Lecture 4



Foundations of Computing II

CSE 312: Foundations of Computing II

Probability Axioms & Equally-Likely Outcomes

Consider the following piece of code:

```
1 while FlipCoin(1/2) ≠ HEADS:
2 print "Hello!"
```

Consider the following piece of code:

```
1 while FlipCoin(1/2) ≠ HEADS:
2 print "Hello!"
```

Does it terminate?

Consider the following piece of code:

```
1 while FlipCoin(1/2) ≠ HEADS:
2 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:

```
1 while FlipCoin(1/2) # HEADS:
2 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 pretty much the rest of the course is about!

Consider the following piece of code:

1 while FlipCoin(1/2) # HEADS: 2 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 pretty much the rest of the course is about!

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

Probability Primitives

For now, assume we have two random generators: FlipCoin(p) returns HEADS with probability p and TAILS otherwise. RollDie(N) returns $x \in [N]$ with probability 1/N. (We will do better later, but for now...)

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
```

if c == d: print c
else: print "Failure!"
(\LEADS, /)

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:

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:

(HEADS, 1), (HEADS, 2), (TAILS, 1), (TAILS, 2)

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)

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} .

S

An event is a subset of the sample space.

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

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

```
c = F
d = R
if c
```

8

```
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!"



An **event** is a subset of the sample space. We call these events, because they 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!" {(HEADS,2),(TAILS,1)}

An **event** is a subset of the sample space. We call these events, because they 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!" {(HEADS,2),(TAILS,1)}
- The code halts.

An **event** is a subset of the sample space. We call these events, because they 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!" {(HEADS,2),(TAILS,1)}

```
The code halts.
{(HEADS,1),(HEADS,2),(TAILS,1),(TAILS,2)}
```

An **event** is a subset of the sample space. We call these events, because they 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!" {(HEADS,2),(TAILS,1)}
- The code halts.
 {(HEADS,1),(HEADS,2),(TAILS,1),(TAILS,2)}
- The coin flip gives HEADS.

An **event** is a subset of the sample space. We call these events, because they 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!" {(HEADS,2),(TAILS,1)}
- The code halts.
 {(HEADS,1),(HEADS,2),(TAILS,1),(TAILS,2)}
- The coin flip gives HEADS. {(HEADS, 1), (HEADS, 2)}

Definition (Probability)

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

Definition (Probability)

Let Ω be a sample space and $E \subseteq \Omega$ be an event. Then, we say Pr(E) is the probability of E. All probability functions must satisfy the following three axioms:



Definition (Probability)

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

All probability functions must satisfy the following three axioms:

- $0 \leq \Pr(E)$
- Pr(Ω) = 1

Definition (Probability)

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

All probability functions must satisfy the following three axioms:

- $0 \leq \Pr(E)$
- Pr(Ω) = 1
- If $E_1, E_2, \dots, E_n \subseteq \Omega$ and are pairwise disjoint events, then $(E_1 \cup E_2 \cup \dots \cup E_n) = \sum_{i=1}^n \Pr(E_i)$

An Axiomatic Proof

Claim

 $\Pr(\overline{E}) = 1 - \Pr(E)$

Proof.

Note that $\Pr(\Omega) = 1$ by the second axiom. Furthermore, $E \cup \overline{E} = \Omega$ for all events E; so, $1 = \Pr(\Omega) = \Pr(E \cup \overline{E}) = \Pr(E) + \Pr(\overline{E})$ by the above and axiom 3. Then, $\Pr(\overline{E}) = 1 - \Pr(E)$.



1	coin1	=	<pre>FlipCoin(1/2)</pre>
2	coin2	=	<pre>FlipCoin(1/2)</pre>

Sample Space (All potential outcomes of the random calls)







Sample Space (All potential outcomes of the random calls)

 $\Omega = \{(\texttt{HEADS}, \texttt{HEADS}), (\texttt{HEADS}, \texttt{TAILS}), (\texttt{TAILS}, \texttt{HEADS}), (\texttt{TAILS}, \texttt{TAILS})\}$

Example Events $(E \subseteq \Omega)$

- ∎ Ø
- {(HEADS, HEADS)}
- coin1 = HEADS

flip = FlipCoin(1/2) 1 <u>Ω</u> =

```
1 flip = FlipCoin(1/2)

\Omega = \{\text{HEADS}, \text{TAILS}\}
```

```
1 flip1 = FlipCoin(1/2)
2 flip2 = FlipCoin(1/2)
\Omega =
```

```
1 flip = FlipCoin(1/2)

\Omega = \{\text{HEADS}, \text{TAILS}\}
```

- 1 flip1 = FlipCoin(1/2)
- 2 flip2 = FlipCoin(1/2)

 $\Omega = \{(\texttt{HEADS}, \texttt{HEADS}), (\texttt{HEADS}, \texttt{TAILS}), (\texttt{TAILS}, \texttt{HEADS}), (\texttt{TAILS}, \texttt{TAILS})\}$

```
1 roll = RollDie(6)
\Omega =
```

```
1 flip = FlipCoin(1/2)

\Omega = \{\text{HEADS}, \text{TAILS}\}
```

- 1 flip1 = FlipCoin(1/2)
- 2 flip2 = FlipCoin(1/2)

 $\Omega = \{(\texttt{HEADS},\texttt{HEADS}),(\texttt{HEADS},\texttt{TAILS}),(\texttt{TAILS},\texttt{HEADS}),(\texttt{TAILS},\texttt{TAILS})\}$

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$


```
1 flip = FlipCoin(1/3)

\Omega = \{\text{HEADS}, \text{TAILS}\}
Pr(HEADS) =
```



1 flip = FlipCoin(1/3)

$$\Omega = \{\text{HEADS}, \text{TAILS}\}$$
Pr(HEADS) = $\left(\frac{1}{3}\right)$
Pr(TAILS) = $\left(\frac{2}{3}\right)$

1 flip = FlipCoin(1/3)

$$\Omega = \{\text{HEADS}, \text{TAILS}\}$$
Pr(HEADS) = $\left(\frac{1}{3}\right)$
Pr(TAILS) = $\left(\frac{2}{3}\right)$

flip = FlipCoin(1/2)
if flip == HEADS:
flip2 = RollDie(2)
else:
flip2 = RollDie(3)

$$\Omega = \{(HEADS, 1), (HEADS, 2), (TAILS, 1), (TAILS, 2), (TAILS, 3)\}$$

Pr((HEADS, 1)) =

flip = FlipCoin(1/3)

$$\Omega = \{\text{HEADS}, \text{TAILS}\}$$

$$\Pr(\text{HEADS}) = \left(\frac{1}{3}\right)$$

$$\Pr(\text{TAILS}) = \left(\frac{2}{3}\right)$$

flip = FlipCoin(1/2)
if flip == HEADS:
flip2 = RollDie(2)
else:
flip2 = RollDie(3)

$$\Omega = \{(HEADS, 1), (HEADS, 2), (TAILS, 1), (TAILS, 2), (TAILS, 3)\}$$

$$\Pr((\text{HEADS}, 1)) = \left(\frac{1}{2}\right)^2$$

Pr((TAILS,1)) =

flip = FlipCoin(1/3)

$$\Omega = \{\text{HEADS}, \text{TAILS}\}$$

$$\Pr(\text{HEADS}) = \left(\frac{1}{3}\right)$$

$$\Pr(\text{TAILS}) = \left(\frac{2}{3}\right)$$

1 flip = FlipCoin(1/2)
2 if flip == HEADS:
3 flip2 = RollDie(2)
4 else:
5 flip2 = RollDie(3)

$$\Omega = \{(HEADS, 1), (HEADS, 2), (TAILS, 1), (TAILS, 2), (TAILS, 3)\}$$

$$\Pr((\text{HEADS}, 1)) = \left(\frac{1}{2}\right)^2$$

$$\Pr((\texttt{TAILS},1)) = \left(\frac{1}{2}\right) \left(\frac{1}{3}\right)$$

We'll stick with equally-likely outcomes for now.

 $Pr(E) = \frac{|E|}{|\Omega|}$ For equally-likely outcomes



We'll stick with equally-likely outcomes for now.

If we have a sample space, $\Omega,$ and all the outcomes are equally-likely,

We'll stick with equally-likely outcomes for now.

If we have a sample space, Ω , and all the outcomes are equally-likely, then the $\Pr(\{x\}) = \frac{1}{|\Omega|}$.

We'll stick with equally-likely outcomes for now.

If we have a sample space, Ω , and all the outcomes are equally-likely, then the $Pr(\{x\}) = \frac{1}{|\Omega|}$. So, $Pr(\{x_1, x_2, \dots, x_n\}) =$

We'll stick with equally-likely outcomes for now.

If we have a sample space, Ω , and all the outcomes are equally-likely, then the $\Pr(\{x\}) = \frac{1}{|\Omega|}$. So, $\Pr(\{x_1, x_2, \dots, x_n\}) = \frac{n}{|\Omega|}$. So, $\Pr(E) =$

We'll stick with equally-likely outcomes for now.

If we have a sample space, Ω , and all the outcomes are equally-likely, then the $\Pr(\{x\}) = \frac{1}{|\Omega|}$. So, $\Pr(\{x_1, x_2, \dots, x_n\}) = \frac{n}{|\Omega|}$. So, $\Pr(E) = \frac{|E|}{|\Omega|}$.

A's Gambling Problem(s)

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

- I bet a fair coin will come up HEADS.
- I bet a six-sided die will be even.
- I bet if each of us rolls a six-sided die, the sum will be 7.

Let's model each of these with code, determine the events and sample space, and evaluate the probabilities of each.

```
1 coin = FlipCoin(1/2)
```

 $\Omega = \{ \text{HEADS}, \text{TAILS} \}$ $E = \{ \text{HEADS} \}$

1 coin = FlipCoin(1/2)

 $\Omega = \{\text{HEADS}, \text{TAILS}\}$ $E = \{\text{HEADS}\}$ $\Pr(E) = \frac{|E|}{|\Omega|} = \frac{1}{2}$

1 roll = RollDie(6)

 $\Omega = \{1, 2, 3, 4, 5, 6\}$ E = $\{2, 4, 6\}$

1 roll = RollDie(6)

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$E = \{2, 4, 6\}$$

$$\Pr(E) = \frac{|E|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}$$

I bet if each of us rolls a six-sided die, the sum will be 7 15

1 adamRoll = RollDie(6)
2 taRoll = RollDie(6)

I bet if each of us rolls a six-sided die, the sum will be 7 15

```
1 adamRoll = RollDie(6)
2 taRoll = RollDie(6)
```

1 adamRoll = RollDie(6)
2 taRoll = RollDie(6)



I bet if each of us rolls a six-sided die, the sum will be 7 15

1 adamRoll = RollDie(6)
2 taRoll = RollDie(6)

I bet if each of us rolls a six-sided die, the sum will be 7 15

1 adamRoll = RollDie(6)
2 taRoll = RollDie(6)

So,
$$Pr(adamRoll + taRoll = 7) = \frac{|E|}{|S|} = \frac{6}{36} = \frac{1}{6}$$

What is the probability that none of *n* people share the same birthday? $\bigcap_{n=0}^{n=0} \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{i=0}^{n} \sum_{i=0}^{n} \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{i=0}^{n} \sum_{i=0}^{n}$

(Jan 1) Jan 1)

What is the probability that none of n people share the same birthday?

- 1 person1 = RollDie(1/365)
- 2 person2 = RollDie(1/365)
- 3 ...
- 4 person365 = RollDie(1/365)

What assumptions have we implicitly made with this code?

What is the probability that none of n people share the same birthday?

- 1 person1 = RollDie(1/365)
- 2 person2 = RollDie(1/365)
- 3 ...
- 4 person365 = RollDie(1/365)

What assumptions have we implicitly made with this code?

Birthdays are equally likely (not actually true... but close enough)

What is the sample space?

 $|\Omega|$ =

What is the probability that none of n people share the same birthday?

- 1 person1 = RollDie(1/365)
- 2 person2 = RollDie(1/365)
- 3 ...
- 4 person365 = RollDie(1/365)

What assumptions have we implicitly made with this code?

Birthdays are equally likely (not actually true... but close enough)

What is the sample space?

$$|\Omega| = |\{(x_1, x_2, \dots, x_n) : 0 < x_i \le 365\}| = 365^n$$

Let E be the event that all people have different birthdays.

What is E?

$$|E| = \frac{(3(5)(3(4)) - (3(5-n))(1)}{(3(5-n))(n)} = \frac{3(5, n)}{(3(5-n))(n)}$$

What is the probability that none of n people share the same birthday?

- 1 person1 = RollDie(1/365)
- 2 person2 = RollDie(1/365)
- 3 ...
- 4 person365 = RollDie(1/365)

What assumptions have we implicitly made with this code?

Birthdays are equally likely (not actually true... but close enough)

What is the sample space?

$$|\Omega| = |\{(x_1, x_2, \dots, x_n) : 0 < x_i \le 365\}| = 365^n$$

Let E be the event that all people have different birthdays.

What is *E*?

 $|E| = 365 \times 364 \times \cdots \times (365 - n + 1)$

What is the probability that none of n people share the same birthday?

Let E be the event that all people have different birthdays.

 $\begin{aligned} |\Omega| &= 365^n \\ |E| &= 365 \times 364 \times \dots \times (365 - n + 1) \\ \text{So, } \Pr(\text{no shared birthdays}) &= \frac{|E|}{|\Omega|} = \end{aligned}$

What is the probability that none of n people share the same birthday?

Let E be the event that all people have different birthdays.



The Situation

- n chips manufactured
- k chips randomly selected for testing
- one chip defective

What is Pr(defective chip in k selected chips)?

The Situation

- n chips manufactured
- k chips randomly selected for testing
- one chip defective

What is Pr(defective chip in k selected chips)?

$$|\Omega| = \binom{n}{k}$$

■ $|E| = {1 \choose 1} {n-1 \choose k-1}$ (choose defective chip; pick remaining chips to test)

Pr(defective chip in k selected chips) =
$$\frac{|E|}{|\Omega|} = \frac{\binom{1}{1}\binom{n-1}{k-1}}{\binom{n}{k}}$$

What is Pr(the *i*th selected chip is defective)?

The Situation

- n chips manufactured
- k chips randomly selected for testing
- one chip defective

What is Pr(defective chip in k selected chips)?

$$|\Omega| = \binom{n}{k}$$

■ $|E| = {\binom{1}{1}} {\binom{n-1}{k-1}}$ (choose defective chip; pick remaining chips to test)

Pr(defective chip in k selected chips) =
$$\frac{|E|}{|\Omega|} = \frac{\binom{1}{1}\binom{n-1}{k-1}}{\binom{n}{k}}$$

What is Pr(the *i*th selected chip is defective)?

It is equally likely that each of the *n* chips is the defective one. So, $\Pr(\text{the }i\text{th selected chip is defective}) = \Pr(E_i) = \frac{1}{n}$ $\Pr(\text{defective chip in }k \text{ selected chips}) = \Pr(E_1) + \Pr(E_2) + \dots + \Pr(E_k) = \frac{k}{n}$.

The Situation

- n chips manufactured
- k chips randomly selected for testing
- d chips defective

What is Pr(at least one defective chip in k selected chips)?

The Situation

- n chips manufactured
- k chips randomly selected for testing
- d chips defective

What is Pr(at least one defective chip in k selected chips)?

Consider Pr(no chip defective in k selected chips) instead.

The Situation

- n chips manufactured
- k chips randomly selected for testing
- d chips defective

What is Pr(at least one defective chip in k selected chips)?

Consider Pr(no chip defective in k selected chips) instead.

$$\bullet |\Omega| = \binom{n}{k}$$

■
$$|E| = \binom{n-d}{k}$$
 (choose from non-defective chips)

Pr(at least one defective in
$$k$$
) = $1 - \frac{|E|}{|\Omega|} = 1 - \frac{\binom{n-d}{k}}{\binom{n}{k}}$