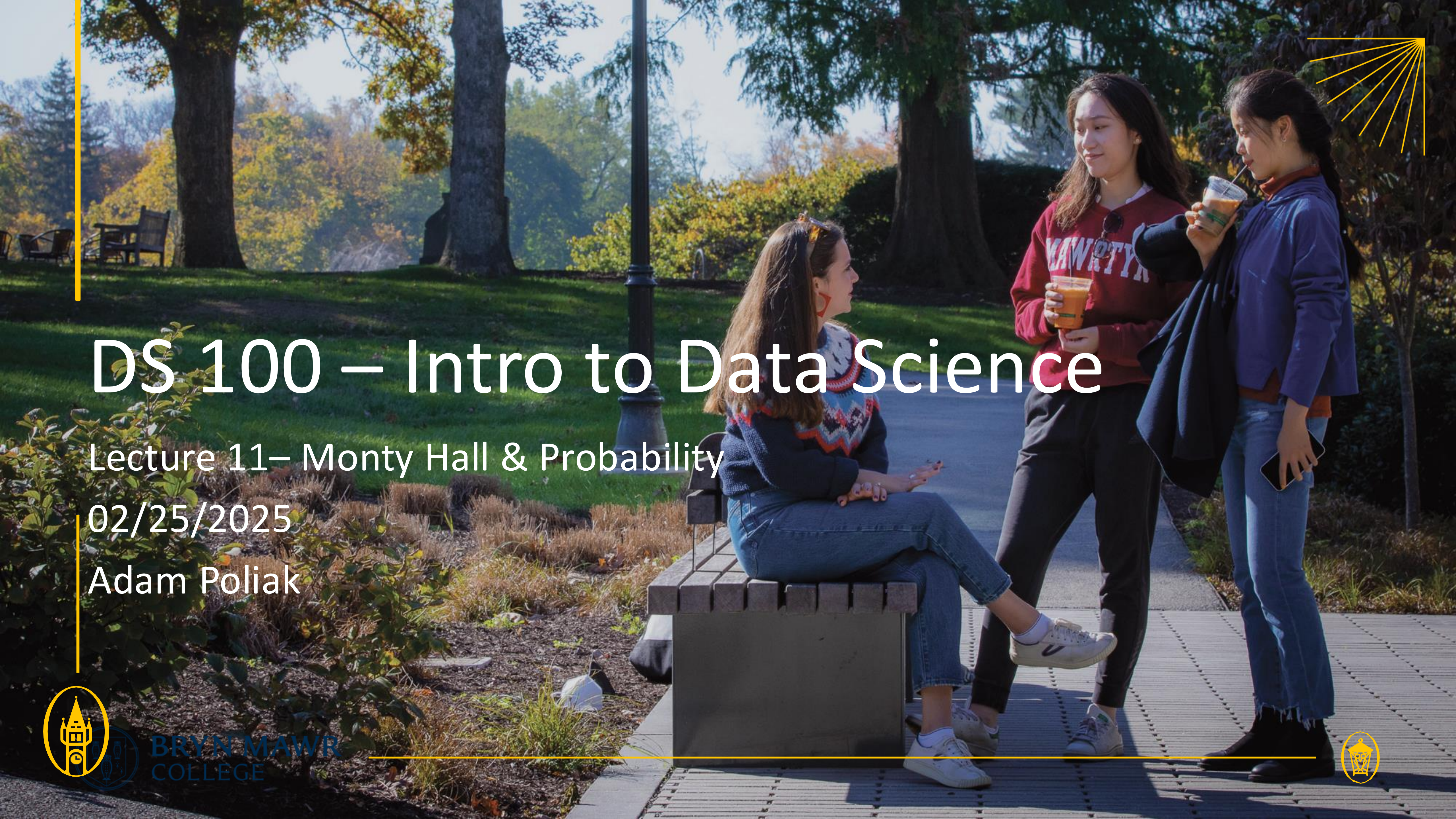


DS 100 – Intro to Data Science

Lecture 11– Monty Hall & Probability

02/25/2025

Adam Poliak

BRYN MAWR COLLEGE

Announcements

Checkpoint/Project 1:

- Paired assignment that covers the previous section of the course material
- Due Friday 02/28

HW05 – Probability, Simulation, Estimation, and Assessing Models

- Due Wednesday (03/05)

Lab & Late Days

Can’t use late days for lab

Lab 0 – 4, if you missed any, let me know (by end of today) and you can submit them with two late days

Experimentation & Simulation

Experimentation

Why do we want to run experiments?

To test a hypothesis

What could a hypothesis be in our previous (coin flipping) experiments?

Is a coin fair or not?

…

How to run an experiment/test a hypothesis

Step 1: Choose a measurement/statistic to study
Textbook lingo: what to simulate

Step 2: Figure out how to compute the measurement
Textbook lingo: figure how out to simulate the statistic

Step 3: Choose how many times to simulate the statistic
Textbook lingo: Number of Repetitions

Step 4: Do it!
Textbook lingo: simulate multiple values

How to run an experiment/test a hypothesis

Step 1: Choose a measurement/statistic to study
Textbook lingo: what to simulate

Step 2: Figure out how to compute the measurement
Textbook lingo: figure how out to simulate the statistic

Step 3: Choose how many times to simulate the statistic
Textbook lingo: Number of Repetitions

Step 4: Do it!
Textbook lingo: simulate multiple values

func for np.append()

How to run an experiment/test a hypothesis

Step 1: Choose a measurement/statistic to study

Textbook lingo: what to simulate

Step 2: Figure out how to compute the measurement

Textbook lingo: figure how out to simulate the statistic

func

Step 3: Choose how many times to simulate the statistic

Textbook lingo: Number of Repetitions

Step 4: Do it!

Textbook lingo: simulate multiple values

for

np.append()

How to run an experiment/test a hypothesis

Step 1: Choose a measurement/statistic to study

Textbook lingo: what to simulate

Step 2: Figure out how to compute the measurement

Textbook lingo: figure how out to simulate the statistic

func

Step 3: Choose how many times to simulate the statistic

Textbook lingo: Number of Repetitions

for

Step 4: Do it!

Textbook lingo: simulate multiple values

np.append()

How to run an experiment/test a hypothesis

Step 1: Choose a measurement/statistic to study
Textbook lingo: what to simulate

Step 2: Figure out how to compute the measurement — func
Textbook lingo: figure how out to simulate the statistic

Step 3: Choose how many times to simulate the statistic — for
Textbook lingo: Number of Repetitions

Step 4: Do it! — np.append()
Textbook lingo: simulate multiple values

How to simulate multiple values

Collection array

empty array to store the simulated values/statistic

make_array()

Create a “repetitions sequence”

A sequence as long as the number of iterations

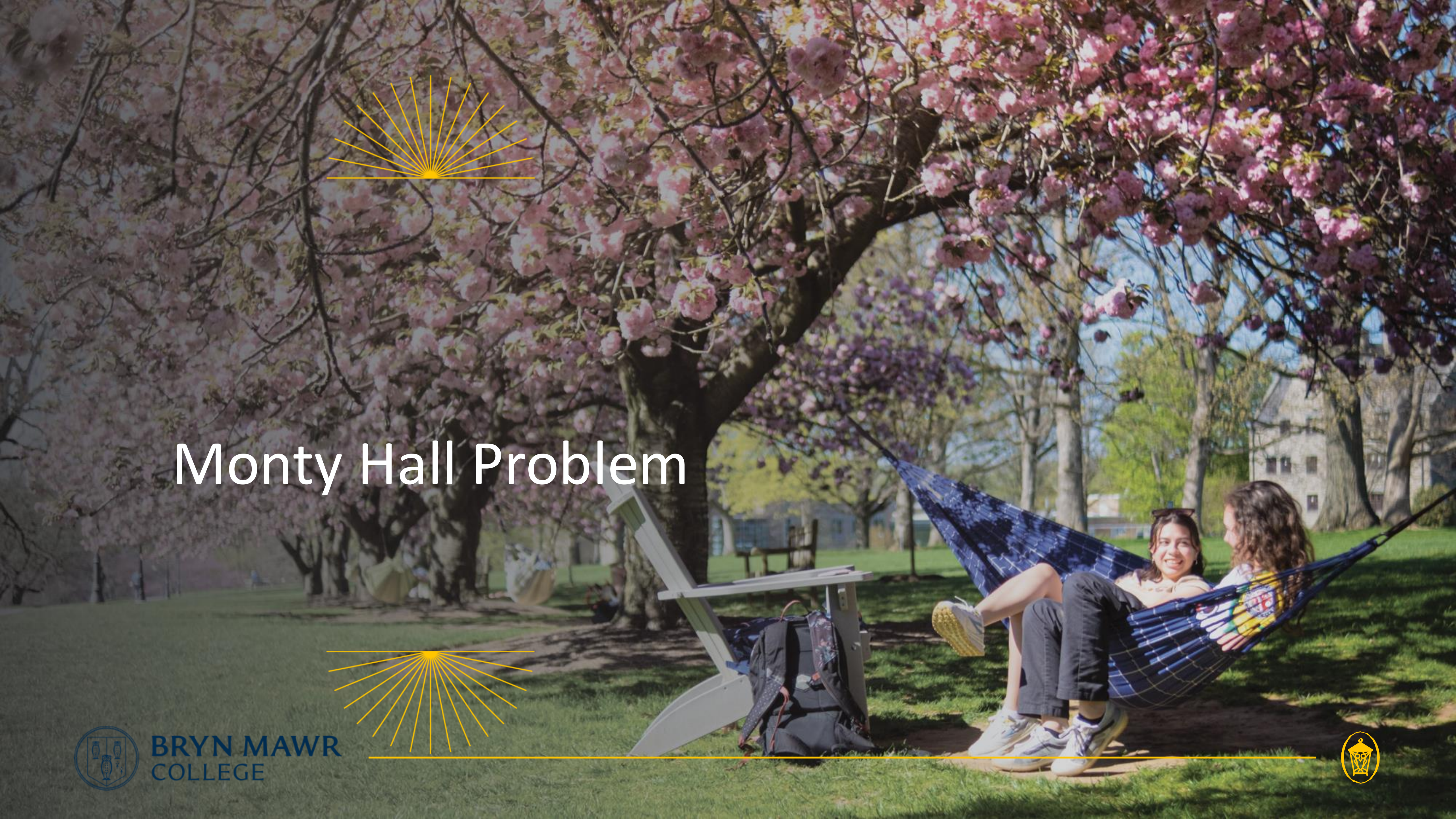
For *n* repetitions, use the sequence np.arange(n)

Create a for loop. For each element:

Simulate *one* value by using the function you wrote in Step 2

Augment the collection array with this simulated value

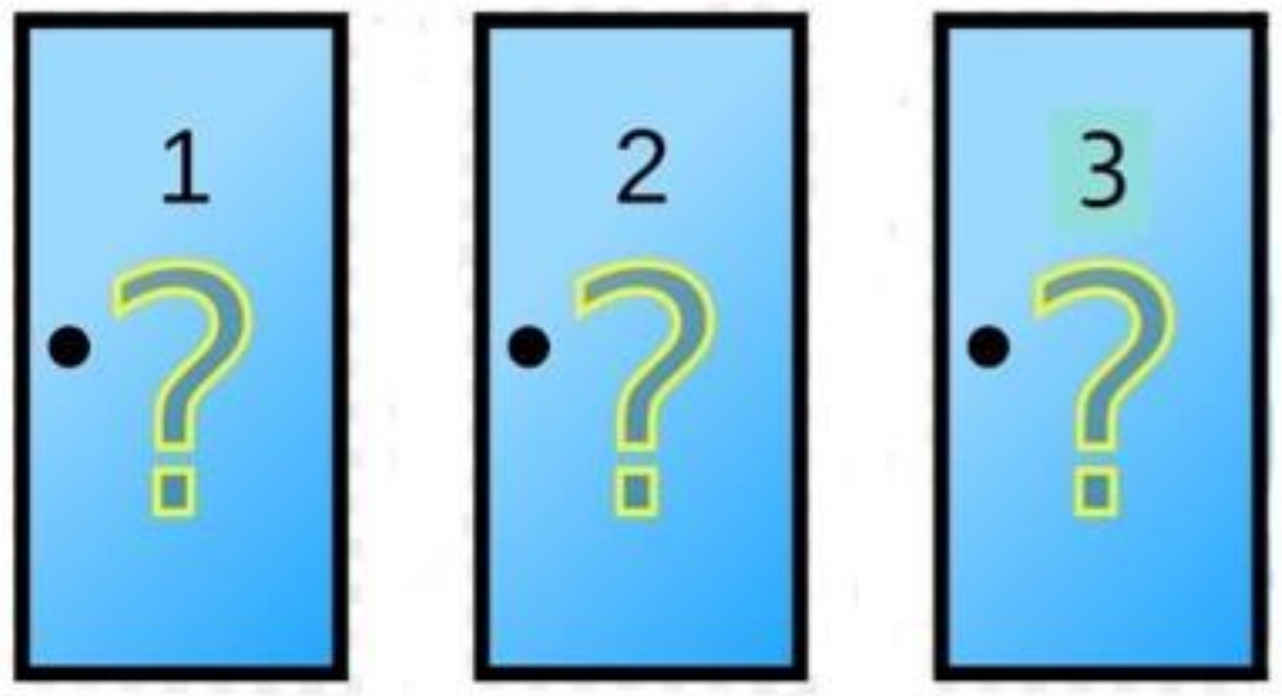
Monty Hall Problem



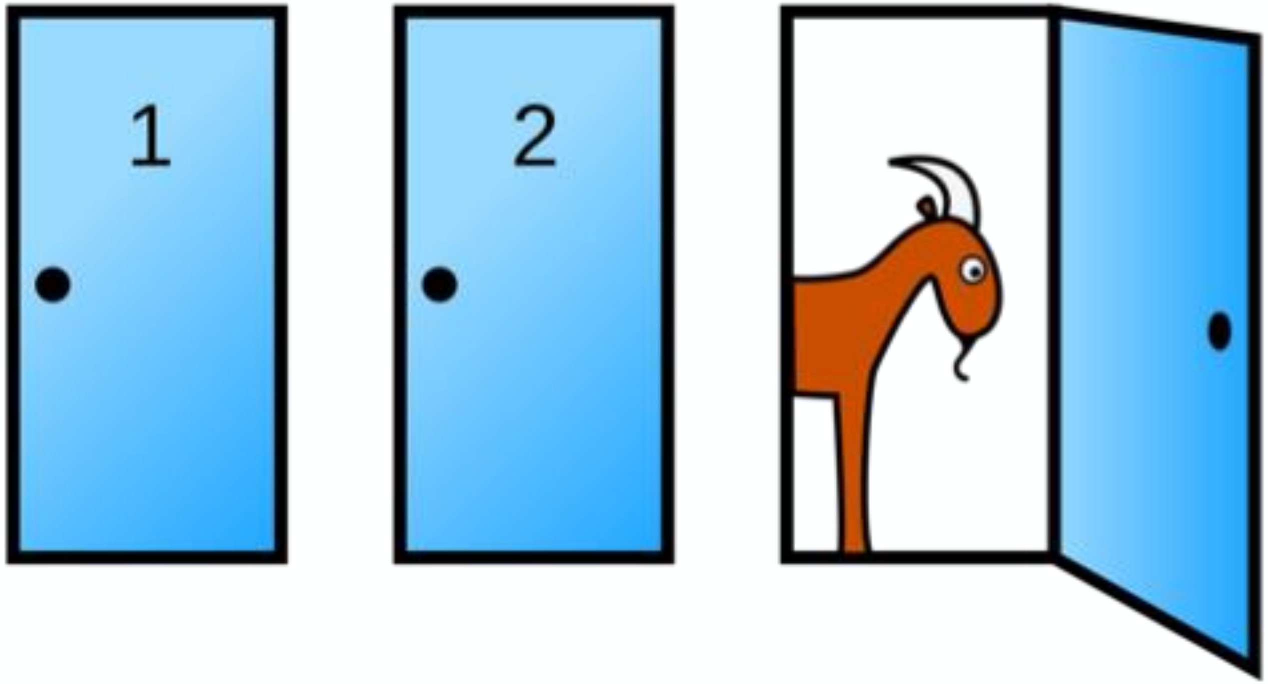
BRYN MAWR COLLEGE


LET'S
MAKE
A
DEAL

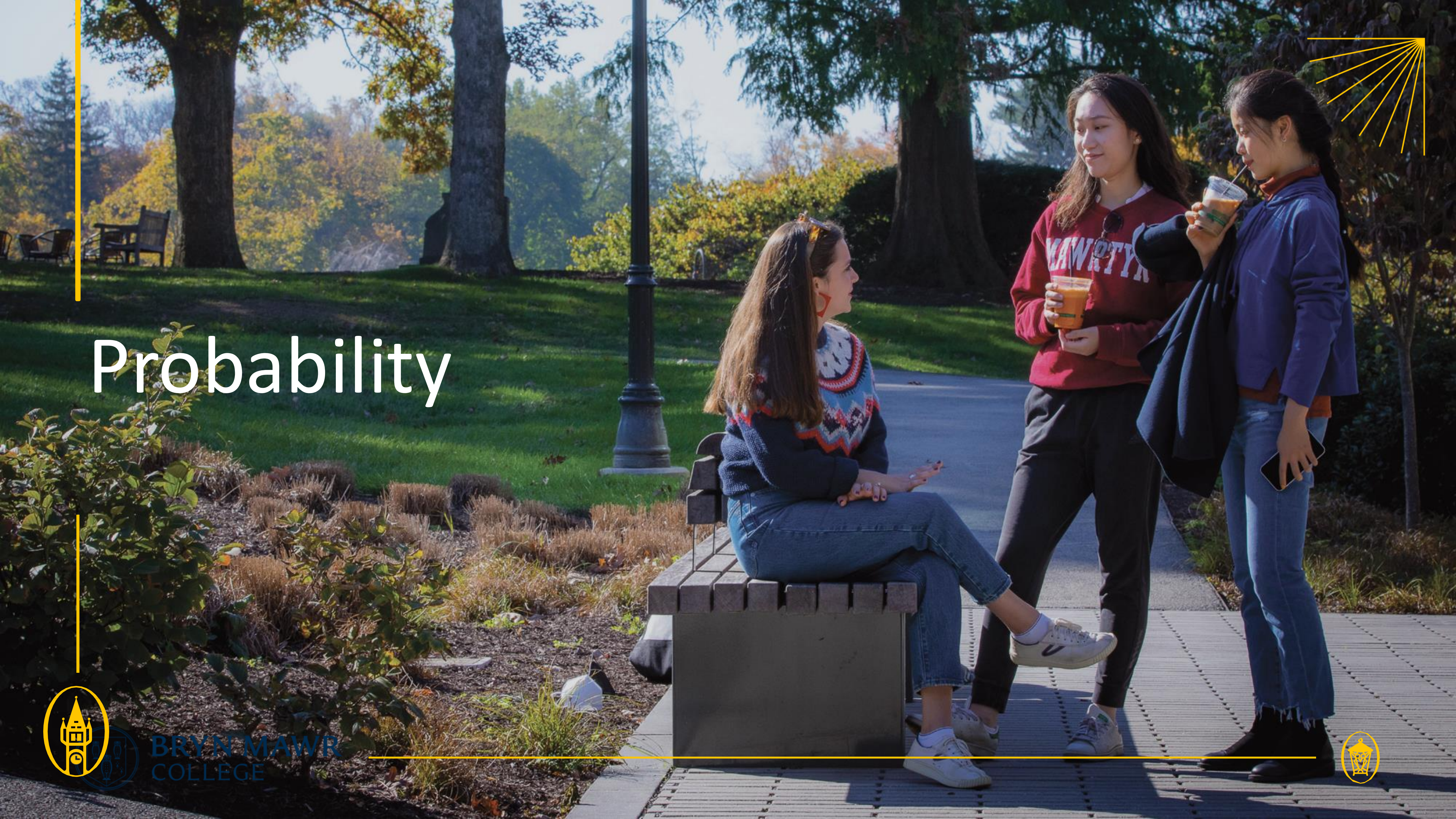
Monty Hall Problem



Monty Hall Problem



Probability



Basics

Lowest value: 0

- Chance of event that is impossible

Highest value: 1 (or 100%)

- Chance of event that is certain

If an event has chance 70%, then the chance that it doesn't happen is:

- 100% - 70% = 30%
- 1 – 0.7 = 0.3
- We call this the **Complement**

Equally Likely Outcomes

Assuming all outcomes are equally likely, the chance of an event A is:

P(A) =

$$\frac{\boldsymbol{number\ of\ outcomes\ that\ make\ A\ happen}}{\boldsymbol{total\ number\ of\ outcomes}}$$

Question

I have 3 cards: **ace of hearts**, **king of diamonds**, and **queen of spades**

I shuffle them and draw two cards at *random without replacement.*

What is the chance that I get the Queen followed by the King?

Approach 1: Enumerate all outcomes

What is the chance that I get the Queen followed by the King?

1. Queen, King
2. Queen, Ace
3. Ace, King
4. Ace, Queen
5. King, Queen
6. King, Ace

Approach 1: Enumerate all outcomes

What is the chance that I get the Queen followed by the King?

1. Queen, King
2. Queen, Ace
3. Ace, King
4. Ace, Queen
5. King, Queen
6. King, Ace

Approach 1: Enumerate all outcomes

What is the chance that I get the Queen followed by the King?

1. Queen, King
2. Queen, Ace
3. Ace, King
4. Ace, Queen
5. King, Queen
6. King, Ace

Answer: 1/ 6

Approach 2: Probabilities of the sequences

What is the chance that I get the Queen followed by the King?

What’s the probability I first draw Queen and what’s the probability I then draw King

Approach 2: Probabilities of the sequences

Step 1:

- Draw Queen from {Ace, King, Queen}
- What's the probability of drawing Queen?

1/3

Step 2:

- Draw King from {King, Ace}
- What’s the probability of drawing King?

1/2

Combining them:

- What’s 1/2 of 1/3? **1/6**

Multiplication Rule

Chance that two events *A* and *B* both happen

= P(*A* happens) x P(*B* happens given that *A* has happened)

- The answer is *less than or equal* to each of the two chances being multiplied
- The more conditions you have to satisfy, the less likely you are to satisfy them all

Addition Rule

If event *A* can happen in *exactly one* of two ways, then

$$P(A) = P(\text{first way}) + P(\text{second way})$$

The answer is *greater than or equal* to the chance of each individual way

Complement: At least one head

What the probability that I flip coins and I get at least one head?

In 3 tosses:

- Any outcome *except* TTT (tails, tails, tails)
- P(TTT) = (1/2) x (1/2) x (1/2) = 1/8
- P(at least one head) = 1 – P(TTT) = 1 – (1/8) = 87.5%

In 10 tosses:

- 1 – (1/2)**10 ≅ 99.9%