events of facing a left-handed pitcher and getting a hit from the last lecture. These events are dependent because, if the pitcher is left-handed, the chances of getting a hit are different (in fact, higher) than they would be ordinarily. Formally, $P(\text{getting a hit}) = 0.3$ but $P(\text{getting a hit}|\text{pitcher is left-handed}) = 0.5$.

2 Complementary Events

Recall that an event is a subset of the sample space. The complement of an event is another event, consisting of the remaining elements of the sample space. If A is an event, we let A^c denote the complementary event. The probability of an event and the probability of its complement are related by $P(A^c) = 1 - P(A)$.

[picture]

For example, from the die roll, consider the event of rolling at least a 5. This has probability $2/5$. The complementary event is rolling less than a 5. This has probability $1 - 2/5 = 3/5$.

3 Mutually Exclusive Events

Events A and B are said to be mutually exclusive if they have no outcomes in common. Equivalently, A and B can not happen at the same time.

For example, from the die roll, consider the two events of rolling a 3 and rolling a 6. These are mutually exclusive because you can not roll a 3 and a 6 at the same time.

Complementary events are mutually exclusive. However, mutually exclusive events need not be complementary.

4 Probability of Intersection of Events

Let A and B be two events. Their intersection, denoted by $A \cap B$ or A and B is the event consisting of outcomes which are common to both A and B. That is, A and B occurs when both A and B occur simultaneously.

[picture]

The fraction of the sample space which is in A is $P(A)$. The fraction of A which is also in B is $P(B|A)$, the conditional probability that B occurs, given that A occurs. Therefore, the fraction of the sample space which is in both A and B is $P(A)P(B|A)$.

We could arrive at another result by interchanging the roles of A and B. The fraction of the sample space which is in B is $P(B)$. The fraction of B which is also in A is $P(A|B)$, the conditional probability that A occurs, given that B occurs. Therefore, the fraction of the sample space which is in both B and A is $P(B)P(A|B)$.

Evidently, we have two equivalent formulas for the probability of the intersection of A and B, $P(A \text{ and } B) = P(A)P(B|A)$ and $P(A \text{ and } B) = P(B)P(A|B)$. Sometimes $P(A|B)$ is known but $P(B|A)$ is not. Use whichever one is known.

If A and B are independent, then $P(B|A) = P(B)$ and $P(A|B) = P(A)$, so these formulas both collapse to $P(A \text{ and } B) = P(A)P(B)$.

As an example, consider a single die roll and the two events $A =$ rolling a 4 or less and $B =$ rolling a 4 or a 5. Since B consists of 2 outcomes, $P(B) = 2/6$. If B happens then the result is either a 4 or a 5, and it is equally likely to be either. Therefore, the probability that A also happens is 1/2, because A happens if the result is 4. That is, $P(A|B) = 1/2$. So, $P(A \text{ and } B) = 2/6 \cdot 1/2 = 1/6$.

Of course there is an easier way to figure this out. A includes the outcomes 1, 2, 3, and 4. B includes the outcomes 4 and 5. Only the outcome 4 is in both A and B. Therefore, $P(A \text{ and } B) = 1/6$.

As an easier example, we find the probability of rolling three dice and getting 6 on all three. Since these events are independent, we find that

$$\begin{aligned} P(\text{die 1} = 6 \text{ and die 2} = 6 \text{ and die 3} = 6) &= P(\text{die 1} = 6)P(\text{die 2} = 6)P(\text{die 3} = 6) \\ &= \frac{1}{6}\frac{1}{6}\frac{1}{6} = \frac{1}{216} \approx 0.0046 = 0.46\%. \end{aligned}$$

The probability of the intersection of several independent events is the product of the probabilities of the events.

5 Probability of Union of Events

Let A and B be two events. Their union, denoted by $A \cup B$ or A or B is the event consisting of outcomes which are in either A or B or both A and B. That is, A or B occurs when at least one of A or B occurs.

[picture]

The fraction of the sample space which is in A is $P(A)$. The fraction of the sample space which is in B is $P(B)$. If we add the probabilities of A and B, we count the probabilities of the events in both A and B twice. Therefore, $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.

If A and B are mutually exclusive, then $P(A \text{ and } B) = 0$, so this formula collapses to $P(A \text{ or } B) = P(A) + P(B)$.

As an example, we find the probability of drawing a king or a spade. $P(\text{king}) = 4/52$ and $P(\text{spade}) = 13/52$. Since there is one card which is both king and spade, $P(\text{king and spade}) = 1/52$. Therefore, $P(\text{king or spade}) = 4/52 + 13/52 - 1/52 = 16/52$.

As an easier example, we find the probability of drawing a king or a queen. This is easier because the two events are mutually exclusive. $P(\text{king}) = P(\text{queen}) = 4/52$. Therefore, $P(\text{king or queen}) = 4/52 + 4/52 = 8/52$.

6 Bayes's Theorem

In the baseball example from the last lecture, it is determined whether the pitcher will be left-handed before it is determined whether the batter will get a hit. Therefore, we typically think of the batter getting a hit depending on the pitcher being left-handed, and not vice versa. The chance that the batter gets a hit, given that the pitcher is left-handed is a prior probability. The chance that the pitcher was left-handed, given that the batter gets a hit is called a posterior probability.

Sometimes one conditional probability, $P(A|B)$, is known but the reverse conditional probability, $P(B|A)$ is desired. Typically the known probability is readily observable and thus a prior probability, and the unknown probability is posterior. The way to convert between the two is known as Bayes's Theorem:

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

Notice that we also need to know the marginal probabilities of A and B. Bayes's Theorem can be easily derived by equating the two equations for $P(A \text{ and } B)$.

As an example, suppose that we have two buckets, numbered 1 and 2, of balls. Some balls are yellow and some balls are green. Bucket 1 has 90% yellow balls and 10$ green balls, and bucket 2 has 40% yellow balls and 60% green balls. These are the prior probabilities. We can find these just by looking in the buckets. We also know that bucket 2 has four times as many balls as bucket 1. Suppose that the buckets are dumped out and a ball is chosen at random. If the ball is yellow, what is the probability that it came from bucket 1? This is a posterior probability.

We can first try to get an estimate. Notice that 90% of the balls in bucket 1 are yellow. Therefore, it might seem very likely that a yellow ball originated in bucket 1. However, even though only 40% of the balls in bucket 2 are yellow, there are many more balls in bucket 2 than there are in bucket 1. So perhaps it is equally likely that a yellow ball originated in bucket 2.

To solve this, we use Bayes's Theorem. We want to know $P(\text{bucket 1}|\text{yellow})$ and we know $P(\text{yellow}|\text{bucket 1}) = 0.9$. The formula is

$$P(\text{bucket 1})\text{yellow} = \frac{P(\text{bucket 1})P(\text{yellow}|\text{bucket 1})}{P(\text{yellow})}.$$

All that remains is to find $P(\text{bucket 1})$ and $P(\text{yellow})$.

First we find $P(\text{bucket 1})$. At first we might think that $1/4$ of the balls are in bucket 1. However, that means that $3/4$ of the balls are in bucket 2. In this case, bucket 2 only has 3 times as many balls as bucket 1 does. So, instead, $1/5$ of the balls are in bucket 1 and $4/5$ of the balls are in bucket 2. Then bucket 2 has 4 times as many balls as bucket 1 does, as desired. Therefore, $P(\text{bucket 1}) = 1/5$.

Now all that remains is to find $P(\text{yellow})$. We are not told this directly. However, we can divide the yellow balls up into two disjoint categories - one which originated in bucket 1 and one which originated in bucket 2. Now we can compute

$$P(\text{yellow and bucket 1}) = P(\text{bucket 1})P(\text{yellow}|\text{bucket 1}) = \frac{1}{5} \cdot \frac{9}{10} = \frac{9}{50} = \frac{18}{100}$$

and

$$P(\text{yellow and bucket 2}) = P(\text{bucket 2})P(\text{yellow}|\text{bucket 2}) = \frac{4}{5} \cdot \frac{4}{10} = \frac{16}{50} = \frac{32}{100}.$$

Now we can write

$$P(\text{yellow}) = P(\text{yellow and bucket 1}) + P(\text{yellow and bucket 2}) = \frac{18}{100} + \frac{32}{100} = \frac{50}{100} = \frac{1}{2}.$$

Putting everything together,

$$P(\text{bucket 1}|\text{yellow}) = \frac{\frac{1}{5}\frac{9}{10}}{\frac{1}{2}} = \frac{\frac{9}{50}}{\frac{1}{2}} = \frac{18}{50} = 36\%.$$

That is, the chance that a yellow ball originated in bucket 1 is 36%.