THE UNIVERSITY OF TEXAS AT AUSTIN
Department of Statistics and Data Sciences
College of Natural Sciences

SDS 321: Introduction to Probability and Statistics

Lecture 4: Bayes rule and independence

Purnamrita Sarkar
Department of Statistics and Data Science
The University of Texas at Austin

www.cs.cmu.edu/∼psarkar/teaching.html

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus (event B) iff she sells more than 10 umbrellas in a day (event W).

- Now $W = B$. We have $P(R) = 0.1$, $P(W|R) = 0.8$ and $P(W|R^c) = 0.25$.
- If you knew that Anita got a bonus, then what is the probability that it rained?

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus (event B) iff she sells more than 10 umbrellas in a day (event W).

- Now $W = B$. We have $P(R) = 0.1$, $P(W|R) = 0.8$ and $P(W|R^c) = 0.25$.
- If you knew that Anita got a bonus, then what is the probability that it rained?
- We are interested in $P(R|B)$, which is the same as $P(R|W)$.

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus (event B) iff she sells more than 10 umbrellas in a day (event W).

- Now $W = B$. We have $P(R) = 0.1$, $P(W|R) = 0.8$ and $P(W|R^c) = 0.25$.
- If you knew that Anita got a bonus, then what is the probability that it rained?
- We are interested in $P(R|B)$, which is the same as $P(R|W)$.
- First we need $P(R \cap W)$ and then we need $P(W)$.
- $P(R \cap W) = P(W|R)P(R) = 0.8 \times 0.1 = 0.08$.

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus (event B) iff she sells more than 10 umbrellas in a day (event W).

- Now $W = B$. We have $P(R) = 0.1$, $P(W|R) = 0.8$ and $P(W|R^c) = 0.25$.
- If you knew that Anita got a bonus, then what is the probability that it rained?
- We are interested in $P(R|B)$, which is the same as $P(R|W)$.
- First we need $P(R \cap W)$ and then we need $P(W)$.
- $P(R \cap W) = P(W|R)P(R) = 0.8 \times 0.1 = 0.08$.
- Now what about $P(W)$?

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus iff she sells more than 10 umbrellas in a day.

- Now what about $P(W)$?

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus iff she sells more than 10 umbrellas in a day.

- Now what about $P(W)$?
  - First write W as an union of two disjoint events. Guesses?

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus iff she sells more than 10 umbrellas in a day.

- Now what about $P(W)$?
  - First write W as an union of two disjoint events. Guesses?
  - $W = W \cap \Omega = W \cap (R \cup R^c) = (W \cap R) \cup (W \cap R^c)$.

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus iff she sells more than 10 umbrellas in a day.

- Now what about $P(W)$?
  - First write W as an union of two disjoint events. Guesses?
  - $W = W \cap \Omega = W \cap (R \cup R^c) = (W \cap R) \cup (W \cap R^c)$.
  - Now additivity gives

$$\begin{aligned} P(W) &= P(W \cap R) + P(W \cap R^c) \\ &= P(W|R)P(R) + P(W|R^c)P(R^c) = 0.08 + 0.25 \times 0.9 \approx 0.3 \end{aligned}$$

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus iff she sells more than 10 umbrellas in a day.

- Now what about $P(W)$?
  - First write W as an union of two disjoint events. Guesses?
  - $W = W \cap \Omega = W \cap (R \cup R^c) = (W \cap R) \cup (W \cap R^c)$.
  - Now additivity gives

$$\begin{aligned} P(W) &= P(W \cap R) + P(W \cap R^c) \\ &= P(W|R)P(R) + P(W|R^c)P(R^c) = 0.08 + 0.25 \times 0.9 \approx 0.3 \end{aligned}$$

$$\begin{aligned} P(R|B) = P(R|W) &= \frac{P(R \cap W)}{P(W)} \\ &= \frac{P(W|R)P(R)}{P(W|R)P(R) + P(W|R^c)P(R^c)} = 0.08/0.3 \approx 0.27 \end{aligned}$$

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus iff she sells more than 10 umbrellas in a day.

- Now what about $P(W)$?
  - First write W as an union of two disjoint events. Guesses?
  - $W = W \cap \Omega = W \cap (R \cup R^c) = (W \cap R) \cup (W \cap R^c)$.
  - Now additivity gives

$$\begin{aligned} P(W) &= P(W \cap R) + P(W \cap R^c) \qquad \text{Theorem of total probability} \\ &= P(W|R)P(R) + P(W|R^c)P(R^c) = 0.08 + 0.25 \times 0.9 \approx 0.3 \end{aligned}$$

$$\begin{aligned} P(R|B) = P(R|W) &= \frac{P(R \cap W)}{P(W)} \\ &= \frac{P(W|R)P(R)}{P(W|R)P(R) + P(W|R^c)P(R^c)} = 0.08/0.3 \approx 0.27 \end{aligned}$$

Revisit Example: Umbrella sales

Anita works for an umbrella company. She gets a bonus iff she sells more than 10 umbrellas in a day.

- Now what about $P(W)$?
  - First write W as an union of two disjoint events. Guesses?
  - $W = W \cap \Omega = W \cap (R \cup R^c) = (W \cap R) \cup (W \cap R^c)$.
  - Now additivity gives

$$
\begin{aligned}
P(W) &= P(W \cap R) + P(W \cap R^c) \qquad \text{Theorem of total probability} \\
&= P(W|R)P(R) + P(W|R^c)P(R^c) = 0.08 + 0.25 \times 0.9 \approx 0.3
\end{aligned}
$$

$$
\begin{aligned}
P(R|B) = P(R|W) &= \frac{P(R \cap W)}{P(W)} \\
&= \frac{P(W|R)P(R)}{P(W|R)P(R) + P(W|R^c)P(R^c)} = 0.08/0.3 \approx 0.27
\end{aligned}
$$

The last step is also known as **Bayes Rule**, which we will study next.

Bayes rule

- Simple rule to get conditional probability of A given B, from the conditional formula of B given A.

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\ &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \end{aligned}$$

- This is very useful for **inferring** hidden causes underlying our observations.

When do you use Bayes' rule?

- You have a homework question asking for $P(A|B)$.
- First check what is given to you.
- If you have $P(A|B)$ or $P(A^c|B)$ then you are all set.
- If not, then you have to use Bayes rule.
- For these you will need to know $P(B|A)$ and $P(B|A^c)$ and $P(A)$. If you don't have these, then either the question is incomplete, or you should read it again.

Typical Bayes rule example

- Consider testing for some latent (hidden/unobservable) disease, that won't become symptomatic until a future time point.
- We can directly observe the outcome of the test.
- Assuming the test isn't 100% accurate, we *can't* directly observe whether we have the disease.
- We have two possible **hidden causes** for a positive test result:
  - We have the disease, and the test is correct.
  - We don't have the disease, and the test is a *false positive*.
- We want to **infer** *which* hidden cause underlies our observation.

Inference: Disease testing

Let's add some numbers to this example. Let's assume:

- The disease affects 2% of the population.
- The false positive rate is 1%.
- The false negative rate is 5%.
- If you take the test and the result is positive, you are really interested in the question: Given that you tested positive, what is the chance you have the disease?

Inference: Disease testing

Let's add some numbers to this example. Let's assume:

- The disease affects 2% of the population.
- The false positive rate is 1%.
- The false negative rate is 5%.
- If you take the test and the result is positive, you are really interested in the question: Given that you tested positive, what is the chance you have the disease?
- Let T be the event "tests positive" and "D" be the event "has disease".

Inference: Disease testing

Let's add some numbers to this example. Let's assume:

- The disease affects 2% of the population.
- The false positive rate is 1%.
- The false negative rate is 5%.
- If you take the test and the result is positive, you are really interested in the question: Given that you tested positive, what is the chance you have the disease?

- Let T be the event "tests positive" and "D" be the event "has disease".
- We know that:

$$P(D) = 0.02 \qquad P(T^c|D) = 0.05 \qquad P(T|D^c) = 0.01$$

Another application of Bayes' rule

If you take the test and the result is positive, you are really interested in the question: Given that you tested positive, what is the chance you have the disease?

i.e. what is $P(D|T)$? Bayes' rule gives us:

$$P(D|T) = \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|D^c)P(D^c)}$$

It lets us get from the conditional probability of an observation given a hidden cause (which we usually know), to the conditional probability of a hidden cause given an observation (which we usually care about!)

Inference: Disease testing

So, let's plug in the numbers. Recall

$$P(D) = 0.02 \qquad P(T^c|D) = 0.05 \qquad P(T|D^c) = 0.01$$

So, $P(T|D) = 1 - 0.05 = 0.95$.

$$\begin{aligned}
P(D|T) =& \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|D^c)P(D^c)} \\
=& \frac{0.95 \times 0.02}{0.95 \times 0.02 + 0.01 \times 0.98} \\
=& \frac{0.019}{0.0288} = .66
\end{aligned}$$

More examples to apply Bayes' Rule (Coding)

Alice is sending a coded message to Bob using "dots" and "dashes", which are known to occur in the proportion of $3:4$ for Morse codes. Because of interference on the transmission line, a dot can be mistakenly received as a dash with probability $1/8$ and vice-versa. If Bob receives a "dot", what is the probability that Alice sent a "dot".

More examples to apply Bayes' Rule (Coding)

Alice is sending a coded message to Bob using "dots" and "dashes", which are known to occur in the proportion of $3:4$ for Morse codes. Because of interference on the transmission line, a dot can be mistakenly received as a dash with probability $1/8$ and vice-versa. If Bob receives a "dot", what is the probability that Alice sent a "dot".

- Notation: $\text{dotS} = \{\text{dot sent}\}$, $\text{dashS} = \{\text{dash sent}\}$, $\text{dotR} = \{\text{dot received}\}$, $\text{dashR} = \{\text{dash received}\}$.
- $P(\text{dotS}) = \frac{3}{7}$, $P(\text{dashS}) = \frac{4}{7}$ and $P(\text{dashR}|\text{dotS}) = P(\text{dotR}|\text{dashS}) = \frac{1}{8}$.

More examples to apply Bayes' Rule (Coding)

Alice is sending a coded message to Bob using "dots" and "dashes", which are known to occur in the proportion of $3:4$ for Morse codes. Because of interference on the transmission line, a dot can be mistakenly received as a dash with probability $1/8$ and vice-versa. If Bob receives a "dot", what is the probability that Alice sent a "dot".

- Notation: $\text{dotS} = \{\text{dot sent}\}$, $\text{dashS} = \{\text{dash sent}\}$, $\text{dotR} = \{\text{dot received}\}$, $\text{dashR} = \{\text{dash received}\}$.
- $P(\text{dotS}) = \frac{3}{7}$, $P(\text{dashS}) = \frac{4}{7}$ and $P(\text{dashR}|\text{dotS}) = P(\text{dotR}|\text{dashS}) = \frac{1}{8}$.
- We want:
  $P(\text{dotS}|\text{dotR})$

More examples to apply Bayes' Rule (Coding)

Alice is sending a coded message to Bob using "dots" and "dashes", which are known to occur in the proportion of $3:4$ for Morse codes. Because of interference on the transmission line, a dot can be mistakenly received as a dash with probability $1/8$ and vice-versa. If Bob receives a "dot", what is the probability that Alice sent a "dot".

- Notation: $\text{dotS} = \{\text{dot sent}\}$, $\text{dashS} = \{\text{dash sent}\}$, $\text{dotR} = \{\text{dot received}\}$, $\text{dashR} = \{\text{dash received}\}$.
- $P(\text{dotS}) = \frac{3}{7}$, $P(\text{dashS}) = \frac{4}{7}$ and $P(\text{dashR}|\text{dotS}) = P(\text{dotR}|\text{dashS}) = \frac{1}{8}$.
- We want:
  $$P(\text{dotS}|\text{dotR}) = \frac{P(\text{dotR}|\text{dotS})P(\text{dotS})}{P(\text{dotR})}$$

More examples to apply Bayes' Rule (Coding)

Alice is sending a coded message to Bob using "dots" and "dashes", which are known to occur in the proportion of $3:4$ for Morse codes. Because of interference on the transmission line, a dot can be mistakenly received as a dash with probability $1/8$ and vice-versa. If Bob receives a "dot", what is the probability that Alice sent a "dot".

- Notation: $\text{dotS} = \{\text{dot sent}\}$, $\text{dashS} = \{\text{dash sent}\}$, $\text{dotR} = \{\text{dot received}\}$, $\text{dashR} = \{\text{dash received}\}$.
- $P(\text{dotS}) = \frac{3}{7}$, $P(\text{dashS}) = \frac{4}{7}$ and $P(\text{dashR}|\text{dotS}) = P(\text{dotR}|\text{dashS}) = \frac{1}{8}$.
- We want:
$$P(\text{dotS}|\text{dotR}) = \frac{P(\text{dotR}|\text{dotS})P(\text{dotS})}{P(\text{dotR})} = \frac{(1-1/8)\times 3/7}{P(\text{dotR})} = \frac{3/8}{P(\text{dotR})}$$

More examples to apply Bayes' Rule (Coding)

Alice is sending a coded message to Bob using "dots" and "dashes", which are known to occur in the proportion of $3:4$ for Morse codes. Because of interference on the transmission line, a dot can be mistakenly received as a dash with probability $1/8$ and vice-versa. If Bob receives a "dot", what is the probability that Alice sent a "dot".

- Notation: $\text{dotS} = \{\text{dot sent}\}$, $\text{dashS} = \{\text{dash sent}\}$, $\text{dotR} = \{\text{dot received}\}$, $\text{dashR} = \{\text{dash received}\}$.
- $P(\text{dotS}) = \frac{3}{7}$, $P(\text{dashS}) = \frac{4}{7}$ and $P(\text{dashR}|\text{dotS}) = P(\text{dotR}|\text{dashS}) = \frac{1}{8}$.
- We want:
$$P(\text{dotS}|\text{dotR}) = \frac{P(\text{dotR}|\text{dotS})P(\text{dotS})}{P(\text{dotR})} = \frac{(1-1/8)\times 3/7}{P(\text{dotR})} = \frac{3/8}{P(\text{dotR})}$$
- $$P(\text{dotR}) = P(\text{dotR}|\text{dotS})P(\text{dotS}) + P(\text{dotR}|\text{dashS})P(\text{dashS})$$
$$= \frac{3}{8} + \frac{1}{8} \times \frac{4}{7} = \frac{25}{56}$$

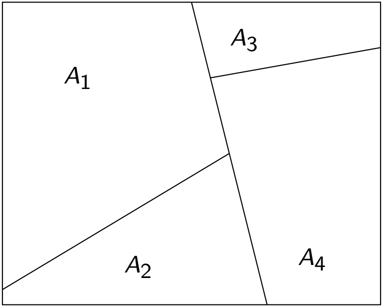
More examples to apply Bayes' Rule (Coding)

Alice is sending a coded message to Bob using "dots" and "dashes", which are known to occur in the proportion of $3:4$ for Morse codes. Because of interference on the transmission line, a dot can be mistakenly received as a dash with probability $1/8$ and vice-versa. If Bob receives a "dot", what is the probability that Alice sent a "dot".

- Notation: $\text{dotS} = \{\text{dot sent}\}$, $\text{dashS} = \{\text{dash sent}\}$, $\text{dotR} = \{\text{dot received}\}$, $\text{dashR} = \{\text{dash received}\}$.
- $P(\text{dotS}) = \frac{3}{7}$, $P(\text{dashS}) = \frac{4}{7}$ and $P(\text{dashR}|\text{dotS}) = P(\text{dotR}|\text{dashS}) = \frac{1}{8}$.
- We want:
$$P(\text{dotS}|\text{dotR}) = \frac{P(\text{dotR}|\text{dotS})P(\text{dotS})}{P(\text{dotR})} = \frac{(1-1/8)\times 3/7}{P(\text{dotR})} = \frac{3/8}{P(\text{dotR})}$$
- $$P(\text{dotR}) = P(\text{dotR}|\text{dotS})P(\text{dotS}) + P(\text{dotR}|\text{dashS})P(\text{dashS})$$
$$= \frac{3}{8} + \frac{1}{8} \times \frac{4}{7} = \frac{25}{56}$$
- So we finally have: $P(\text{dotS}|\text{dotR}) = \frac{3/8}{25/56} = \frac{21}{25}$.

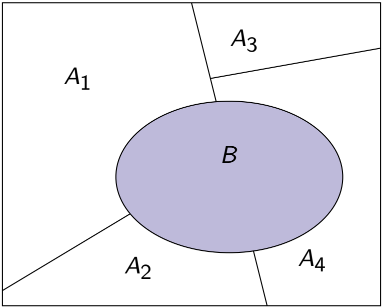
Total Probability Theorem

- Next, we're going to look at ways of obtaining the probability of a subset, using conditional probabilities.
- Let $A_1, \ldots, A_n$ be a partition of Ω, such that $P(A_i) > 0$ for all A_i.



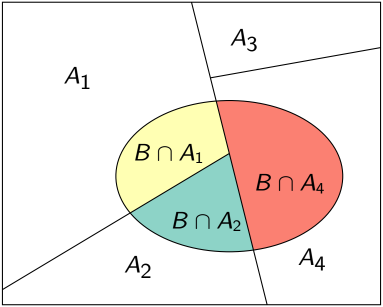
Total Probability Theorem

- Next, we're going to look at ways of obtaining the probability of a subset, using conditional probabilities.
- Let $A_1, \ldots, A_n$ be a partition of Ω, such that $P(A_i) > 0$ for all A_i.
- Let B be an event.



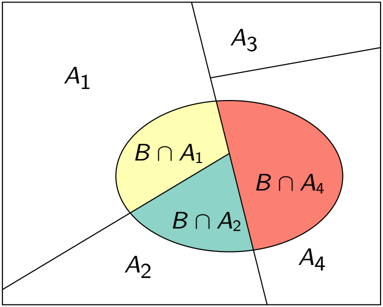
Total Probability Theorem

- Let B be an event.
- Note that $B = \cup_i (A_i \cap B)$.
- Therefore, $P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_3 \cap B)$.



Total Probability Theorem

- By the multiplication rule, $P(A_i \cap B) = P(A_i)P(B|A_i)$.
- So, $P(B) = P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n)$.
- This is known as the **Total Probability Theorem**



Bayes' Rule

Figure: Thomas Bayes, 1701-1761. English statistician, philosopher and Presbyterian minister

Bayes' rule: Let A_1, $A_2, \ldots, A_n$, be a partition of the sample space, and let B be any set. Then, for each $i = 1, 2, \ldots, n$,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_j)P(A_j)}$$

Example: The Monty Hall problem

You are a contestant on a game show, where you have to pick one of three boxes (say A, B, and C) to open. One of the three boxes contains money, the rest are empty. Assume the host knows which box contains the money.

You pick a box, say A. To build suspense, the host opens one of the other two boxes (say B) revealing it is empty. He asks, do you want to stick with your existing box or switch?

What do you do? Does it make any difference?

The Monty Hall problem

- Our outcome consists of *two* random variables: where the money is, and which box the host opens.
- We will observe which box is opened; we want to infer where the money is.
- Let's write M_A for the event "money in A", M_B for "money in B", M_C for "money in C".
- H_A for "host opens A", H_B for "host opens B", H_C for "host opens C".
- What is $P(M_A)$?

The Monty Hall problem

- Our outcome consists of *two* random variables: where the money is, and which box the host opens.
- We will observe which box is opened; we want to infer where the money is.
- Let's write M_A for the event "money in A", M_B for "money in B", M_C for "money in C".
- H_A for "host opens A", H_B for "host opens B", H_C for "host opens C".
- What is $P(M_A)$?
- $P(M_A) = P(M_B) = P(M_C) = 1/3$.

The Monty Hall problem

- You picked box A (without loss of generality).
- For every possible location of the money, we can calculate the conditional probability of the host opening a given box.
  - We assume that the host opened a box he knew to be empty.
  - We know he's not going to open box A – that's the box we picked.
- *If the money is in A, what is the probability that he opens box B?*

The Monty Hall problem

- You picked box A (without loss of generality).
- For every possible location of the money, we can calculate the conditional probability of the host opening a given box.
  - We assume that the host opened a box he knew to be empty.
  - We know he's not going to open box A – that's the box we picked.

- *If the money is in A, what is the probability that he opens box B?*
- $P(H_B|M_A) = 1/2$ (and similarly, $P(H_C|M_A) = 1/2$).
- *If the money is in B, what is the probability that he opens box B?*

The Monty Hall problem

- You picked box A (without loss of generality).
- For every possible location of the money, we can calculate the conditional probability of the host opening a given box.
  - We assume that the host opened a box he knew to be empty.
  - We know he's not going to open box A – that's the box we picked.

- *If the money is in A, what is the probability that he opens box B?*
- $P(H_B|M_A) = 1/2$ (and similarly, $P(H_C|M_A) = 1/2$).
- *If the money is in B, what is the probability that he opens box B?*
- $P(H_B|M_B) = 0$ (and similarly, $P(H_C|M_B) = 1$).
- *If the money is in C, what is the probability that he opens box B?*

The Monty Hall problem

- You picked box A (without loss of generality).
- For every possible location of the money, we can calculate the conditional probability of the host opening a given box.
  - We assume that the host opened a box he knew to be empty.
  - We know he's not going to open box A – that's the box we picked.

- *If the money is in A, what is the probability that he opens box B?*
- $P(H_B|M_A) = 1/2$ (and similarly, $P(H_C|M_A) = 1/2$).
- *If the money is in B, what is the probability that he opens box B?*
- $P(H_B|M_B) = 0$ (and similarly, $P(H_C|M_B) = 1$).
- *If the money is in C, what is the probability that he opens box B?*
- $P(H_B|M_C) = 1$ (and similarly, $P(H_C|M_B) = 0$).

The Monty Hall problem

- So, if the host opens box B, what's the probability that the money is in box C?
- By Bayes' Rule,

$$P(M_C|H_B) = \frac{P(H_B|M_C)P(M_C)}{P(H_B)}$$

The Monty Hall problem

- So, if the host opens box B, what's the probability that the money is in box C?
- By Bayes' Rule,

$$P(M_C|H_B) = \frac{P(H_B|M_C)P(M_C)}{P(H_B)}$$

- We know that $P(M_C) = 1/3$, and $P(H_B|M_C) = 1$.
- By the law of total probability,

$$\begin{aligned} P(H_B) =& P(H_B|M_A)P(M_A) + P(H_B|M_B)P(M_B) + P(H_B|M_C)P(M_C) \\ =& \frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} \\ =& \frac{1}{2} \end{aligned}$$

The Monty Hall problem

- So,

$$P(M_C|H_B) = \frac{1/3 \times 1}{1/2} = 2/3$$

- In other words, given the **partial information** that the host has opened box B, the probability that the money is in box C is 2/3.

- So, we should switch!
- What if I don't specify your choice A and the host's choice B?

Statistical Independence

We know that $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$. In general this is not the same as $P(A)$.

Statistical Independence

We know that $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$. In general this is not the same as $P(A)$.

- What if $P(A|B) = P(A)$?
- So, any evidence about B does not affect our belief about A.

Statistical Independence

We know that $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$. In general this is not the same as $P(A)$.

- What if $P(A|B) = P(A)$?
- So, any evidence about B does not affect our belief about A.
- Say two fair coins are tossed together. We say H_i for $i \in \{1, 2\}$ is the event that the i^{th} coin gives H. Similarly define T_1 and T_2.

Statistical Independence

We know that $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$. In general this is not the same as $P(A)$.

- What if $P(A|B) = P(A)$?
- So, any evidence about B does not affect our belief about A.
- Say two fair coins are tossed together. We say H_i for $i \in \{1, 2\}$ is the event that the i^{th} coin gives H. Similarly define T_1 and T_2.
  - What is $P(H_1|H_2)$?

Statistical Independence

We know that $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$. In general this is not the same as $P(A)$.

- What if $P(A|B) = P(A)$?
- So, any evidence about B does not affect our belief about A.
- Say two fair coins are tossed together. We say H_i for $i \in \{1, 2\}$ is the event that the i^{th} coin gives H. Similarly define T_1 and T_2.
  - What is $P(H_1|H_2)$?
  - $P(H_1|H_2) = \dfrac{P(H_1 \cap H_2)}{P(H_2)}$.

Statistical Independence

We know that $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$. In general this is not the same as $P(A)$.

- What if $P(A|B) = P(A)$?
- So, any evidence about B does not affect our belief about A.
- Say two fair coins are tossed together. We say H_i for $i \in \{1, 2\}$ is the event that the i^{th} coin gives H. Similarly define T_1 and T_2.
  - What is $P(H_1|H_2)$?
  - $P(H_1|H_2) = \dfrac{P(H_1 \cap H_2)}{P(H_2)} = \dfrac{1/4}{1/2} = 1/2 = P(H_1)$.
  - Knowing H_2 does not give me additional information about H_1.

Pairwise Independence

- If $P(A|B) = P(A)$, we say the events A and B are **independent**.
- In other words, knowing B tells us nothing about the probability of event A!
- We can rewrite our definition by writing $P(A|B) = P(A \cap B)/P(B)$:

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

- We generally prefer this definition... why?
- We know that $P(A \cap B) = P(B \cap A)$... so if A is independent of B, then B is independent of A.
- **Definition** Two events A and B are independent if $P(A \cap B) = P(A)P(B)$.

The gambler

A gambler is rolling 4 fair dice. What is the probability that there is at least one 6 in 4 rolls.

- Each roll is independent. Let X_i denote the event that there is no six in the i^{th} roll.
- $P(\text{at least 1 six in 4 rolls}) =$

The gambler

A gambler is rolling 4 fair dice. What is the probability that there is at least one 6 in 4 rolls.

- Each roll is independent. Let X_i denote the event that there is no six in the i^{th} roll.
- $P(\text{at least 1 six in 4 rolls}) = 1 - P(\text{no sixes in 4 rolls})$

The gambler

A gambler is rolling 4 fair dice. What is the probability that there is at least one 6 in 4 rolls.

- Each roll is independent. Let X_i denote the event that there is no six in the i^{th} roll.
- $$\begin{aligned} P(\text{at least 1 six in 4 rolls}) &= 1 - P(\text{no sixes in 4 rolls}) \\ &= 1 - P(X_1 \cap X_2 \cap X_3 \cap X_4) \end{aligned}$$

The gambler

A gambler is rolling 4 fair dice. What is the probability that there is at least one 6 in 4 rolls.

- Each roll is independent. Let X_i denote the event that there is no six in the i^{th} roll.
- $$\begin{aligned} P(\text{at least 1 six in 4 rolls}) &= 1 - P(\text{no sixes in 4 rolls}) \\ &= 1 - P(X_1 \cap X_2 \cap X_3 \cap X_4) \\ &= 1 - P(X_1)P(X_2)P(X_3)P(X_4) \end{aligned}$$

The gambler

A gambler is rolling 4 fair dice. What is the probability that there is at least one 6 in 4 rolls.

- Each roll is independent. Let X_i denote the event that there is no six in the i^{th} roll.
- $$\begin{aligned} P(\text{at least 1 six in 4 rolls}) &= 1 - P(\text{no sixes in 4 rolls}) \\ &= 1 - P(X_1 \cap X_2 \cap X_3 \cap X_4) \\ &= 1 - P(X_1)P(X_2)P(X_3)P(X_4) \\ &= 1 - \left(\frac{5}{6}\right)^4 = 0.518 \end{aligned}$$

Some ground rules

Theorem. If A and B are independent ($A \perp\!\!\!\perp B$), then so are

- A and B^c
- A^c and B
- A^c and B^c

Some ground rules

Theorem. If A and B are independent ($A \perp\!\!\!\perp B$), then so are

- A and B^c
    - $P(A \cap B^c)$
- A^c and B
- A^c and B^c

Some ground rules

Theorem. If A and B are independent ($A \perp\!\!\!\perp B$), then so are

- A and B^c
  - $P(A \cap B^c) = P(A) - P(A \cap B)$
- A^c and B
- A^c and B^c

Some ground rules

Theorem. If A and B are independent ($A \perp\!\!\!\perp B$), then so are

- A and B^c
  - $$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) \\ &= P(A) - P(A)P(B) \end{aligned}$$
- A^c and B
- A^c and B^c

Some ground rules

Theorem. If A and B are independent ($A \perp\!\!\!\perp B$), then so are

- A and B^c
  - $$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) \\ &= P(A) - P(A)P(B) \\ &= P(A)(1 - P(B)) = P(A)P(B^c) \end{aligned}$$
- A^c and B
- A^c and B^c

Mutual independence

Q: I have three events A, B and C, s.t. $P(A \cap B \cap C) = P(A)P(B)P(C)$. Are A, B, C mutually independent?

- You are tossing two fair dice. $A = \{\text{First roll is odd}\}$, $B = \{\text{First roll is} \leq 3\}$ and $C = \{\text{The sum is } 9\}$.

Mutual independence

Q: I have three events A, B and C, s.t. $P(A \cap B \cap C) = P(A)P(B)P(C)$. Are A, B, C mutually independent?

- You are tossing two fair dice. $A = \{\text{First roll is odd}\}$, $B = \{\text{First roll is} \leq 3\}$ and $C = \{\text{The sum is } 9\}$.
- $A = \{\text{First roll is } 1, 3, 5\}$. $P(A) = 1/2$.

Mutual independence

Q: I have three events A, B and C, s.t. $P(A \cap B \cap C) = P(A)P(B)P(C)$. Are A, B, C mutually independent?

- You are tossing two fair dice. $A = \{\text{First roll is odd}\}$, $B = \{\text{First roll is} \leq 3\}$ and $C = \{\text{The sum is } 9\}$.
- $A = \{\text{First roll is } 1, 3, 5\}$. $P(A) = 1/2$.
- $B = \{\text{First roll is } 1, 2, 3\}$. $P(B) = 1/2$.

Mutual independence

Q: I have three events A, B and C, s.t. $P(A \cap B \cap C) = P(A)P(B)P(C)$. Are A, B, C mutually independent?

- You are tossing two fair dice. $A = \{\text{First roll is odd}\}$, $B = \{\text{First roll is} \leq 3\}$ and $C = \{\text{The sum is } 9\}$.
- $A = \{\text{First roll is } 1, 3, 5\}$. $P(A) = 1/2$.
- $B = \{\text{First roll is } 1, 2, 3\}$. $P(B) = 1/2$.
- $C = \{(3, 6), (4, 5), (5, 4), (6, 3)\}$. $P(C) = 4/36$.

Mutual independence

Q: I have three events A, B and C, s.t. $P(A \cap B \cap C) = P(A)P(B)P(C)$. Are A, B, C mutually independent?

- You are tossing two fair dice. $A = \{\text{First roll is odd}\}$, $B = \{\text{First roll is} \leq 3\}$ and $C = \{\text{The sum is } 9\}$.
- $A = \{\text{First roll is } 1, 3, 5\}$. $P(A) = 1/2$.
- $B = \{\text{First roll is } 1, 2, 3\}$. $P(B) = 1/2$.
- $C = \{(3, 6), (4, 5), (5, 4), (6, 3)\}$. $P(C) = 4/36$.
- $A \cap B \cap C = \{(3, 6)\}$. $P(A \cap B \cap C) = 1/36$.

Mutual independence

Q: I have three events A, B and C, s.t. $P(A \cap B \cap C) = P(A)P(B)P(C)$. Are A, B, C mutually independent?

- You are tossing two fair dice. $A = \{\text{First roll is odd}\}$, $B = \{\text{First roll is} \leq 3\}$ and $C = \{\text{The sum is } 9\}$.
- $A = \{\text{First roll is } 1, 3, 5\}$. $P(A) = 1/2$.
- $B = \{\text{First roll is } 1, 2, 3\}$. $P(B) = 1/2$.
- $C = \{(3, 6), (4, 5), (5, 4), (6, 3)\}$. $P(C) = 4/36$.
- $A \cap B \cap C = \{(3, 6)\}$. $P(A \cap B \cap C) = 1/36$.
- $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Mutual independence

Q: I have three events A, B and C, s.t. $P(A \cap B \cap C) = P(A)P(B)P(C)$. Are A, B, C mutually independent?

- You are tossing two fair dice. $A = \{\text{First roll is odd}\}$, $B = \{\text{First roll is} \leq 3\}$ and $C = \{\text{The sum is } 9\}$.
- $A = \{\text{First roll is } 1, 3, 5\}$. $P(A) = 1/2$.
- $B = \{\text{First roll is } 1, 2, 3\}$. $P(B) = 1/2$.
- $C = \{(3, 6), (4, 5), (5, 4), (6, 3)\}$. $P(C) = 4/36$.
- $A \cap B \cap C = \{(3, 6)\}$. $P(A \cap B \cap C) = 1/36$.
- $P(A \cap B \cap C) = P(A)P(B)P(C)$.
- How about $P(A \cap B)$?

Mutual independence

Q: I have three events A, B and C, s.t. $P(A \cap B \cap C) = P(A)P(B)P(C)$. Are A, B, C mutually independent?

- You are tossing two fair dice. $A = \{\text{First roll is odd}\}$, $B = \{\text{First roll is} \leq 3\}$ and $C = \{\text{The sum is } 9\}$.
- $A = \{\text{First roll is } 1, 3, 5\}$. $P(A) = 1/2$.
- $B = \{\text{First roll is } 1, 2, 3\}$. $P(B) = 1/2$.
- $C = \{(3, 6), (4, 5), (5, 4), (6, 3)\}$. $P(C) = 4/36$.
- $A \cap B \cap C = \{(3, 6)\}$. $P(A \cap B \cap C) = 1/36$.
- $P(A \cap B \cap C) = P(A)P(B)P(C)$.

- How about $P(A \cap B)$?
  - $P(A \cap B) = P(\{\text{First roll is 1 or 3}\}) =$

Mutual independence

Q: I have three events A, B and C, s.t. $P(A \cap B \cap C) = P(A)P(B)P(C)$. Are A, B, C mutually independent?

- You are tossing two fair dice. $A = \{\text{First roll is odd}\}$, $B = \{\text{First roll is} \leq 3\}$ and $C = \{\text{The sum is } 9\}$.
- $A = \{\text{First roll is } 1, 3, 5\}$. $P(A) = 1/2$.
- $B = \{\text{First roll is } 1, 2, 3\}$. $P(B) = 1/2$.
- $C = \{(3, 6), (4, 5), (5, 4), (6, 3)\}$. $P(C) = 4/36$.
- $A \cap B \cap C = \{(3, 6)\}$. $P(A \cap B \cap C) = 1/36$.
- $P(A \cap B \cap C) = P(A)P(B)P(C)$.
- How about $P(A \cap B)$?
  - $P(A \cap B) = P(\{\text{First roll is 1 or 3}\}) = 12/36 = 1/3$.
  - But $P(A) \times P(B) = 1/4$. So $P(A \cap B) \neq P(A)P(B)$.

Mutual independence

$P(A \cap B \cap C) = P(A)P(B)P(C)$ **is too weak for mutual independence.**

- **Definition.** Events $A_1, \ldots A_n$ are mutually independent if for any subset S of $\{1, \ldots, n\}$ we have $P(\cap_{i \in S} A_i) = \prod_{i \in S} P(A_i)$.
- Also, any combination of a set of events and the complements of each the remaining events are mutually independent too. i.e. if A, B, C are mutually independent, then so are A^c, B^c, C and A, B^c, C or A^c, B, C^c etc.
- Mutual independence implies pairwise dependence.
- Does the converse hold?

Pairwise independence $\not\rightarrow$ mutual independence

- You are tossing two fair coins. $A = \{\text{First toss if head}\}$, $B = \{\text{Second toss is head}\}$, and $C = \{\text{The outcomes are same}\}$

Pairwise independence $\not\to$ mutual independence

- You are tossing two fair coins. $A = \{\text{First toss if head}\}$, $B = \{\text{Second toss is head}\}$, and $C = \{\text{The outcomes are same}\}$
- $A = \{\text{HH,HT}\}$, $B = \{HH, TH\}$ and $C = \{HH, TT\}$.

Pairwise independence $\not\to$ mutual independence

- You are tossing two fair coins. $A = \{\text{First toss if head}\}$, $B = \{\text{Second toss is head}\}$, and $C = \{\text{The outcomes are same}\}$
- $A = \{\text{HH,HT}\}$, $B = \{HH, TH\}$ and $C = \{HH, TT\}$.
- $P(A) = 1/2$, $P(B) = 1/2$ and $P(C) = 1/2$.

Pairwise independence $\not\to$ mutual independence

- You are tossing two fair coins. $A = \{\text{First toss if head}\}$, $B = \{\text{Second toss is head}\}$, and $C = \{\text{The outcomes are same}\}$
- $A = \{\text{HH,HT}\}$, $B = \{HH, TH\}$ and $C = \{HH, TT\}$.
- $P(A) = 1/2$, $P(B) = 1/2$ and $P(C) = 1/2$.
- $P(A \cap B) = 1/4 = P(A)P(B)$. So $A \perp\!\!\!\perp B$

Pairwise independence $\not\Rightarrow$ mutual independence

- You are tossing two fair coins. $A = \{\text{First toss if head}\}$, $B = \{\text{Second toss is head}\}$, and $C = \{\text{The outcomes are same}\}$
- $A = \{\text{HH,HT}\}$, $B = \{HH, TH\}$ and $C = \{HH, TT\}$.
- $P(A) = 1/2$, $P(B) = 1/2$ and $P(C) = 1/2$.
- $P(A \cap B) = 1/4 = P(A)P(B)$. So $A \perp\!\!\!\perp B$
- $P(C \cap B) = 1/4 = P(C)P(B)$. So $C \perp\!\!\!\perp B$

Pairwise independence $\not\to$ mutual independence

- You are tossing two fair coins. $A = \{\text{First toss if head}\}$, $B = \{\text{Second toss is head}\}$, and $C = \{\text{The outcomes are same}\}$
- $A = \{\text{HH,HT}\}$, $B = \{HH, TH\}$ and $C = \{HH, TT\}$.
- $P(A) = 1/2$, $P(B) = 1/2$ and $P(C) = 1/2$.
- $P(A \cap B) = 1/4 = P(A)P(B)$. So $A \perp\!\!\!\perp B$
- $P(C \cap B) = 1/4 = P(C)P(B)$. So $C \perp\!\!\!\perp B$
- $P(A \cap C) = 1/4 = P(A)P(C)$. So $A \perp\!\!\!\perp C$

Pairwise independence $\not\Rightarrow$ mutual independence

- You are tossing two fair coins. $A = \{\text{First toss if head}\}$, $B = \{\text{Second toss is head}\}$, and $C = \{\text{The outcomes are same}\}$
- $A = \{\text{HH,HT}\}$, $B = \{HH, TH\}$ and $C = \{HH, TT\}$.
- $P(A) = 1/2$, $P(B) = 1/2$ and $P(C) = 1/2$.
- $P(A \cap B) = 1/4 = P(A)P(B)$. So $A \perp\!\!\!\perp B$
- $P(C \cap B) = 1/4 = P(C)P(B)$. So $C \perp\!\!\!\perp B$
- $P(A \cap C) = 1/4 = P(A)P(C)$. So $A \perp\!\!\!\perp C$
- $P(A \cap B \cap C) = 1/4 \neq P(A)P(B)P(C)$

Mutual independence

Bob, Alice and Maya live in separate houses. They do grocery shopping once a week. They decide whether to shop or not independently. If I tell you Alice or Maya didn't go grocery shopping a week, is Bob more likely to go grocery shopping?

- Intuitively? If all events are independent, it shouldn't matter to Bob whether Alice or Maya are going.
- More formally: for mutually independent set of events, the occurrence or non-occurrence of any number of events from that set carries no information for the remaining events. In other words, if A, B, C, D are mutually independent, then you have pairwise independence of
  - $A \cap B$ and $C \cup D$
  - $A \cap B$ and $C \cap D^c$, etc.

Conditional Independence

Bob and Alice live in the same apartment. They do grocery shopping once a week. Given Alice didn't go grocery shopping a week, is Bob more likely to go?

- Given the event {stay in the same apartment}, it is more likely that Bob will do grocery shopping if Alice didn't go in a week.
- This brings us to conditional independence.
- **Definition** Two events A and B are conditionally independent given another event C if $P(A \cap B|C) = P(A|C)P(B|C)$
- We write this as $A \perp\!\!\!\perp B|C$

Conditional Independence

Bob, and Alice mostly go to their 9am probability class when the weather is sunny. Are the events {Bob goes to class} and {Alice goes to class} independent events?

- **No.** If I know Bob went to class. Then its likely that its sunny. This makes it likely that Alice goes too.
- Given the event {its sunny}, {Bob went to class} does not give us any information about {Alice went to class}.
- So {Bob goes to class} and {Alice goes to class} are **conditionally independent** given {its sunny}.

Conditional Independence

- Recall, we said two events A and B were independent if

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

- If $P(B) > 0$, this means that $P(A|B) = P(A)$ - knowing B tells us nothing about the probability of A.

- We can extend this definition to conditional probabilities. We say two events A and B are *conditionally* independent *given* some event C if

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

- We write this as $A \perp\!\!\!\perp B|C$.
- How do we interpret this?

Conditional Independence

- Conditional independence: $P(A \cap B|C) = P(A|C)P(B|C)$
- Intuitively, what we are thinking is, $P(A|B \cap C) = P(A|C)$. Is this true? Assume that $P(B \cap C) > 0$.

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

Conditional Independence

- Conditional independence: $P(A \cap B|C) = P(A|C)P(B|C)$
- Intuitively, what we are thinking is, $P(A|B \cap C) = P(A|C)$. Is this true? Assume that $P(B \cap C) > 0$.

$$
\begin{aligned}
P(A|B \cap C) &= \frac{P(A \cap B \cap C)}{P(B \cap C)} \\
&= \frac{P(A \cap B|C)P(C)}{P(B|C)P(C)}
\end{aligned}
$$

Conditional Independence

- Conditional independence: $P(A \cap B|C) = P(A|C)P(B|C)$
- Intuitively, what we are thinking is, $P(A|B \cap C) = P(A|C)$. Is this true? Assume that $P(B \cap C) > 0$.

$$\begin{aligned} P(A|B \cap C) &= \frac{P(A \cap B \cap C)}{P(B \cap C)} \\ &= \frac{P(A \cap B|C)P(C)}{P(B|C)P(C)} \\ &= \frac{P(A|C)P(B|C)P(C)}{P(B|C)P(C)} \end{aligned}$$

Conditional Independence

- Conditional independence: $P(A \cap B|C) = P(A|C)P(B|C)$
- Intuitively, what we are thinking is, $P(A|B \cap C) = P(A|C)$. Is this true? Assume that $P(B \cap C) > 0$.

$$
\begin{aligned}
P(A|B \cap C) &= \frac{P(A \cap B \cap C)}{P(B \cap C)} \\
&= \frac{P(A \cap B|C)P(C)}{P(B|C)P(C)} \\
&= \frac{P(A|C)P(B|C)P(C)}{P(B|C)P(C)} \\
&= P(A|C)
\end{aligned}
$$

Conditional Independence

- Conditional independence: $P(A \cap B|C) = P(A|C)P(B|C)$
- Intuitively, what we are thinking is, $P(A|B \cap C) = P(A|C)$. Is this true? Assume that $P(B \cap C) > 0$.

$$
\begin{aligned}
P(A|B \cap C) =& \frac{P(A \cap B \cap C)}{P(B \cap C)} \\
=& \frac{P(A \cap B|C)P(C)}{P(B|C)P(C)} \\
=& \frac{P(A|C)P(B|C)P(C)}{P(B|C)P(C)} \\
=& P(A|C)
\end{aligned}
$$

- So, provided $P(B \cap C) > 0$, we can write

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

Conditional Independence

- Conditional independence: $P(A \cap B|C) = P(A|C)P(B|C)$
- Intuitively, what we are thinking is, $P(A|B \cap C) = P(A|C)$. Is this true? Assume that $P(B \cap C) > 0$.

$$\begin{aligned} P(A|B \cap C) =& \frac{P(A \cap B \cap C)}{P(B \cap C)} \\ =& \frac{P(A \cap B|C)P(C)}{P(B|C)P(C)} \\ =& \frac{P(A|C)P(B|C)P(C)}{P(B|C)P(C)} \\ =& P(A|C) \end{aligned}$$

- So, provided $P(B \cap C) > 0$, we can write

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

- **Given we know C, also knowing B tells us nothing about A**

Conditional Independence: Urn example

- Consider two urns, each containing 100 balls.
- The first urn contains all red balls.
- The second urn contains all blue balls.
- We select an urn at random. Let A be the event that the first urn is chosen.
- We select a ball from the urn, note it's color, and put it back. We then select another ball from the urn, note it's color, and put it back.
- Let A be the event that the first urn was chosen, let R_1 be the event that the first ball was red, and let R_2 be the event that the second ball was red.
- *Are R_1 and R_2 independent?*

Conditional Independence: Urn example

Think about it intuitively first.

- I tell you that the first ball is red.

Conditional Independence: Urn example

Think about it intuitively first.

- I tell you that the first ball is red.
- Then you know for sure that the first urn is picked.

Conditional Independence: Urn example

Think about it intuitively first.

- I tell you that the first ball is red.
- Then you know for sure that the first urn is picked.
- So the second ball has to be red as well.

Conditional Independence: Urn example

Think about it intuitively first.

- I tell you that the first ball is red.
- Then you know for sure that the first urn is picked.
- So the second ball has to be red as well.
- Knowing about the first ball tells you a lot about the color of the second ball.

Conditional Independence: Urn example

Think about it intuitively first.

- I tell you that the first ball is red.
- Then you know for sure that the first urn is picked.
- So the second ball has to be red as well.
- Knowing about the first ball tells you a lot about the color of the second ball.
- Clearly they are not independent!

Conditional Independence: Urn example

- We need to compare $P(R_1 \cap R_2)$ with $P(R_1)P(R_2)$
- By the law of total probability,

Conditional Independence: Urn example

- We need to compare $P(R_1 \cap R_2)$ with $P(R_1)P(R_2)$
- By the law of total probability,

$$P(R_1) = P(R_1|A)P(A) + P(R_1|A^c)P(A^c)$$

Conditional Independence: Urn example

- We need to compare $P(R_1 \cap R_2)$ with $P(R_1)P(R_2)$
- By the law of total probability,

$$
\begin{aligned}
P(R_1) =& P(R_1|A)P(A) + P(R_1|A^c)P(A^c) \\
=& 1 \times 0.5 + 0 \times 0.5 = 0.5
\end{aligned}
$$

Conditional Independence: Urn example

- We need to compare $P(R_1 \cap R_2)$ with $P(R_1)P(R_2)$
- By the law of total probability,

$$
\begin{aligned}
P(R_1) =& P(R_1|A)P(A) + P(R_1|A^c)P(A^c) \\
=& 1 \times 0.5 + 0 \times 0.5 = 0.5
\end{aligned}
$$

- What is $P(R_2)$? Its also 0.5.

Conditional Independence: Urn example

- We need to compare $P(R_1 \cap R_2)$ with $P(R_1)P(R_2)$
- By the law of total probability,

$$\begin{aligned} P(R_1) =& P(R_1|A)P(A) + P(R_1|A^c)P(A^c) \\ =& 1 \times 0.5 + 0 \times 0.5 = 0.5 \end{aligned}$$

- What is $P(R_2)$? Its also 0.5.
  - Again, by the law of total probability,

$$\begin{aligned} P(R_1 \cap R_2) =& P(R_1 \cap R_2|A)P(A) + P(R_1 \cap R_2|A^c)P(A^c) \\ =& 1 \times 1 \times 0.5 + 0 \times 0 \times 0.5 = 0.5 \end{aligned}$$

  - Let's compare: $P(R_1)P(R_2) = 0.5 \times 0.5 = 0.25$

Conditional Independence: Urn example

- We need to compare $P(R_1 \cap R_2)$ with $P(R_1)P(R_2)$
- By the law of total probability,

$$
\begin{aligned}
P(R_1) =& P(R_1|A)P(A) + P(R_1|A^c)P(A^c) \\
=& 1 \times 0.5 + 0 \times 0.5 = 0.5
\end{aligned}
$$

- What is $P(R_2)$? Its also 0.5.
  - Again, by the law of total probability,

$$
\begin{aligned}
P(R_1 \cap R_2) =& P(R_1 \cap R_2|A)P(A) + P(R_1 \cap R_2|A^c)P(A^c) \\
=& 1 \times 1 \times 0.5 + 0 \times 0 \times 0.5 = 0.5
\end{aligned}
$$

  - Let's compare: $P(R_1)P(R_2) = 0.5 \times 0.5 = 0.25$
  - So, $P(R_1 \cap R_2) \neq P(R_1)P(R_2)$ – i.e. $R_1 \not\perp R_2$.

Conditional Independence: Urn example

- What if we condition on the chosen urn? Are the two colors now independent?

Conditional Independence: Urn example

- What if we condition on the chosen urn? Are the two colors now independent?
- Our gut says yes... let's just double check the math.

$$P(R_1|A) = P(R_2|A) = 1$$

$$P(R_1 \cap R_2|A) = 1 \times 1 = 1 = P(R_1|A)P(R_2|A)$$

- So, $R_1 \perp\!\!\!\perp R_2|A$
- Knowing which urn was used tells us something about how likely it is that they are both red!
- **Conditional independence does not imply independence!**

Conditional Independence: Urn example

- What if we condition on the chosen urn? Are the two colors now independent?
- Our gut says yes... let's just double check the math.

$$P(R_1|A) = P(R_2|A) = 1$$

$$P(R_1 \cap R_2|A) = 1 \times 1 = 1 = P(R_1|A)P(R_2|A)$$

- So, $R_1 \perp\!\!\!\perp R_2|A$
- Knowing which urn was used tells us something about how likely it is that they are both red!
- **Conditional independence does not imply independence!**
- What if both urns were identical?

Mutual independence $\overset{?}{\rightarrow}$ Conditional Independence

We know that two events which are conditionally independent given another event, need not be independent.

- I toss a fair coin twice. $H_i = \{i^{th}\text{toss is H}\}$. $H_1 \perp\!\!\!\perp H_2$.
- Now I tell you that the two tosses have different outcomes. Call this event E. Is $H_1 \perp\!\!\!\perp H_2 | E$ true?

Mutual independence $\stackrel{?}{\rightarrow}$ Conditional Independence

We know that two events which are conditionally independent given another event, need not be independent.

- I toss a fair coin twice. $H_i = \{i^{th}\text{toss is H}\}$. $H_1 \perp\!\!\!\perp H_2$.
- Now I tell you that the two tosses have different outcomes. Call this event E. Is $H_1 \perp\!\!\!\perp H_2 | E$ true?
- Conditioned on H_1 and E, you know that the tosses are different and the first toss is a H. Does this tell you anything about the second toss?

Mutual independence $\stackrel{?}{\rightarrow}$ Conditional Independence

We know that two events which are conditionally independent given another event, need not be independent.

- I toss a fair coin twice. $H_i = \{i^{th}\text{toss is H}\}$. $H_1 \perp\!\!\!\perp H_2$.
- Now I tell you that the two tosses have different outcomes. Call this event E. Is $H_1 \perp\!\!\!\perp H_2|E$ true?
- Conditioned on H_1 and E, you know that the tosses are different and the first toss is a H. Does this tell you anything about the second toss?
- Of course! The second toss **has to** be a T! So intuitively, H_1 and H_2 should not be independent given E.

Mutual independence does not imply Conditional Independence

- Consider an experiment involving the family names of two randomly selected people.
- It makes sense to assume that the names are independent. e.g. (First family name is Smith) $\perp\!\!\!\perp$ (Second family name is Smith)

Mutual independence does not imply Conditional Independence

- Consider an experiment involving the family names of two randomly selected people.
- It makes sense to assume that the names are independent. e.g. (First family name is Smith) $\perp\!\!\!\perp$ (Second family name is Smith)
- What if we then find out the two people are brothers? Are the probabilities still independent?

Mutual independence does not imply Conditional Independence

- Consider an experiment involving the family names of two randomly selected people.
- It makes sense to assume that the names are independent. e.g. (First family name is Smith) $\perp\!\!\!\perp$ (Second family name is Smith)
- What if we then find out the two people are brothers? Are the probabilities still independent?
- (First person is Smith) $\not\!\perp\!\!\!\perp$ (Second person is Smith)$|$brothers
- **Mutual independence does not imply conditional independence!**

Announcements

- HW2 is now available.
- Please turn in HW1.
- Today we will practice some problems in class!

Practice problems

Q1. Bob and Alice are playing a game. They alternatively keep tossing a coin and the first one to get a H wins 5$. The coin is fair.

- What is the sample space? *Hint: the game stops the moment someone gets a head.*
- What are the associated probabilities of the elements in the sample space? Do they sum to one?
- Alice insists she should toss first. Why?

Q2. A, B, C, D are mutually independent events. Prove that $A \cap B \perp\!\!\!\perp C \cup D$. You can assume that A, B, C^c, D^c are also mutually independent.

Practice problem: Alice and Bob play a game

Bob and Alice are playing a game. They alternatively keep tossing a coin and the first one to get a H wins 5$. The coin is fair.

- Hint 1: the game stops the moment someone gets a head.

Practice problem: Alice and Bob play a game

Bob and Alice are playing a game. They alternatively keep tossing a coin and the first one to get a H wins 5$. The coin is fair.

- Hint 1: the game stops the moment someone gets a head.
- All outcomes have exactly one H, and stops with a H
- $\Omega = \{H, TH, TTH, TTTH, \dots\}$. Its countably infinite.
- What are the probabilities of these events?
  - $P(H) = 1/2$. $P(TH) = (1/2)^2 . P(TTH) = (1/2)^3, \dots$
  - But do they sum to one? $\sum_{i=1}^{\infty} (1/2)^i = 1$.
- Alice insists she should toss first. Why?

Practice problem: solution

Bob and Alice are playing a game. They alternatively keep tossing a coin and the first one to get a H wins 5\$. What is the sample space? The coin is fair.

- Who should play first? Does the person who plays first have a better chance at winning?

Practice problem: solution

Bob and Alice are playing a game. They alternatively keep tossing a coin and the first one to get a H wins 5$. What is the sample space? The coin is fair.

- Who should play first? Does the person who plays first have a better chance at winning?
- When does the first person win?
  - The sequences are $H, TTH, TTTTH \ldots$
  - The probability that the first person wins is
    $$(1/2) + (1/2)^3 + (1/2)^5 + \ldots = 1/2 \sum_{i=1}^{\infty} (1/4)^i = 2/3.$$
  - So 2 out of 3 times this game is played, Alice will win.
  - We used $\sum_{i=0}^{\infty} p^i = 1/(1-p)$ for some $p < 1$.

Practice problem: Mutual independence. Roadmap

A, B, C, D are mutually independent events. Prove that $A \cap B \perp\!\!\!\perp C \cup D$.

- Well, I just want to show that $P((A \cap B) \cap (C \cup D)) = P(A \cap B)P(C \cup D)$.
- What do I have? A, B, C, D are mutually independent events. We can also show that any combination of a set of events and the complements of each the remaining events are mutually independent too. i.e. A, B, C^c, D^c are mutually independent.
- I should somehow bring things down to an intersection of events (or their complements). Then we will be in business!

Practice problem: Mutual independence

A, B, C, D are mutually independent events. Prove that $A \cap B \perp\!\!\!\perp C \cup D$.

- Well, I just want to show that $P((A \cap B) \cap (C \cup D)) = P(A \cap B)P(C \cup D)$.

Practice problem: Mutual independence

A, B, C, D are mutually independent events. Prove that $A \cap B \perp\!\!\!\perp C \cup D$.

- Well, I just want to show that $P((A \cap B) \cap (C \cup D)) = P(A \cap B)P(C \cup D)$.
- $$\begin{aligned} P((A \cap B) \cap (C \cup D)) &= P(A \cap B) - P((A \cap B) \cap (C \cup D)^c) \\ \text{(Use De-morgan's laws)} \quad &= P(A \cap B) - P(A \cap B \cap C^c \cap D^c) \\ &= P(A \cap B) - P(A)P(B)P(C^c)P(D^c) \\ &= P(A \cap B) - P(A \cap B)P(C^c \cap D^c) \\ &= P(A \cap B)(1 - P(C^c \cap D^c)) \\ &= P(A \cap B)P((C^c \cap D^c)^c) \\ \text{(De-morgan's law again!)} \quad &= P(A \cap B)P(C \cup D) \end{aligned}$$

Practice problem: Mutual independence does not imply Conditional Independence

We know that two events can be conditionally independent given another event, but they may not be independent.

- I toss a fair coin twice. $H_i = \{i^{th}\text{toss is H}\}$. $H_1 \perp\!\!\!\perp H_2$.
- Now I tell you that the two tossed have different outcomes. Call this event E. Is $H_1 \perp\!\!\!\perp H_2 | E$ true?

Practice problem: Mutual independence does not imply Conditional Independence

We know that two events can be conditionally independent given another event, but they may not be independent.

- I toss a fair coin twice. $H_i = \{i^{th}\text{toss is H}\}$. $H_1 \perp\!\!\!\perp H_2$.
- Now I tell you that the two tossed have different outcomes. Call this event E. Is $H_1 \perp\!\!\!\perp H_2|E$ true?
- $P(H_1) = P(H_2) = 1/2$ and $P(E) = 1/2$. How about $P(H_1 \cap H_2|E)$?

Practice problem: Mutual independence does not imply Conditional Independence

We know that two events can be conditionally independent given another event, but they may not be independent.

- I toss a fair coin twice. $H_i = \{i^{th}\text{toss is H}\}$. $H_1 \perp\!\!\!\perp H_2$.
- Now I tell you that the two tossed have different outcomes. Call this event E. Is $H_1 \perp\!\!\!\perp H_2 | E$ true?
- $P(H_1) = P(H_2) = 1/2$ and $P(E) = 1/2$. How about $P(H_1 \cap H_2 | E)$?
- $P(H_1 \cap H_2 | E) = P(H_1 \cap H_2 \cap E)/P(E) = 0$.

Practice problem: Mutual independence does not imply Conditional Independence

We know that two events can be conditionally independent given another event, but they may not be independent.

- I toss a fair coin twice. $H_i = \{i^{th}\text{toss is H}\}$. $H_1 \perp\!\!\!\perp H_2$.
- Now I tell you that the two tossed have different outcomes. Call this event E. Is $H_1 \perp\!\!\!\perp H_2 | E$ true?
- $P(H_1) = P(H_2) = 1/2$ and $P(E) = 1/2$. How about $P(H_1 \cap H_2 | E)$?
- $P(H_1 \cap H_2 | E) = P(H_1 \cap H_2 \cap E)/P(E) = 0$.
- What is $P(H_1|E) =?$. This is $P(H_1 \cap E)/P(E) = P(\{(HT)\})/(1/2) = 1/2$.
- Similarly, $P(H_2|E) = 1/2$.

Practice problem: Mutual independence does not imply Conditional Independence

We know that two events can be conditionally independent given another event, but they may not be independent.

- I toss a fair coin twice. $H_i = \{i^{th}\text{toss is H}\}$. $H_1 \perp\!\!\!\perp H_2$.
- Now I tell you that the two tossed have different outcomes. Call this event E. Is $H_1 \perp\!\!\!\perp H_2|E$ true?
- $P(H_1) = P(H_2) = 1/2$ and $P(E) = 1/2$. How about $P(H_1 \cap H_2|E)$?
- $P(H_1 \cap H_2|E) = P(H_1 \cap H_2 \cap E)/P(E) = 0$.
- What is $P(H_1|E) =?$. This is $P(H_1 \cap E)/P(E) = P(\{(HT)\})/(1/2) = 1/2$.
- Similarly, $P(H_2|E) = 1/2$.
- $P(H_1 \cap H_2|E) \neq P(H_1|E)P(H_2|E)$
- So mutual independence does not necessarily imply conditional independence.