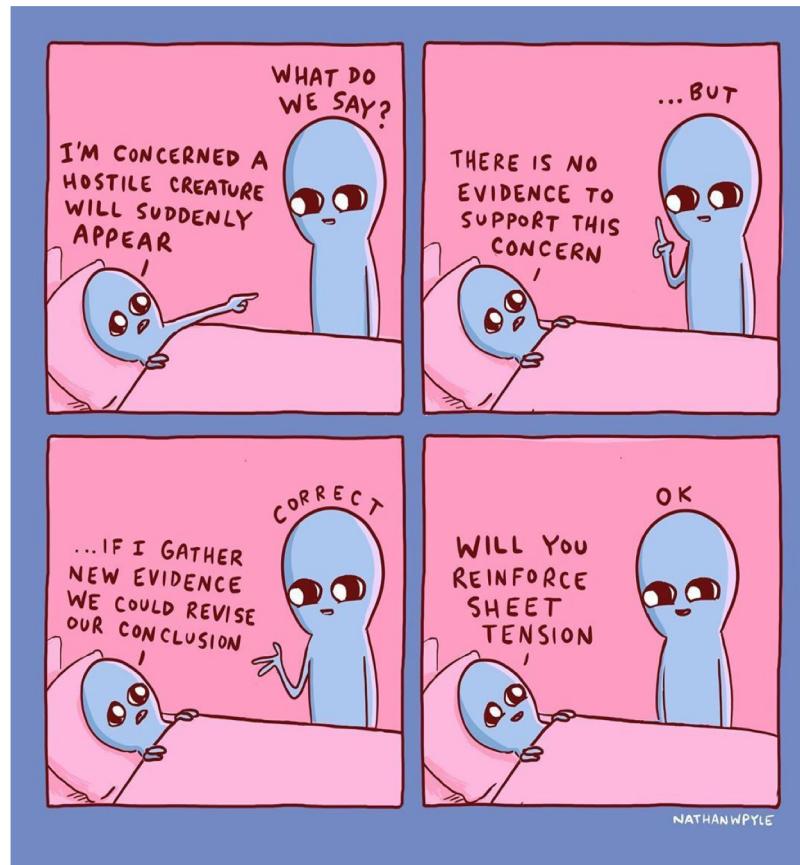


L26: Bayes' Theorem



3 End-of-Term Surveys

- Surveys due ***Wed. April 20th @11:59pm:***
 - Foundational Course Initiative (FCI) end of term survey.
 - Computing Cares Exit survey.
 - These are both assignments and part of your course grade
- Extra Credit
 - Complete your **Course Evaluations** for both lecture and discussion
 - Submit a screenshot of the confirmation page to Gradescope to earn extra credit added to your overall course grade.

Learning Objectives

After today's lecture (and associated readings, discussion & homework), you should be able to:

- State the relationship between Bernoulli Trials and Binomial Experiments, and between Bernoulli Trials and Geometric (aka “Waiting Time”) Experiments
- Calculate probabilities and expected values for Binomial and Geometric random variables.
- Conditional probability and how to use it.
- Reasoning from inconclusive evidence.
- Using Bayes’ theorem.
- Terminology: ***prior*** (distribution), ***likelihood***, ***posterior*** (distribution).

Overview

- **First part of lecture:**
 - Review on how to solve probability problems
 - Overview of two common “probability distributions”
- **Second part of lecture:**
 - How to use probability to *update your beliefs*
 - Bayes’ Law and applications

Outline

- **Geometric Distribution**
- Binomial Distribution
- Bayes' Theorem
 - Ex: Diagnosing a Rare Disease
- Law of Total Probability with Bayes'
 - Total Probability
 - Ex: Soccer Team Drug Testing
 - Ex: Pocket full of Dice
- Monty Hall Problem

Geometric Random Variables

- Suppose you have a biased coin comes up Heads with probability $p \in (0,1)$ and Tails with probability $1 - p$.
- You flip repeatedly until you get the first Head. Let X be the total number of flips (including the last Head).
- What is the probability distribution, $p(X = k)$?
 - What is the sample space S ?
 - What is $p(s)$ and $X(s)$ for each outcome?

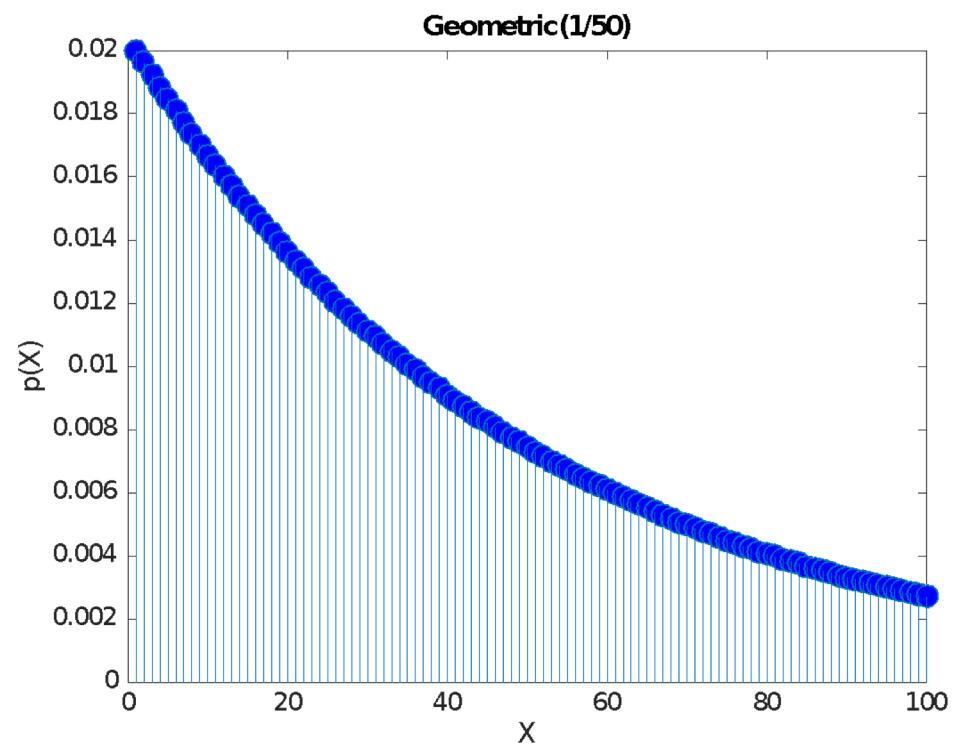
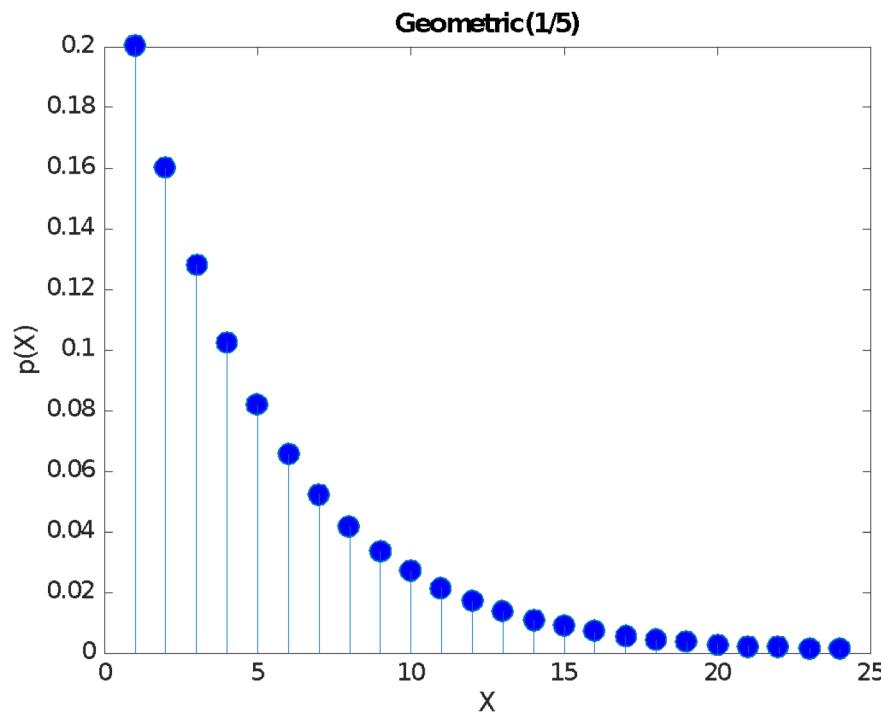
Geometric Random Variables

- Suppose you have a biased coin comes up Heads with probability $p \in (0,1)$ and Tails with probability $1 - p$.
- You flip repeatedly until you get the first Head. Let X be the total number of flips (including the last Head).
- What is the probability distribution, $p(X = k)$?
 - What is the sample space S ?
 - What is $p(s)$ and $X(s)$ for each outcome?
- We have a (countably) infinite sample space:

<u>s</u>	<u>p(s)</u>	<u>$X(s) = \# \text{ of flips until first Head}$</u>
– H	p	1
– TH	$(1 - p)p$	2
– TTH	$(1 - p)^2 p$	3
– ...		
– $T^{k-1}H$	$(1 - p)^{k-1} p$	k
— ...		

$$p(X = k) = (1 - p)^{k-1} p$$

Geometric Random Variables



Geometric Random Variables

- Suppose you have a biased coin comes up Heads with probability $p \in (0,1)$ and Tails with probability $1 - p$.
- You flip repeatedly until you get the first Head. Let X be the total number of flips (including the last Head).
- The probability distribution is $p(X = k) = (1 - p)^{k-1} p$
- What is $E(X)$?
- Consider the first flip. It's either heads or tails.
 - If it's tails, then $\underline{X = 1}$.
 - If it's heads, then we add 1 to our total and repeat from scratch.

$$E(X) = p \cdot 1 + (1 - p)(E(X) + 1)$$

Solve for $E(X)$ and you get...

$$E(X) = 1/p$$

Geometric Random Variables

- Suppose you have a biased coin comes up Heads with probability $p \in (0,1)$ and Tails with probability $1 - p$.
- You flip repeatedly until you get the first Head. Let X be the total number of flips (including the last Head).
 - Geometrically Distributed

A r.v. X has a ***geometric distribution*** with parameter p if:

$$p(X = k) = (1 - p)^{k-1} p$$

If X is geometrically distributed with parameter p then:

$$E(X) = 1/p$$

Watching Seinfeld Reruns

- Every night a Seinfeld episode is drawn *uniformly at random* from the 180 shows and broadcast.
- What is the *expected number of nights* you need to watch to see all episodes?
- This is a tricky problem if you don't start right!
 - Use linearity of expectations.
 - You know the expectation of a geometric distribution.

Seinfeld Reruns

- Let $X_{\underline{i} \rightarrow j}$ be the number of days you have to watch to go from having watched i distinct shows to having watched j distinct shows. Then,

$$\underline{X_{0 \rightarrow 180}} = X_{0 \rightarrow 1} + X_{1 \rightarrow 2} + \cdots + X_{178 \rightarrow 179} + X_{179 \rightarrow 180}$$

By Linearity of Expectations...

$\underline{X_{k \rightarrow k+1}}$ is a geometric r.v. with success probability $P = \frac{180-k}{180}$

$$E[X_{k \rightarrow k+1}] = \frac{180}{180-k}$$

$$E[X_{0 \rightarrow 180}] = \frac{180}{180} + \frac{180}{179} + \frac{180}{178} + \cdots + \frac{180}{2} + \frac{180}{1}$$

$$\approx 180 \cdot 5.77 < 1039$$

Outline

- Geometric Distribution
- **Binomial Distribution**
- Bayes' Theorem
 - Ex: Diagnosing a Rare Disease
- Law of Total Probability with Bayes'
 - Total Probability
 - Ex: Soccer Team Drug Testing
 - Ex: Pocket full of Dice
- Monty Hall Problem

Coin-toss Example

$$p(\text{heads}) = \frac{2}{3}$$

A coin whose probability of getting heads is $2/3$ is tossed 8 times.

What is the probability of exactly 3 heads in the 8 tosses?



“success” = heads

$$p(\text{success}) = \frac{2}{3}$$

One sequence with 3 heads:

T H H T T H T T

Probability of this sequence:

$$\left(\frac{1}{3} \frac{2}{3} \frac{2}{3} \frac{1}{3} \frac{1}{3} \frac{2}{3} \frac{1}{3} \frac{1}{3}\right) = (2/3)^3 * (1/3)^5$$

- So, what is the probability of any particular sequence with 3 heads? $(2/3)^3 * (1/3)^5$
- How many sequences with 3 heads are there? $\binom{8}{3}$

Probability of getting 3 heads in the 8 tosses:

$$P(3 \text{ heads}) = \binom{8}{3} \left(\frac{2}{3}\right)^3 \left(\frac{1}{3}\right)^5 \left\{ p(K \text{ heads}) = \binom{n}{K} p^K (1-p)^{n-K} \right.$$

what if $p(\text{success}) = p$

Coin-toss Example

A coin whose probability of getting heads is 2/3 is tossed 8 times.

What is the probability of exactly 3 heads in the 8 tosses?

One sequence with 3 heads:

Probability of this sequence: $\frac{1}{3} \frac{2}{3} \frac{2}{3} \frac{1}{3} \frac{1}{3} \frac{2}{3} \frac{1}{3} \frac{1}{3} = (2/3)^3 * (1/3)^5$

- So, what is the probability of any particular sequence with 3 heads? $(2/3)^3 * (1/3)^5$
- How many of sequences with 3 heads are there?

Probability of getting 3 heads:

$$P(3 \text{ heads}) = \binom{8}{3} (2/3)^3 (1/3)^5$$

Take-aways from the example:

- The *Experiment* consists of repeating a simpler experiment n times.
- This *simpler experiment* is “simpler” because its sample space consists of two outcomes.
- Events of interest may consist of the *number of times* one of the two simpler experimental outcomes occurs.

Bernoulli Trials, Binomial Experiment, and Binomial Distribution

- Bernoulli Trial
 - An experiment that has exactly 2 outcomes, $S = \{\text{success, failure}\}$
 - $p(\text{success}) = p$
 - $p(\text{failure}) = q = (1-p)$
- Binomial Experiment
 - Repeat the Bernoulli trial n times where each trial has the same probability of success and all trials are mutually independent
- Binomial Distribution
 - Let event be that k successes occur in n Bernoulli trials. Then

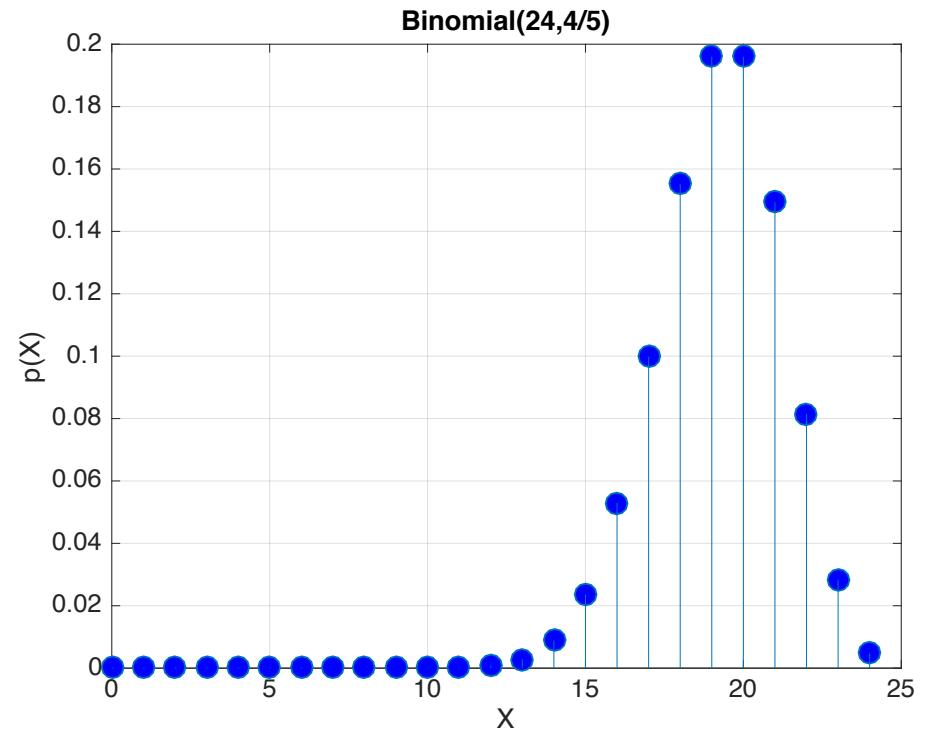
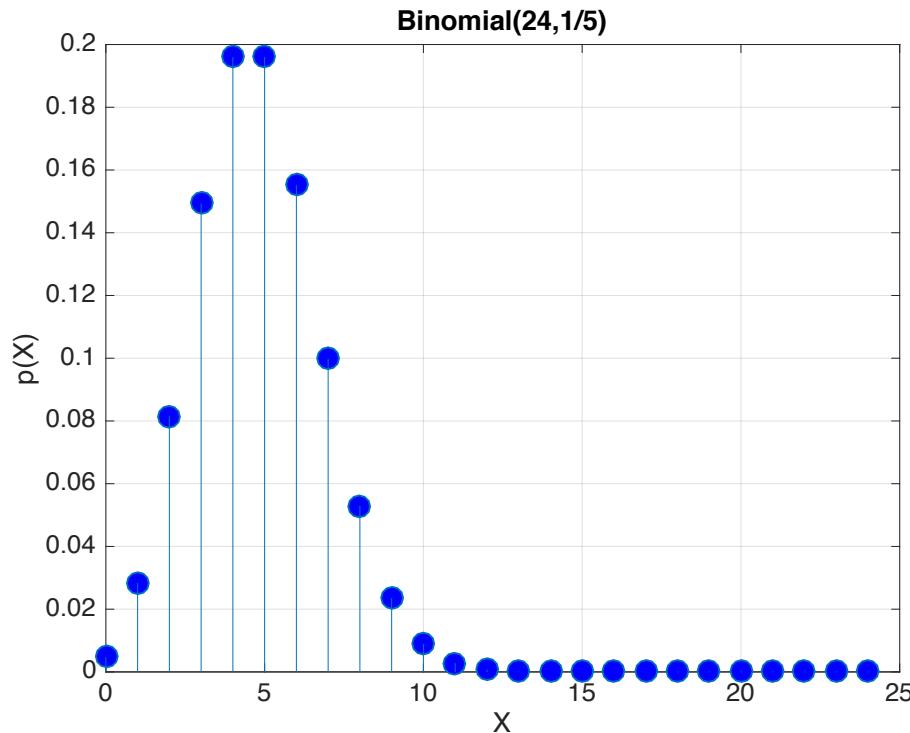
$$p(\text{exactly } k \text{ successes in } n \text{ trials}) = \binom{n}{k} p^k q^{n-k}$$

\downarrow
 $(1-p)^{n-k}$

Random Variable: Binomial Distribution

- X is the number of successes in n Bernoulli trials where probability of a success is p

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$



Lecture 26 Handout: Bernoulli Trials, Binomial Distribution & Bayes' Theorem

Bernoulli Trial

Experiment that has

exactly 2 outcomes,

- $S = \{\text{success, failure}\}$
- $p(\text{success}) = p$
- $p(\text{failure}) = q = 1 - p$

Binomial Experiment

Repeat the Bernoulli trial n times, where

- Each trial has the same probability of success
- All trials are mutually independent

Binomial Distribution:

The probability of exactly k successes in n independent (and identically distributed) Bernoulli trials is

$$\binom{n}{k} p^k (1-p)^{n-k}$$

Outline

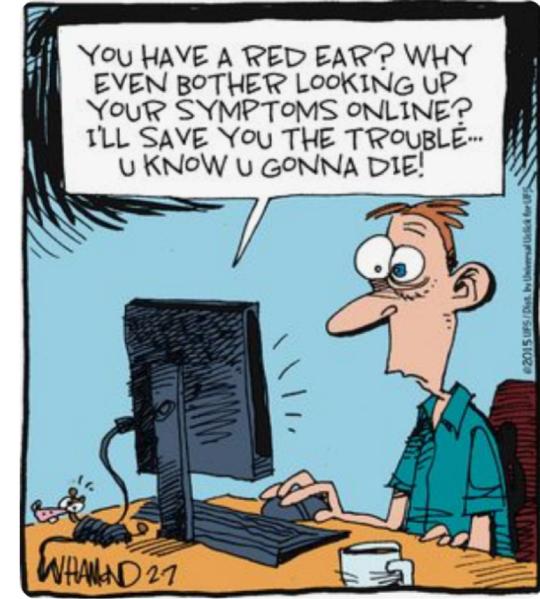
- Geometric Distribution
- Binomial Distribution
- Bayes' Theorem
 - Ex: Diagnosing a Rare Disease
- Law of Total Probability with Bayes'
 - Total Probability
 - Ex: Soccer Team Drug Testing
 - Ex: Pocket full of Dice
- Monty Hall Problem

Reasoning & Probability

- Applications of probability:
 - Reasoning about *games* and their value. If X is your lottery winnings, is $E(X) >$ the cost of the ticket?
 - Reasoning about *randomized algorithms* (more in EECS 281, 376, 477, etc.)
 - Reasoning about partial/inconclusive evidence. Using statistical evidence to make more accurate guesses about reality.
 - *Bayes' Theorem* is the most basic tool.
 - Critical in modern ML (more in EECS 492, 445)

Conditional Probabilities – Diagnosing a Rare Disease

- Meningitis is rare: $p(m) = \frac{1}{50,000} = 0.00002$
- Meningitis causes stiff neck: $p(s|m) = 0.5$
- Stiff neck is not so rare: $p(s) = \frac{1}{20} = 0.05$
- You have a stiff neck.
- What is $p(m|s)$, the probability of meningitis if you have a stiff neck?



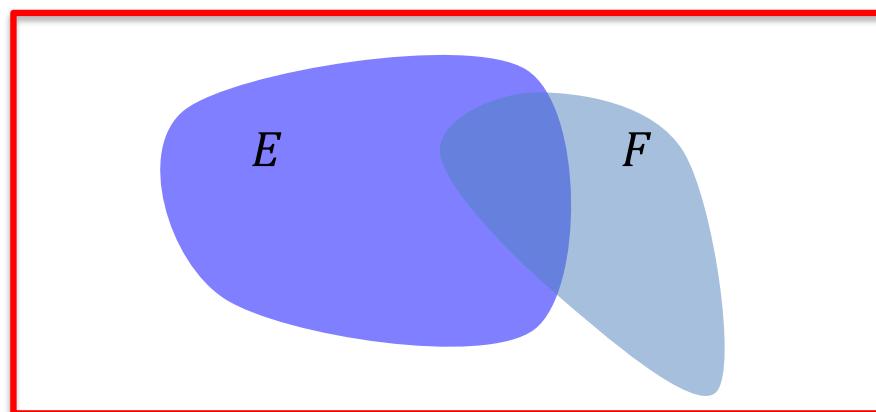
$$P(m|s) = \frac{P(s \cap m)}{P(s)}$$

↙ ↘ ↗

$$P(m|s) = \frac{P(s \cap m)}{P(m)}$$

Conditional Probability Review

- E, F are **events**.
- $p(F|E) = \frac{p(F \cap E)}{p(E)}$ is the **conditional probability** of F holding, given that E holds. (Only sensible if $p(E) > 0$.)
- Symmetrically: $p(E|F) = \frac{p(E \cap F)}{p(F)}$.



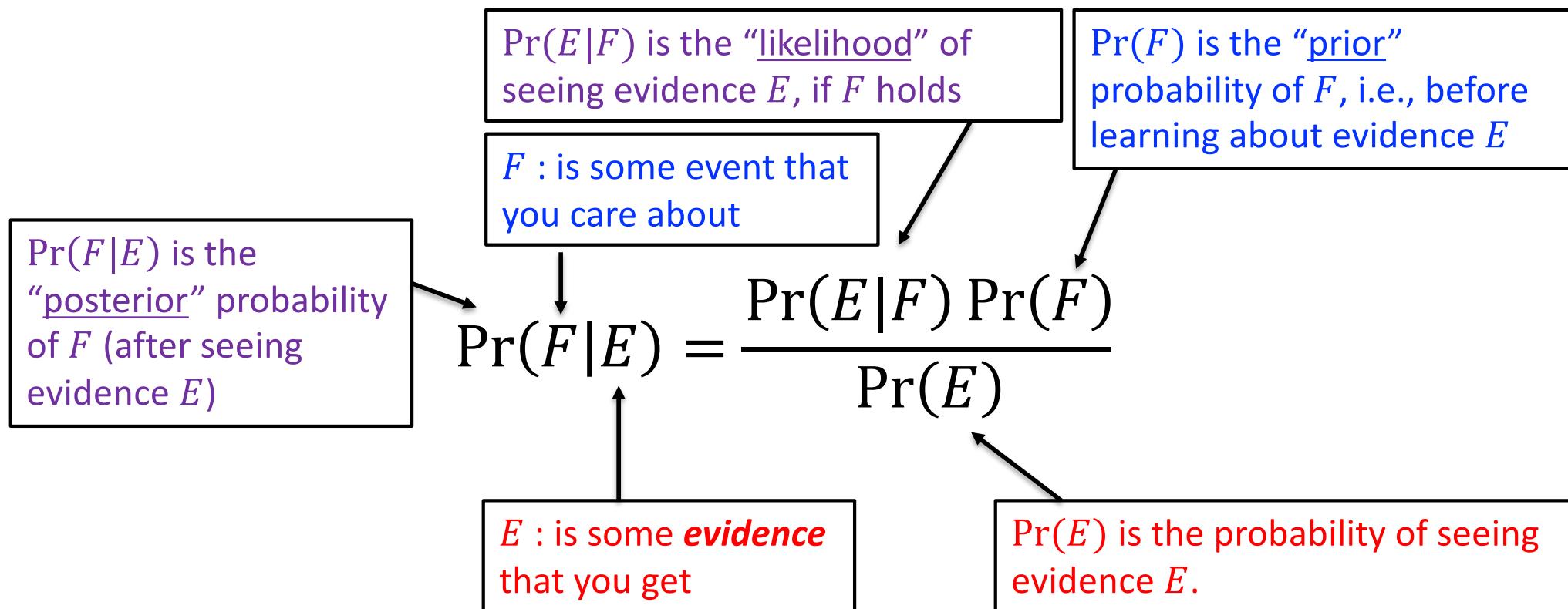
Bayes' Theorem

- **Bayes' Theorem:** (*is just rearranging terms in conditional probability*)

$$p(F|E) = \frac{p(E|F) p(F)}{p(E)}$$

Bayes' Theorem

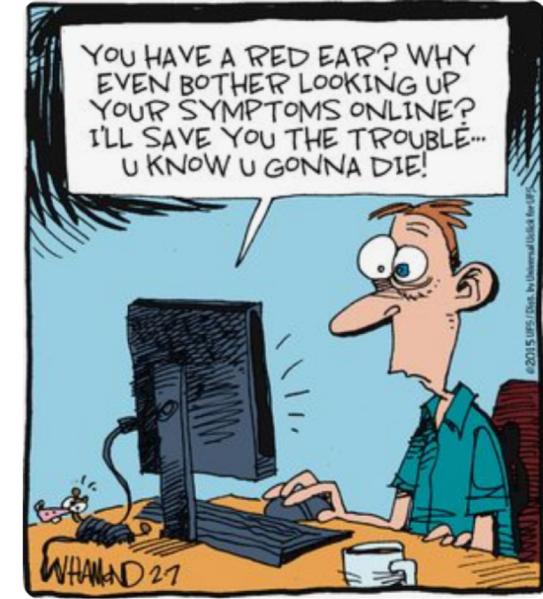
- **Bayes' Theorem:** (*is just rearranging terms in conditional probability*)



Conditional Probabilities – Diagnosing a Rare Disease

- Meningitis is rare: $p(m) = \frac{1}{50,000} = 0.00002$
- Meningitis **causes** stiff neck: $p(s|m) = 0.5$
- Stiff neck is not so rare: $p(s) = \frac{1}{20} = 0.05$
- You have a stiff neck.
- What is $p(m|s)$, the probability of meningitis if you have a stiff neck?

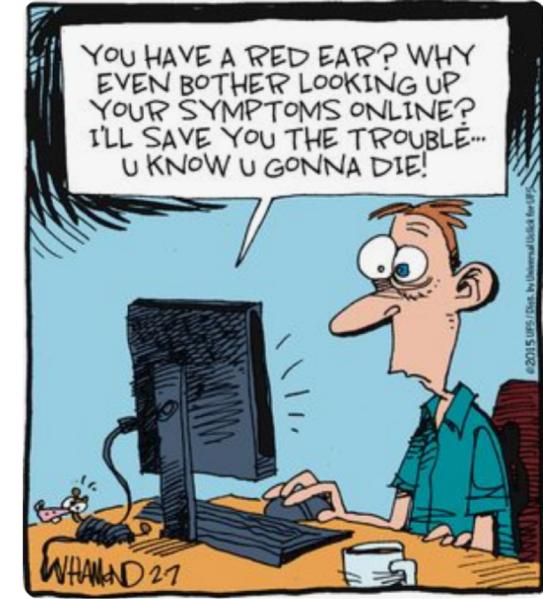
$$\begin{aligned} p(m|s) &= \frac{p(s|m)p(m)}{p(s)} \\ &= \frac{(0.5)(\frac{1}{50,000})}{\frac{1}{20}} = \frac{1}{5,000} \end{aligned}$$



Conditional Probabilities – Diagnosing a Rare Disease

- Meningitis is rare: $p(m) = \frac{1}{50,000} = 0.00002$
 - Meningitis **causes** stiff neck: $p(s|m) = 0.5$
 - Stiff neck is not so rare: $p(s) = \frac{1}{20} = 0.05$
 - You have a stiff neck.
-
- What is $p(m|s)$, the probability of meningitis if you have a stiff neck?
 - Solution: Apply Bayes' Theorem

$$p(m|s) = \frac{p(s|m)p(m)}{p(s)} = \frac{0.5 \times 0.00002}{0.05} = \frac{1}{5,000} = 0.0002$$



Having the additional information that you have a stiff neck increases the chances you also have meningitis by a factor of 10, but it's still really unlikely.

Bayes' Theorem

Suppose that E and F are events from a sample space S such that $p(E) > 0$ and $p(F) > 0$. Then

$$p(F|E) = \frac{p(E|F)p(F)}{p(E)}$$

Alternative form of Bayes', with expanded denominator:

$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\bar{F})p(\bar{F})}$$

Helpful tips for Bayes':

- $p(E|F) + p(\underline{\hspace{2cm}}) = 1$
- $p(E) + p(\underline{\hspace{2cm}}) = 1$
- w22 • $p(E|\bar{F}) + p(\bar{E}|\bar{F}) = \underline{\hspace{2cm}}$
- $p(\underline{\hspace{2cm}}) + p(\bar{F}) = 1$

Why is Bayes' Rule so Useful?

- Bayes' Rule = **How to Update Your Beliefs, Given New Evidence**
- In many situations
 - Diagnostic evidence $P(\text{disease} \mid \text{symptom})$ is often hard to get.
 - But it's usually what you really want.
 - This is the “posterior” probability of having the disease (after seeing the evidence of the symptom)
 - Causal evidence $P(\text{symptom} \mid \text{disease})$ is often easier to get.
 - This is the “likelihood” of seeing the evidence (the symptom), if someone has the disease

Outline

- Geometric Distribution
- Binomial Distribution
- Bayes' Theorem
 - Ex: Diagnosing a Rare Disease
- **Law of Total Probability with Bayes'**
 - Total Probability
 - Ex: Soccer Team Drug Testing
 - Ex: Pocket full of Dice (*time permitting*)
- Monty Hall Problem

Soccer Team Drug Testing

- When a test for steroids is given to soccer players
 - 98% of the players taking steroids test positive
 - 12% of the players not taking steroids test positive
- Suppose 5% of soccer players take steroids.
- What is the probability that a soccer player who tests positive takes steroids?
Find $P(S|T)$
 - Let S = event that a player takes Steroids,
Let T = event that a player Tested positive

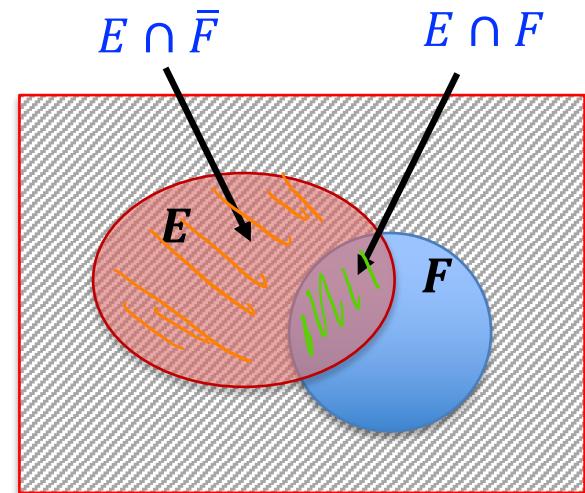
$$\begin{aligned} P(S|T) &= \frac{P(T|S) P(S)}{P(T)} \\ &= \frac{P(T|S) P(S)}{P(T|S) P(S) + P(T|\bar{S}) P(\bar{S})} \\ &= \frac{(0.98)(0.05)}{(0.98)(0.05) + (0.12)(0.95)} \\ &\approx .301 \end{aligned}$$

Law of Total Probability

For two events E and F , the **total probability** of E is

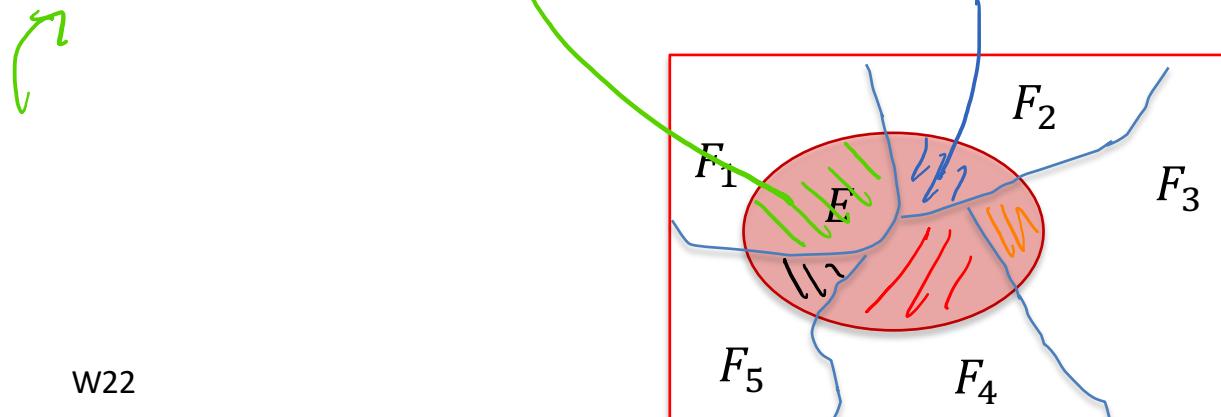
$$p(E) = p(E \cap F) + p(E \cap \bar{F})$$

$$p(E) = p(E|F)p(F) + p(E|\bar{F})p(\bar{F})$$



If events F_1, F_2, \dots, F_n partition the sample space, then the **total probability** of event E is

$$p(E) = p(E|F_1)p(F_1) + p(E|F_2)p(F_2) + \dots + p(E|F_n)p(F_n)$$



Bayes' Theorem

- **Bayes' Theorem:** (*is just rearranging terms in conditional probability*)

$$p(F|E) = \frac{p(E|F)p(F)}{p(E)}$$

- Sometimes you don't have direct access to the probability in the denominator.
 - Either F or \bar{F} holds. $E = E \cap (F \cup \bar{F}) = (E \cap F) \cup (E \cap \bar{F})$
 - $p(E) = p(E \cap F) + p(E \cap \bar{F}) = p(E|F)p(F) + p(E|\bar{F})p(\bar{F})$

$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\bar{F})p(\bar{F})}$$

$p(E)$

Bayes' Theorem

Suppose that E and F are events from a sample space S such that $p(E) > 0$ and $p(F) > 0$. Then

$$p(F|E) = \frac{p(E|F)p(F)}{p(E)}$$

Alternative form of Bayes', with expanded denominator:

$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\bar{F})p(\bar{F})}$$

Helpful tips for Bayes':

- $p(E|F) + p(\underline{\hspace{2cm}}) = 1$
- $p(E) + p(\underline{\hspace{2cm} E}) = 1$
- $p(E|\bar{F}) + p(\bar{E}|\bar{F}) = \underline{\hspace{2cm}}$
- $p(\underline{\hspace{2cm} F}) + p(\bar{F}) = 1$

Soccer Team Drug Testing

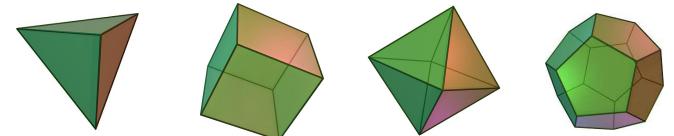
- When a test for steroids is given to soccer players
 - 98% of the players taking steroids test positive
 - 12% of the players not taking steroids test positive
- Suppose 5% of soccer players take steroids.
- What is the probability that a soccer player who tests positive takes steroids?
 - Let S = event that a player takes Steroids,
Let T = event that a player Tested positive

Soccer Team Drug Testing

- When a test for steroids is given to soccer players
 - 98% of the players taking steroids test positive
 - 12% of the players not taking steroids test positive
- Suppose 5% of soccer players take steroids.
- What is the probability that a soccer player who tests positive takes steroids?
 - Let S = event that a player takes Steroids,
Let T = event that a player Tested positive

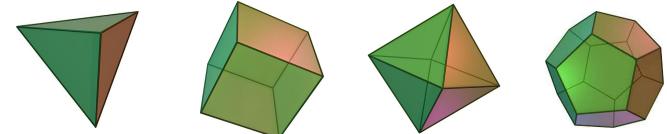
$$\begin{aligned} p(S|T) &= \frac{p(T|S)p(S)}{p(T|S)p(S) + p(T|\bar{S})p(\bar{S})} \\ &= \frac{p(T|S)p(S)}{p(T|S)p(S) + p(T|\bar{S})(1 - p(S))} \\ &= \frac{(0.98)(0.05)}{(0.98)(0.05)+(0.12)(0.95)} \approx 0.301 \end{aligned}$$

A Pocket Full of Dice



- In my pocket I have a 4-sided, 6-sided, 8-sided, and 12-sided die. I pick one and roll it twice and tell you that it showed the same number (**S**) both times. What's the probability that it is a **T**(etrahedron), **C**(ube), **O**(ctahedron), **D**(odecahedron)?

A Pocket Full of Dice



- In my pocket I have a 4-sided, 6-sided, 8-sided, and 12-sided die. I pick one and roll it twice and tell you that it showed the same number (S) both times. What's the probability that it is a T (etrahedron), C (ube), O (ctahedron), D (odecahedron)?
- What are the priors?
 - $\Pr(T), \Pr(C), \Pr(O), \Pr(D) = \frac{1}{4}$

- What are

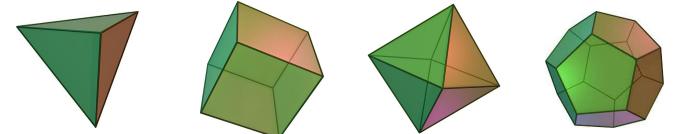
- $\Pr(S|T) = 1/4$
- $\Pr(S|C) = 1/6$
- $\Pr(S|O) = 1/8$
- $\Pr(S|D) = 1/12$

$$\begin{aligned} \Pr(S) &= \Pr(S|T) \Pr(T) + \Pr(S|C) \Pr(C) + \Pr(S|O) \Pr(O) + \Pr(S|D) \Pr(D) \\ &= \frac{1}{4} \cdot \left(\frac{1}{4} + \frac{1}{6} + \frac{1}{8} + \frac{1}{12} \right) = \frac{1}{4} \cdot \frac{15}{24} = \frac{5}{32} \end{aligned}$$

$$\Pr(T|S) = \frac{\Pr(S|T) \Pr(T)}{\Pr(S)} = \frac{\frac{1}{4} \cdot \frac{1}{4}}{\frac{5}{32}} = \frac{2}{5}$$

$$\Pr(C|S) = \frac{\Pr(S|C) \Pr(C)}{\Pr(S)} = \frac{\frac{1}{6} \cdot \frac{1}{4}}{\frac{5}{32}} = \frac{4}{15}$$

A Pocket Full of Dice



- In my pocket I have a 4-sided, 6-sided, 8-sided, and 12-sided die. I pick one and roll it twice and tell you that it showed the same number (S) both times. What's the probability that it is a T (etrahedron), C (ube), O (ctahedron), D (odecahedron)?
- What are the priors?
 - $\Pr(T), \Pr(C), \Pr(O), \Pr(D) = \frac{1}{4}$

- What are

- $\Pr(S|T) = 1/4$
- $\Pr(S|C) = 1/6$
- $\Pr(S|O) = 1/8$
- $\Pr(S|D) = 1/12$

$$\begin{aligned} \Pr(S) &= \Pr(S|T)\Pr(T) + \Pr(S|C)\Pr(C) + \Pr(S|O)\Pr(O) + \Pr(S|D)\Pr(D) \\ &= \frac{1}{4} \cdot \left(\frac{1}{4} + \frac{1}{6} + \frac{1}{8} + \frac{1}{12} \right) = \frac{1}{4} \cdot \frac{15}{24} = \frac{5}{32} \end{aligned}$$

$$\Pr(O|S) = \frac{\Pr(S|O)\Pr(O)}{\Pr(S)} = \frac{\frac{1}{8} \cdot \frac{1}{4}}{\frac{5}{32}} = \frac{1}{5}$$

$$\Pr(D|S) = \frac{\Pr(S|D)\Pr(D)}{\Pr(S)} = \frac{\frac{1}{12} \cdot \frac{1}{4}}{\frac{5}{32}} = \frac{2}{15}$$

Outline

- Geometric Distribution
- Binomial Distribution
- Bayes' Theorem
 - Ex: Diagnosing a Rare Disease
- Law of Total Probability with Bayes'
 - Total Probability
 - Ex: Soccer Team Drug Testing
 - Ex: Pocket full of Dice
- Monty Hall Problem
 - Let's Make a Deal



Monty Hall Problem

From an old game show, “Let’s Make a Deal”



- You pick a door: A, B, C.
- Behind one is a new car. Behind the other two are goats.
 - Goats are cool but assume for this problem you want the car
- Monty opens one of the other two doors that ***does not*** have the car and shows you a goat.
- He then asks whether you'd like to switch doors.
 - (a) Doesn't matter, it's 50-50 either way.
 - (b) Always switch.
 - (c) Never switch.



Monty Hall Problem

From an old game show, “Let’s Make a Deal”



- You pick a door: A, B, C.
- Behind one is a new car. Behind the other two are goats.
 - Goats are cool but assume for this problem you want the car
- Monty opens one of the other two doors that *does not* have the car and shows you a goat.
- He then asks whether you'd like to switch doors.
 - (a) Doesn't matter, it's 50-50 either way.
 - **(b) Always switch.**
 - (c) Never switch.

Monty Hall Problem



Suppose you initially pick curtain “A,” and Monty reveals a goat behind curtain “B”. (Using *without loss of generality*, this covers all cases)

You have **new evidence** (goat behind door B), so how should you **update your beliefs**?

Bayes' will help us

We're going to set up a grid to better visualize the problem

Four Possibilities:	Car behind A	Car not behind A
Monte reveals B		
Monte reveals C		

Monty Hall Problem



Suppose you initially pick curtain “A,” and Monty reveals a goat behind curtain “B”

Four Possibilities:	Car behind A	Car not behind A
Monte reveals B		
Monte reveals C		

$$\Pr[\text{Car A and Reveal B}] = \Pr[\text{Reveal B} \mid \text{Car A}] \cdot \Pr[\text{Car A}]$$

$$= \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$$

$$\Pr[\text{Car A}] = \frac{1}{3}$$

$$\Pr[\text{Reveal B} \mid \text{Car A}] = \frac{1}{2}$$

Monty Hall Problem



Suppose you initially pick curtain “A,” and Monte reveals a goat behind curtain “B”

Four Possibilities:	Car behind A	Car not behind A
Monte reveals B	$\frac{1}{6}$	
Monte reveals C		

$\Pr[\text{Car not A and Reveal B}]$

$$= \Pr[\text{Reveal B} | \text{Car not A}] \cdot \Pr[\text{Car not A}]$$

$$= \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$$

$$\Pr[\text{Car not A}] = \frac{2}{3}$$

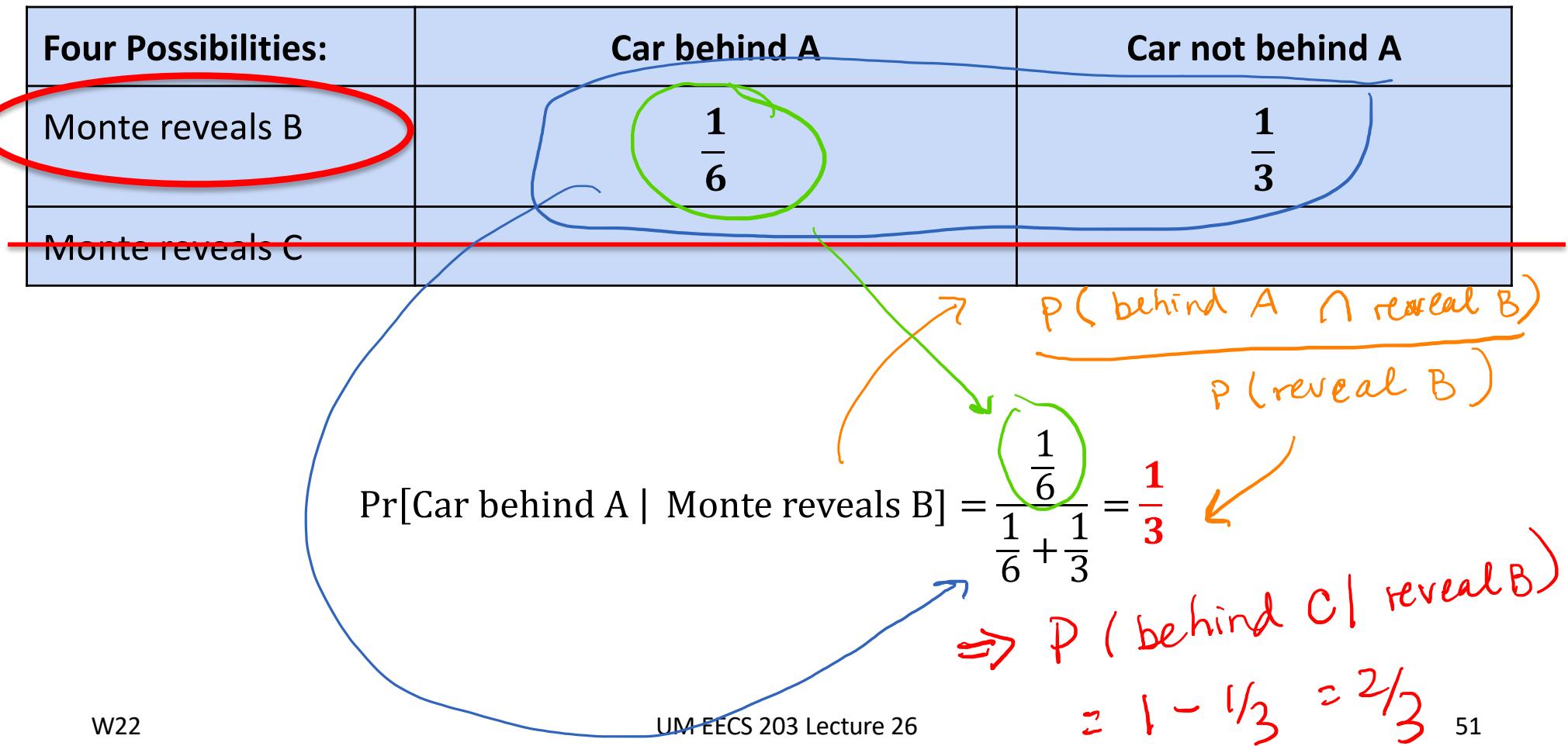
- Car in B or C
- So Monte reveals B iff car in C

$$\Pr[\text{Reveal B} | \text{Car not A}] = \frac{1}{2}$$

Monty Hall Problem



Suppose you initially pick curtain “A,” and Monte reveals a goat behind curtain “B”



That's a wrap on Probability

- Final two lectures will be on algorithms and how fast they run