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
- Answer questions you are interested in

Motivation

- Apply the skills and tools you have learned
- 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

When is it Due?

- **October 12:** Submit project idea and question or spoken with me about the project. Submit names of people in your group.
- **November 16:** (Optional) Submit rough draft for comments.
- **December 7:** Hand in project.

Examples

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

Examples

- R Tutorials provide examples of the kinds of exploration I am looking for
- CEA blog posts on jobs and GDP and other publications (like the Economic Report of the President)

Examples

- R Tutorials provide examples of the kinds of exploration I am looking for
- CEA blog posts on jobs and GDP and other publications (like the Economic Report of the President)
- 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

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
- 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

“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.*

“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
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)

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

Probability

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

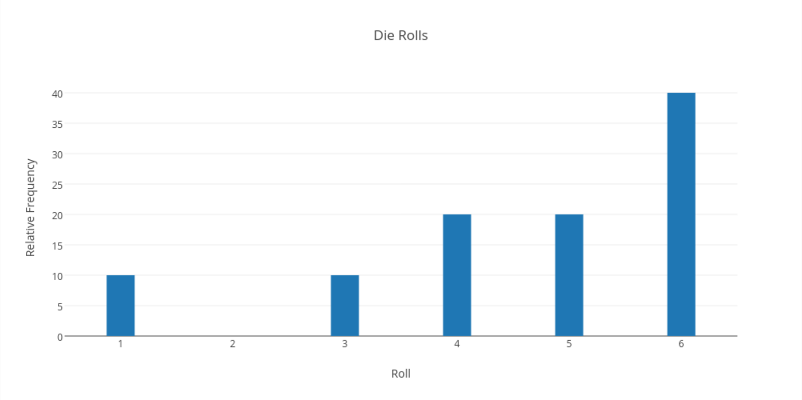
- Using information about the population to predict properties of a sample
- Deductive: "safe" argument
    - All ducks waddle, swim, and quack. Donald is a duck. Donald must waddle, swim, and quack.
- It turns out that we're really bad at this

Our Definition of Probability for this Course

Probability: The long-run relative frequency

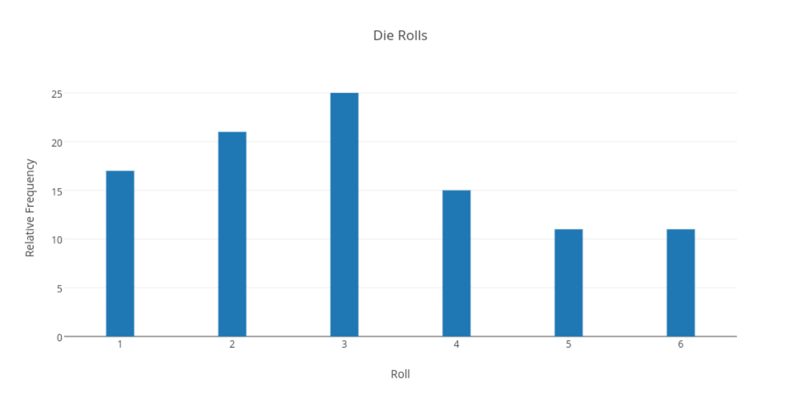
That is, relative frequencies settle down to probabilities if we carry out an experiment over, and over, and over...

10 Die Rolls

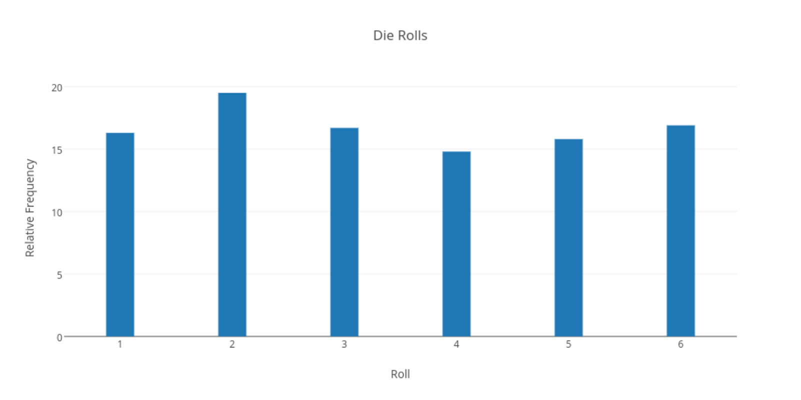


Die Rolls
Relative Frequency
40
35
30
25
20
15
10
5
0
1
2
3
4
5
6
Roll

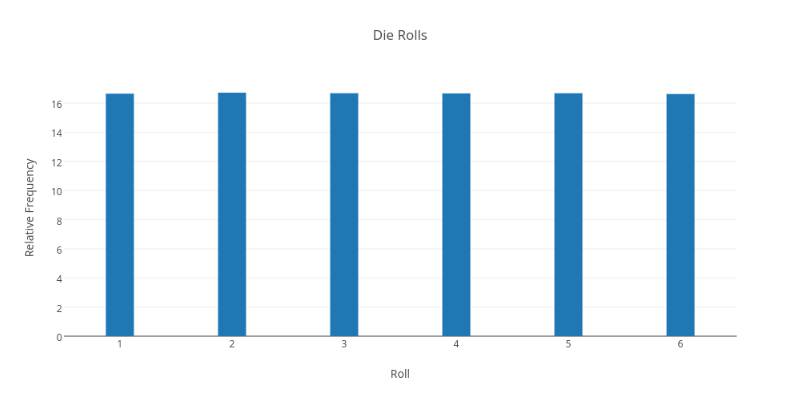
100 Die Rolls



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.

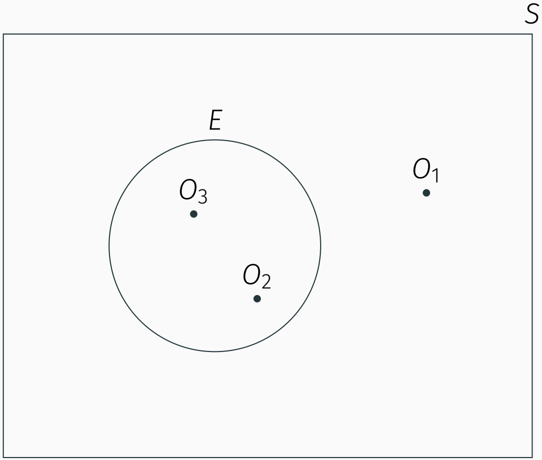
Definitions

- **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 $E \subseteq S$.

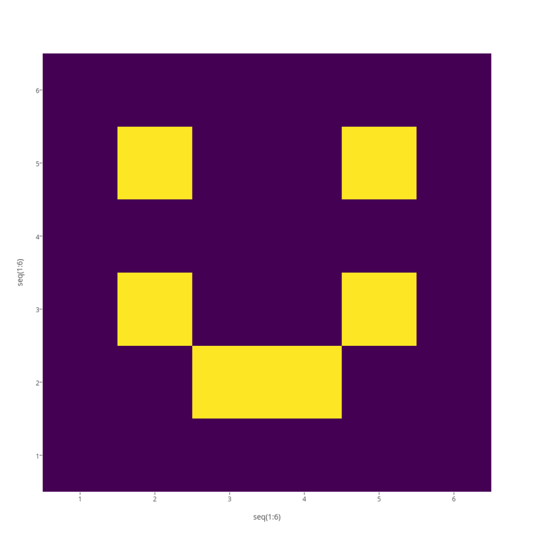
Examples

- **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

Dice Problem

Set Theory

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

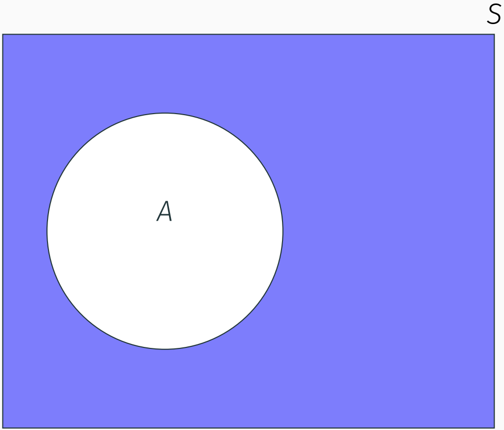


Figure 1: The complement A^c of an event $A \subseteq S$ is the collection of all basic outcomes from S not contained in A.

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

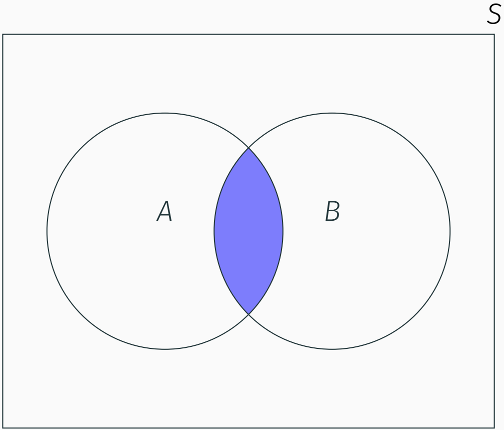


Figure 2: The intersection $A \cap B$ of two events $A, B \subseteq S$ is the collection of all basic outcomes from S contained in both A and B

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

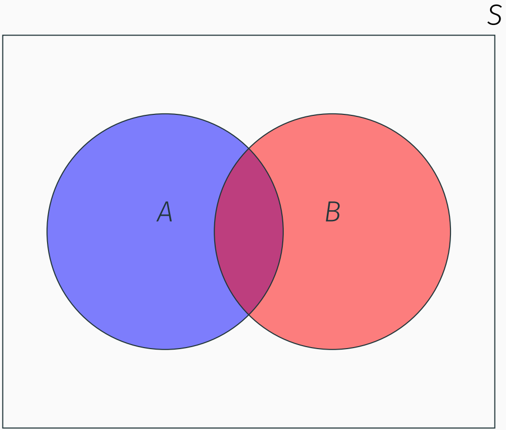


Figure 3: The union $A \cup B$ of two events $A, B \subseteq S$ is the collection of all basic outcomes from S contained in A, B or both.

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*

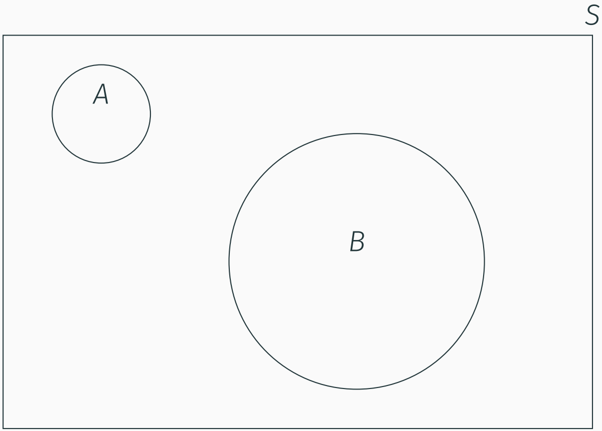


Figure 4: Although A and B don't overlap, they also don't cover S.

Collectively Exhaustive but *not Mutually Exclusive*

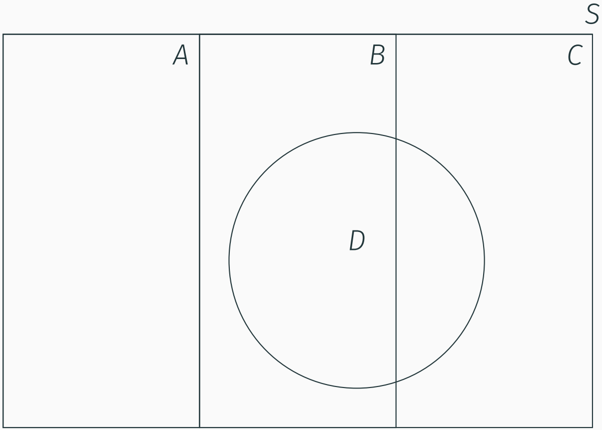


Figure 5: Together A, B, C and D cover S, but D overlaps with B and C.

Collectively Exhaustive *and* Mutually Exclusive

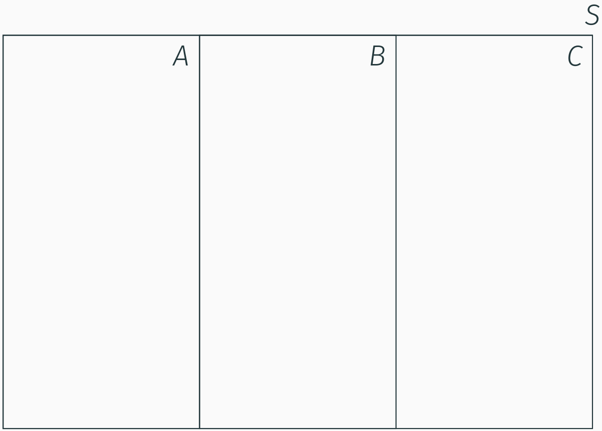


Figure 6: A, B, and C cover S and don't overlap.

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

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

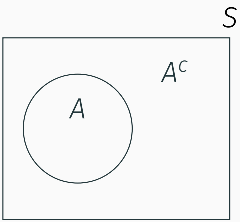


Figure 7: $A \cap A^c = \emptyset$, $A \cup A^c = S$

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

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

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

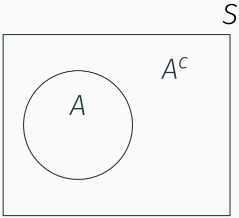


Figure 7: $A \cap A^c = \emptyset$, $A \cup A^c = S$

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

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

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

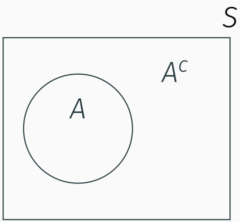


Figure 7: $A \cap A^c = \emptyset$, $A \cup A^c = S$

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

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

Rearranging:

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

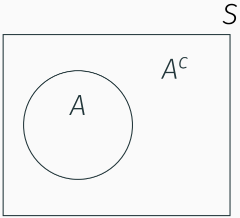


Figure 7: $A \cap A^c = \emptyset$, $A \cup A^c = S$

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 Birthday Problem

What is the least number of persons required if the probability exceeds 1/2 that two or more of them have the same birthday? (Year of birth need not match.)

[Hint: Use the Complement Rule.]

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 $P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \ldots$

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

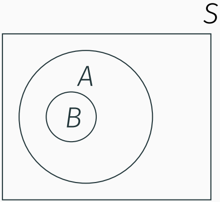


Figure 8: $B = A \cap B$, and $A = B \cup (A \cap B^c)$

Permutations and Combinations

“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)

Coin Flipping, Cards, Tennis, and Obama

Conditional Probability

Conditional Probability – Reduced Sample Space

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

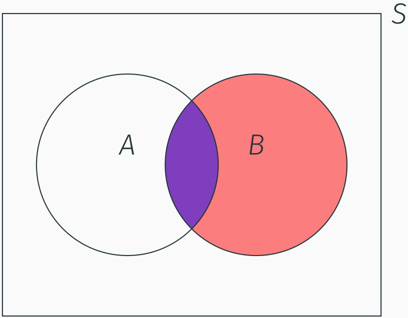


Figure 9: B becomes the "new sample space" so we need to re-scale by $P(B)$ to keep probabilities between zero and one.

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

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

Independence and The Multiplication Rule

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

Independence and The Multiplication Rule

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.

The Law of Total Probability

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 $k = 2$

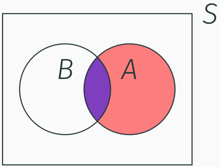


Figure 10:
$A = (A \cap B) \cup (A \cap B^c)$,
$(A \cap B) \cap (A \cap B^c) = \emptyset$

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,

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

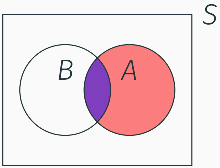


Figure 10: $A = (A \cap B) \cup (A \cap B^c)$, $(A \cap B) \cap (A \cap B^c) = \emptyset$

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:

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

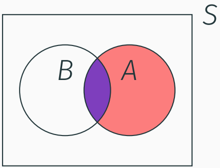


Figure 10: $A = (A \cap B) \cup (A \cap B^c)$, $(A \cap B) \cap (A \cap B^c) = \emptyset$

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,

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

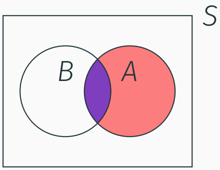


Figure 10: $A = (A \cap B) \cup (A \cap B^c)$, $(A \cap B) \cap (A \cap B^c) = \emptyset$

Prediction Markets

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 ⇒ buy for close to $1.
Trump very unlikely to win ⇒ buy for close to $0.

Sellers – Sell Obligation to Pay

Trump very likely to win ⇒ sell for close to $1.
Trump very unlikely to win ⇒ 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 $0.36 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

The Lie Detector

Four Volunteers Please!

The Lie Detector Problem

From accounting records, we know that 10% of employees in the store are stealing merchandise.

The managers want to fire the thieves, but their only tool in distinguishing is a lie detector test that is 80% accurate:

$$\begin{aligned} \text{Innocent} \quad &\Rightarrow \quad \text{Pass test with 80\% Probability} \\ \text{Thief} \quad &\Rightarrow \quad \text{Fail test with 80\% Probability} \end{aligned}$$

The Lie Detector Problem

From accounting records, we know that 10% of employees in the store are stealing merchandise.

The managers want to fire the thieves, but their only tool in distinguishing is a lie detector test that is 80% accurate:

$$\begin{aligned} \text{Innocent} &\quad\Rightarrow\quad \text{Pass test with 80\% Probability} \\ \text{Thief} &\quad\Rightarrow\quad \text{Fail test with 80\% Probability} \end{aligned}$$

What is the probability that someone is a thief *given* that she has failed the lie detector test?

Monte Carlo Simulation – Roll a 10-sided Die Twice

Managers will split up and visit employees. Employees roll the die twice but keep the results secret!

First Roll – Thief or not?

$0 \Rightarrow$ Thief, $1 - 9 \Rightarrow$ Innocent

Second Roll – Lie Detector Test

$0, 1 \Rightarrow$ Incorrect Test Result, $2 - 9$ Correct Test Result

	0 or 1	2–9
Thief	Pass	Fail
Innocent	Fail	Pass

What percentage of those who failed the test are guilty?

Who Failed Lie Detector Test:

Of Thieves Among Those Who Failed:

Base Rate Fallacy – Failure to Consider Prior Information

Base Rate – Prior Information

Before the test we know that 10% of Employees are stealing.

People tend to focus on the fact that the test is 80% accurate and ignore the fact that only 10% of the employees are theives.

Thief (Y/N), Lie Detector (P/F)

	0	1	2	3	4	5	6	7	8	9
0	YP	YP	YF	YF	YF	YF	YF	YF	YF	YF
1	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
2	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
3	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
4	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
5	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
6	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
7	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
8	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP
9	NF	NF	NP	NP	NP	NP	NP	NP	NP	NP

Table 1: Each outcome in the table is equally likely. The 26 given in red correspond to failing the test, but only 8 of these (YF) correspond to being a thief.

Base Rate of Thievery is 10%

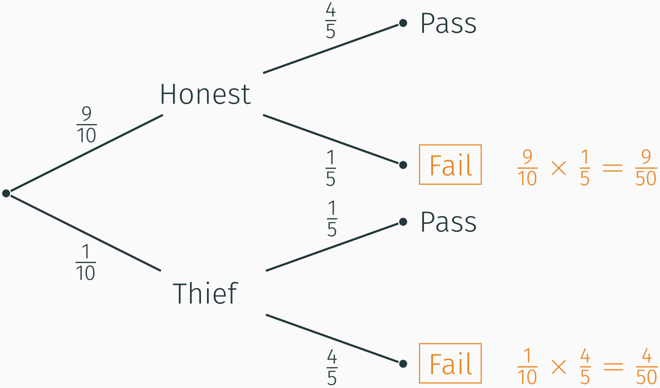


Figure 11: Although $\frac{9}{50} + \frac{4}{50} = \frac{13}{50}$ fail the test, only $\frac{4/50}{13/50} = \frac{4}{13} \approx 0.31$ are actually theives!

Bayes' Rule

Deriving Bayes' Rule

Intersection is symmetric: $A \cap B = B \cap A$ so $P(A \cap B) = P(B \cap A)$

Deriving Bayes' Rule

Intersection is symmetric: $A \cap B = B \cap A$ so $P(A \cap B) = P(B \cap A)$

By the definition of conditional probability,

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

Deriving Bayes' Rule

Intersection is symmetric: $A \cap B = B \cap A$ so $P(A \cap B) = P(B \cap A)$
By the definition of conditional probability,

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

And by the multiplication rule:

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

Deriving Bayes' Rule

Intersection is symmetric: $A \cap B = B \cap A$ so $P(A \cap B) = P(B \cap A)$

By the definition of conditional probability,

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

And by the multiplication rule:

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

Finally, combining these

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$\boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)}}$$

Reversing the Conditioning

Express $P(A|B)$ in terms of $P(B|A)$. *Relative magnitudes* of the two conditional probabilities determined by the ratio $P(A)/P(B)$.

Base Rate

$P(A)$ is called the "base rate" or the "prior probability."

Denominator

Typically, we calculate $P(B)$ using the law of total probability

In General $P(A|B) \neq P(B|A)$

Question

Most college students are Democrats. Does it follow that most Democrats are college students? (A = YES, B = NO)

In General $P(A|B) \neq P(B|A)$

Question

Most college students are Democrats. Does it follow that most Democrats are college students? (A = YES, B = NO)

Answer

There are many more Democracts than college students:

$$P(\text{Dem}) > P(\text{Student})$$

so $P(\text{Student}|\text{Dem})$ is small even though $P(\text{Dem}|\text{Student})$ is large.

Solving the Lie Detector Problem with Bayes' Rule

$\boxed{T = \text{Employee is a Thief}, F = \text{Employee Fails Lie Detector Test}}$

$$P(T|F) = \frac{P(F|T)P(T)}{P(F)}$$

Solving the Lie Detector Problem with Bayes' Rule

$\boxed{T = \text{Employee is a Thief, } F = \text{Employee Fails Lie Detector Test}}$

$$P(T|F) = \frac{P(F|T)P(T)}{P(F)}$$

$$P(F) \quad = \quad P(F|T)P(T) + P(F|T^c)P(T^c)$$

Solving the Lie Detector Problem with Bayes' Rule

$\boxed{T = \text{Employee is a Thief, } F = \text{Employee Fails Lie Detector Test}}$

$$P(T|F) = \frac{P(F|T)P(T)}{P(F)}$$

$$\begin{aligned} P(F) &= P(F|T)P(T) + P(F|T^c)P(T^c) \\ &= 0.8 \times 0.1 + 0.2 \times 0.9 \end{aligned}$$

Solving the Lie Detector Problem with Bayes' Rule

$\boxed{T = \text{Employee is a Thief}, F = \text{Employee Fails Lie Detector Test}}$

$$P(T|F) = \frac{P(F|T)P(T)}{P(F)}$$

$$\begin{aligned} P(F) &= P(F|T)P(T) + P(F|T^c)P(T^c) \\ &= 0.8 \times 0.1 + 0.2 \times 0.9 \\ &= 0.08 + 0.18 = 0.26 \end{aligned}$$

Solving the Lie Detector Problem with Bayes' Rule

$\boxed{T = \text{Employee is a Thief}, F = \text{Employee Fails Lie Detector Test}}$

$$P(T|F) = \frac{P(F|T)P(T)}{P(F)}$$

$$\begin{aligned} P(F) &= P(F|T)P(T) + P(F|T^c)P(T^c) \\ &= 0.8 \times 0.1 + 0.2 \times 0.9 \\ &= 0.08 + 0.18 = 0.26 \end{aligned}$$

$$P(T|F) = \frac{0.08}{0.26} =$$

Solving the Lie Detector Problem with Bayes' Rule

$\boxed{T = \text{Employee is a Thief, } F = \text{Employee Fails Lie Detector Test}}$

$$P(T|F) = \frac{P(F|T)P(T)}{P(F)}$$

$$\begin{aligned} P(F) &= P(F|T)P(T) + P(F|T^c)P(T^c) \\ &= 0.8 \times 0.1 + 0.2 \times 0.9 \\ &= 0.08 + 0.18 = 0.26 \end{aligned}$$

$$P(T|F) = \frac{0.08}{0.26} = \frac{8}{26} = \frac{4}{13} \approx 0.31$$

“Odd” Question # 5

There are two kinds of taxis: green cabs and blue cabs. Of all the cabs on the road, 85% are green cabs. 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.

Given that the witness said she saw a blue cab, what is the probability that a blue cab was the sideswiper?

Solving The Taxi Problem

G = Taxi is Green, $P(G) = 0.85$
B = Taxi is Blue, $P(B) = 0.15$
W_B = Witness says Taxi is Blue, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

Solving The Taxi Problem

G = Taxi is Green, $P(G) = 0.85$
B = Taxi is Blue, $P(B) = 0.15$
W_B = Witness says Taxi is Blue, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

Solving The Taxi Problem

$G =$ Taxi is Green, $P(G) = 0.85$
$B =$ Taxi is Blue, $P(B) = 0.15$
$W_B =$ Witness says Taxi is Blue, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

$$P(W_B) = P(W_B|B)P(B) + P(W_B|G)P(G)$$

Solving The Taxi Problem

$G =$ Taxi is Green, $P(G) = 0.85$
$B =$ Taxi is Blue, $P(B) = 0.15$
$W_B =$ Witness says Taxi is Blue, $P(W_B|B) = 0.8, P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

$$\begin{aligned} P(W_B) &= P(W_B|B)P(B) + P(W_B|G)P(G) \\ &= 0.8 \times 0.15 + 0.2 \times 0.85 \end{aligned}$$

Solving The Taxi Problem

$G =$ Taxi is Green, $P(G) = 0.85$
$B =$ Taxi is Blue, $P(B) = 0.15$
$W_B =$ Witness says Taxi is Blue, $P(W_B|B) = 0.8$, $P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

$$\begin{aligned} P(W_B) &= P(W_B|B)P(B) + P(W_B|G)P(G) \\ &= 0.8 \times 0.15 + 0.2 \times 0.85 \\ &= 0.12 + 0.17 = 0.29 \end{aligned}$$

Solving The Taxi Problem

$G =$ Taxi is Green, $P(G) = 0.85$
$B =$ Taxi is Blue, $P(B) = 0.15$
$W_B =$ Witness says Taxi is Blue, $P(W_B|B) = 0.8$, $P(W_B|G) = 0.2$

$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

$$\begin{aligned} P(W_B) &= P(W_B|B)P(B) + P(W_B|G)P(G) \\ &= 0.8 \times 0.15 + 0.2 \times 0.85 \\ &= 0.12 + 0.17 = 0.29 \end{aligned}$$

$$P(B|W_B) = 0.12/0.29 = 12/29 \approx 0.41$$

Solving The Taxi Problem

G = Taxi is Green, $P(G) = 0.85$
B = Taxi is Blue, $P(B) = 0.15$
W_B = Witness says Taxi is Blue, $P(W_B|B) = 0.8$, $P(W_B|G) = 0.2$

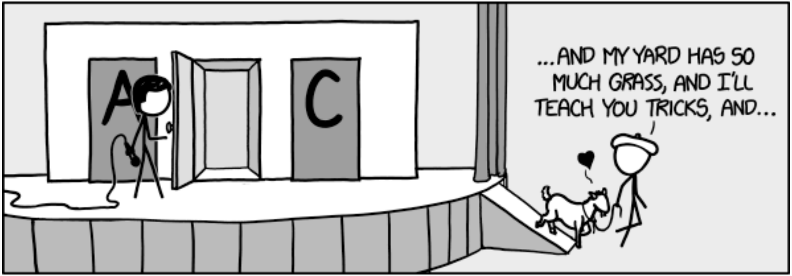
$$P(B|W_B) = P(W_B|B)P(B)/P(W_B)$$

$$\begin{aligned} P(W_B) &= P(W_B|B)P(B) + P(W_B|G)P(G) \\ &= 0.8 \times 0.15 + 0.2 \times 0.85 \\ &= 0.12 + 0.17 = 0.29 \end{aligned}$$

$$\begin{aligned} P(B|W_B) &= 0.12/0.29 = 12/29 \approx 0.41 \\ P(G|W_B) &= 1 - (12/19) \approx 0.59 \end{aligned}$$

The Monty Hall Problem!

What is the probability that you win if you switch?



Key Point – Monte doesn't choose a door randomly: he *always* shows you a goat.

Without loss of generality, suppose you chose door #1