CS 13

Mathematical Foundations of Computing

Probability

Some Code

Consider the following piece of code:

```
while FlipCoin() ≠ HEADS:
    print "Hello!"
```

Consider the following piece of code:

```
while FlipCoin() ≠ HEADS:
    print "Hello!"
```

Does it terminate?

Consider the following piece of code:

```
while FlipCoin() ≠ HEADS:
    print "Hello!"
```

Does it terminate?

Despite the answer seeming like an **obvious** "almost certainly", we don't really have the tools to give a real answer...

Consider the following piece of code:

```
while FlipCoin() ≠ HEADS:
    print "Hello!"
```

Does it terminate?

Despite the answer seeming like an **obvious** "almost certainly", we don't really have the tools to give a real answer...

That's what today is about!

Some Code

Consider the following piece of code:

```
while FlipCoin() ≠ HEADS:
    print "Hello!"
```

Does it terminate?

Despite the answer seeming like an **obvious** "almost certainly", we don't really have the tools to give a real answer...

That's what today is about!

To drive the point home, we are going to only use probability as a way of **analyzing code with randomized components**.

We will always assume we have the following random primitives:

- **FlipCoin**(p) returns HEADS with probability p and TAILS otherwise.
- **RollDie**(N) returns $x \in [N]$ with probability $1/N$.

Suppose we have a piece of random code $\mathcal{R}$. We want to be able to reason about $\mathcal{R}$ formally

Suppose we have a piece of random code $\mathcal{R}$. We want to be able to reason about $\mathcal{R}$ formally; so, we use the following definitions:

Definition (Outcome)

Suppose we have a piece of random code $\mathcal{R}$. We want to be able to reason about $\mathcal{R}$ formally; so, we use the following definitions:

Definition (Outcome)

An **outcome** for $\mathcal{R}$ is a sequence of values for all random calls in $\mathcal{R}$.

Suppose we have a piece of random code $\mathcal{R}$. We want to be able to reason about $\mathcal{R}$ formally; so, we use the following definitions:

Definition (Outcome)

An **outcome** for $\mathcal{R}$ is a sequence of values for all random calls in $\mathcal{R}$. For example, if our code is

```
c = FlipCoin(1/2)
d = RollDie(2)

if c == HEADS: c = 1
else: c = 2

if c == d: print c
else: print "Failure!"
```

Suppose we have a piece of random code $\mathcal{R}$. We want to be able to reason about $\mathcal{R}$ formally; so, we use the following definitions:

Definition (Outcome)

An **outcome** for $\mathcal{R}$ is a sequence of values for all random calls in $\mathcal{R}$. For example, if our code is

```
c = FlipCoin(1/2)
d = RollDie(2)

if c == HEADS: c = 1
else: c = 2

if c == d: print c
else: print "Failure!"
```

The possible outcomes are:

Definitions

Suppose we have a piece of random code $\mathcal{R}$. We want to be able to reason about $\mathcal{R}$ formally; so, we use the following definitions:

Definition (Outcome)

An **outcome** for $\mathcal{R}$ is a sequence of values for all random calls in $\mathcal{R}$. For example, if our code is

```
c = FlipCoin(1/2)
d = RollDie(2)

if c == HEADS: c = 1
else: c = 2

if c == d: print c
else: print "Failure!"
```

The possible outcomes are:

$$(\texttt{HEADS}, 1), (\texttt{HEADS}, 2), (\texttt{TAILS}, 1), (\texttt{TAILS}, 2)$$

Suppose we have a piece of random code $\mathcal{R}$. We want to be able to reason about $\mathcal{R}$ formally; so, we use the following definitions:

Definition (Outcome)

An **outcome** for $\mathcal{R}$ is a sequence of values for all random calls in $\mathcal{R}$.

Definition (Sample Space)

More Definitions

Suppose we have a piece of random code $\mathcal{R}$. We want to be able to reason about $\mathcal{R}$ formally; so, we use the following definitions:

Definition (Outcome)

An **outcome** for $\mathcal{R}$ is a sequence of values for all random calls in $\mathcal{R}$.

Definition (Sample Space)

The **Sample Space** of $\mathcal{R}$ is the set of all possible outcomes of running $\mathcal{R}$.

Definition (Event)

An **event** is a subset of the sample space.

Definition (Event)

An **event** is a subset of the sample space. We call these events, because the represent things that can happen.

Definition (Event)

An **event** is a subset of the sample space. We call these events, because the represent things that can happen. Consider the code from before:

```
c = FlipCoin(1/2)
d = RollDie(2)

if c == HEADS: c = 1
else: c = 2

if c == d: print c
else: print "Failure!"
```

Here's some events:

- The code prints "Failure!"

Even More Definitions

Definition (Event)

An **event** is a subset of the sample space. We call these events, because the represent things that can happen. Consider the code from before:

```
c = FlipCoin(1/2)
d = RollDie(2)

if c == HEADS: c = 1
else: c = 2

if c == d: print c
else: print "Failure!"
```

Here's some events:

- The code prints "Failure!"
  $\{(\texttt{HEADS}, 2), (\texttt{TAILS}, 1)\}$

Even More Definitions

Definition (Event)

An **event** is a subset of the sample space. We call these events, because the represent things that can happen. Consider the code from before:

```
c = FlipCoin(1/2)
d = RollDie(2)

if c == HEADS: c = 1
else: c = 2

if c == d: print c
else: print "Failure!"
```

Here's some events:

- The code prints "Failure!"
  $\{(\texttt{HEADS}, 2), (\texttt{TAILS}, 1)\}$
- The code halts.

Even More Definitions

Definition (Event)

An **event** is a subset of the sample space. We call these events, because the represent things that can happen. Consider the code from before:

```
c = FlipCoin(1/2)
d = RollDie(2)

if c == HEADS: c = 1
else: c = 2

if c == d: print c
else: print "Failure!"
```

Here's some events:

- The code prints "Failure!"
  $\{(\texttt{HEADS}, 2), (\texttt{TAILS}, 1)\}$
- The code halts.
  $\{(\texttt{HEADS}, 1), (\texttt{HEADS}, 2), (\texttt{TAILS}, 1), (\texttt{TAILS}, 2)\}$

Definition (Event)

An **event** is a subset of the sample space. We call these events, because the represent things that can happen. Consider the code from before:

```
c = FlipCoin(1/2)
d = RollDie(2)

if c == HEADS: c = 1
else: c = 2

if c == d: print c
else: print "Failure!"
```

Here's some events:

- The code prints "Failure!"
  $\{(\texttt{HEADS}, 2), (\texttt{TAILS}, 1)\}$
- The code halts.
  $\{(\texttt{HEADS}, 1), (\texttt{HEADS}, 2), (\texttt{TAILS}, 1), (\texttt{TAILS}, 2)\}$
- The coin flip gives HEADS.

Definition (Event)

An **event** is a subset of the sample space. We call these events, because the represent things that can happen. Consider the code from before:

```
c = FlipCoin(1/2)
d = RollDie(2)

if c == HEADS: c = 1
else: c = 2

if c == d: print c
else: print "Failure!"
```

Here's some events:

- The code prints "Failure!"
  $\{(\mathtt{HEADS},2),(\mathtt{TAILS},1)\}$
- The code halts.
  $\{(\mathtt{HEADS},1),(\mathtt{HEADS},2),(\mathtt{TAILS},1),(\mathtt{TAILS},2)\}$
- The coin flip gives HEADS.
  $\{(\mathtt{HEADS},1),(\mathtt{HEADS},2)\}$

Definition (Probability)

Let S be a sample space and $E \subseteq S$ be an event. Then, we say $\Pr(E)$ is the probability of E.

Definition (Probability)

Let S be a sample space and $E \subseteq S$ be an event. Then, we say $\Pr(E)$ is the probability of E.

Some facts about $\Pr(E)$:

Definition (Probability)

Let S be a sample space and $E \subseteq S$ be an event. Then, we say $\Pr(E)$ is the probability of E.

Some facts about $\Pr(E)$:

- $\Pr(S) = 1$

Definition (Probability)

Let S be a sample space and $E \subseteq S$ be an event. Then, we say $\Pr(E)$ is the probability of E.
Some facts about $\Pr(E)$:

- $\Pr(S) = 1$
- $0 \leq \Pr(E) \leq 1$

Definition (Probability)

Let S be a sample space and $E \subseteq S$ be an event. Then, we say $\Pr(E)$ is the probability of E.
Some facts about $\Pr(E)$:

- $\Pr(S) = 1$
- $0 \leq \Pr(E) \leq 1$
- If $E_1, E_2, \ldots, E_n \subseteq S$ and are pairwise disjoint, then

$$\Pr(E_1 \cup E_2 \cup \cdots \cup E_n) = \sum_{i=1}^{n} \Pr(E_i)$$

Definition (Probability)

Let S be a sample space and $E \subseteq S$ be an event. Then, we say $\Pr(E)$ is the probability of E.
Some facts about $\Pr(E)$:

- $\Pr(S) = 1$
- $0 \leq \Pr(E) \leq 1$
- If $E_1, E_2, \ldots, E_n \subseteq S$ and are pairwise disjoint, then

$$\Pr(E_1 \cup E_2 \cup \cdots \cup E_n) = \sum_{i=1}^{n} \Pr(E_i)$$

Okay, enough of that. Let's do some problems!

One of the TAs has a gambling problem. They make all of the following bets with Adam:

- I bet a coin will come up `HEADS`.
- I bet a six-sided die will be even.
- I bet if each of us rolls a four-sided die, the sum will be 4.
- I bet if I roll a four-sided die, and you roll a die with that many sides, the result will be even.

Which of these bets has a more than 50% chance of paying off?

One of the TAs has a gambling problem. They make all of the following bets with Adam:

- I bet a coin will come up `HEADS`.
- I bet a six-sided die will be even.
- I bet if each of us rolls a four-sided die, the sum will be 4.
- I bet if I roll a four-sided die, and you roll a die with that many sides, the result will be even.

Which of these bets has a more than 50% chance of paying off?

The first two are relatively easy to analyze:

- I bet a coin will come up `HEADS`.

One of the TAs has a gambling problem. They make all of the following bets with Adam:

- I bet a coin will come up HEADS.
- I bet a six-sided die will be even.
- I bet if each of us rolls a four-sided die, the sum will be 4.
- I bet if I roll a four-sided die, and you roll a die with that many sides, the result will be even.

Which of these bets has a more than 50% chance of paying off?

The first two are relatively easy to analyze:

- I bet a coin will come up HEADS.
  $\Pr(\texttt{HEADS}) = 1/2$

One of the TAs has a gambling problem. They make all of the following bets with Adam:

- I bet a coin will come up HEADS.
- I bet a six-sided die will be even.
- I bet if each of us rolls a four-sided die, the sum will be 4.
- I bet if I roll a four-sided die, and you roll a die with that many sides, the result will be even.

Which of these bets has a more than 50% chance of paying off?

The first two are relatively easy to analyze:

- I bet a coin will come up HEADS.
  $\Pr(\texttt{HEADS}) = 1/2$
- I bet a six-sided die will be even.

One of the TAs has a gambling problem. They make all of the following bets with Adam:

- I bet a coin will come up `HEADS`.
- I bet a six-sided die will be even.
- I bet if each of us rolls a four-sided die, the sum will be 4.
- I bet if I roll a four-sided die, and you roll a die with that many sides, the result will be even.

Which of these bets has a more than 50% chance of paying off?

The first two are relatively easy to analyze:

- I bet a coin will come up `HEADS`.
  $\Pr(\texttt{HEADS}) = 1/2$
- I bet a six-sided die will be even.
  $\Pr(\text{even roll}) = 3/6 = 1/2$

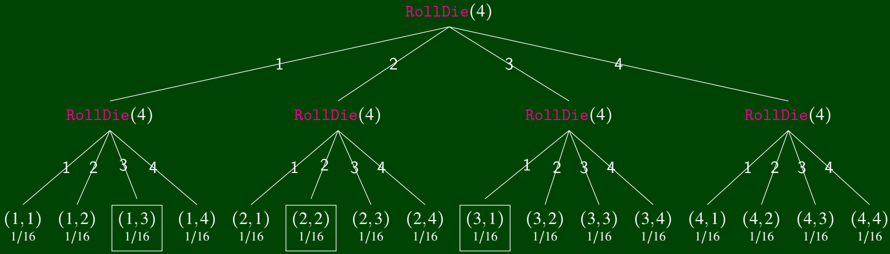
"I bet if each of us rolls a four-sided die, the sum will be 4."

"I bet if each of us rolls a four-sided die, the sum will be 4."

```
ta_die = RollDie(4)
adam_die = RollDie(4)
if ta_die + adam_die == 4:
   print "TA Wins!"
else:
   print "Adam Wins!"
```

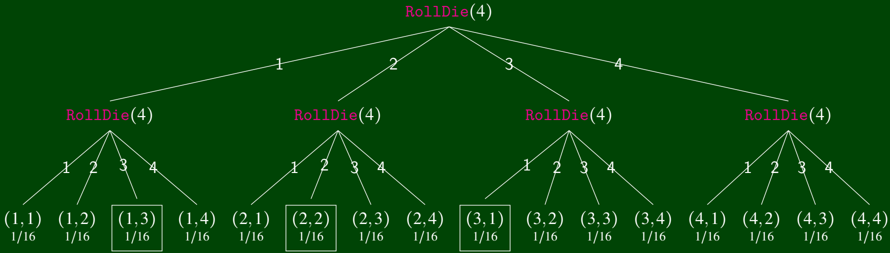
"I bet if each of us rolls a four-sided die, the sum will be 4."

```
ta_die = RollDie(4)
adam_die = RollDie(4)
if ta_die + adam_die == 4:
    print "TA Wins!"
else:
    print "Adam Wins!"
```



"I bet if each of us rolls a four-sided die, the sum will be 4."

```
ta_die = RollDie(4)
adam_die = RollDie(4)
if ta_die + adam_die == 4:
    print "TA Wins!"
else:
    print "Adam Wins!"
```



$$\Pr(\text{sum is 4}) = \Pr(\{(1,3),(2,2),(3,1)\}) = \frac{3}{16}$$

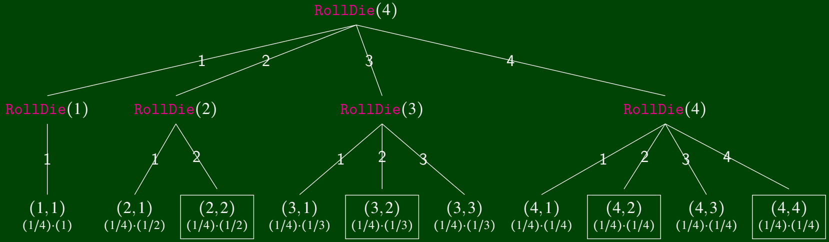
"I bet if I roll a four-sided die, and you roll a die with that many sides, the result will be even."

"I bet if I roll a four-sided die, and you roll a die with that many sides, the result will be even."

```
ta_die = RollDie(4)
result = RollDie(ta_die)
if result % 2 == 0:
   print "TA Wins!"
else:
   print "Adam Wins!"
```

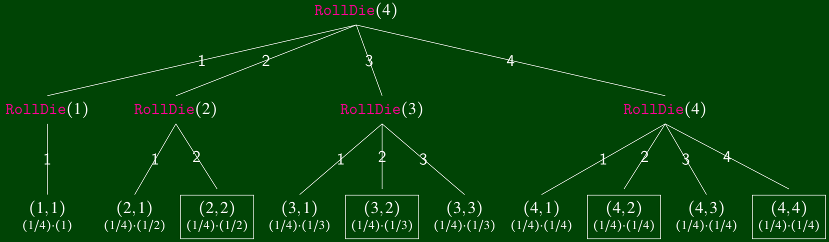
"I bet if I roll a four-sided die, and you roll a die with that many sides, the result will be even."

```
ta_die = RollDie(4)
result = RollDie(ta_die)
if result % 2 == 0:
   print "TA Wins!"
else:
   print "Adam Wins!"
```



"I bet if I roll a four-sided die, and you roll a die with that many sides, the result will be even."

```
ta_die = RollDie(4)
result = RollDie(ta_die)
if result % 2 == 0:
   print "TA Wins!"
else:
   print "Adam Wins!"
```



$$\Pr(\text{result is even}) = \Pr(\{(2,2),(3,2),(4,2),(4,4)\}) = \frac{1}{8} + \frac{1}{12} + \frac{1}{8} = \frac{1}{3}$$

What is the probability that 1313 flips of a fair coin are all `HEADS`?

What is the probability that 1313 flips of a fair coin are all HEADS?

Answer: Each flip has probability $1/2$ of coming up heads;

What is the probability that 1313 flips of a fair coin are all `HEADS`?

Answer: Each flip has probability $1/2$ of coming up heads; so, **multiplying them all together**, we get $\left(\frac{1}{2}\right)^{1313}$.

What is the probability that 1313 flips of a fair coin are all `HEADS`?

Answer: Each flip has probability $1/2$ of coming up heads; so, **multiplying them all together**, we get $\left(\frac{1}{2}\right)^{1313}$.

Why did we multiply them together? What allowed us to do that?

What is the probability that 1313 flips of a fair coin are all `HEADS`?

Answer: Each flip has probability $1/2$ of coming up heads; so, **multiplying them all together**, we get $\left(\frac{1}{2}\right)^{1313}$.

Why did we multiply them together? What allowed us to do that?
We will come back to this argument later when we have more powerful machinery.

What is the probability that 1313 flips of a fair coin are all HEADS?

Answer: Each flip has probability $1/2$ of coming up heads; so, **multiplying them all together**, we get $\left(\frac{1}{2}\right)^{1313}$.

Why did we multiply them together? What allowed us to do that?
We will come back to this argument later when we have more powerful machinery.

More generally:

What is the probability that 1313 flips of a fair coin are all `HEADS`?

Answer: Each flip has probability $1/2$ of coming up heads; so, **multiplying them all together**, we get $\left(\frac{1}{2}\right)^{1313}$.

Why did we multiply them together? What allowed us to do that?
We will come back to this argument later when we have more powerful machinery.

More generally: The probability of any sequence of 1313 flips of a fair coin is $\left(\frac{1}{2}\right)^{1313}$.

What is the probability that 1313 flips of a fair coin are all HEADS?

Answer: Each flip has probability $1/2$ of coming up heads; so, **multiplying them all together**, we get $\left(\frac{1}{2}\right)^{1313}$.

Why did we multiply them together? What allowed us to do that?
We will come back to this argument later when we have more powerful machinery.

More generally: The probability of any sequence of 1313 flips of a fair coin is $\left(\frac{1}{2}\right)^{1313}$. They are all equally likely!

A dating service has a $1/2$ probability of each couple it matches working out. Suppose that the dating service matches 1313 couples.

A dating service has a $1/2$ probability of each couple it matches working out. Suppose that the dating service matches 1313 couples.
What is the probability that exactly 13 couples work out?

A dating service has a $1/2$ probability of each couple it matches working out. Suppose that the dating service matches 1313 couples.
What is the probability that exactly 13 couples work out?
First, model it as code:

```
good_matches = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      good_matches++
if good_matches == 13:
   print "Yay!"
```

Dating Service

A dating service has a $1/2$ probability of each couple it matches working out. Suppose that the dating service matches 1313 couples.
What is the probability that exactly 13 couples work out?
First, model it as code:

```
good_matches = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      good_matches++
if good_matches == 13:
   print "Yay!"
```

- Let E be the event "the program prints "Yay!"".

A dating service has a $1/2$ probability of each couple it matches working out. Suppose that the dating service matches 1313 couples.
What is the probability that exactly 13 couples work out?
First, model it as code:

```
good_matches = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      good_matches++
if good_matches == 13:
   print "Yay!"
```

- Let E be the event "the program prints "Yay!"".
- The sample space, $S = \{\texttt{HEADS}, \texttt{TAILS}\}^{1313}$.

A dating service has a $1/2$ probability of each couple it matches working out. Suppose that the dating service matches 1313 couples.
What is the probability that exactly 13 couples work out?
First, model it as code:

```
good_matches = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      good_matches++
if good_matches == 13:
   print "Yay!"
```

- Let E be the event "the program prints "Yay!"".
- The sample space, $S = \{\texttt{HEADS}, \texttt{TAILS}\}^{1313}$.
- Every outcome is equally likely, because we showed it above!

A dating service has a $1/2$ probability of each couple it matches working out. Suppose that the dating service matches 1313 couples.
What is the probability that exactly 13 couples work out?
First, model it as code:

```
good_matches = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      good_matches++
if good_matches == 13:
   print "Yay!"
```

- Let E be the event "the program prints "Yay!"".
- The sample space, $S = \{\texttt{HEADS}, \texttt{TAILS}\}^{1313}$.
- Every outcome is equally likely, because we showed it above!
- So, $\Pr(E) = \frac{|E|}{|S|}$.

A dating service has a $1/2$ probability of each couple it matches working out. Suppose that the dating service matches 1313 couples.
What is the probability that exactly 13 couples work out?
First, model it as code:

```
good_matches = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      good_matches++
if good_matches == 13:
   print "Yay!"
```

- Let E be the event "the program prints "Yay!"".
- The sample space, $S = \{\texttt{HEADS}, \texttt{TAILS}\}^{1313}$.
- Every outcome is equally likely, because we showed it above!
- So, $\Pr(E) = \frac{|E|}{|S|}$. (Take the number of ways to do what we want and divide by the number of outcomes.)

Dating Service

A dating service has a $1/2$ probability of each couple it matches working out. Suppose that the dating service matches 1313 couples.
What is the probability that exactly 13 couples work out?
First, model it as code:

```
good_matches = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      good_matches++
if good_matches == 13:
   print "Yay!"
```

- Let E be the event "the program prints "Yay!"".
- The sample space, $S = \{\texttt{HEADS}, \texttt{TAILS}\}^{1313}$.
- Every outcome is equally likely, because we showed it above!
- So, $\Pr(E) = \frac{|E|}{|S|} = \frac{\binom{1313}{13}}{2^{1313}}$.

How many people do we need before there's a 50% chance that two of them share a birthday?

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space.

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space. We have 23 people each with one of 366 birthdays; so, the outcomes are length 23 strings of birthdays.

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space. We have 23 people each with one of 366 birthdays; so, the outcomes are length 23 strings of birthdays.
Let E be the event that at least two people have the same birthday.

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space. We have 23 people each with one of 366 birthdays; so, the outcomes are length 23 strings of birthdays.
Let E be the event that at least two people have the same birthday. If we can calculate $|E|$ and $|S|$, we're done.

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space. We have 23 people each with one of 366 birthdays; so, the outcomes are length 23 strings of birthdays.
Let E be the event that at least two people have the same birthday. If we can calculate $|E|$ and $|S|$, we're done.

Since $|E|$ is hard to count, we count $|\overline{E}|$ instead!

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space. We have 23 people each with one of 366 birthdays; so, the outcomes are length 23 strings of birthdays.
Let E be the event that at least two people have the same birthday. If we can calculate $|E|$ and $|S|$, we're done.

Since $|E|$ is hard to count, we count $|\overline{E}|$ instead! $\overline{E}$ is the event that no two people share a birthday.

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space. We have 23 people each with one of 366 birthdays; so, the outcomes are length 23 strings of birthdays.
Let E be the event that at least two people have the same birthday. If we can calculate $|E|$ and $|S|$, we're done.

Since $|E|$ is hard to count, we count $|\overline{E}|$ instead! $\overline{E}$ is the event that no two people share a birthday. In other words, the event that the string has no duplicate entries.

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space. We have 23 people each with one of 366 birthdays; so, the outcomes are length 23 strings of birthdays.
Let E be the event that at least two people have the same birthday. If we can calculate $|E|$ and $|S|$, we're done.

Since $|E|$ is hard to count, we count $|\overline{E}|$ instead! $\overline{E}$ is the event that no two people share a birthday. In other words, the event that the string has no duplicate entries. Using the Rule of Product, we see that

$$|\overline{E}| = (366)(365)\cdots(344)$$

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space. We have 23 people each with one of 366 birthdays; so, the outcomes are length 23 strings of birthdays.
Let E be the event that at least two people have the same birthday. If we can calculate $|E|$ and $|S|$, we're done.

Since $|E|$ is hard to count, we count $|\overline{E}|$ instead! $\overline{E}$ is the event that no two people share a birthday. In other words, the event that the string has no duplicate entries. Using the Rule of Product, we see that

$$|\overline{E}| = (366)(365)\cdots(344)$$

We also know that $|S| = 366^{23}$ (also by the Rule of Product).

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space. We have 23 people each with one of 366 birthdays; so, the outcomes are length 23 strings of birthdays.
Let E be the event that at least two people have the same birthday. If we can calculate $|E|$ and $|S|$, we're done.

Since $|E|$ is hard to count, we count $|\overline{E}|$ instead! $\overline{E}$ is the event that no two people share a birthday. In other words, the event that the string has no duplicate entries. Using the Rule of Product, we see that

$$|\overline{E}| = (366)(365)\cdots(344)$$

We also know that $|S| = 366^{23}$ (also by the Rule of Product). Then, $\frac{|\overline{E}|}{|S|} = 0.494$

How many people do we need before there's a 50% chance that two of them share a birthday?

It's actually only **23**! Let's prove it:
Let S be the sample space. We have 23 people each with one of 366 birthdays; so, the outcomes are length 23 strings of birthdays.
Let E be the event that at least two people have the same birthday. If we can calculate $|E|$ and $|S|$, we're done.

Since $|E|$ is hard to count, we count $|\overline{E}|$ instead! $\overline{E}$ is the event that no two people share a birthday. In other words, the event that the string has no duplicate entries. Using the Rule of Product, we see that

$$|\overline{E}| = (366)(365)\cdots(344)$$

We also know that $|S| = 366^{23}$ (also by the Rule of Product). Then, $\frac{|\overline{E}|}{|S|} = 0.494$ So,

$$\frac{|E|}{|S|} = 0.506$$

It is sometimes useful to discuss how often an event A occurs assuming that another event B has already occurred.

It is sometimes useful to discuss how often an event A occurs assuming that another event B has already occurred.

Definition (Conditional Probability)

$\Pr(A \mid B)$ is pronounced "the probability of A given B", and it's defined as:

It is sometimes useful to discuss how often an event A occurs assuming that another event B has already occurred.

Definition (Conditional Probability)

$\Pr(A \mid B)$ is pronounced "the probability of A given B", and it's defined as:

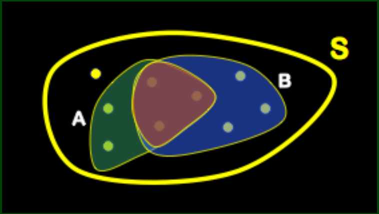
$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

It is sometimes useful to discuss how often an event A occurs assuming that another event B has already occurred.

Definition (Conditional Probability)

$\Pr(A \mid B)$ is pronounced "the probability of A given B", and it's defined as:

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$



Think of $\Pr(A \mid B)$ as restricting the sample space to B and considering how often A occurs.

Suppose we roll a black die and a white die. What is the probability that the white die is 1 given that the sum is 7?

Suppose we roll a black die and a white die. What is the probability that the white die is 1 given that the sum is 7?

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

We can calculate this two ways:

Suppose we roll a black die and a white die. What is the probability that the white die is 1 given that the sum is 7?

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

We can calculate this two ways:

Way 1: Let W be the event that the white die is 1 and S be the event that the sum is 7. We see that $\Pr(S) = \frac{6}{36}$ and $\Pr(S \cap W) = \frac{1}{36}$. Using the formula for conditional probability, we get $\Pr(W \mid S) = \frac{\frac{1}{36}}{\frac{6}{36}} = \frac{1}{6}$.

Suppose we roll a black die and a white die. What is the probability that the white die is 1 given that the sum is 7?

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

We can calculate this two ways:

Way 1: Let W be the event that the white die is 1 and S be the event that the sum is 7. We see that $\Pr(S) = \frac{6}{36}$ and $\Pr(S \cap W) = \frac{1}{36}$. Using the formula for conditional probability, we get $\Pr(W \mid S) = \frac{\frac{1}{36}}{\frac{6}{36}} = \frac{1}{6}$.

Way 2: If we restrict our sample space to the six outcomes where the sum is 7, only one of them has the white die as 1. So, the probability we're looking for is $\frac{1}{6}$.

We have three bags with coins: one has two gold coins, one has two silver coins, and the third has one gold coin and one silver coin.

We randomly choose one of the bags, and we randomly select a coin from the bag. It turns out to be gold.
What is the probability that the other coin in that bag is gold?

We have three bags with coins: one has two gold coins, one has two silver coins, and the third has one gold coin and one silver coin.

We randomly choose one of the bags, and we randomly select a coin from the bag. It turns out to be gold.
What is the probability that the other coin in that bag is gold?

Let G_1 be the event that the coin we chose first is gold.

We have three bags with coins: one has two gold coins, one has two silver coins, and the third has one gold coin and one silver coin.

We randomly choose one of the bags, and we randomly select a coin from the bag. It turns out to be gold.
What is the probability that the other coin in that bag is gold?

Let G_1 be the event that the coin we chose first is gold. Let G_2 be the event that the second chosen coin is gold.

We have three bags with coins: one has two gold coins, one has two silver coins, and the third has one gold coin and one silver coin.

We randomly choose one of the bags, and we randomly select a coin from the bag. It turns out to be gold.
What is the probability that the other coin in that bag is gold?

Let G_1 be the event that the coin we chose first is gold. Let G_2 be the event that the second chosen coin is gold. The problem is asking us to find $\Pr(G_2 \mid G_1)$:

- $\Pr(G_1) = 3/6 = 1/2$ (three of the six coins are gold)
- $\Pr(G_1 \cap G_2) = 1/3$ (the coins must be in the gold-gold bag)
- $\Pr(G_2 \mid G_1) = \frac{\Pr(G_1 \cap G_2)}{\Pr(G_1)} = \frac{2}{3}$

Definition (Independence)

We say that two events A and B are **independent** iff

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

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:
Let C_P be the event that Pete comes to 98-000 class.

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:
Let C_P be the event that Pete comes to 98-000 class. Let C_S be the event that Sandy comes to 98-000 class.

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:
Let C_P be the event that Pete comes to 98-000 class. Let C_S be the event that Sandy comes to 98-000 class. Empirically, we know that:

- $\Pr(C_P) = 1$

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:
Let C_P be the event that Pete comes to 98-000 class. Let C_S be the event that Sandy comes to 98-000 class. Empirically, we know that:

- $\Pr(C_P) = 1$
- $\Pr(C_S) = \frac{1}{8}$

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:

Let C_P be the event that Pete comes to 98-000 class. Let C_S be the event that Sandy comes to 98-000 class. Empirically, we know that:

- $\Pr(C_P) = 1$
- $\Pr(C_S) = \frac{1}{8}$
- $\Pr(C_P \cap C_S) = \frac{1}{8}$

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:
Let C_P be the event that Pete comes to 98-000 class. Let C_S be the event that Sandy comes to 98-000 class. Empirically, we know that:

- $\Pr(C_P) = 1$
- $\Pr(C_S) = \frac{1}{8}$
- $\Pr(C_P \cap C_S) = \frac{1}{8}$

Since $\frac{1}{8} = \frac{1}{8}$, $\Pr(C_P)$ and $\Pr(C_S)$ are independent!

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:
Let C_P be the event that Pete comes to 98-000 class. Let C_S be the event that Sandy comes to 98-000 class. Empirically, we know that:

- $\Pr(C_P) = 1$
- $\Pr(C_S) = \frac{1}{8}$
- $\Pr(C_P \cap C_S) = \frac{1}{8}$

Since $\frac{1}{8} = \frac{1}{8}$, $\Pr(C_P)$ and $\Pr(C_S)$ are independent!
What does this mean?

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:
Let C_P be the event that Pete comes to 98-000 class. Let C_S be the event that Sandy comes to 98-000 class. Empirically, we know that:

- $\Pr(C_P) = 1$
- $\Pr(C_S) = \frac{1}{8}$
- $\Pr(C_P \cap C_S) = \frac{1}{8}$

Since $\frac{1}{8} = \frac{1}{8}$, $\Pr(C_P)$ and $\Pr(C_S)$ are independent!
What does this mean?
Tempting Answer: Pete and Sandy's decisions to go to class aren't reliant on each other.

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:
Let C_P be the event that Pete comes to 98-000 class. Let C_S be the event that Sandy comes to 98-000 class. Empirically, we know that:

- $\Pr(C_P) = 1$
- $\Pr(C_S) = \frac{1}{8}$
- $\Pr(C_P \cap C_S) = \frac{1}{8}$

Since $\frac{1}{8} = \frac{1}{8}$, $\Pr(C_P)$ and $\Pr(C_S)$ are independent!
What does this mean?
Tempting Answer: Pete and Sandy's decisions to go to class aren't reliant on each other.
In Actuality: There are a total of 8 sessions of 98-000.

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:
Let C_P be the event that Pete comes to 98-000 class. Let C_S be the event that Sandy comes to 98-000 class. Empirically, we know that:

- $\Pr(C_P) = 1$
- $\Pr(C_S) = \frac{1}{8}$
- $\Pr(C_P \cap C_S) = \frac{1}{8}$

Since $\frac{1}{8} = \frac{1}{8}$, $\Pr(C_P)$ and $\Pr(C_S)$ are independent!

What does this mean?

Tempting Answer: Pete and Sandy's decisions to go to class aren't reliant on each other.

In Actuality: There are a total of 8 sessions of 98-000. Sandy met Pete at the first class which they both attended.

Definition (Independence)

We say that two events A and B are **independent** iff

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

Here's an example:

Let C_P be the event that Pete comes to 98-000 class. Let C_S be the event that Sandy comes to 98-000 class. Empirically, we know that:

- $\Pr(C_P) = 1$
- $\Pr(C_S) = \frac{1}{8}$
- $\Pr(C_P \cap C_S) = \frac{1}{8}$

Since $\frac{1}{8} = \frac{1}{8}$, $\Pr(C_P)$ and $\Pr(C_S)$ are independent!

What does this mean?

Tempting Answer: Pete and Sandy's decisions to go to class aren't reliant on each other.

In Actuality: There are a total of 8 sessions of 98-000. Sandy met Pete at the first class which they both attended. For all future sessions, Pete took notes for Sandy, and she never showed up again.

What is the probability that 1313 flips of a fair coin are all `HEADS`?

What is the probability that 1313 flips of a fair coin are all HEADS?

```
num_heads = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      num_heads++
if num_heads == 1313:
   print "Yay!"
```

What is the probability that 1313 flips of a fair coin are all HEADS?

```
num_heads = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      num_heads++
if num_heads == 1313:
   print "Yay!"
```

Let C_i (for $1 \leq i \leq 1313$) be the event that the ith coin flip is HEADS.

What is the probability that 1313 flips of a fair coin are all HEADS?

```
num_heads = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      num_heads++
if num_heads == 1313:
   print "Yay!"
```

Let C_i (for $1 \leq i \leq 1313$) be the event that the ith coin flip is HEADS.

We see that none of the results of the coin flips in the procedure interact with each other; so, each of them is independent.

What is the probability that 1313 flips of a fair coin are all HEADS?

```
num_heads = 0
for i=1 to 1313:
   if FlipCoin(1/2) == HEADS:
      num_heads++
if num_heads == 1313:
   print "Yay!"
```

Let C_i (for $1 \leq i \leq 1313$) be the event that the ith coin flip is HEADS.
We see that none of the results of the coin flips in the procedure interact with each other; so, each of them is independent. It follows that

$$\Pr\left(\bigcap_{i=1}^{1313} C_i\right) = \prod_{i=1}^{1313} \Pr(C_i) = \left(\frac{1}{2}\right)^{1313}$$

```
while FlipCoin() ≠ HEADS:
    print "Hello!"
```

Does this code terminate?
Let E be the event that the code terminates.

```
while FlipCoin() ≠ HEADS:
    print "Hello!"
```

Does this code terminate?

Let E be the event that the code terminates. Note that if E_i is the event that the code terminates on the ith run through the loop, we have:

$$E = \bigcup_{i=0}^{\infty} E_i$$

Then, we know $\Pr(E) = \Pr(\bigcup_{i=0}^{\infty} E_i) = \sum_{i=1}^{\infty} \Pr(E_i)$. Note that $\Pr(E_i)$ is the probability that the code does not terminate on the first $i-1$ iterations and does on the ith iteration. Then, $\Pr(E_i) = \left(\frac{1}{2}\right)^i$. So,

$$\Pr(E) = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i = \frac{1}{1-\frac{1}{2}} - 1 = 1$$