Conditional Joint Distributions

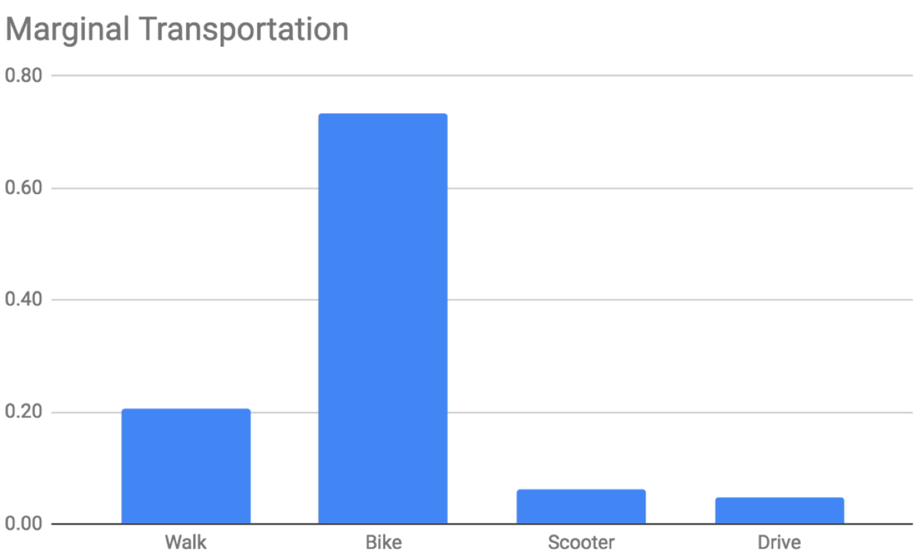
Chris Piech
CS109, Stanford University

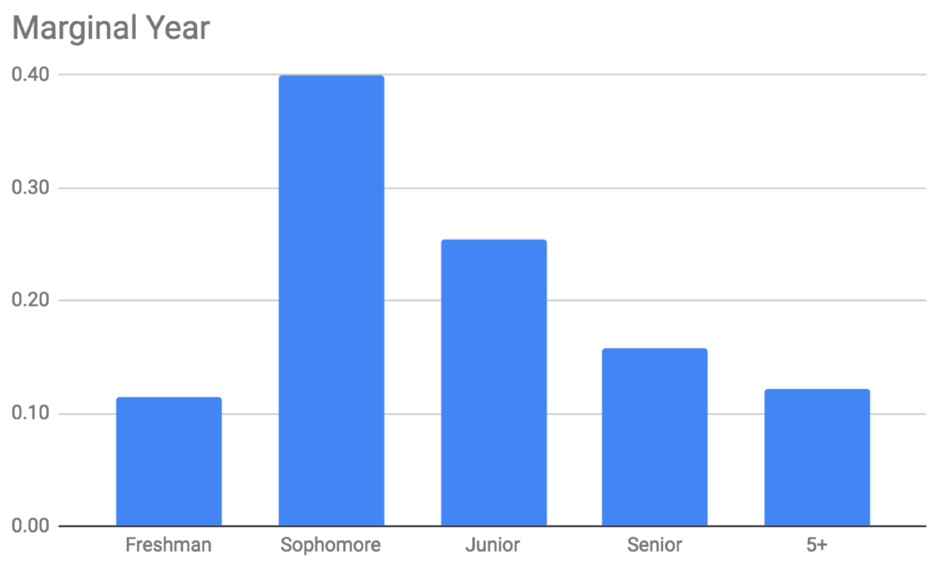
Joint Random Variables

- [x] Use a joint table, density function or CDF to solve probability question
- [x] Think about **conditional** probabilities with joint variables (which might be continuous)
- [ ] Use and find **independence** of random variables
- [ ] Use and find **expectation** of random variables
- [ ] What happens when you **add** random variables?

Joint Probability Table

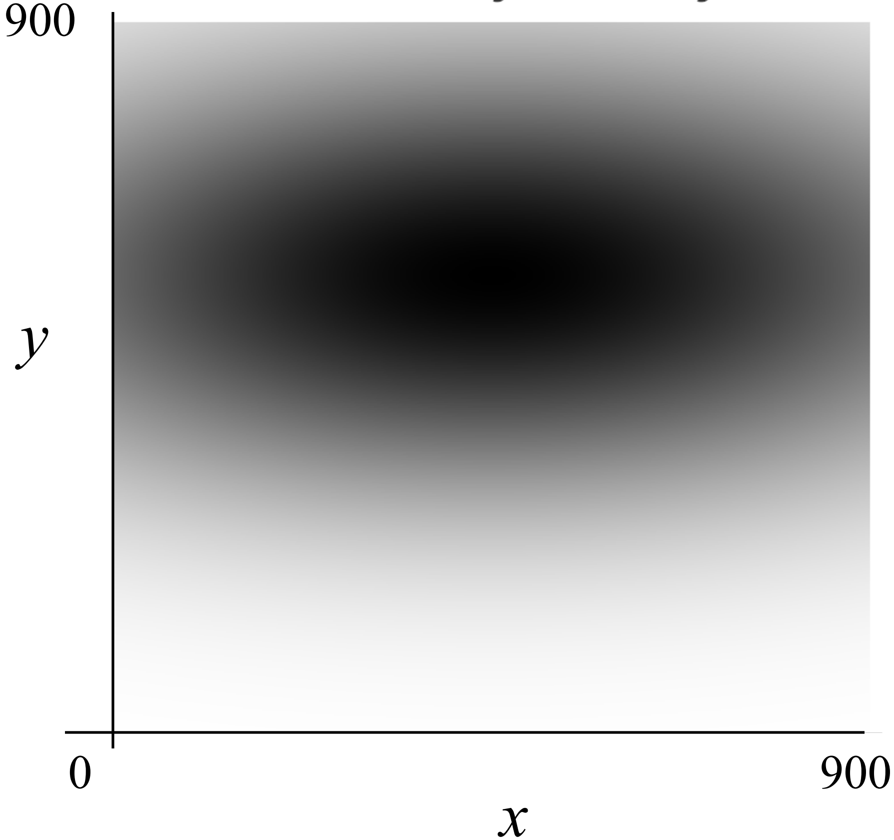
	Walk	Bike	Scooter	Drive	**Marginal Year**
Freshman	0.04	0.04	0.01	0.03	0.12
Sophomore	0.03	0.34	0.03	0.00	0.40
Junior	0.04	0.21	0.01	0.00	0.25
Senior	0.07	0.08	0.01	0.00	0.16
5+	0.04	0.07	0.00	0.02	0.12
Marginal Mode	0.21	0.73	0.06	0.05	





Continuous Joint Random Variables

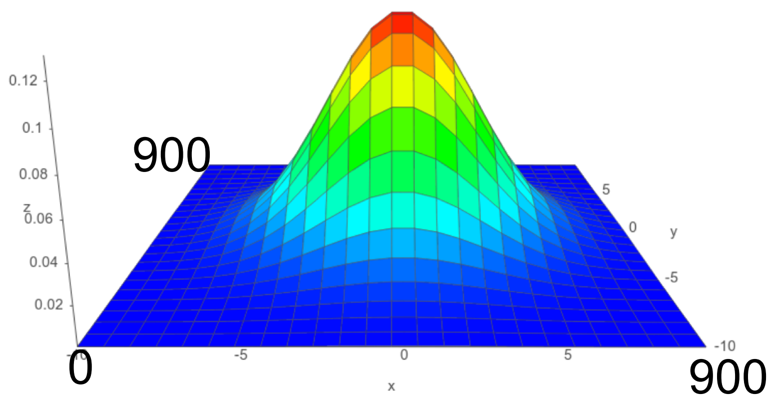
Dart Probability Density

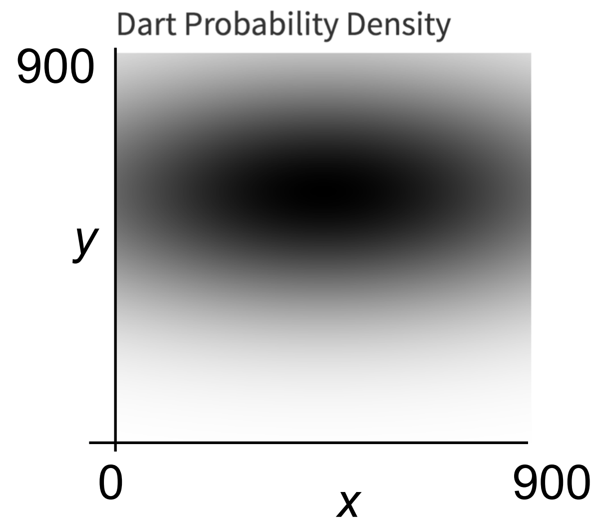


Dart Results

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.





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

Jointly Continuous

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

- Cumulative Density Function (CDF):

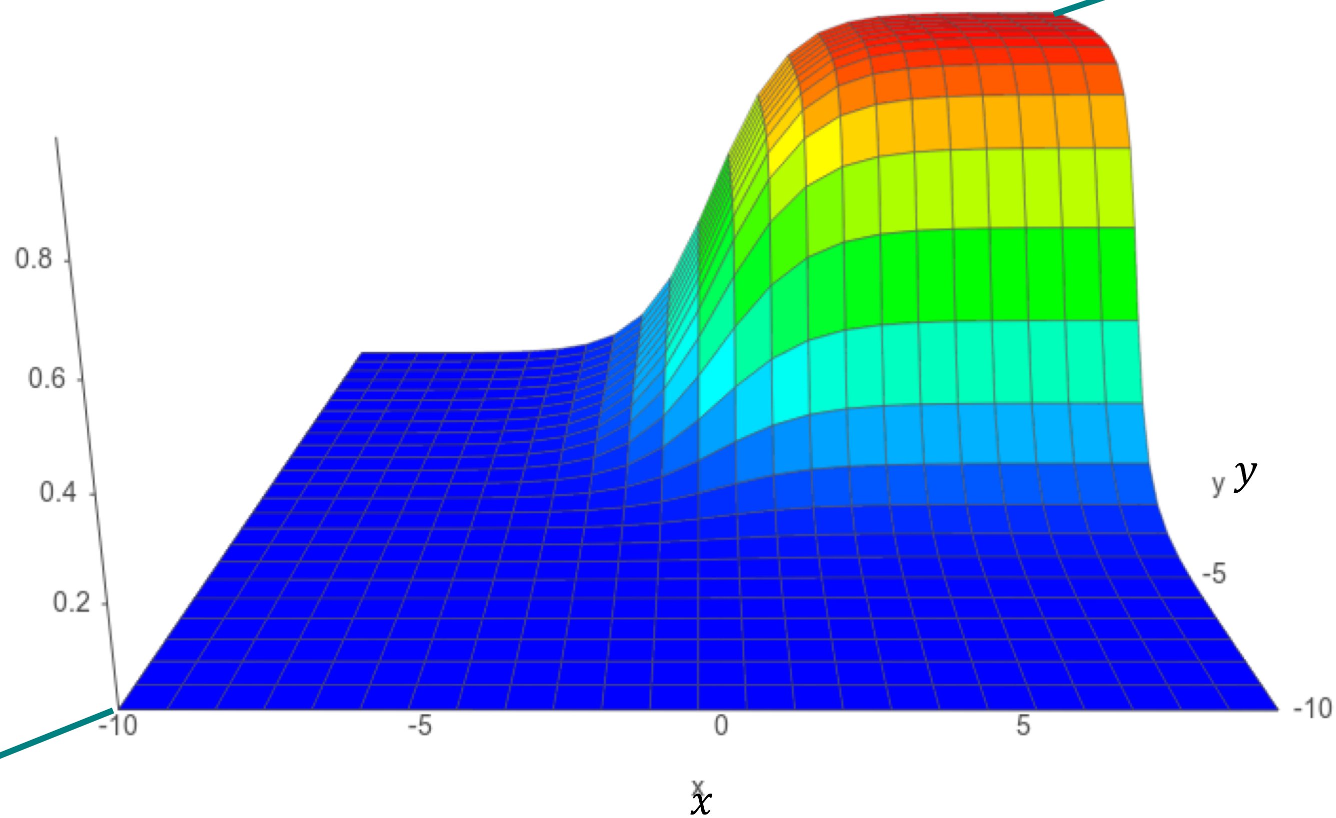
$$F_{X,Y}(a,b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x, y)\, dy\, dx$$

$$f_{X,Y}(a,b) = \frac{\partial^2}{\partial a\, \partial b} F_{X,Y}(a,b)$$

Jointly CDF

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$$

to 1 as
x → +∞,
y → +∞

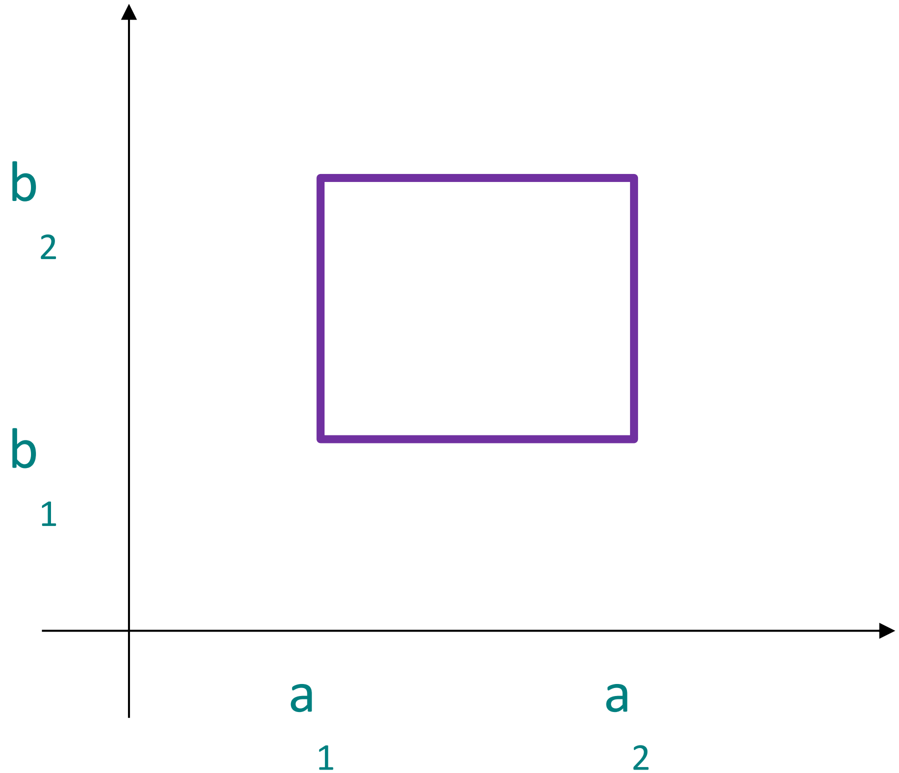


to 0 as
x → -∞,
y → -∞

plot by [Academo]

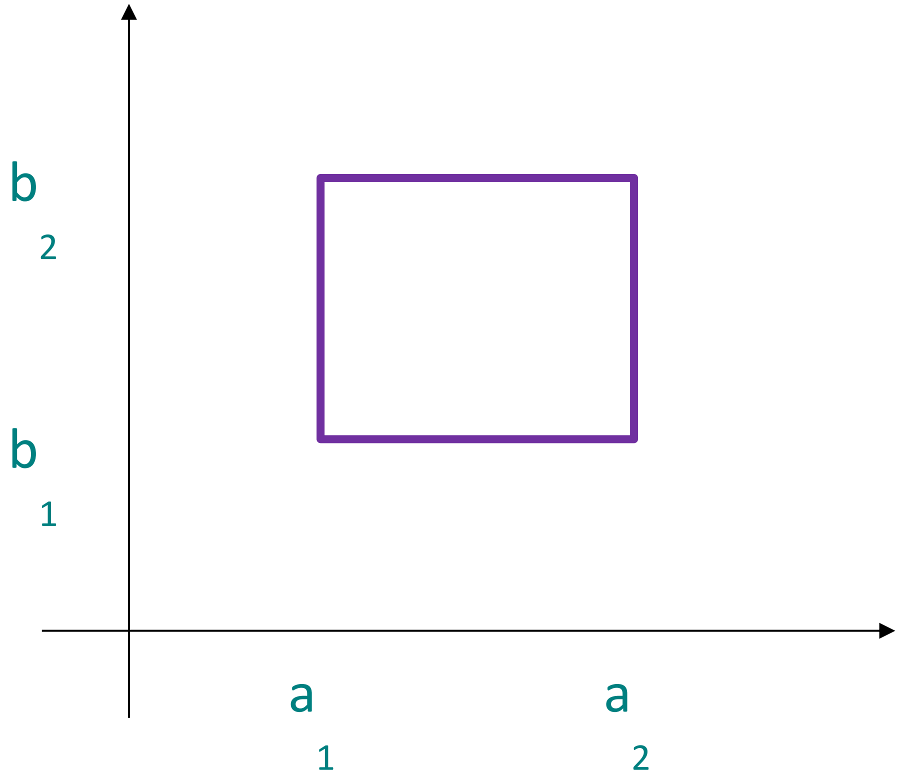
Probabilities from Joint CDF

$$P\left(a_1 < X \leq a_{2,} b_1 < Y \leq b_2\right)$$



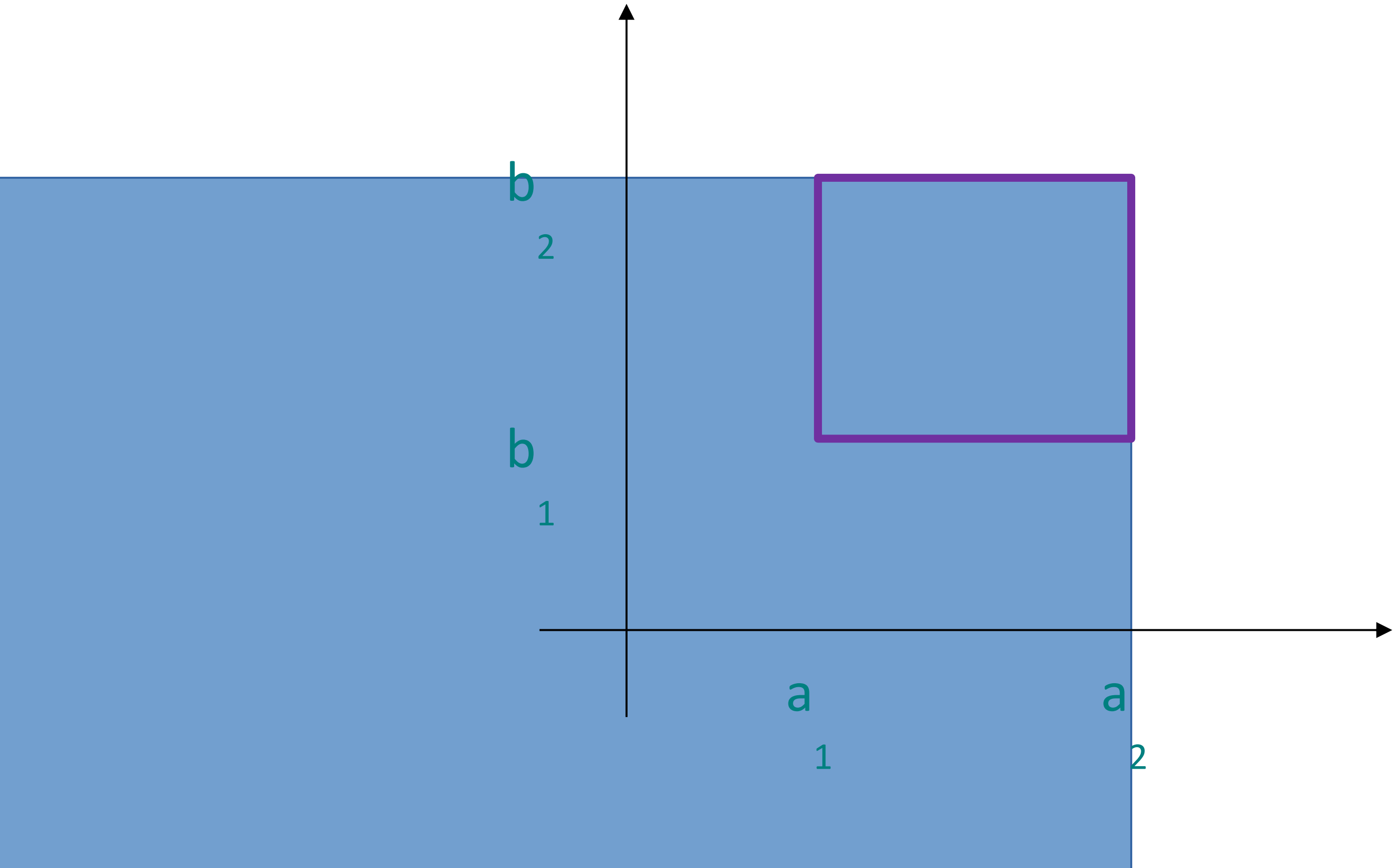
Probabilities from Joint CDF

$$P(a_1 < X \le a_{2,} b_1 < Y \le b_2) = F_{X,Y}(a_{2,} b_2)$$



Probabilities from Joint CDF

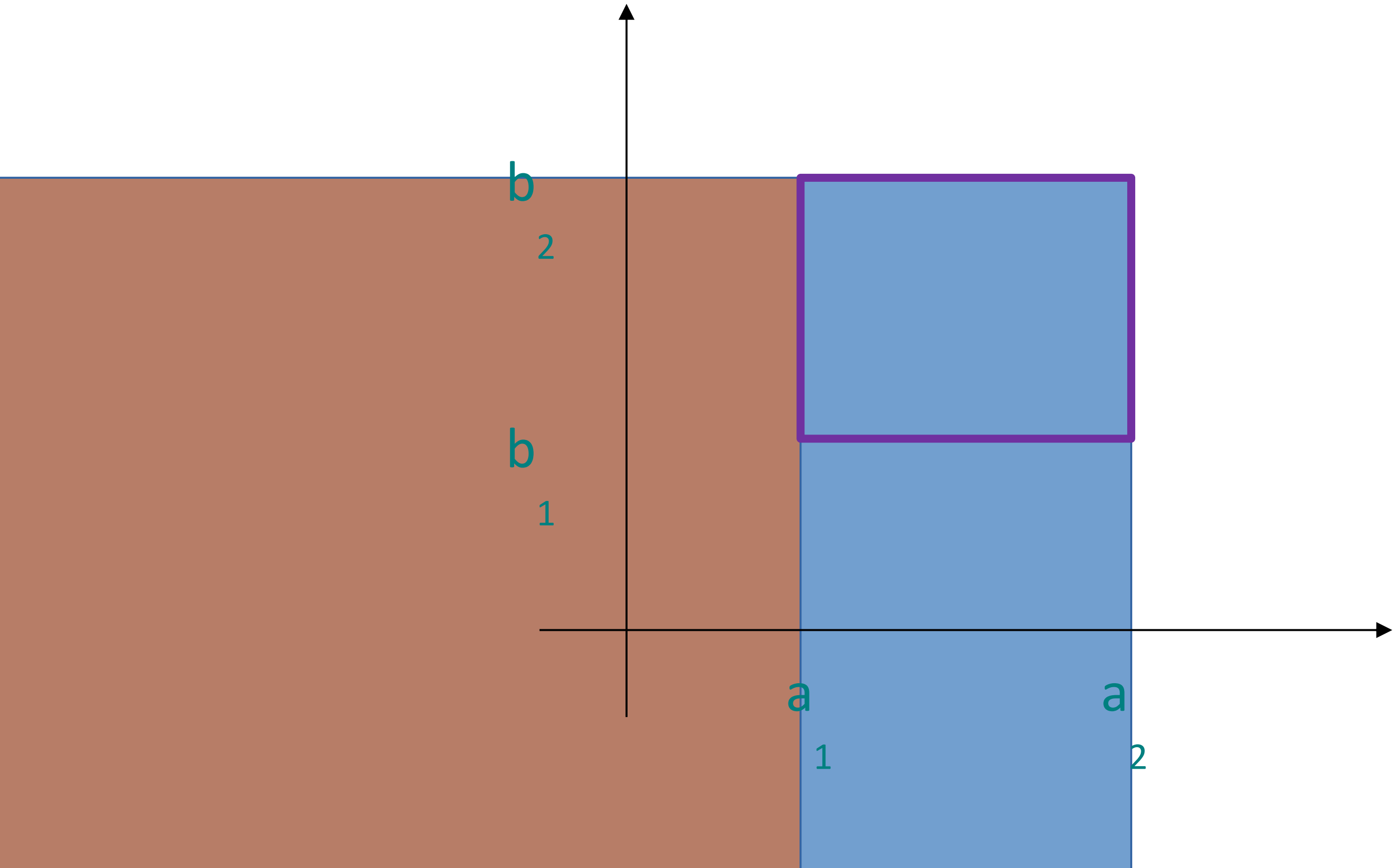
$$P\left(a_1 < X \le a_2, b_1 < Y \le b_2\right) = F_{X,Y}\left(a_2, b_2\right)$$



Probabilities from Joint CDF

$$P(a_1 < X \leq a_{2,} b_1 < Y \leq b_2) = F_{X,Y}(a_{2,} b_2)$$

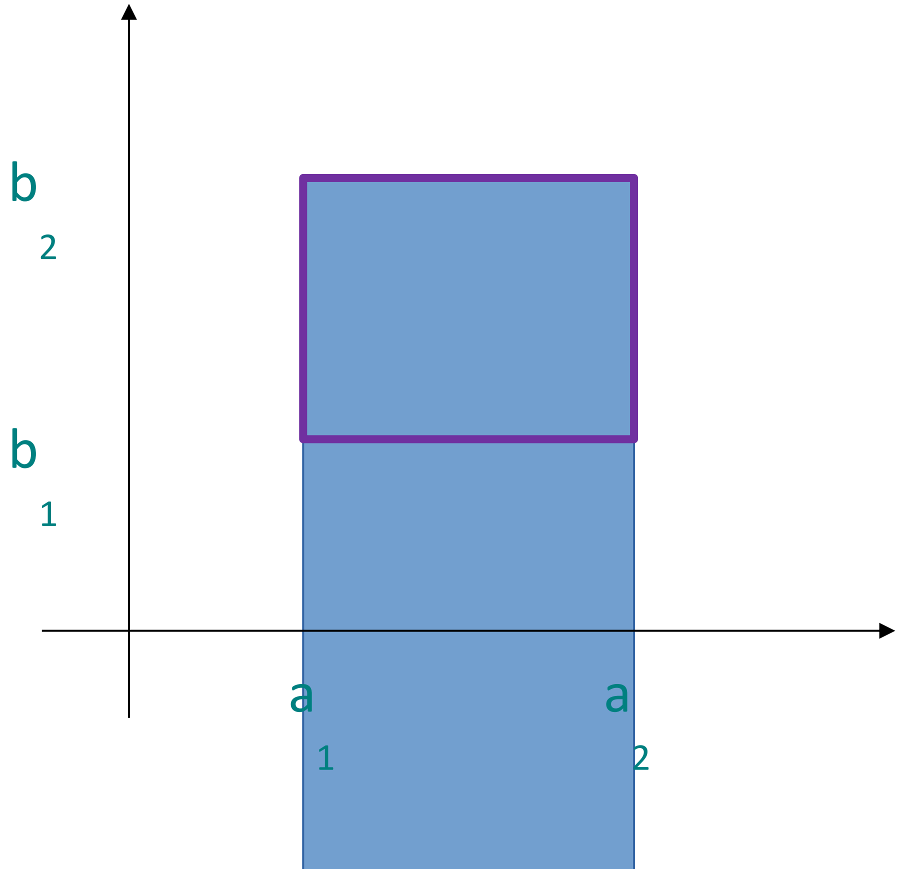
$$-F_{X,Y}(a_{1,} b_2)$$



Probabilities from Joint CDF

$$P\left(a_1 < X \le a_{2,} b_1 < Y \le b_2\right) = F_{X,Y}\left(a_{2,} b_2\right)$$

$$-F_{X,Y}\left(a_{1,} b_2\right)$$

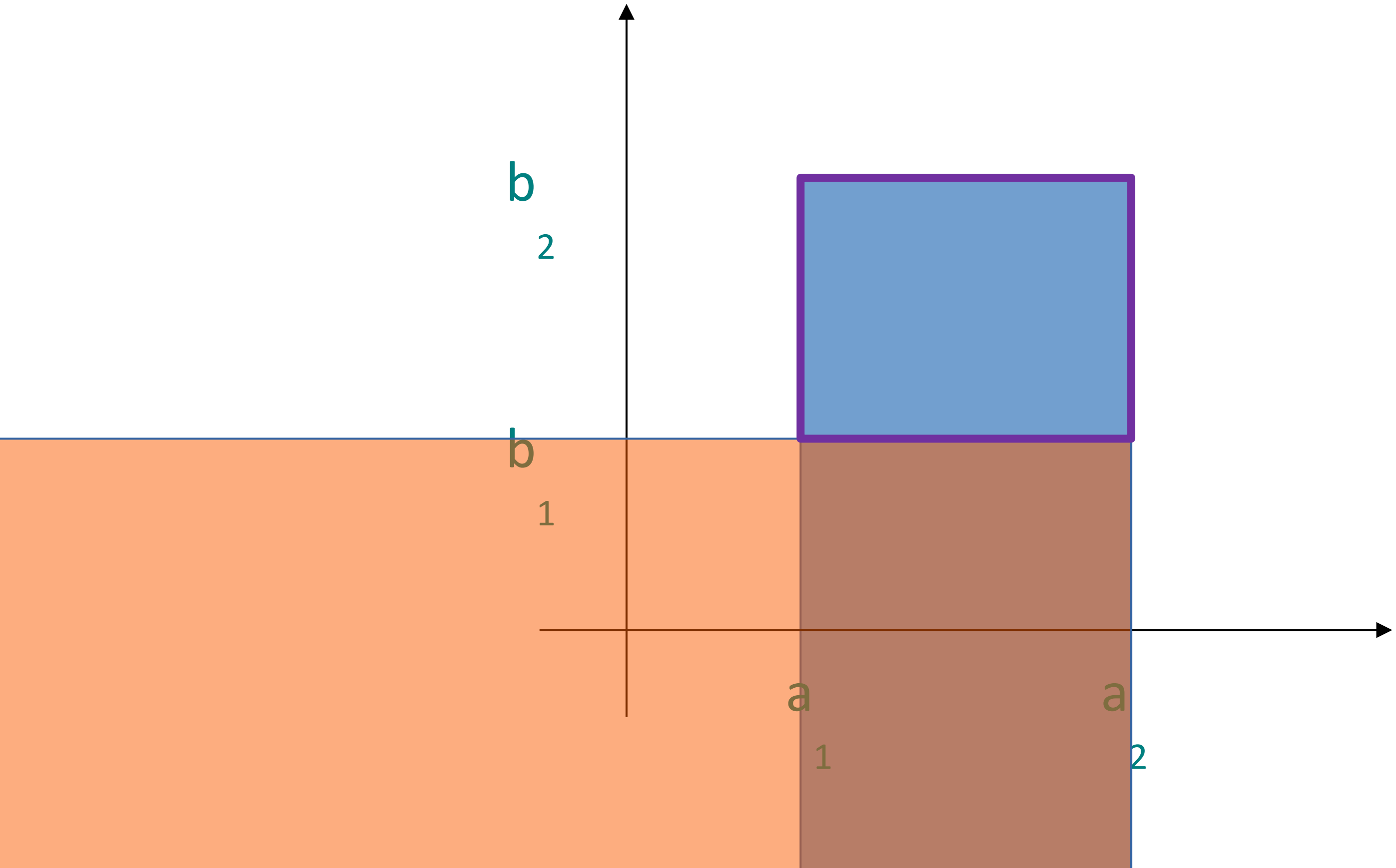


Probabilities from Joint CDF

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F_{X,Y}(a_2, b_2)$$

$$-F_{X,Y}(a_1, b_2)$$

$$-F_{X,Y}(a_2, b_1)$$

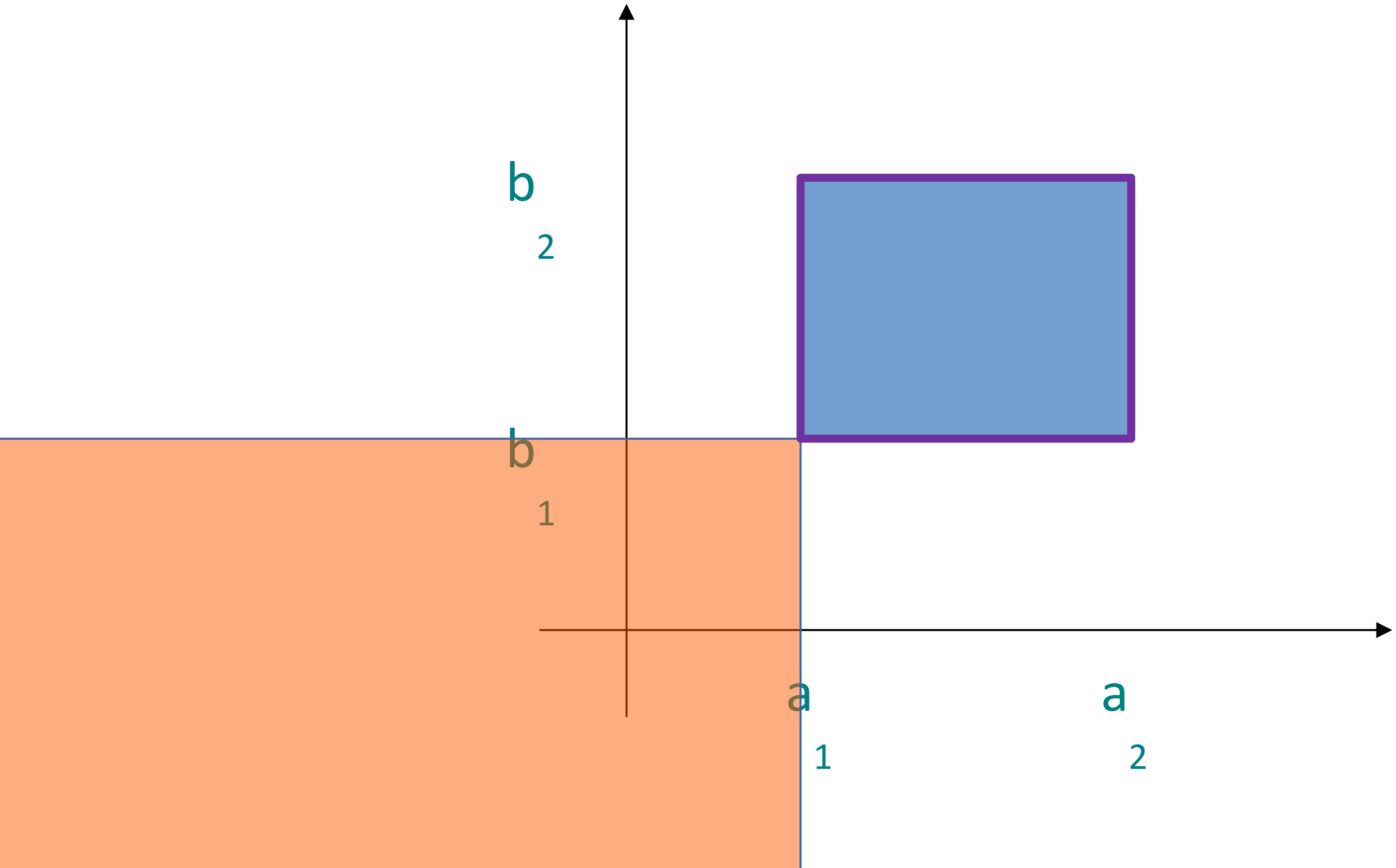


Probabilities from Joint CDF

$$P\left(a_1 < X \leq a_{2,} b_1 < Y \leq b_2\right) = F_{X,Y}\left(a_{2,} b_2\right)$$

$$-F_{X,Y}\left(a_{1,} b_2\right)$$

$$-F_{X,Y}\left(a_{2,} b_1\right)$$



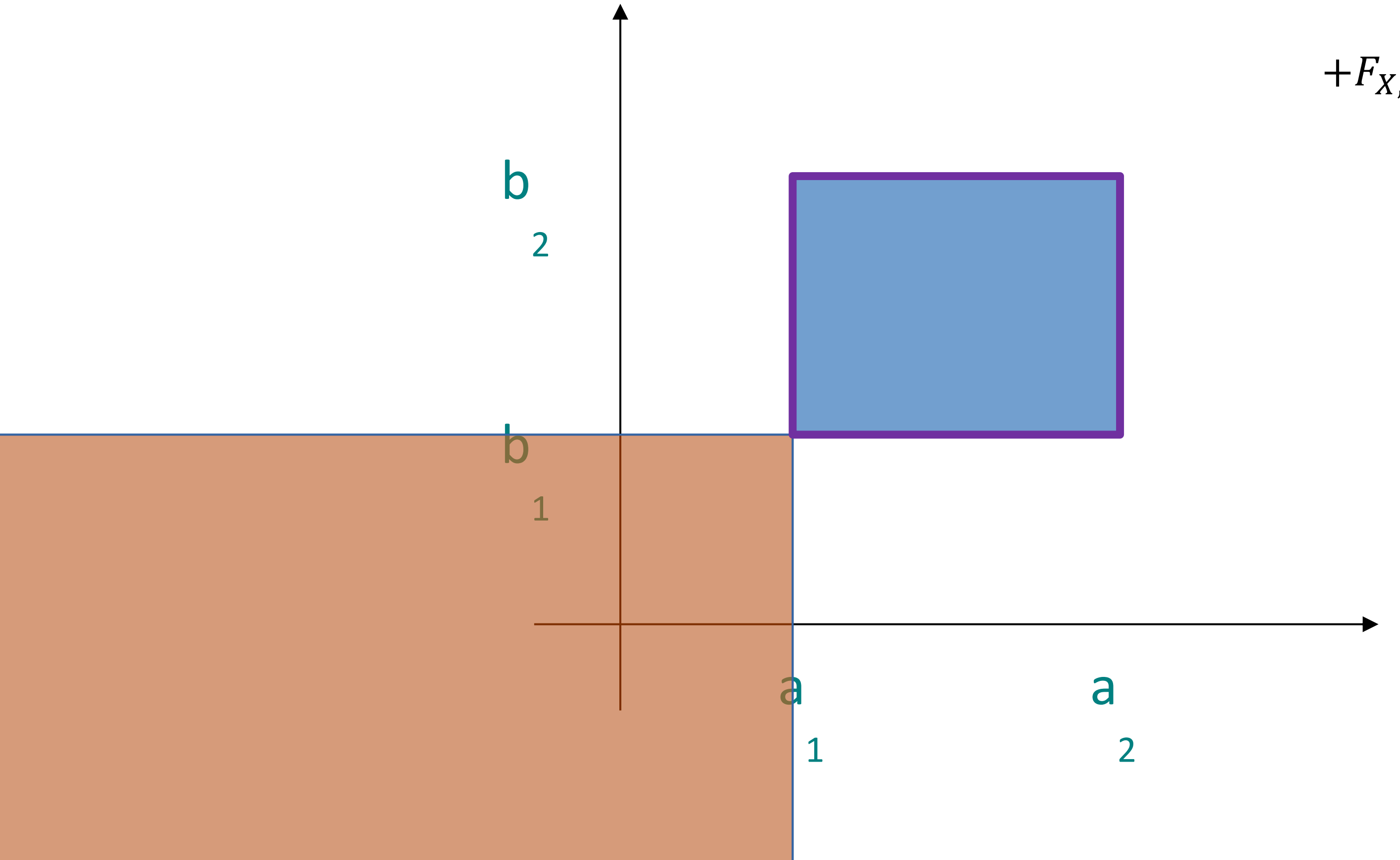
Probabilities from Joint CDF

$$P(a_1 < X \le a_2, b_1 < Y \le b_2) = F_{X,Y}(a_2, b_2)$$

$$-F_{X,Y}(a_1, b_2)$$

$$-F_{X,Y}(a_2, b_1)$$

$$+F_{X,Y}(a_1, b_1)$$



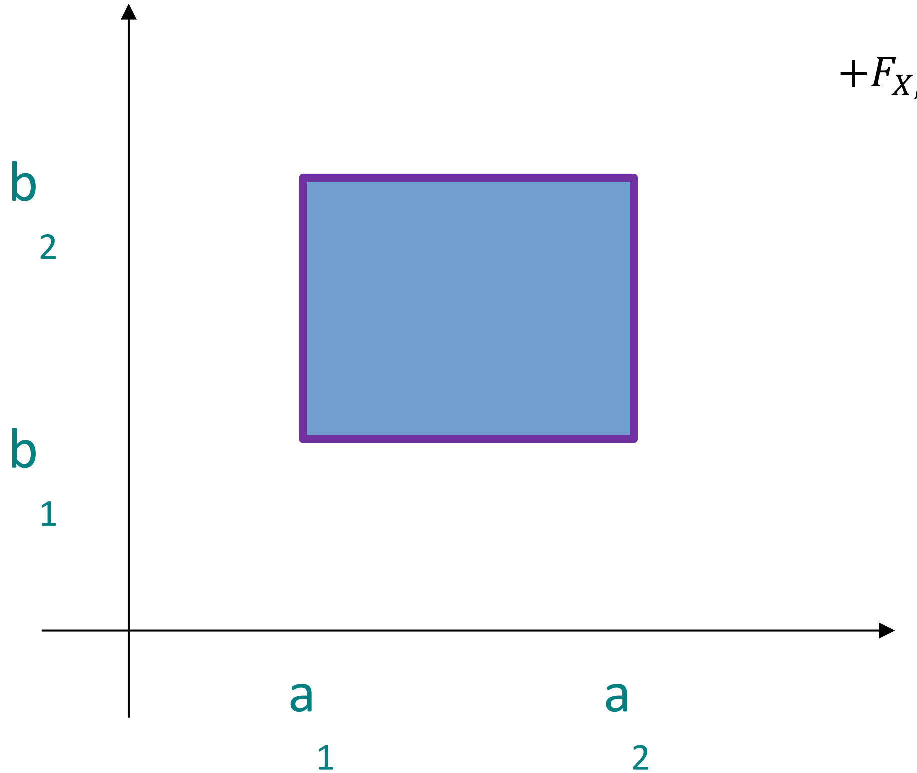
Probabilities from Joint CDF

$$P(a_1 < X \leq a_{2,} b_1 < Y \leq b_2) = F_{X,Y}(a_{2,} b_2)$$

$$-F_{X,Y}(a_{1,} b_2)$$

$$-F_{X,Y}(a_{2,} b_1)$$

$$+F_{X,Y}(a_{1,} b_1)$$

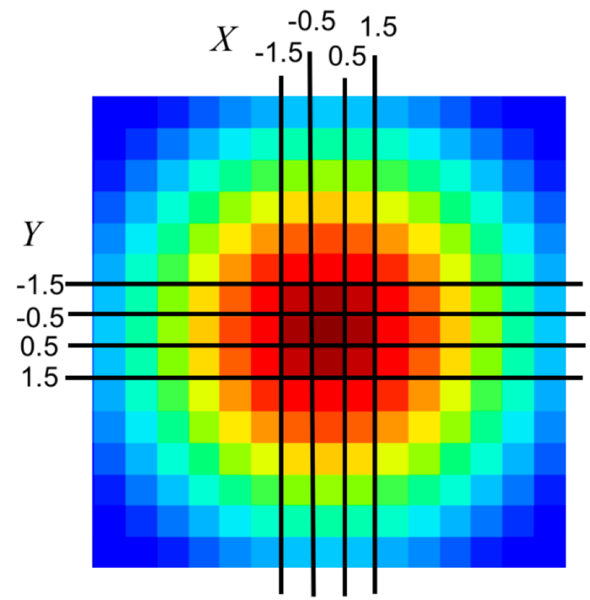


Probability for Instagram!

Gaussian Blur

In image processing, a Gaussian blur is the result of blurring an image by a Gaussian function. It is a widely used effect in graphics software, typically to reduce image noise.

0.0000	0.0000	0.0000	0.0001	0.0001	(
0.0000	0.0001	0.0005	0.0020	0.0032	(
0.0000	0.0005	0.0052	0.0206	0.0326	(
0.0001	0.0020	0.0206	0.0821	0.1300	(
0.0001	0.0032	0.0326	0.1300	**0.2060**	(
0.0001	0.0020	0.0206	0.0821	0.1300	(
0.0000	0.0005	0.0052	0.0206	0.0326	(
0.0000	0.0001	0.0005	0.0020	0.0032	(
0.0000	0.0000	0.0000	0.0001	0.0001	(



Gaussian Blur

In image processing, a Gaussian blur is the result of blurring an image by a Gaussian function. It is a widely used effect in graphics software, typically to reduce image noise.

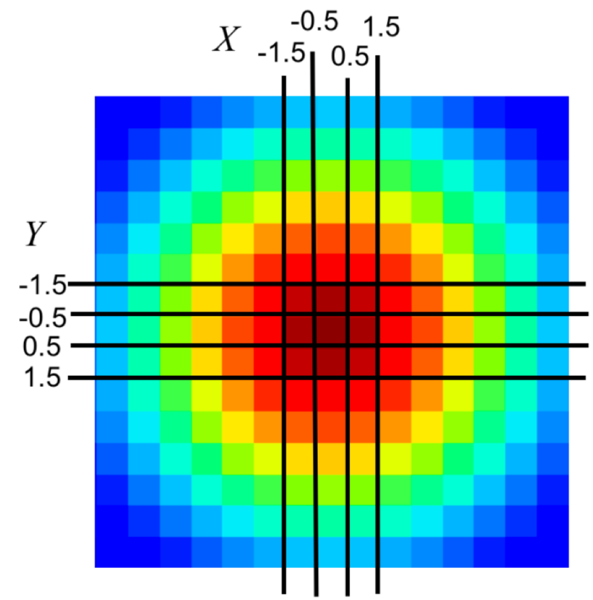
Gaussian blurring with StDev = 3, is based on a joint probability distribution:

Joint PDF

$$f_{X,Y}(x,y) = \frac{1}{2\pi \cdot 3^2} e^{-\frac{x^2+y^2}{2\cdot 3^2}}$$

Joint CDF

$$F_{X,Y}(x,y) = \Phi\left(\frac{x}{3}\right) \cdot \Phi\left(\frac{y}{3}\right)$$



Used to generate this weight matrix

Gaussian Blur

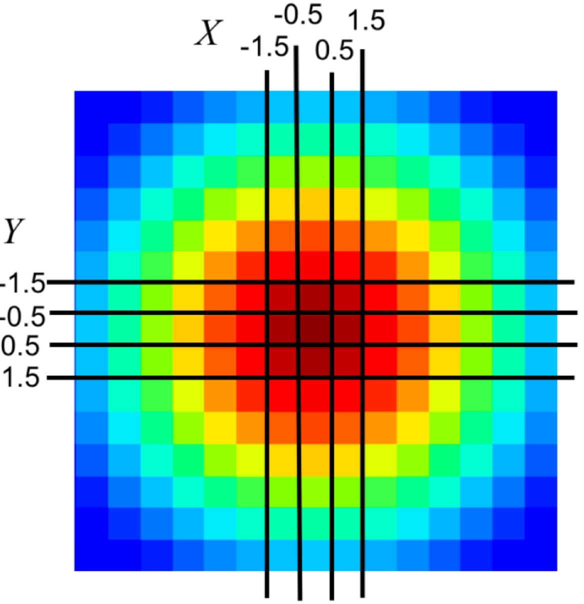
Joint PDF

$$f_{X,Y}(x,y) = \frac{1}{2\pi \cdot 3^2} e^{-\frac{x^2+y^2}{2\cdot 3^2}}$$

Joint CDF

$$F_{X,Y}(x,y) = \Phi\left(\frac{x}{3}\right) \cdot \Phi\left(\frac{y}{3}\right)$$

Weight Matrix



Each pixel is given a weight equal to the probability that X and Y are both within the pixel bounds. The center pixel covers the area where

$-0.5 \le x \le 0.5$ and $-0.5 \le y \le 0.5$

What is the weight of the center pixel?

$$P(-0.5 < X < 0.5, -0.5 < Y < 0.5)$$
$$= P(X < 0.5, Y < 0.5) - P(X < 0.5, Y < -0.5)$$
$$- P(X < -0.5, Y < 0.5) + P(X < -0.5, Y < -0.5)$$
$$= \phi\left(\frac{0.5}{3}\right) \cdot \phi\left(\frac{0.5}{3}\right) - 2\phi\left(\frac{0.5}{3}\right) \cdot \phi\left(\frac{-0.5}{3}\right)$$
$$+ \phi\left(\frac{-0.5}{3}\right) \cdot \phi\left(\frac{-0.5}{3}\right)$$
$$= 0.5662^2 - 2 \cdot 0.5662 \cdot 0.4338 + 0.4338^2 = 0.206$$

Properties of Joint Distributions

Boolean Operation on Variable = Event

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

$$P(X \leq 5)$$

$$P(Y = 6)$$

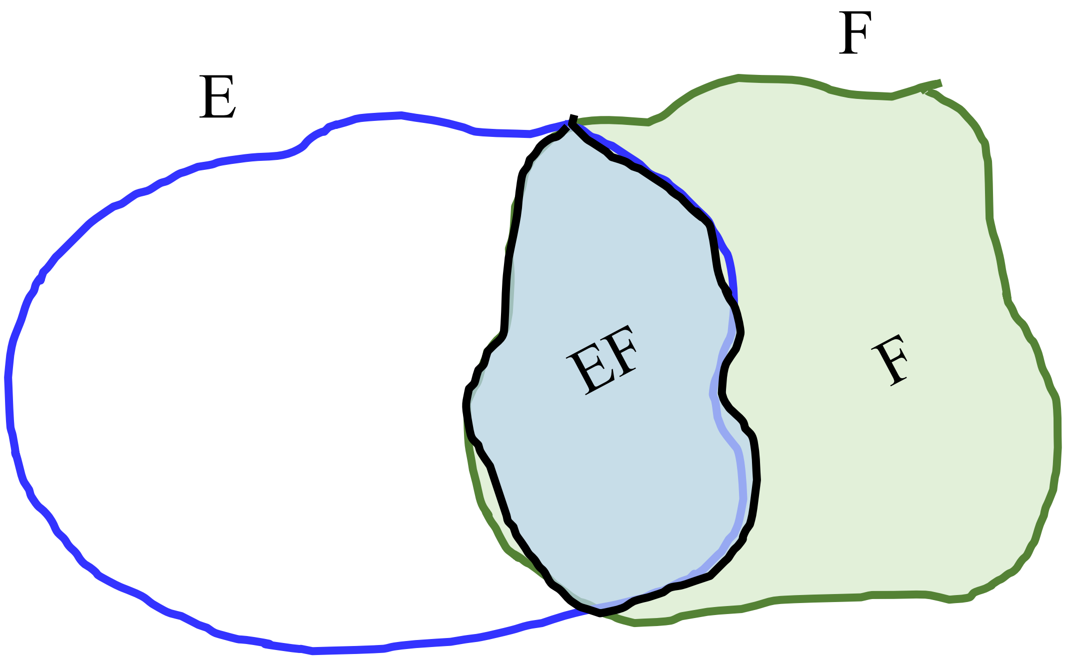
$$P(5 \leq Z \leq 10)$$

Conditionals with multiple variables

Discrete Conditional Distribution

- Recall that for *events* E and F:

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



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 (where $p_Y(y) > 0$):

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

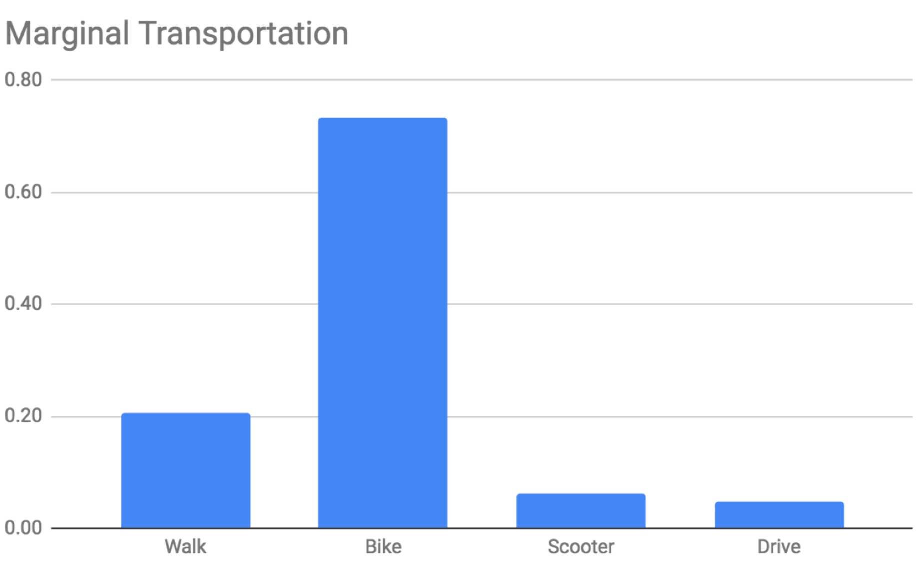
  - Conditional CDF of X given Y (where $p_Y(y) > 0$):

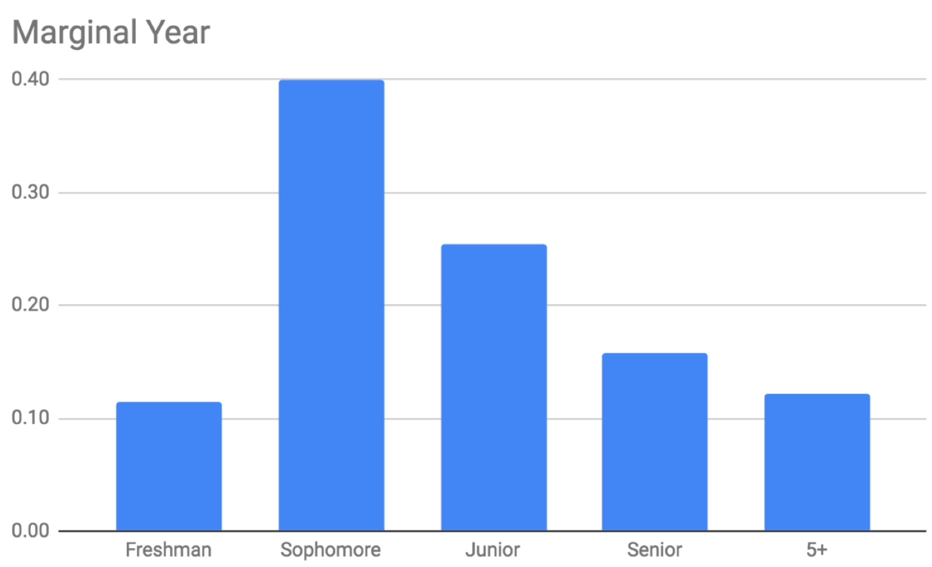
$$F_{X|Y}(a \mid y) = P(X \le a \mid Y = y) = \frac{P(X \le a, Y = y)}{P(Y = y)}$$

$$= \frac{\sum_{x \le a} p_{X,Y}(x, y)}{p_Y(y)} = \sum_{x \le a} p_{X|Y}(x \mid y)$$

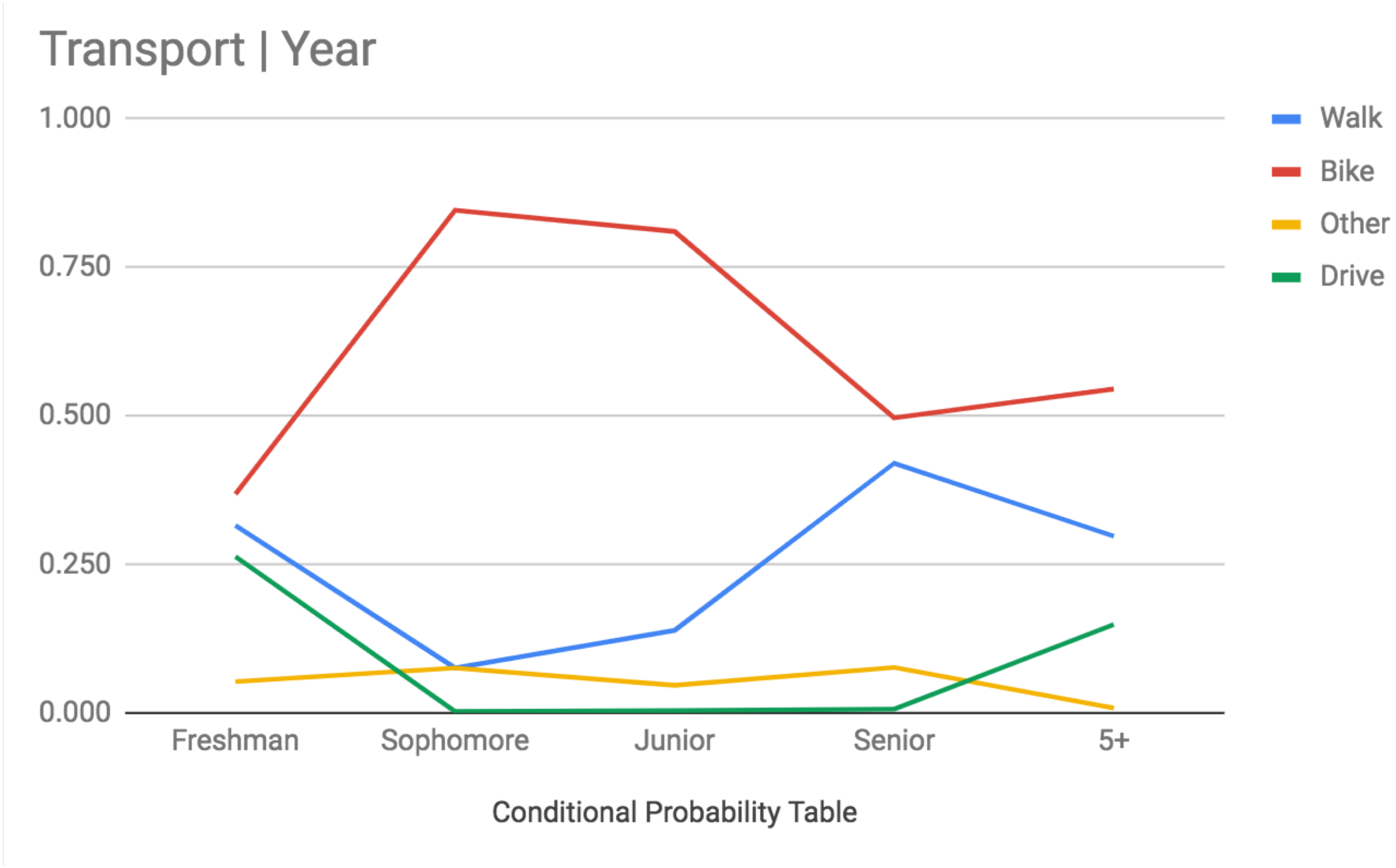
Joint Probability Table

	Walk	Bike	Scooter	Drive	**Marginal Year**
Freshman	0.04	0.04	0.01	0.03	0.12
Sophomore	0.03	0.34	0.03	0.00	0.40
Junior	0.04	0.21	0.01	0.00	0.25
Senior	0.07	0.08	0.01	0.00	0.16
5+	0.04	0.07	0.00	0.02	0.12
Marginal Mode	0.21	0.73	0.06	0.05	

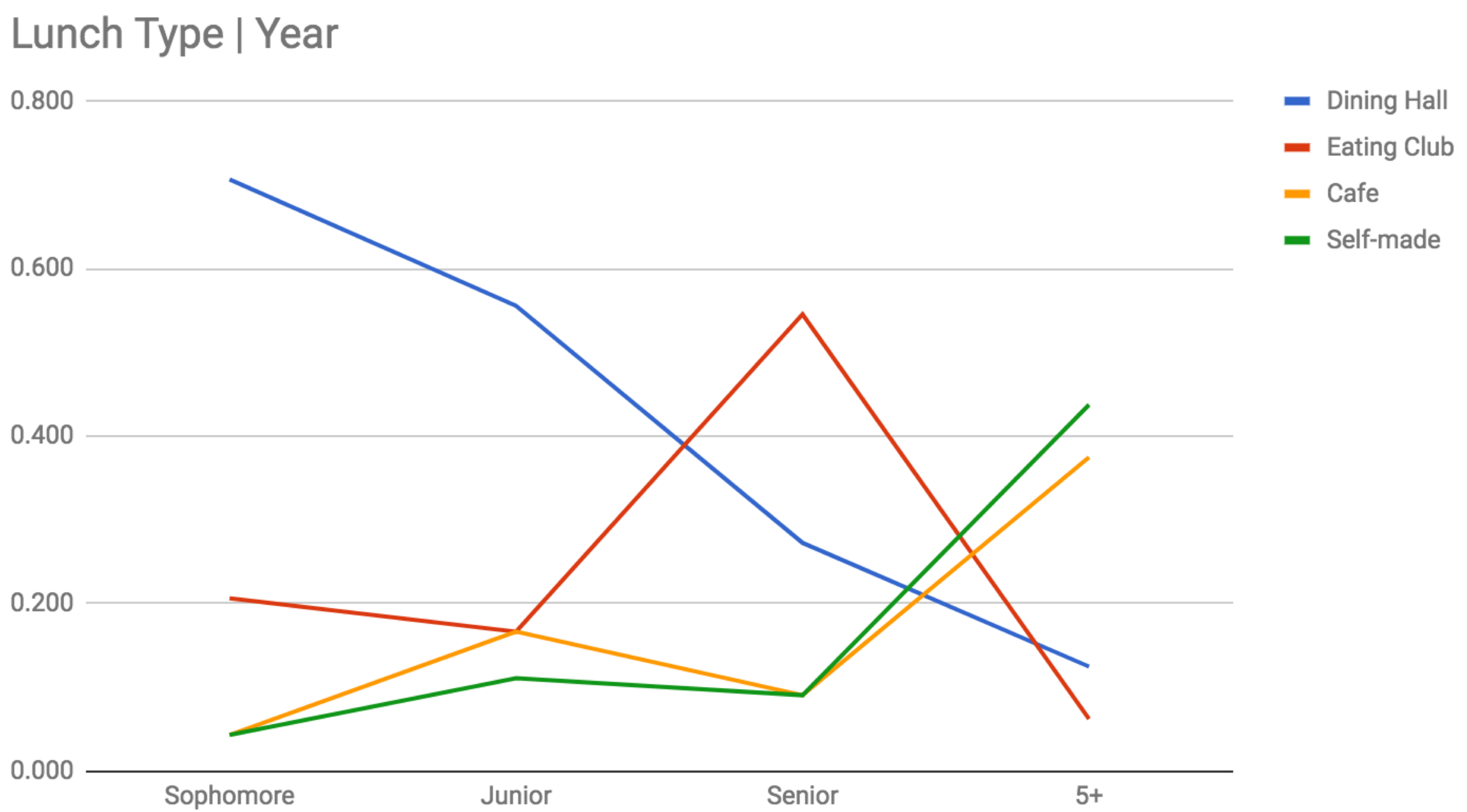




Transport | Year

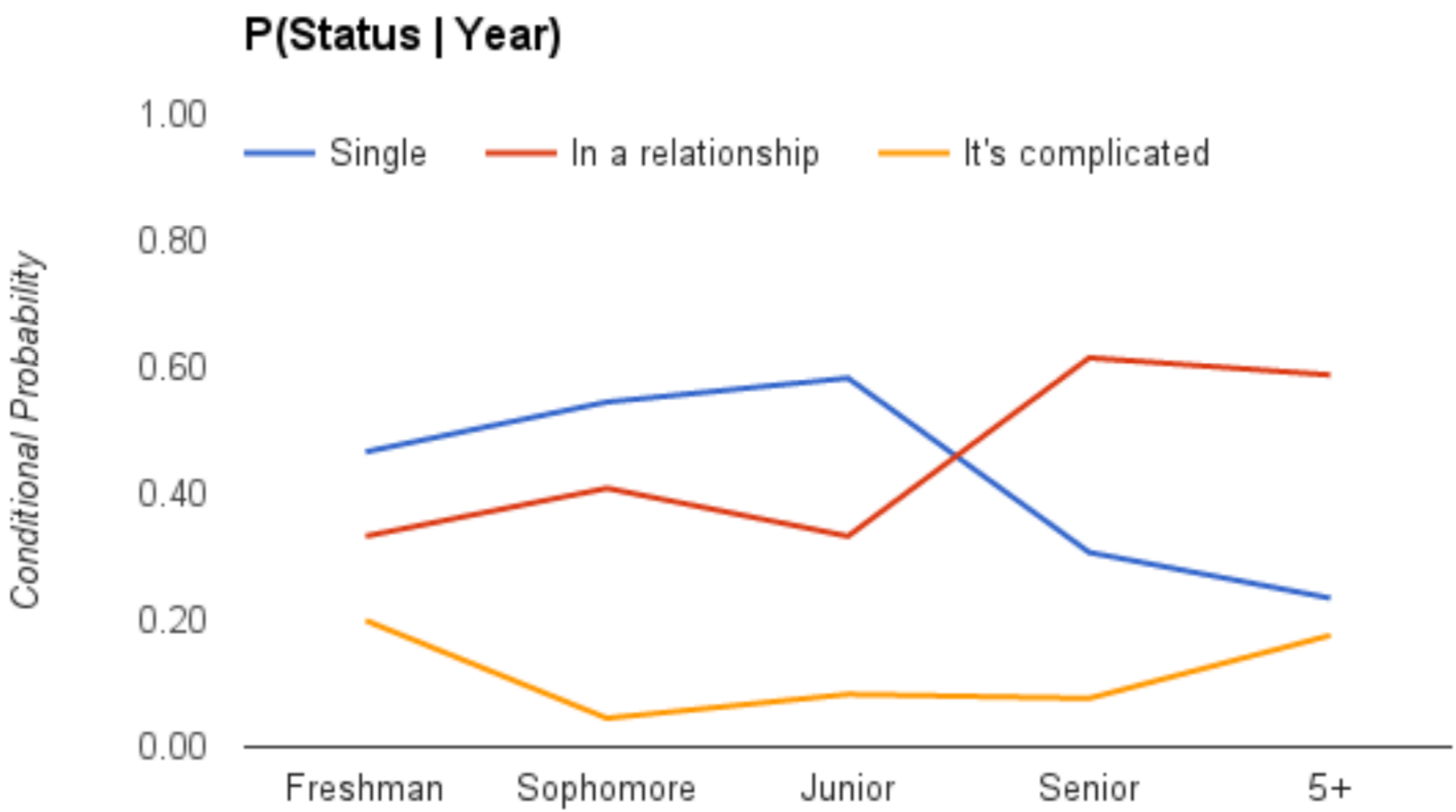


Lunch | Year



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



And It Applies to Books Too

P(Buy Book Y | Bought Book X)

Continuous Conditional Distributions

Let X and Y be continuous random variables

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

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

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

Warmup: Bayes Revisited

Posterior belief

Likelihood of evidence

Prior belief

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

Normalization constant

Mixing Discrete and Continuous

Let X be a continuous random variable

Let N be a discrete random variable

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

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

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

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

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

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

All Continuous

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

Warmup: Bayes Revisited

Posterior belief

Likelihood of evidence

Prior belief

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

Normalization constant

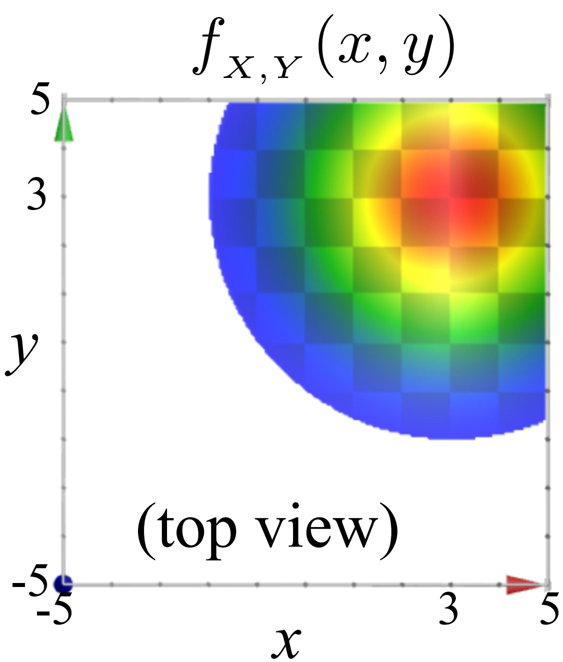
Warmup: Bivariate Normal

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

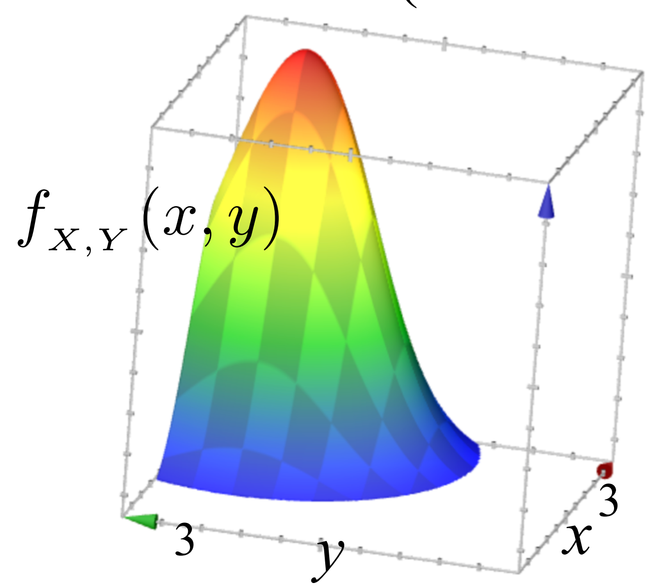
$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{[(x-\mu_x)^2+(y-\mu_y)^2]}{2\cdot\sigma^2}}$$

Here is an example where:

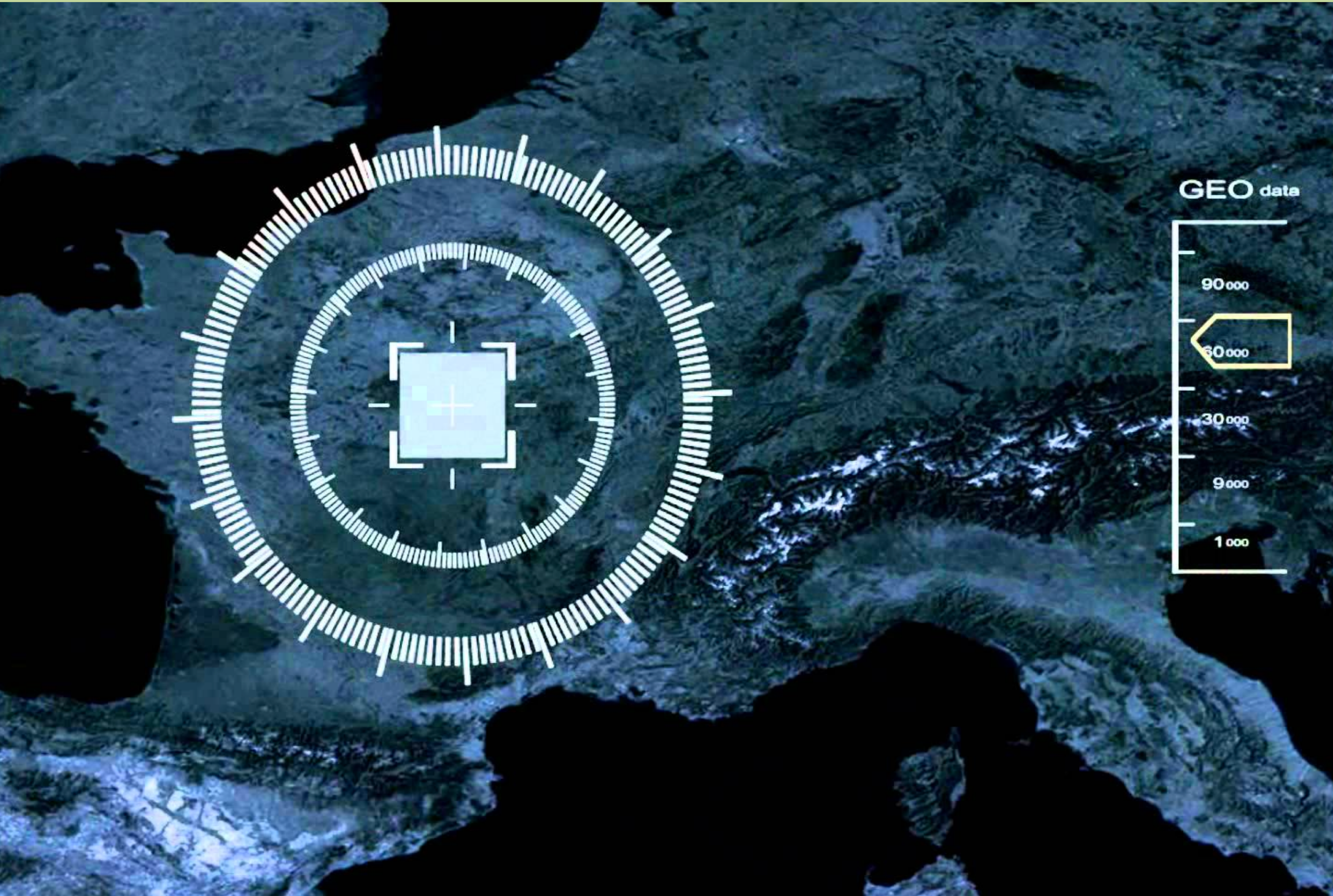
$$\mu_x = 3$$
$$\mu_y = 3$$
$$\sigma = 2$$



(side view)

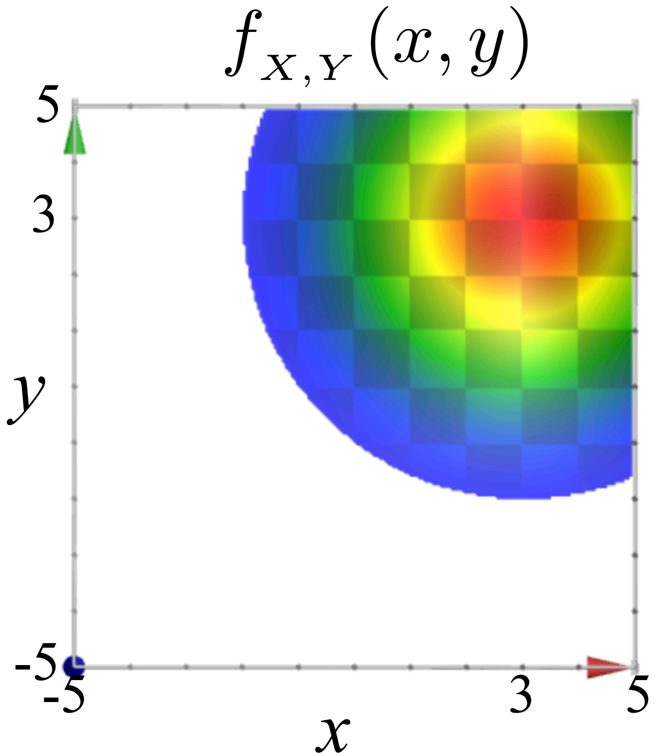


Tracking in 2D Space?



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)

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

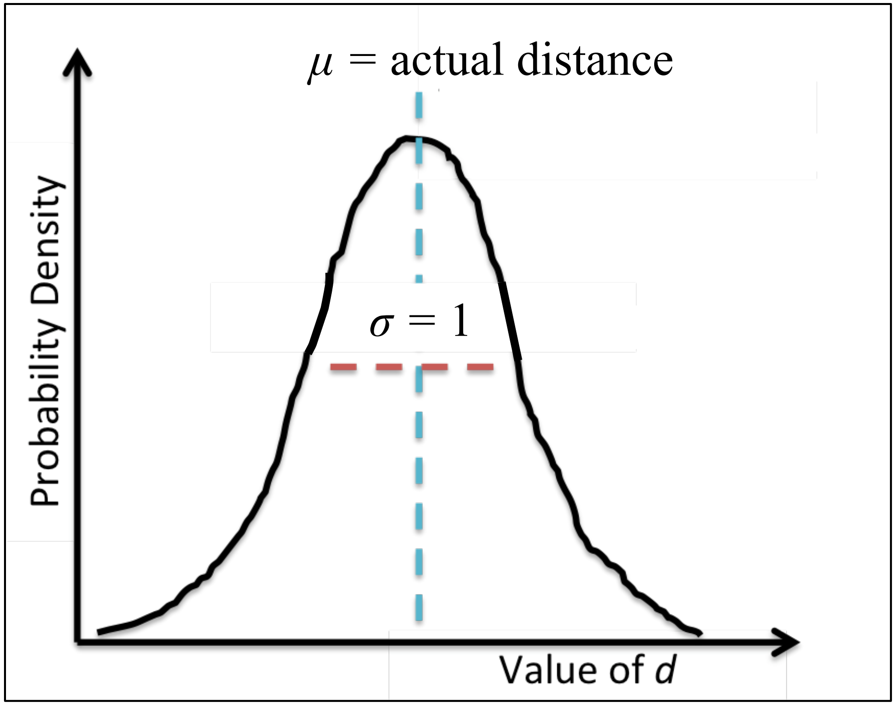
Prior belief with K: $$f_{X,Y}(x,y) = K \cdot e^{-\frac{[(x-3)^2+(y-3)^2]}{8}}$$

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…

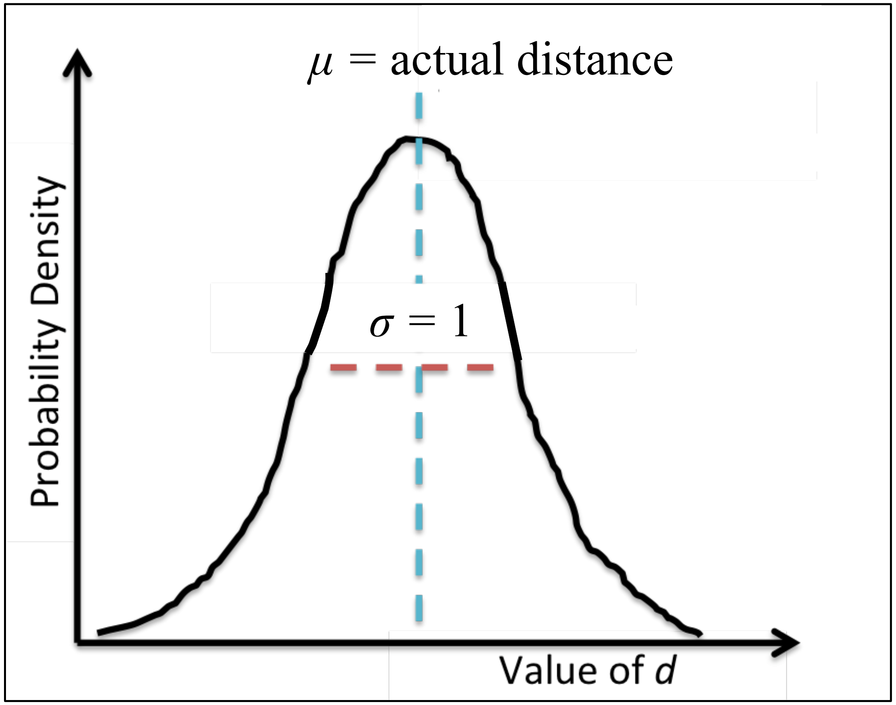
P(D | X, Y) is knowable!



Tracking in 2D Space: Observation!

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

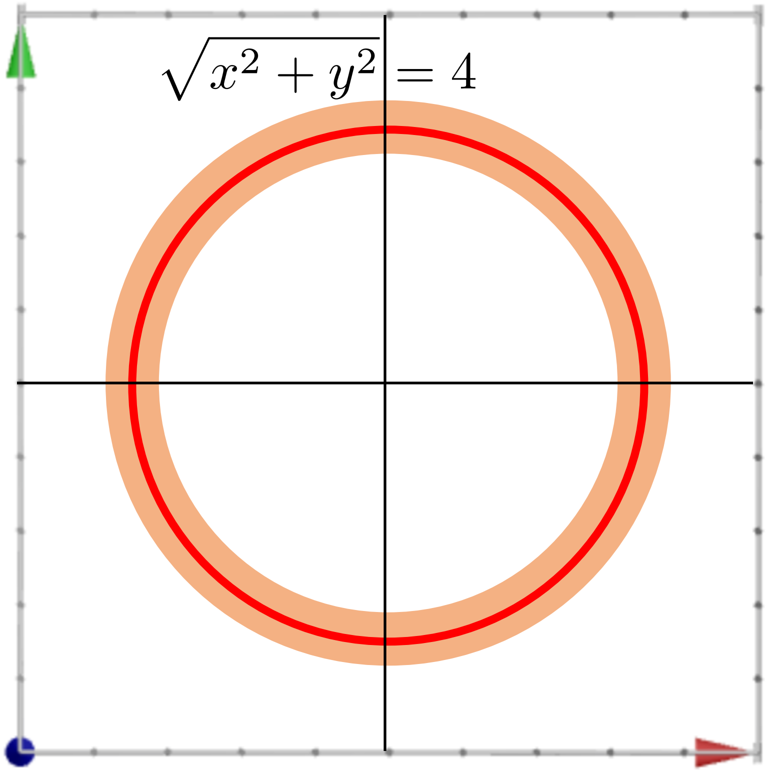
$$D|X, Y \sim N(\mu = \sqrt{x^2 + y^2}, \sigma^2 = 1)$$

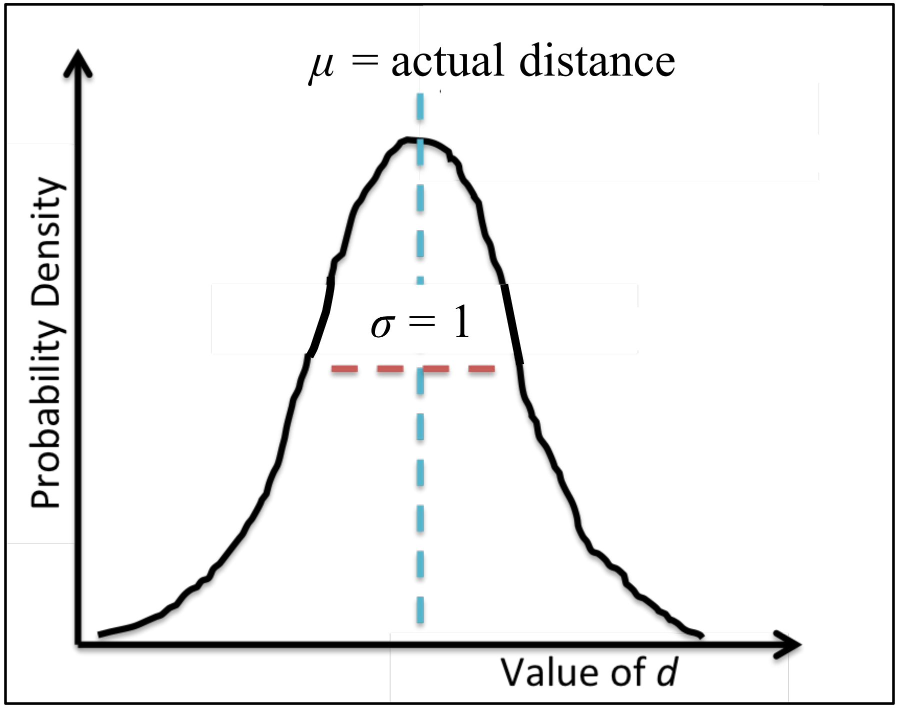


Know that the distance of a ping is normal with respect to the true distance.

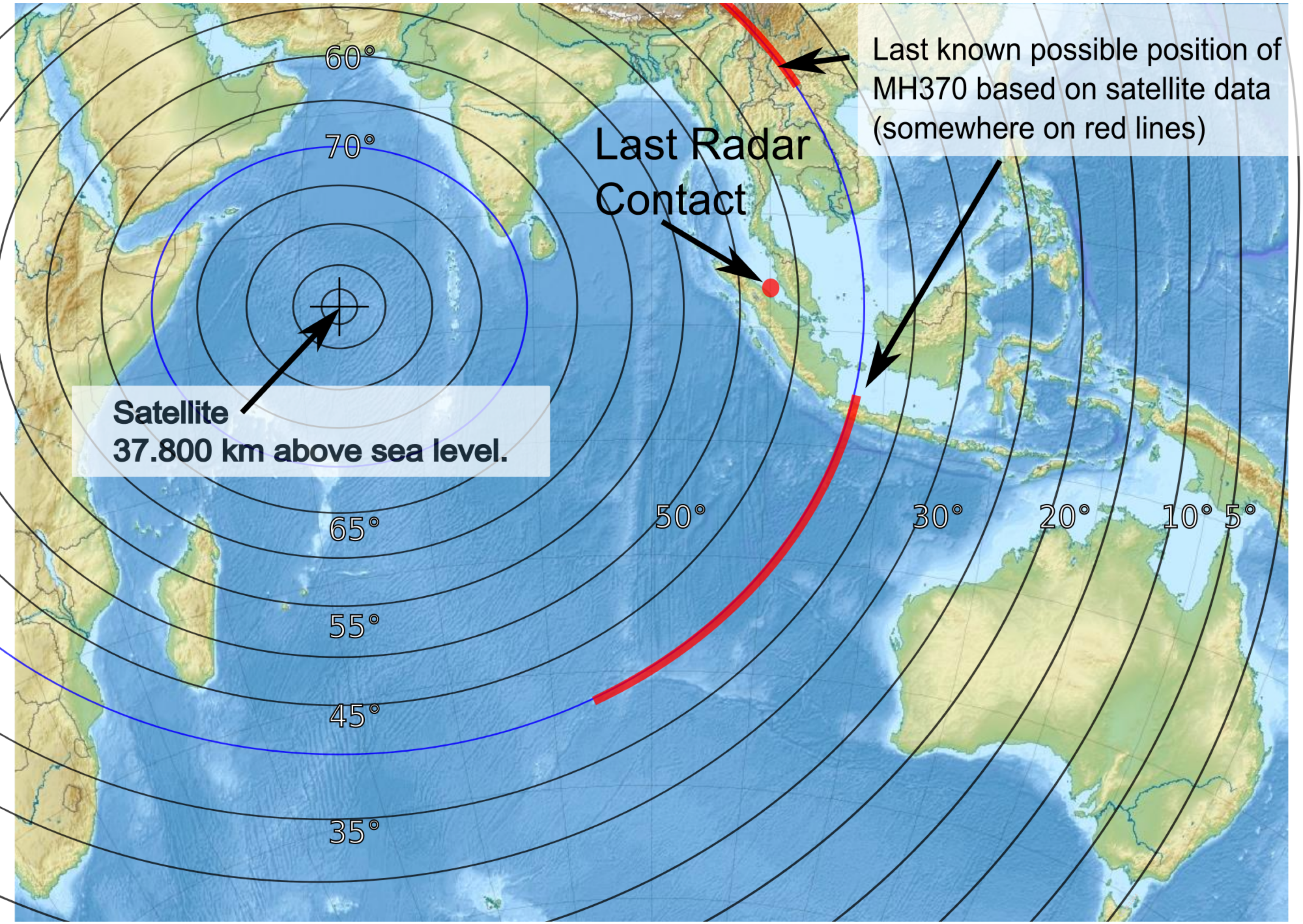
Tracking in 2D Space: Observation!

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



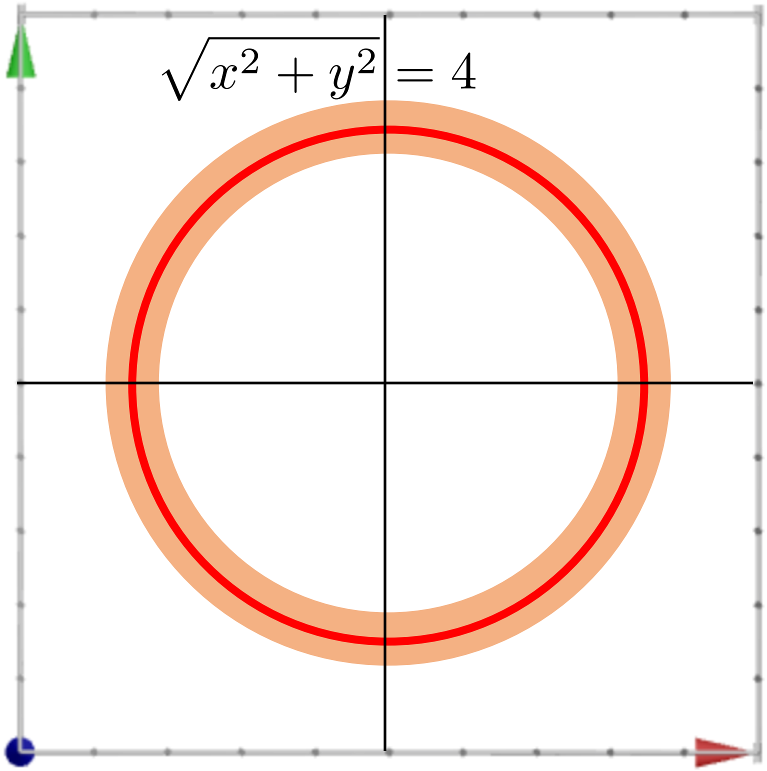


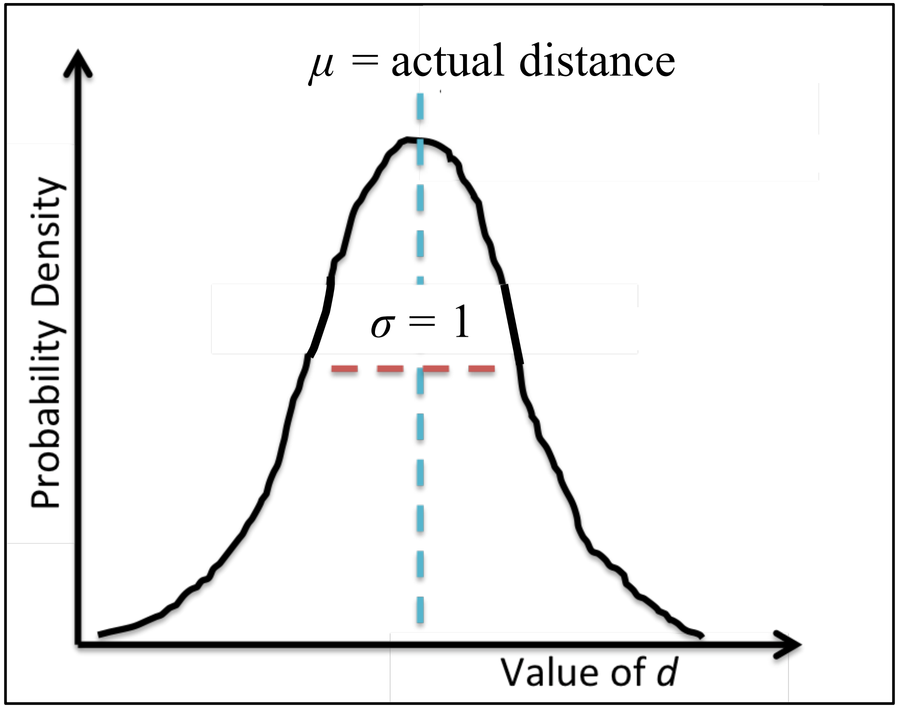
Know that the distance of a ping is normal with respect to the true distance


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!

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

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

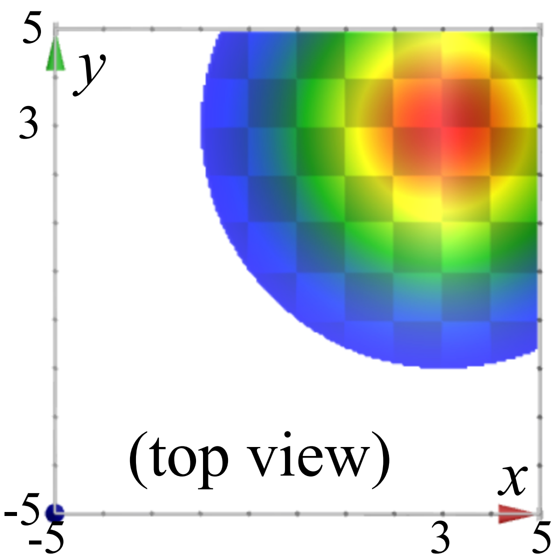
$$= \frac{1}{\sqrt{2\pi}} e^{\frac{-(d-\mu)^2}{2}}$$

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

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

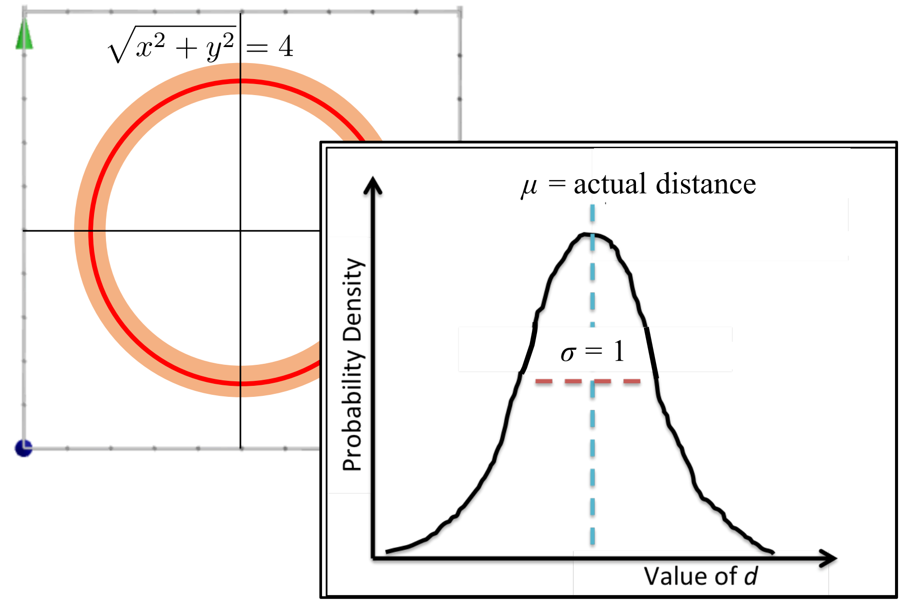
Tracking in 2D Space: New Belief

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



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

Tracking in 2D Space: New Belief

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

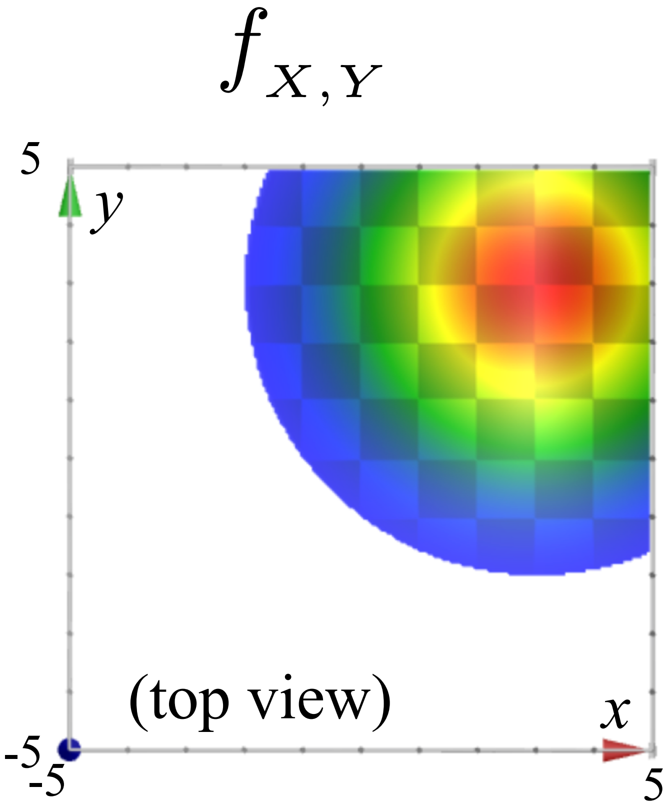
$$=\frac{K_1\cdot e^{-\frac{[4-\sqrt{x^2+y^2})^2]}{2}}\cdot K_2\cdot e^{-\frac{[(x-3)^2+(y-3)^2]}{8}}}{f(D=4)}$$

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

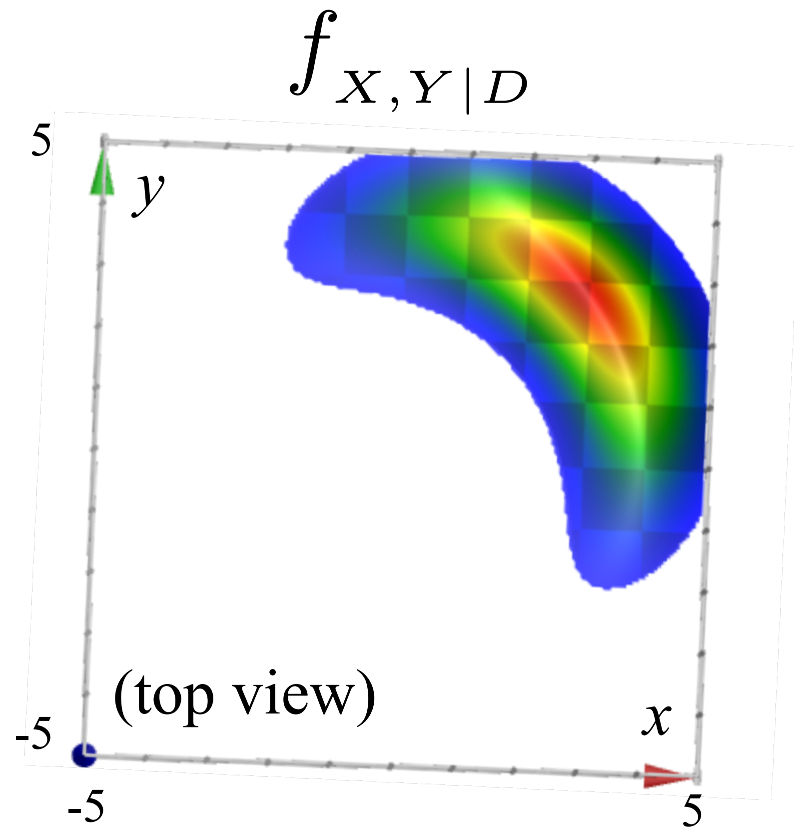
$$=K_4\cdot e^{-\left[\frac{(4-\sqrt{x^2+y^2})^2}{2}+\frac{[(x-3)^2+(y-3)^2]}{8}\right]}$$

For your notes…

Tracking in 2D Space: Posterior

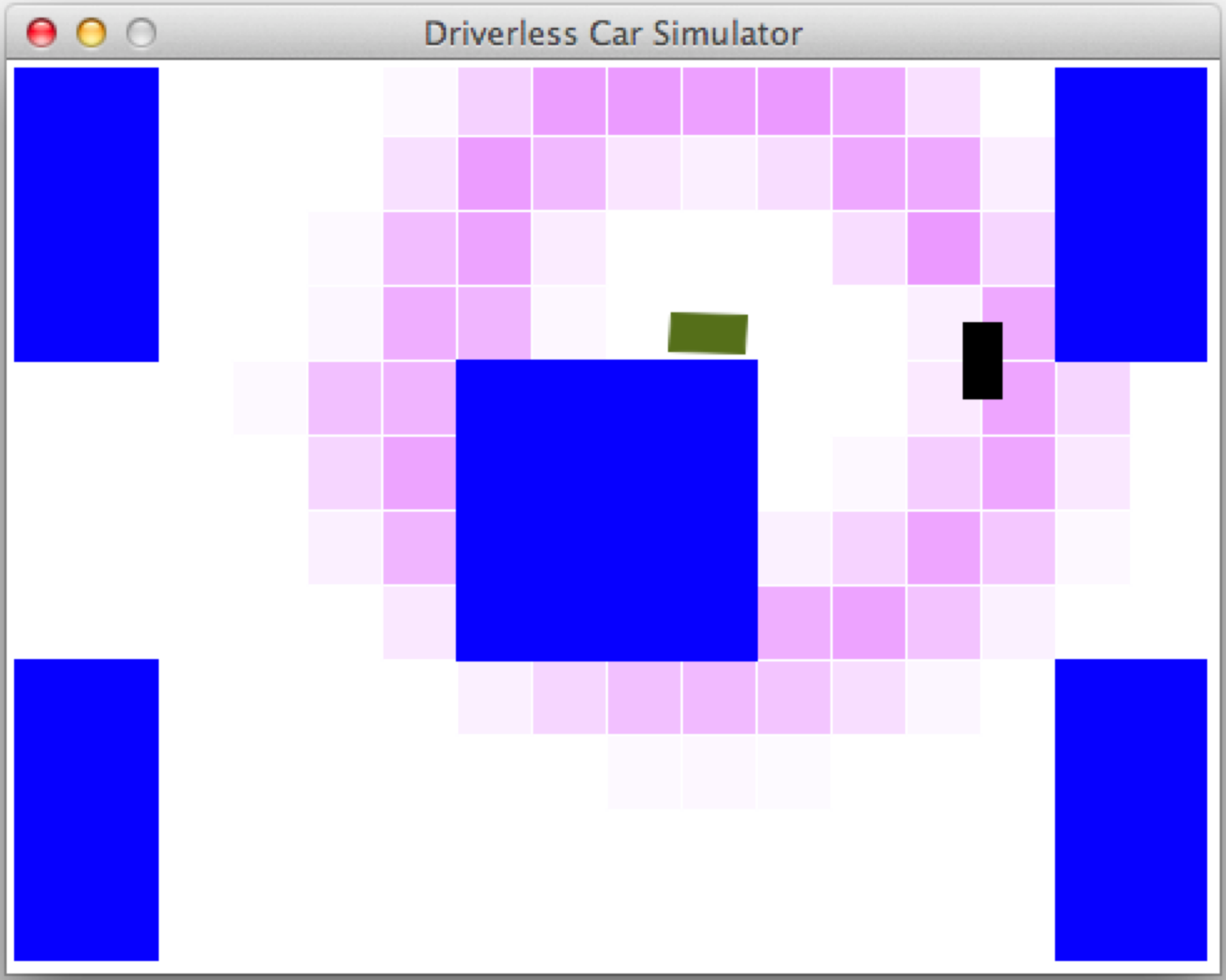


Prior



Posterior

Tracking in 2D Space: CS221



Independence and Random Variables