

Probability Spaces

...

PL SC 309

14 January 2019

Review

Last week we learned:

- Types of data: continuous; discrete; categorical
- Data is composed of observations and variables
- Data can be represented in tabular or matrix form
- Data-generating process (DGP)

Preview

Today we'll be talking about:

- Probability space
- Everything that could happen (potential outcomes)
- Different sets of outcomes (events)
- Formalize this with sigma-notation
 - Ω : sample space
 - F : set of events
 - P : probability measure

Humans are bad at probabilistic thinking



Before probability, there were heuristics

- *Heuristic* is a fancy way of saying, if this, then that

IF

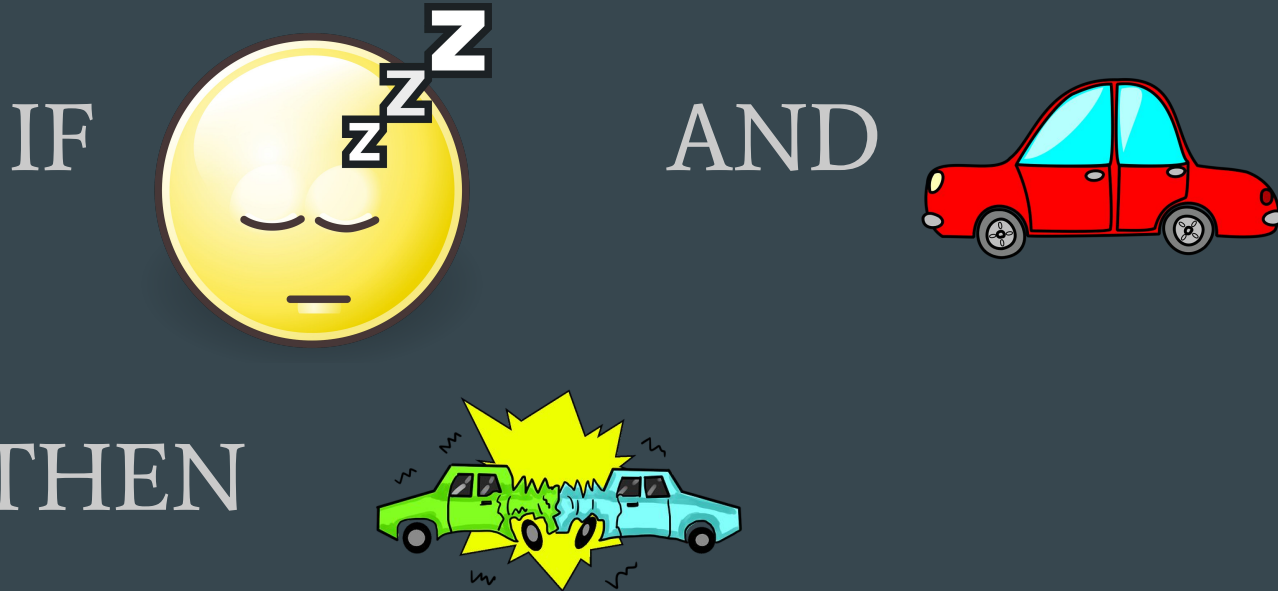


THEN



Before probability, there were heuristics

- Heuristics can be more complex, if-then statements



Before probability, there were heuristics

- Heuristics evolved when we didn't have much time to think about the future

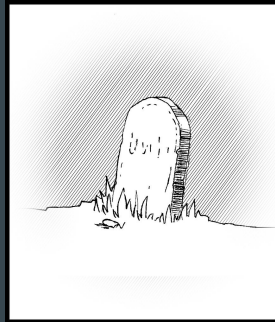
IF



AND



THEN



Problems with heuristics

- Availability heuristic - whatever is on your mind is the answer
- Representativeness heuristic - basing decisions on small, personal samples
- Affect heuristic - what you're feeling affects your predictions

Common thread: based too much on narrow information (personal experience), instead of broad information (data!)

What is probability?

- What's going to happen?
 - a. One or more things will happen
 - b. Given some information about the world...
 - c. ...what is the chance that each thing, or some combination of things, will happen?
 - d. Express that chance in a number between 0 and 1
 - 0 means no chance of it happening
 - 1 means it is guaranteed

What is statistics?

Statistics is the application of probability to quantified information.

What is statistics?

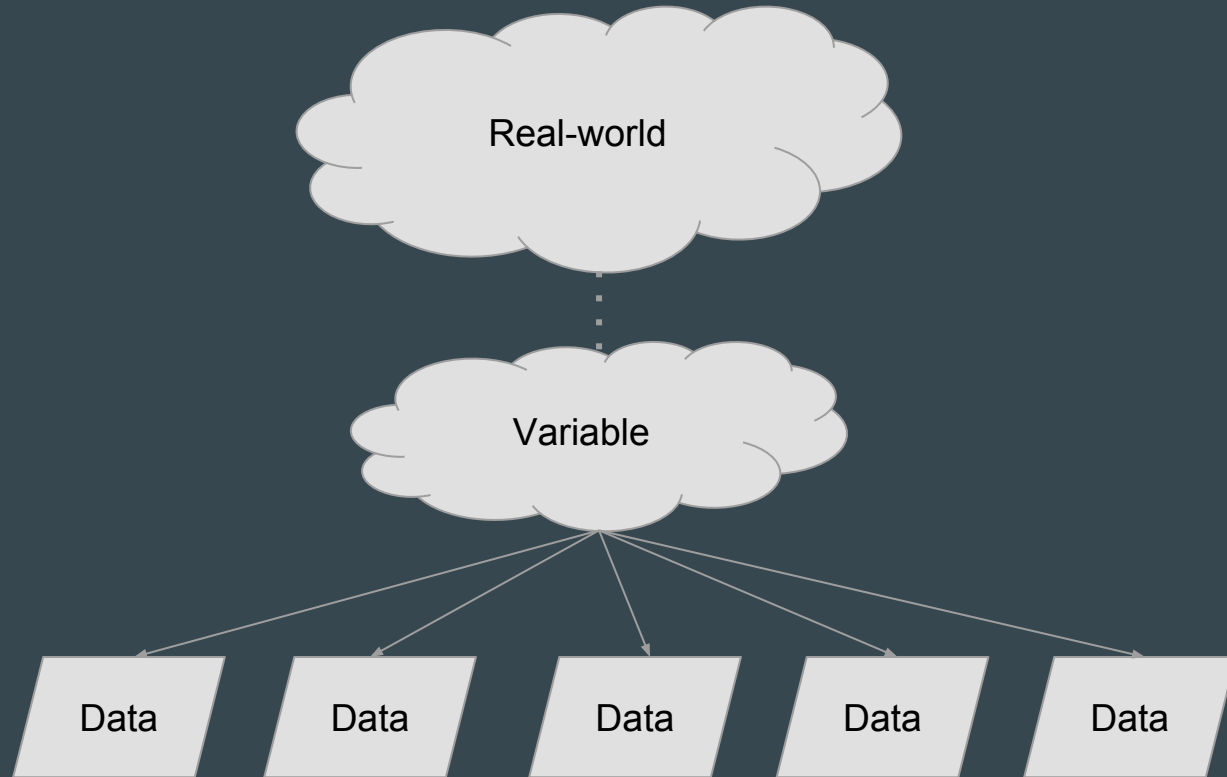
Statistics is the application of probability to quantified information.



What will happen in the future...

Hypothetically, if we did this a thousand times...

What's a variable?



Let's go back to observations and variables



Let's go back to observations and variables



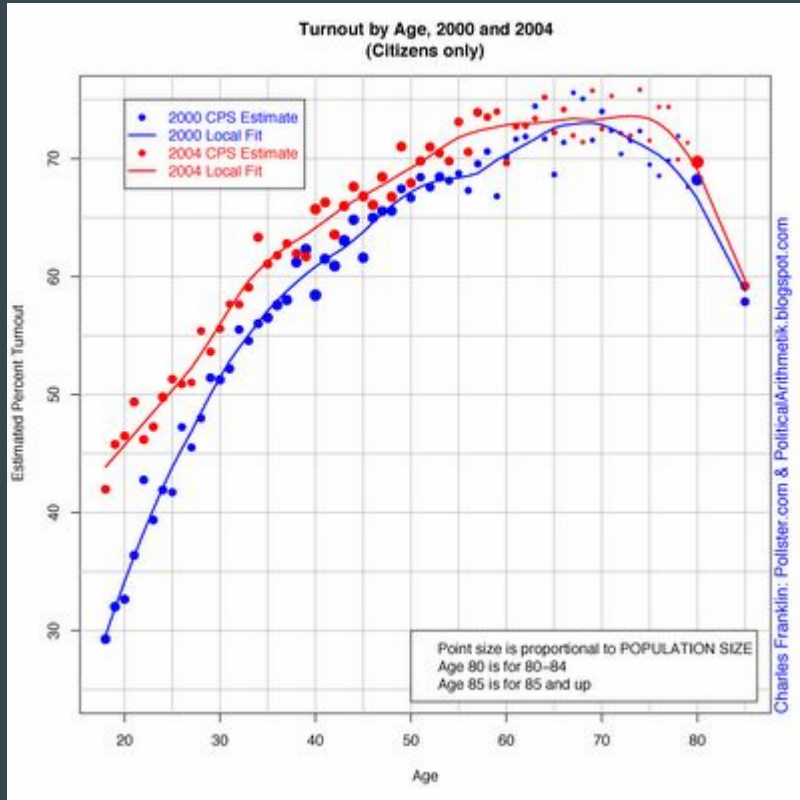
Let's go back to observations and variables



Let's go back to observations and variables

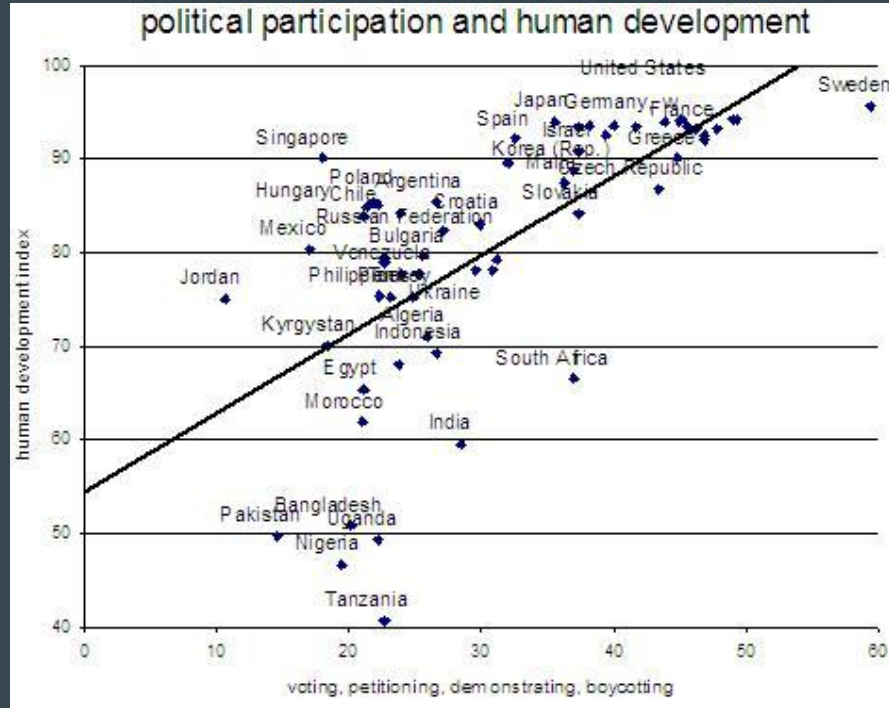


Let's go back to observations and variables



- Statistics is about thinking probabilistically, which is about thinking *infinitely*
- In other words, each point (or observation) is an expression of a variable, which could possibly exist at any location, with differing degrees of probability

Let's go back to observations and variables



- We start out with the idea that there is some thing called voting, and some thing called human development
- And that these can take a certain range of values
- We then *observe*, for some units, a value for each of these *variables*

Let's go back to observations and variables

Observations are each **point**, variables are the **value of that point**

Observations are an **observed outcome**, variables are the **value of that outcome**

Potential outcomes and observations

If observations are *observed outcomes*, they are part of a larger set of potential outcomes...



Heads

Tails

Potential
outcomes

Potential outcomes and observations

If observations are *observed outcomes*, they are part of a larger set of potential outcomes...

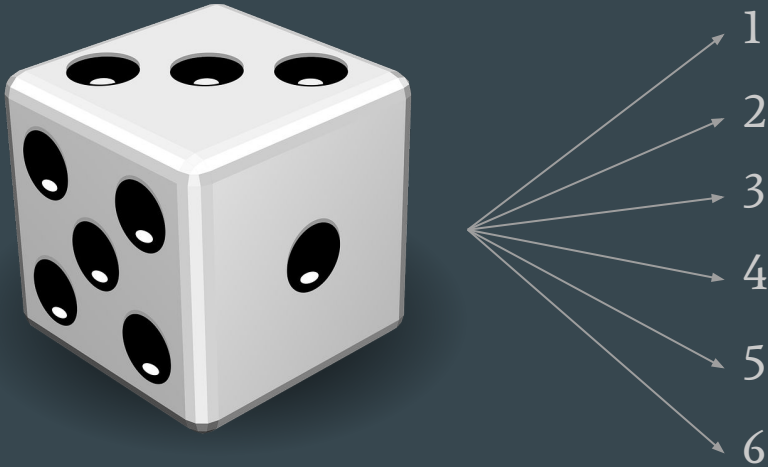


Flip no.	Result
1	H
2	H
3	T
4	H
5	T
6	T
7	T

Observations

Outcomes

We can write outcomes as a tree diagram



Outcomes

Or we can write outcomes as a set


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

Probability Spaces

A defined space that encompasses all possibilities and their likelihood

Ω : sample space

F : set of events

P : probability measure

Ω : sample space

All possible outcomes, or *elementary events*



1	2
3	4
5	6

F : set of events

Any event or combination of events



Roll a 1

1	2
3	4
5	6

F : set of events

Any event or combination of events



Roll less than 3

1	2
3	4
5	6

F : set of events

Any event or combination of events



1	2
3	4
5	6

Roll an even number

P: probability

A number between 0 and 1 for each F



1 ($\frac{1}{6}$)	2 ($\frac{1}{6}$)
3 ($\frac{1}{6}$)	4 ($\frac{1}{6}$)
5 ($\frac{1}{6}$)	6 ($\frac{1}{6}$)

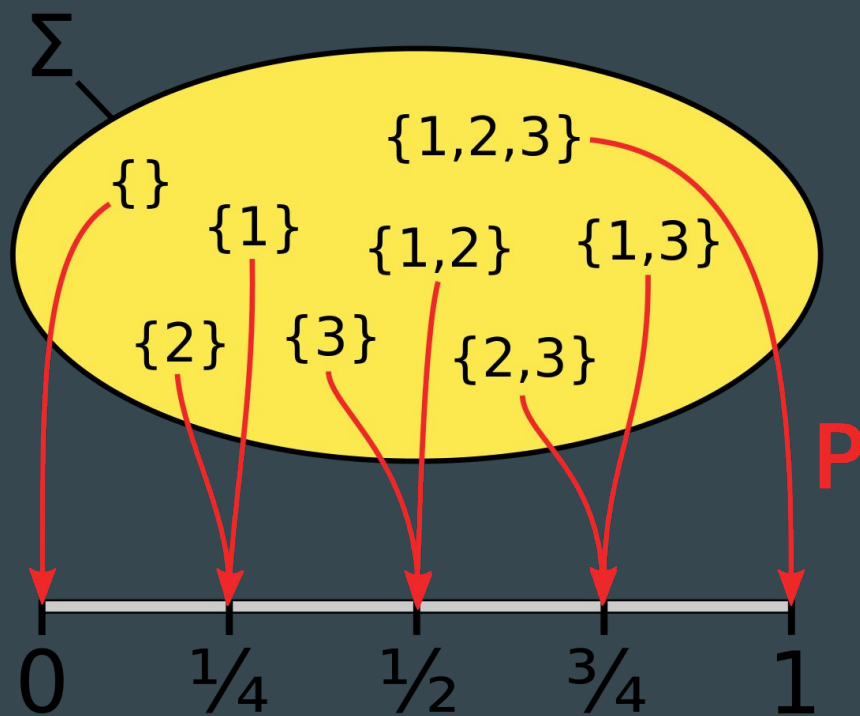
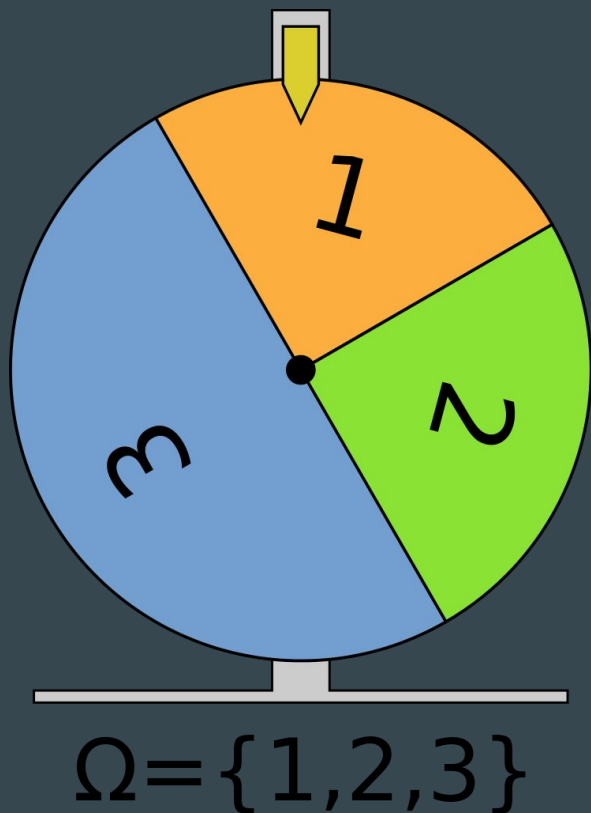
P: probability

A number between 0 and 1 for each F







Odd ($\frac{1}{2}$)	Even ($\frac{1}{2}$)
-----------------------	------------------------

Example of a probability space



Example of a probability space

3:05 p.m. Eastern		ELO POINT SPREAD	WIN PROB.	SCORE	6:40 p.m.		ELO POINT SPREAD	WIN PROB.	SCORE
	L.A. Rams		36%			New England		39%	
	New Orleans	- 4	64%			Kansas City	- 3	61%	

Ω : {rams win, saints win}

F : {rams win, saints win}

P: predictions for team's victory

Ω : {chiefs win, patriots win}

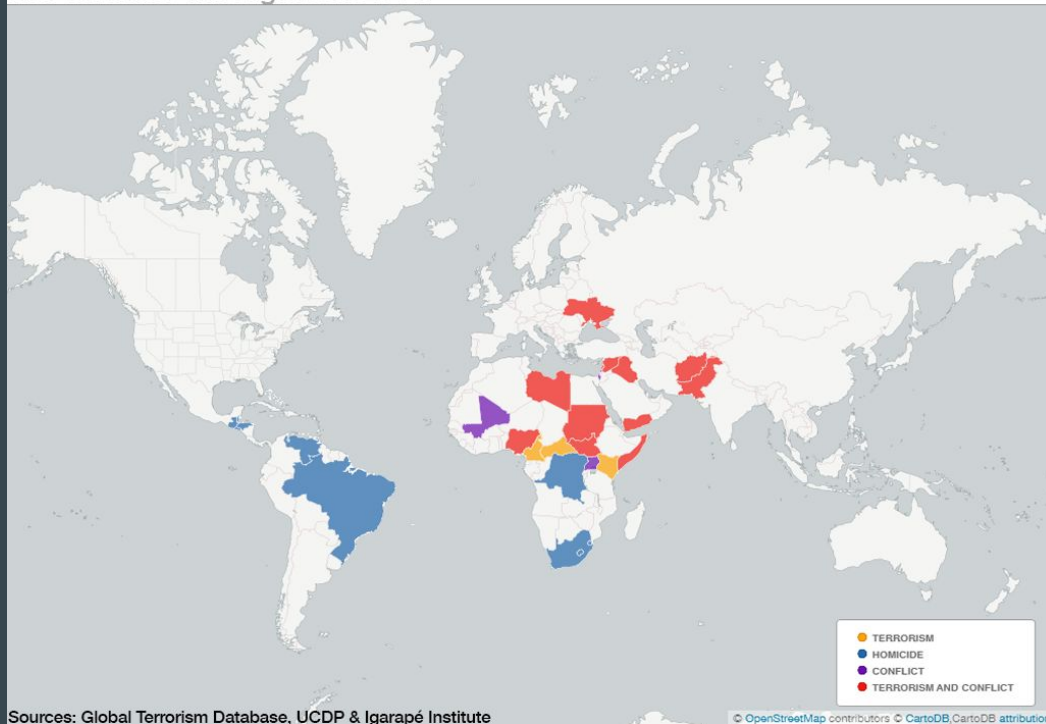
F : {chiefs win, patriots win}

P: predictions for team's victory

Example of a probability space

Where is the risk of lethal violence highest?

These are the countries where your risk of death from terrorism, conflict and homicide was highest in 2014.



Sources: Global Terrorism Database, UCDP & Igarapé Institute

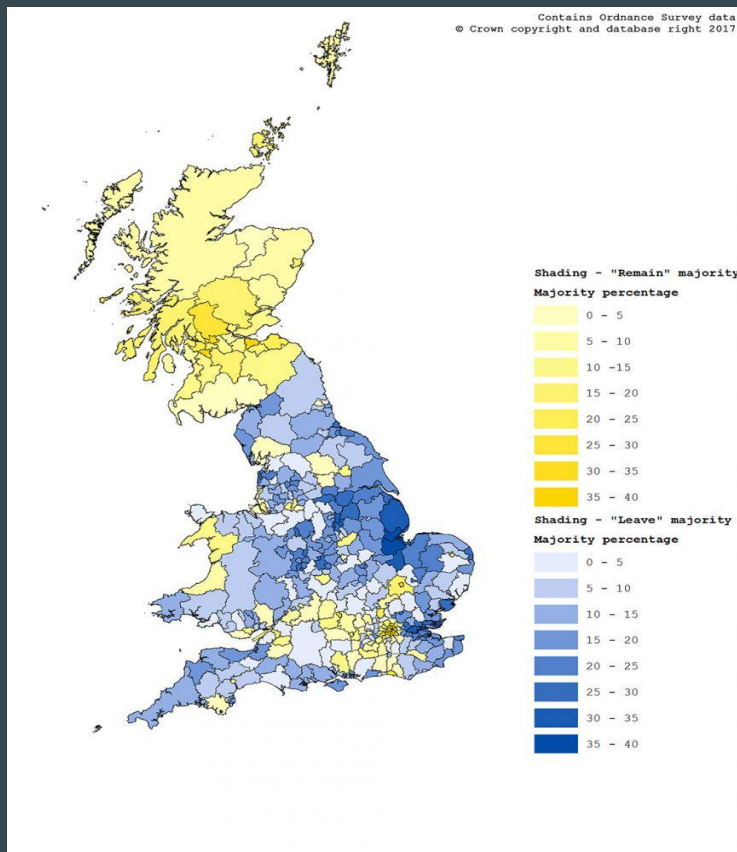
© OpenStreetMap contributors © CartoDB CartoDB attribution

Ω : each country's risk of violence

F : each different possible combination of countries with the highest risk

P : probability of each country's risk

Example of a probability space

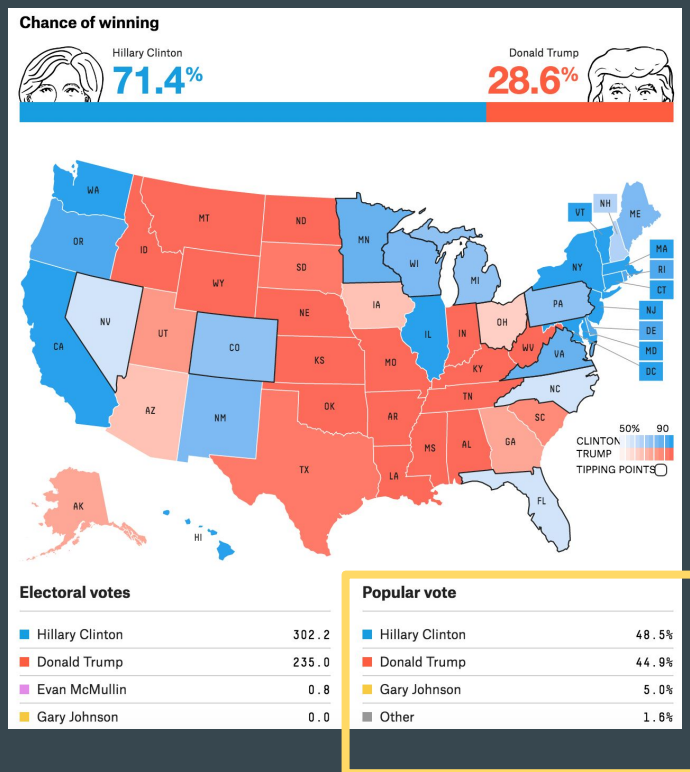


Ω : percentage of people voting
leave, remain

F : all combinations of voting

P : probability of vote

Example of a probability space

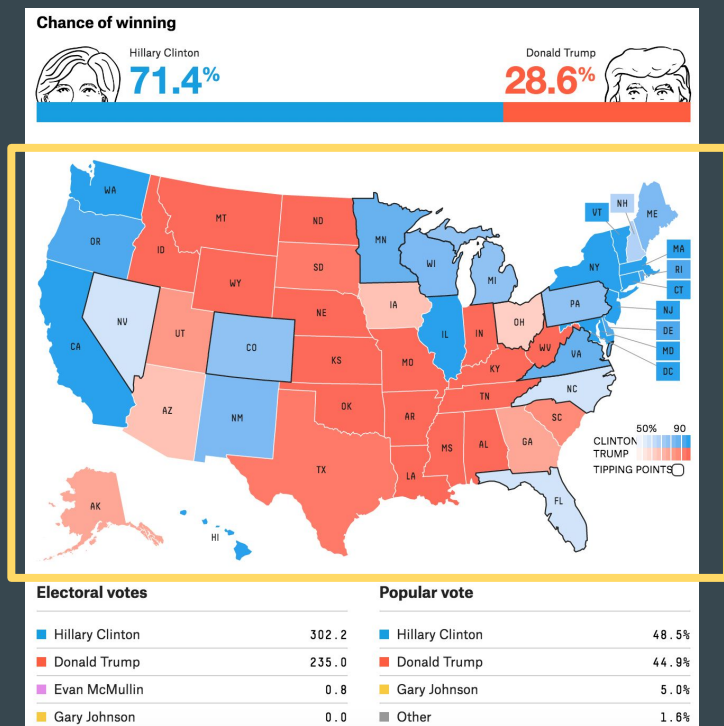


Ω : trump or clinton national vote percentage

F : all possible national votes

P : model predictions for country-wide victory

Example of a probability space



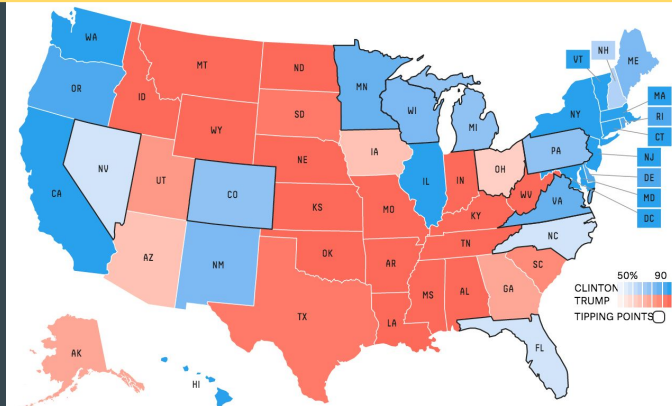
Ω : trump or clinton victory in each state

F : all possible state level results

P : model predictions for state-wide victory

Example of a probability space

Chance of winning



Electoral votes

Hillary Clinton	302.2
Donald Trump	235.0
Evan McMullin	0.8
Gary Johnson	0.0

Popular vote

Hillary Clinton	48.5%
Donald Trump	44.9%
Gary Johnson	5.0%
Other	1.6%

Ω : trump or clinton victory in each state

F : combination of state victories to win electoral college

P : model predictions for winning electoral college

Review

- Probability assigns a measure to how likely certain sets of outcomes are to happen
- Statistics is the application of probability to quantified information
- Probabilistic thinking is about thinking in terms of all possible outcomes, repeated infinitely
- Sigma-notation
 - Ω : sample space
 - F : set of events
 - P : probability measure