17: Continuous Joint Distributions (II)

Lisa Yan and Jerry Cain

October 21, 2020

Quick slide reference

Convolution: Sum of independent Uniform RVs

Today's lecture

Take what we've seen in discrete joint distributions…

…and translate them to continuous joint distributions!

For the most part, this is easy. For example:

Marginal distributions	$p_X(a) = \sum_y p_{X,Y}(a,y)$	$f_X(a) = \int_{-\infty}^{\infty} f_{X,Y}(a,y)dy$
Independent RVs	$p_{X,Y}(x,y) = p_X(x)p_Y(y)$	$f_{X,Y}(x,y) = f_X(x)f_Y(y)$

But some concepts, while mathematically straightforward to write, are harder to implement in practice.

We'll focus on these today.

Goal of CS109 continuous joint distributions unit: **build mathematical maturity**

Dance, Dance, Convolution

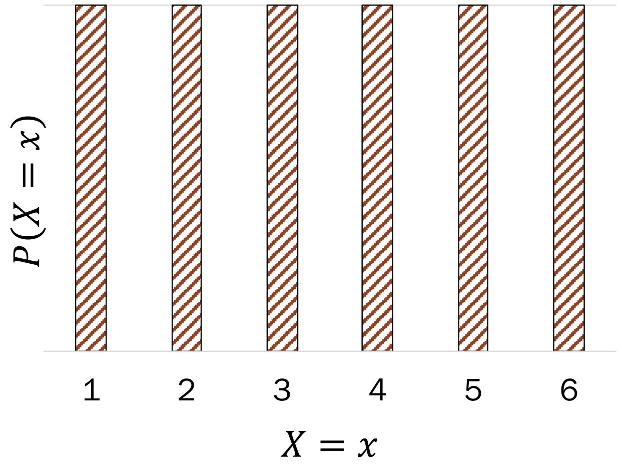
Review

Recall that for independent discrete random variables X and Y:

$P_{X+Y}(n) = \sum_k P_X(k) P_Y(n-k)$

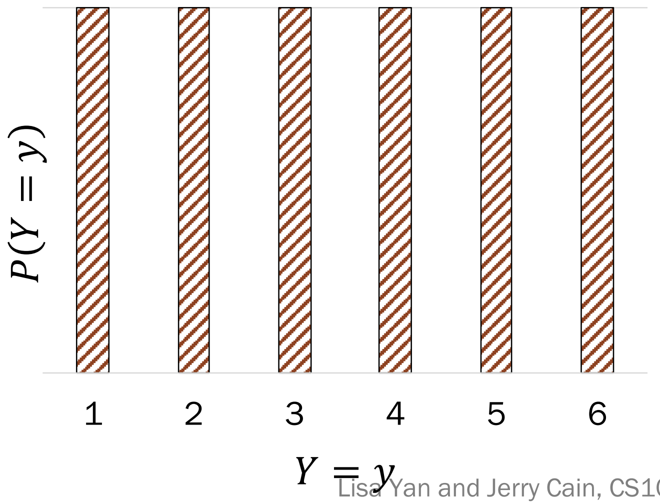
$$P(X+Y=n) = \sum_k P(X=k)P(Y=n-k)$$

the convolution of p_X and p_Y

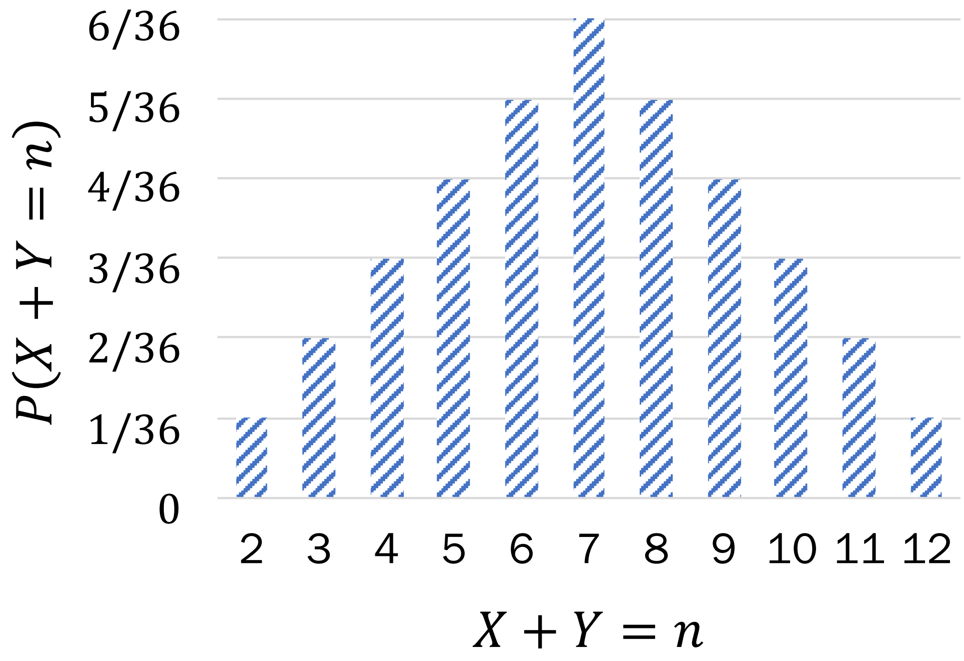


Independent X, Y

+



=



Dance, Dance, Convolution

Recall that for independent discrete random variables X and Y:

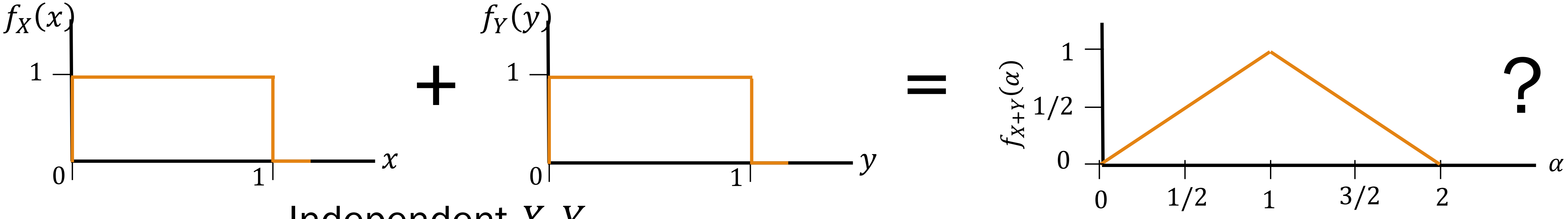
$$P(X+Y=n)=\sum_k P(X=k)P(Y=n-k)$$

the convolution of p_X and p_Y

For independent continuous random variables X and Y:

$$f_{X+Y}(\alpha)=\int_{-\infty}^{\infty} f_X(x)f_Y(\alpha-x)dx$$

the convolution of f_X and f_Y



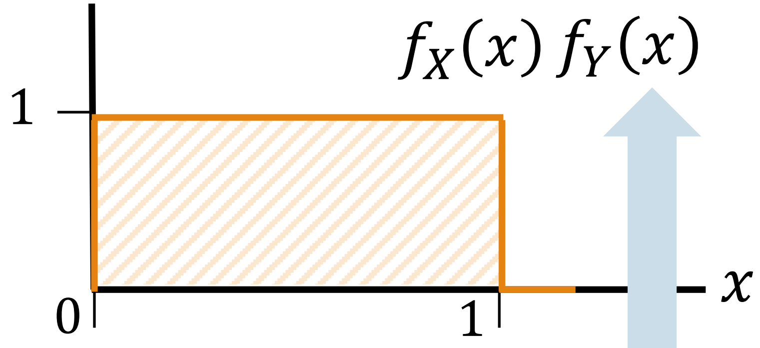
Independent X, Y

Dance, Dance, Convolution Extreme

Sum of independent Uniforms

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs.
What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

$$f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x) dx$$



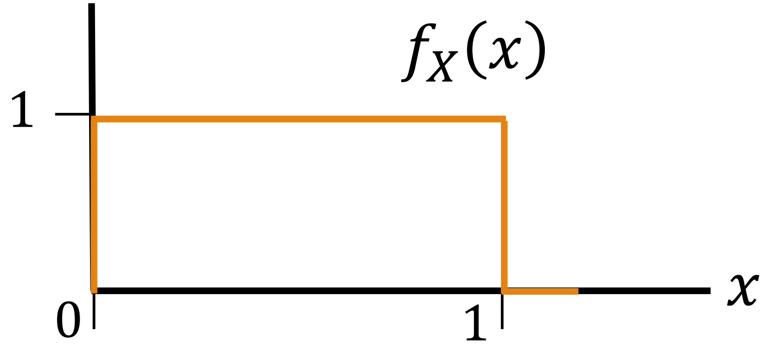
Not so fast…

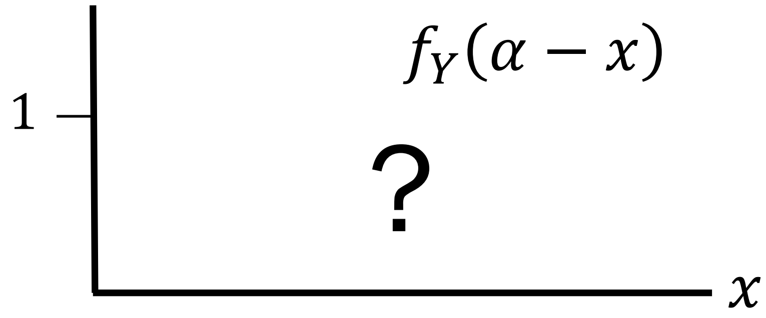
Isn't this just one??

Sum of independent Uniforms

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs.
What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

$$f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x)\, dx$$





$$f_X(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(\alpha - x) = \begin{cases} 1 & \text{if } 0 \le \alpha - x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} 1 & \text{if } \alpha - 1 \le x \le \alpha \\ 0 & \text{otherwise} \end{cases}$$

$-\alpha \le -x \le 1-\alpha$

$\alpha - 1 \le x \le \alpha$

α is a constant in the integral w.r.t. x.

Sum of independent Uniforms

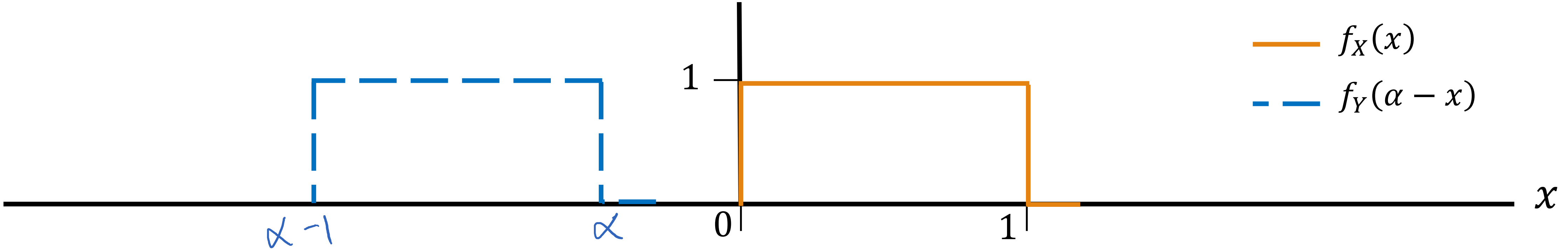
X and Y independent + continuous $\quad f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x)\, dx$

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(\alpha - x) = \begin{cases} 1 & \text{if } \alpha - 1 \le x \le \alpha \\ 0 & \text{otherwise} \end{cases}$$

1. $\alpha \le 0$ $\qquad 0$



Sum of independent Uniforms

X and Y independent + continuous $\quad f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x)\, dx$

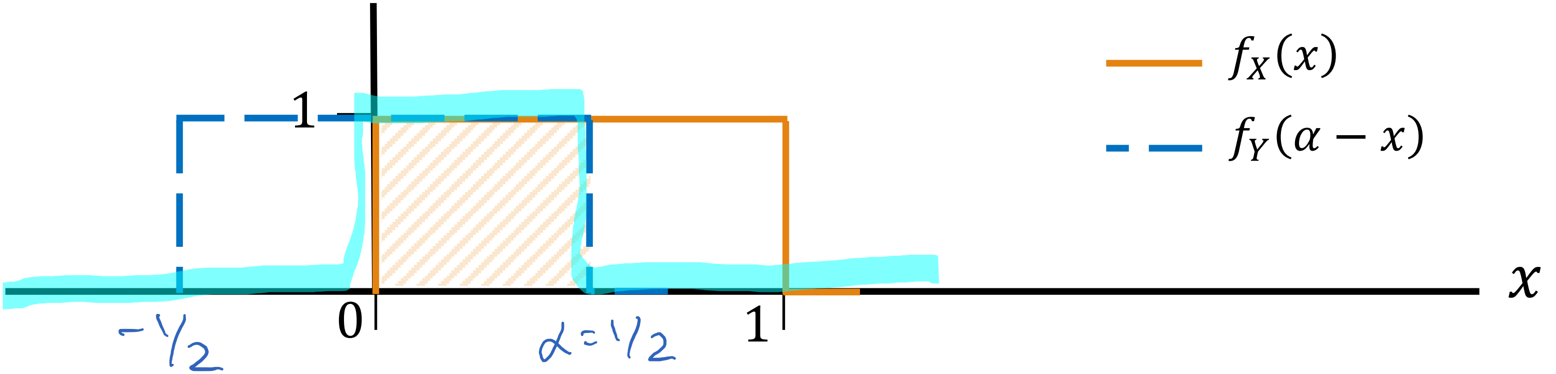
Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(\alpha - x) = \begin{cases} 1 & \text{if } \alpha - 1 \le x \le \alpha \\ 0 & \text{otherwise} \end{cases}$$

1. $\alpha \le 0$ $\quad 0$

2. $\alpha = 1/2$ $\quad 1/2$



Integral = area under the curve
This curve = product of 2 functions of x

Sum of independent Uniforms

X and Y independent + continuous $\quad f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x)\, dx$

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs.
What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(\alpha - x) = \begin{cases} 1 & \text{if } \alpha - 1 \le x \le \alpha \\ 0 & \text{otherwise} \end{cases}$$

1. $\alpha \le 0$ $\quad 0$
2. $\alpha = 1/2$ $\quad 1/2$
3. $\alpha = 1$
4. $\alpha = 3/2$
5. $\alpha \ge 2$

Sum of independent Uniforms

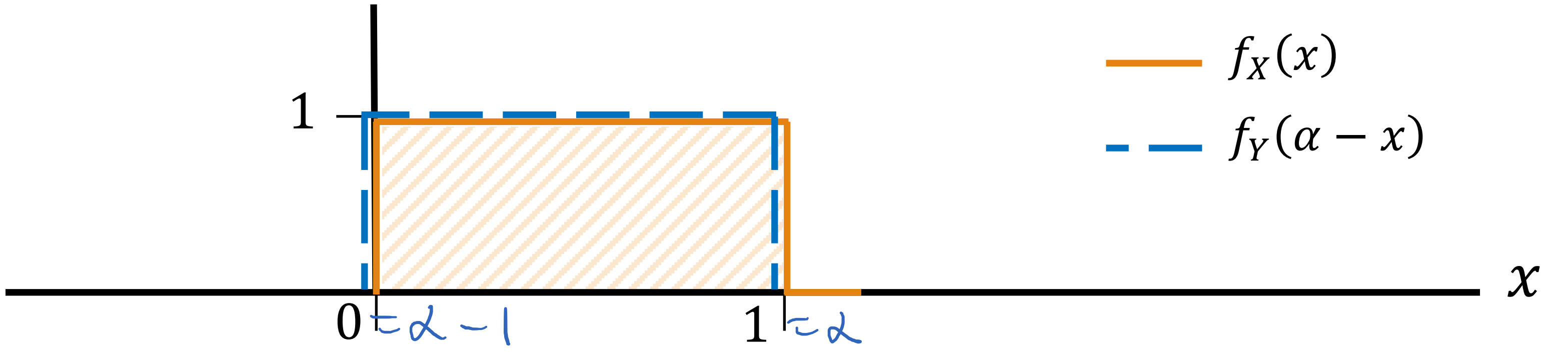
X and Y independent + continuous $\quad f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x)\, dx$

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X+Y$, $f_{X+Y}(\alpha)$?

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(\alpha - x) = \begin{cases} 1 & \text{if } \alpha - 1 \le x \le \alpha \\ 0 & \text{otherwise} \end{cases}$$

1. $\alpha \le 0$ 0
2. $\alpha = 1/2$ $1/2$
3. $\alpha = 1$ 1
4. $\alpha = 3/2$
5. $\alpha \ge 2$



Sum of independent Uniforms

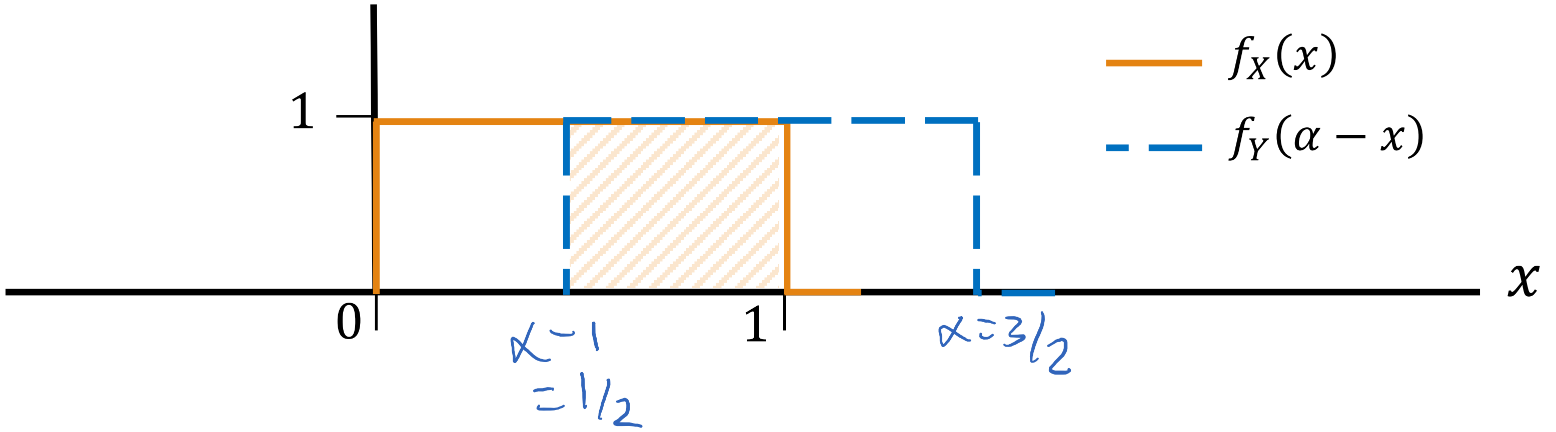
X and Y independent + continuous $\quad f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x)\, dx$

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs. What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(\alpha - x) = \begin{cases} 1 & \text{if } \alpha - 1 \le x \le \alpha \\ 0 & \text{otherwise} \end{cases}$$

1. $\alpha \le 0$ $\quad 0$
2. $\alpha = 1/2$ $\quad 1/2$
3. $\alpha = 1$ $\quad 1$
4. $\alpha = 3/2$ $\quad 1/2$
5. $\alpha \ge 2$



Sum of independent Uniforms

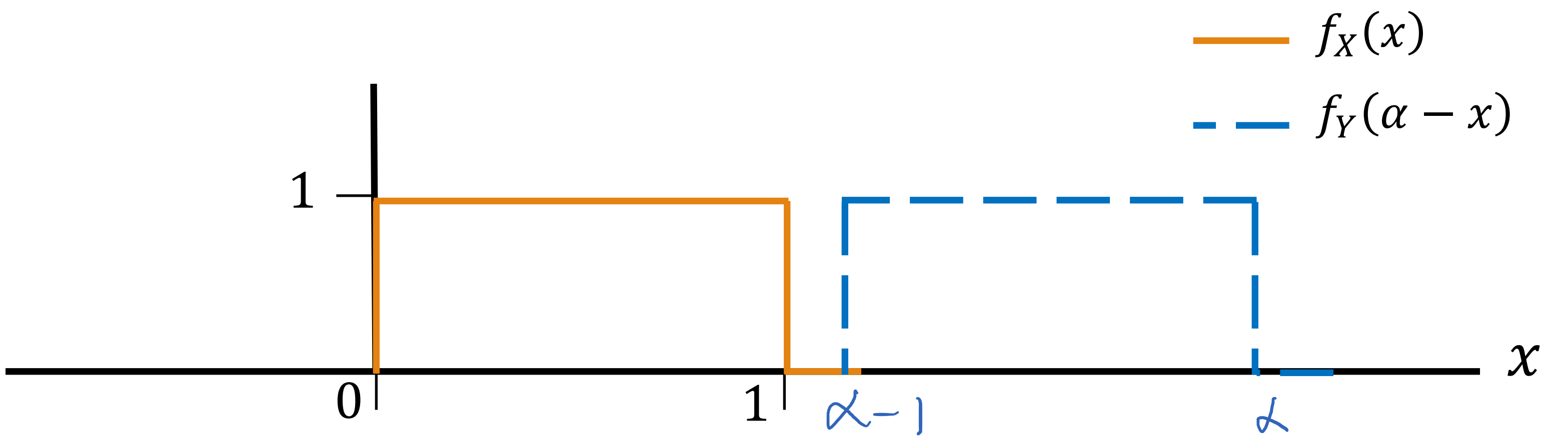
X and Y independent + continuous $f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x)\, dx$

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs.
What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(\alpha - x) = \begin{cases} 1 & \text{if } \alpha - 1 \le x \le \alpha \\ 0 & \text{otherwise} \end{cases}$$

1. $\alpha \le 0$ $\quad 0$
2. $\alpha = 1/2$ $\quad 1/2$
3. $\alpha = 1$ $\quad 1$
4. $\alpha = 3/2$ $\quad 1/2$
5. $\alpha \ge 2$ $\quad 0$

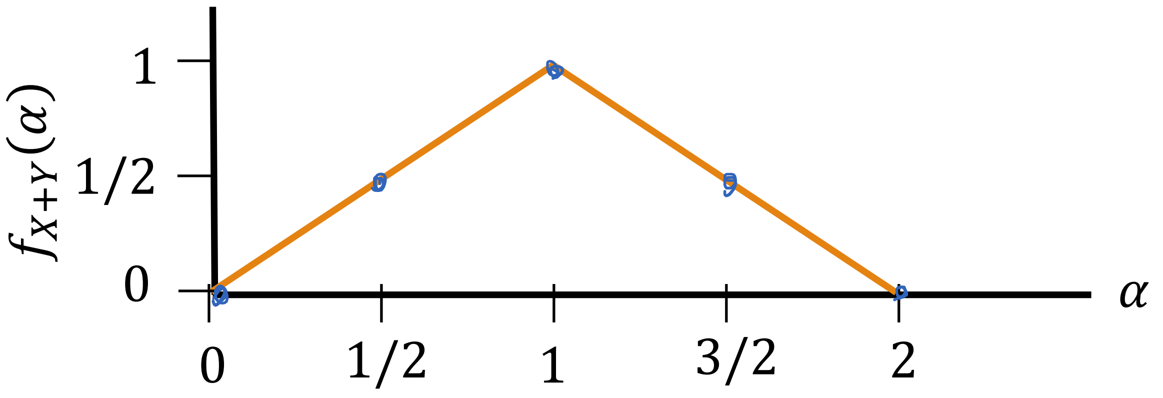


Sum of independent Uniforms

X and Y independent + continuous $\quad f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x) f_Y(\alpha - x)\, dx$

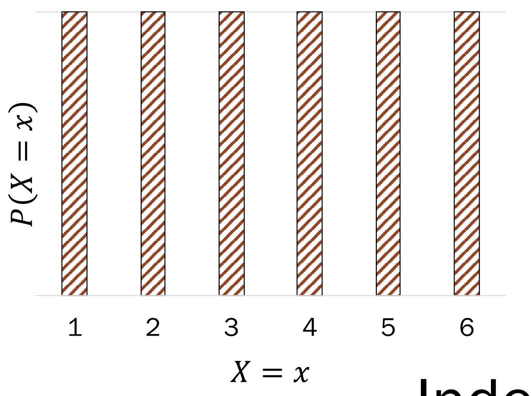
Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent RVs.
What is the distribution of $X + Y$, $f_{X+Y}(\alpha)$?

1. $\alpha \le 0$ $\quad 0$
2. $\alpha = 1/2$ $\quad 1/2$
3. $\alpha = 1$ $\quad 1$
4. $\alpha = 3/2$ $\quad 1/2$
5. $\alpha \ge 2$ $\quad 0$

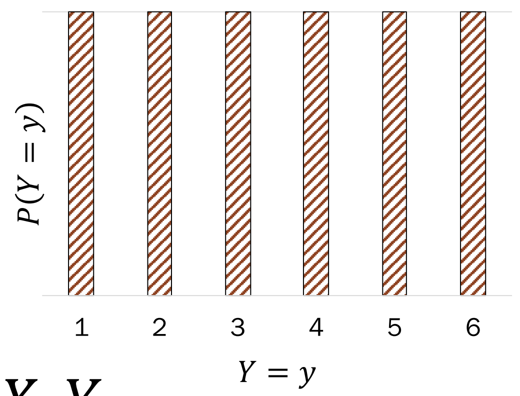


$$f_{X+Y}(\alpha) = \begin{cases} \alpha & 0 \le \alpha \le 1 \\ 2 - \alpha & 1 \le \alpha \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Dance, Dance, Convolution Extreme

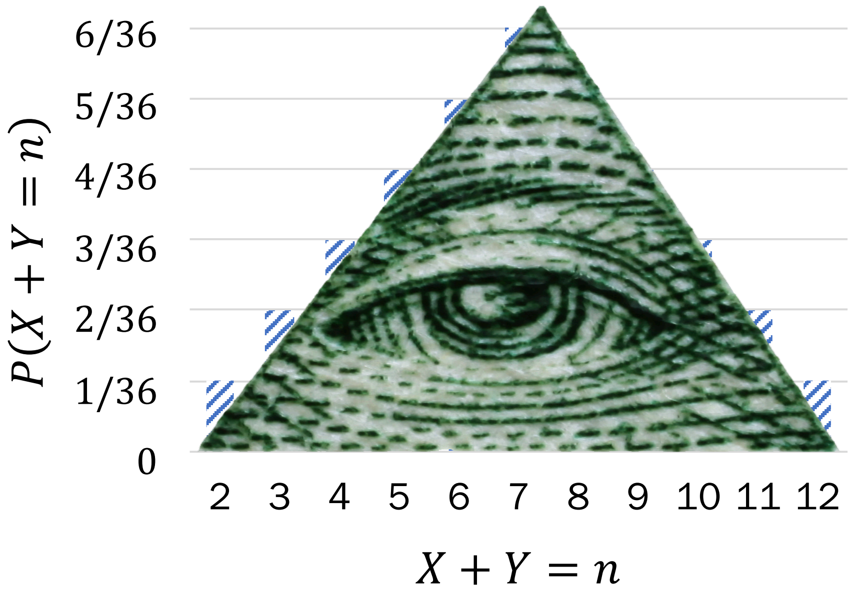


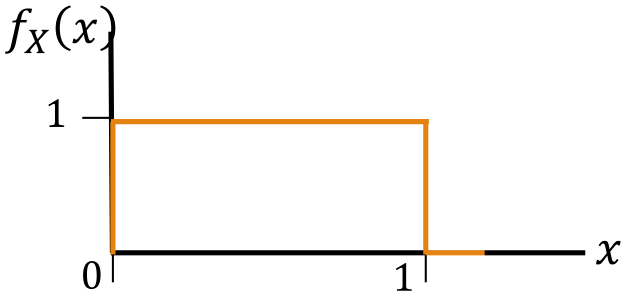
+



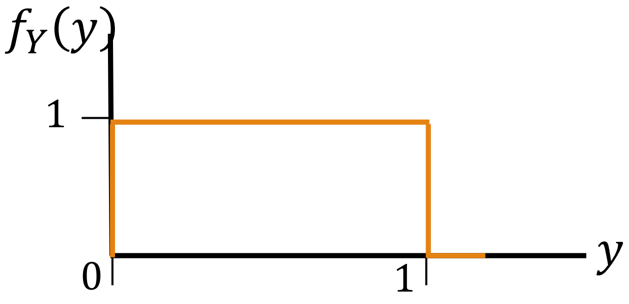
Independent X, Y

=



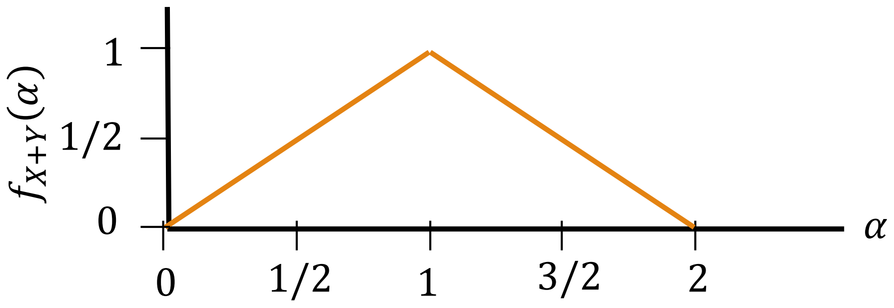


+



Independent X, Y

=



Dance, Dance, Convolution Extreme

Phew....that was a mental workout.

In practice, we try to avoid convolution where possible, by choosing “nice” distributions.

Ready for something truly useful? Stay tuned!

Sums of independent Normal RVs

Sum of independent Normals

$X \sim \mathcal{N}(\mu_1, \sigma_1^2)$,
$Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$
X, Y independent

$$\Rightarrow \quad X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

(proof left to Wikipedia)

Holds in general case:

$X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$
X_i independent for $i = 1, \dots, n$

$$\Rightarrow \quad \sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

Back for another playoffs game

What is the probability that the Warriors win?
How do you model zero-sum games?

$$P(A_W > A_B)$$

This is a probability of an event involving ***two*** random variables!

We will compute:

$$P(A_W - A_B > 0)$$

Motivating idea: Zero sum games

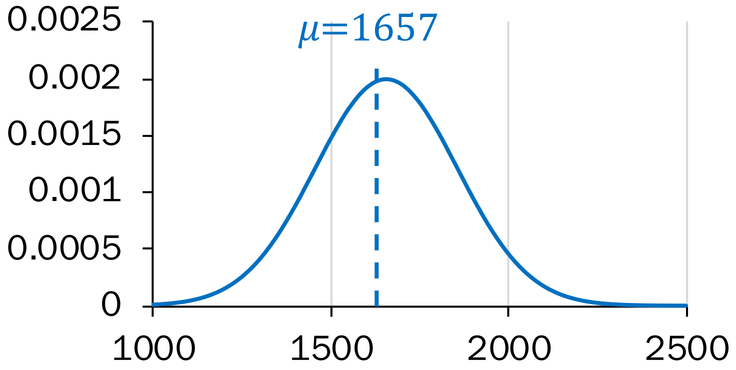
Want: $P(\text{Warriors win}) = P(A_W - A_B > 0)$

Assume A_W, A_B are independent.
Let $D = A_W - A_B$.

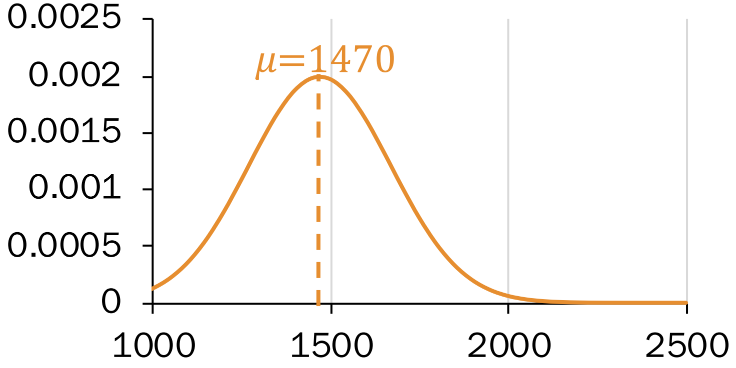
What is the distribution of D?

A. $D \sim \mathcal{N}(1657 - 1470,\ 200^2 - 200^2)$
B. $D \sim \mathcal{N}(1657 - 1470,\ 200^2 + 200^2)$
C. $D \sim \mathcal{N}(1657 + 1470,\ 200^2 + 200^2)$
D. Dance, Dance, Convolution
E. None/other

Warriors $A_W \sim \mathcal{N}(S = 1657, 200^2)$



Opponents $A_B \sim \mathcal{N}(S = 1470, 200^2)$



Motivating idea: Zero sum games

Want: $P(\text{Warriors win}) = P(A_W - A_B > 0)$

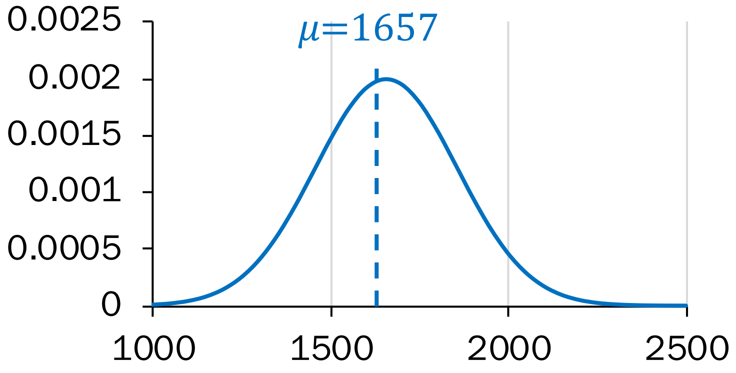
Assume A_W, A_B are independent.
Let $D = A_W - A_B$.

$+ -A_B \sim N(-1470, +200^2)$

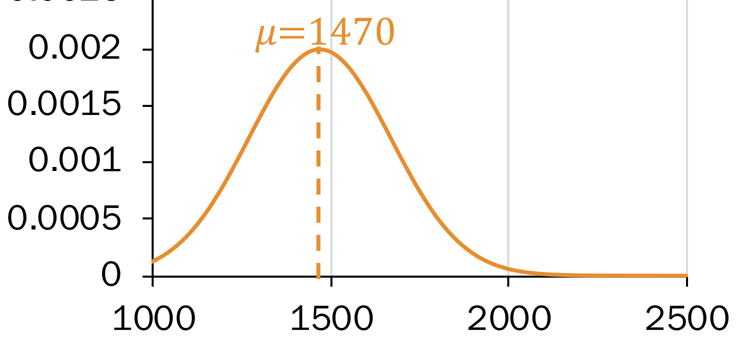
What is the distribution of D?

A. $D \sim \mathcal{N}(1657 - 1470,\ 200^2 - 200^2)$
B. $D \sim \mathcal{N}(1657 - 1470,\ 200^2 + 200^2)$
C. $D \sim \mathcal{N}(1657 + 1470,\ 200^2 + 200^2)$
D. Dance, Dance, Convolution
E. None/other

Warriors $A_W \sim \mathcal{N}(S = 1657, 200^2)$



Opponents $A_B \sim \mathcal{N}(S = 1470, 200^2)$



If $X \sim \mathcal{N}(\mu, \sigma^2)$,
then $(-X) \sim \mathcal{N}(-\mu, (-1)^2\sigma^2 = \sigma^2)$.

Motivating idea: Zero sum games

Want: $P(\text{Warriors win}) = P(A_W - A_B > 0)$

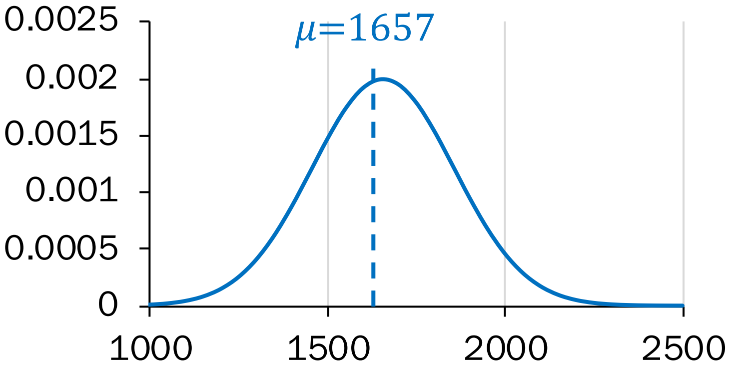
Assume A_W, A_B are independent.
Let $D = A_W - A_B$.

$$D \sim \mathcal{N}(1657 - 1470, \;\; 200^2 + 200^2)$$
$$\sim \mathcal{N}(187, 2 \cdot 200^2) \qquad \sigma \approx 283$$

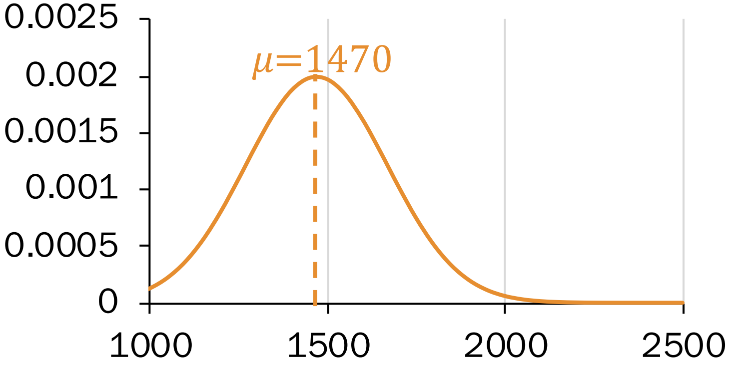
$$P(D > 0) = 1 - F_D(0) = 1 - \Phi\left(\frac{0 - 187}{283}\right)$$
$$\approx 0.7454$$

Compare with 0.7488, calculated by sampling!

Warriors $A_W \sim \mathcal{N}(S = 1657, 200^2)$



Opponents $A_B \sim \mathcal{N}(S = 1470, 200^2)$



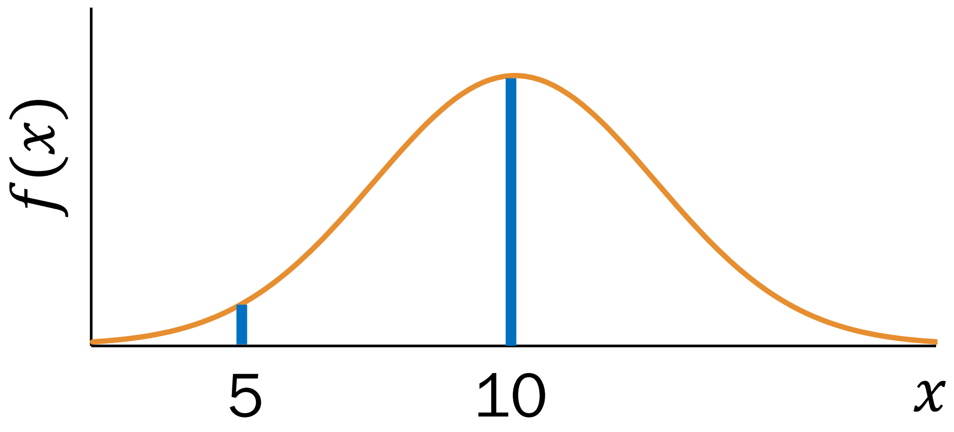
Ratio of PDFs

Relative probabilities of continuous random variables

Let X = time to finish problem set 4.

Suppose $X \sim \mathcal{N}(10,2)$.

How much *more likely* are you to complete in 10 hours than 5 hours?



$$\frac{P(X = 10)}{P(X = 5)} =$$

A. 0/0 = undefined

B. $\frac{f(10)}{f(5)}$

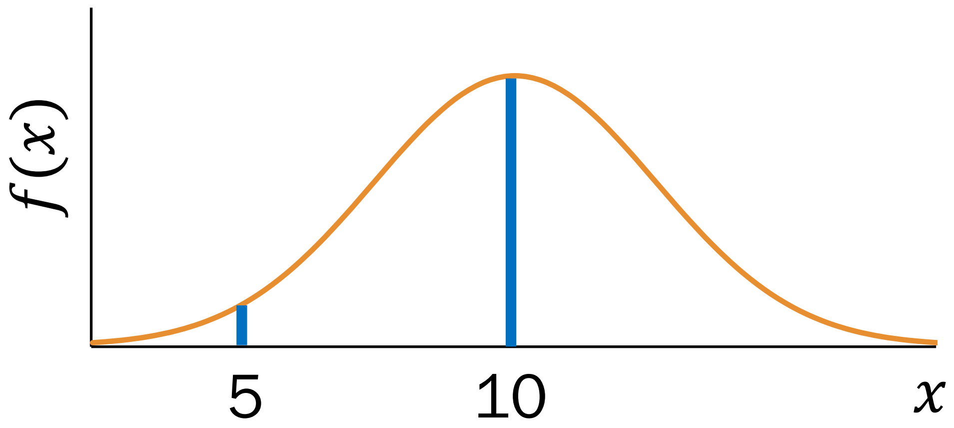
C. stay healthy

Relative probabilities of continuous random variables

Let X = time to finish problem set 4.

Suppose $X \sim \mathcal{N}(10, 2)$.

How much *more likely* are you to complete in 10 hours than 5 hours?



$$\frac{P(X = 10)}{P(X = 5)} =$$

A. 0/0 = undefined

B. $\dfrac{f(10)}{f(5)}$

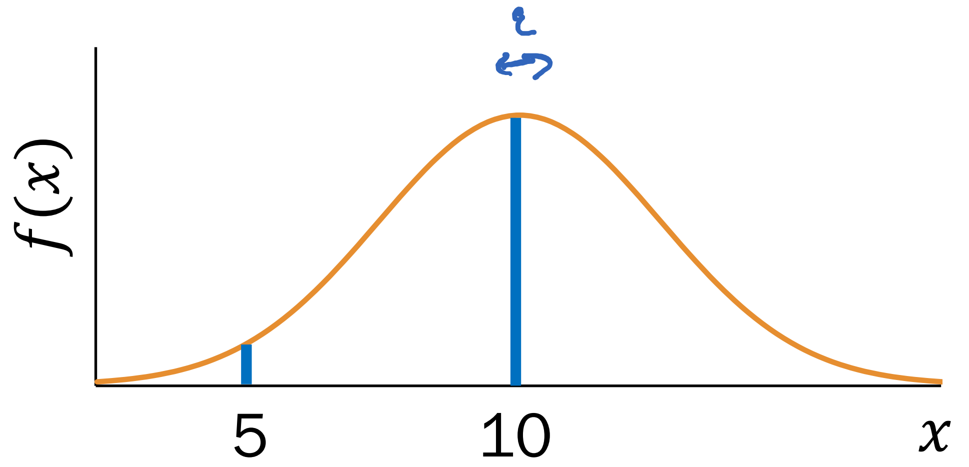
C. stay healthy

Relative probabilities of continuous random variables

Let X = time to finish problem set 4.

Suppose $X \sim \mathcal{N}(10, 2)$.

How much *more likely* are you to complete in 10 hours than 5 hours?



$$\frac{P(X=10)}{P(X=5)} = \frac{f(10)}{f(5)}$$

$$P(X=a) = P\left(a - \frac{\varepsilon}{2} \le X \le a + \frac{\varepsilon}{2}\right) = \int_{a-\frac{\varepsilon}{2}}^{a+\frac{\varepsilon}{2}} f(x)dx \approx \varepsilon f(a)$$

Therefore $\frac{P(X=a)}{P(X=b)} = \frac{\varepsilon f(a)}{\varepsilon f(b)} = \frac{f(a)}{f(b)}$

$\lim_{\varepsilon \to 0} \frac{\varepsilon f(a)}{\varepsilon f(b)}$

$$= \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(10-\mu)^2}{2\sigma^2}}}{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(5-\mu)^2}{2\sigma^2}}} = \frac{e^{-\frac{(10-10)^2}{2\cdot 2}}}{e^{-\frac{(5-10)^2}{2\cdot 2}}} = \frac{e^0}{e^{-\frac{25}{4}}} = 518$$

Ratios of PDFs are meaningful!!

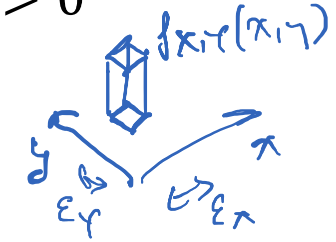
Continuous conditional distributions

Continuous conditional distributions

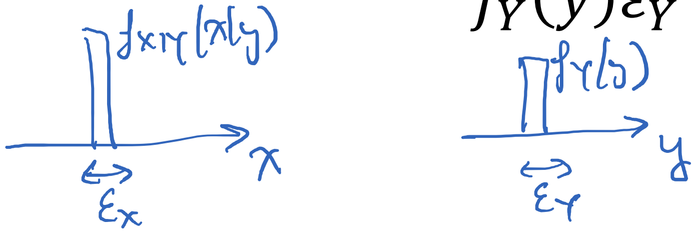
For continuous RVs X and Y, the conditional PDF of X given Y is

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

where $f_Y(y) > 0$



Intuition: $P(X = x|Y = y) = \dfrac{P(X = x, Y = y)}{P(Y = y)}$ ⟺ $f_{X|Y}(x|y)\varepsilon_X = \dfrac{f_{X,Y}(x,y)\varepsilon_X\varepsilon_Y}{f_Y(y)\varepsilon_Y}$



Note that conditional PDF $f_{X|Y}$ is a "true" density:

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx = \int_{-\infty}^{\infty} \frac{f_{X,Y}(x,y)}{f_Y(y)}dx = \frac{f_Y(y)}{f_Y(y)} = 1$$

$$\sum_x p_{X|Y}(x|y) = 1$$

(live) 17: Continuous Joint Distributions (II)

Lisa Yan and Jerry Cain

October 21, 2020

Why sums of random variables?

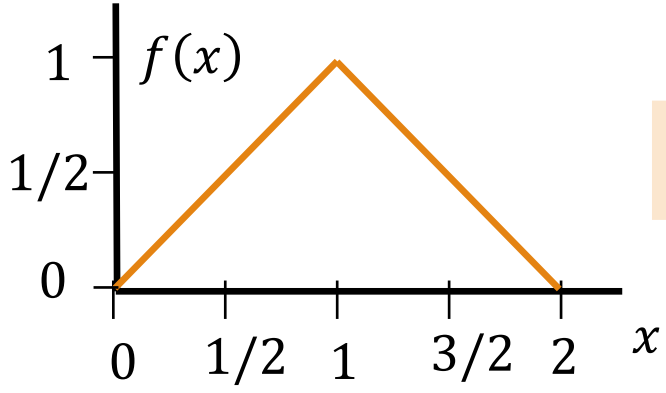
$X_1 + X_2$

Sometimes modeling and <u>understanding</u> a complex X is hard.

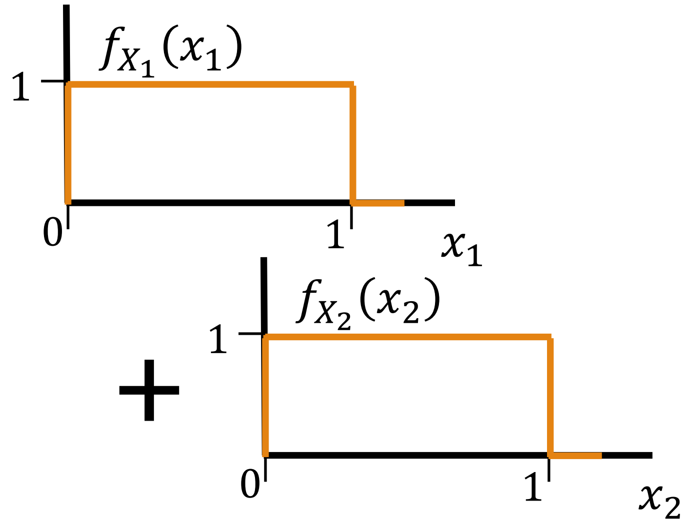
But if we can decompose X into the **sum of independent simpler** RVs,

- We can then compute distributions on X.
- We can then to understand how X changes when its parts change.

What can we model with a triangular PDF?



Sum of uniforms!



We're covering the reverse direction for now; the forward direction will come next time

Discussion

Slide 3~~6~~4 has a question to discuss together.

Post any clarifications here or in chat!

https://us.edstem.org/courses/2678/discussion/153772

Think by yourself: 1 min

Discuss (as a class, in chat): 3 min

(discuss)

Virus infections

Suppose you are working with the WHO to plan a response to the initial conditions of a virus. There are two exposed groups:

- G1: 200 people, each independently infected with $p_1 = 0.1$
- G2: 100 people, each independently infected with $p_2 = 0.4$

What is $P(\text{people infected} \geq 55)$? An approximation is okay.

1. Define RVs
   & state goal

Let A = # infected in G1.

$A \sim \text{Bin}(200, 0.1)$

B = # infected in G2.

$B \sim \text{Bin}(100, 0.4)$

Want: $P(A + B \geq 55)$

Strategy:

A. Dance, Dance, Convolution
B. Sum of indep. Binomials
C. (approximate) Sum of indep. Poissons
D. (approximate) Sum of indep. Normals
E. None/other

1 min think, a few min discuss

$\sum_k p(k)\,p(n-k)$

$\text{Bin}(n_1, p) + \text{Bin}(n_2, p)$

large n, small p

large enough variance $np(1-p) > 10$

(discuss)

Virus infections

Suppose you are working with the WHO to plan a response to the initial conditions of a virus. There are two exposed groups:

- G1: 200 people, each independently infected with $p_1 = 0.1$
- G2: 100 people, each independently infected with $p_2 = 0.4$

What is $P(\text{people infected} \geq 55)$? An approximation is okay.

1. Define RVs & state goal

Let $A = \#$ infected in G1.

$A \sim \text{Bin}(200, 0.1)$

$B = \#$ infected in G2.

$B \sim \text{Bin}(100, 0.4)$

Want: $P(A + B \geq 55)$

Strategy:

A. Dance, Dance, Convolution
B. Sum of indep. Binomials
C. (approximate) Sum of indep. Poissons
D. (approximate) Sum of indep. Normals
E. None/other

Virus infections

approx Bin w/ Normal
discrete → continuity correction

Suppose you are working with the WHO to plan a response to the initial conditions of a virus. There are two exposed groups:

- G1: 200 people, each independently infected with $p_1 = 0.1$
- G2: 100 people, each independently infected with $p_2 = 0.4$

What is $P(\text{people infected} \geq 55)$?

1. Define RVs & state goal

Let A = # infected in G1.
$A \sim \text{Bin}(200, 0.1)$
B = # infected in G2.
$B \sim \text{Bin}(100, 0.4)$

Want: $P(A + B \geq 55)$

2. Approximate as sum of Normals

$$A \approx X \sim \mathcal{N}(20, 18) \qquad B \approx Y \sim \mathcal{N}(40, 24)$$

$$P(A + B \geq 55) \approx P(X + Y \geq 54.5)$$ continuity correction

3. Solve

Virus infections

Suppose you are working with the WHO to plan a response to the initial conditions of a virus. There are two exposed groups:

- G1: 200 people, each independently infected with $p_1 = 0.1$
- G2: 100 people, each independently infected with $p_2 = 0.4$

What is $P(\text{people infected} \geq 55)$?

1. Define RVs & state goal

Let A = # infected in G1.

$A \sim \text{Bin}(200, 0.1)$

B = # infected in G2.

$B \sim \text{Bin}(100, 0.4)$

Want: $P(A + B \geq 55)$

2. Approximate as sum of Normals

$A \approx X \sim \mathcal{N}(20, 18)$ $\quad B \approx Y \sim \mathcal{N}(40, 24)$

$P(A + B \geq 55) \approx P(X + Y \geq 54.5)$ continuity correction

3. Solve

Let $W = X + Y \sim \mathcal{N}(20 + 40 = 60, 18 + 24 = 42)$

$$P(W \geq 54.5) = 1 - \Phi\left(\frac{54.5 - 60}{\sqrt{42}}\right) \approx 1 - \Phi(-0.85)$$

$$\approx 0.8023$$

A conceptual review

Everything* in probability is a sum or a product (or both)

*except conditional probability (a ratio)

Sum of values that can be considered separately (possibly weighted by prob. of happening)

$$E[X] = \sum_x \underbrace{xp(x)}_{\text{weight}}$$

$$E[X|Y = y] = \int_{-\infty}^{\infty} x \underbrace{f_{X|Y}(x|y)}_{\text{weight}} dx$$

$$P(E) = \sum_{i=1}^{n} P(E|F_i)\underbrace{P(F_i)}_{\text{weight}}$$

Law of Total Probability

$$P(E) = \sum_{i=1}^{n} P(E_i)$$

Axiom 3, $E = E_1 \cup \cdots \cup E_n$

$E_i \cap E_j = \emptyset$

Product of values that can each be considered in sequence

$$P(E \cap F \cap G) = P(E)P(F|E)P(G|EF)$$

Chain Rule

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

Independent cont. RVs

$$P(X + Y = n) = \sum_k P(X = k)P(Y = n - k)$$

Sum of indep. discrete RVs (convolution)

Conditional probability and Bayes' Theorem

Definition

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

Scaling to the correct sample space

Independence

E, F independent

$$P(F|E) = P(F)$$

Sample space doesn't need to be scaled

Bayes' Theorem

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

Prior: some prob. of event F

Likelihood

Posterior: prob. of F knowing that E happened

Scaling to the correct sample space

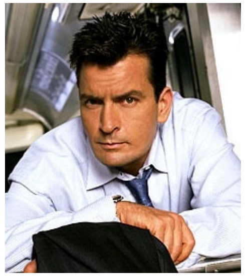
Multiple Bayes' Theorems

with events

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

with discrete RVs

$$p_{Y|X}(y|x) = \frac{p_Y(y)p_{X|Y}(x|y)}{p_X(x)}$$



with continuous RVs

You are given this value...

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

...so this is just a scalar

Really all the same idea!

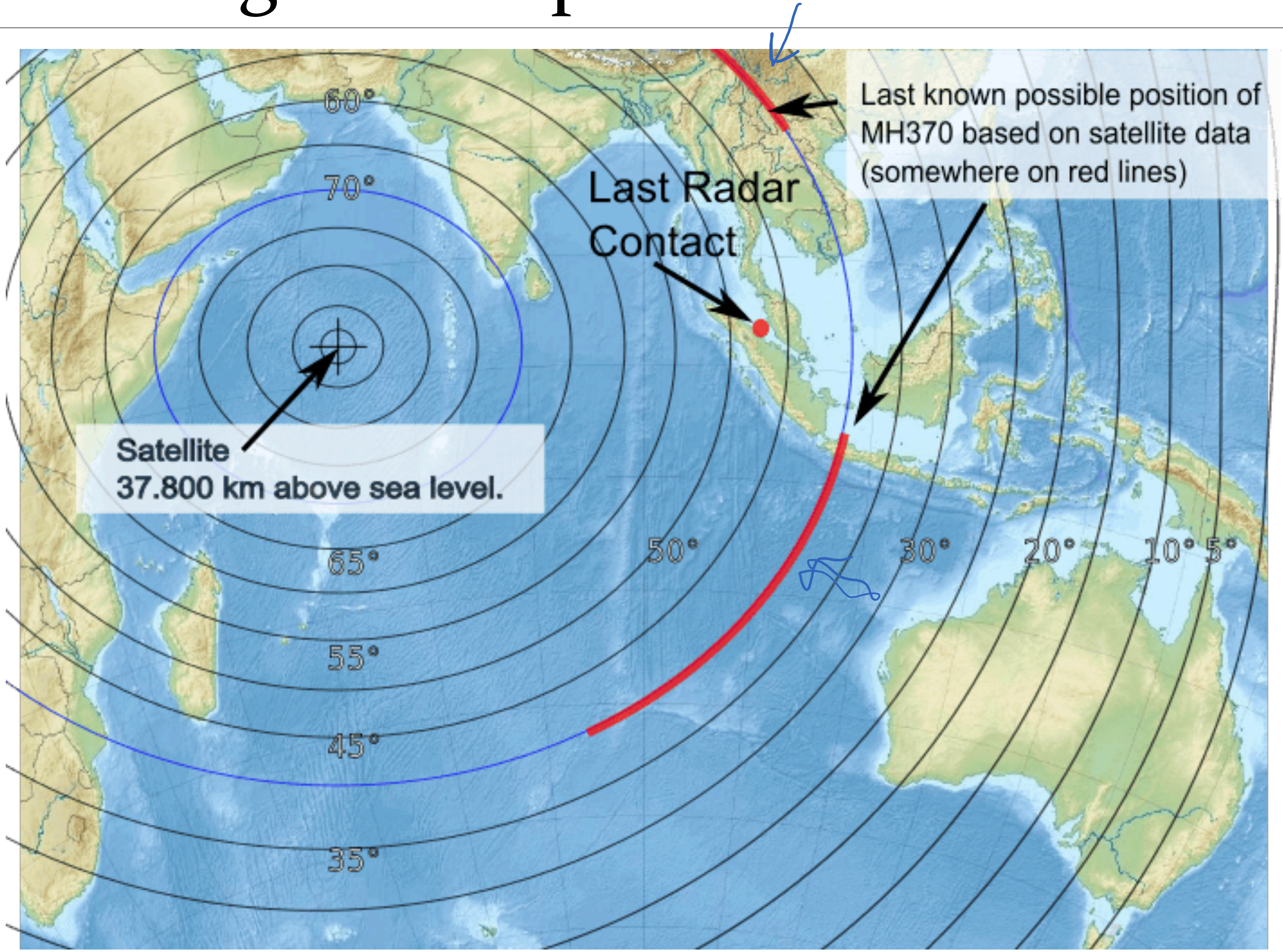
Extra fun in lecture today

We’ve gotten very far in our ability to model different situations.

Let’s test our mettle to analyze an important application that involves:

- Conditional densities
- Bayes’ Theorem
- A computer
- Normalization constants

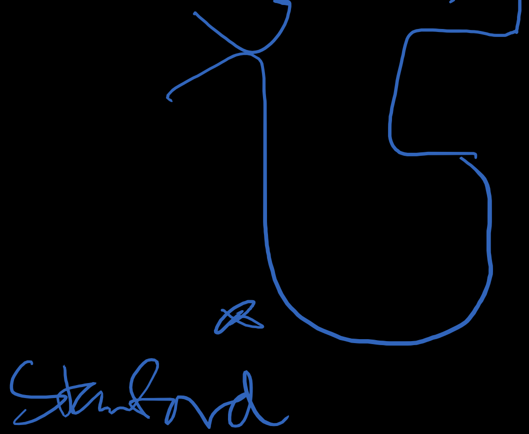
Tracking in 2-D space

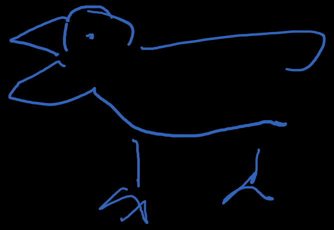


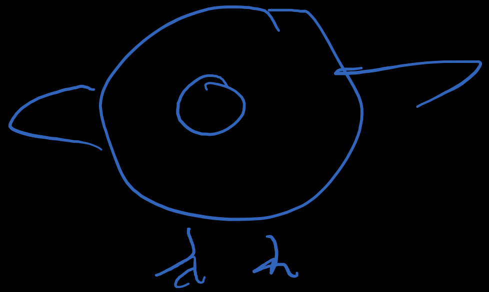
You want to know the 2-D location of an object.

Your satellite ping gives you a noisy 1-D measurement of the distance of the object from the satellite (0,0).

Using the satellite measurement, where is the object?







Interlude for jokes/announcements

Announcements

Quiz #2

Time frame: Wednesday 10/28 2:00pm – Friday 10/30 12:59pm PT
Covers: Up to end of Week 5 (including Lecture 15). PS3+PS4
Info and practice: (to be posted soon)

Homework parties

Saturdays 9am-11am PT
Sundays 2pm-4pm PT

Designated student group work time on Nooks (no CAs)

Office Hours/Mid-quarter feedback update

Thanks for your feedback! We are working on updating our OH to help more students learn

feedback Friday form

Accessing old concept checks

PDF released after late deadline passes

Week 5

MON OCT 12

13 Joint RV Statistics

- Coupon Collecting Problems
- Covariance
- Variance for Independent RVs
- Correlation

Slides (Blank) (Annotated)

Read: Ch 6.4-6.5

Concept Check

Lecture Notes

WED OCT 14

14 Conditional Expectation

- Conditional distributions
- Conditional expectation
- Law of Total Expectation
- Analyzing Recursive Code

Slides (Blank) (Annotated)

Read: Ch 7.1-7.2

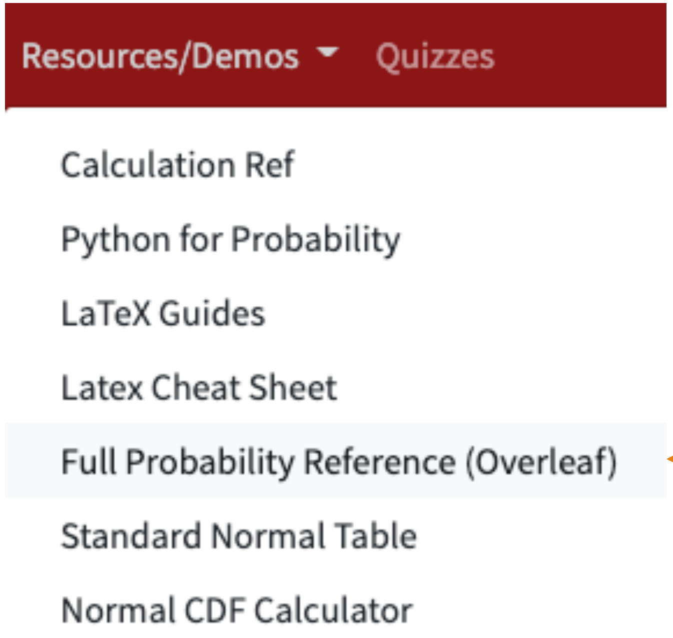
Concept Check

Lecture Notes

Still live on Gradescope

Week 1
Week 2
Week 3
Week 4
Week 5
Week 6
Week 7
Week 8
Week 9
Week 10

New handout



A summary of alllllllllllllll the things we've learned so far.

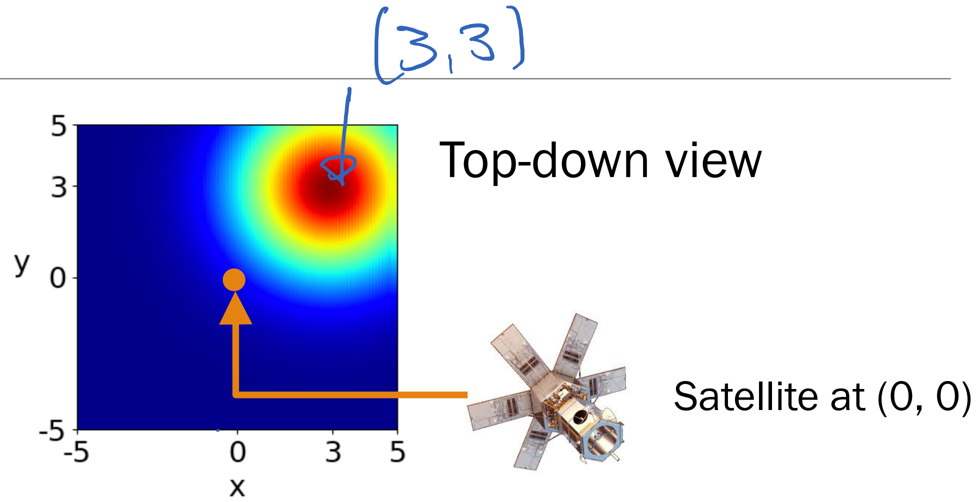
- Many equations look the same.
- ...because they're all built on the same principles!
- Overleaf, so LaTeX-friendly

Also recommended:

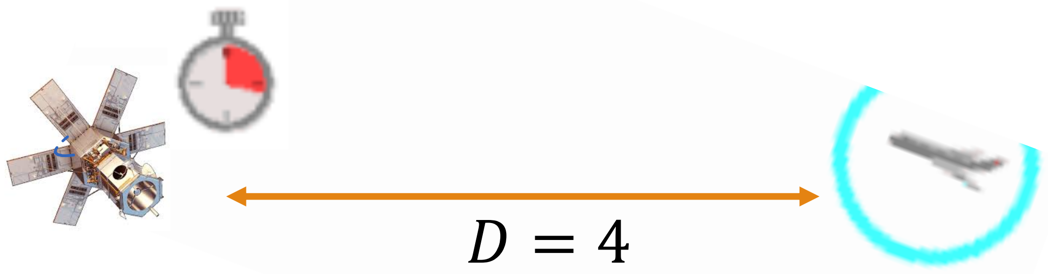
- Lecture Notes (generally shorter than slides)
- A previous CA's midterm review

Tracking in 2-D space

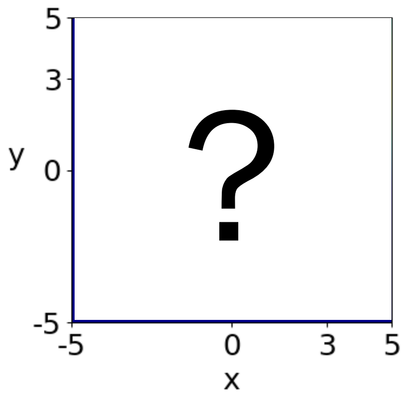
- Before measuring, we have some **prior belief** about the 2-D location of an object, (X, Y).



- We observe some noisy **measurement** $D = 4$, the Euclidean distance of the object to a satellite.



- After the measurement, what is our **updated (posterior) belief** of the 2-D location of the object?



Tracking in 2-D space

- You have a prior belief about the 2-D location of an object, (X, Y).
- You observe a noisy distance measurement, $D = 4$.
- What is your updated (posterior) belief of the 2-D location of the object after observing the measurement?

Recall Bayes terminology:

posterior belief

likelihood (of evidence)

prior belief

$$f_{X,Y|D}(x, y|d) = \frac{f_{D|X,Y}(d|x, y)\, f_{X,Y}(x, y)}{f_D(d)}$$

normalization constant

1. Define prior

$$f_{X,Y|D}(x,y|d) = \frac{f_{D|X,Y}(d|x,y)\, f_{X,Y}(x,y)}{f_D(d)}$$

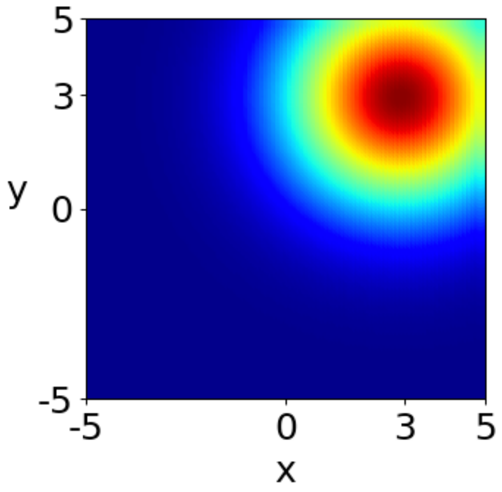
You have a prior belief about the 2-D location of an object, (X, Y).

$\mu = \begin{bmatrix} 3 \\ 3 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 2^2 & 0 \\ 0 & 2^2 \end{bmatrix}$

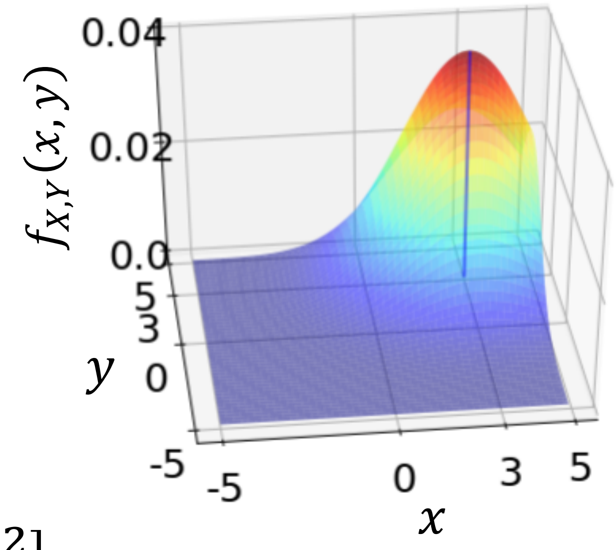
Let (X, Y) = object's 2-D location.
(your satellite is at (0,0))

Suppose the prior distribution is a symmetric bivariate normal distribution:

Top-down view



3-D view

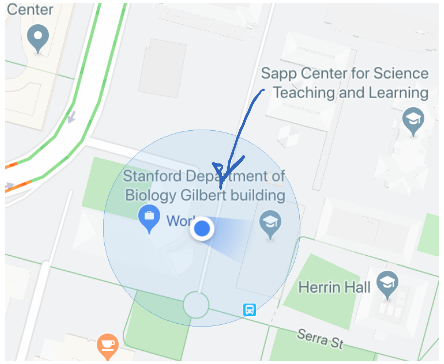


$$f_{X,Y}(x,y) = \frac{1}{2\pi 2^2} e^{-\frac{[(x-3)^2+(y-3)^2]}{2(2^2)}} = K_1 \cdot e^{-\frac{[(x-3)^2+(y-3)^2]}{8}}$$

normalizing constant

2. Define likelihood

$$f_{X,Y|D}(x,y|d) = \frac{f_{D|X,Y}(d|x,y)\, f_{X,Y}(x,y)}{f_D(d)}$$

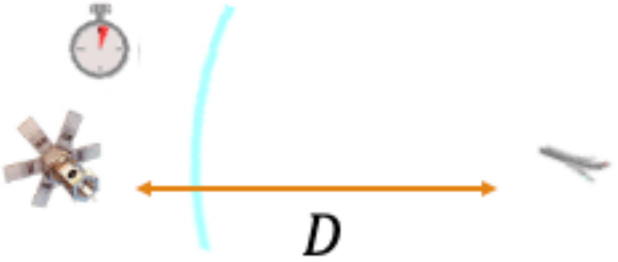


You observe a noisy distance measurement, $D = 4$.

If you knew your actual location (x, y), you could say how likely a measurement $D = 4$ is:

Let D = distance from the satellite (radially).
Suppose you knew your actual position: (x, y).

- D is still noisy! Suppose noise is **standard normal.**
- On average, D is your true Euclidean distance: $\sqrt{x^2 + y^2}$



Think

Check out the question on the next slide (Slide 54). Post any clarifications here!

https://us.edstem.org/courses/2678/discussion/153772

Think by yourself: 2 min

Post your interpretation in the chat.

(by yourself)

2. Define likelihood

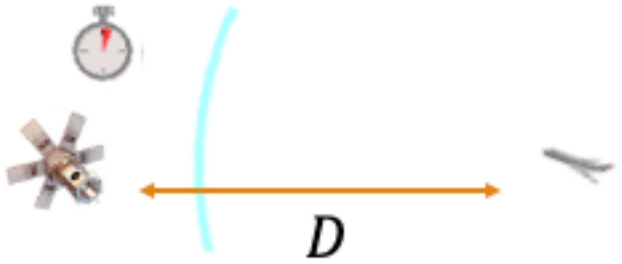
$$f_{X,Y|D}(x,y|d) = \frac{f_{D|X,Y}(d|x,y)\, f_{X,Y}(x,y)}{f_D(d)}$$

You observe a noisy distance measurement, $D = 4$.

If you knew your actual location (x, y), you could say how likely a measurement $D = 4$ is:

Let D = distance from the satellite (radially).
Suppose you knew your actual position: (x, y).

- D is still noisy! Suppose noise is **standard normal.**
- On average, D is your true Euclidean distance: $\sqrt{x^2 + y^2}$



$$D|X,Y \sim N(\mu = \ (A)\ \ , \sigma^2 = \ (B)\ )$$

$$f_{D|X,Y}(D = d|X = x, Y = y) = \frac{1}{(C)\ \sqrt{2\pi}} e^{\{\ \ (D)\ \ \}}$$

(by yourself)

2. Define likelihood

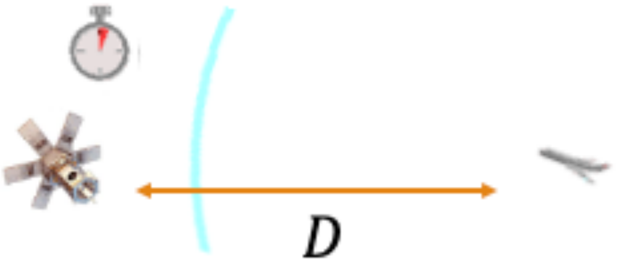
$$f_{X,Y|D}(x,y|d) = \frac{f_{D|X,Y}(d|x,y)\, f_{X,Y}(x,y)}{f_D(d)}$$

You observe a noisy distance measurement, $D = 4$.

If you knew your actual location (x, y), you could say how likely a measurement $D = 4$ is:

Let D = distance from the satellite (radially).
Suppose you knew your actual position: (x, y).

- D is still noisy! Suppose noise is **standard normal.**
- On average, D is your true Euclidean distance: $\sqrt{x^2 + y^2}$



$D = \sqrt{x^2+y^2} + Z, \quad Z \sim N(0,1)$

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

$$f_{D|X,Y}(D = d|X = x, Y = y) = \frac{1}{1 \cdot \sqrt{2\pi}} e^{\frac{-\left(d-\sqrt{x^2+y^2}\right)^2}{2 \cdot 1^2}} = K_2 \cdot e^{\frac{-\left(d-\sqrt{x^2+y^2}\right)^2}{2}}$$

normalizing constant

3. Compute posterior

$$f_{X,Y|D}(x,y|d) = \frac{f_{D|X,Y}(d|x,y)\, f_{X,Y}(x,y)}{f_D(d)}$$

What is your updated (posterior) belief of the 2-D location of the object after observing the measurement?

Compute:

Posterior belief

$$f_{X,Y|D}(x,y|4) = f_{X,Y|D}(X=x, Y=y|D=4)$$

Breakout Rooms

Check out the question on the next slide (Slide 58). Post any clarifications here!

https://us.edstem.org/courses/2678/discussion/153772

Breakout rooms: 3 min

3. Compute posterior

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

What is your updated (posterior) belief of the 2-D location of the object after observing the measurement?

Compute:

Posterior belief

$$f_{X,Y|D}(x,y|4) = f_{X,Y|D}(X=x, Y=y|D=4)$$

Know:

Prior belief

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

Observation likelihood

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

Tips

- Use Bayes' Theorem!
- $f_D(4)$ is just a scaling constant. Why?
- How can we approximate the final scaling constant with a computer?

Deep breath

Tracking in 2-D space

What is your updated (posterior) belief of the 2-D location of the object after observing the measurement?

likelihood of $D = 4$ — prior belief

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

Bayes' Theorem

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

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

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

$f(x) = \lambda e^{-\lambda x}$

$f(3) = \lambda e^{-\lambda 3}$

$f(5) = \lambda e^{-\lambda 3}$

Key: Once we know the part dependent on x, y, we can computationally approximate K_4 such that $f_{X,Y|D}$ is a valid PDF.

For your notes...

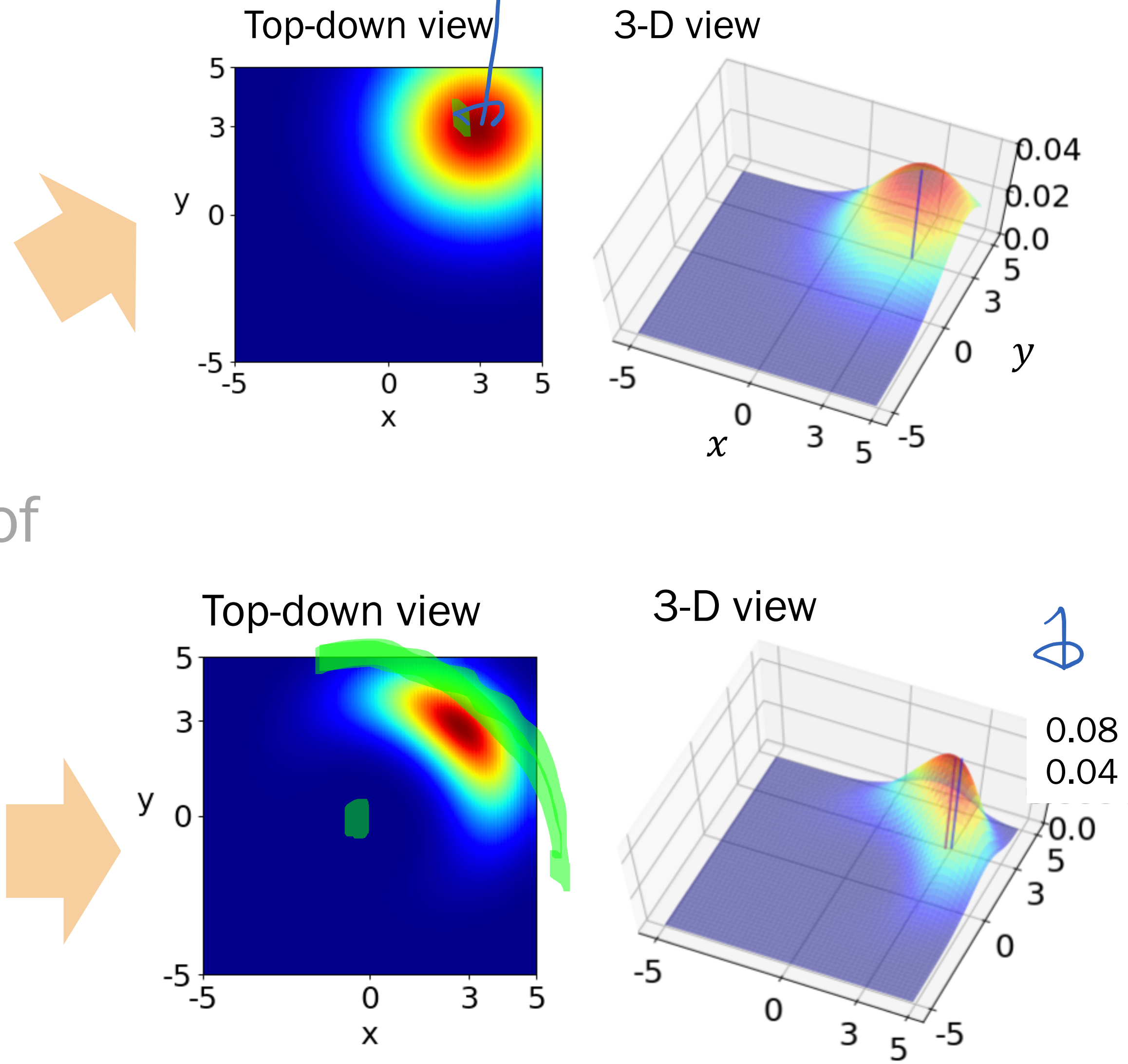
Tracking in 2-D space

$\sqrt{3^2+3^2} = 3\sqrt{2} = 4.2$

$(3,3)$

With this continuous version of Bayes' theorem, we can explore new domains.

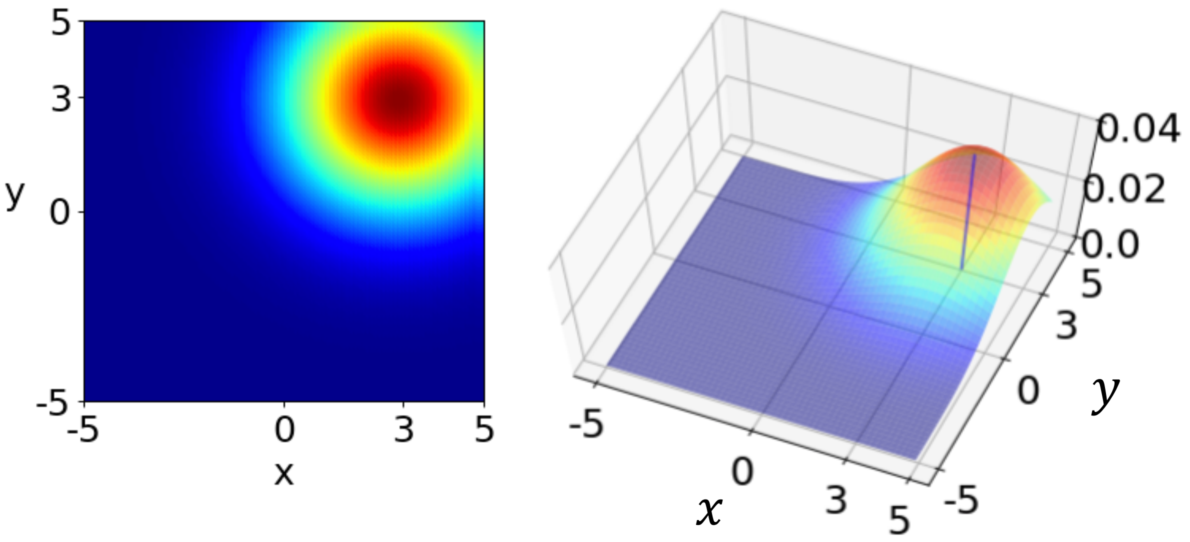
- Before measuring, we have some **prior belief** about the 2-D location of an object, (X, Y).
- We observe some noisy **measurement** of the distance of the object to a satellite.
- After the measurement, what is our **updated (posterior) belief** of the 2-D location of the object?



Tracking in 2-D space: Posterior belief

Prior belief

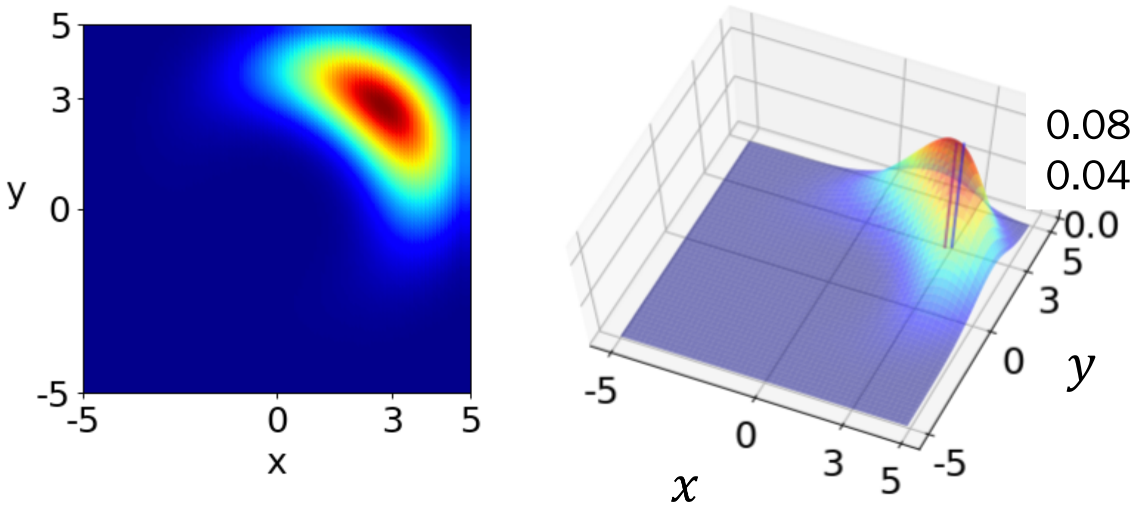
Top-down view 3-D view



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

Posterior belief

Top-down view 3-D view



$$f_{X,Y|D}(x,y|4) =$$

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

How'd you compute that K_4?

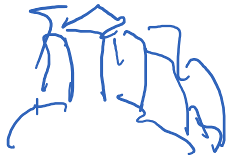
To be a valid conditional PDF, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y|D}(x,y|4)\,dx\,dy = 1$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_4 \cdot e^{-\left[\frac{\left(4-\sqrt{x^2+y^2}\right)^2}{2}+\frac{[(x-3)^2+(y-3)^2]}{8}\right]} dx\,dy = 1$$

$$\Rightarrow \frac{1}{K_4} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\left[\frac{\left(4-\sqrt{x^2+y^2}\right)^2}{2}+\frac{[(x-3)^2+(y-3)^2]}{8}\right]} dx\,dy$$

(pull out K_4, divide)

Approximate:



$$\frac{1}{K_4} \approx \sum_x \sum_y e^{-\left[\frac{\left(4-\sqrt{x^2+y^2}\right)^2}{2}+\frac{[(x-3)^2+(y-3)^2]}{8}\right]} \Delta x \Delta y$$

Use a computer!


Give yourself a pat
on the back after this!

(no video)

Extra slides

Conditional densities

Let X and Y be continuous RVs with joint PDF:

$$f_{X,Y}(x,y) = \begin{cases} \frac{12}{5}x(2-x-y) & 0 < x, y < 1 \\ 0 & \text{otherwise} \end{cases}$$

1. What is the conditional density $f_{X|Y}(x|y)$?
2. Are X and Y independent?

(discuss)

Conditional densities

Let X and Y be continuous RVs with joint PDF:

$$f_{X,Y}(x,y) = \begin{cases} \frac{12}{5}x(2-x-y) & 0 < x,y < 1 \\ 0 & \text{otherwise} \end{cases}$$

1. What is the conditional density $f_{X|Y}(x|y)$?
2. Are X and Y independent?

1. $$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{f_{X,Y}(x,y)}{\int_0^1 f_{X,Y}(x,y)dx} = \frac{\frac{12}{5}x(2-x-y)}{\int_0^1 \frac{12}{5}x(2-x-y)dx} = \frac{x(2-x-y)}{\int_0^1 x(2-x-y)dx}$$

$$= \frac{x(2-x-y)}{\left[x^2 - \frac{x^3}{3} - \frac{x^2 y}{2}\right]_0^1} = \frac{x(2-x-y)}{\frac{2}{3} - \frac{y}{2}} = \frac{6x(2-x-y)}{4-3y}$$

2. No, X and Y are dependent.

Follow up:
What is $f_{X|Y}\left(x|\frac{1}{2}\right)$?