Econ 103 – Statistics for Economists

Chapter 3: Probability

Mallick Hossain

University of Pennsylvania

R Project

Motivation

- Apply the skills and tools you have learned

Motivation

- Apply the skills and tools you have learned
    - Sidebar on me and metrics

Motivation

- Apply the skills and tools you have learned
  - Sidebar on me and metrics
- Answer questions you are interested in

Motivation

- Apply the skills and tools you have learned
    - Sidebar on me and metrics
- Answer questions you are interested in
- Head-start on honors thesis?

What Subjects Can I Explore?

- ANYTHING

What Subjects Can I Explore?

- ANYTHING, seriously.
  - Political or ideological agendas are welcome

What Subjects Can I Explore?

- ANYTHING, seriously.
  - Political or ideological agendas are welcome
  - Guns, gender, inequality, race, climate change, …

What Subjects Can I Explore?

- ANYTHING, seriously.
    - Political or ideological agendas are welcome
    - Guns, gender, inequality, race, climate change, ...
    - Google trends, Twitter, financial data, macro data, ...

What Do I Turn In?

- Summary of question

What Do I Turn In?

- Summary of question
- Summary of data and why it is relevant for the question

What Do I Turn In?

- Summary of question
- Summary of data and why it is relevant for the question
- Tables and charts of summary stats of the data

What Do I Turn In?

- Summary of question
- Summary of data and why it is relevant for the question
- Tables and charts of summary stats of the data
- Data visualization/hypothesis testing

What Do I Turn In?

- Summary of question
- Summary of data and why it is relevant for the question
- Tables and charts of summary stats of the data
- Data visualization/hypothesis testing
- Discussion of results

What Do I Turn In?

- Summary of question
- Summary of data and why it is relevant for the question
- Tables and charts of summary stats of the data
- Data visualization/hypothesis testing
- Discussion of results
- Criticism of your findings. What are the biggest flaws in your analysis or in the underlying data?

What Do I Turn In?

- Summary of question
- Summary of data and why it is relevant for the question
- Tables and charts of summary stats of the data
- Data visualization/hypothesis testing
- Discussion of results
- Criticism of your findings. What are the biggest flaws in your analysis or in the underlying data?
- Suggestions for further analysis or extension to the project

Examples

- R Tutorials provide some examples of the kinds of exploration I am looking for

Examples

- R Tutorials provide some examples of the kinds of exploration I am looking for
- CEA blog posts on jobs and GDP

Examples

- R Tutorials provide some examples of the kinds of exploration I am looking for
- CEA blog posts on jobs and GDP
- FiveThirtyEight's report on gun deaths

Do I Need to Discover Something New?

- Not looking for a Nobel-prize-winning discovery

Do I Need to Discover Something New?

- Not looking for a Nobel-prize-winning discovery
- You should learn something new

Do I Need to Discover Something New?

- Not looking for a Nobel-prize-winning discovery
- You should learn something new
- Hopefully I’ll learn something new as well

Do I Need to Discover Something New?

- Not looking for a Nobel-prize-winning discovery
- You should learn something new
- Hopefully I'll learn something new as well
- Be honest

Odd Questions

“Odd Question” # 1

Is the following likely true or false and why?

> *The counties with the* ***lowest*** *incidence of kidney cancer are mostly rural, sparsely populated, and located in traditionally Republican states in the Midwest, the South, and the East.*

“Odd Question” # 1

Is the following likely true or false and why?

> *The counties with the* ***lowest*** *incidence of kidney cancer are mostly rural, sparsely populated, and located in traditionally Republican states in the Midwest, the South, and the East.*

- Less water or air pollution
- Lower stress
- Healthier food

Is the following likely true or false and why?

> *The counties with the* ***highest*** *incidence of kidney cancer are mostly rural, sparsely populated, and located in traditionally Republican states in the Midwest, the South, and the East.*

“Odd Question” # 1

Is the following likely true or false and why?

> *The counties with the* ***highest*** *incidence of kidney cancer are mostly rural, sparsely populated, and located in traditionally Republican states in the Midwest, the South, and the East.*

- Higher stress because of poverty
- Drink alcohol
- Higher tobacco use

“Odd Question” # 1

Which statement is right?

1. Lower rate in rural counties

“Odd Question” # 1

Which statement is right?

1. Lower rate in rural counties
2. Higher rate in rural counties

Which statement is right?

1. Lower rate in rural counties
2. Higher rate in rural counties
3. Both

“Odd Question” # 1

Which statement is right?

1. Lower rate in rural counties
2. Higher rate in rural counties
3. Both
4. Neither

“Odd Question” # 1

Which statement is right?

1. Lower rate in rural counties
2. Higher rate in rural counties
3. Both
4. Neither

How?

From *Thinking Fast and Slow* by Daniel Kahneman

> *Something is wrong, of course. The rural lifestyle cannot explain both very high and very low incidence of kidney cancer. The key factor is not that the counties were rural or predominately Republican. It is that rural counties have small populations.*

“Odd Question” # 2

Pia is thirty-one years old, single, outspoken, and smart. She was a philosophy major. When a student, she was an ardent supporter of Native American rights, and she picketed a department store that had no facilities for nursing mothers.

Rank the following statements in order from most probable to least probable.

(a) Pia is a bank teller.
(b) Pia is a bank teller and an active feminist.

The Conjunction Fallacy

When it is assumed that specific conditions are more probable than a single general one.

Think Venn diagrams (we'll see this formally later in the lecture)

“Odd Question” # 3

In Lotto 6/49, a standard government-run lottery, you choose 6 out of 49 numbers (1 through 49). You win if these 6 are drawn. (The prize money is divided between all those who choose the lucky numbers. If no one wins, then most of the prize money is put back into next weeks lottery.)

Suppose your aunt offers you, *free*, a choice between two ticket in the lottery, with numbers as shown:

I. You win if 1, 2, 3, 4, 5, and 6 are drawn.
II. You win if 39, 36, 32, 21, 14, and 3 are drawn.

Do you prefer I, II, or are you indifferent between the two?

(a) Prefer I
(b) Prefer II
(c) Indifferent

To throw a total of 7 with a pair of dice, you have to get a 1 and a 6, or a 2 and a 5, or a 3 and a 4. To throw a total of 6 with a pair of dice, you have to get a 1 and a 5, or a 2 and a 4, or a 3 and another 3. With two fair dice, you would expect:

(a) To throw 7 more frequently than 6.
(b) To throw six more frequently than 7.
(c) To throw 6 and 7 equally often.

“Odd Question” # 5

“Imitate” a coin. That is, write down a sequence of 100 H (for heads) and T (for tails) without tossing a coin–but a sequence that you think will fool everyone into thinking it is the reporting of tossing a fair coin.

Which of these is a real sequence of coin flips?

Exhibit A

H H H H T H H T T T H H T T T H T T T H T T H H T
T T T H T T T H T H T T H H H H T T T T T H H H H
H T T H T T H H T T H H H H H T H H T H T H T H T
T H H T H H T T T H T T T T T T T T T H H T T T T

Exhibit B

H H T H T T T H H T H H H T H T T T H T H H T T T
T H H T T T H H H T H T T T H T T H H T H H T H T
T T H H H T H T T H T H H T T H H H T H T T H H H
T T H H H H T H T T H H T T T H H T H H H T T H T

How could we tell which are the real coin flips?

Hardly anyone making up a sequence of 10 tosses puts in a run of 7 heads in a row. It is true that the chance of getting 7 heads in a row with a fair coin is only 1/64. But in tossing a coin 100 times, you have at least 93 chances to start tossing 7 heads in a row, because each of the first 93 tosses could begin a run of 7. It is more probable than not, in 100 tosses, that you will get 7 heads in a row. It is certainly more probable than not, that you will get at least 6 heads in a row. Yet almost no one writes down a pretend sequence, in which there are even 6 heads in a row.

Remember!

Probability: Population $\rightarrow$ Sample

- Using information about the population to predict properties of a sample
- Deductive: “safe” argument
  - All ducks waddle, swim, and quack. Donald is a duck. Donald must waddle, swim, and quack.

Remember!

Probability: Population $\rightarrow$ Sample

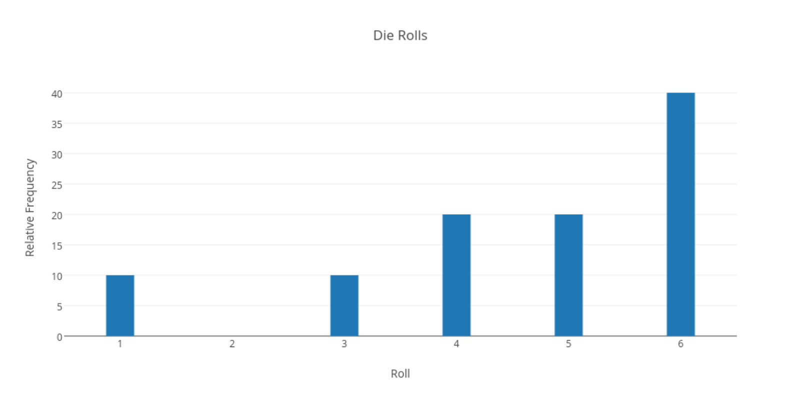
- Using information about the population to predict properties of a sample
- Deductive: “safe” argument
  - All ducks waddle, swim, and quack. Donald is a duck. Donald must waddle, swim, and quack.
- It turns out that we’re really bad at this

Our Definition of Probability for this Course

Probability: The long-run relative frequency

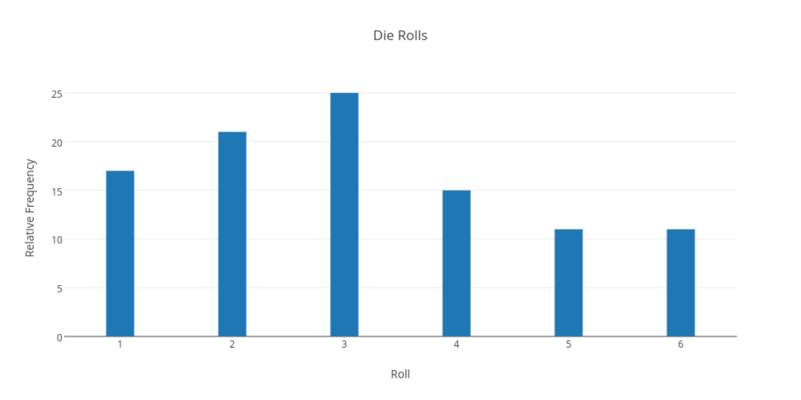
That is, relative frequencies settle down to probabilities if we carry out an experiment over, and over, and over...

10 Die Rolls

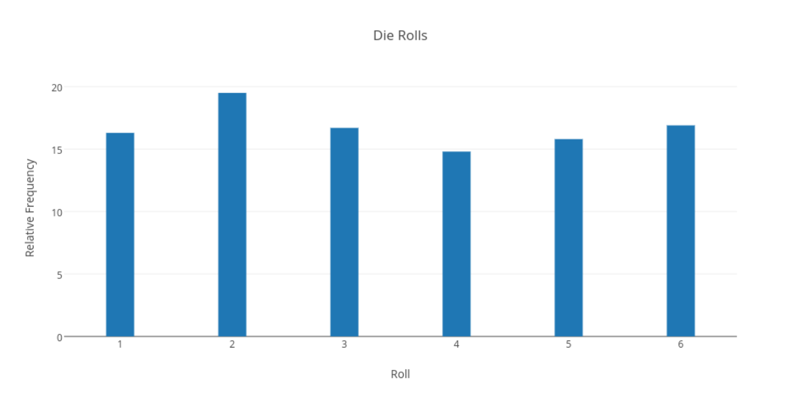


Die Rolls
Relative Frequency
40
35
30
25
20
15
10
5
0
1
2
3
4
5
6
Roll

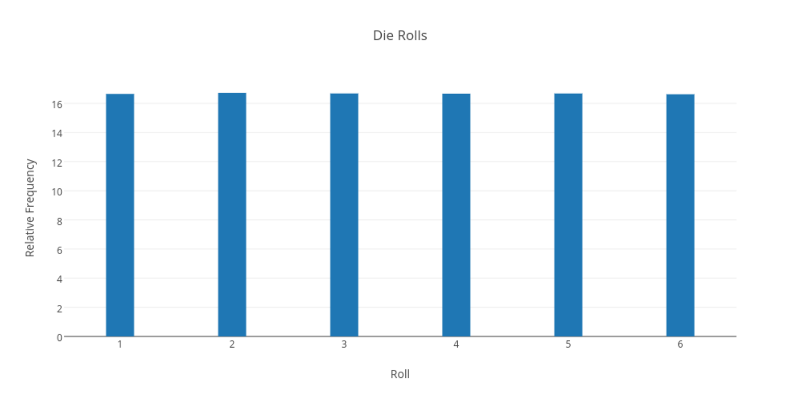
100 Die Rolls



1,000 Die Rolls



1,000,000 Die Rolls



What do you think of this argument?

- The probability of flipping heads is 1/2: if we flip a coin many times, about half of the time it will come up heads.
- The last ten throws in a row the coin has come up heads.
- The coin is bound to come up tails next time – it would be very rare to get 11 heads in a row.

Relative frequencies settle down to probabilities, but this does not mean that the trials are dependent.

Dependent: "Memory" of previous trials

Independent: No "memory" of previous trials

Another Argument

Lucie visits Albert. As she enters, he rolls four dice and shouts "Hooray!" for he has just rolled four sixes. Lucie: "I bet you've been rolling the dice for a long time to get that result!" Now, Lucie may have many reasons for saying this – perhaps Albert is a lunatic dice-roller. But simply on the evidence that he has just rolled four sixes, is her conclusion reasonable?

The *Inverse* Gambler's Fallacy

This is true:
Albert is more likely to get four sixes if he rolls many times than if he rolls only once.

However:
Regardless of how long Albert has been rolling, the probability that he gets four sixes on the particular roll that Lucie observes is unchanged.

The outcome of that roll doesn't tell us anything about whether he has rolled the dice before, just like six heads in a row doesn't mean we're "due" for a tails.

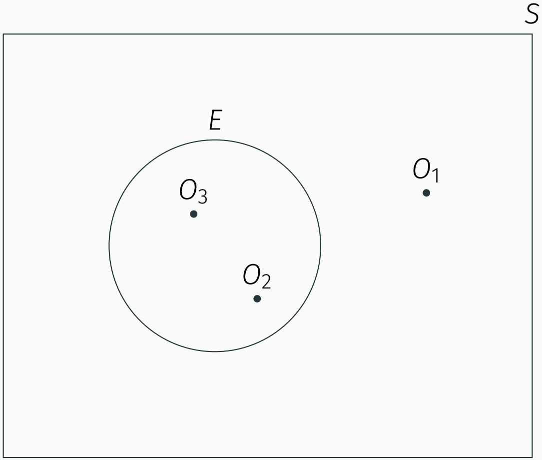
Definitions

- **Random Experiment:** An experiment whose outcomes are random.
- **Basic Outcomes:** Possible outcomes (mutually exclusive) of random experiment.
- **Sample Space (S):** Set of all basic outcomes of a random experiment.
- **Event (E):** A subset of the *sample space*. In set notation we write $E \subseteq S$.

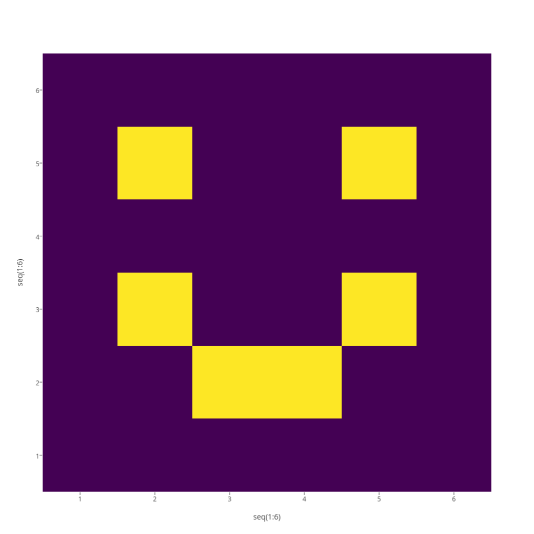
Examples

- **Experiment:** Tossing a pair of dice.
- **Basic Outcome:** An ordered pair (a, b) where $a, b \in \{1, 2, 3, 4, 5, 6\}$, e.g. $(2, 5)$
- **Sample Space:** $S = All$ ordered pairs (a, b) where $a, b \in \{1, 2, 3, 4, 5, 6\}$
- **Event:** $E = \{(2, 5), (5, 5), (2, 3), (3, 2), (4, 2), (5, 3)\}$

Visual Representation



Our Example



Probability is Defined on *Sets*, and Events are Sets

Set Theory

Complement of an Event: $A^c = \text{not } A$

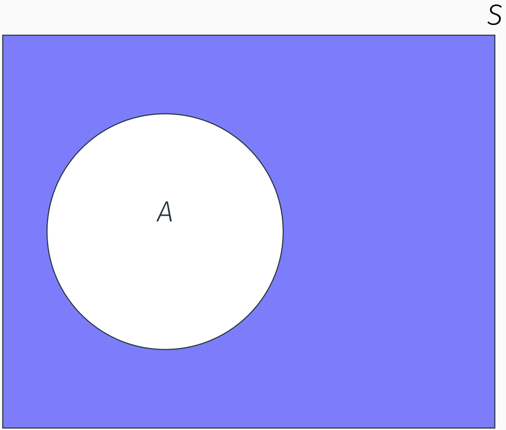


Figure 1: The complement A^c of an event $A \subseteq S$ is the collection of all basic outcomes from S not contained in A.

Intersection of Events: $A \cap B = A$ and B

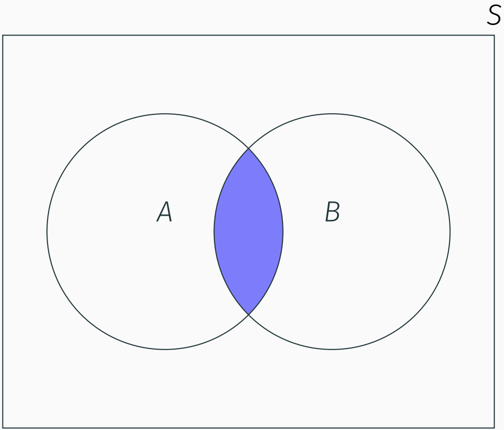


Figure 2: The intersection $A \cap B$ of two events $A, B \subseteq S$ is the collection of all basic outcomes from S contained in both A and B

Union of Events: $A \cup B = A$ or B

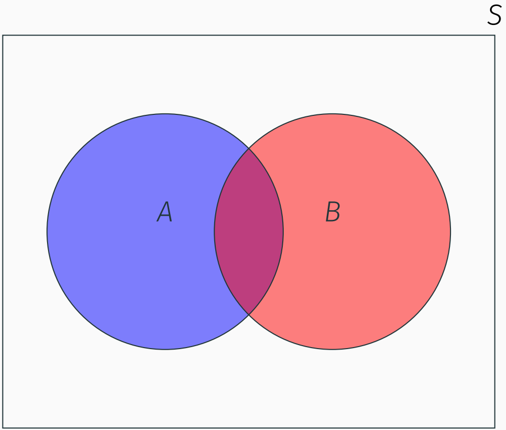


Figure 3: The union $A \cup B$ of two events $A, B \subseteq S$ is the collection of all basic outcomes from S contained in A, B or both.

Quick Set Theory Result

What is the formula for $A \cup B$?

Quick Set Theory Result

What is the formula for $A \cup B$?

$$A \cup B = A + B - (A \cap B)$$

Mutually Exclusive and Collectively Exhaustive

Mutually Exclusive Events

A collection of events $E_1, E_2, E_3, \ldots$ is *mutually exclusive* if $E_i \cap E_j$ of *any two different events* is empty (formally $E_i \cap E_j = \emptyset$ for any $i \neq j$).

Collectively Exhaustive Events

A collection of events $E_1, E_2, E_3, \ldots$ is *collectively exhaustive* if, taken together, they contain *all of the basic outcomes in S* (formally $E_1 \cup E_2 \cup \ldots = S$).

Implications

Mutually Exclusive Events

If one of the events occurs, then none of the others did.

Collectively Exhaustive Events

One of these events *must* occur.

Can you come up with examples of

- Mutually exclusive events that are not collectively exhaustive?
- Collectively exhaustive events that are not mutually exclusive?

Mutually Exclusive but *not Collectively Exhaustive*

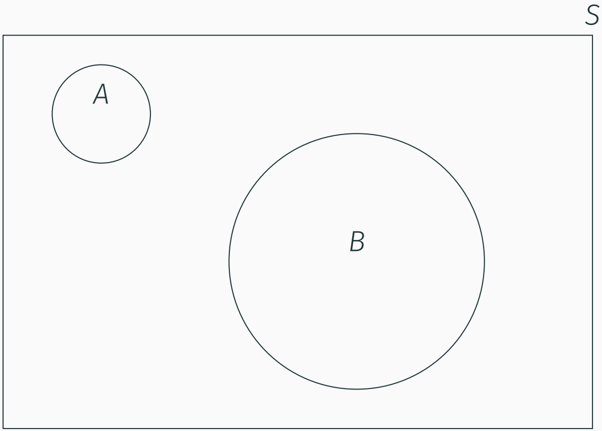


Figure 4: Although A and B don't overlap, they also don't cover S.

Collectively Exhaustive but *not Mutually Exclusive*

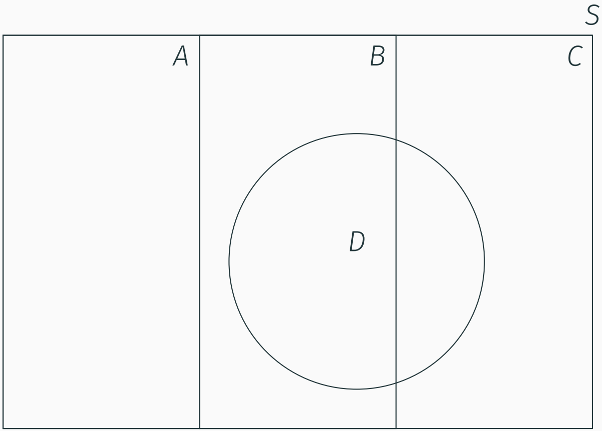


Figure 5: Together A, B, C and D cover S, but D overlaps with B and C.

Collectively Exhaustive *and* Mutually Exclusive

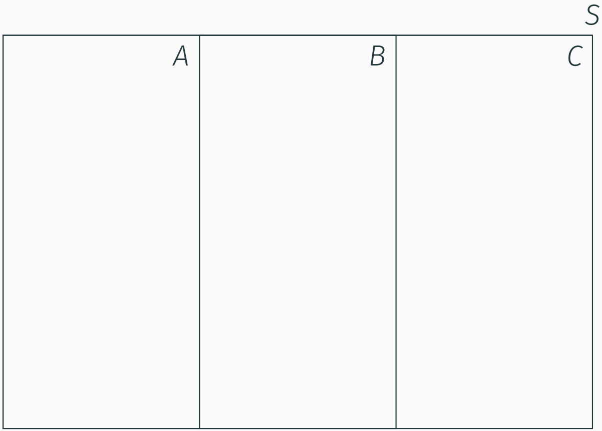


Figure 6: A, B, and C cover S and don't overlap.

Axioms of Probability

We assign every event A in the sample space S a real number $P(A)$ called the probability of A such that:

Axiom 1 $0 \leq P(A) \leq 1$

Axiom 2 $P(S) = 1$

Axiom 3 If $A_1, A_2, A_3, \ldots$ are mutually exclusive events, then $P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \ldots$

“Classical” Probability

When all of the basic outcomes are equally likely, calculating the probability of an event is simply a matter of counting – count up all the basic outcomes that make up the event, and divide by the total number of basic outcomes.

Recall from High School Math:

Multiplication Rule for Counting

n_1 ways to make first decision, n_2 ways to make second, $\ldots$, n_k ways to make kth $\Rightarrow n_1 \times n_2 \times \cdots \times n_k$ total ways to decide.

Corollary – Number of Possible Orderings

$k \times (k-1) \times (k-2) \times \cdots \times 2 \times 1 = k!$

Permutations – Order n people in k slots

$P_k^n = \frac{n!}{(n-k)!}$ (Order Matters)

Combinations – Choose committee of k from group of n

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$, where $0! = 1$ (Order Doesn't Matter)

Poker – Deal 5 Cards, Order Doesn't Matter

Basic Outcomes

$\binom{52}{5}$ possible hands

How Many Hands have Four Aces?

Basic Outcomes

$\binom{52}{5}$ possible hands

How Many Hands have Four Aces?

48 (# of ways to choose the single card that is not an ace)

Probability of Getting Four Aces

$48/\binom{52}{5} \approx 0.00002$

Poker – Deal 5 Cards, Order Doesn't Matter

What is the probability of getting 4 of a kind?

- 13 ways to choose *which* card we have four of
- 48 ways to choose the last card in the hand
- $13 \times 48 = 624$

$$624/\binom{52}{5} \approx 0.00024$$

A Fairly Ridiculous Example

Roger Federer and Novak Djokovic have agreed to play in a tennis tournament against six Penn professors. Each player in the tournament is randomly allocated to one of the eight rungs in the ladder (next slide). Federer always beats Djokovic and, naturally, either of the two pros always beats any of the professors. What is the probability that Djokovic gets second place in the tournament?

Rung 1

Quarterfinalist 1

Rung 2

Semifinalist 1

Rung 3

Quarterfinalist 2

Rung 4

Winner

Rung 5

Quarterfinalist 3

Rung 6

Semifinalist 2

Rung 7

Quarterfinalist 4

Rung 8

Solution: Order Matters!

Denominator

8! basic outcomes – ways to arrange players on tournament ladder.

Numerator

Sequence of three decisions:

1. Which rung to put Federer on? (8 possibilities)
2. Which rung to put Djokovic on?
   - For any given rung that Federer is on, only 4 rungs prevent Djokovic from meeting him until the final.
3. How to arrange the professors? (6! ways)

$$\frac{8 \times 4 \times 6!}{8!} = \frac{8 \times 4}{7 \times 8} = 4/7 \approx 0.57$$

Even if the basic outcomes are equally likely, the events of interest may not be...

“Odd Question” # 4

To throw a total of 7 with a pair of dice, you have to get a 1 and a 6, or a 2 and a 5, or a 3 and a 4. To throw a total of 6 with a pair of dice, you have to get a 1 and a 5, or a 2 and a 4, or a 3 and another 3. With two fair dice, you would expect:

(a) To throw 7 more frequently than 6.
(b) To throw six more frequently than 7.
(c) To throw 6 and 7 equally often.

Basic Outcomes Equally Likely, Events of Interest Aren’t

		Second Die					
		1	2	3	4	5	6
	1	2	3	4	5	6	7
	2	3	4	5	6	7	8
First	3	4	5	6	7	8	9
Die	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

Table 1: There are 36 equally likely basic outcomes, of which 5 correspond to a sum of six and 6 correspond to a sum of seven.

$P(7) = 6/36 = 1/6$
$P(6) = 5/36$

More Odd Questions

“Odd Question” # 6

You have been called to jury duty in a town where there are two taxi companies, Green Cab Ltd. and Blue Taxi Inc. Blue taxi uses cars painted blue; Green Cabs uses green cars. Green cabs dominates the market, with 85% of the taxis on the road. On a misty winter night a taxi sideswiped another car and drove off. A witness says it was a blue cab. The witness is tested under conditions like those on the night of the accident, and 80% of the time she correctly reports the color of the cab that is seen. That is, regardless of whether she is shown a blue or a green cab in misty evening light, she gets the color right 80% of the time.

You conclude, on the basis of this information:

(a) The probability that the sideswiper was blue is 0.8.

(b) It is more likely that the sideswiper was blue, but the probability is less than 0.8.

(c) It is just as probable that the sideswiper was green as that it was blue.

(d) It is more likely than not that the sideswiper was green.

“Odd Question” # 7

You are a physician. You think it is quite likely that one of your patients has strep throat, but you aren’t sure. You take some swabs from the throat and send them to a lab for testing. The test is (like nearly all lab tests) not perfect.

- If the patient has strep throat, then 70% if the time the lab says yes. But 30% of the time it says NO.
- If the patient does not have strep throat, then 90% of the time the lab says NO. But 10% of the time it says YES.

You send five successive swabs to the lab, from the same patient. and get back these results in order: YES, NO, YES, NO, YES. You conclude:

(a) These results are worthless.
(b) It is likely that the patient does not have strep throat.
(c) It is slightly more likely that the patient has strep throat.
(d) It is much more likely that the patient has strep throat.

Recall: Axioms of Probability

Let S be the sample space. With each event $A \subseteq S$ we associate a real number $P(A)$ called the probability of A, satisfying the following conditions:

Axiom 1 $0 \leq P(A) \leq 1$

Axiom 2 $P(S) = 1$

Axiom 3 If $A_1, A_2, A_3, \ldots$ are mutually exclusive events, then $P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \ldots$

The Complement Rule: $P(A^c) = 1 - P(A)$

Since A, A^c are mutually exclusive and collectively exhaustive:

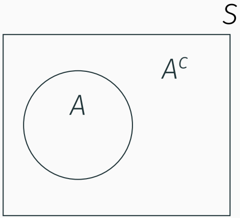


Figure 7: $A \cap A^c = \emptyset$, $A \cup A^c = S$

Since A, A^c are mutually exclusive and collectively exhaustive:

$$P(A \cup A^c) = P(A) + P(A^c) =$$

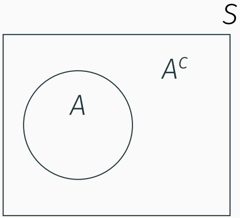


Figure 7: $A \cap A^c = \emptyset$, $A \cup A^c = S$

The Complement Rule: $P(A^c) = 1 - P(A)$

Since A, A^c are mutually exclusive and collectively exhaustive:

$$P(A \cup A^c) = P(A) + P(A^c) = P(S) = 1$$

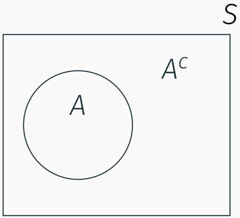


Figure 7: $A \cap A^c = \emptyset$, $A \cup A^c = S$

The Complement Rule: $P(A^c) = 1 - P(A)$

Since A, A^c are mutually exclusive and collectively exhaustive:

$$P(A \cup A^c) = P(A) + P(A^c) = P(S) = 1$$

Rearranging:

$$P(A^c) = 1 - P(A)$$

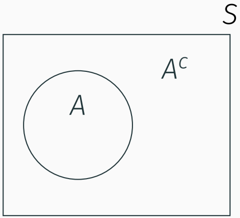


Figure 7: $A \cap A^c = \emptyset$, $A \cup A^c = S$

Another Important Rule – Equivalent Events

If A and B are Logically Equivalent, then $P(A) = P(B)$.

In other words, if A and B contain exactly the same basic outcomes, then $P(A) = P(B)$.

Although this seems obvious it's important to keep in mind, especially later in the course…

The Logical Consequence Rule

If B Logically Entails A, then $P(B) \leq P(A)$

In other words, $B \subseteq A \Rightarrow P(B) \leq P(A)$

Why is this so?

If $B \subseteq A$, then all the basic outcomes in B are also in A.

Proof of Logical Consequence Rule (Optional)

Since $B \subseteq A$, we have $B = A \cap B$ and $A = B \cup (A \cap B^c)$. Combining these,

$$A = (A \cap B) \cup (A \cap B^c)$$

Now since $(A \cap B) \cap (A \cap B^c) = \emptyset$,

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^c) \\ &= P(B) + P(A \cap B^c) \\ &\geq P(B) \end{aligned}$$

because $0 \leq P(A \cap B^c) \leq 1$.

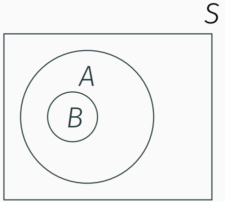


Figure 8: $B = A \cap B$, and $A = B \cup (A \cap B^c)$

“Odd Question” # 2

Pia is thirty-one years old, single, outspoken, and smart. She was a philosophy major. When a student, she was an ardent supporter of Native American rights, and she picketed a department store that had no facilities for nursing mothers. Rank the following statements in order from most probable to least probable.

(A) Pia is an active feminist.
(B) Pia is a bank teller.
(C) Pia works in a small bookstore.
(D) Pia is a bank teller and an active feminist.
(E) Pia is a bank teller and an active feminist who takes yoga classes.
(F) Pia works in a small bookstore and is an active feminist who takes yoga classes.

Using the Logical Consequence Rule...

(A) Pia is an active feminist.

(B) Pia is a bank teller.

(C) Pia works in a small bookstore.

(D) Pia is a bank teller and an active feminist.

(E) Pia is a bank teller and an active feminist who takes yoga classes.

(F) Pia works in a small bookstore and is an active feminist who takes yoga classes.

Any Correct Ranking Must Satisfy:

$$P(\mathrm{A}) \geq P(\mathrm{D}) \geq P(\mathrm{E})$$

$$P(\mathrm{B}) \geq P(\mathrm{D}) \geq P(\mathrm{E})$$

$$P(\mathrm{A}) \geq P(\mathrm{F})$$

$$P(\mathrm{C}) \geq P(\mathrm{F})$$

Throw a Fair Die Once

$E =$ roll an even number

What are the basic outcomes?

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

Throw a Fair Die Once

$E =$ roll an even number

What are the basic outcomes?

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

What is $P(E)$?

Throw a Fair Die Once

$E =$ **roll an even number**

What are the basic outcomes?

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

What is $P(E)$?

$E = \{2, 4, 6\}$ and the basic outcomes are equally likely (and mutually exclusive), so

$$P(E) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2$$

Throw a Fair Die Once

$E =$ roll an even number $\qquad M =$ roll a 1 or a prime number

What is $P(E \cup M)$?

Throw a Fair Die Once

$E =$ **roll an even number** $\quad M =$ **roll a 1 or a prime number**

What is $P(E \cup M)$?

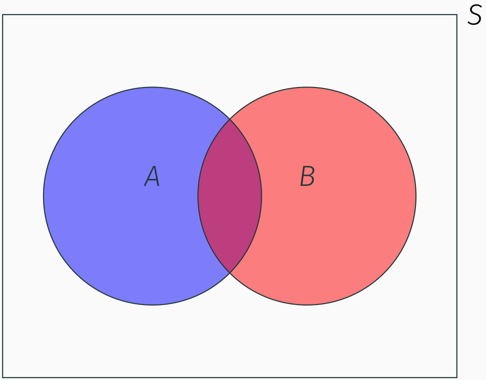
Key point: E and M are not mutually exclusive!

$$\begin{aligned} P(E \cup M) &= P(\{1,2,3,4,5,6\}) = 1 \\ P(E) &= P(\{2,4,6\}) = 1/2 \\ P(M) &= P(\{1,2,3,5\}) = 4/6 = 2/3 \end{aligned}$$

$$P(E) + P(M) = 1/2 + 2/3 = 7/6 \neq P(E \cup M) = 1$$

The Addition Rule – Don't Double-Count!

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Construct a formal proof as an optional homework problem.

Who’s on the other side?

Three Cards, Each with a Face on the Front and Back

1. Gaga/Gaga
2. Obama/Gaga
3. Obama/Obama

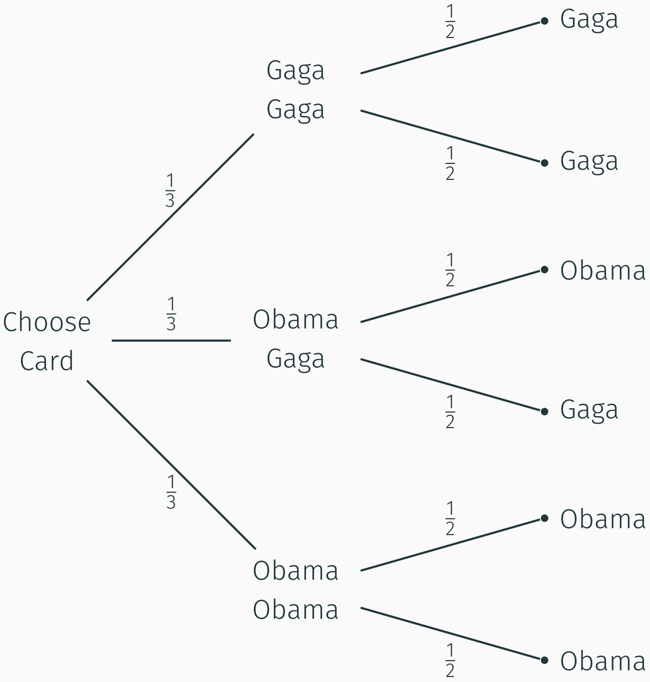
I draw a card at random and look at one side: it's Obama. What is the probability that the other side is also Obama?

Let’s Try The Method of Monte Carlo...

Procedure

1. Close your eyes and thoroughly shuffle your cards.
2. Keeping eyes closed, draw a card and place it on your desk.
3. Stand if Obama is face-up on your chosen card.
4. We’ll count those standing and call the total N
5. Of those standing, sit down if Obama is *not* on the back of your chosen card.
6. We’ll count those *still* standing and call the total m.

Monte Carlo Approximation of Desired Probability $= \frac{m}{N}$


Choose
Card
1/3
1/3
1/3
Gaga
Gaga
Obama
Gaga
Obama
Obama
1/2
1/2
1/2
1/2
1/2
1/2
Gaga
Gaga
Obama
Gaga
Obama
Obama

Conditional Probability – Reduced Sample Space

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{provided } P(B) > 0$$

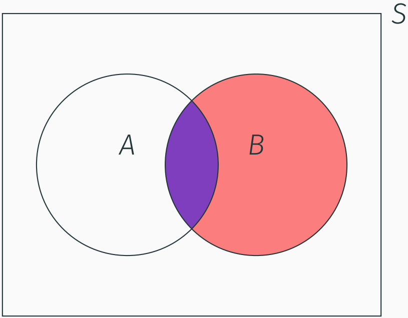


Figure 9: B becomes the "new sample space" so we need to re-scale by $P(B)$ to keep probabilities between zero and one.

Who’s on the other side?

Let F be the event that Obama is on the front of the card of the card we draw and B be the event that he is on the back.

$$P(B|F) = \frac{P(B \cap F)}{P(F)} =$$

Who's on the other side?

Let F be the event that Obama is on the front of the card of the card we draw and B be the event that he is on the back.

$$P(B|F) = \frac{P(B \cap F)}{P(F)} = \frac{1/3}{1/2} =$$

Who's on the other side?

Let F be the event that Obama is on the front of the card of the card we draw and B be the event that he is on the back.

$$P(B|F) = \frac{P(B \cap F)}{P(F)} = \frac{1/3}{1/2} = 2/3$$

Conditional Versions of Probability Axioms

1. $0 \leq P(A|B) \leq 1$
2. $P(B|B) = 1$
3. If $A_1, A_2, A_3, \ldots$ are mutually exclusive events, then $P(A_1 \cup A_2 \cup A_3 \cup \cdots |B) = P(A_1|B) + P(A_2|B) + P(A_3|B) \ldots$

Conditional Versions of Other Probability Rules

- $P(A|B) = 1 - P(A^c|B)$
- A_1 logically equivalent to $A_2 \iff P(A_1|B) = P(A_2|B)$
- $A_1 \subseteq A_2 \implies P(A_1|B) \leq P(A_2|B)$
- $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2|B)$

However: $P(A|B) \neq P(B|A)$ and $P(A|B^c) \neq 1 - P(A|B)$!

The Multiplication Rule

Rearrange the definition of conditional probability:

$P(A \cap B) = P(A|B)P(B)$

The Multiplication Rule

Rearrange the definition of conditional probability:

$P(A \cap B) = P(A|B)P(B)$

Statistical Independence

$P(A \cap B) = P(A)P(B)$

The Multiplication Rule

Rearrange the definition of conditional probability:

$P(A \cap B) = P(A|B)P(B)$

Statistical Independence

$P(A \cap B) = P(A)P(B)$

By the Multiplication Rule

Independence $\iff P(A|B) = P(A)$

The Multiplication Rule

Rearrange the definition of conditional probability:

$P(A \cap B) = P(A|B)P(B)$

Statistical Independence

$P(A \cap B) = P(A)P(B)$

By the Multiplication Rule

Independence $\iff P(A|B) = P(A)$

Interpreting Independence

Knowledge that B has occurred tells nothing about whether A will.

Will Having 5 Children Guarantee a Boy?

A couple plans to have five children. Assuming that each birth is independent and male and female children are equally likely, what is the probability that they have at least one boy?

Will Having 5 Children Guarantee a Boy?

A couple plans to have five children. Assuming that each birth is independent and male and female children are equally likely, what is the probability that they have at least one boy?

By Independence and the Complement Rule,

$$\begin{aligned} P(\text{no boys}) &= P(\text{5 girls}) \\ &= 1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2 \\ &= 1/32 \end{aligned}$$

Will Having 5 Children Guarantee a Boy?

A couple plans to have five children. Assuming that each birth is independent and male and female children are equally likely, what is the probability that they have at least one boy?

By Independence and the Complement Rule,

$$\begin{aligned} P(\text{no boys}) &= P(\text{5 girls}) \\ &= 1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2 \\ &= 1/32 \end{aligned}$$

$$\begin{aligned} P(\text{at least 1 boy}) &= 1 - P(\text{no boys}) \\ &= 1 - 1/32 = 31/32 = 0.97 \end{aligned}$$

The Law of Total Probability

If $E_1, E_2, \ldots, E_k$ are mutually exclusive, collectively exhaustive events and A is another event, then

$$P(A) = P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + \ldots + P(A|E_k)P(E_k)$$

Example of Law of Total Probability

Define the following events:

$$
\begin{aligned}
F &= \text{Obama on front of card} \\
A &= \text{Draw card with two Gagas} \\
B &= \text{Draw card with two Obamas} \\
C &= \text{Draw card with BOTH Obama and Gaga}
\end{aligned}
$$

$$P(F) = P(F|A)P(A) + P(F|B)P(B) + P(F|C)P(C)$$

Example of Law of Total Probability

Define the following events:

$$\begin{aligned}
F &= \text{Obama on front of card} \\
A &= \text{Draw card with two Gagas} \\
B &= \text{Draw card with two Obamas} \\
C &= \text{Draw card with BOTH Obama and Gaga}
\end{aligned}$$

$$\begin{aligned}
P(F) &= P(F|A)P(A) + P(F|B)P(B) + P(F|C)P(C) \\
&= 0 \times 1/3 + 1 \times 1/3 + 1/2 \times 1/3 \\
&= 1/2
\end{aligned}$$

Deriving the Law of Total Probability For $k = 2$

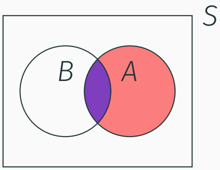


Figure 10:
$A = (A \cap B) \cup (A \cap B^c)$,
$(A \cap B) \cap (A \cap B^c) = \emptyset$

Deriving the Law of Total Probability For $k = 2$

Since $A \cap B$ and $A \cap B^c$ are mutually exclusive and their union equals A,

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

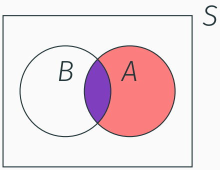


Figure 10: $A = (A \cap B) \cup (A \cap B^c)$, $(A \cap B) \cap (A \cap B^c) = \emptyset$

Deriving the Law of Total Probability For $k = 2$

Since $A \cap B$ and $A \cap B^c$ are mutually exclusive and their union equals A,

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

But by the multiplication rule:

$$\begin{aligned} P(A \cap B) &= P(A|B)P(B) \\ P(A \cap B^c) &= P(A|B^c)P(B^c) \end{aligned}$$

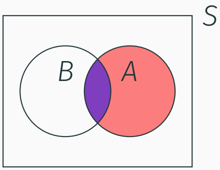


Figure 10: $A = (A \cap B) \cup (A \cap B^c)$, $(A \cap B) \cap (A \cap B^c) = \emptyset$

Deriving the Law of Total Probability For $k = 2$

Since $A \cap B$ and $A \cap B^c$ are mutually exclusive and their union equals A,

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

But by the multiplication rule:

$$\begin{aligned} P(A \cap B) &= P(A|B)P(B) \\ P(A \cap B^c) &= P(A|B^c)P(B^c) \end{aligned}$$

Combining,

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

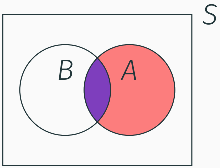


Figure 10: $A = (A \cap B) \cup (A \cap B^c)$, $(A \cap B) \cap (A \cap B^c) = \emptyset$

How do prediction markets work?

THIS CERTIFICATE ENTITLES THE BEARER TO \$1 IF TRUMP WINS THE 2016 US PRESIDENTIAL ELECTION.

Buyers – Purchase Right to Collect

Trump very likely to win $\Rightarrow$ buy for close to \$1.

Trump very unlikely to win $\Rightarrow$ buy for close to \$0.

Sellers – Sell Obligation to Pay

Trump very likely to win $\Rightarrow$ sell for close to \$1.

Trump very unlikely to win $\Rightarrow$ sell for close to \$0.

Probabilities from Beliefs

Market price of contract encodes market participants' beliefs in the form of probability:

$$\text{Price/Payout} \approx \text{Subjective Probability}$$

Probabilities from Beliefs

Market price of contract encodes market participants' beliefs in the form of probability:

$$\text{Price}/\text{Payout} \approx \text{Subjective Probability}$$

Statistical Arbitrage

If the probabilities implied by prediction market prices violate *any* of our probability rules there is a *pure arbitrage opportunity*: a way make to make a guaranteed, risk-free profit.

Probabilities from Beliefs

Market price of contract encodes market participants' beliefs in the form of probability:

$$\text{Price}/\text{Payout} \approx \text{Subjective Probability}$$

Statistical Arbitrage

If the probabilities implied by prediction market prices violate *any* of our probability rules there is a *pure arbitrage opportunity*: a way make to make a guaranteed, risk-free profit.

Last I checked a $1 contract on Trump winning was going for $0.29 on `predictit.org`

A Simple Example of Statistical Arbitrage

November 5th, 2012

- \$2.30 for contract paying \$10 if Romney wins on BetFair
- \$6.58 for contract paying \$10 if Obama wins on InTrade

Implied Probabilities

- BetFair: $P(Romney) \approx 0.23$
- InTrade: $P(Obama) \approx 0.66$

What's Wrong with This?

A Simple Example of Statistical Arbitrage

November 5th, 2012

- \$2.30 for contract paying \$10 if Romney wins on BetFair
- \$6.58 for contract paying \$10 if Obama wins on InTrade

Implied Probabilities

- BetFair: $P(Romney) \approx 0.23$
- InTrade: $P(Obama) \approx 0.66$

What's Wrong with This?

Violates complement rule! $P(Obama) = 1 - P(Romney)$ but the implied probabilities here don't sum up to one!

A Simple Example of Statistical Arbitrage

November 5th, 2012

- \$2.30 for contract paying \$10 if Romney wins on BetFair
- \$6.58 for contract paying \$10 if Obama wins on InTrade

Arbitrage Strategy

Buy Equal Numbers of Each

- Cost = \$2.30 + \$6.58 =\$8.88 per pair
- Payout if Romney Wins: \$10
- Payout if Obama Wins: \$10
- Guaranteed Profit: \$10 - \$8.88 = \$1.12 per pair