Recall the definition for independence

- So we can *suppose* events are independent and compute probabilities
- Or we can *test* to see if two events are independent

**Useful Facts: 4.4  Independent events**

Two events $A$ and $B$ are **independent** if and only if

$$P(A \cap B) = P(A)P(B)$$
Recall permutations

- How many different strings can we create by rearranging the letters in the word “horse”?
  - $5! = 120$
- How about the word “Illinois”?
  - $\frac{8!}{3!2!} = 3360$
Coin flips

- If I flip a coin $N$ times, how many outcomes have exactly $k$ heads?
- Think of this as a string of $(N-k)$ Ts and $k$ Hs that is $N$ long.
- Every re-arrangement of such a string is a valid run of this experiment.
- The number of such re-arrangements is “$N$ choose $k$”

\[
\binom{N}{k} = \frac{N!}{k!(N-k)!}
\]
An airline has a regular flight with 6 seats. They always sell 7 tickets for this flight. If passengers show up independently with probability $p$ what is the probability that the flight is overbooked?

Can think of each individual as making a biased coin-flip. With probability $p$ the person comes up $S$ which means they show and with probability $(1-p)$ they come up $N$ or no-show.

There’s only one way to write a string of 7 $S$’s.

So our probability is going to just be $p^7$. 

---

**Overbooking 1**

- An airline has a regular flight with 6 seats. They always sell 7 tickets for this flight. If passengers show up independently with probability $p$ what is the probability that the flight is overbooked?
- Can think of each individual as making a biased coin-flip. With probability $p$ the person comes up $S$ which means they show and with probability $(1-p)$ they come up $N$ or no-show.
- There’s only one way to write a string of 7 $S$’s.
- So our probability is going to just be $p^7$. 

An airline has a flight with 6 seats and it sells 8 tickets. Ticket holders show up independently with probability $p$. What is the probability that exactly 6 passengers show up?

Let’s think about the event “6 passengers show up”, what kinds of outcomes are in this set?

Things like SSSSSSNN, SSSSSNSN, etc
Recall this axiom

The probability of disjoint events is additive: if for all $i$ and $j$ we have $A_i \cap A_j = \emptyset$ then

$$P(\bigcup_i A_i) = \sum_i P(A_i)$$

So we can write the event “six people show up” in a way that lets us use this axiom. The set of outcomes corresponding to “six people show up”, $E$, is

$$E = \{SSSSSSNN\} \cup \{SSSSSNSN\} \cup \ldots \cup \{NNSSSSSSS\}$$

Thus

$$P(E) = P(\{SSSSSSNN\}) + P(\{SSSSSNSN\}) + \ldots + P(\{NNSSSSSSS\})$$

So we need to know the value of each term and how many terms there are above
Overbooking 2

Each disjoint event where 6 passengers shows up occurs with what probability? I.e. what is the probability of the event SSSSSSNN?

\[ p^6(1 - p)^2 \]

And how many such events are there?

\[ \binom{8}{6} = \frac{8!}{6!2!} \]

So the probability of the event that exactly six people show up is

\[ \frac{8!}{6!2!} p^6(1 - p)^2 \]
An airline has a flight with 6 seats and it sells 8 tickets. Ticket holders show up independently with probability $p$. What is the probability that more than 6 people show up?

\[
P(\text{overbooked}) = P(7S's \cup 8S's) = P(7S's) + P(8S's) = \binom{8}{7} p^7 (1 - p) + \binom{8}{8} p^8 = 8p^7 + p^8
\]
An airline has a flight with $s$ seats. They sell $t$ tickets for this flight. Each person shows up independently with probability $p$. What is the probability that $u$ passengers show up?

How many disjoint events can we think of this event as consisting of?

Each with probability

\[
\binom{t}{u} = \frac{t!}{u!(t-u)!}
\]

Giving a probability of

\[
\frac{t!}{u!(t-u)!} p^u (1-p)^{t-u}
\]
An airline has a flight with $s$ seats. They sell $t$ tickets for this flight. Each person shows up independently with probability $p$. What is the probability that too many passengers show up?

We are looking for

$$P(\{s + 1 \text{ show up}\} \cup \{s + 2 \text{ show up}\} \cup \ldots \cup \{t \text{ show up}\})$$

$$P(\{s + 1 \text{ show up}\}) + P(\{s + 2 \text{ show up}\}) + \ldots + P(\{t \text{ show up}\})$$
Or we could write this as

\[ \sum_{i=s+1}^{t} P(\{i \text{ show up}\}) \]

Or if we use our formula from the last example, we get a probability of overbooking given by

\[ \sum_{i=s+1}^{t} \frac{t!}{i!(t-i)!} p^i (1-p)^{t-i} \]
Conditional probability
Conditional probability

- Suppose we roll two dice and are interested in the probability that the sum is less than 6
- The probability of this event is $\frac{10}{36}$
- If someone tells us that one of the dice rolled was a 4, this probability goes down to $\frac{1}{6}$
- If someone tells us that instead one of the dice rolled was a 1, the probability would increase to $\frac{2}{3}$
Conditional probability

- Knowing that an event has occurred might change the probability that we compute for some other event we haven’t yet observed.
- The probability of an event $B$ given an event $A$, written $P(B \mid A)$ and called the conditional probability of $B$ given $A$ is how we capture this notion.
Conditional probability

- Since event $A$ is known to have occurred, the space of possible outcomes for the experiment, or the sample space, are only those in the event $A$.

- The experiment outcome lies in $A$ so $P(B \mid A)$ is the probability that it also lies in $B \cap A$.

- So we have

$$P(B \mid A) = cP(B \cap A)$$
Total probability

Notice that

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

And that $A \cap B$ and $A \cap B^c$ are disjoint events

Which means we can rewrite

$$P(A) = P(A \cap B) + P(A \cap B^c)$$
Conditional probability

\[ P(B|A) = cP(B \cap A) \]

Let’s figure out what \( c \) is

For the event \( B \), either it occurred or didn’t even if we only consider the case where \( A \) occurred

\[ P(B|A) + P(B^c|A) = 1 \]

Rewriting

\[
\begin{align*}
  cP(B \cap A) + cP(B^c \cap A) & = 1 \\
  c(P(B \cap A) + P(B^c \cap A)) & = 1 \\
  cP(A) & = 1
\end{align*}
\]

So we get

\[ c = \frac{1}{P(A)} \]

\[ P(B|A) = \frac{P(B \cap A)}{P(A)} \]
Conditional probability

If we mess around with our original expression a little

\[ P(B|A) = \frac{P(B \cap A)}{P(A)} \]

We get

\[ P(B|A)P(A) = P(B \cap A) \]

Or if we did this with \( P(A|B) \)

\[ P(A|B)P(B) = P(B \cap A) \]

And this allows us to write our expression for conditional probability in the following useful way

\[ P(B|A) = \frac{P(A|B)P(B)}{P(A)} \]
There are two car factories, A and B. Factory A produces 1000 cars, of which 10 are lemons. Factory B produces 2 cars, and both are lemons. They all go to your local car dealership.

If you buy a car, what is the probability that it is a lemon?

\[ P(L) = \frac{12}{1002} \]

What is the probability a car came from factory B?

\[ P(B) = \frac{2}{1002} \]
Car factories

We had $P(L) = 12/1002$ and $P(B) = 2/1002$

Suppose you bought a car that was a lemon. What is the probability it came from factory B? I.e. what is $P(B \mid L)$?

\[
P(B \mid L) = \frac{P(L \mid B)P(B)}{P(L)}
\]

\[
P(B \mid L) = \frac{(1)(2/1002)}{12/1002}
\]

So $P(B \mid L) = 1/6$
Total probability

Notice that

\[ A = (A \cap B) \cup (A \cap B^c) \]

And that \( A \cap B \) and \( A \cap B^c \) are disjoint events

Which means we can rewrite

\[ P(A) = P(A \cap B) + P(A \cap B^c) \]

Using the definition of conditional probability

\[ P(A) = P(A|B)P(B) + P(A|B^c)P(B^c) \]

More generally if some set of disjoint events “cover” \( A \), e.g.

\[ A = A \cap (\bigcup_i B_i) \]

Then

\[ P(A) = \sum_i P(A|B_i)P(B_i) \]
Conditional probability, alternate formula

Using the result from the last side

\[ P(A) = P(A|B)P(B) + P(A|B^C)P(B^C) \]

We can rewrite

\[ P(B|A) = \frac{P(A|B)P(B)}{P(A)} \]

As

\[ P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^C)P(B^C)} \]

Which is known as Bayes’ theorem
False positives

- Suppose there is a blood test for a rare disease. The disease occurs in 1 in every 100,000 people. If you have the disease, the test will say so with probability 0.95. If you do not have the disease, the test will report a false positive with probability 0.001.

- If you get a positive test result, what is the probability that you actually have the disease?
False positives

Suppose there is a blood test for a rare disease. The disease occurs in 1 in every 100,000 people. If you have the disease, the test will say so with probability 0.95. If you do not have the disease, the test will report a false positive with probability 0.001

We have a positive test result and want to know the probability we are actually sick

Let $S$ be the event we are sick and $R$ be the event we get a positive result. We want to know $P(S | R)$

$$P(S | R) = \frac{P(R | S)P(S)}{P(R)}$$

$$P(S | R) = \frac{P(R | S)P(S) + P(R | S^c)P(S^c)}{P(R | S)P(S) + P(R | S^c)P(S^c)}$$
False positives

Suppose there is a blood test for a rare disease. The disease occurs in 1 in every 100,000 people. If you have the disease, the test will say so with probability 0.95. If you do not have the disease, the test will report a false positive with probability 0.001

\[
P(S|R) = \frac{P(R|S)P(S)}{P(R|S)P(S) + P(R|S^c)P(S^c)}
\]

\[
= \frac{0.95 \times 0.00001}{0.95 \times 0.00001 + 0.001 \times 0.99999}
\]

\[
P(S|R) \approx 0.0094
\]
Independence and conditional probability

Two independent events A and B have

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

Think about how this interacts with the definition of conditional probability

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

If A and B are independent we will have

$$P(A|B) = P(A)$$

Giving us an interpretation of independence: Knowing that event B occurs tells us nothing about event A
Independence with more than two events

If we have more than two events, there are a couple of notions of independence to be mindful of.

Pairwise independence: events $A_1, A_2, \ldots, A_n$ are pairwise independent if each pair of events is independent.

Independence: events $A_1, A_2, \ldots, A_n$ are independent if

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2)\ldots P(A_n)$$

Independence is a much stronger assumption.
Cards and independence

- Draw a card from a shuffled deck, replace it, shuffle again, draw again, shuffle again draw again. So we have three cards drawn “with replacement”

- Let A be the event that card 1 and card 2 have the same suit, B be the event that card 2 and card 3 have the same suit, and C be the event that card 1 and card 3 have the same suit
3 cards, drawn with replacement

Event A: card 1 and 2 are the same suit
Event B: card 2 and 3 are the same suit
Event C: card 1 and 3 are the same suit

So A, B, and C are pairwise independent

However, if any two of the events occurred, the third has as well, so

\[ P(A \cap B \cap C) = \frac{1}{16} \]

We have

\[ P(A) = P(B) = P(C) = \frac{4(13)^2}{52^2} = \frac{1}{4} \]

And

\[ P(A \cap B) = P(B \cap C) = P(A \cap C) = \frac{4(13)^3}{52^3} = \frac{1}{16} \]

But

\[ P(A)P(B)P(C) = \frac{1}{4^3} \]

So A, B, and C are only pairwise independent but not simply independent.
Another notion we will use is that of conditional independence. We say that events $A_1, A_2, \ldots, A_n$ are conditionally independent given event B if

$$P(A_1 \cap A_2 \cap \ldots \cap A_n | B) = P(A_1 | B)P(A_2 | B)\ldots P(A_n | B)$$
We’ve seen that conditional probability can easily mislead the intuition.

In a trial, if a prosecutor has evidence $E$ against a suspect, they may try to say that the probability of the evidence given that the person is innocent is very low.

The quantity of relevance for justice to be served isn’t how likely the evidence is, but how likely innocence is given the evidence.

\[
P(I|E) = \frac{P(E|I)P(I)}{P(E)}
\]

Quite possible for $P(I|E)$ to be close to 1 even when $P(E|I)$ is small.
Monty hall problem

- Recall the setup, there are 3 doors, behind two of them are indistinguishable goats, behind one is a car. You pick a door and win what’s behind it. You prefer to win a car to a goat.

- Let’s suppose you pick a door at random and before you open it, Monty announces that he will now open a door and show you a goat from among the doors you didn’t pick.

- After the does this, should you switch doors from your original pick to the one that you didn’t pick that is still closed?
Monty hall

- Let’s call the door you picked door #1, the one the host opened door #2, and the one that you didn’t pick that is still closed door #3.
- Let $C_i$ be the event that the car is behind door $i$ and $H_j$ be the event that the host opened door $j$.
- We want to compute $P(C_1 | H_2)$ and compare it to $P(C_3 | H_2)$ to see if we should switch.
Monty hall

First we compute $P(C_1|H_2)$

\[
P(C_1|H_2) = \frac{P(H_2|C_1)P(C_1)}{P(H_2|C_1)P(C_1) + P(H_2|C_2)P(C_2) + P(H_2|C_3)P(C_3)}
\]

\[
\begin{array}{ccc}
1/2 & 1/3 & \\
1/3 & 0 & 1/3 \\
1 & 1/3 &
\end{array}
\]

$= 1/3$

Now let’s compute $P(C_3|H_2)$

\[
P(C_3|H_2) = \frac{P(H_2|C_3)P(C_3)}{P(H_2|C_1)P(C_1) + P(H_2|C_2)P(C_2) + P(H_2|C_3)P(C_3)}
\]

\[
\begin{array}{ccc}
1/2 & 1/3 & \\
1/3 & 0 & 1/3 \\
1 & 1/3 &
\end{array}
\]

$= 2/3$
Takeaway

- See the text for other ways to set up Monty Hall and why it matters
- Conditional probabilities can be quite counterintuitive