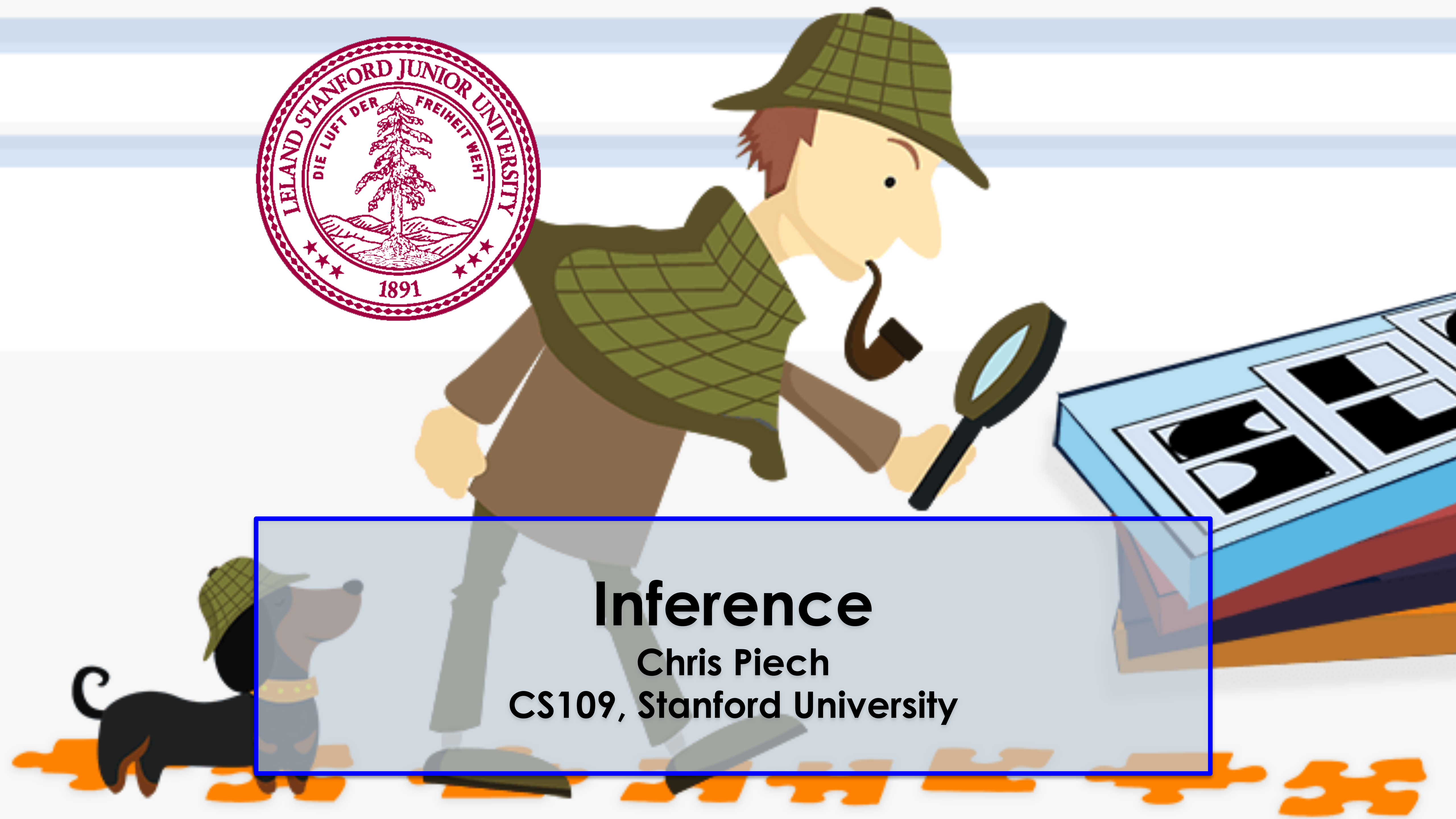


Inference

Chris Piech
CS109, Stanford University

Announcements

- Midterm next Tuesday.
  - 7-9p in Hewlett 200. Oct 26th. OAE? Conflict? Email by 9a Th
  - Covers material on PSet 1- 3 and Sections through this week.
  - Practice released starting tonight.
  - Closed computer. 10 pages of notes.
  - No class the day before the midterm!
  - Midterm review Thurs 7p.
  - Reader frozen (Part 1, Part 2, Part 3) by Thursday.

Where are we in CS109?

YOU ARE HERE

Overview of Topics

- Counting Theory
- Core Probability
- Random Variables
- Probabilistic Models
- Uncertainty Theory
- Machine Learning

Where are we locally?

Discrete Models: General Case, Multinomial

Continuous Models: General Case, Multi-Gauss

Inference Conclusions from Observations

Modelling: Make your own!

General Inference: Use computers to infer

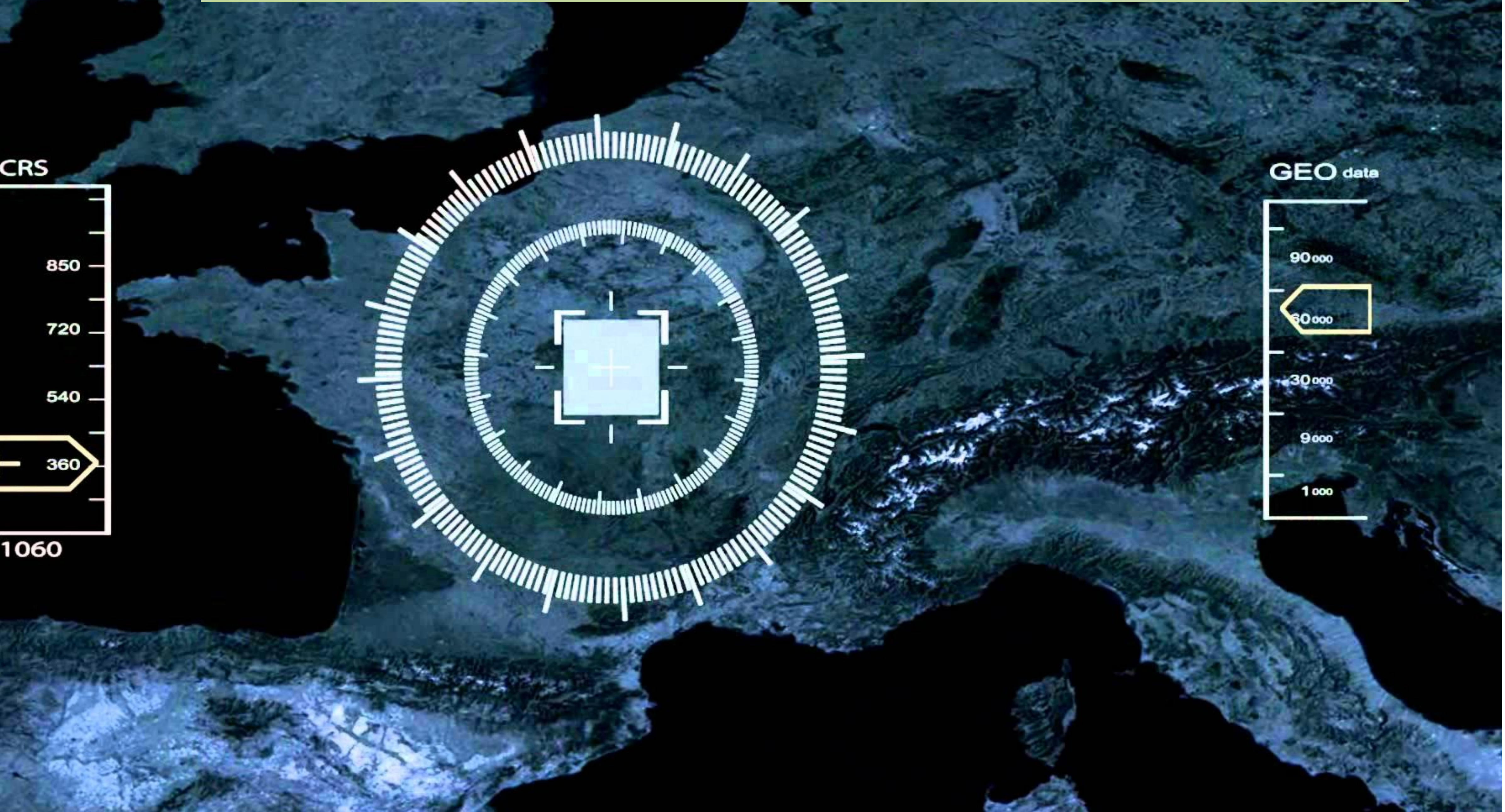
Today: Inference

Inference *noun*

An updated belief about a random variable (or multiple) based on conditional knowledge regarding another random variable (or multiple) in a probabilistic model.

TLDR: conditional probability with random variables.

Tracking in 2D Space?



And It Applies to Books Too

P(Buy Book Y | Bought Book X)

Joint Random Variables

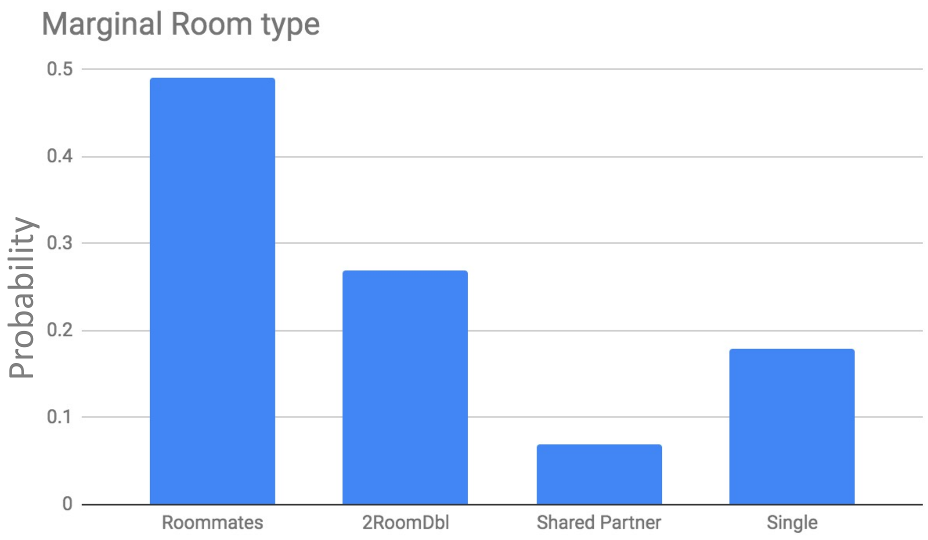
- [x] Use a joint table, density function or CDF to solve probability question
- [ ] Think about **conditional** probabilities with joint variables (which might be continuous)

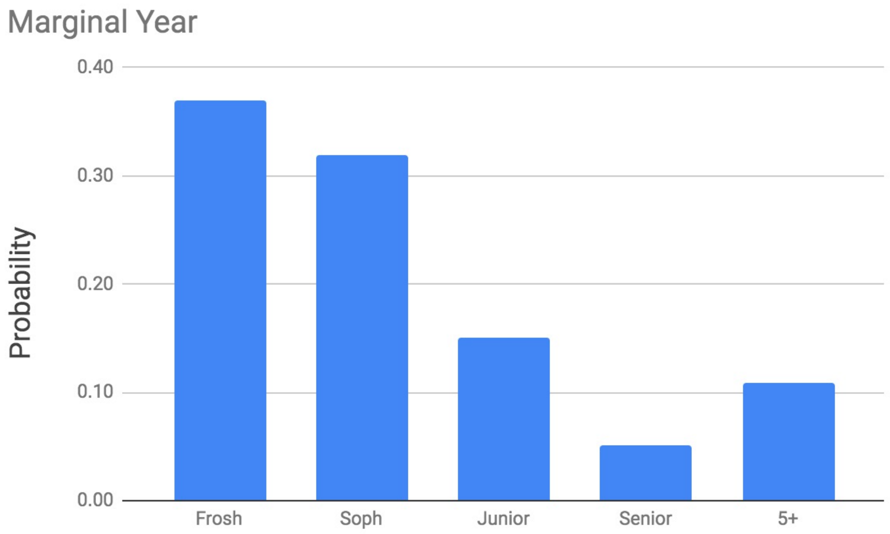
Joint Random Variables

- Use a joint table, density function or CDF to solve probability question
- Think about **conditional** probabilities with joint variables (which might be continuous)

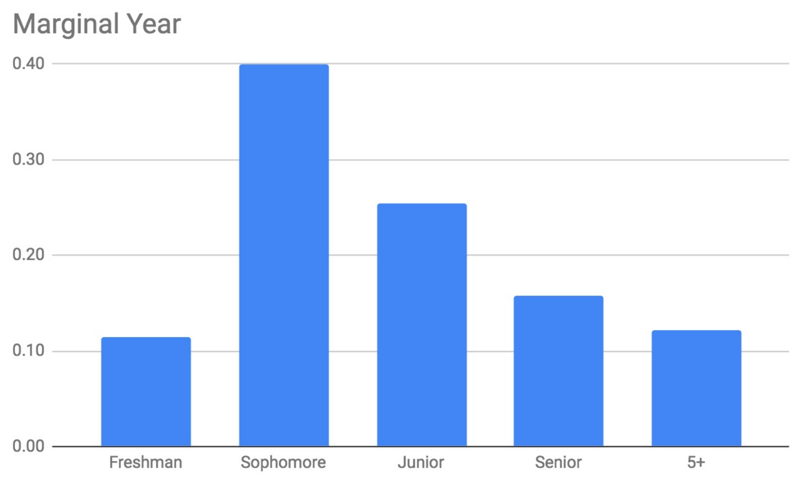
Joint Probability Table

	Roommates	2RoomDbl	Shared Partner	Single	
Frosh	0.30	0.07	0.00	0.00	0.37
Soph	0.12	0.18	0.00	0.03	0.32
Junior	0.04	0.01	0.00	0.10	0.15
Senior	0.01	0.02	0.02	0.01	0.05
5+	0.02	0.00	0.05	0.04	0.11
	0.49	0.27	0.07	0.18	1.00

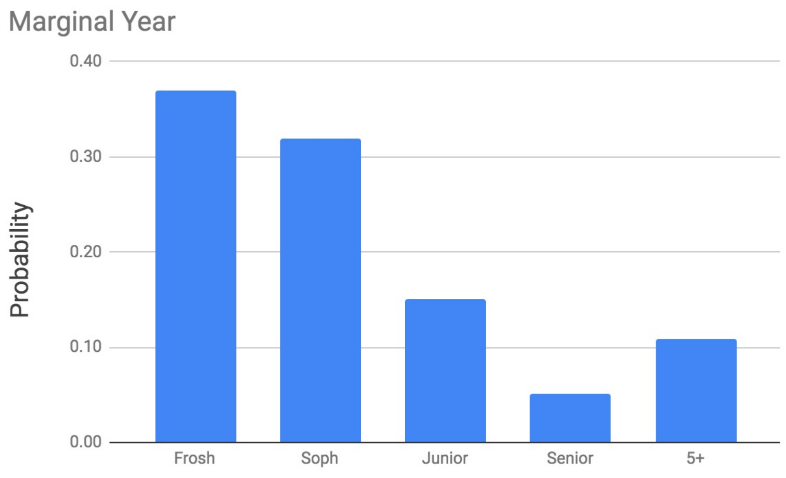




Change in Marginal Year

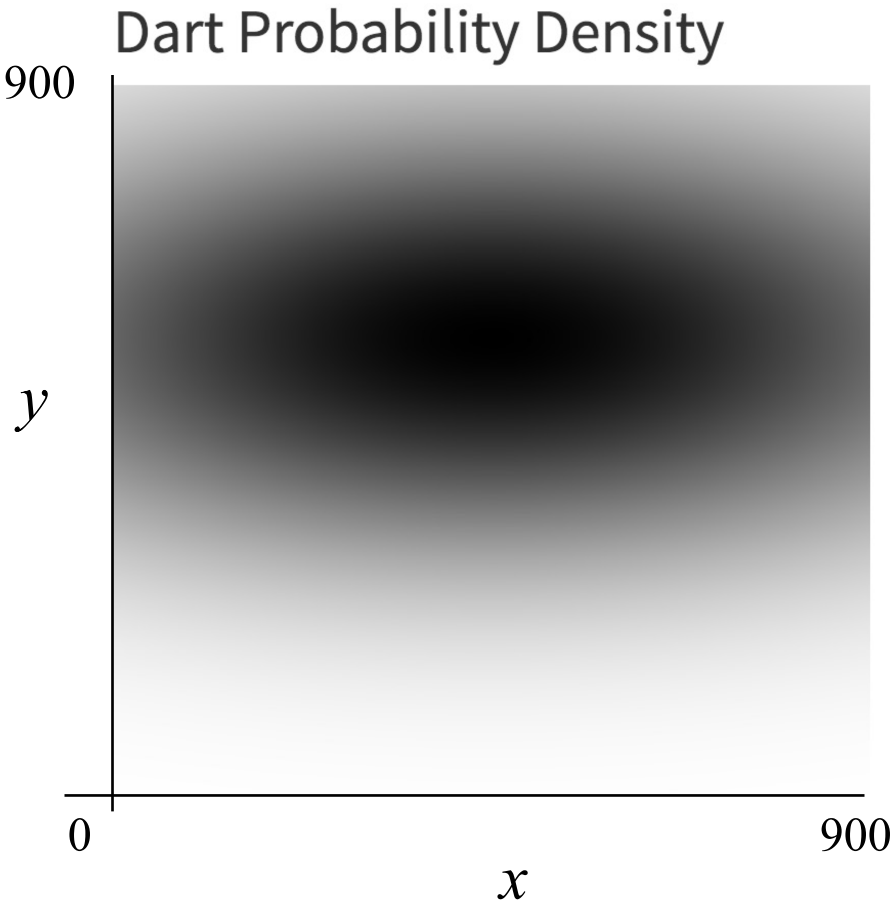


Fall quarter ‘18



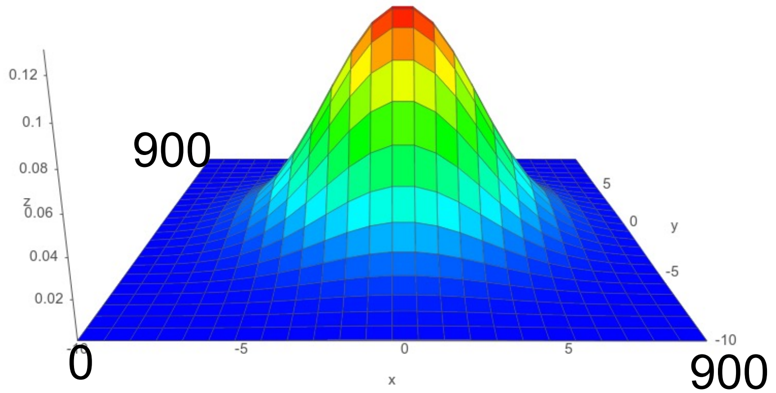
Spr quarter ‘19

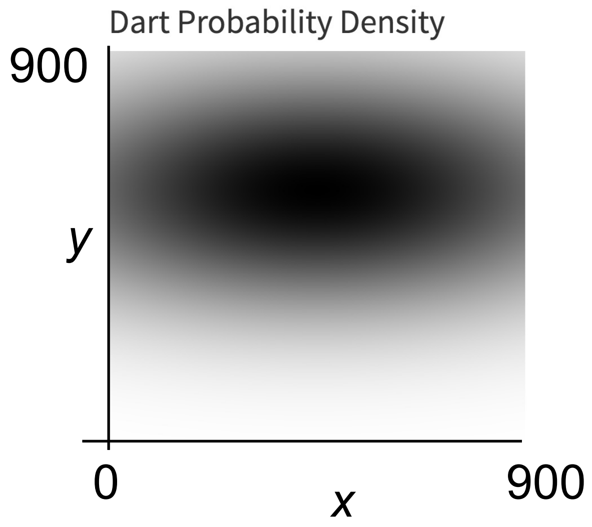
Continuous Joint Random Variables



Joint Probability Density Function

A **joint probability density function** gives the relative likelihood of **more than one** continuous random variable **each** taking on a specific value.





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

Boolean Operation on Variable = Event

Recall: any boolean question about a random variable makes for an event. For example:

$$P(X \leq 5)$$

$$P(Y = 6)$$

$$P(5 \leq Z \leq 10)$$

End Review

Conditionals with multiple variables

Aka Inference

Discrete Conditional Distributions

- Recall that for *events* E and F:

$$P(E \mid F) = \frac{P(EF)}{P(F)} \quad \text{where} \quad P(F) > 0$$

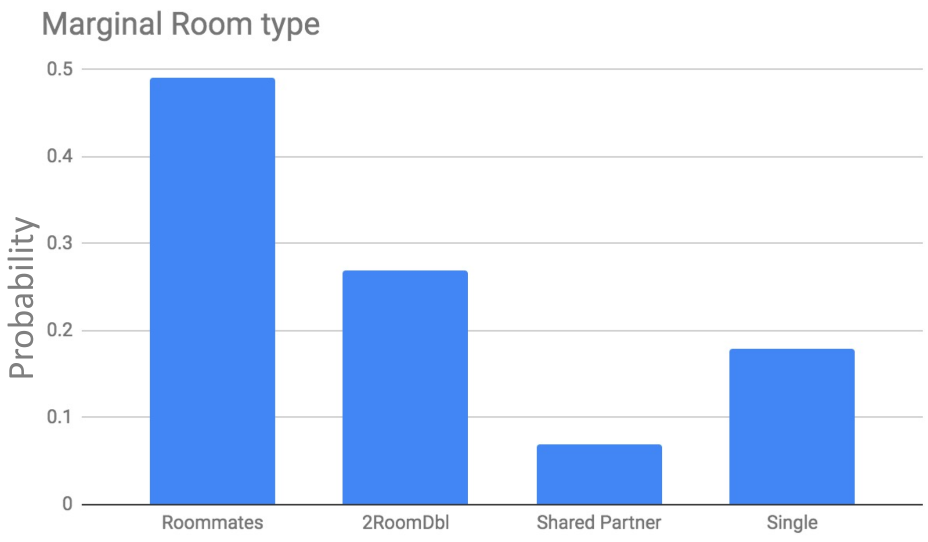
- Now, have X and Y as **discrete** random variables
  - **Conditional PMF** of X given Y:

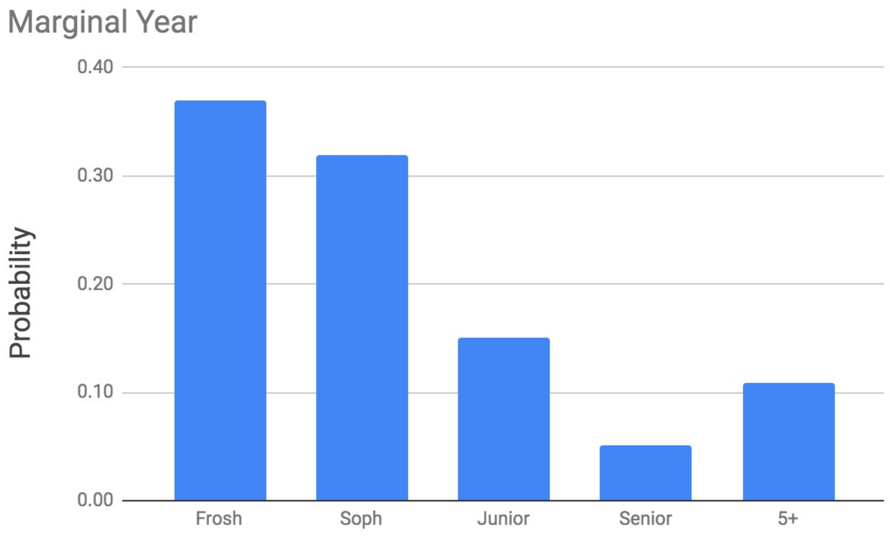
$$P_{X|Y}(x \mid y) = P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

Different notations, same idea.

Joint Probability Table

	Roommates	2RoomDbl	Shared Partner	Single	
Frosh	0.30	0.07	0.00	0.00	0.37
Soph	0.12	0.18	0.00	0.03	0.32
Junior	0.04	0.01	0.00	0.10	0.15
Senior	0.01	0.02	0.02	0.01	0.05
5+	0.02	0.00	0.05	0.04	0.11
	0.49	0.27	0.07	0.18	1.00

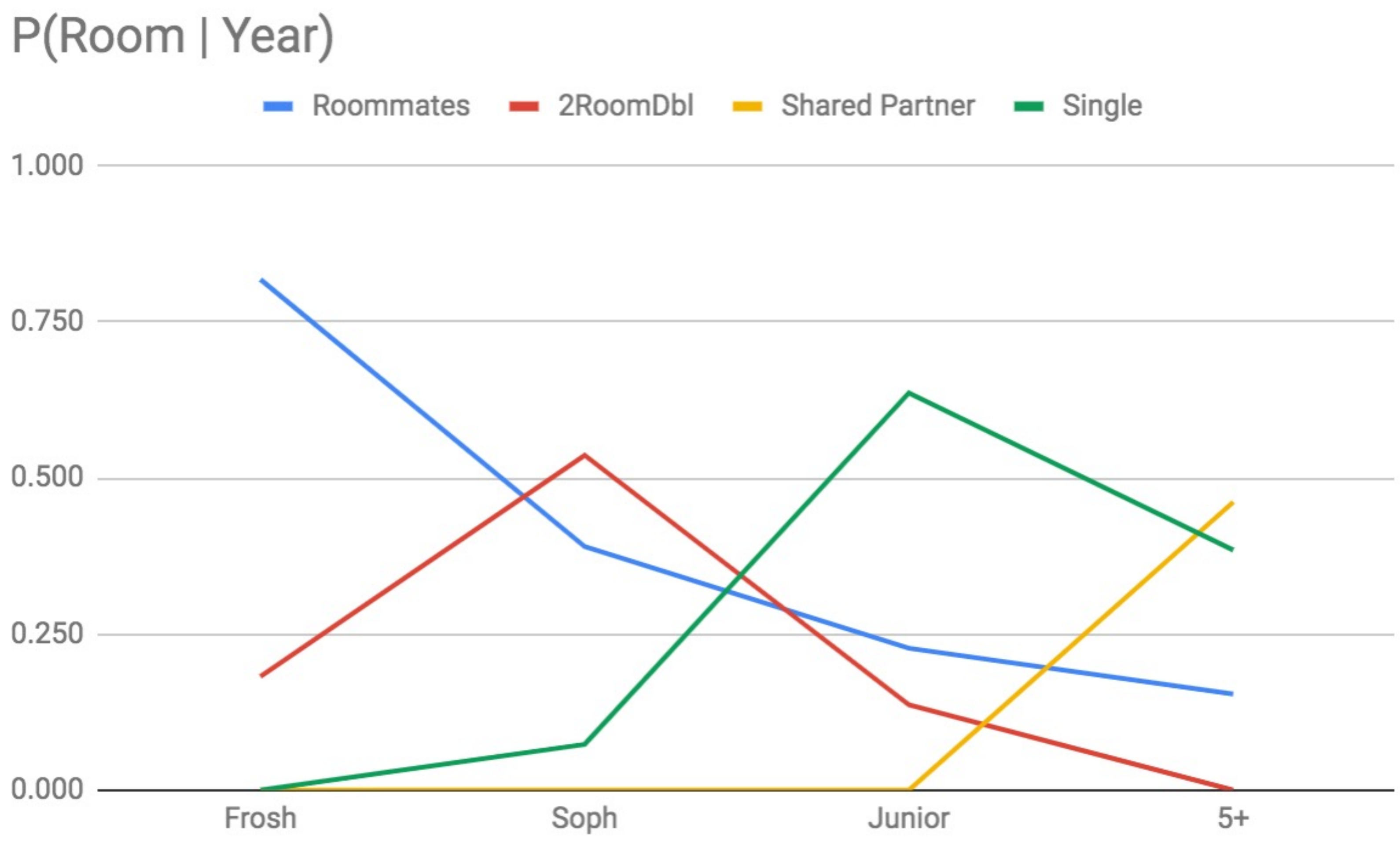




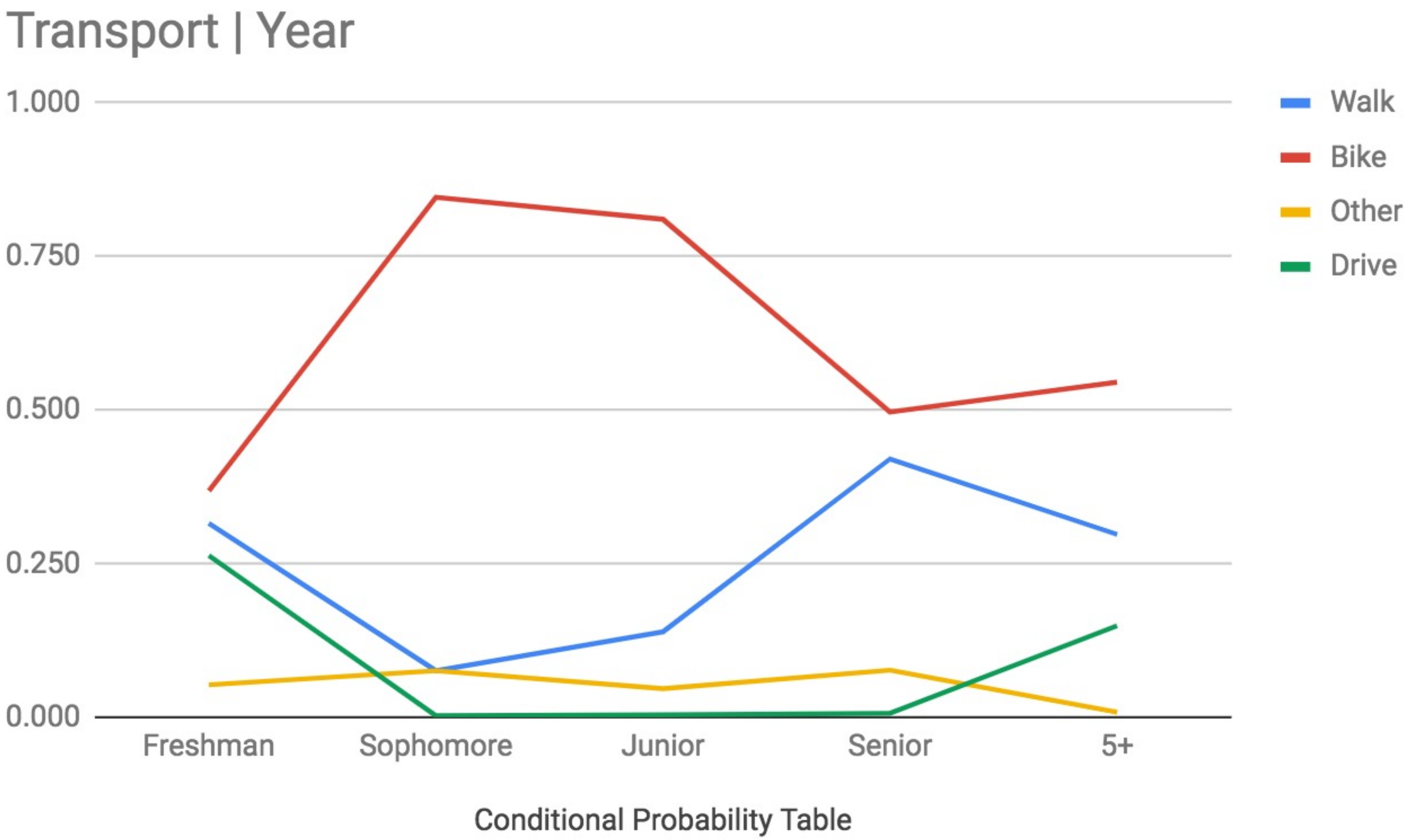
Warmup Inference

Q: What is the probability that someone has a single, given that they are a senior?

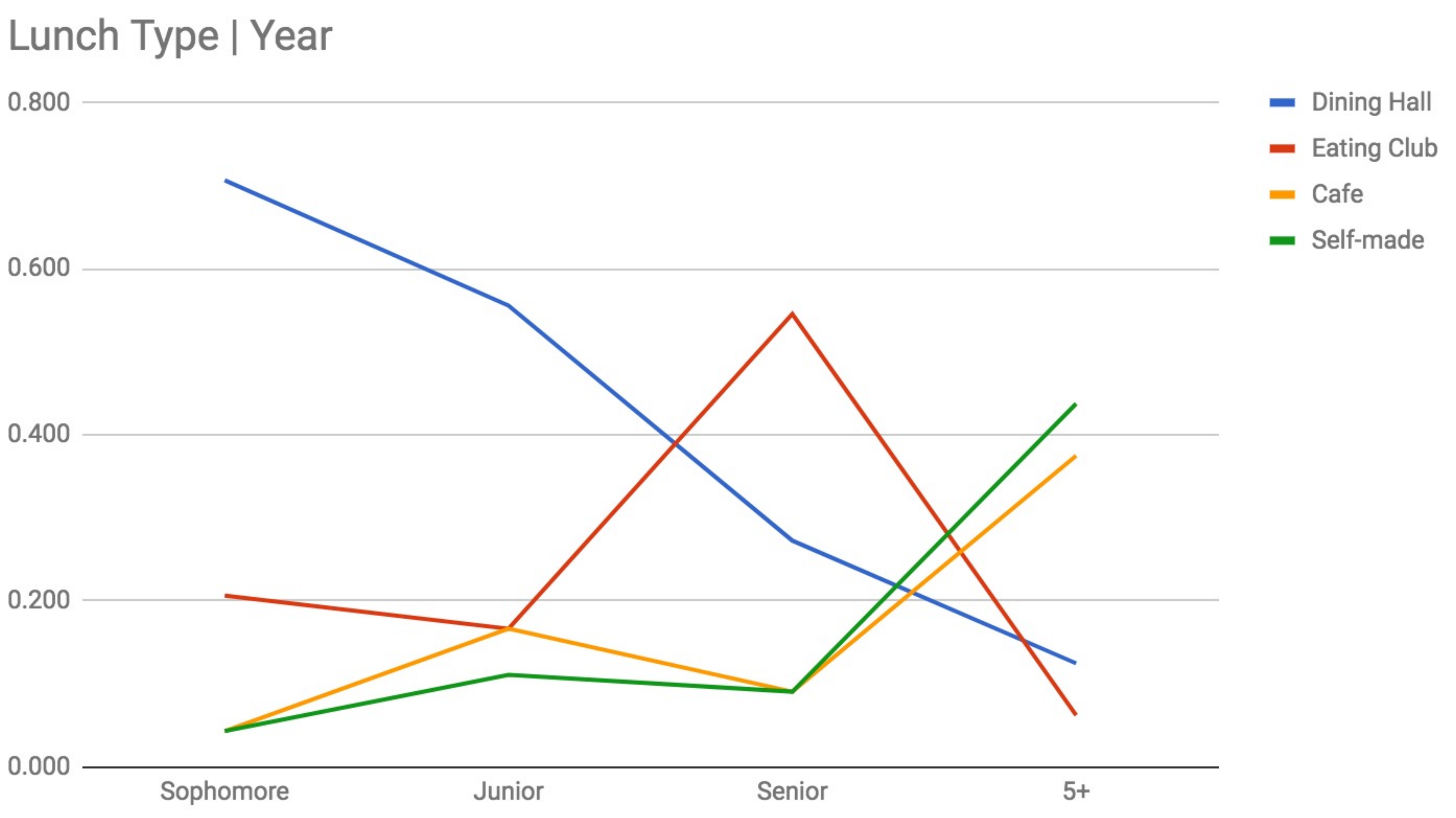
Room | Year


P(Room | Year)
Roommates
2RoomDbl
Shared Partner
Single
1.000
0.750
0.500
0.250
0.000
Frosh
Soph
Junior
5+

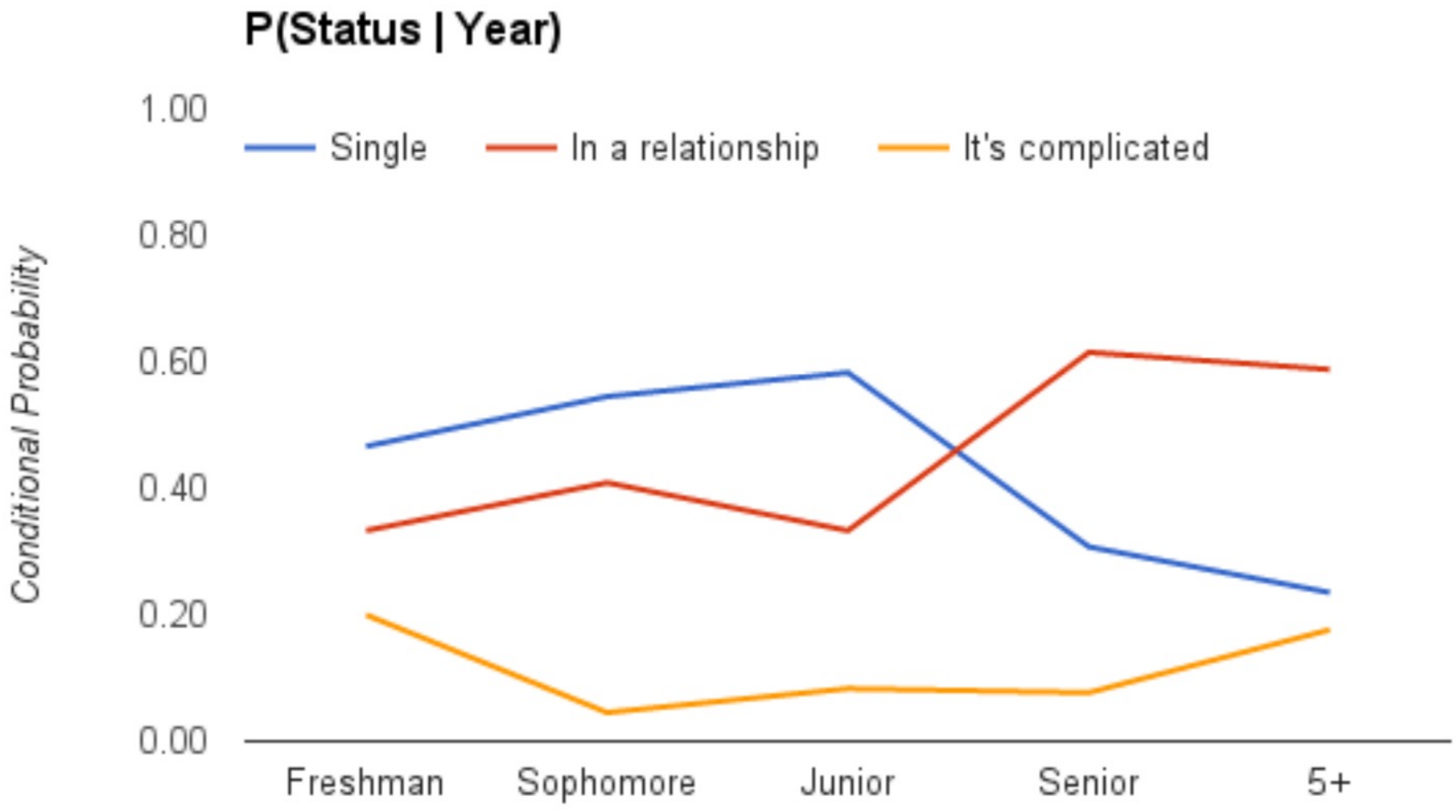
Transport | Year



Lunch | Year


Lunch Type | Year
0.800
0.600
0.400
0.200
0.000
Sophomore
Junior
Senior
5+
Dining Hall
Eating Club
Cafe
Self-made

Relationship Status | Year



Number or Function?

$$P(X = 2 | Y = 5)$$

Number

Number or Function?

$$P(X = x|Y = 2)$$

Random Variable

(also a function or 1D table)

Number or Function?

$$P(X = x|Y = y)$$

2D Function

(or 2D table)

Warmup: Bayes Revisited

Posterior belief

Likelihood of evidence

Prior belief

$$P(B|E) = \frac{P(E|B) \quad P(B)}{P(E)}$$

Normalization constant

Medium Inference

Q: At birth, girl elephant weights are distributed as a Gaussian with mean = 160kg, std = 7kg. At birth, boy elephant weights are distributed as a Gaussian with mean = 165kg, std = 3kg. All you know about a newborn elephant is that it is 163kg. What is the probability that it is a girl?

Medium Inference

Q: At birth, girl elephant weights are distributed as a Gaussian with mean = 160kg, std = 7kg. At birth, boy elephant weights are distributed as a Gaussian with mean = 165kg, std = 3kg. All you know about a newborn elephant is that it is 163kg. What is the probability that it is a girl?

Model:

Let G be an indicator that the elephant is a girl. G is Bern($p = 0.5$)

Let X be the distribution of weight of the elephant.

$X \mid G = 1$ is $N(\mu = 160, \sigma^2 = 7^2)$

$X \mid G = 0$ is $N(\mu = 165, \sigma^2 = 3^2)$

Medium Inference

Q: What is $P(G = 1 \mid X = 163)$

Let G be an indicator that the elephant is a girl. G is $\text{Bern}(p = 0.5)$

Let X be the distribution of weight of the elephant.

$X \mid G = 1$ is $N(\mu = 160, \sigma^2 = 7^2)$

$X \mid G = 0$ is $N(\mu = 165, \sigma^2 = 3^2)$

Mixing Discrete and Continuous

Let X be a continuous random variable

Let N be a discrete random variable

$$P(N = n|X = x) = \frac{P(X = x|N = n)P(N = n)}{P(X = x)}$$

$$P(N = n|X = x) = \frac{f(X = x|N = n) \cdot \epsilon \cdot P(N = n)}{f(X = x) \cdot \epsilon}$$

$$P(N = n|X = x) = \frac{f(X = x|N = n) \cdot P(N = n)}{f(X = x)}$$

Mixing Discrete and Continuous

Let X be a continuous random variable

Let N be a discrete random variable

$$P(X = x|N = n) = \frac{P(N = n|X = x)P(X = x)}{P(N = n)}$$

$$P_{X|N}(x|n) = \frac{P_{N|X}(n|x)P_X(x)}{P_N(n)}$$

$$f_{X|N}(x|n) \cdot \epsilon_x = \frac{P_{N|X}(n|x)f_X(x) \cdot \epsilon_x}{P_N(n)}$$

$$f_{X|N}(x|n) = \frac{P_{N|X}(n|x)f_X(x)}{P_N(n)}$$

All the Bayes Belong to Us

M,N are discrete. X, Y are continuous

OG Bayes

$$p_{M|N}(m|n) = \frac{P_{N|M}(n|m)p_M(m)}{p_N(n)}$$

Mix Bayes #1

$$f_{X|N}(x|n) = \frac{P_{N|X}(n|x)f_X(x)}{P_N(n)}$$

Mix Bayes #2

$$p_{N|X}(n|x) = \frac{f_{X|N}(x|n)p_N(n)}{f_X(x)}$$

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}$$

LOTP? Chain Rule? You can play too!

N is discrete. X is continuous

$$f(N = n, X = x) = f(X = x | N = n) P(N = n)$$

$$f(X = x) = \sum_{n} f(X = x | N = n) P(N = n)$$

Medium Inference

Q: What is $P(G = 1 \mid X = 163)$

Let G be an indicator that the elephant is a girl. G is $\text{Bern}(p = 0.5)$

Let X be the distribution of weight of the elephant.

$X \mid G = 1$ is $N(\mu = 160, \sigma^2 = 7^2)$

$X \mid G = 0$ is $N(\mu = 165, \sigma^2 = 3^2)$

Medium Inference

Q: What is P(G = 1 | X = 163)

Let G be an indicator that the elephant is a girl. G is Bern($p = 0.5$)

Let X be the distribution of weight of the elephant.

X | G = 1 is N($\mu = 160, \sigma^2 = 7^2$)

X | G = 0 is N($\mu = 165, \sigma^2 = 3^2$)

Joint Distribution is Implied:

$$f(G = g, X = x) = f(X = x | G = g) P(G = g)$$

Medium Inference

Q: What is $P(G = 1 \mid X = 163)$

Let G be an indicator that the elephant is a girl. G is $\text{Bern}(p = 0.5)$

Let X be the distribution of weight of the elephant.

$X \mid G = 1$ is $N(\mu = 160, \sigma^2 = 7^2)$

$X \mid G = 0$ is $N(\mu = 165, \sigma^2 = 3^2)$

Continuous Conditional Distributions

Let X and Y be continuous random variables

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

$$f_{X|Y}(x|y) \cdot \epsilon_x = \frac{f_{X,Y}(x, y) \cdot \epsilon_x \cdot \epsilon_y}{f_Y(y) \cdot \epsilon_y}$$

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

Warmup: Bayes Revisited

Posterior belief

Likelihood of evidence

Prior belief

$$P(B|E) = \frac{P(E|B) \quad P(B)}{P(E)}$$

Normalization constant

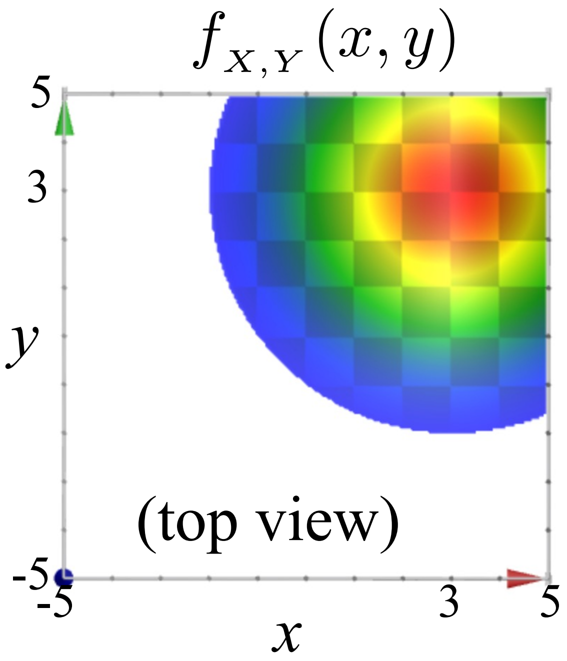
Warmup: Bivariate Normal

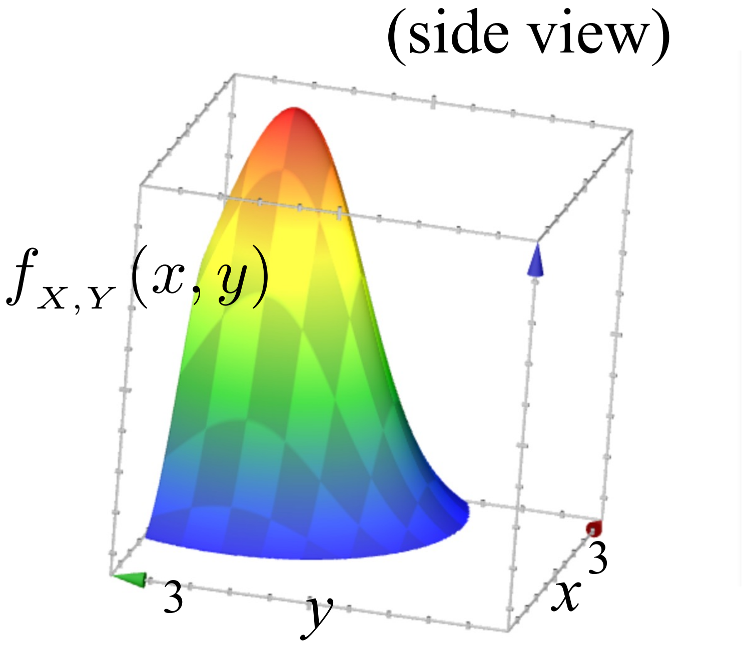
- X, Y follow a symmetric bivariate normal distribution if they have joint PDF:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{[(x-\mu_x)^2+(y-\mu_y)^2]}{2\cdot\sigma^2}}$$

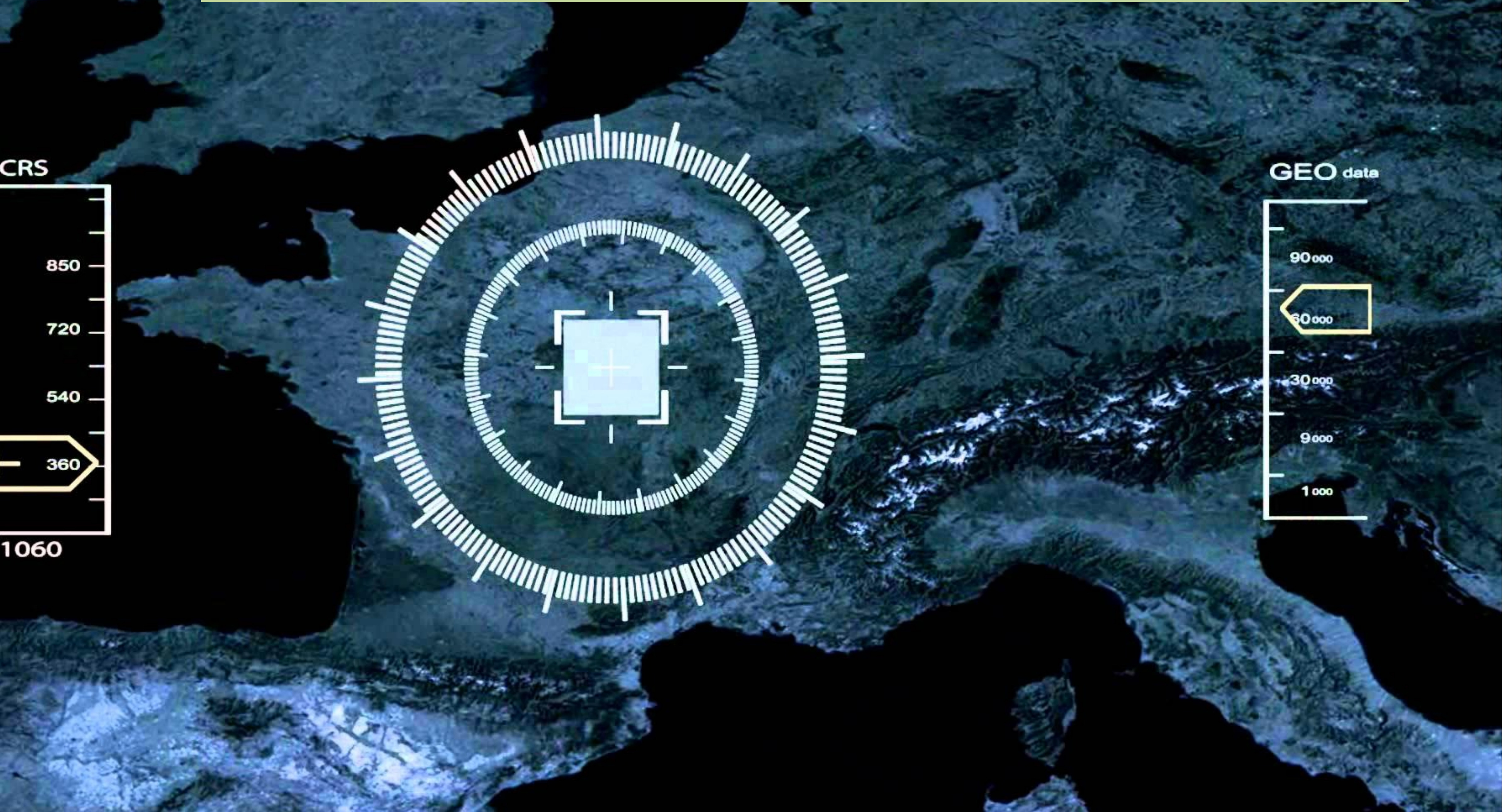
Here is an example where:

$$\mu_x = 3$$
$$\mu_y = 3$$
$$\sigma = 2$$



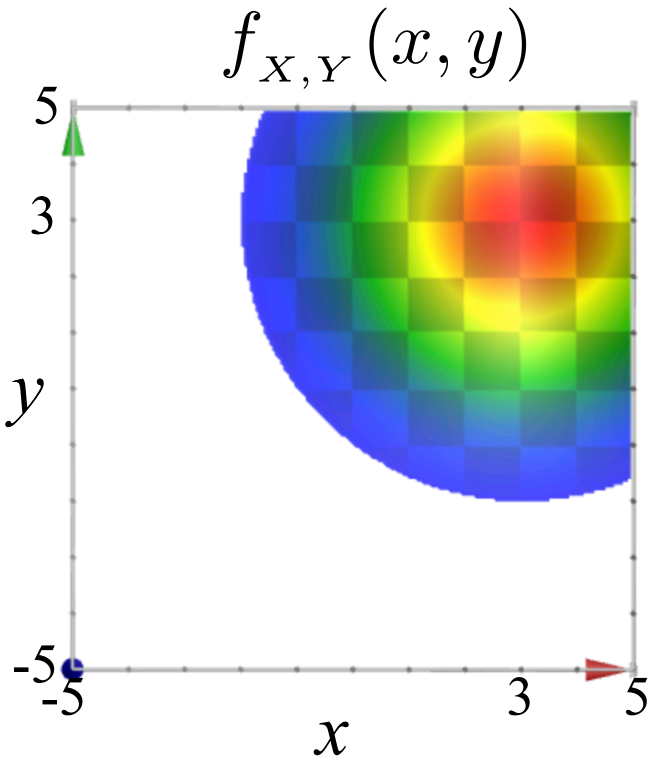


Tracking in 2D Space?



Tracking in 2D Space: Prior

Prior belief: $$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{[(x-\mu_x)^2+(y-\mu_y)^2]}{2\cdot\sigma^2}}$$



Satellite is at (0, 0)

$$\mu_x = 3$$

$$\mu_y = 3$$

$$\sigma = 2$$

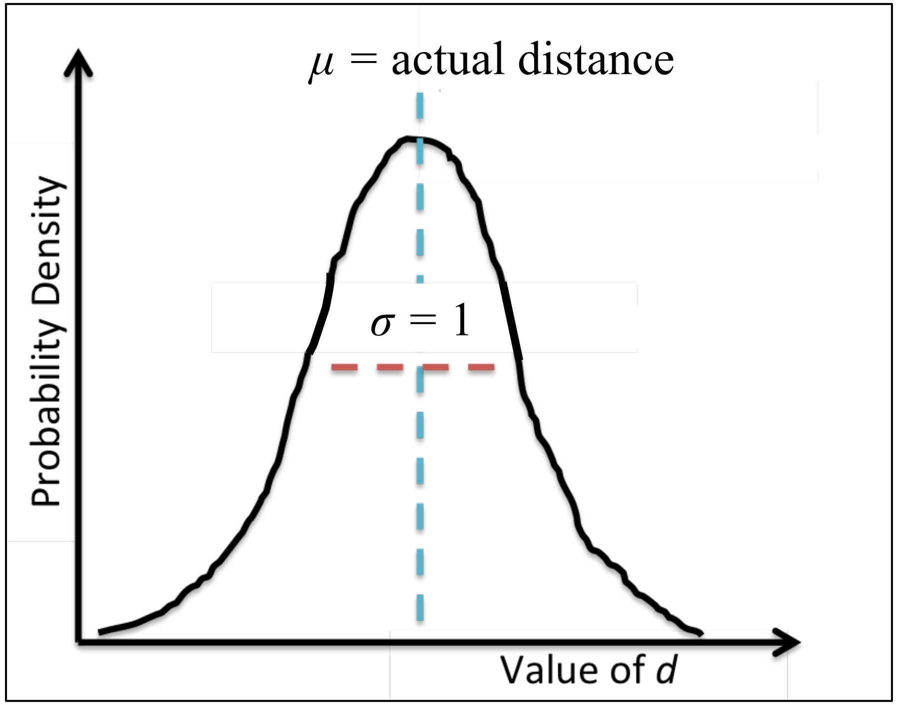
Prior belief with K: $$f_{X,Y}(x,y) = K \cdot e^{-\frac{[(x-3)^2+(y-3)^2]}{8}}$$

Tracking in 2D Space: Observation!

You now observe a noisy distance reading.
It says that your object is distance D away

We can say how likely that reading is if we know the actual location of the object…

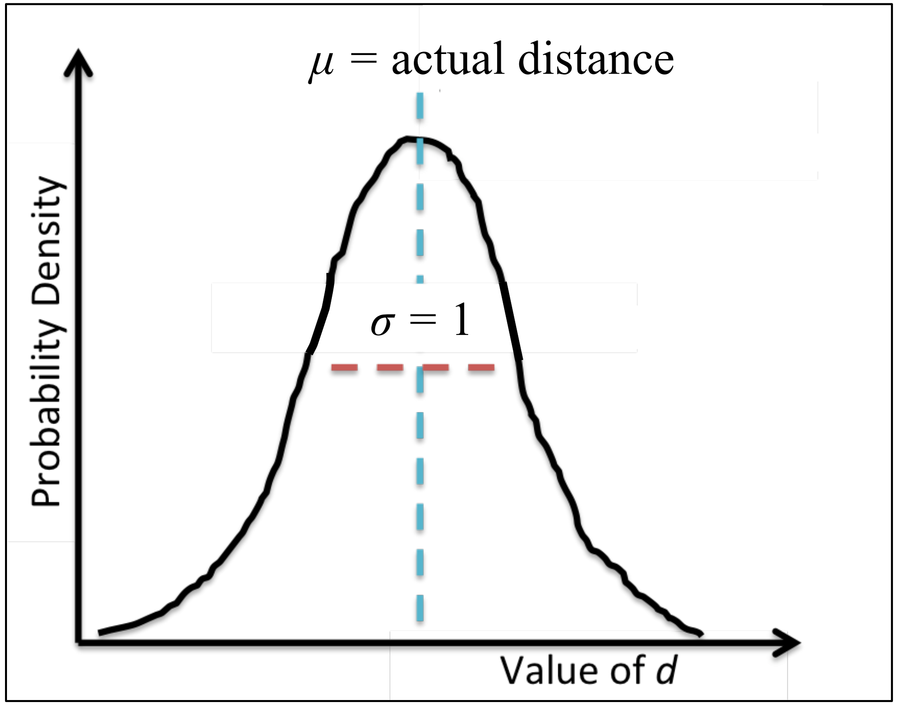
P(D | X, Y) is knowable!



Tracking in 2D Space: Observation!

Observe a ping of the object that is distance D away from satellite!

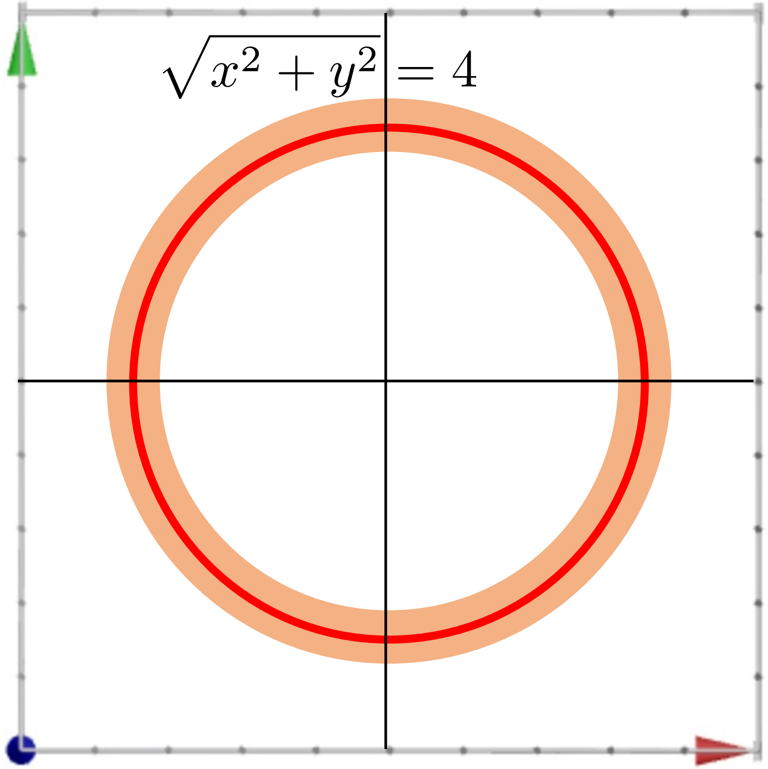
$$D|X,Y \sim N(\mu = \sqrt{x^2 + y^2}, \sigma^2 = 1)$$

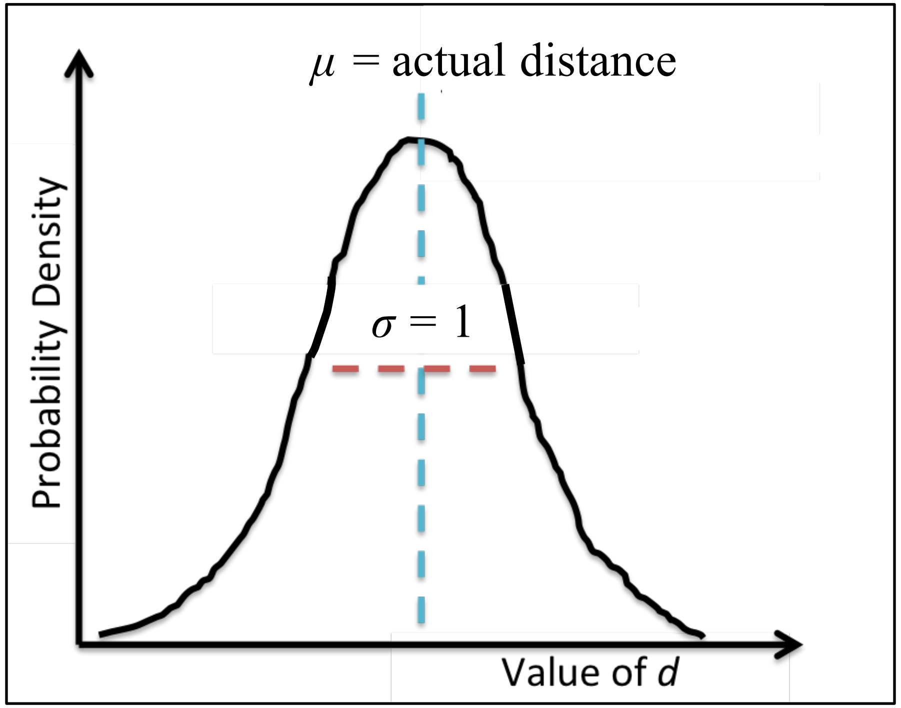


Know that the distance of a ping is normal with respect to the true distance.

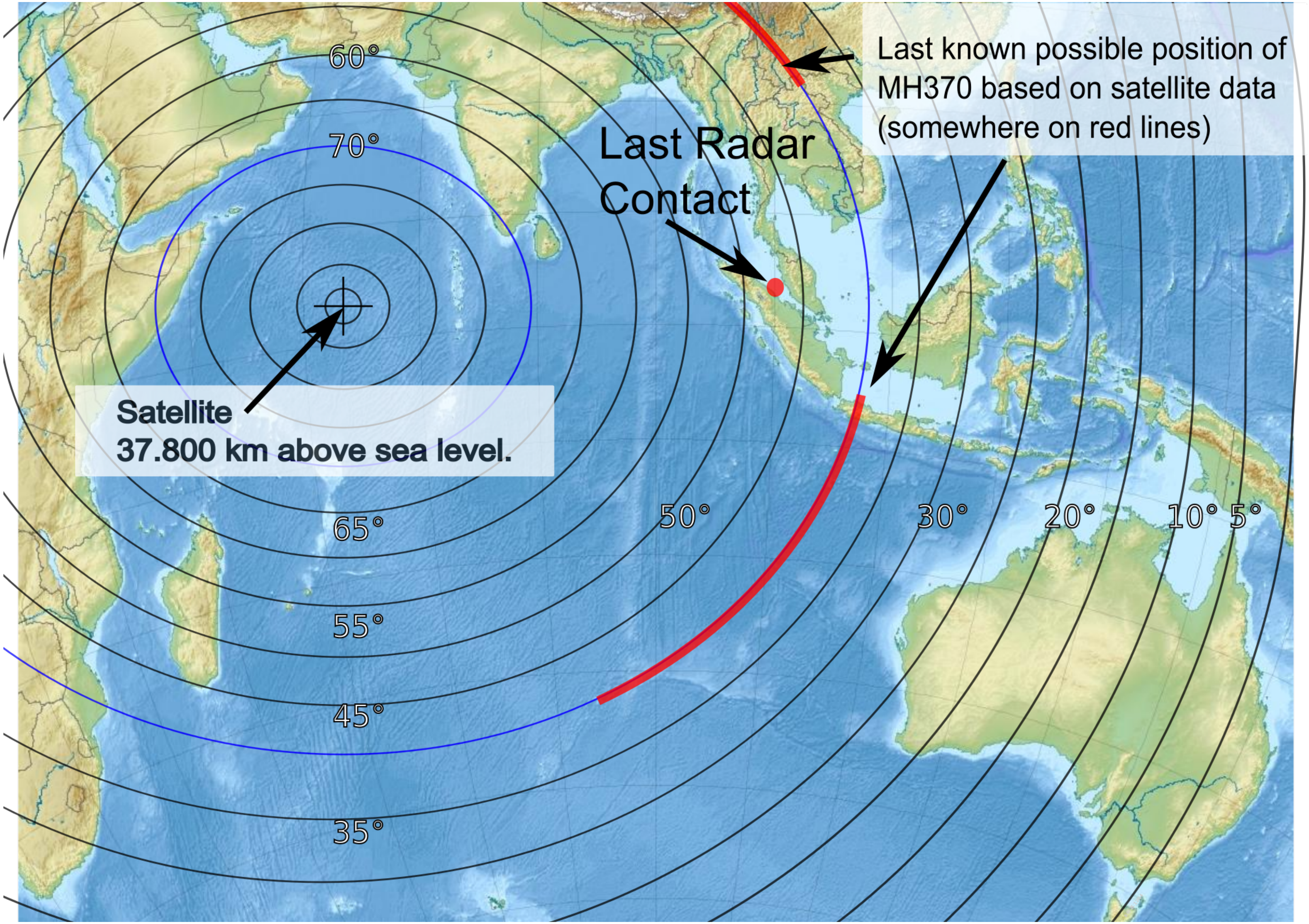
Tracking in 2D Space: Observation!

Observe a ping of the object that is distance $D = 4$ away!



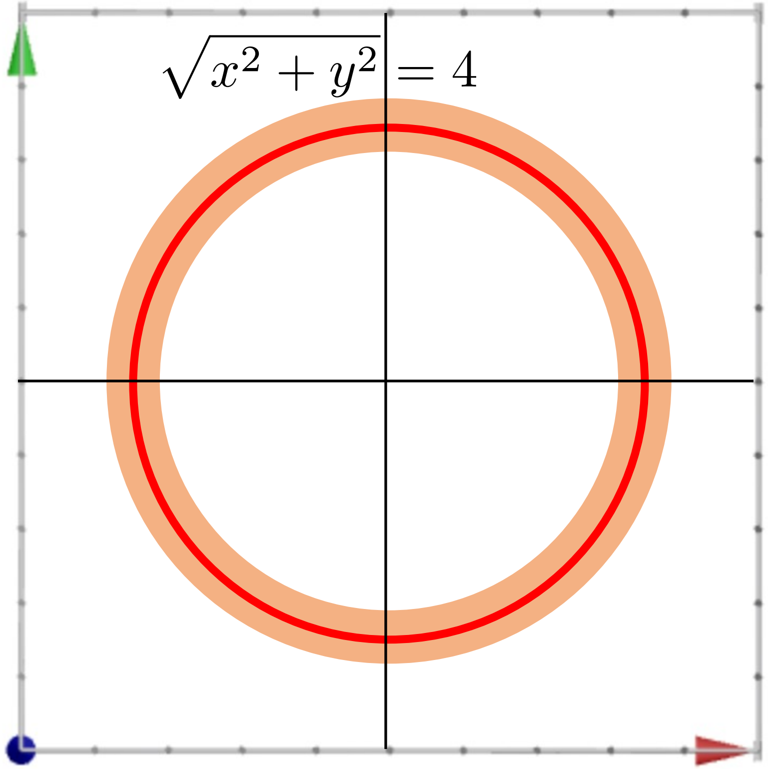


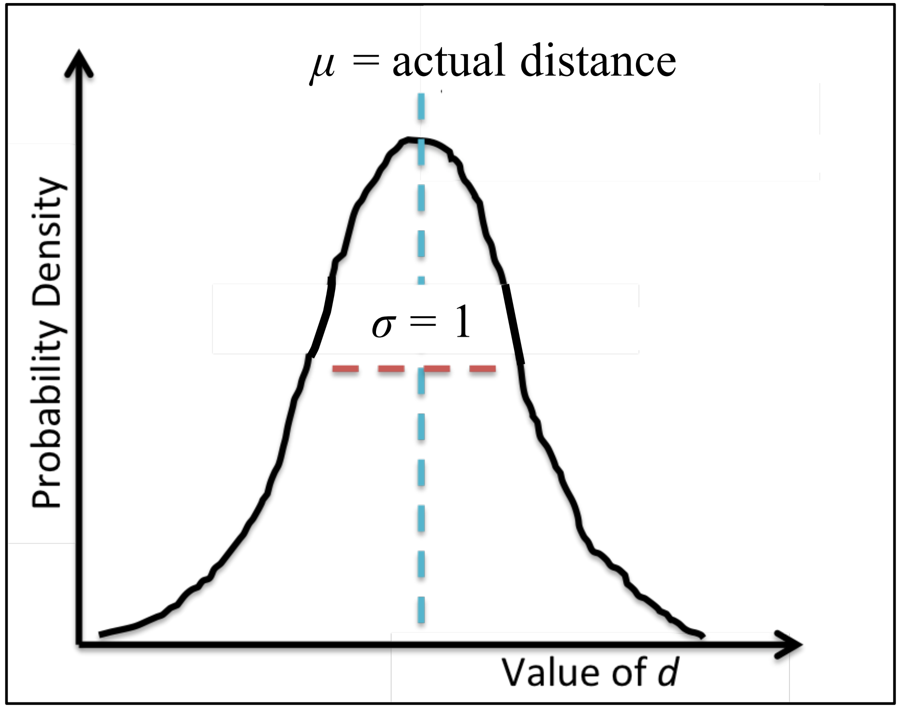
Know that the distance of a ping is normal with respect to the true distance


Last known possible position of
MH370 based on satellite data
(somewhere on red lines)
Last Radar
Contact
Satellite
37.800 km above sea level.
60°
70°
65°
55°
45°
35°
50°
30°
20°
10°
5°

Tracking in 2D Space: Observation!

Observe a ping of the object that is distance $D = 4$ away!





Know that the distance of a ping is normal with respect to the true distance

Tracking in 2D Space: Observation!

Observe a ping of the object that is distance $D = 4$ away from satellite!

$$D|X, Y \sim N(\mu = \sqrt{x^2 + y^2}, \sigma^2 = 1)$$

$$f(D = d|X = x, Y = y) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(d-\mu)^2}{2\sigma^2}}$$

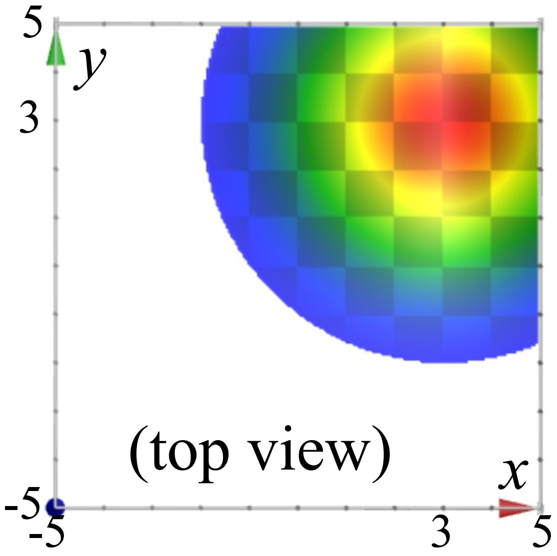
$$= \frac{1}{\sqrt{2\pi}} e^{\frac{-(d-\mu)^2}{2}}$$

$$= K_2 \cdot e^{\frac{-(d-\mu)^2}{2}}$$

$$= K_2 \cdot e^{\frac{-(d-\sqrt{x^2+y^2})^2}{2}}$$

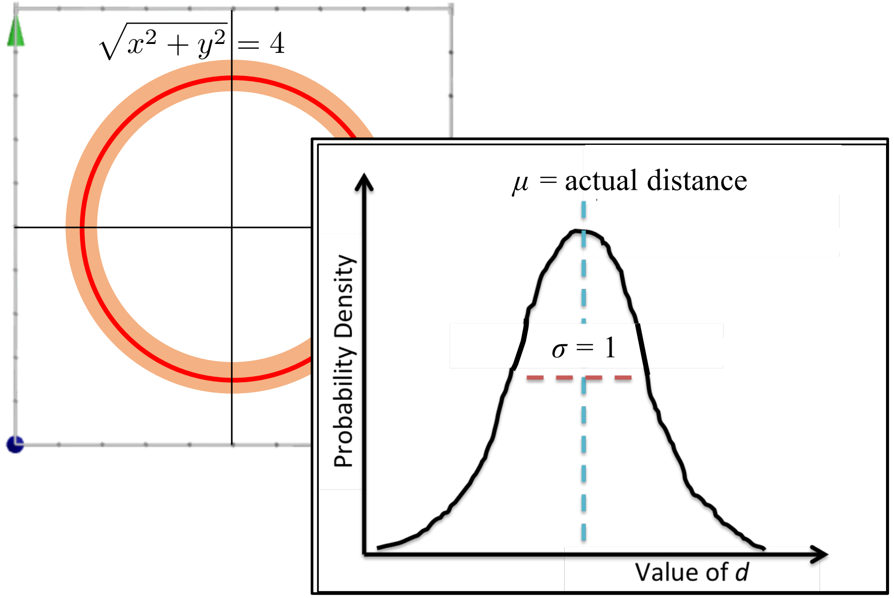
Tracking in 2D Space: New Belief

$$f(X = x, Y = y) = K_1 \cdot e^{-\frac{[(x-3)^2 + (y-3)^2]}{8}}$$



Prior

Observation



$$f(D = d | X = x, Y = y) = K_2 \cdot e^{-\frac{[d - \sqrt{x^2 + y^2}]^2}{2}}$$

What is your ***new*** belief for the location of the object being tracked?
Your joint probability density function can be expressed with a constant

Prior

$$f(X = x, Y = y) = K \cdot e^{-\frac{[(x-3)^2 + (y-3)^2]}{8}}$$

Observation

$$f(D = d | X = x, Y = y) = K \cdot e^{-[d - \sqrt{x^2 + y^2}]^2}$$

Tracking in 2D Space: New Belief

$$f(X = x, Y = y | D = 4) = \frac{f(D = 4 | X = x, Y = y) \cdot f(X = x, Y = y)}{f(D = 4)}$$

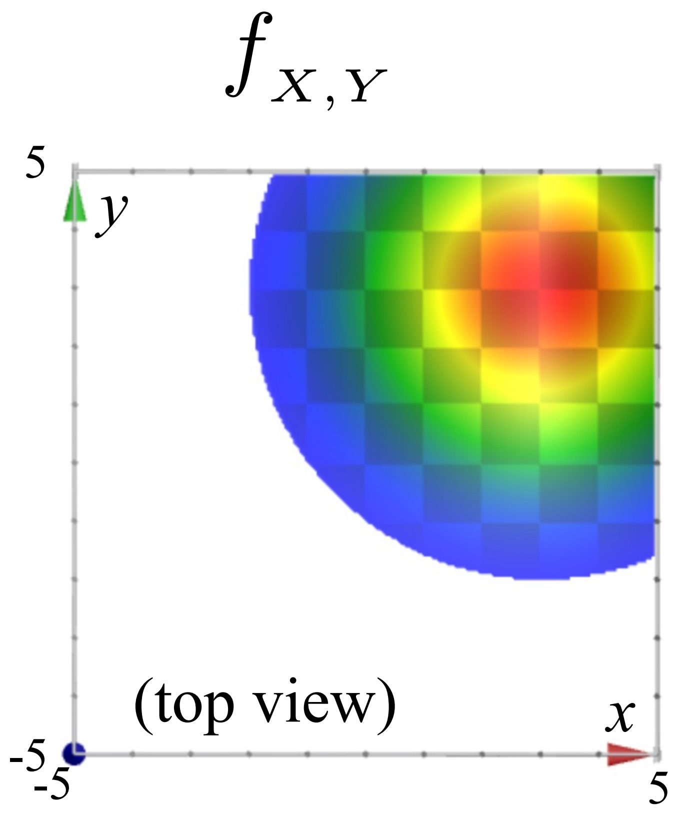
$$= \frac{K_1 \cdot e^{-\frac{[4 - \sqrt{x^2 + y^2})^2]}{2}} \cdot K_2 \cdot e^{-\frac{[(x-3)^2 + (y-3)^2]}{8}}}{f(D = 4)}$$

$$= \frac{K_3 \cdot e^{-\left[\frac{[4 - \sqrt{x^2 + y^2})^2]}{2} + \frac{[(x-3)^2 + (y-3)^2]}{8}\right]}}{f(D = 4)}$$

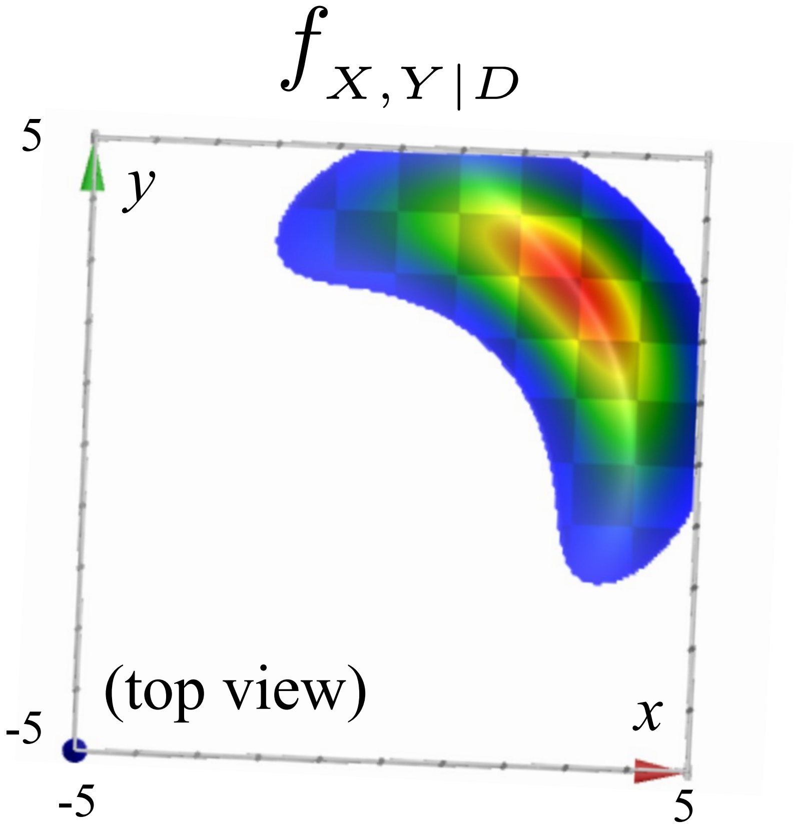
$$= K_4 \cdot e^{-\left[\frac{(4 - \sqrt{x^2 + y^2})^2}{2} + \frac{[(x-3)^2 + (y-3)^2]}{8}\right]}$$

For your notes…

Tracking in 2D Space: Posterior

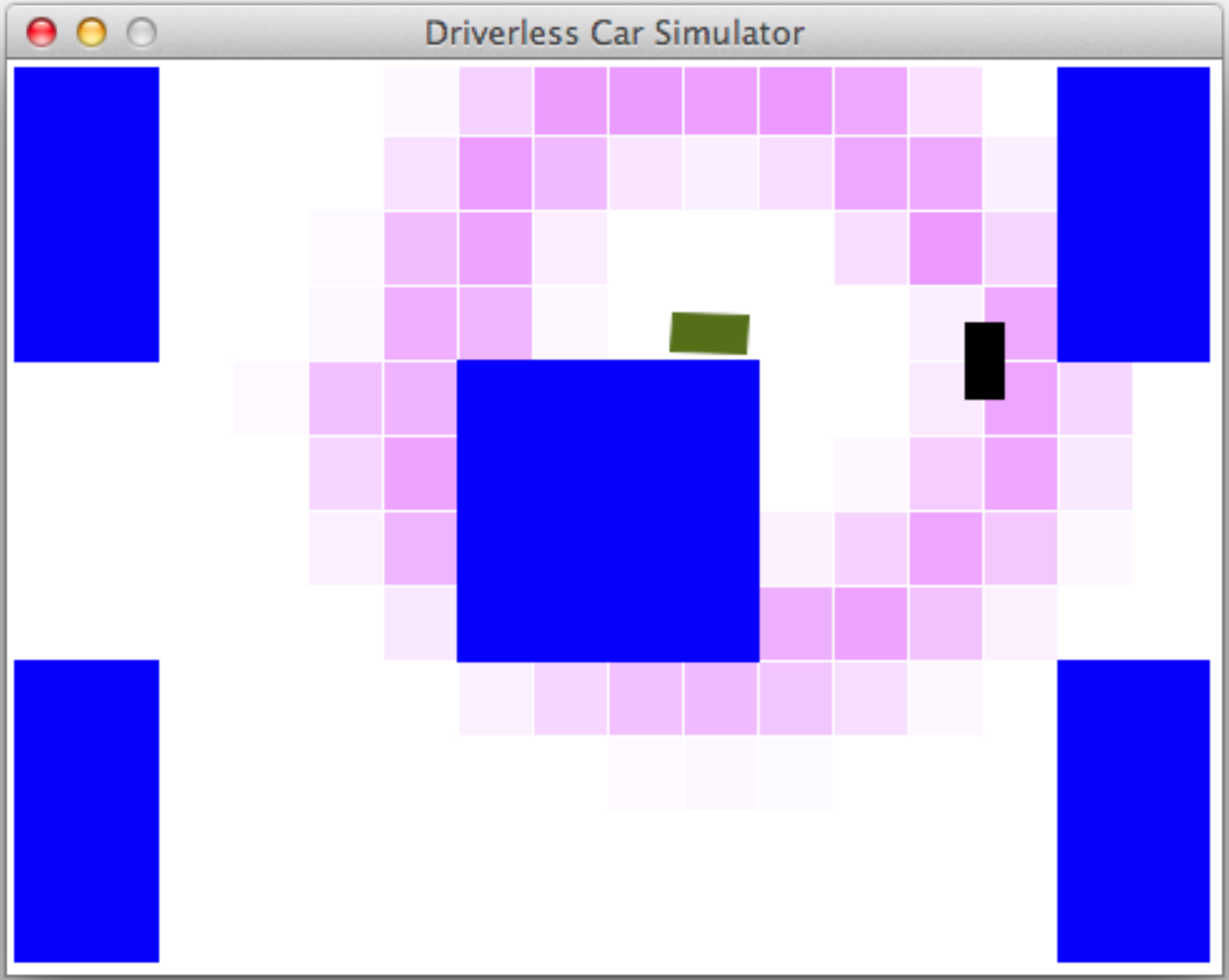


Prior



Posterior

Tracking in 2D Space: CS221



See you Wednesday.