Political Science 209 - Fall 2018

Probability

Florian Hollenbach

26th October 2018

Why probability?

- Probability rules our lives
- It is everywhere!

Why probability?

- Humans are really bad at interpreting probabilities
- Even worse at calculating (estimating) probabilities

Why probability?

The Media Has A Probability Problem

The media's demand for certainty — and its lack of statistical rigor — is a bad match for our complex world.

By Nate Silver
Filed under The Real Story Of 2016
Published Sep. 21, 2017

Why probability?

- What are the chances it rains tomorrow?

Why probability?

- What are the chances it rains tomorrow?
- What are the chances you win the lottery?

Why probability?

- What are the chances it rains tomorrow?
- What are the chances you win the lottery?
- What is the probabilty of getting an A in pols 209?

Why probability?

- We use probability to express and calculate uncertainty
- *Preview*: later we will use probability to make statements about the uncertainty in our data analysis

Two fundamental concepts of probability

- Frequentist: long-run frequency of events
  - ratio between the number of times the event occurs and the number of trials
  - example: coin flips

Two fundamental concepts of probability

- Frequentist: long-run frequency of events
    - ratio between the number of times the event occurs and the number of trials
    - example: coin flips
- Bayesian: belief about the likelihood of event occurrence
    - evidence based belief
    - often more sensible philosophy in political world

Important Terms

1. Experiment: an action or a set of actions that produce stochastic events of interest

Important Terms

1. Experiment: an action or a set of actions that produce stochastic events of interest

1. sample space: a set of all possible outcomes of the experiment, typically denoted by Ω

Important Terms

1. Experiment: an action or a set of actions that produce stochastic events of interest

1. sample space: a set of all possible outcomes of the experiment, typically denoted by Ω

1. event: a subset of the sample space

(Imai - QSS)

Example

What is the experiment, sample space, and one event for coin flips or pulling a single card out of a deck of 52?

Defining Probability

$$\text{Probability of event A} = P(A) = \frac{\text{number of elements in A}}{\text{number of elements in sample space}}$$

Defining Probability

$$\text{Probability of event A} = P(A) = \frac{\text{number of elements in A}}{\text{number of elements in sample space}}$$

$$\text{Probability of Head} = P(H) = \frac{1}{2}$$

Example

What is the probability of 3 head in 3 flips?

Sample space?

Example

What is the probability of 3 head in 3 flips?

Sample space?

$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

Example

What is the probability of 3 head in 3 flips?

Sample space?

$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

What is the event space we are interested in?

Example

What is the probability of 3 head in 3 flips?

Sample space?

$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

What is the event space we are interested in?

$\{HHH\}$

Example

What is the probability of 3 head in 3 flips?

Example

What is the probability of 3 head in 3 flips?

$P(HHH) = \frac{1}{8}$

Example

What is the probability of 2 head in 3 flips?

$\Omega = \{$HHH,HHT,HTH,THH, HTT, THT, TTH, TTT$\}$

What is the event space we are interested in?

Example

What is the probability of 2 head in 3 flips?

$\Omega = \{$HHH,HHT,HTH,THH, HTT, THT, TTH, TTT$\}$

What is the event space we are interested in?

$\{$HHT, HTH, THH$\}$

Example

What is the probability of 2 head in 3 flips?

$\Omega = \{\text{HHH,HHT,HTH,THH, HTT, THT, TTH, TTT}\}$

What is the event space we are interested in?

$\{\text{HHT, HTH, THH}\}$

$\text{P(2 H)} = \frac{3}{8}$

Axioms (rules) of Probability

- the probability of any event A is at least 0
  - $P(A) \geq 0$

Axioms (rules) of Probability

- the probability of any event A is at least 0
  - $P(A) \geq 0$

- The total sum of all possible outcomes in the sample space must be 1
  - $P(\Omega) = 1$

Axioms (rules) of Probability

- the probability of any event A is at least 0
    - $P(A) \geq 0$

- The total sum of all possible outcomes in the sample space must be 1
    - $P(\Omega) = 1$

- If A and B are mutually exclusive (meaning only one or the other can happen), then $P(A \text{ or } B) = P(A) + P(B)$

Axioms (rules) of Probability

A^c - complement to A, i.e. part of sample space not in A

Sometimes it is easier to calculate the probability of an event by using its complement

Using the complement:

What is the probability of having at least one Tail on three coin flips?

$\Omega = \{$HHH,HHT,HTH,THH, HTT, THT, TTH, TTT$\}$

Using the complement:

What is the probability of having at least one Tail on three coin flips?

Ω = {HHH,HHT,HTH,THH, HTT, THT, TTH, TTT}

P(at least one T) = $\frac{7}{8}$

P(at least one T) = 1 - P(HHH) = 1 - $\frac{1}{8}$

Example of simple probability

What is the probability of getting a Queen as the first card from a full deck?

$\Omega = \{?\}$

Event space $= \{?\}$

Example of simple probability

What is the probability of getting a Queen as the first card from a full deck?

$\Omega = \{?\}$

Event space $= \{?\}$

p(Queen) $= \frac{4}{52} = \frac{1}{13}$

How to quickly count the sample space when order matters: permutations

- Often we do not want to or can't write out all possible combinations by hand
- How many possibilities are there to arrange letters A,B,C?

How to quickly count the sample space when order matters: permutations

- Often we do not want to or can't write out all possible combinations by hand
- How many possibilities are there to arrange letters A,B,C?

Three outcomes: A, B, C & three draws

How to quickly count the sample space when order matters: permutations

- Often we do not want to or can't write out all possible combinations by hand
- How many possibilities are there to arrange letters A,B,C?

Three outcomes: A, B, C & three draws

First draw: A,B, or C

Second draw: two possibilities

Third draw: one left

3 x 2 x 1 possibilities

How to quickly count the sample space when order matters: permutations

Permutations count many ways we can order k objects out of a set of n unique objects

$_nP_k = n \times (n-1) \times (n-2) \times ... \times (n-k+1) = \frac{n!}{(n-k)!}$

What does n! stand for?

How to quickly count the sample space when order matters: permutations

Permutations count many ways we can order k objects out of a set of n unique objects

${}_nP_k = n \times (n-1) \times (n-2) \times ... \times (n-k+1) = \frac{n!}{(n-k)!}$

What does n! stand for?

n! = n-factorial = $n \times (n-1) \times (n-2) \times ... \times (n-n+1)$

$3! = 3 \times 2 \times 1$

Note: 0! = 1

Permutation Example:

How many ways can we arrange four cards out of a the 13 spades in our card deck?

first draw: ?

Permutation Example:

How many ways can we arrange four cards out of a the 13 spades in our card deck?

first draw: ?

$13 \times 12 \times 11 \times 10$

Permutation Example:

How many ways can we arrange four cards out of a the 13 spades in our card deck?

first draw: ?

$13 \times 12 \times 11 \times 10$

$\frac{13!}{(13-4)!} = \frac{13!}{9!} = \frac{13 \times 12 \times 11 \times \ldots \times 2 \times 1}{9 \times 8 \times \ldots \times 2 \times 1} = 13 \times 12 \times 11 \times 10 = 17,160$

Birthday Problem

Impress your family over Thanksgiving!

Birthday Problem

Impress your family over Thanksgiving!

What is the probability that at least two people in this room have the same birthday?

How could we figure that out?

Birthday Problem

Can the law of total probabilities and complement help us?

Birthday Problem

Can the law of total probabilities and complement help us?

Yes, P(at least two share bday) = 1 - P(nobody shares bday)

Birthday Problem

P(nobody shares bday)?

What is the event space?

Birthday Problem

P(nobody shares bday)?

What is the event space?

Event space: everyone has a unique birthday. How many different possibilities?

Birthday Problem

P(nobody shares bday)?

What is the event space?

Event space: everyone has a unique birthday. How many different possibilities?

How many possibilities for birthdays in a year?

Birthday Problem

P(nobody shares bday)?

What is the event space?

Event space: everyone has a unique birthday. How many different possibilities?

How many possibilities for birthdays in a year?

365

Birthday Problem

P(nobody shares bday)?

What is the event space?

Event space: everyone has a unique birthday. How many different possibilities?

How many possibilities for birthdays in a year?

365

How many unique arrangements would we need for nobody to share the birthday?

number of people in room - k

Birthday Problem

1. ${}_{365}P_k = \frac{365!}{(365-k)!}$ possibilities to arrange k unique birthdays over 365 days
2. What is the sample space? All the different possibilities for k birthdays (even non-unique).

Birthday Problem

1. ${}_{365}P_k = \frac{365!}{(365-k)!}$ possibilities to arrange k unique birthdays over 365 days
2. What is the sample space? All the different possibilities for k birthdays (even non-unique).

365^k

Birthday Problem

P(at least two share bday) = 1 - P(nobody shares bday) = 1 - $\frac{365!}{(365-k)! \times 365^k}$

Birthday Problem

P(at least two share bday) = 1 - P(nobody shares bday) = 1 - $\frac{365!}{(365-k)! \times 365^k}$

P(at least two share bday):

k = 10; 0.116,

k = 23; 0.504,

and k = 68; 0.999.

Combinations

Combinations are similar to permutations, except that the ordering doesn't matter

So with respect to combinations of 3 out of 26 letters, ABC, BAC, CAB, etc are the same

Combinations

Combinations are similar to permutations, except that the ordering doesn't matter

So with respect to combinations of 3 out of 26 letters, ABC, BAC, CAB, etc are the same

There are always fewer combinations than permutations

Combinations vs. Permutations

Draw 2 out of letters ABC

Permutations:

Combinations vs. Permutations

Draw 2 out of letters ABC

Permutations:

AB, AC, BA, BC, CA, CB $= \frac{3!}{1!}$

Combinations:

Combinations vs. Permutations

Draw 2 out of letters ABC

Permutations:

AB, AC, BA, BC, CA, CB $= \frac{3!}{1!}$

Combinations:

AB, AC, BC

How to Calculate Combinations

Calculate permutations and then account for the fact that we overcounted due to ordering

Get rid of counts of different arrangements of same combination: divide by k!

$_nC_k = \binom{n}{k} = \frac{_nP_k}{k!} = \frac{n!}{k!(n-k)!}$

How to Calculate Combinations

Calculate permutations and then account for the fact that we overcounted due to ordering

Get rid of counts of different arrangements of same combination: divide by k!

$_nC_k = \binom{n}{k} = \frac{_nP_k}{k!} = \frac{n!}{k!(n-k)!}$

Why divide by k! ?

How to Calculate Combinations

Calculate permutations and then account for the fact that we overcounted due to ordering

Get rid of counts of different arrangements of same combination: divide by k!

$_nC_k = \binom{n}{k} = \frac{_nP_k}{k!} = \frac{n!}{k!(n-k)!}$

Why divide by k! ?

for two sampled elements, we have $2!(= 2 \times 1 = 2)$: A, B = AB, BA

Lottery

What is the probability of winning (simplified) Mega Millions?

Pick five numbers between 1 and 70

Probability of getting 5 correct?

Lottery

Probability of getting 5 correct?

What is the size of the event space?

Lottery

Probability of getting 5 correct?

What is the size of the event space?

1 ticket

Lottery

Pick five numbers between 1 and 70

Sample space?

Lottery

Pick five numbers between 1 and 70

Sample space?

$\binom{70}{5} = \frac{70!}{5! \times (70-5)!} = \frac{70!}{5! \times 65!}$

Lottery

Pick five numbers between 1 and 70

Sample space?

$\binom{70}{5} = \frac{70!}{5! \times (70-5)!} = \frac{70!}{5! \times 65!}$

12,103,014

Lottery

$\binom{n}{k}$ in R

choose(n,k)

```
choose(70,5)
```

```
[1] 12103014
```

Samping *with* and *without* Replacement

Two ways to sample (draw) data:

- with replacement: put draw back in box
- without replacement: keep draw, ticket can not be drawn again

Samping *with* and *without* Replacement

Two ways to sample (draw) data:

- with replacement: put draw back in box
- without replacement: keep draw, ticket can not be drawn again

If we are sampling for a survey, what technique do we use?

Simulating the birthday problem in *R*

- Instead of calculating probabilities, we can often simulate them in *R*
- Use *R* to draw k birthdays and see whether any duplicates exist

Simulating the birthday problem in *R*

- Instead of calculating probabilities, we can often simulate them in *R*
- Use *R* to draw k birthdays and see whether any duplicates exist

- We repeat the experiment over and over (~ 1000 times). The share of experiments in which we found duplicates, will represent P(at least one shared bday)

Simulating the birthday problem in *R*

```
k <- 23 # number of people
sims <- 1000 # number of simulations
event <- 0 # counter
for (i in 1:sims) {
    days <- sample(1:365, k, replace = TRUE)
    days.unique <- unique(days) # unique birthdays
    if (length(days.unique) < k) {
        event <- event + 1 } }
event / sims
```

```
[1] 0.499
```

Simulating the birthday problem in *R*

The larger the number of simulation iterations, the better the accuracy

```
sims <- 10000 # number of simulations
event <- 0 # counter
for (i in 1:sims) {
    days <- sample(1:365, k, replace = TRUE)
    days.unique <- unique(days) # unique birthdays
    if (length(days.unique) < k) {
        event <- event + 1  }}
event / sims
```

```
[1] 0.5181
```