Probability

September 26, 2022

POLS 300: Research Methods in Politics and Public Policy
Dr. Isaac Hale
Fall Semester, 2022

OXY Occidental College

Outline

1. The basics of probability
2. Conditional probability
3. Some words of caution

The Basics of Probability

Probability

- Most events can’t be predicted with certainty
- The statistical tools you will acquire in this class will help you assess the likelihood of certain outcomes
- This is probability

Classic Examples

- When you toss a coin, what is the likelihood of each outcome (expressed as a fraction)?

- How about when you roll a six-sided die?

- Generally speaking, the probability of an event is equal to:

The number of ways an event can occur / N outcomes

- When you roll a single die, there is only way a “3” can be rolled, and only 1 die. So: (1/6) / 1 = **1/6**

Sample Size Matters

- Sample size (or “N-size”) is a central characteristic of statistics
- Think about coin flipping
- How strange would it be to flip a coin 10 times and get 6 heads?
- How strange would it be to flip a coin 1,000 times and get 600 heads?
- To illustrate, lets have Stata flip some coins for us

Distributions & Probability

- Imagine we are rolling a six-sided die
- What are the odds of getting each face?
- What kind of distribution of die faces should we expect?
- **Uniform** distribution
- What happens as we roll the die more times?

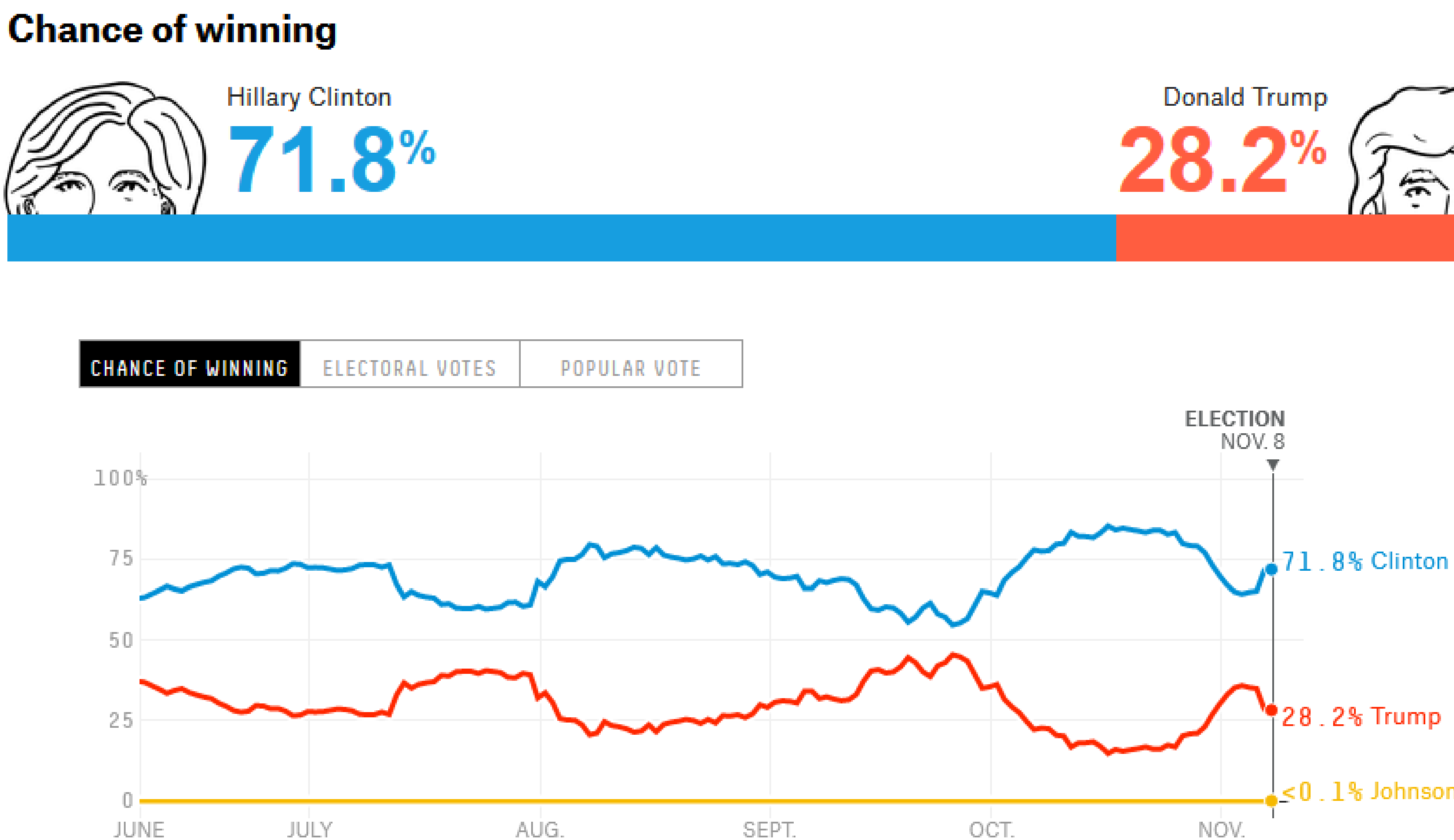
Distributions & Probability

- Instead of rolling a single D6, what if we roll 4?
- What should be the modal outcome when you add the dice together?
- 14
- Let’s do it in Stata, for fun
- What kind of distribution does this look like?
- **Normal** distribution

Probability & Politics

- Unlike with dice and coins, we don’t *a priori* know the probability of a political outcome (e.g., a candidate winning an election)
- Probability tells us what is likely to happen
- Unlike with coins or dice, a given political outcome only happens once
- We can use data to assign probabilities to outcomes, but with only one “trial,” it’s not surprising when a relatively unlikely outcome occurs

Case In Point



Expected Value

- Expected value (EV): the expected total payoff from an event
  - Calculated by summing all outcomes, with each **weighted** by its probability and payoff

- Think of the Illinois lottery example from *Naked Statistics*
  - 1 in 10 odds of free ticket (value: $1)
  - 1 in 15 odds of $2 prize
  - 1 in 42.86 odds of $4 prize
  - 1 in 75 odds of $5 prize
  - Etc., with the largest prize of $1,000 having 1 in 40,000 odds

Expected Value

- EV(lottery ticket) = (1/10) X 1 + (1/15) X 2 + (1/42.86) X 4 + (1/75) X 5 …. + (1/40,000) X 1,000
- EV(lottery ticket) = $0.56
- Shocker: you are unlikely to make money on the lottery
- What happens as you buy more tickets?
- Your money made will **converge** on 0.56x the amount you spent
- If you spend a million dollars on these tickets, the probability of your total payout will be **normally distributed** around $560,000

POV: I’m Buying a Lottery Ticket

Melanie H
@hamnox

RIP to everyone killed by the gods for their hubris but im different. and better. maybe even better than the gods

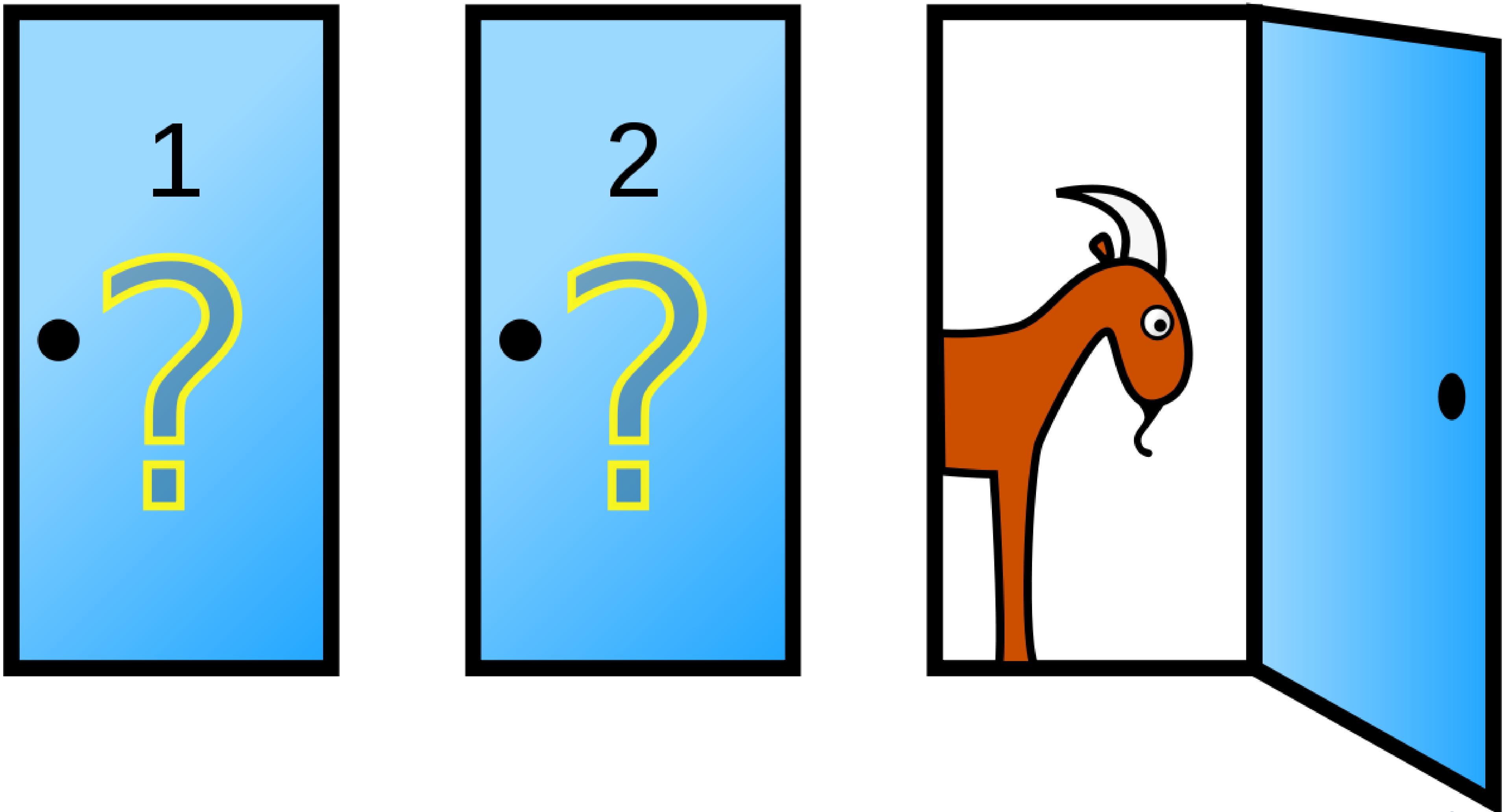
5:13 PM · Jan 14, 2022 · Twitter Web App

Conditional Probability

Conditional Probability

- Imagine two **related** possible event outcomes, A and B
- If B happens, we should **revise** our expectations about the probability of A also occurring
- $P(A \mid B)$ = the probability that event A occurs given that B has *already occurred*
- $P(A \cap B)$ = the probability that both A **and** B occur
- $\mathbf{P(A \mid B) = P(A \cap B) / P(B)}$

The Monty Hall Problem



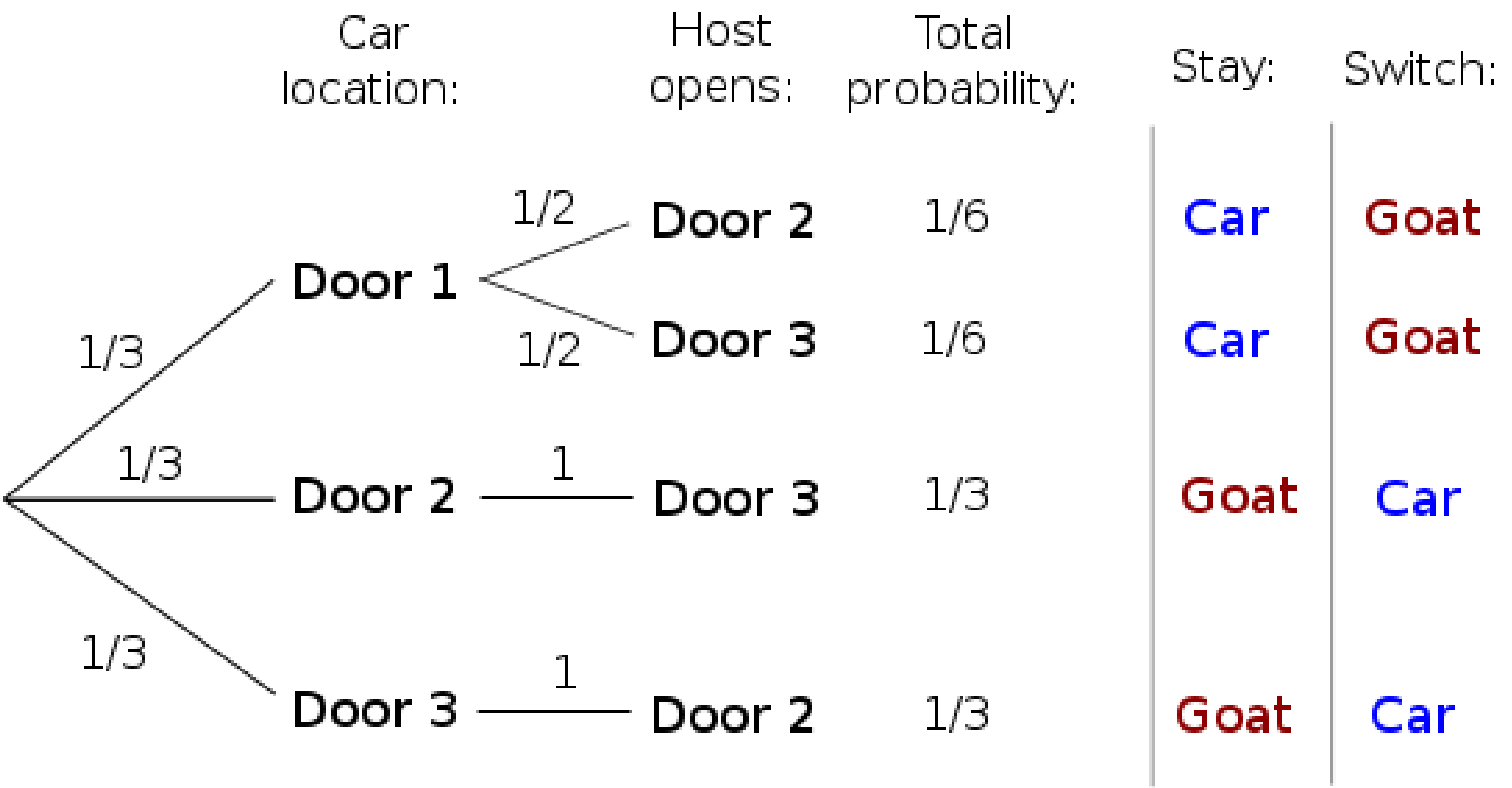
Mathing it Out

- P(prize each door) = ⅓
- You pick **door 1**. Monty now reveals a goat from door 2 or 3
- Possible states of the universe:
  - P(prize door 1, Monty opens door 2) = 1/3 X 1/2 = 1/6
  - P(prize door 1, Monty opens door 3) = 1/3 X 1/2 = 1/6
  - P(prize door 2, Monty opens door 3) = 1/3 X 1 = 1/3
  - P(prize door 3, Monty opens door 2) = 1/3 X 1 = 1/3

Mathing it Out

- Monty opens door 3, revealing a goat
- You now have two options: keep door 1 or switch to 2
- **$P(A \mid B) = P(A \cap B) / P(B)$**
- P(Keep door 1 and win) = P(prize door 1 | host door 3) = 1/6 / ½ = **1/3**
- P(Keep door 1 and lose) = P(prize door 1 | host door 3) = 1/3 / ½ = **2/3**

A Decision Tree



Conditional Probability & Politics

- Let's look at the 2022 midterm elections as an example
- Election outcomes are **related** (aka correlated) events
- If we assume Democrats win a tough Senate race in Georgia, we should assume that they will also do well in other Senate races
  - Especially in demographically similar states
- **https://projects.fivethirtyeight.com/2022-flip-senate-house/**

Limitations of Statistics & Probability

Common Mistakes with Probability

- Assuming related events are independent
- Assuming independent events are related
- Clustering
- Reversion to the mean
- Statistical discrimination
- … and more!

Alternatives to Statistics & Probabilty

- Statistics and probability are massively powerful tools

- They are **not** a replacement for other kinds of valuable research:
  - Interviews
  - Observational research
  - Archival work
  - Etc.

- Qualitative methods are often superior for **theory-building** and **exploratory** work

Matching Method to Goals

- Statistics are a great way to assess the relationships between variables and make **generalizable** claims about the world
- Statistics are only as smart as the people who use them

Q&A