CMPU 100 · Programming with Data

Simulation and Probability

Class 19



Monty Hall hosted the gameshow *Let's Make a Deal* (1963–1986), and one of the decisions contestants were asked to make is a classic thought experiment.



Three doors



Three doors



Behind one door is a fabulous prize – a brand new car!

Three doors



Behind two doors… not a prize!

You select a door.

The host then opens one of the two doors you didn't select to reveal a goat.

There's always a goat, even if you picked one.

You are asked: "Would you like to switch to the other unopened door?"

Should you…

Stay?

Switch?

Doesn't matter!

Monty Hall Problem

Marilyn vos Savant posed the question in *Parade* magazine in 1990.

- Received 10000 answers
- 1000 from people with PhDs
- Majority disagreed with the correct answer

So, who's right?



Behind Monty Hall's Doors: Puzzle, Debate and Answer?

By JOHN TIERNEY
Special to The New York Times

BEVERLY HILLS, Calif., July 20 — Perhaps it was only an illusion, but for a moment here it seemed that an end might be in sight to the debate raging among mathematicians, readers of Parade magazine and fans of the television game show "Let's Make a Deal."

They began arguing last September after Marilyn vos Savant published a puzzle in Parade. As readers of her "Ask Marilyn" column are reminded each week, Ms. vos Savant is listed in the Guinness Book of World Records Hall of Fame for "Highest I.Q.," but that credential did not impress the public when she answered this question from a reader:

"Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the other doors, opens another door, say No. 3, which has a goat. He then says to you, 'Do you want to pick door No. 2?' Is it to your advantage to take the switch?"

Since she gave her answer, Ms. vos Savant estimates she has received 10,000 letters, the great majority disagreeing with her. The most vehement criticism has come from mathematicians and scientists, who have alternated between gloating at her ("You are the goat!") and lamenting the nation's innumeracy.

Her answer — that the contestant should switch doors — has been debated in the halls of the Central Intelligence Agency and the barracks of fighter pilots in the Persian Gulf. It has been analyzed by mathematicians at the Massachusetts Institute of Technology and computer programmers at Los Alamos National Laboratory in New Mexico. It has been tested in classes from second grade to graduate level at more than 1,000 schools across the country.

But it was not until Thursday after-

Continued on Page 20, Column 1

RECYCLE CONGRESS – LIMIT TO ONE TERM – Career Ambition inhibits judgment and stifles conscience. Seymour Durst, 1133 6th Ave. – ADVT

If you choose: 1

2

And Monty opens: 3

Should You Stick Or Switch?

Solution is on page 20.

The New York Times

1991 New York Times article

Monty Hall Problem

Marilyn vos Savant posed the question in *Parade* magazine in 1990.

- Received 10000 answers
- 1000 from people with PhDs
- Majority disagreed with the correct answer

So, who's right?

By the end of class we'll settle it the way the mathematicians were convinced – with a simulation!



Behind Monty Hall's Doors: Puzzle, Debate and Answer?

By JOHN TIERNEY
Special to The New York Times

BEVERLY HILLS, Calif., July 20 — Perhaps it was only an illusion, but for a moment here it seemed that an end might be in sight to the debate raging among mathematicians, readers of Parade magazine and fans of the television game show "Let's Make a Deal."

They began arguing last September after Marilyn vos Savant published a puzzle in Parade. As readers of her "Ask Marilyn" column are reminded each week, Ms. vos Savant is listed in the Guinness Book of World Records Hall of Fame for "Highest I.Q.," but that credential did not impress the public when she answered this question from a reader:

"Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the other doors, opens another door, say No. 3, which has a goat. He then says to you, 'Do you want to pick door No. 2?' Is it to your advantage to

er answer, Ms. vos she has received great majority dis- The most vehement from mathemati- s, who have alter- ting at her ("You

LIMIT TO ONE TERM – judgment and stifles con- 1133 6th Ave. – ADVT

If you choose: 1 2 3

And Monty opens: 3

Should You Stick Or Switch?

Solution is on page 20.

The New York Times

are the goat!") and lamenting the nation's innumeracy.

Her answer — that the contestant should switch doors — has been debated in the halls of the Central Intelligence Agency and the barracks of fighter pilots in the Persian Gulf. It has been analyzed by mathematicians at the Massachusetts Institute of Technology and computer programmers at Los Alamos National Laboratory in New Mexico. It has been tested in classes from second grade to graduate level at more than 1,000 schools across the country.

But it was not until Thursday after-

Continued on Page 20, Column 1

1991 New York Times article

Probability is a very mathematical topic, but it's one with a lot of importance to computer science and data science.

We'll just see some of the basics!

Probability

P(*A*): Chance that an event *A* will occur



Never — *P*(*A*) = 0 (or 0%)

Half — *P*(*A*) = 0.5 (or 50%)

Always — *P*(*A*) = 1 (or 100%)

***The intuition of probability**: If we repeat an event (e.g., flipping a coin) many times (say, 1 million) … what fraction of the times would we see the outcome we care about (heads)?*

How to calculate an event's probability?

Computers (simulation)

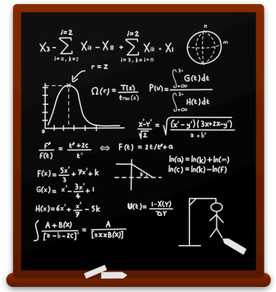
```
N = 1000000 #Roll the dice 1 million times
option_a = np.random.choice(dice, N) + np.random.c
option_b = 2 * np.random.choice(dice, N)

print("Option A Mean: ", np.mean(option_a))
print("Option B Mean: ", np.mean(option_b))

Option A Mean:  7.003198
Option B Mean:  6.99884

samples = Table().with_columns("Option A", option_
samples.hist("Option A", bins=np.arange(0,14))
samples.hist("Option B", bins=np.arange(0,14))
```



*Rooted in **algorithms***

- ✗ Approximate solutions
- ✓ Often convincing
- ✓ Non-trivial problems potentially captured cleanly with code

Math (analytical)





*Rooted in **rules (axioms)***

- ✓ Exact solutions
- ✓ Straightforward for simple problems
- ✗ Non-trivial problems potentially difficult to analyze/verify

→ Notebook

How to calculate an event's probability?

Computers (simulation)

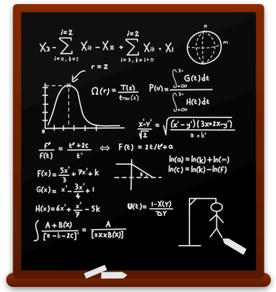
```
N = 1000000 #Roll the dice 1 million times
option_a = np.random.choice(dice, N) + np.random.c
option_b = 2 * np.random.choice(dice, N)

print("Option A Mean: ", np.mean(option_a))
print("Option B Mean: ", np.mean(option_b))

Option A Mean:  7.003198
Option B Mean:  6.99884

samples = Table().with_columns("Option A", option_
samples.hist("Option A", bins=np.arange(0,14))
samples.hist("Option B", bins=np.arange(0,14))
```



*Rooted in **algorithms***

- ✗ Approximate solutions
- ✓ Often convincing
- ✓ Non-trivial problems potentially captured cleanly with code

Math (analytical)





*Rooted in **rules (axioms)***

- ✓ Exact solutions
- ✓ Straightforward for simple problems
- ✗ Non-trivial problems potentially difficult to analyze/verify

How to calculate an event's probability?

Monty Hall Problem

Computers (simulation)

```
N = 1000000 #Roll the dice 1 million times
option_a = np.random.choice(dice, N) + np.random.c
option_b = 2 * np.random.choice(dice, N)

print("Option A Mean: ", np.mean(option_a))
print("Option B Mean: ", np.mean(option_b))

Option A Mean:  7.003198
Option B Mean:  6.99884

samples = Table().with_columns("Option A", option_
samples.hist("Option A", bins=np.arange(0,14))
samples.hist("Option B", bins=np.arange(0,14))
```



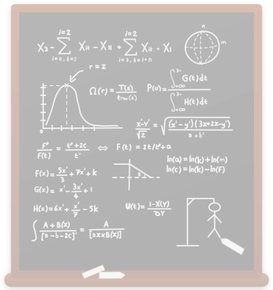
Rooted in algorithms

✗ Approximate solutions

✓ Often convincing

✓ Non-trivial problems potentially captured cleanly with code

Math (analytical)





Rooted in rules (axioms)

✓ Exact solutions

✓ Straightforward for simple problems

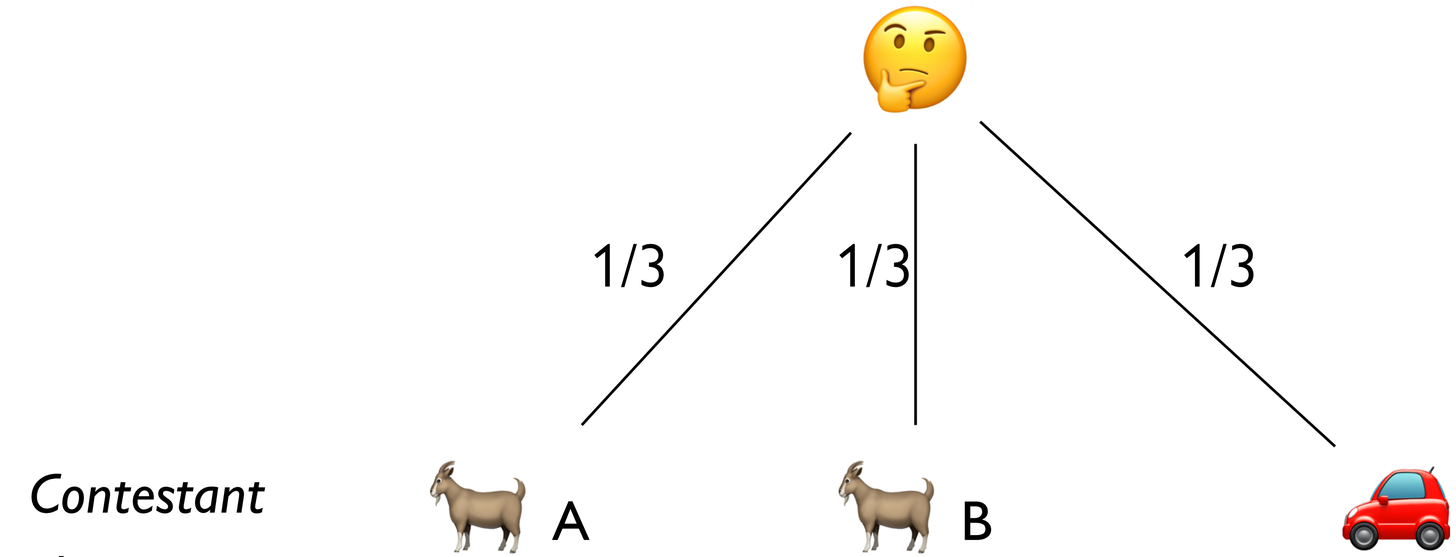
✗ Non-trivial problems potentially difficult to analyze/verify

Practice

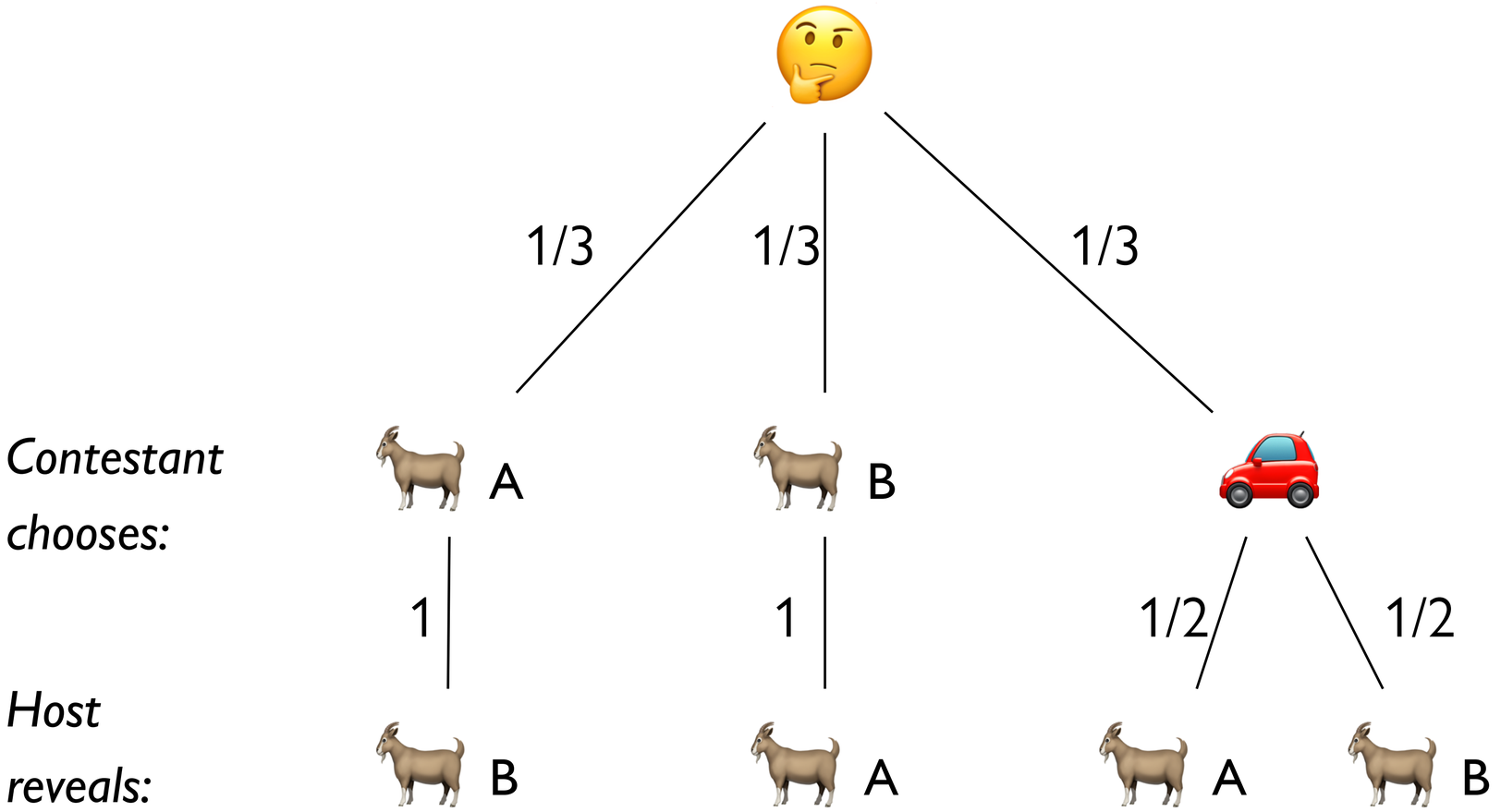
Use the probability rules to answer the Monty Hall Problem:

P(stay and win) = ?

P(switch and win) = ?



Contestant chooses:



	1/3	1/3	1/6	1/6
Stay	*A*	*B*		
Switch			*B*	*A*

Strategies

Don't switch:

$P(🚗) = 1/3 \quad P(🐐) = 2/3$

Always switch:

If you picked wrong and switch, you always win.

If you picked right and switch, you always lose

But you were more likely to pick a wrong door initially; switching improves your odds.

$P(🚗) = 2/3 \quad P(🐐) = 1/3$

How to simulate

1. Write a function for one trial that returns its outcome
2. **`simulate`** it many times
3. Analyze the outcomes

You'll get a chance to practice this for new problems in lab on Friday!

Acknowledgments

This class incorporates material from:

Data 8, University of California, Berkeley

Stephen Freund and Katie Keith, Williams College

