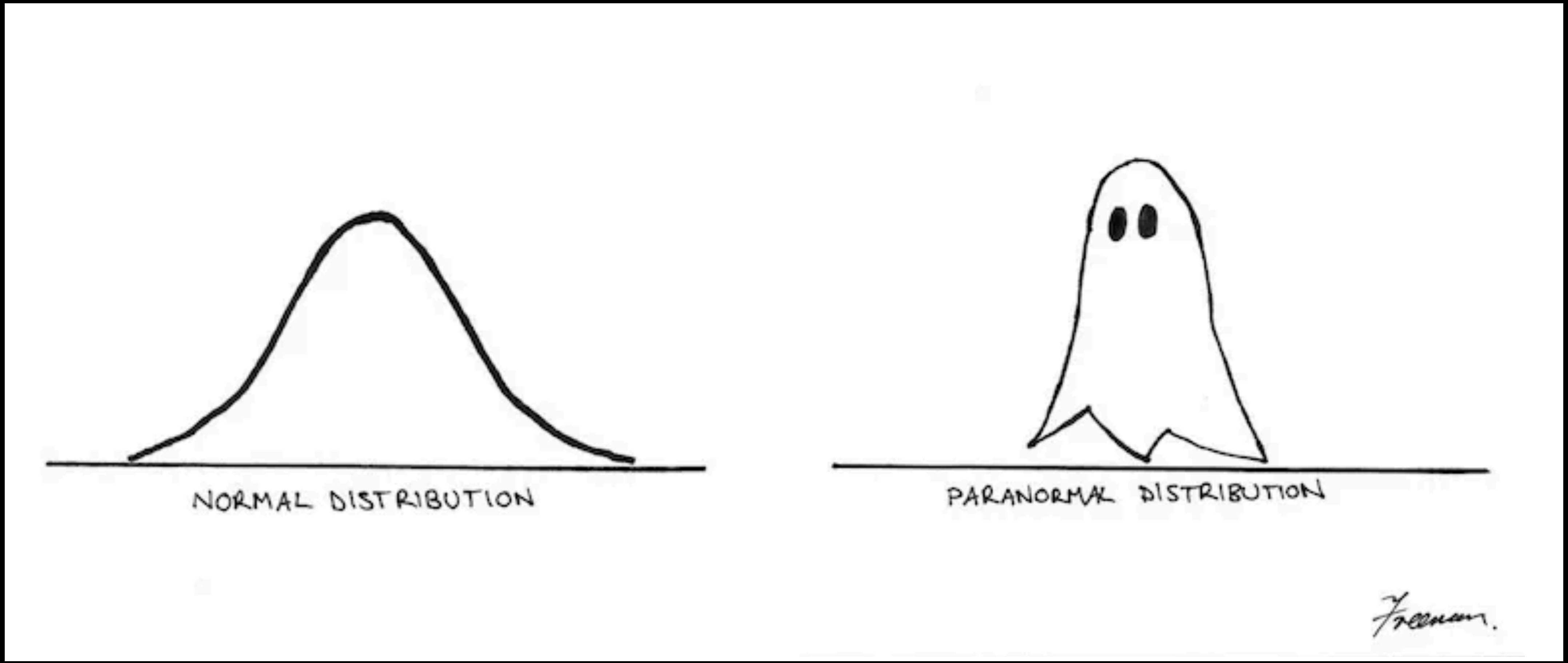


CS 109 Midterm Review

Julia Daniel, 10/27/2018

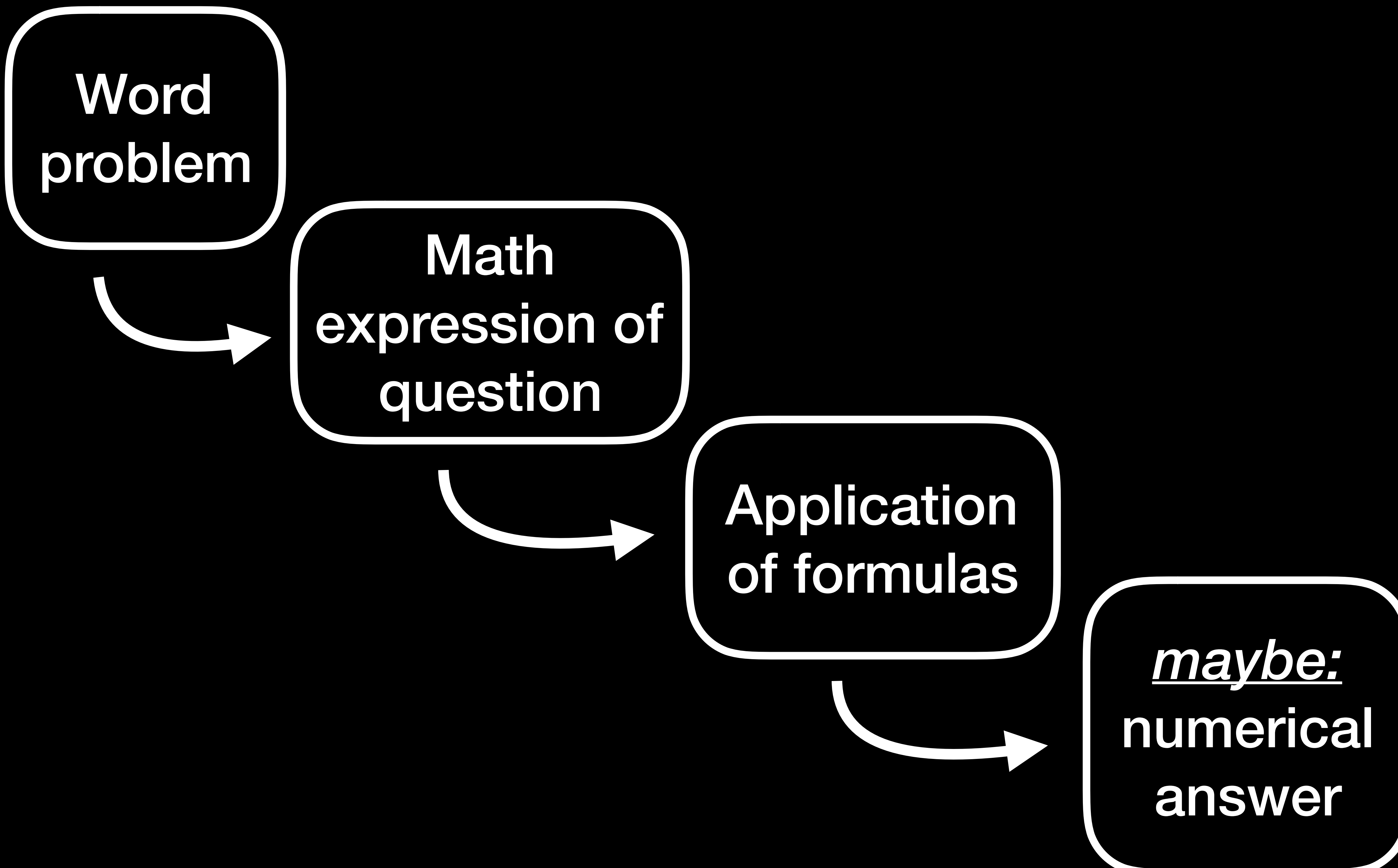
Outline

- General Strategies
- Counting and Events
- Probability Rules
- Random Variables

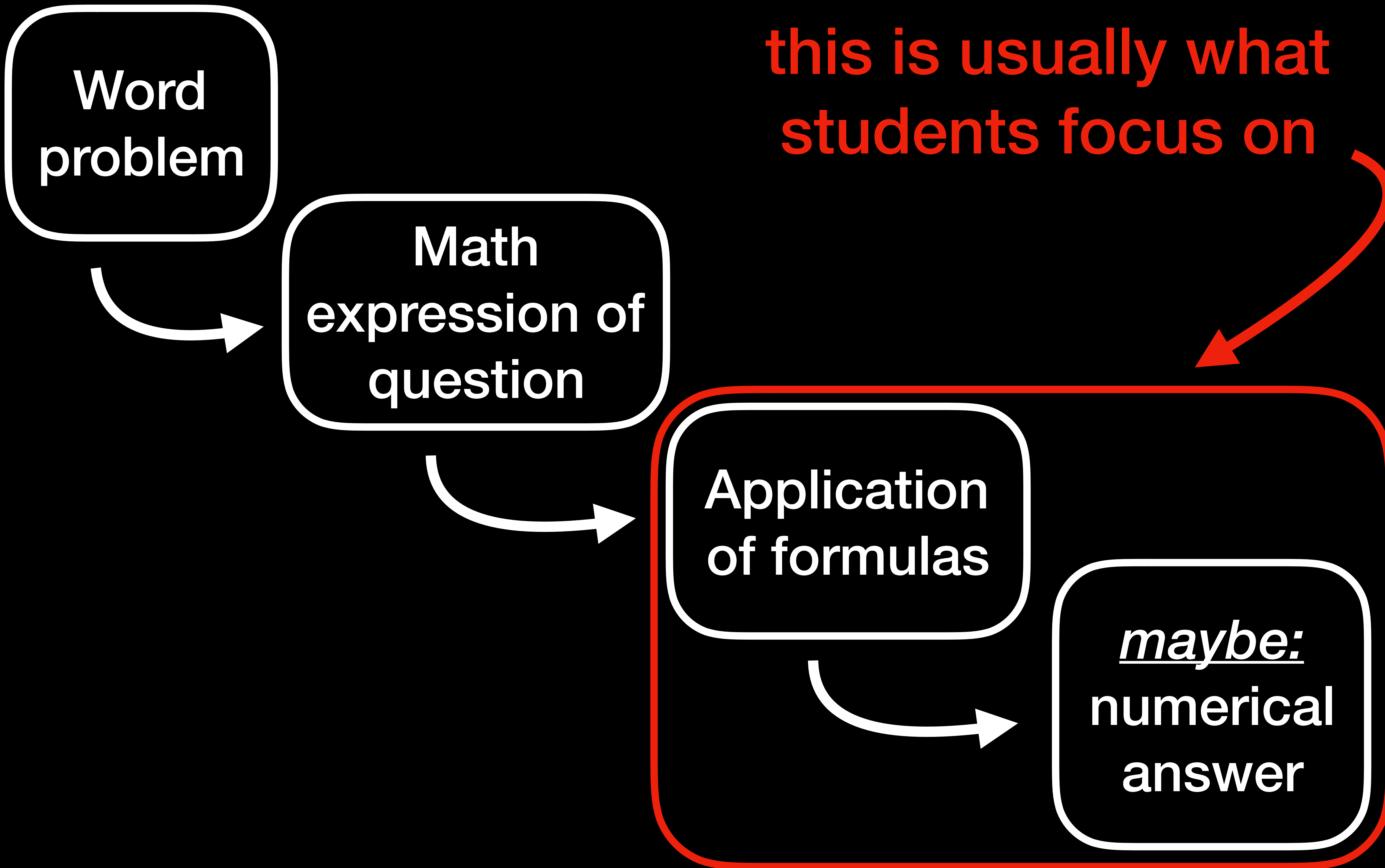
Outline

- General Strategies
- Counting and Events
- Probability Rules
- Random Variables

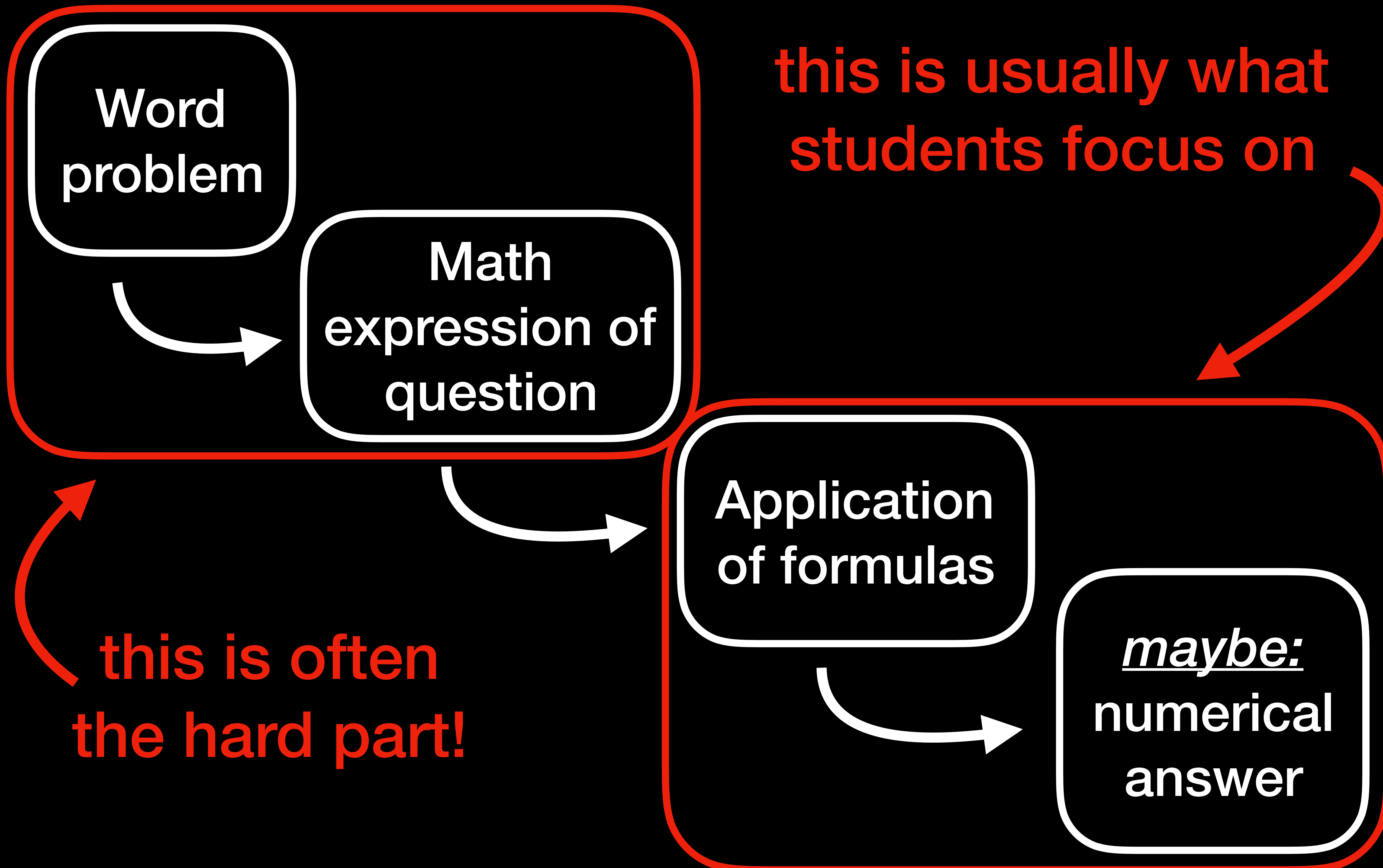
Solving a CS109 problem



Solving a CS109 problem



Solving a CS109 problem



Step 1: Defining Your Terms

- What's a 'success'? What's the sample space?
- What does each random variable actually represent, in English? Every definition of an event or a random variable should have a **verb** in it. (' = ' is a verb)
- Make sure units match - particularly important for λ

Step 1: Defining Your Terms

- What’s a ‘success’? What’s the sample space?
- What does each random variable actually represent, in English? Every definition of an event or a random variable should have a **verb** in it. (‘ = ’ is a verb)
- Make sure units match - particularly important for λ

Translating English to Probability

What the problem asks:	**What you should immediately think:**
“What’s the probability of _____ ”	**P( )**
“___ given ___”, “___ if ___”	**___ | ___**
“at least ___”	***could we use what we know about everything less than __?***
“approximate ____.”	***use an approximation!***
“How many ways…”	***combinatorics***

these are just a few, and these are why practice is the best way to prepare for the exam!

Translating English to Probability

People can have blue or brown eyes. What’s the probability John has blue eyes if his mother has brown eyes?

Translating English to Probability

People can have blue or brown eyes.
What's the probability John has blue eyes if his mother has brown eyes?

1. **What events are we given?**
2. **What are we asked to solve?**

Outline

- General Strategies
- Counting and Events
- Probability Rules
- Random Variables

Counting

Sum Rule
$outcomes = \|A\| + \|B\|$ $if\ \|A \cap B\| = 0$
I can choose to dress up as one of 5 superheroes or one of 4 farm animals. How many costume choices?

Counting

Sum Rule	Inclusion-Exclusion Principle
$outcomes = \|A\| + \|B\|$ $if\ \|A \cap B\| = 0$	$\|A\| + \|B\| - \|A \cap B\|$ $for\ any\ \|A \cap B\|$
I can choose to dress up as one of 5 superheroes or one of 4 farm animals. How many costume choices?	**I can choose to dress up as one of 5 superheroes or one of 6 strong female movie leads. 2 of the superheroes are female movie leads. How many costume choices?**

Counting

Product Rule
$$outcomes = \|A\| \times \|B\|$$ **if all outcomes of B are possible regardless of the outcome of A**
I can choose to go to one of 3 parties and then trick-or-treat in one of 5 neighborhoods. How many different ways to celebrate?

Counting

Product Rule	Pigeonhole Principle
$outcomes = \|A\| \times \|B\|$ **if all outcomes of B are possible regardless of the outcome of A**	**If m objects are placed into n buckets, then at least one bucket has at least *ceiling(m / n)* objects.**
I can choose to go to one of 3 parties and then trick-or-treat in one of 5 neighborhoods. How many different ways to celebrate?	**If you have an infinite number of red, white, blue, and green socks in a drawer, how many must you pull out before being guaranteed a pair?**

Combinatorics: Arranging Items

	Permutations (ordered)	Combinations (unordered)
Distinct	$n!$	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$
Indistinct	$\frac{n!}{k_1!k_2!\ldots k_n!}$	$\binom{n+r-1}{r-1}$ the divider method!

Outline

- General Strategies
- Counting and Events
- Probability Rules
- Random Variables

Probability basics

$$P(E) = \lim_{x->\infty} \frac{n(E)}{n}$$

in the general case

Probability basics

$$P(E) = \lim_{x->\infty} \frac{n(E)}{n}$$

in the general case

$$\text{Probability} = \frac{\text{Event space}}{\text{Sample space}}$$

if all outcomes are equally likely!

(use counting with distinct objects)

Probability basics

$$P(E) = \lim_{x -> \infty} \frac{n(E)}{n}$$

in the general case

$$\text{Probability} = \frac{\text{Event space}}{\text{Sample space}}$$

if all outcomes are equally likely!

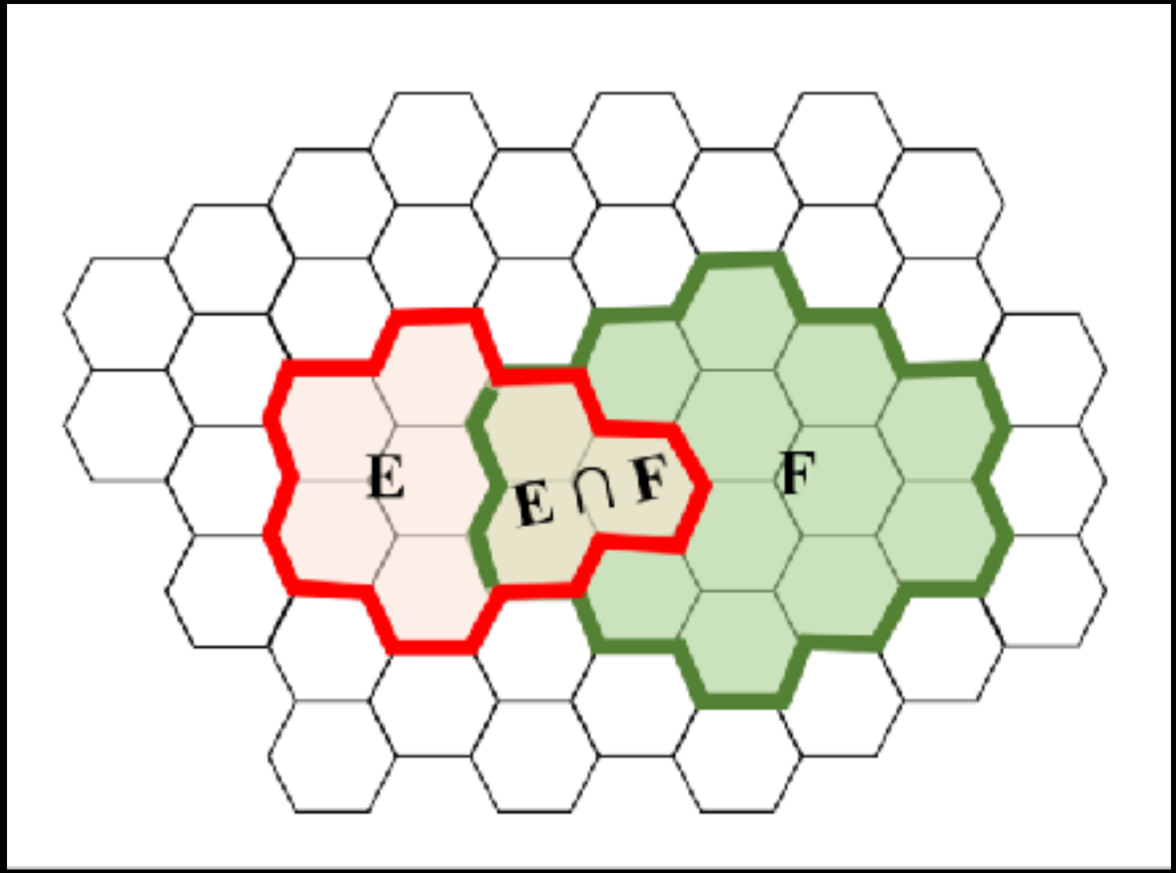
(use counting with distinct objects)

Axioms: $0 \leq P(E) \leq 1$ $\quad P(S) = 1$ $\quad P(E^C) = 1 - P(E)$

Conditional Probability

definition:

$$P(E\,|\,F) = \frac{P(EF)}{P(F)}$$



* $P(EF) = P(E \cap F)$

Chain Rule:

$$P(EF) = P(E\,|\,F)P(F)$$

Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

We can either walk to class, or we can bike.

If we walk to class, we have a 75% chance of being late.
If we bike, we have a 10% chance of being late.

We walk if we can't find our bike key, which happens 30% of the time.

What's our probability of being late to class?

Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

We can either walk to class, or we can bike.
If we walk to class, we have a 50% chance of being late.
If we bike, we have a 10% chance of being late.
We walk if we can't find our bike key, which happens 30% of the time. What's our probability of being late to class?

Event W = we walk to class. Event B = we bike = W^C.

Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

We can either walk to class, or we can bike.
If we walk to class, we have a 50% chance of being late.
If we bike, we have a 10% chance of being late.
We walk if we can't find our bike key, which happens 30% of the time. What's our probability of being late to class?

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.

Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

We can either walk to class, or we can bike.
If we walk to class, we have a 50% chance of being late.
If we bike, we have a 10% chance of being late.
We walk if we can't find our bike key, which happens 30% of the time. What's our probability of being late to class?

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.

Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

We can either walk to class, or we can bike.
If we walk to class, we have a 50% chance of being late.
If we bike, we have a 10% chance of being late.
We walk if we can't find our bike key, which happens 30% of the time. What's our probability of being late to class?

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?

Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?

Law of Total Probability

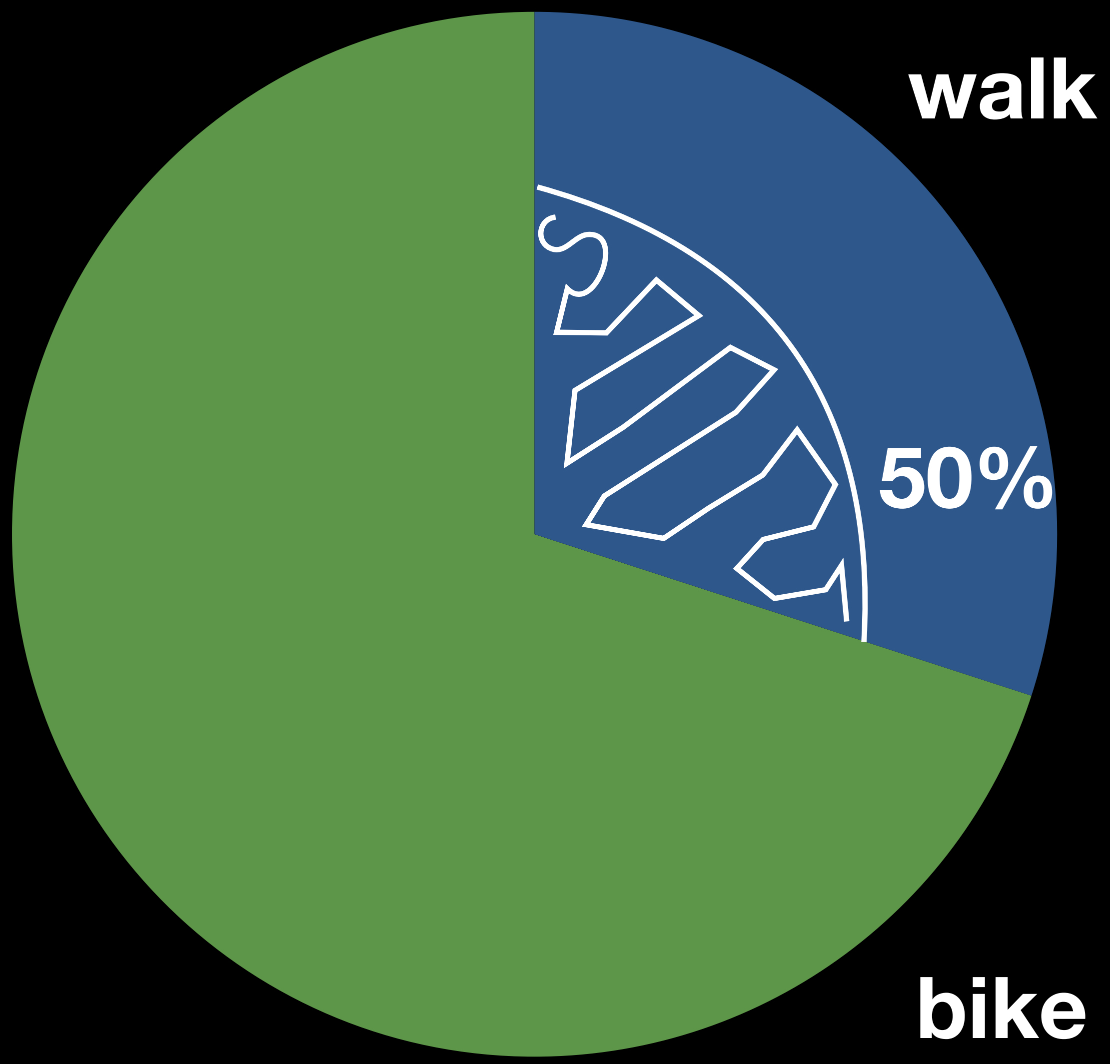
$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?

Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

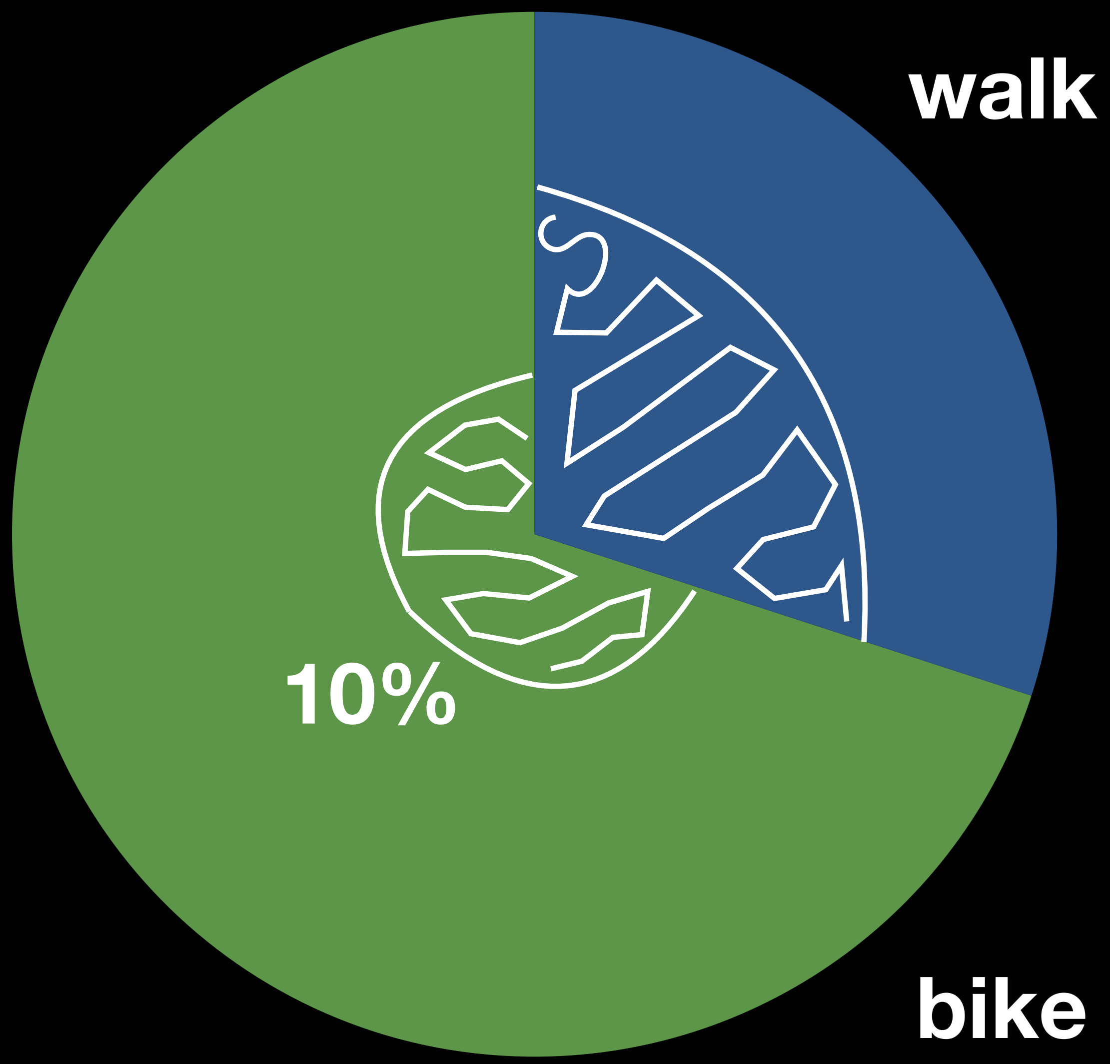
Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?



Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

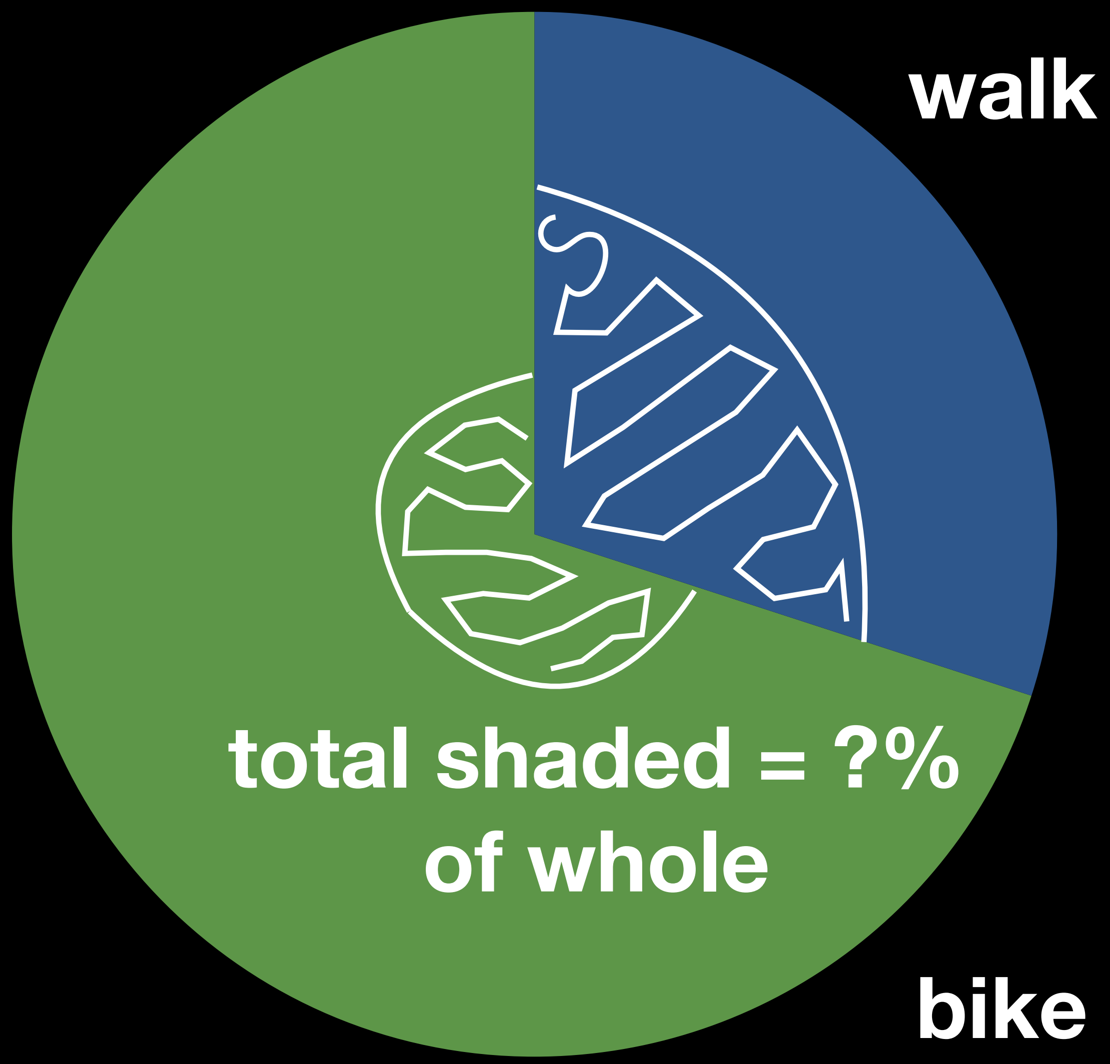
Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?



Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?

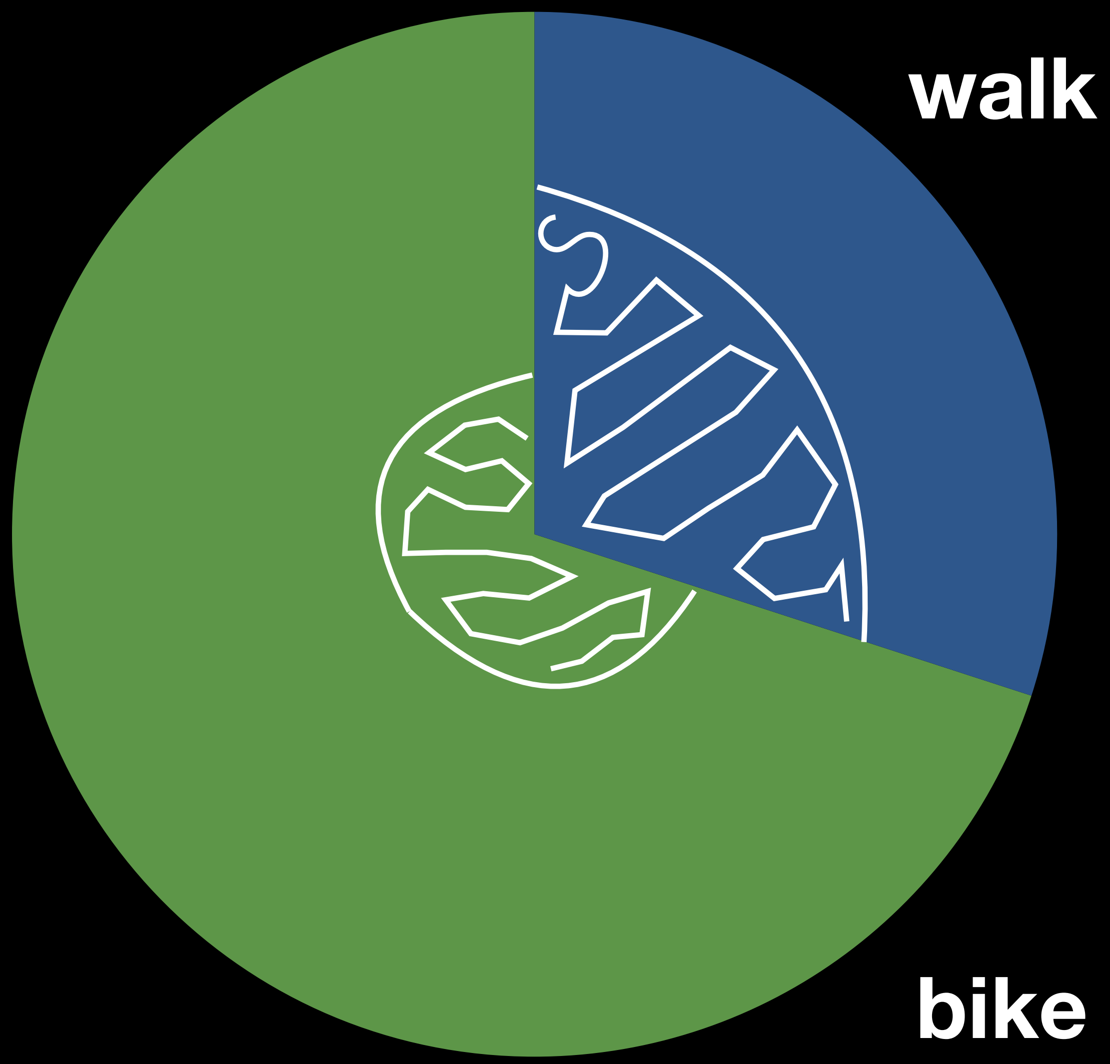


Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?

$$P(L) = P(L \mid W)P(W) + P(L \mid W^C)P(W^C)$$



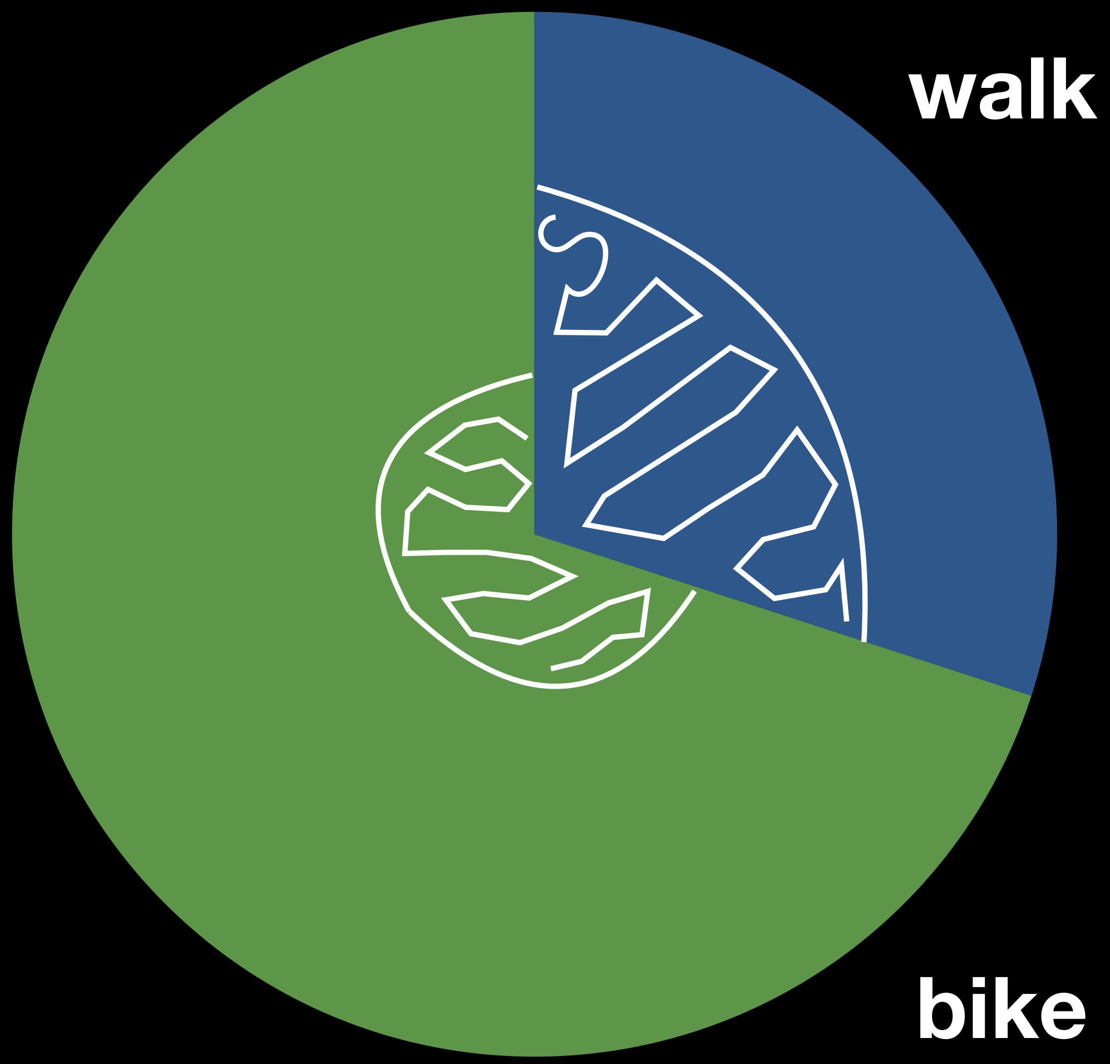
Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?

$$P(L) = P(L \mid W)P(W) + P(L \mid W^C)P(W^C)$$

$$= (0.5)(0.3) + (0.1)(0.7)$$



Law of Total Probability

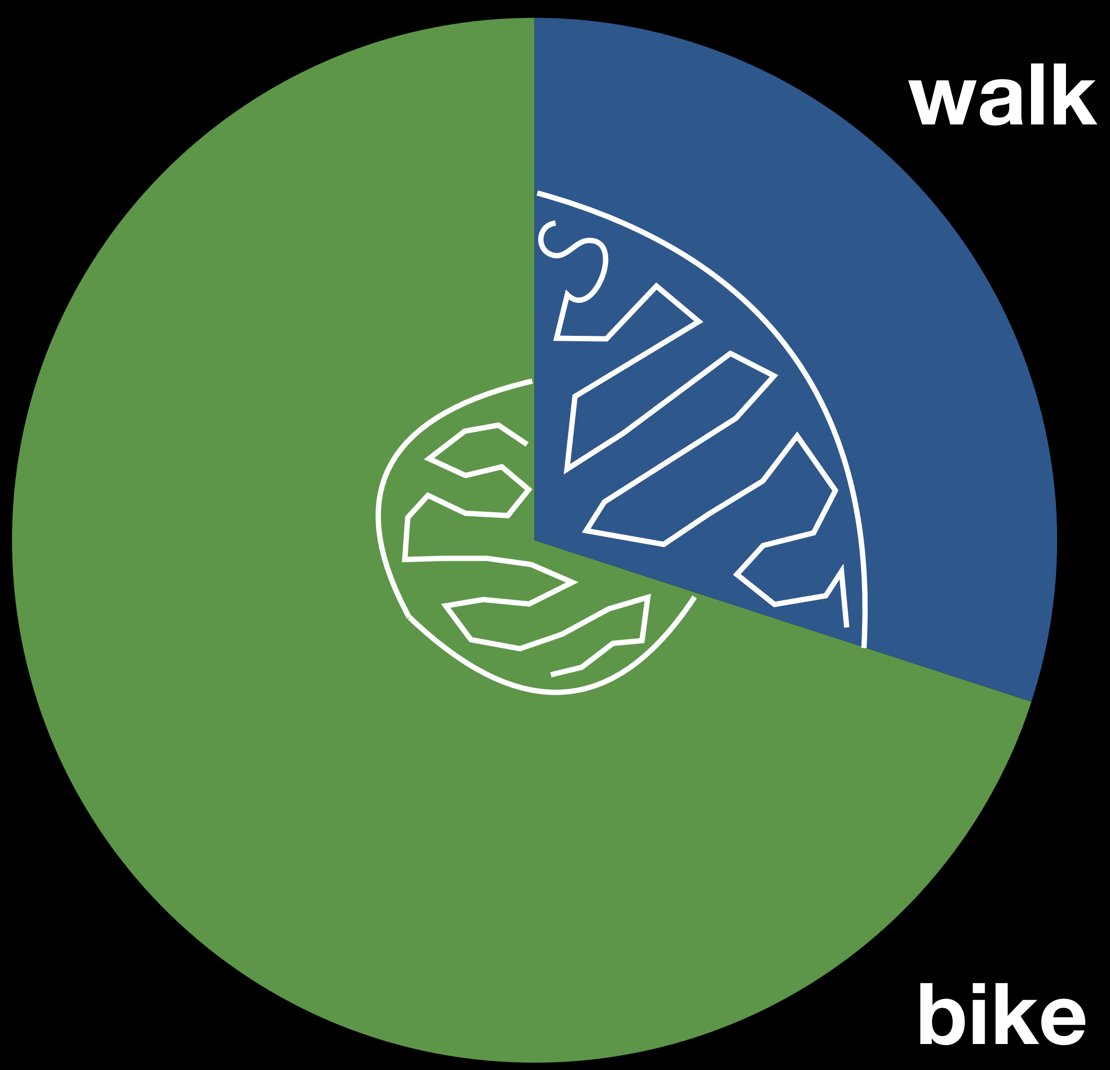
$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?

$$P(L) = P(L \mid W)P(W) + P(L \mid W^C)P(W^C)$$

$$= (0.5)(0.3) + (0.1)(0.7)$$

$$= 0.22$$

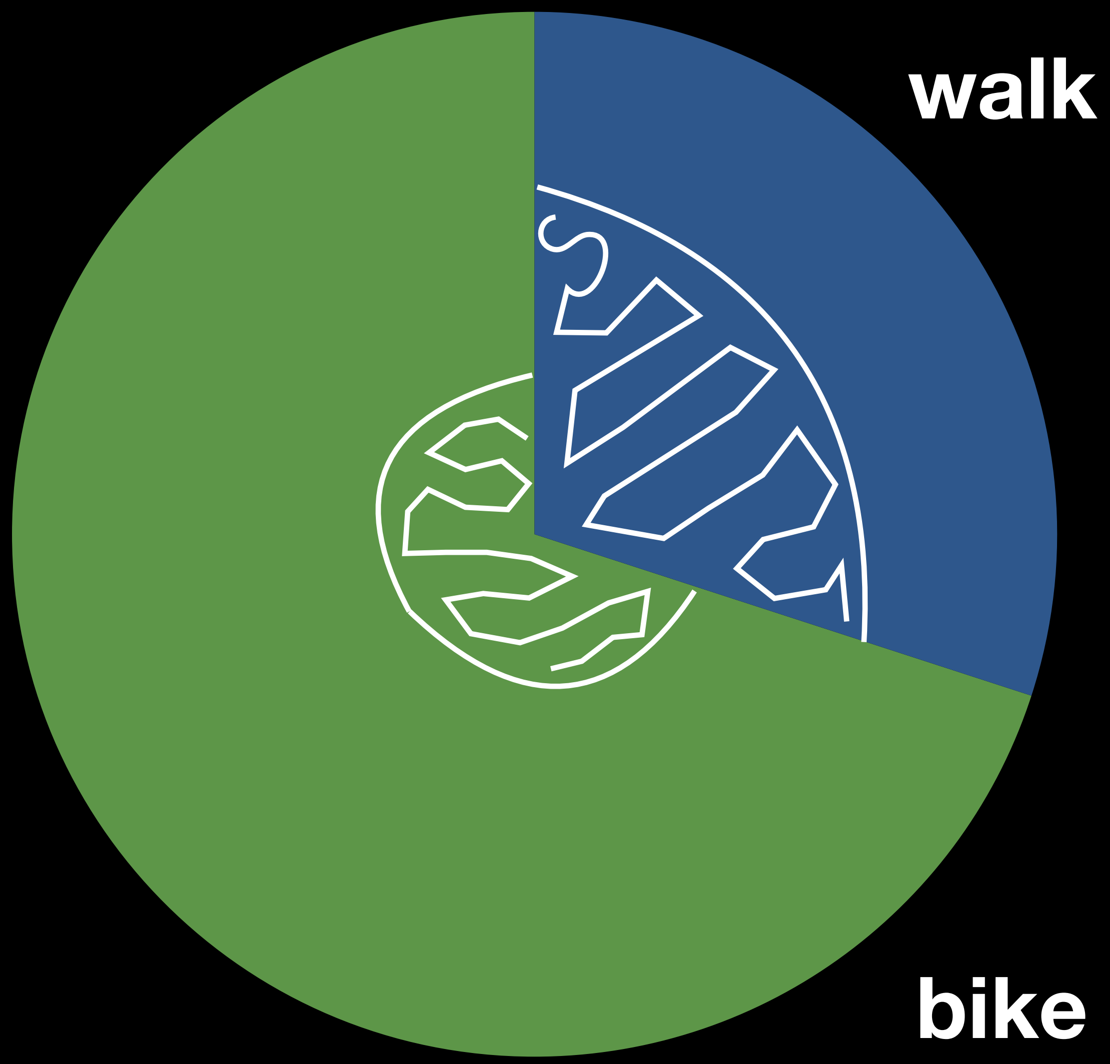


Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?

what if we can bike, walk, or
take the Marguerite (> 2 options)?



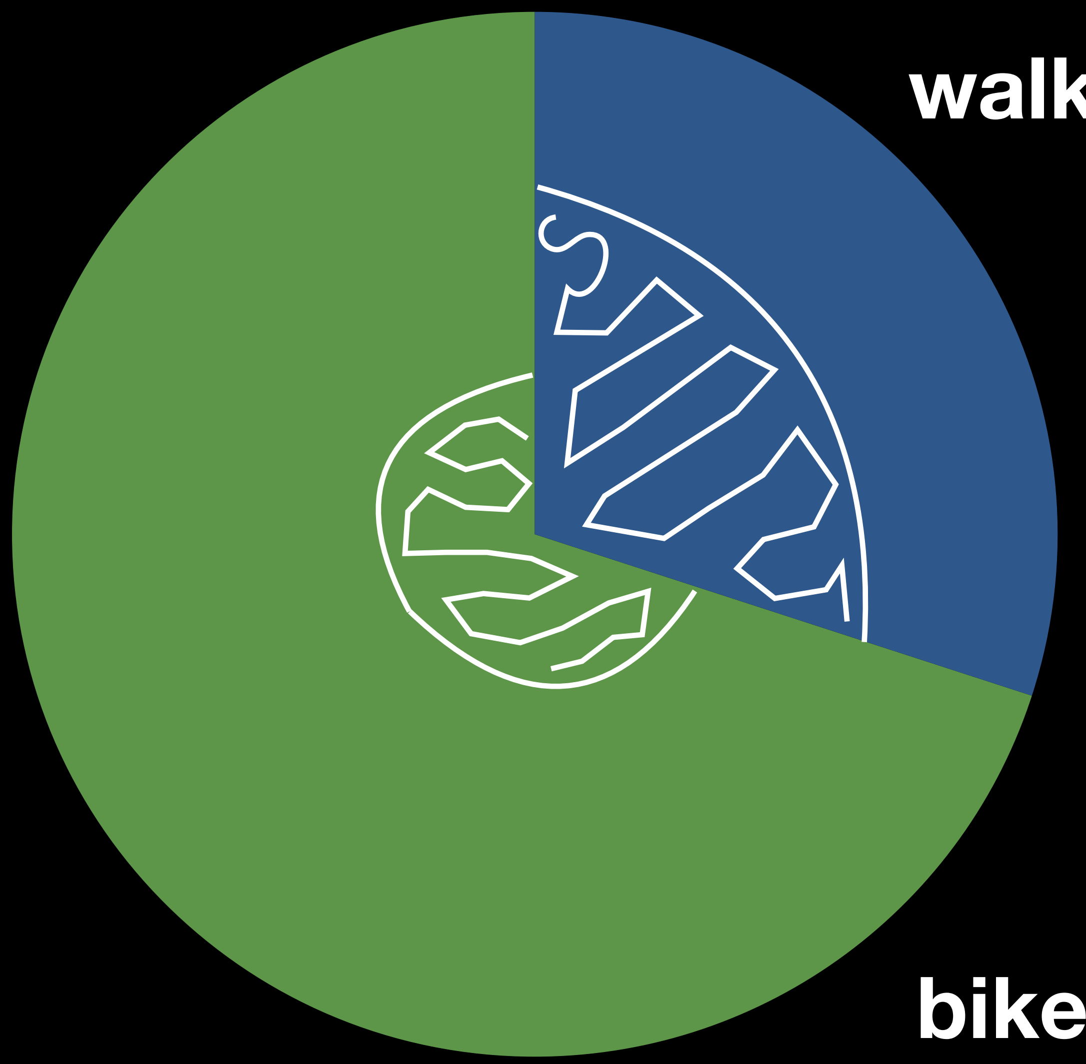
Law of Total Probability

$$P(A) = P(A \mid B)P(B) + P(A \mid B^C)P(B^C)$$

Event W = we walk to class. Event B = we bike = W^C.
Event L = we are late to class.
P(L | W) = 0.5, P(L | B) = 0.1.
P(W) = 0.3.
P(L) = ?

what if we can bike, walk, or
take the Marguerite (> 2 options)?

events for “scale factors” must be:
- **mutually exclusive**, and
- **exhaustive**



Bayes' Rule

$$P(E \mid F) = \frac{P(F \mid E)P(E)}{P(F)}$$

Bayes' Rule

posterior — likelihood — prior

$$P(E \mid F) = \frac{P(F \mid E)P(E)}{P(F)}$$

normalization constant

Bayes' Rule

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

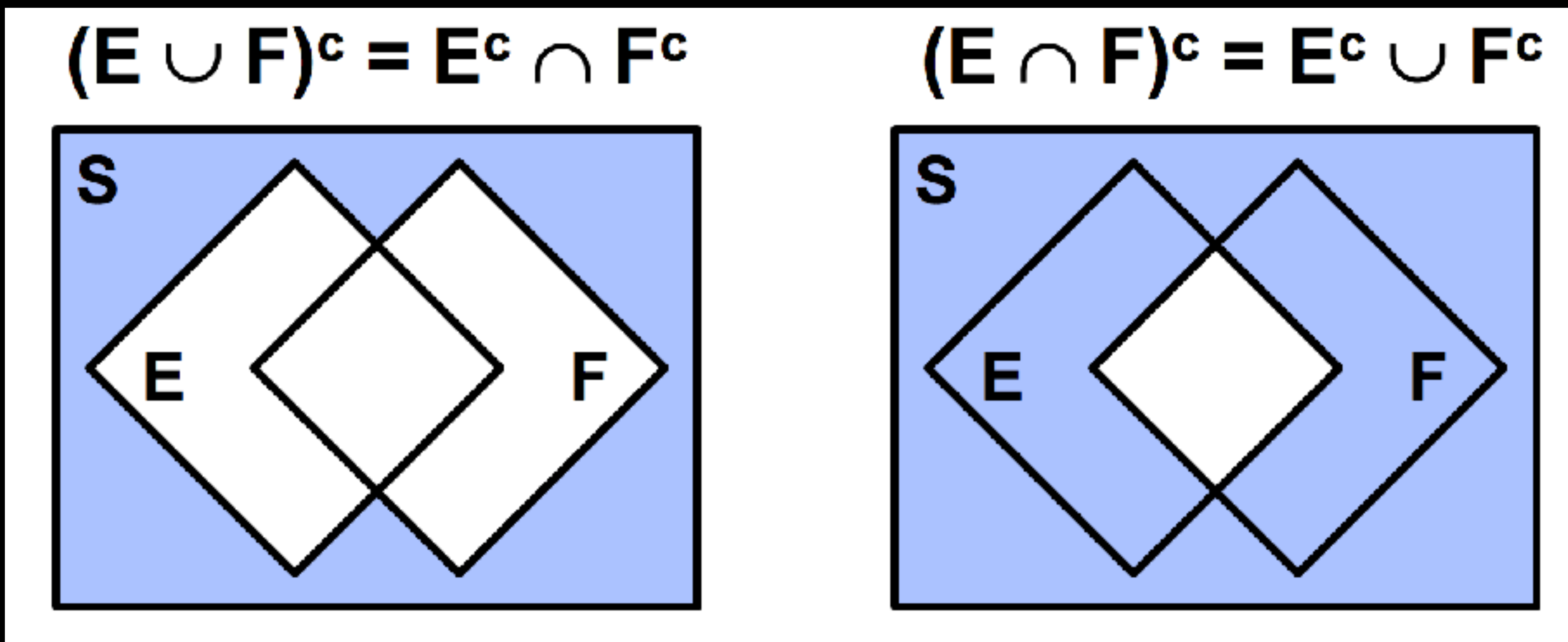
$$P(F|E)P(E) + P(F|E^C)P(E^C)$$

divide the event F into all the possible ways it can happen; use LoTP

Old Principles, New Tricks

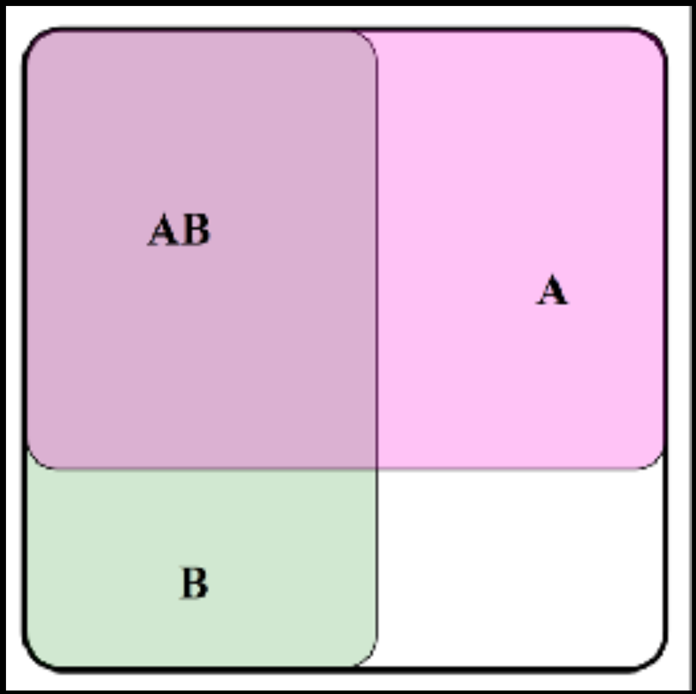
Name of Rule	**Original Rule**	**Conditional Rule**
First axiom of probability	$0 \leq P(E) \leq 1$	$0 \leq P(E \mid G) \leq 1$
Complement Rule	$P(E) = 1 - P(E^C)$	$P(E \mid G) = 1 - P(E^C \mid G)$
Chain Rule	$P(EF) = P(E \mid F)P(F)$	$P(EF \mid G) = P(E \mid FG)P(F \mid G)$
Bayes Theorem	$P(E \mid F) = \frac{P(F\|E)P(E)}{P(F)}$	$P(E \mid FG) = \frac{P(F\|EG)P(E\|G)}{P(F\|G)}$

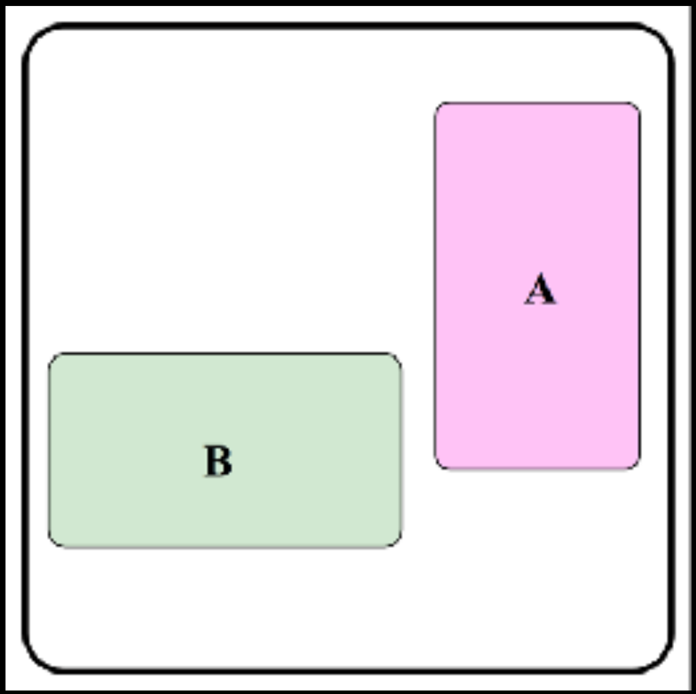
DeMorgan’s Laws



Independence

Independence	Mutual Exclusion
$P(EF) = P(E)P(F)$	$\lvert E \cap F \rvert = 0$
"AND"	**"OR"**





Independence

Independence	Conditional Independence
$P(EF) = P(E)P(F)$	$P(EF \mid G) = P(E \mid G)P(F \mid G)$ $P(E \mid FG) = P(E \mid G)$
"AND"	**"AND [if]"**

If E and F are independent…..

…..that does not mean they'll be independent if another event happens!

& vice versa

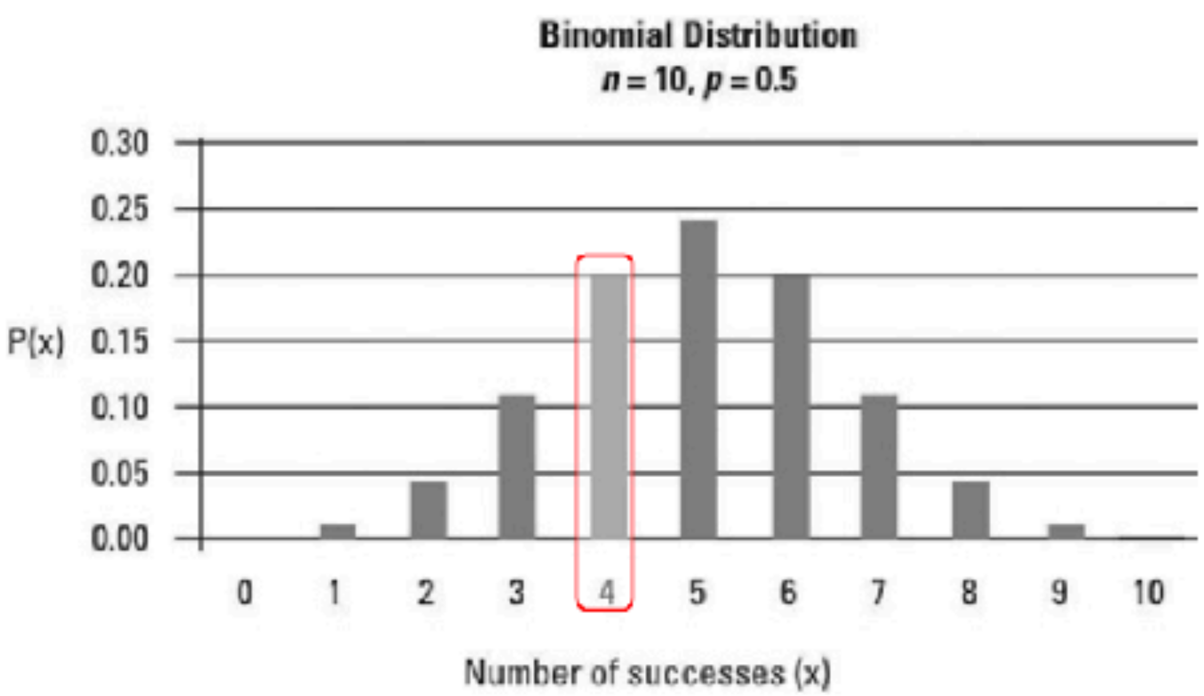
Outline

- **General Strategies**
- **Counting and Events**
- **Probability Rules**
- **Random Variables**

Probability Distributions

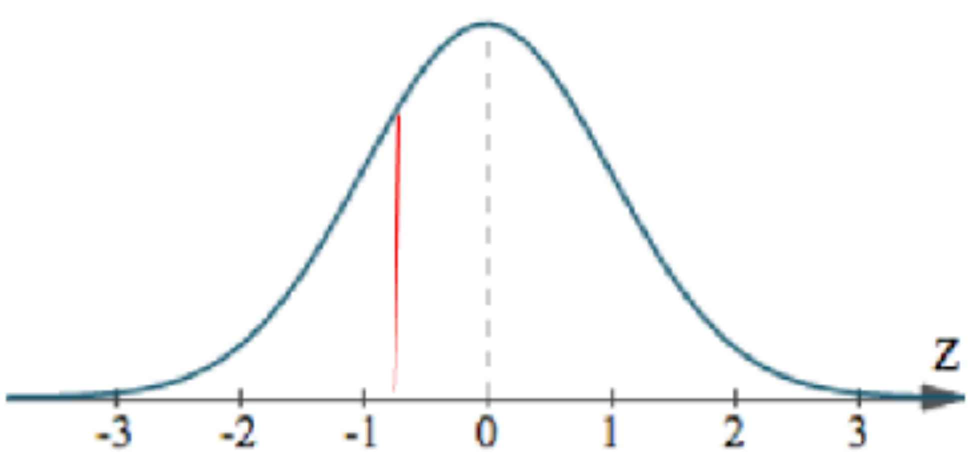
Discrete

PMF:



Continuous

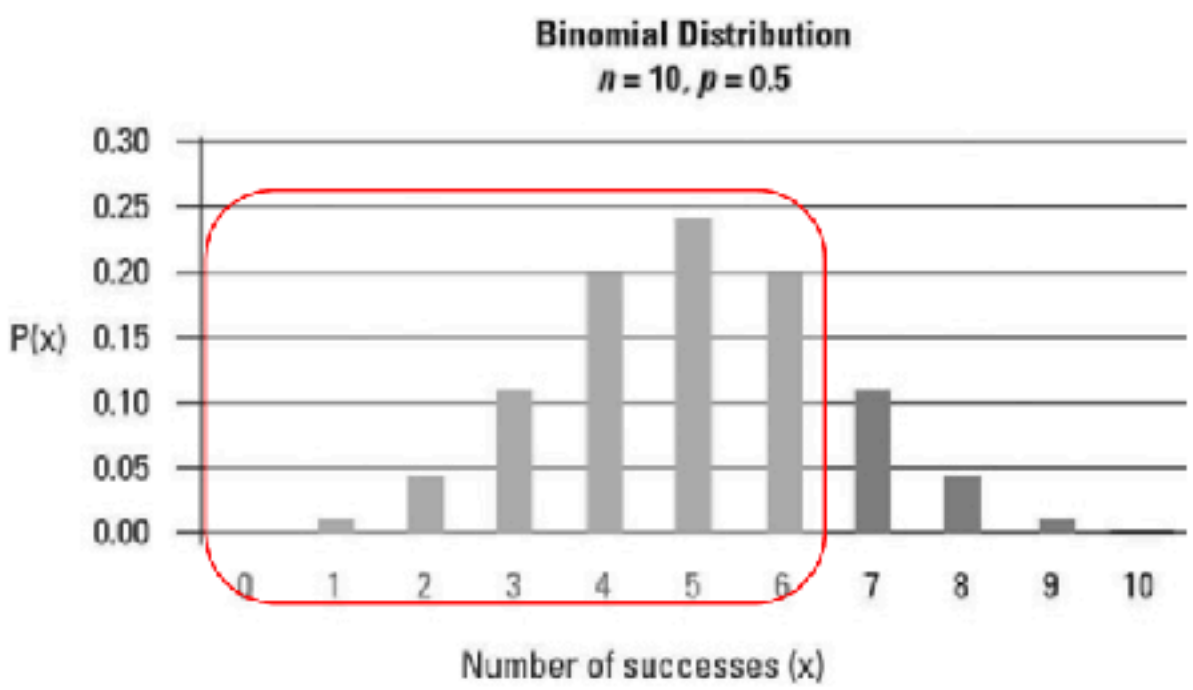
PDF:



Probability Distributions

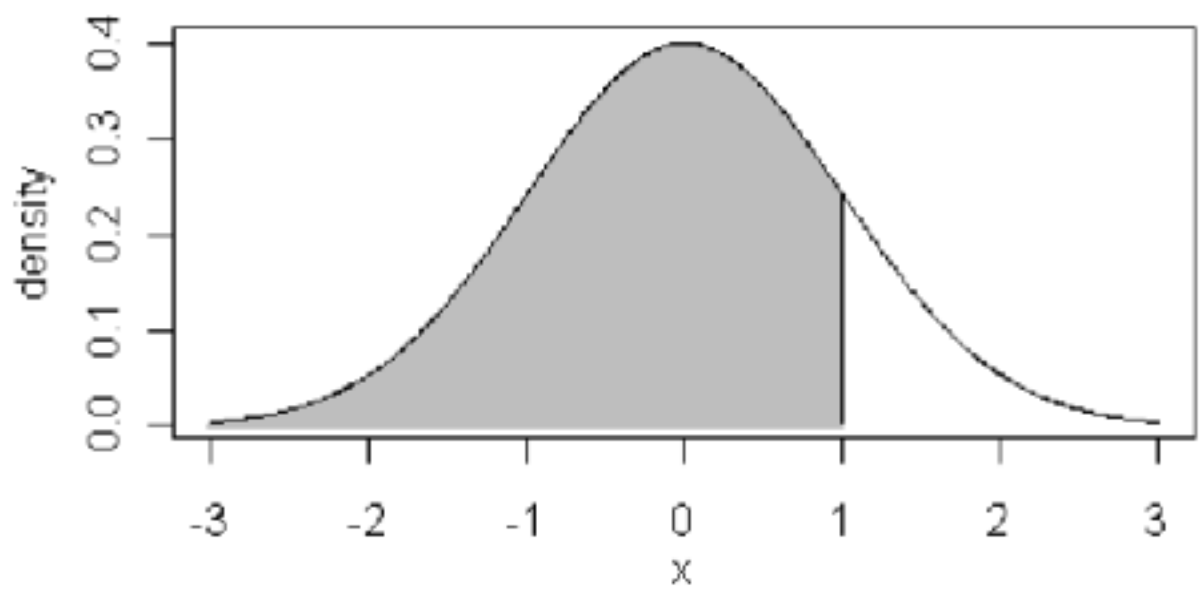
Discrete

CDF:



Continuous

CDF:



Expectation & Variance

Discrete definition

$$E[X] = \sum_{x:P(x)>0} x * P(x)$$

Continuous definition

$$E[X] = \int_x x * p(x)dx$$

Expectation & Variance

Discrete definition

$$E[X] = \sum_{x:P(x)>0} x * P(x)$$

Continuous definition

$$E[X] = \int_x x * p(x)dx$$

Properties of Expectation

$$E[X + Y] = E[X] + E[Y]$$

$$E[aX + b] = aE[X] + b$$

$$E[g(X)] = \sum_x g(x) * p_X(x)$$

Properties of Variance

$$Var(X) = E[(X - \mu)^2]$$

$$Var(X) = E[X^2] - E[X]^2$$

$$Var(aX + b) = a^2Var(X)$$

All our (discrete) friends

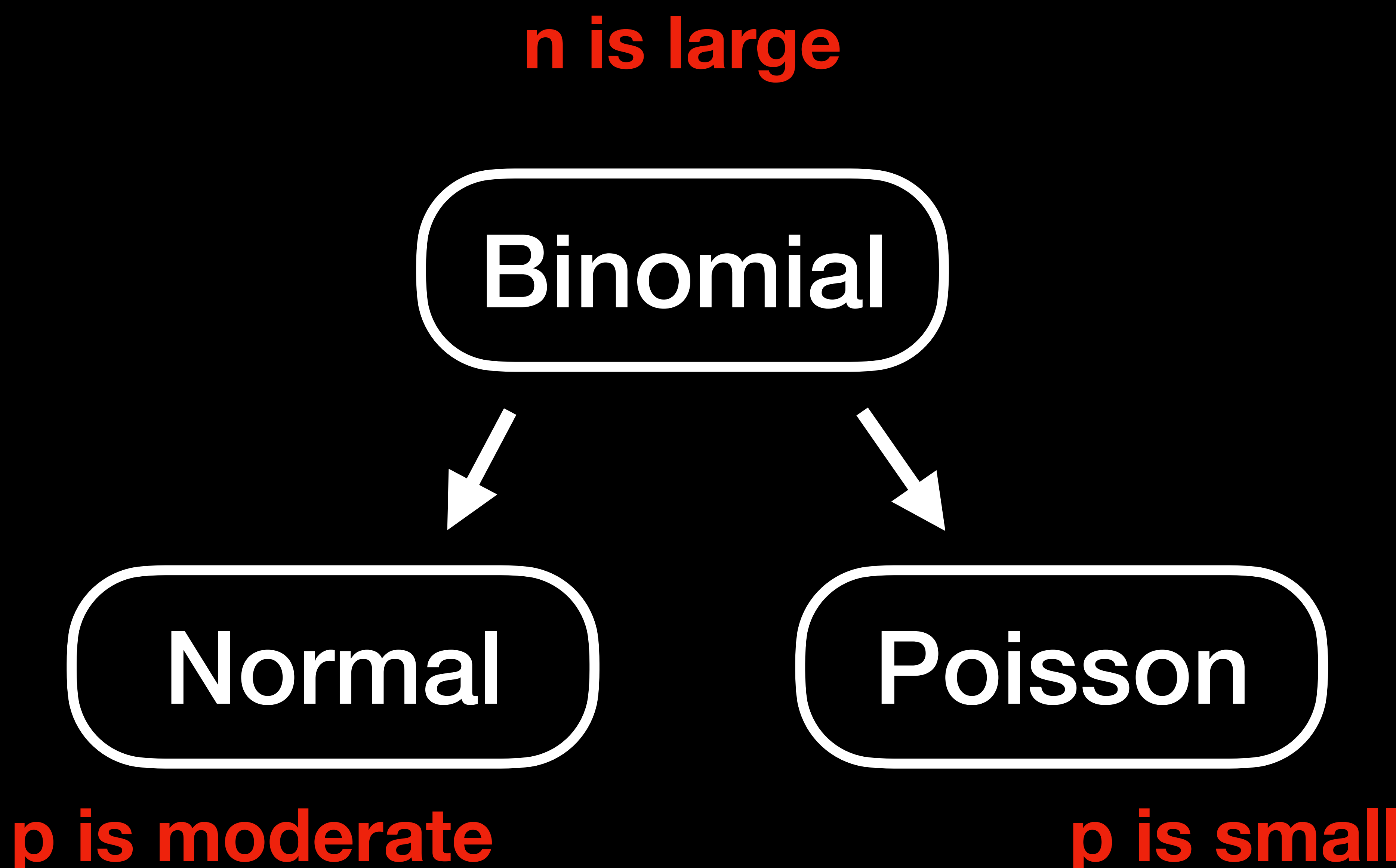
Ber(p)	Bin(n, p)	Poi(λ)	Geo(p)	NegBin (r, p)
P(X) = p	$\binom{n}{k} p^k (1-p)^{n-k}$	$\frac{\lambda^k e^{-\lambda}}{k!}$	$(1-p)^{k-1} p$	$\binom{k-1}{r-1} p^r (1-p)^{k-r}$
E[X] = p	E[X] = np	E[X] = λ	E[X] = 1 / p	E[X] = r / p
Var(X) = p(1-p)	Var(X) = np(1-p)	Var(X) = λ	$\frac{1-p}{p^2}$	$\frac{r(1-p)}{p^2}$
Getting candy or not at a random house	# houses out of 20 that give out candy	# houses in an hour that give out candy	# houses to visit before getting candy	# houses to visit before getting candy 3 times

All our (continuous) friends

Uni(α, β)	Exp(λ)	N(μ, σ)
$f(x) = \frac{1}{\beta - \alpha}$	$f(x) = \lambda e^{-\lambda x}$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$
$P(a \leq X \leq b) = \frac{b-a}{\beta - \alpha}$	$F(x) = 1 - e^{-\lambda x}$	$F(x) = \Phi(\frac{x-\mu}{\sigma})$
$E(x) = \frac{\alpha + \beta}{2}$	E[x] = 1 / λ	E[x] = μ
$Var(x) = \frac{(\beta - \alpha)^2}{12}$	$Var(x) = \frac{1}{\lambda^2}$	$Var(x) = \sigma^2$
thickness of sidewalk pavement between houses	time until feet get too sore to trick or treat	weight of filled candy baskets

Approximations

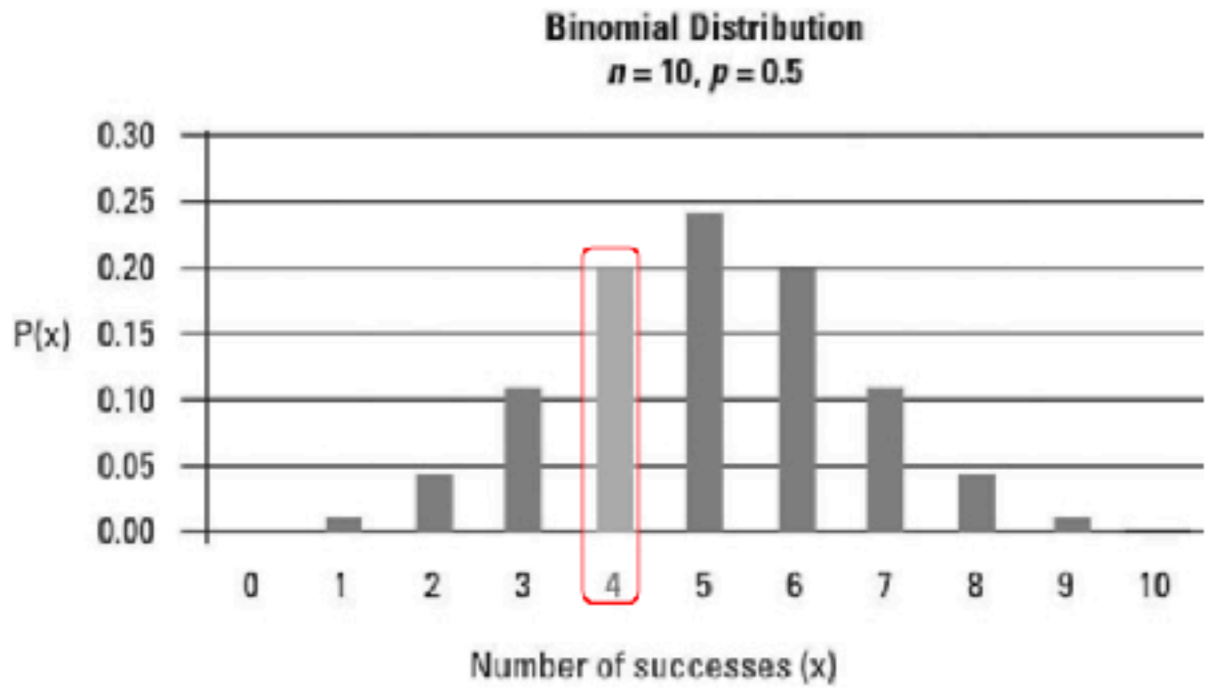
When can we approximate a binomial?



Continuity correction

Discrete

PMF:



Continuous

PDF:

- Only applies to PDF - why?

Joint Distributions

- Discrete case: $p_{x,y}(a,b) = P(X = a, Y = b)\,.\,P_x(a) = \sum_y P_{x,y}(a,y)$
- Continuous case:

$$P(a_1 < x \leq a_2, b_1 < y \leq b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{X,Y}(x,y)dydx$$

$$f_X(a) = \int_{-\infty}^{\infty} f_{X,Y}(a,y)dy$$

- For joint distributions to be independent, both their joint probability density function must be factorable and the bounds of the variables must be separable.

Convolutions

$$X \sim Bin(n_1, p), Y \sim Bin(n_2, p) => X + Y \sim Bin(n_1 + n_2, p)$$

$$X \sim Poi(\lambda_1), Y \sim Poi(\lambda_2) => X + Y \sim Poi(\lambda_1 + \lambda_2)$$

$$X \sim \mathrm{N}(\mu_1, \sigma_1^2), Y \sim \mathrm{N}(\mu_2, \sigma_2^2) => X + Y \sim \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

$$f_{X+Y}(a) = \int_{y=-\infty}^{\infty} f_X(a - y) f_Y(y)\, dy$$

(general case)

Practice Problems

How many ways are there to rearrange the letters of the alphabet such that none of the 5 vowels are next to each other?

Think of the 5 vowels as dividers for buckets, and imagine at least one consonant must go in the middle four buckets. That means that there are $\binom{17+6-1}{6-1}$ arrangements of vowels and consonants (17 consonants into 6 buckets). Then there are 21! ways to arrange the consonants and 5! ways to arrange the vowels, so our final answer is

$$21!\,5!\binom{17+6-1}{6-1}$$

Assume SAT scores are normally distributed, with mean 500 and variance 1000. If two students take the exam, what is the probability that their combined score is greater than 1020?

Score of student 1~N(500,1000)

Score of student 2~N(500,1000)

Score of both students varies as $N(500{,}1000) + N(500{,}1000) = N(1000, 2000)$.

$$P(N(1000, 2000) > 1050) = 1 - P(N(1000{,}2000) < 1020.5))$$

$$P(N(1000{,}2000) < 1020.5)) = P(Z < \frac{1020.5 - 1000}{\sqrt{2000}})$$

$$= P(Z < .458) = .677$$

So our final probability is 1-.677 or .323

Are Hogwarts house and favorite pet independent?

	Dog	Cat	Fish
Gryffindor	**.12**	**.12**	**.06**
Slytherin	.04	.04	.02
Ravenclaw	.16	.16	.08
Hufflepuff	.8	.2	.04