16: Continuous Joint Distributions

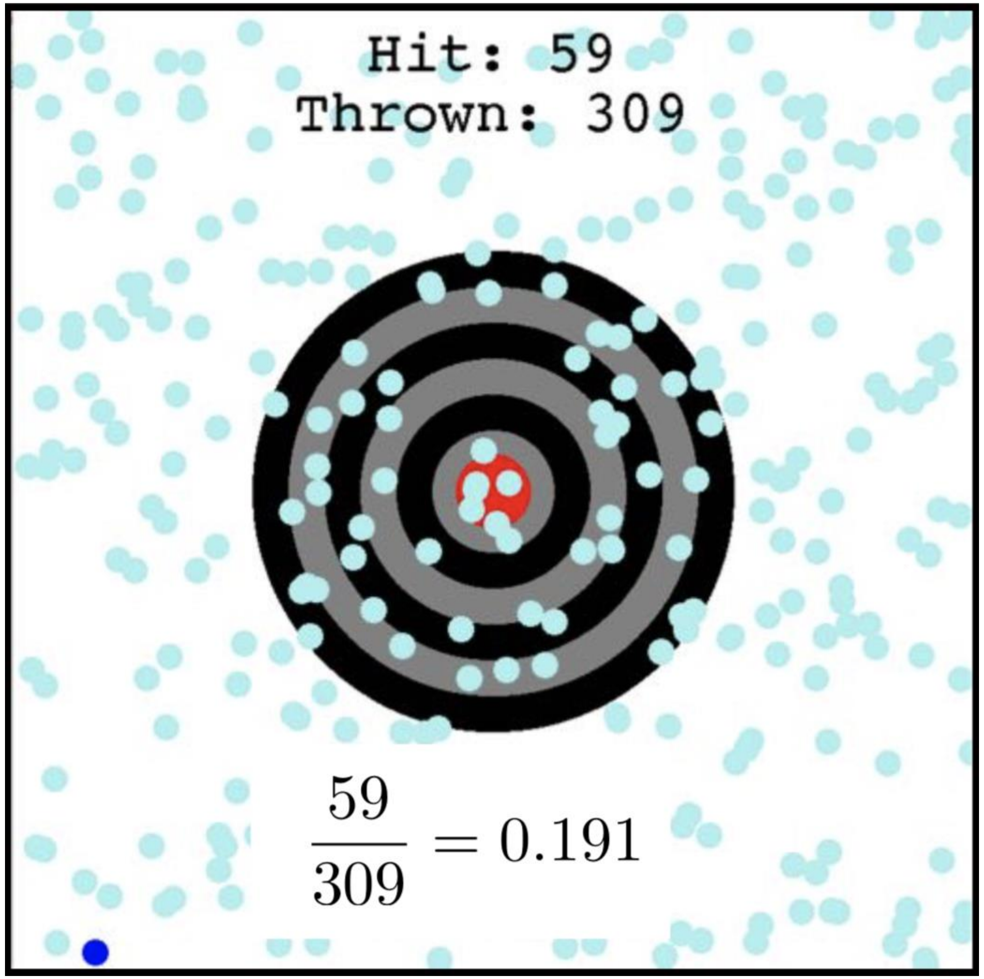
Lisa Yan

May 11, 2020

Quick slide reference

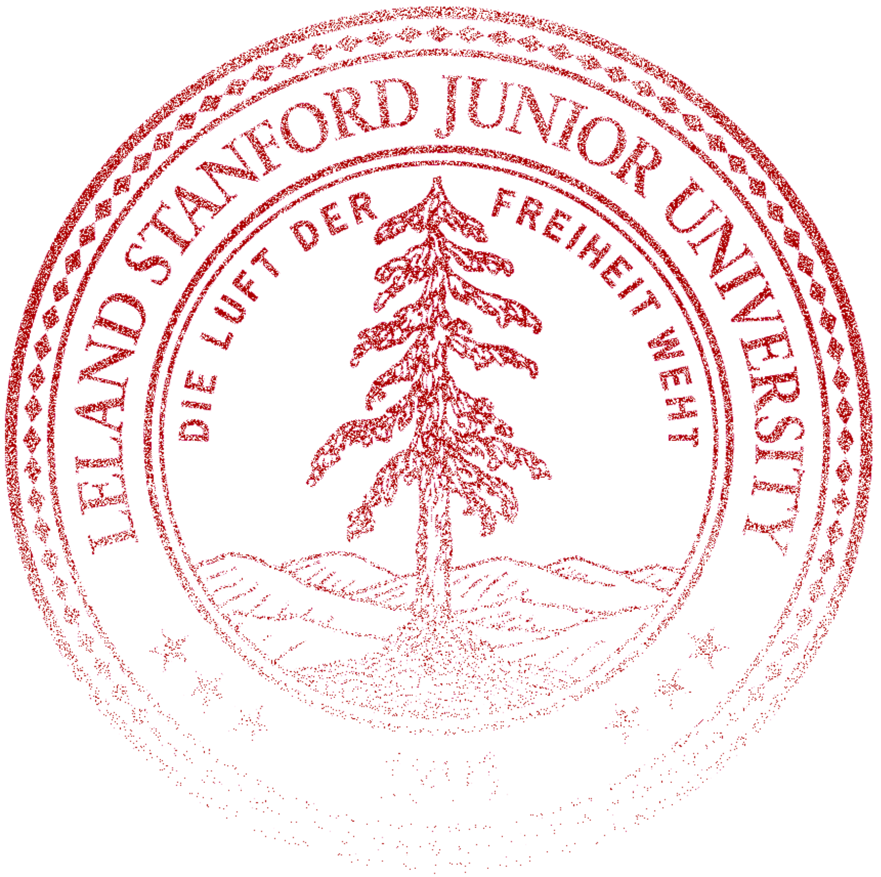
Continuous joint distributions

Remember target?



Good times…

CS109 logo with darts



The CS109 logo was created by throwing 500,000 darts according to a joint distribution.

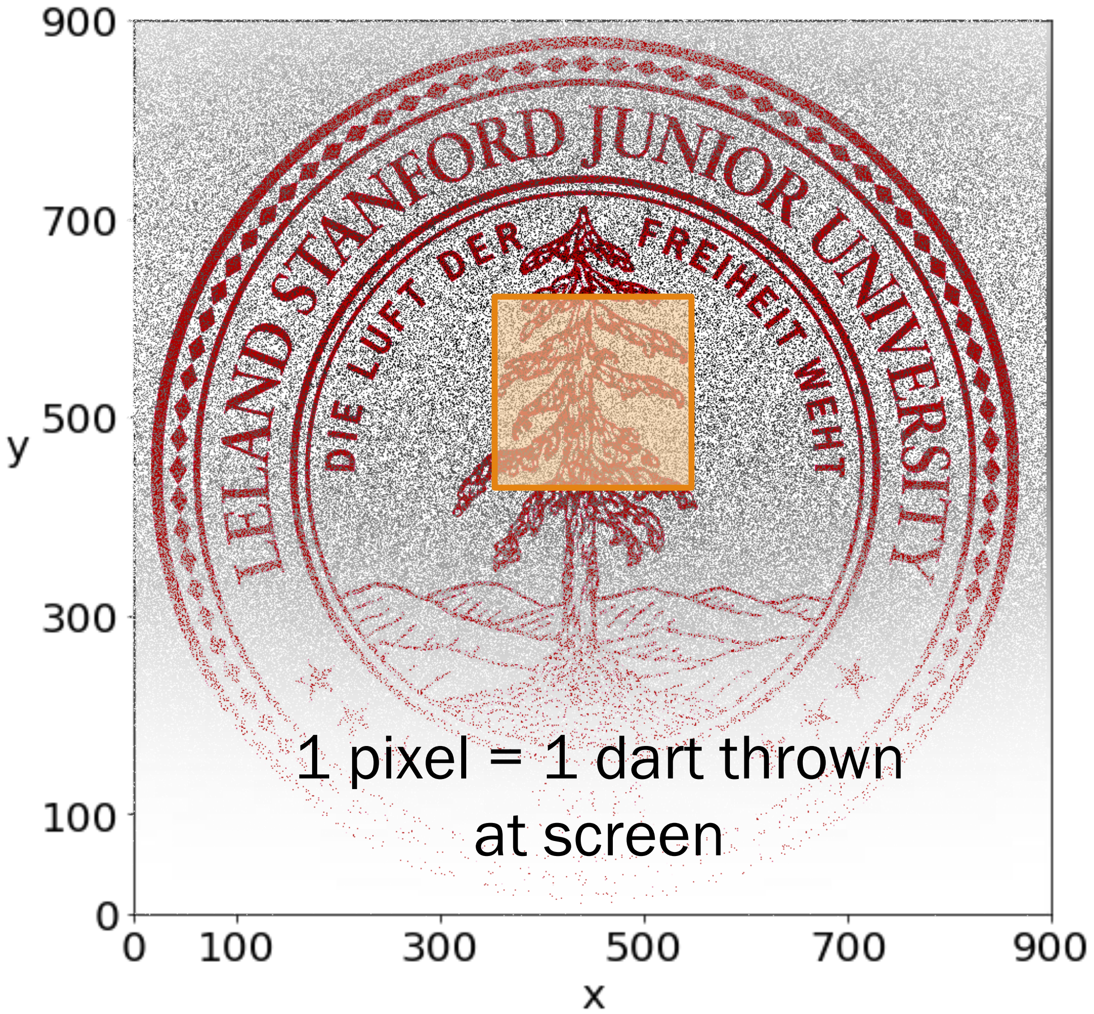
If we throw another dart according to the same distribution, what is

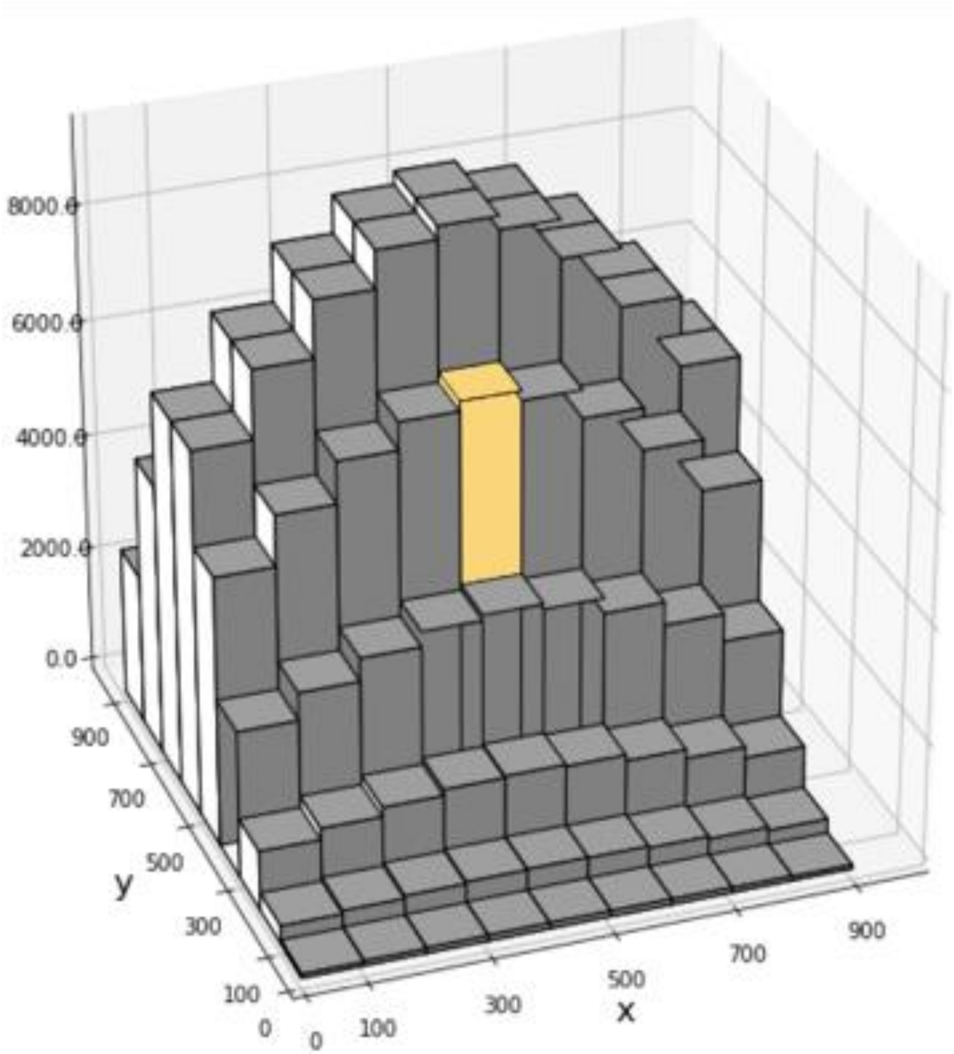
P(dart hits within r pixels of center)?

Quick check: What is the probability that a dart hits at (456.2344132343, 532.1865739012)?

CS109 logo with darts

P(dart hits within r pixels of center)?

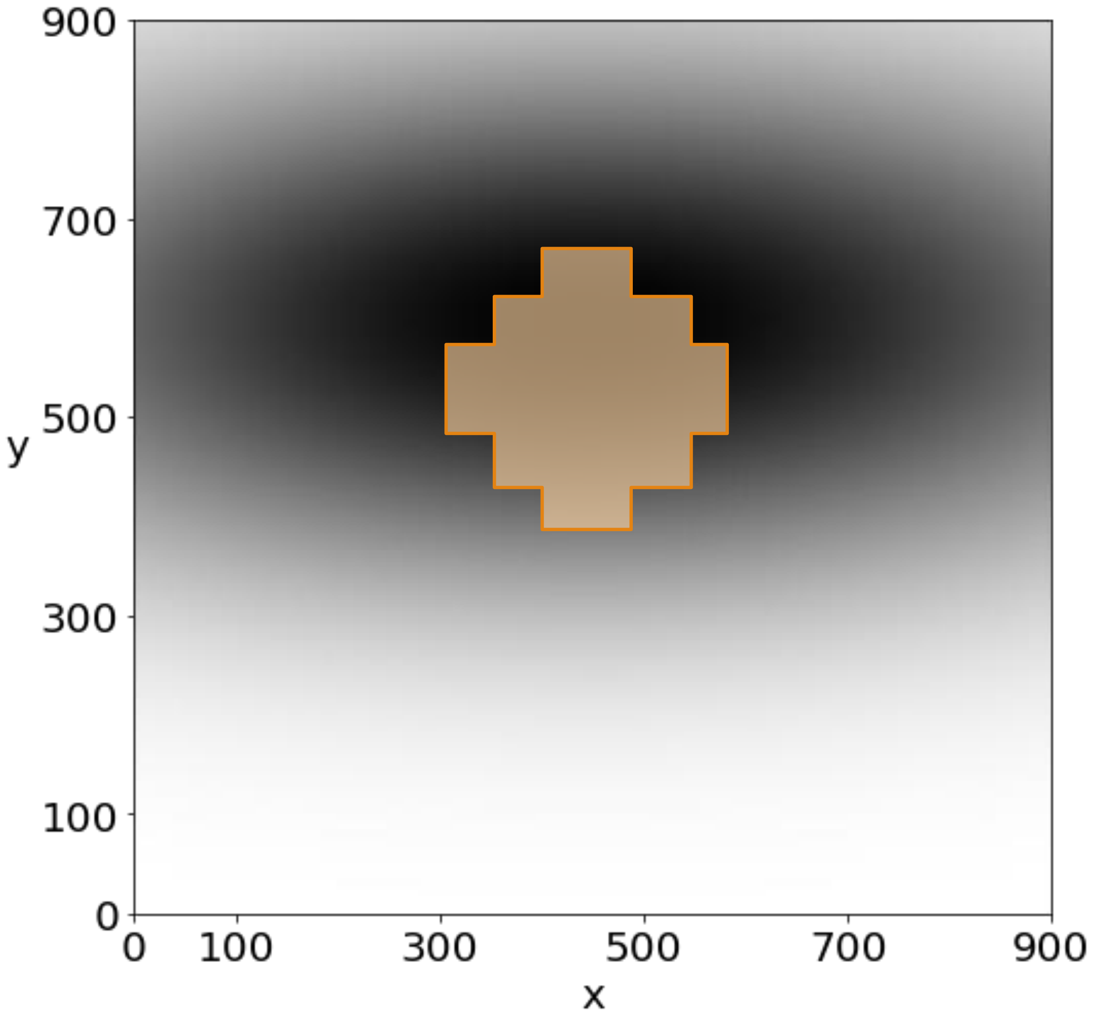


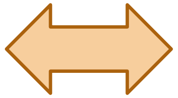


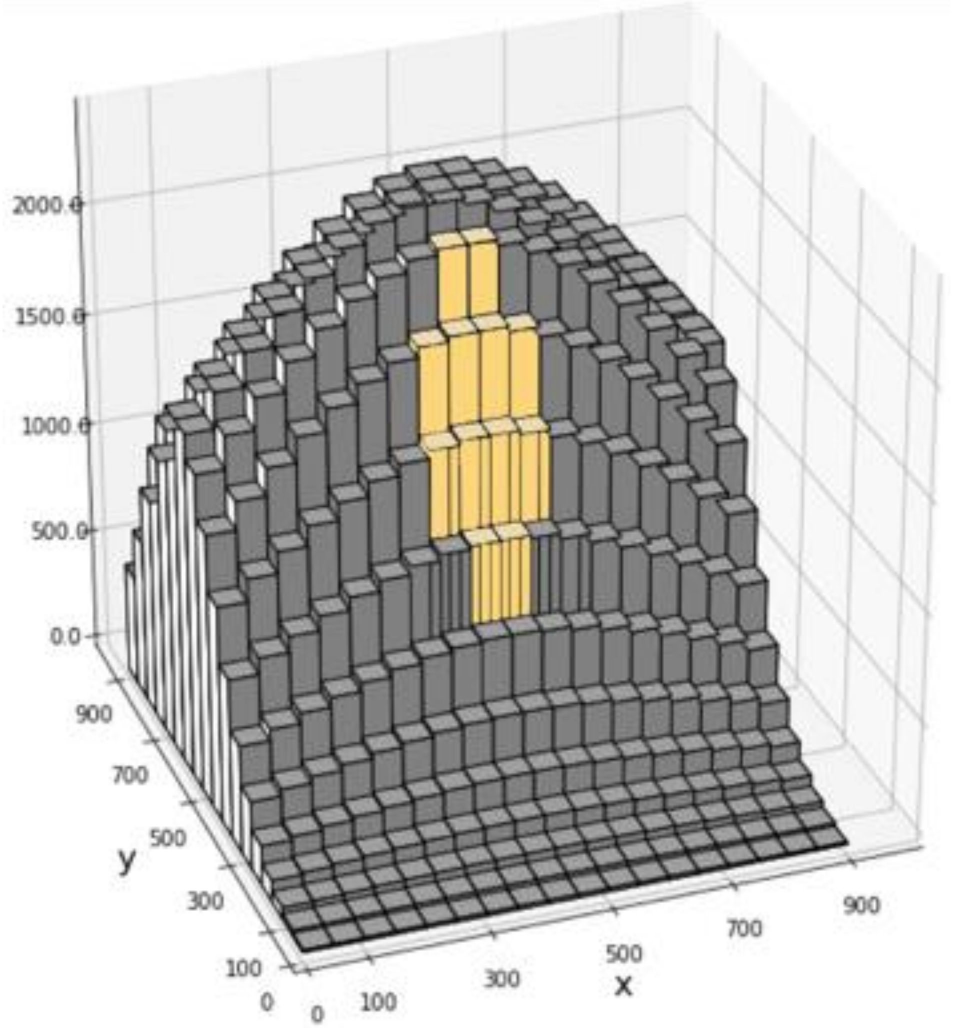
Possible dart counts (in 100x100 boxes)

CS109 logo with darts

P(dart hits within r pixels of center)?



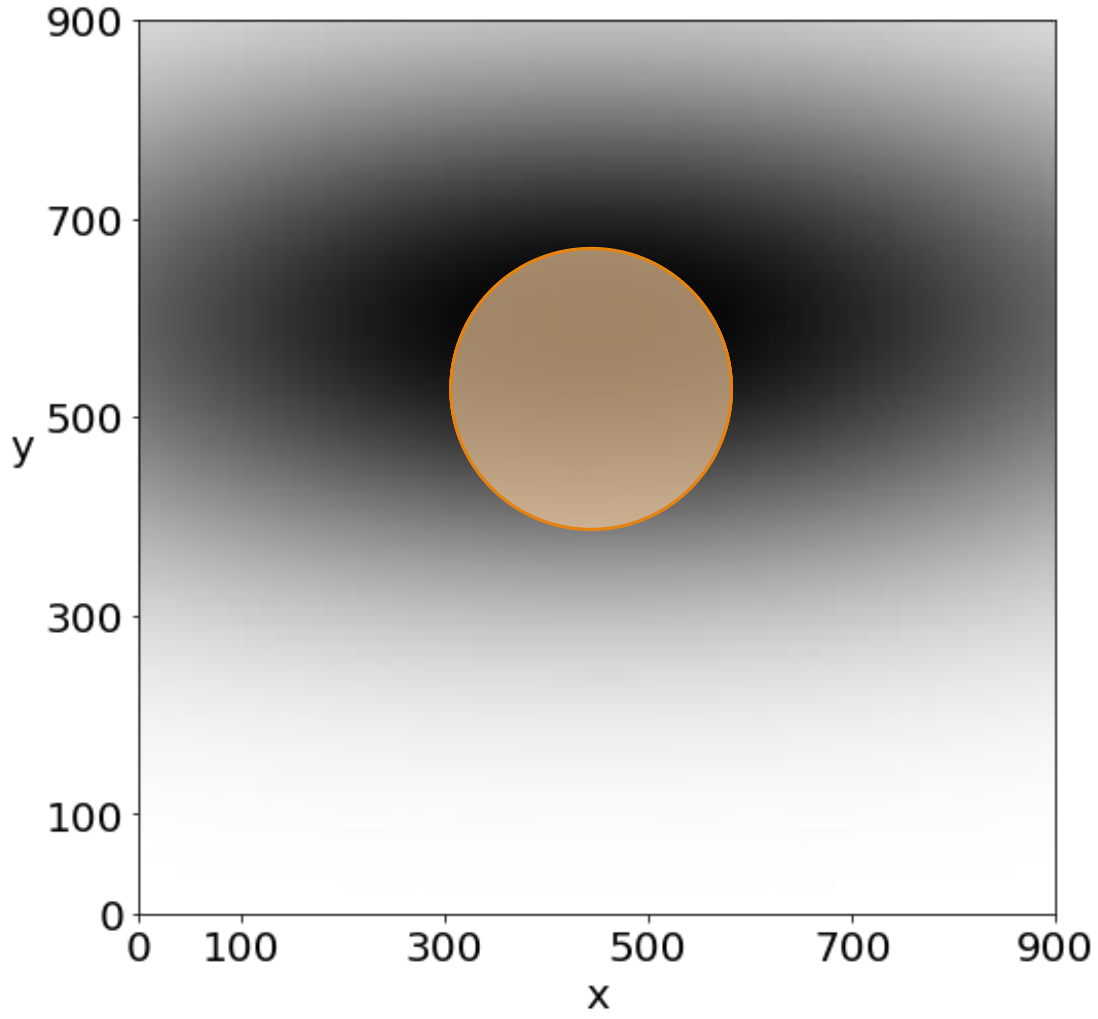


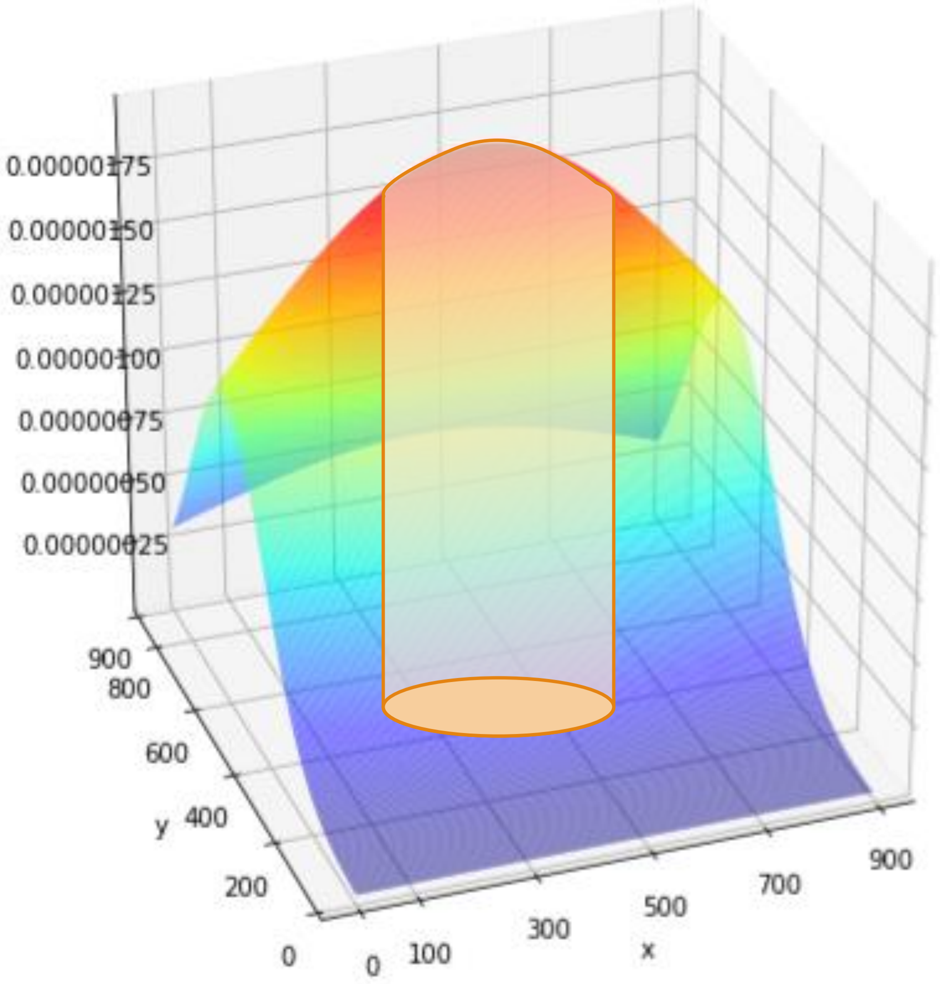


Possible dart counts (in 50x50 boxes)

CS109 logo with darts

P(dart hits within r pixels of center)?





Possible dart counts
(in infinitesimally small boxes)

Continuous joint probability density functions

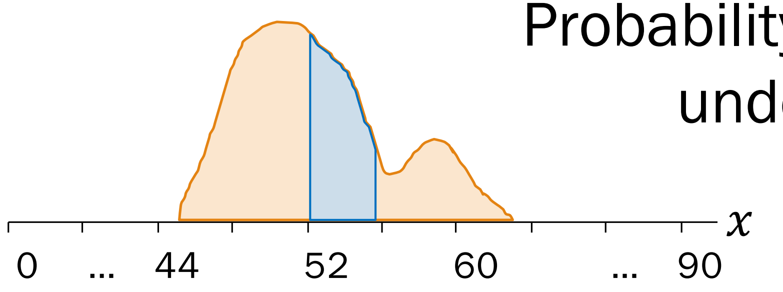
If two random variables X and Y are jointly continuous, then there exists a joint probability density function $f_{X,Y}$ defined over $-\infty < x, y < \infty$ such that:

$$P\left(a_1 \leq X \leq a_{2,}\ \ b_1 \leq Y \leq b_2\right) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{X,Y}(x, y) dy\, dx$$

From one continuous RV to jointly continuous RVs

Single continuous RV X

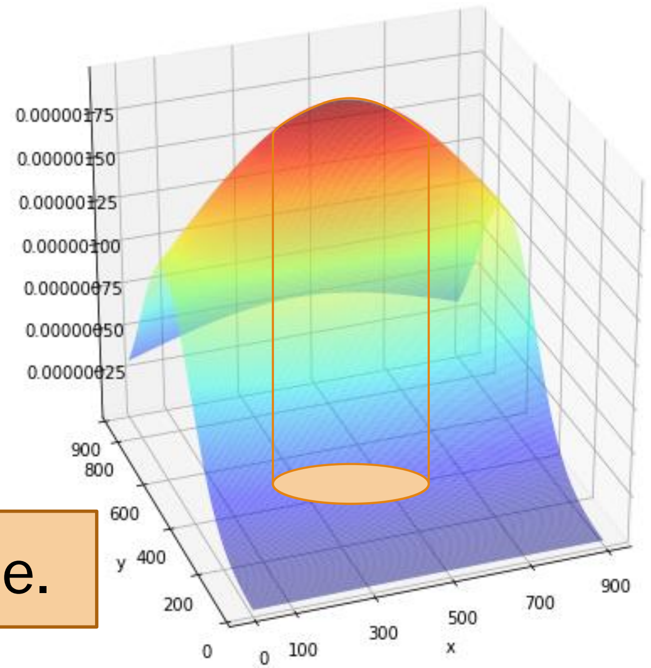
- PDF f_X such that $\int_{-\infty}^{\infty} f_X(x)dx = 1$
- Integrate to get probabilities



Probability = **area** under curve

Jointly continuous RVs X and Y

- PDF $f_{X,Y}$ such that $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy\,dx = 1$
- Double integrate to get probabilities



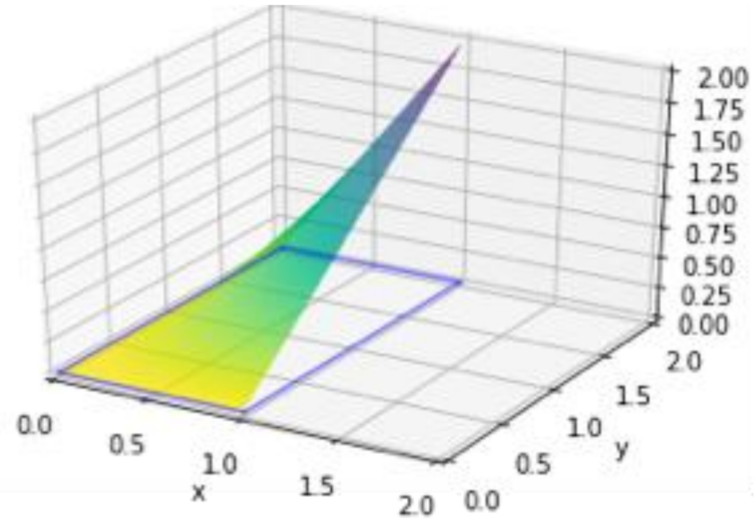
Probability for jointly continuous RVs is **volume** under a surface.

Double integrals without tears

Let X and Y be two continuous random variables.

- Support: $0 \leq X \leq 1,\ 0 \leq Y \leq 2$.

Is $g(x, y) = xy$ a valid joint PDF over X and Y?



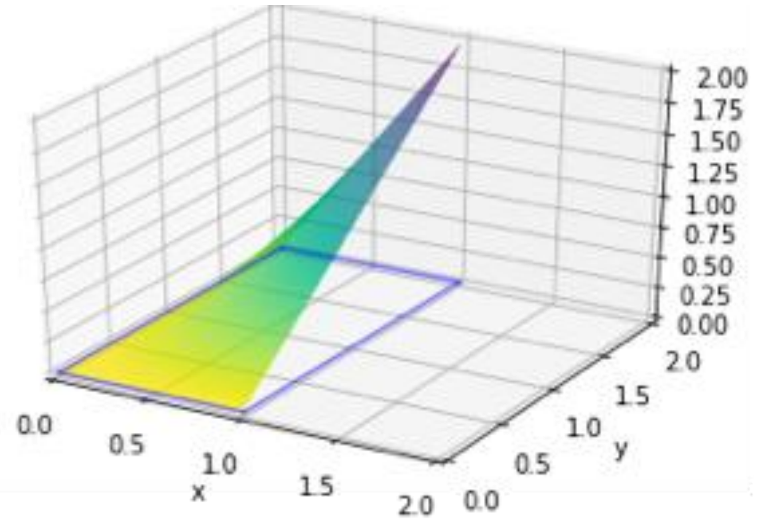
Write down the definite double integral that must integrate to 1:

Double integrals without tears

Let X and Y be two continuous random variables.

- Support: $0 \le X \le 1,\ 0 \le Y \le 2$.

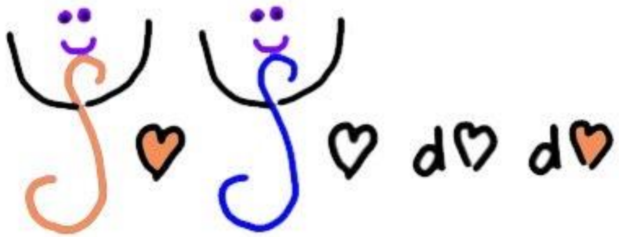
Is $g(x, y) = xy$ a valid joint PDF over X and Y?



Write down the definite double integral that must integrate to 1:

$$\int_{y=0}^{2} \int_{x=0}^{1} xy \, dx \, dy = 1 \qquad \text{or} \qquad \int_{x=0}^{1} \int_{y=0}^{2} xy \, dy \, dx = 1$$

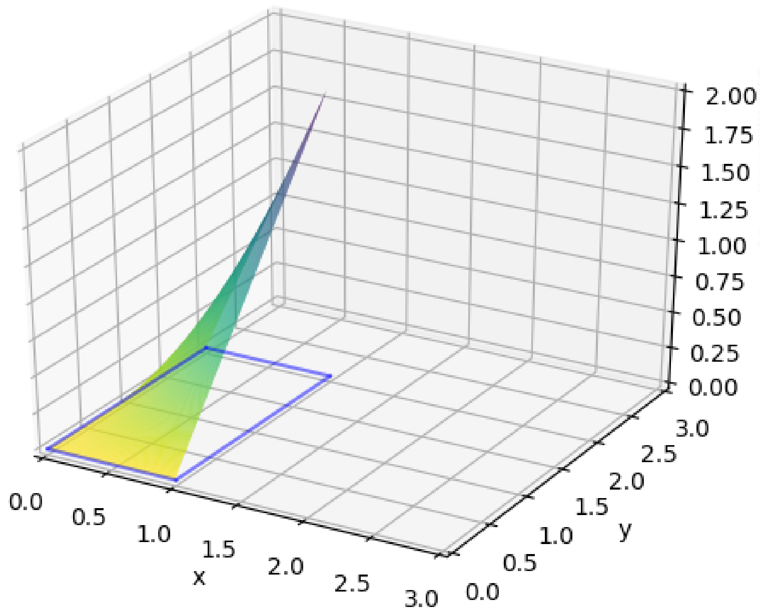
(used in next slide)



Double integrals without tears

Let X and Y be two continuous random variables.

- Support: $0 \leq X \leq 1$, $0 \leq Y \leq 2$.



Is $g(x, y) = xy$ a valid joint PDF over X and Y?

0. Set up integral:

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) dx\, dy = \int_{y=0}^{2}\int_{x=0}^{1} xy\; dx\; dy$$

1. Evaluate inside integral by treating y as a constant:

$$\int_{y=0}^{2}\left(\int_{x=0}^{1} xy\; dx\right) dy = \int_{y=0}^{2} y\left(\int_{x=0}^{1} x\; dx\right) dy = \int_{y=0}^{2} y\left[\frac{x^2}{2}\right]_0^1 dy = \int_{y=0}^{2} y\frac{1}{2} dy$$

2. Evaluate remaining (single) integral:

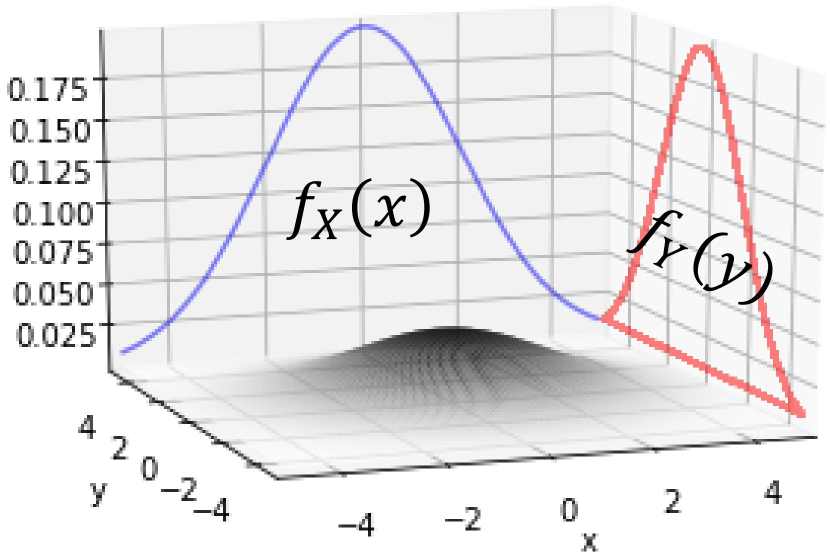
$$\int_{y=0}^{2} y\frac{1}{2} dy = \left[\frac{y^2}{4}\right]_{y=0}^{2} = 1 - 0 = 1$$

Yes, $g(x, y)$ is a valid joint PDF because it integrates to 1.

Marginal distributions

Suppose X and Y are continuous random variables with joint PDF:

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

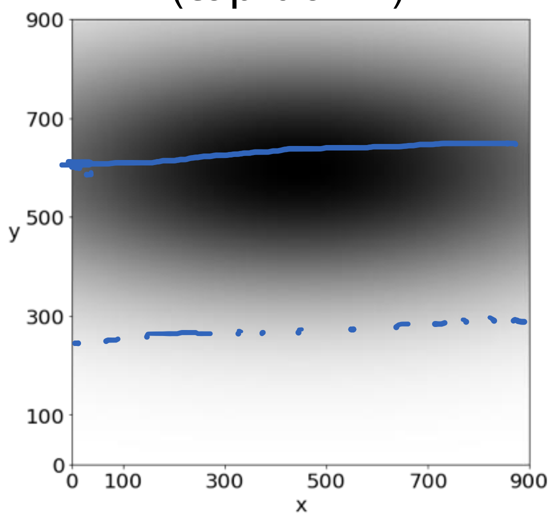


The marginal density functions (marginal PDFs) are therefore:

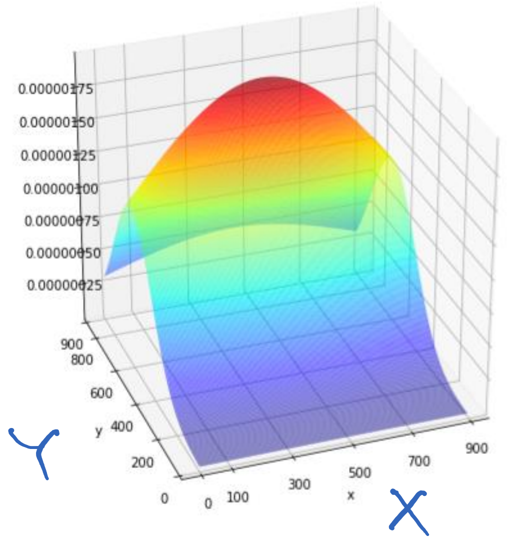
$$f_X(a) = \int_{-\infty}^{\infty} f_{X,Y}(a,y)\,dy \qquad f_Y(b) = \int_{-\infty}^{\infty} f_{X,Y}(x,b)\,dx$$

Back to darts!

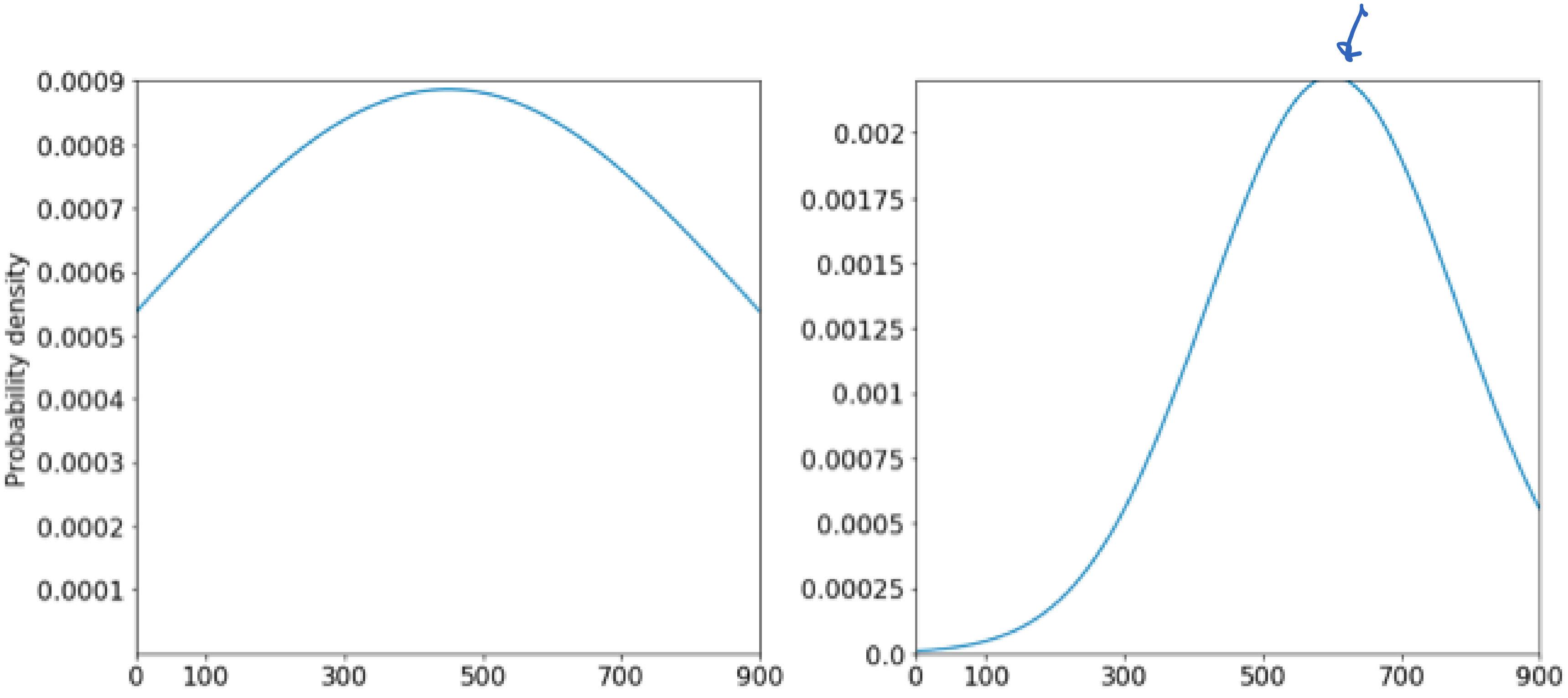
(top-down)



(side view)

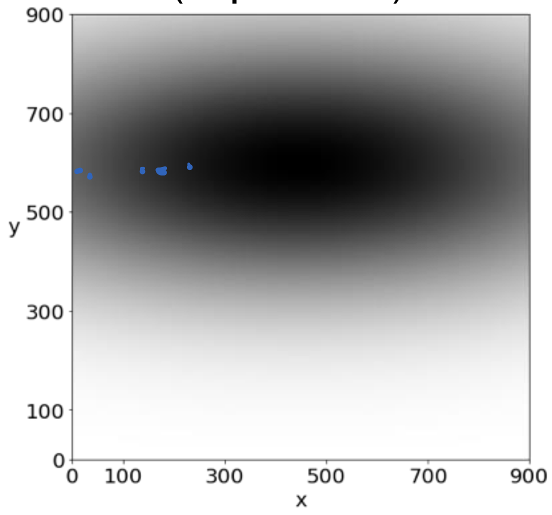


Match X and Y to their respective marginal PDFs:

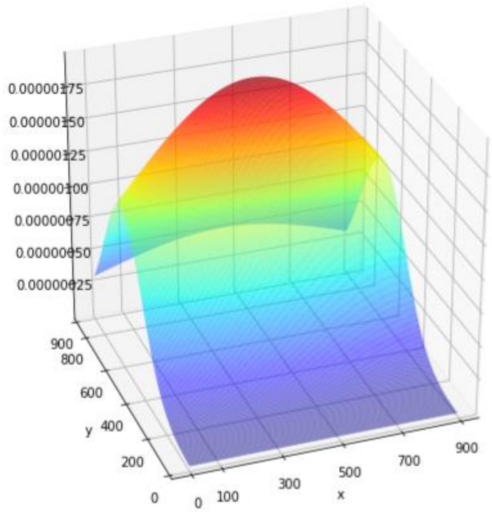


Back to darts!

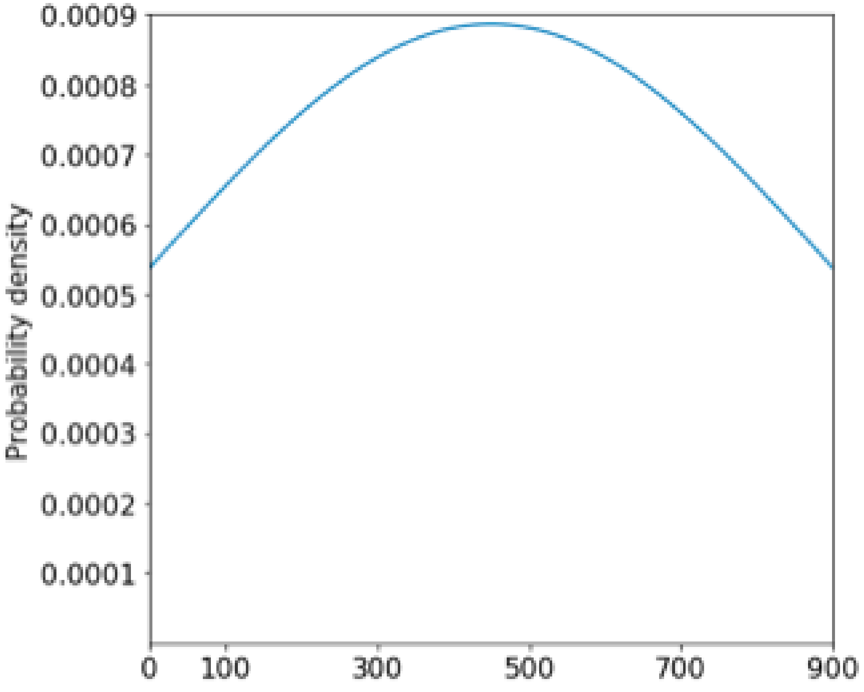
(top-down)

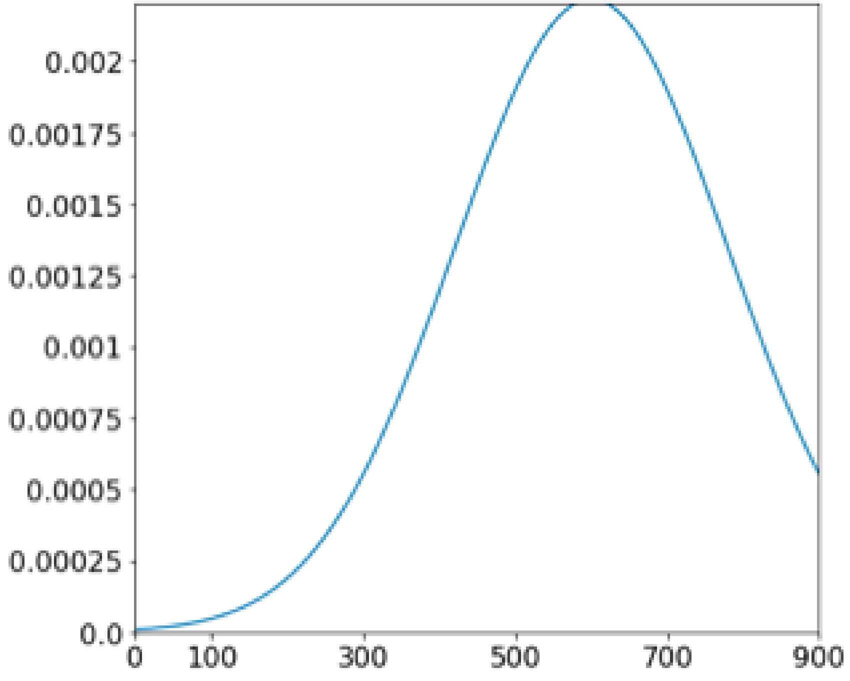


(side view)



Match X and Y to their respective marginal PDFs:





pixel x

pixel y

Extra slides

If you want more practice with double integrals,

I've included two exercises at the end of this lecture.

Joint CDFs

An observation: Connecting CDF to PDF

For a continuous random variable X with PDF f, the CDF (cumulative distribution function) is

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

The density f is therefore the derivative of the CDF, F:

$$f(a) = \frac{d}{da} F(a)$$

(Fundamental Theorem of Calculus)

Joint cumulative distribution function

For two random variables X and Y, there can be a **joint cumulative distribution function** $F_{X,Y}$:

$$F_{X,Y}(a,b) = P(X \leq a, Y \leq b)$$

For discrete X and Y:

$$F_{X,Y}(a,b) = \sum_{x \leq a} \sum_{y \leq b} p_{X,Y}(x,y)$$

For continuous X and Y:

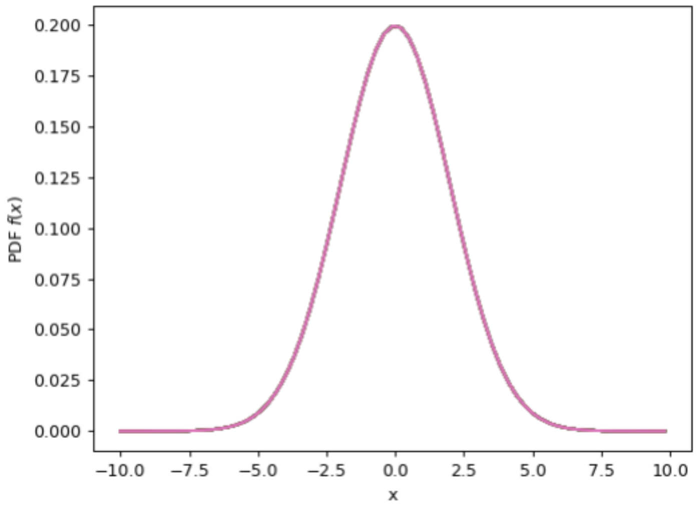
$$F_{X,Y}(a,b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x,y) dy\, dx$$

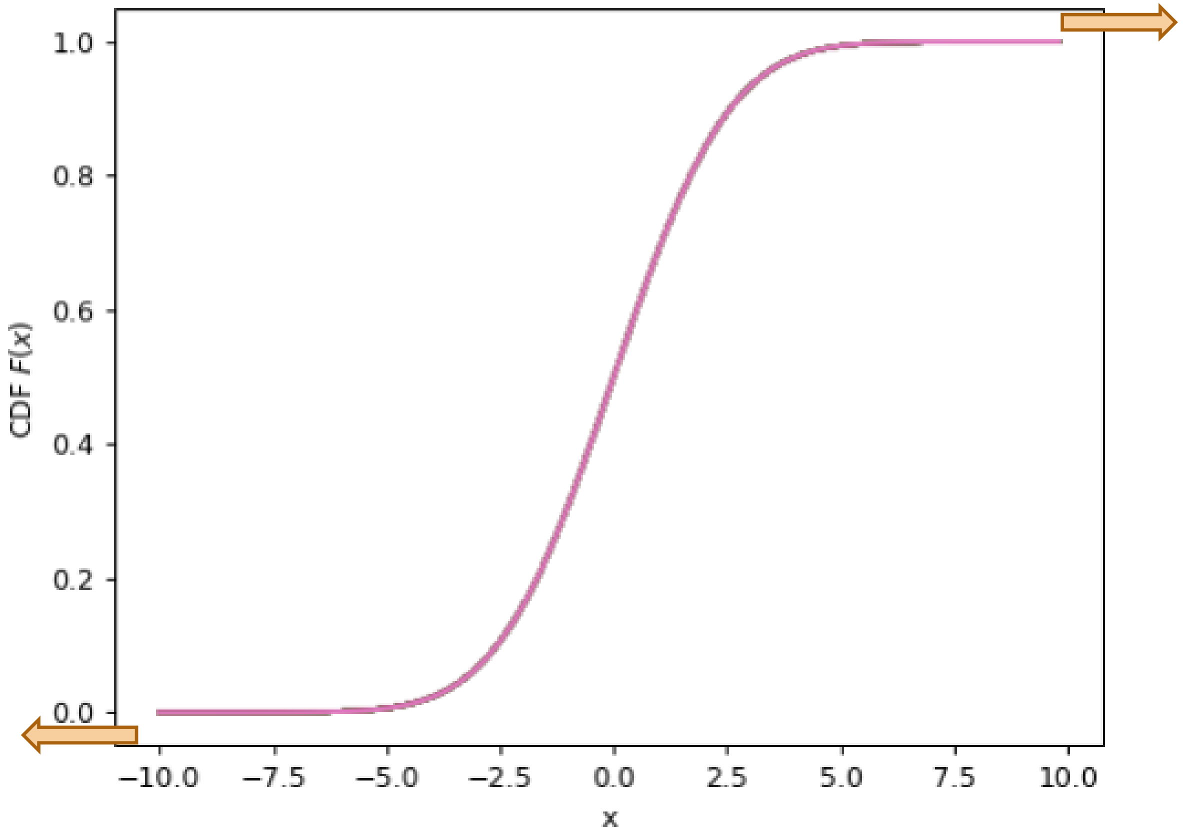
$$f_{X,Y}(a,b) = \frac{\partial^2}{\partial a\, \partial b} F_{X,Y}(a,b)$$

Single variable CDF, graphically

Review

$$\lim_{x \to +\infty} F_X(x) = 1$$



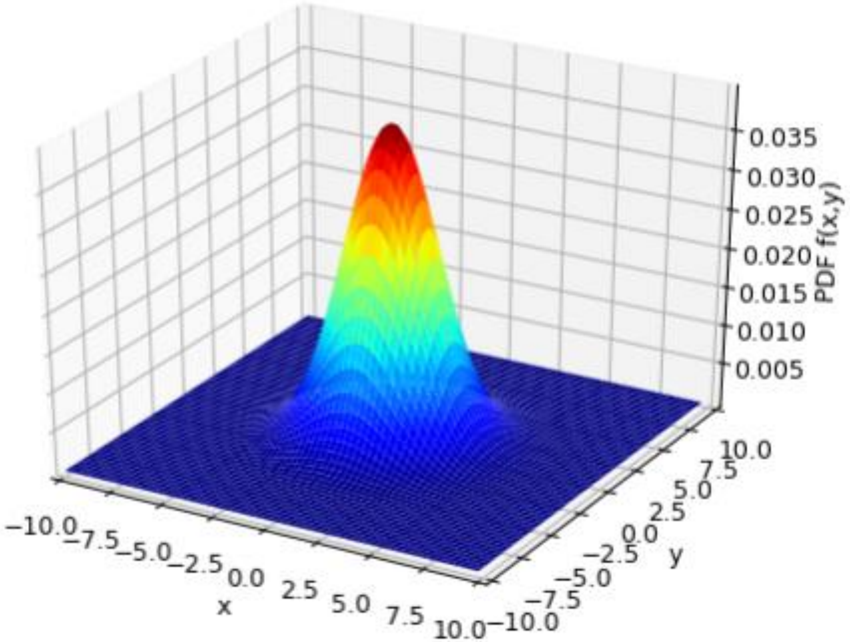


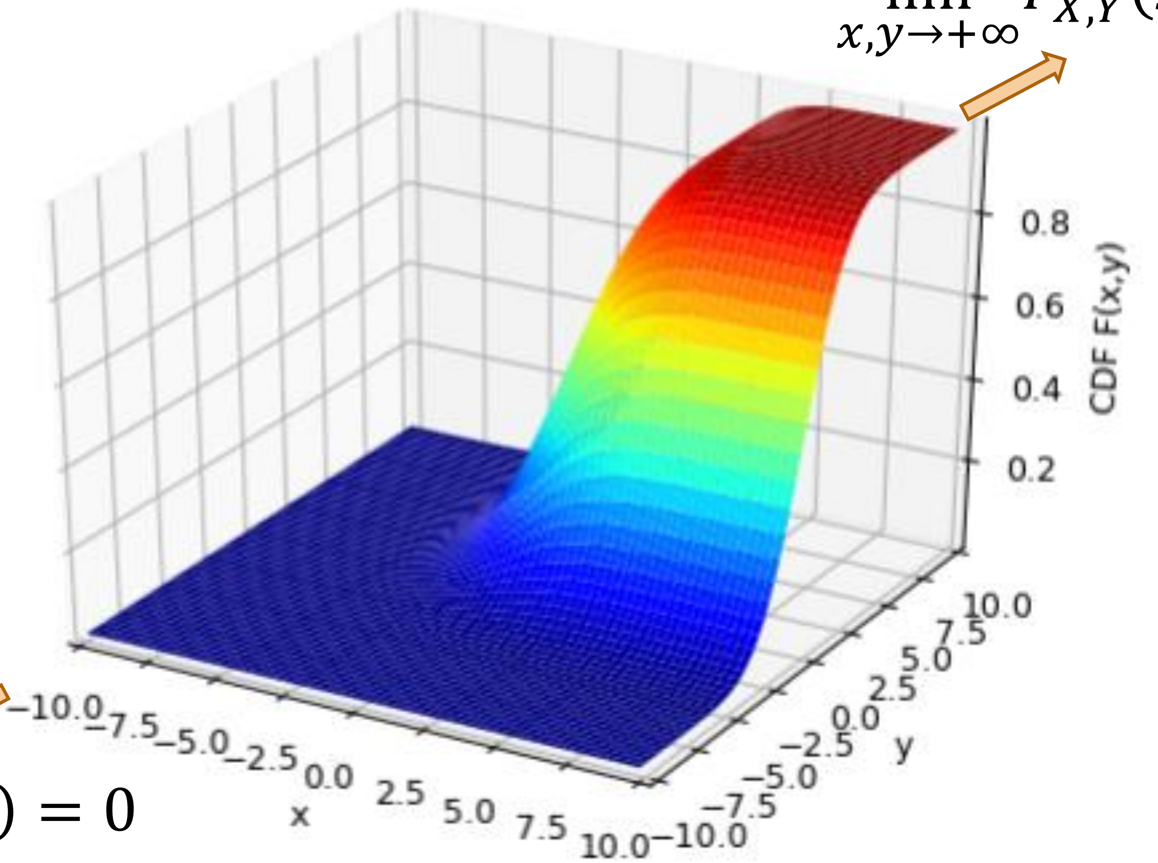
$$\lim_{x \to -\infty} F_X(x) = 0$$

$$f_X(x)$$

$$F_X(x) = P(X \leq x)$$

Joint CDF, graphically





$$\lim_{x,y \to +\infty} F_{X,Y}(x,y) = 1$$

$$\lim_{x,y \to -\infty} F_{X,Y}(x,y) = 0$$

$$f_{X,Y}(x,y)$$

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

Independent Continuous RVs

Independent continuous RVs

Two continuous random variables X and Y are **independent** if:

$$P(X \le x, Y \le y) = P(X \le x)P(Y \le y) \qquad \forall x, y$$

Equivalently:

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) \qquad \forall x, y$$
$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

Proof of PDF:

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x\, \partial y} F_{X,Y}(x,y) = \frac{\partial^2}{\partial x\, \partial y} F_X(x)F_Y(y)$$

$$= \frac{\partial}{\partial x}\frac{\partial}{\partial y} F_X(x)F_Y(y) \qquad = \frac{\partial}{\partial x} F_X(x) \frac{\partial}{\partial y} F_Y(y)$$

$$= f_X(x)f_Y(y)$$

Independent continuous RVs

Two continuous random variables X and Y are **independent** if:

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

Equivalently:

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

More generally, X and Y are independent if joint density factors separately:

$$f_{X,Y}(x,y) = g(x)h(y), \text{ where } -\infty < x, y < \infty$$

Pop quiz! (just kidding)

$f_{X,Y}(x,y) = g(x)h(y)$, where $-\infty < x, y < \infty$ → independent X and Y

Are X and Y independent in the following cases?

1. $f_{X,Y}(x,y) = 6e^{-3x}e^{-2y}$
   where $0 < x, y < \infty$

2. $f_{X,Y}(x,y) = 4xy$
   where $0 < x, y < 1$

3. $f_{X,Y}(x,y) = 24xy$
   where $0 < x + y < 1$

Pop quiz! (just kidding)

$f_{X,Y}(x,y) = g(x)h(y)$, where $-\infty < x, y < \infty$ ➡ independent X and Y

Are X and Y independent in the following cases?

☑ 1. $f_{X,Y}(x,y) = 6e^{-3x}e^{-2y}$ where $0 < x, y < \infty$

Separable functions: $g(x) = 3e^{-3x}$, $h(y) = 2e^{-2y}$

☑ 2. $f_{X,Y}(x,y) = 4xy$ where $0 < x, y < 1$

Separable functions: $g(x) = 2x$, $h(y) = 2y$

✕ 3. $f_{X,Y}(x,y) = 24xy$ where $0 < x + y < 1$

Cannot capture constraint on $x + y$ into factorization!

If you can factor densities over all of the support, you have independence.

Bivariate Normal Distribution

Bivariate Normal Distribution

X_1 and X_2 follow a bivariate normal distribution if their joint PDF f is

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}$$

Can show that $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ (Ross chapter 6, example 5d)

Often written as:

$$\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

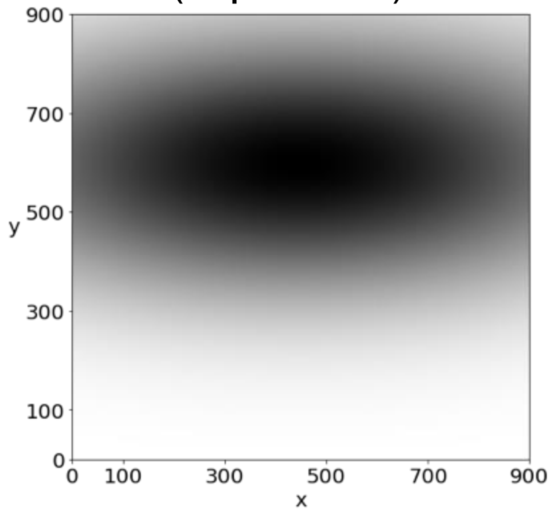
- Vector $\boldsymbol{X} = (X_1, X_2)$
- Mean vector $\boldsymbol{\mu} = (\mu_1, \mu_2)$, Covariance matrix: $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$

Recall correlation: $\rho = \frac{\text{Cov}(X_1, X_2)}{\sigma_1\sigma_2}$

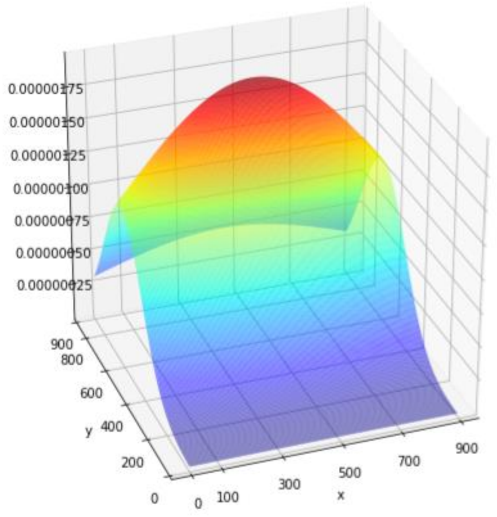
We will focus on understanding the **shape** of a bivariate Normal RV.

Back to darts

(top-down)



(side view)



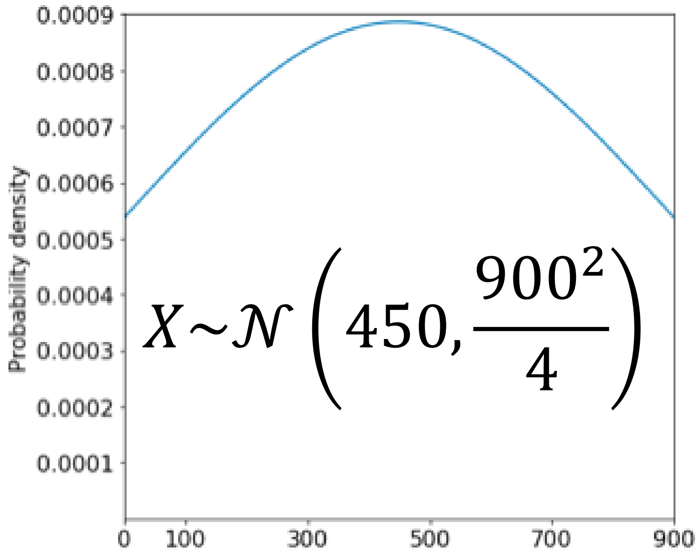
These darts were actually thrown according to a bivariate normal distribution:

$$(X, Y) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\mu} = (450, 600)$$

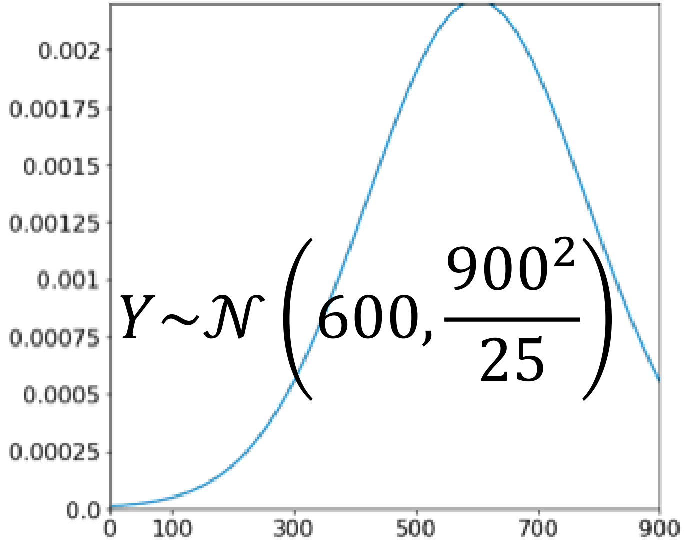
$$\boldsymbol{\Sigma} = \begin{bmatrix} 900^2/4 & 0 \\ 0 & 900^2/25 \end{bmatrix}$$

Marginal PDFs:



$$X \sim \mathcal{N}\left(450, \frac{900^2}{4}\right)$$

pixel x



$$Y \sim \mathcal{N}\left(600, \frac{900^2}{25}\right)$$

pixel y

A diagonal covariance matrix

Let $\boldsymbol{X} = (X_1, X_2)$ follow a bivariate normal distribution $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = (\mu_1, \mu_2), \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

Are X_1 and X_2 independent?

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}$$

$$= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}$$

(Note covariance: $\rho\sigma_1\sigma_2 = 0$)

$$= \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x_1-\mu_1)^2/2\sigma_1^2} \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(x_2-\mu_2)^2/2\sigma_2^2}$$

✓

X_1 and X_2 are **independent** with marginal distributions $X_1 \sim \mathcal{N}(\mu_1\ \sigma_1^2), X_2 \sim \mathcal{N}(\mu_2\ \sigma_2^2)$

(live)
16: Continuous Joint Distributions (I)

Slides by Lisa Yan
July 24, 2020

Jointly continuous RVs

Review

X and Y are jointly continuous if they have a joint PDF:

$$f_{X,Y}(x,y) \text{ such that } \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\,dx = 1$$

Most things we've learned about discrete joint distributions translate:

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	$p_{X,Y}(x,y) = p_X(x)p_Y(y)$	$f_{X,Y}(x,y) = f_X(x)f_Y(y)$
Expectation (e.g., LOTUS)	$E[g(X,Y)] = \sum_x \sum_y g(x,y)p_{X,Y}(x,y)$	$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{X,Y}(x,y)dy\,dx$

...etc.

Think

Slide 35 has a question to go over by yourself.

Post any clarifications here!

https://us.edstem.org/courses/109/discussion/60584

Think by yourself: 2 min

(by yourself)

Warmup exercise

X and Y have the following joint PDF:

$$f_{X,Y}(x,y) = 3e^{-3x}$$
where $0 < x < \infty, 1 < y < 2$

1. Are X and Y independent?

2. What is the marginal PDF of X? Of Y?

3. What is $E[X + Y]$?

(by yourself)

Warmup exercise

X and Y have the following joint PDF:

$$f_{X,Y}(x,y) = 3e^{-3x}$$
where $0 < x < \infty, 1 < y < 2$

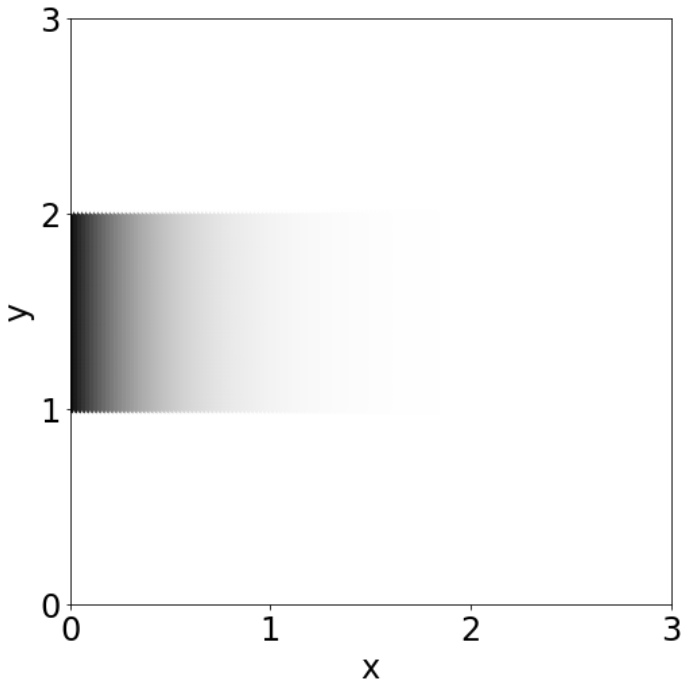
1. Are X and Y independent? ☑

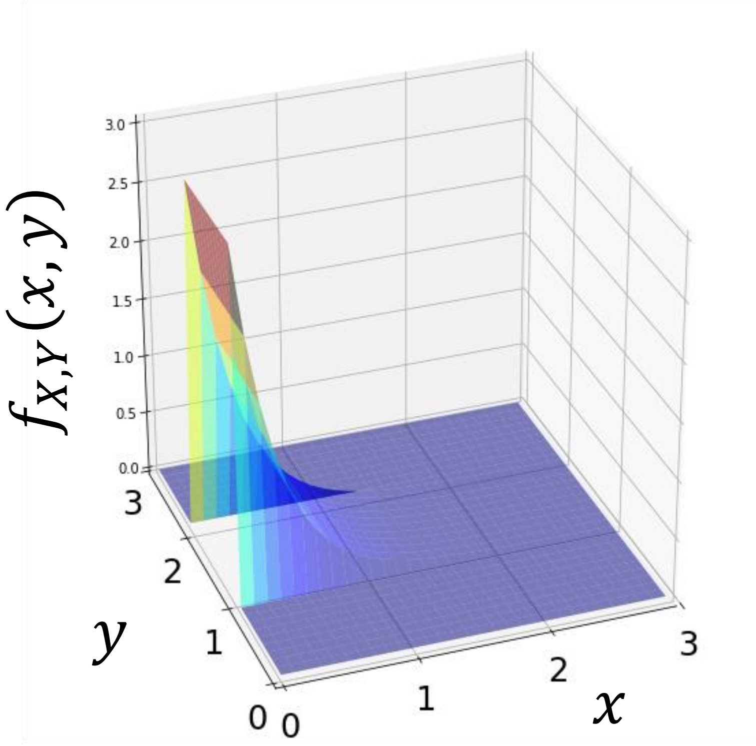
$g(x) = 3Ce^{-3x}, 0 < x < \infty$
$h(y) = 1/C, \quad 1 < y < 2$

C is a constant

2. What is the marginal PDF of X? Of Y?

3. What is $E[X+Y]$?





Warmup exercise

X and Y have the following joint PDF:

$$f_{X,Y}(x,y) = 3e^{-3x}$$
where $0 < x < \infty, 1 < y < 2$

1. Are X and Y independent? ✓

   $g(x) = 3Ce^{-3x}, 0 < x < \infty$
   $h(y) = 1/C, \quad 1 < y < 2$
   C is a constant

2. What is the marginal PDF of X? Of Y?

3. What is $E[X + Y]$?

Breakout Rooms

Check out the question on the next slide.
Post any clarifications here!

https://us.edstem.org/courses/667/discussion/94992

Breakout rooms: 4 min. Introduce yourself!

The joy of meetings

Two people set up a meeting time. Each arrives independently at a time uniformly distributed between 12pm and 12:30pm.

Define X = # minutes past 12pm that person 1 arrives. $X \sim \text{Unif}(0, 30)$
Y = # minutes past 12pm that person 2 arrives. $Y \sim \text{Unif}(0, 30)$

What is the probability that the first to arrive waits >10 mins for the other?

Compute: $P(X + 10 < Y) + P(Y + 10 < X) = 2P(X + 10 < Y)$ (by symmetry)

1. What is "symmetry" here?
2. How do we integrate to compute this probability?

The joy of meetings

Two people set up a meeting time. Each arrives independently at a time uniformly distributed between 12pm and 12:30pm.

Define X = # minutes past 12pm that person 1 arrives. $X \sim \text{Unif}(0, 30)$
Y = # minutes past 12pm that person 2 arrives. $Y \sim \text{Unif}(0, 30)$

What is the probability that the first to arrive waits >10 mins for the other?

Compute: $P(X + 10 < Y) + P(Y + 10 < X) = 2P(X + 10 < Y)$ (by symmetry)

$$= 2 \cdot \iint_{x+10<y} f_{X,Y}(x, y)dxdy = 2 \cdot \iint_{\substack{x+10<y, \\ 0 \le x,y, \le 30}} (1/30)^2 dxdy$$ (independence)

$$= \frac{2}{30^2} \int_{10}^{30} \int_{0}^{y-10} dxdy = \frac{2}{30^2} \int_{10}^{30} (y - 10)dy = \dots = \frac{4}{9}$$

Interlude for jokes/announcements

Announcements

Grades

Pset 3 by: Tonight
Midterm by: End of weekend

Mid-quarter feedback form

Link (Stanford account login)
Open until: Wednesday

Problem Set 4 due soon

Due: Monday 7/27 1pm

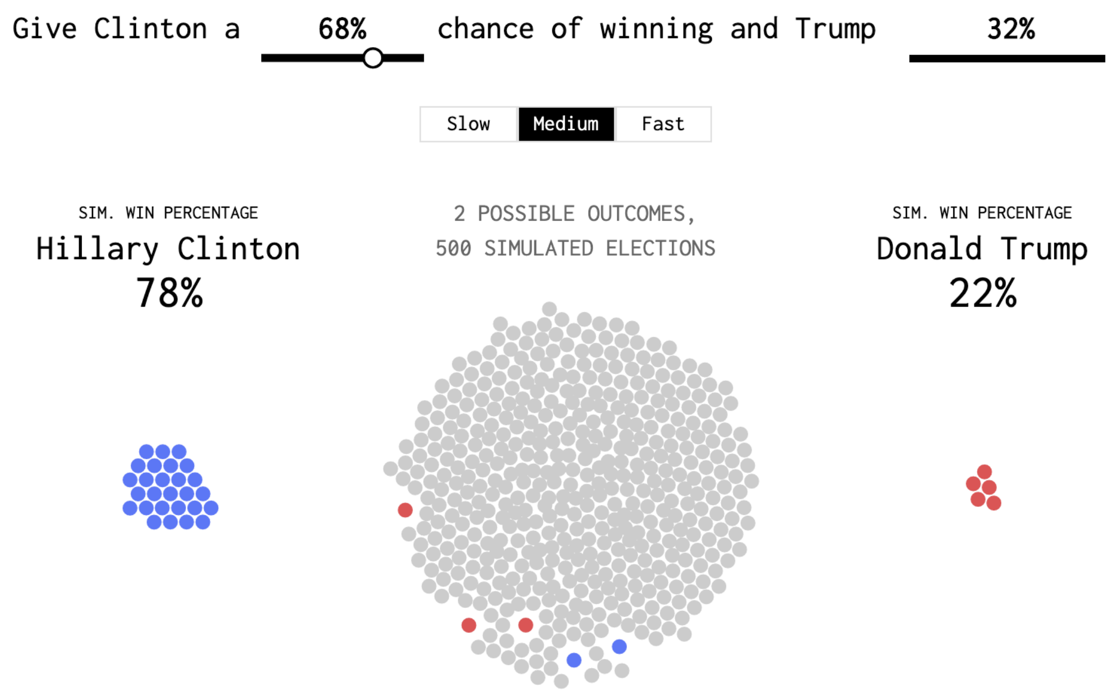
Problem Set 5 out tonight

Due: Monday 8/3 1pm
Covers: Up to Wed's lecture (16)

Interesting probability news

What That Election Probability Means

Even when you shift the probability far left or far right, the opposing candidate still gets some wins. That doesn't mean a forecast was wrong. That's just randomness and uncertainty at play. **The probability estimates the percentage of times you get an outcome if you were to do something multiple times.**



https://flowingdata.com/2016/07/28/what-that-election-probability-means/

Ethics in probability: Utilization and Fairness under Uncertainty

"Suppose there is a remote stretch of coastline with two small villages, A and B, each with a small number of houses."

"In any particular week, the probability distribution over the number of houses C impacted by power outages in each village is as follows:"

$$P_A(C=c) = \begin{cases} 0.6 & c=0 \\ 0.4 & c=2 \\ 0 & \text{otherwise} \end{cases} \qquad P_B(C=c) = \begin{cases} 0.3 & c=0 \\ 0.7 & c=3 \\ 0 & \text{otherwise} \end{cases}$$

Suppose you are trying to assign generators to these villages permanently.

Utilization: E[# of generators that are actually used]
Fairness: E[village A houses in need get generators] ~= E[village B houses in need get generators]

How do you choose an allocation that optimizes **utilization** subject to our **fairness** constraint?
=> With **probability** :-)

https://dl.acm.org/doi/abs/10.1145/3351095.3372847 ACM FAT* Best Paper Award 2020

Bivariate Normal Distribution

Review

The bivariate normal distribution of $\boldsymbol{X} = (X_1, X_2)$:

$$\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

- Mean vector $\boldsymbol{\mu} = (\mu_1, \mu_2)$
- Covariance matrix: $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$ $\quad \text{Cov}(X_1, X_2) = \text{Cov}(X_2, X_1) = \rho\sigma_1\sigma_2$
- Marginal distributions: $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$
- For bivariate normals in particular, $\text{Cov}(X_1, X_2) = 0$ implies X_1, X_2 independent.

We will focus on understanding the **shape** of a bivariate Normal RV.

Breakout Rooms

Check out the question on the next slide.
Post any clarifications here!

https://us.edstem.org/courses/667/discussion/94992

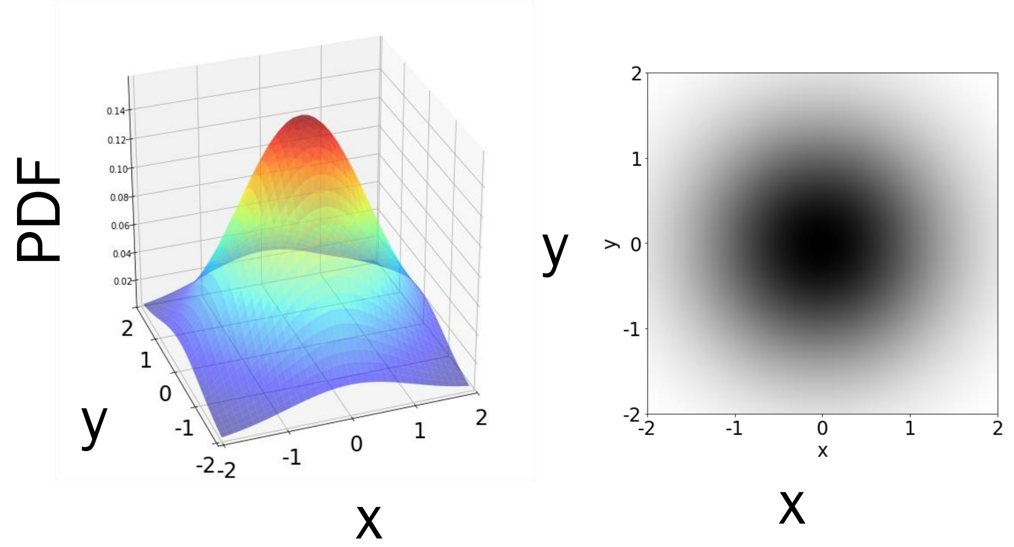
Breakout rooms: 3 min. Introduce yourself!

(X, Y) Matching (all have $\boldsymbol{\mu} = (0, 0)$)

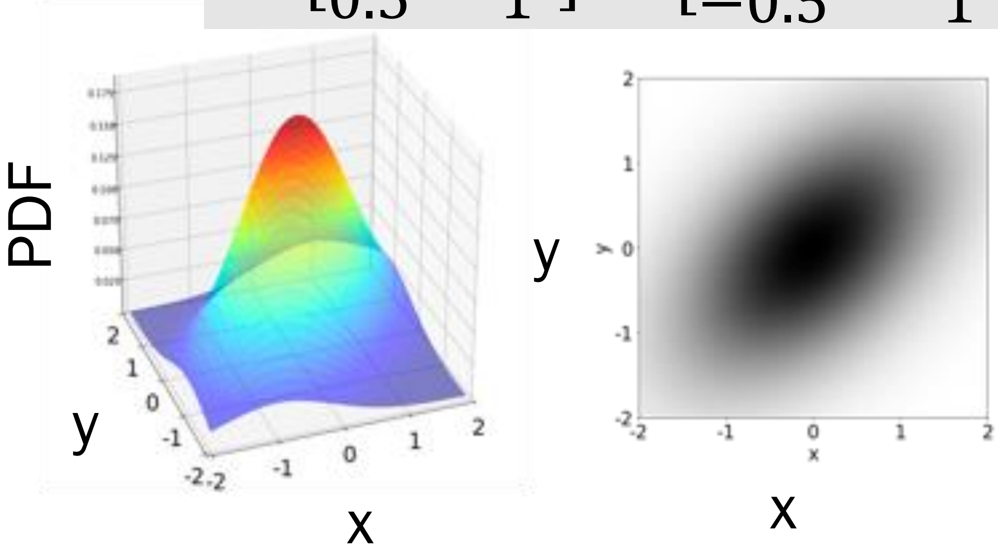
A. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ B. $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

C. $\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ D. $\begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$

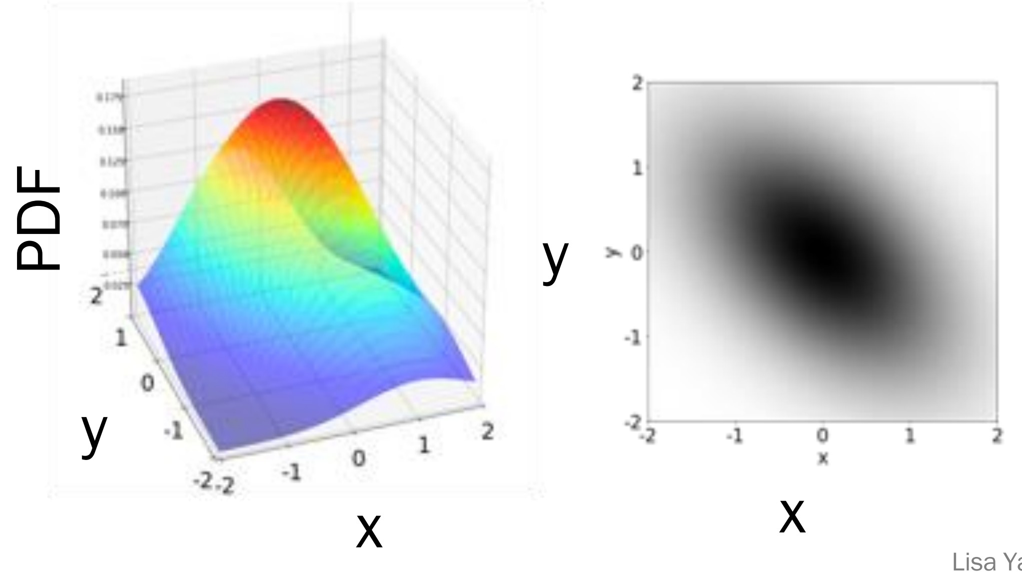
1.



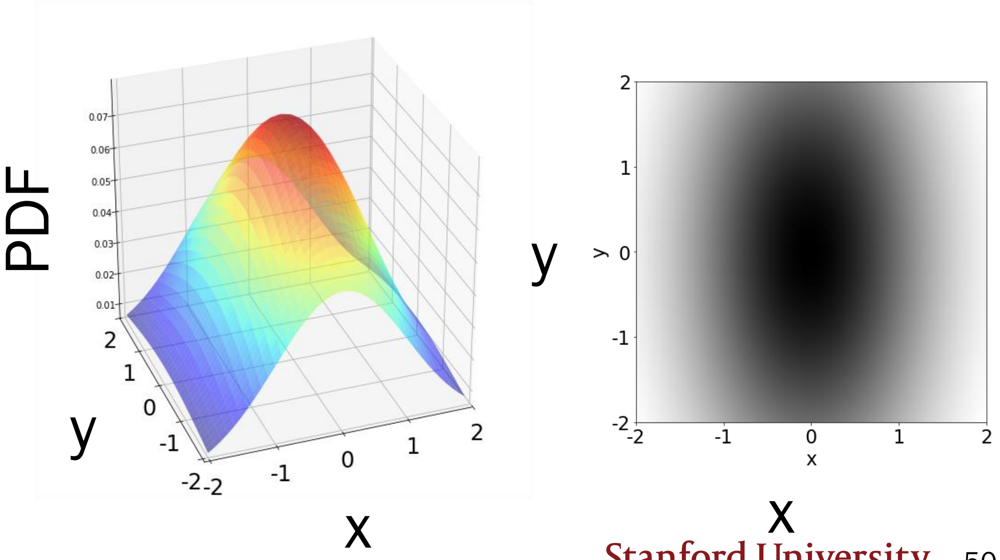
2.



3.



4.

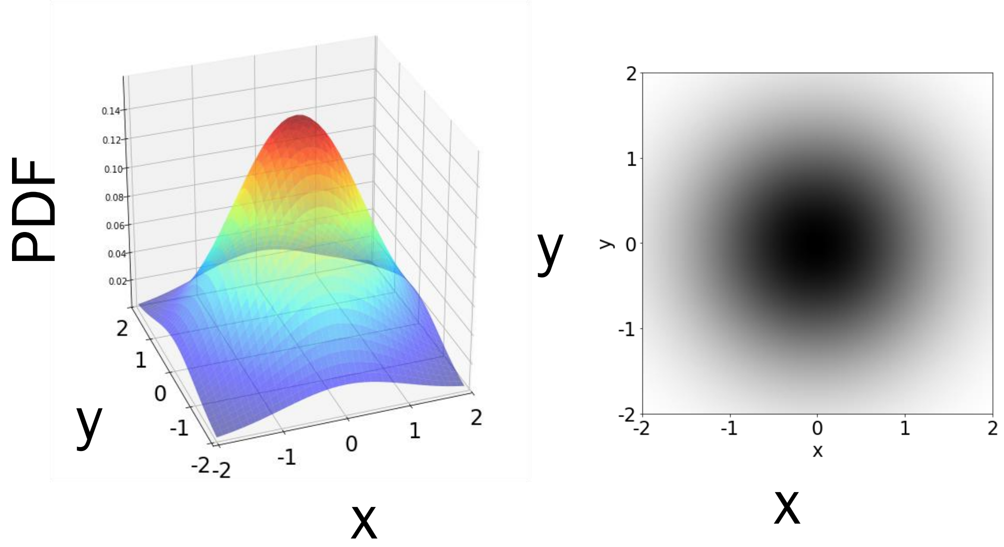


(X, Y) Matching (all have $\boldsymbol{\mu} = (0, 0)$)

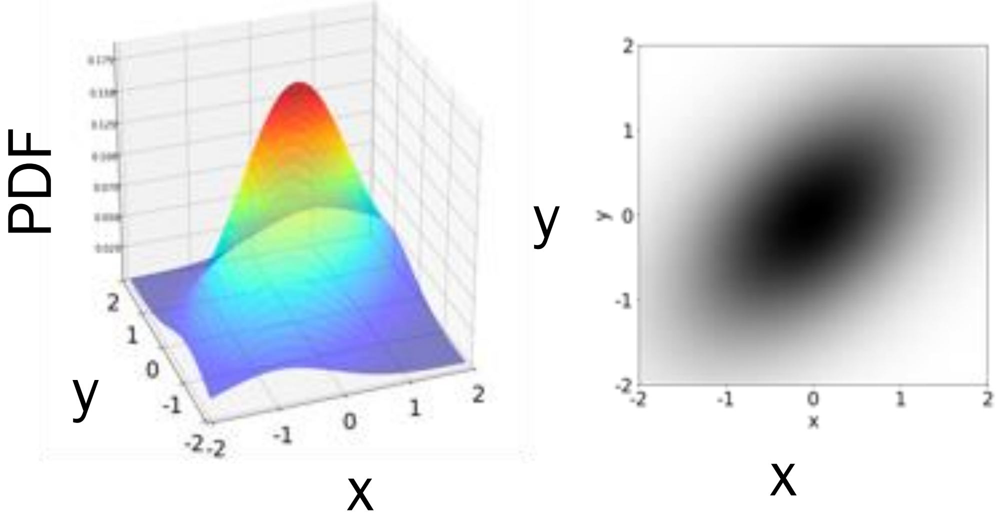
A. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ B. $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

C. $\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ D. $\begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$

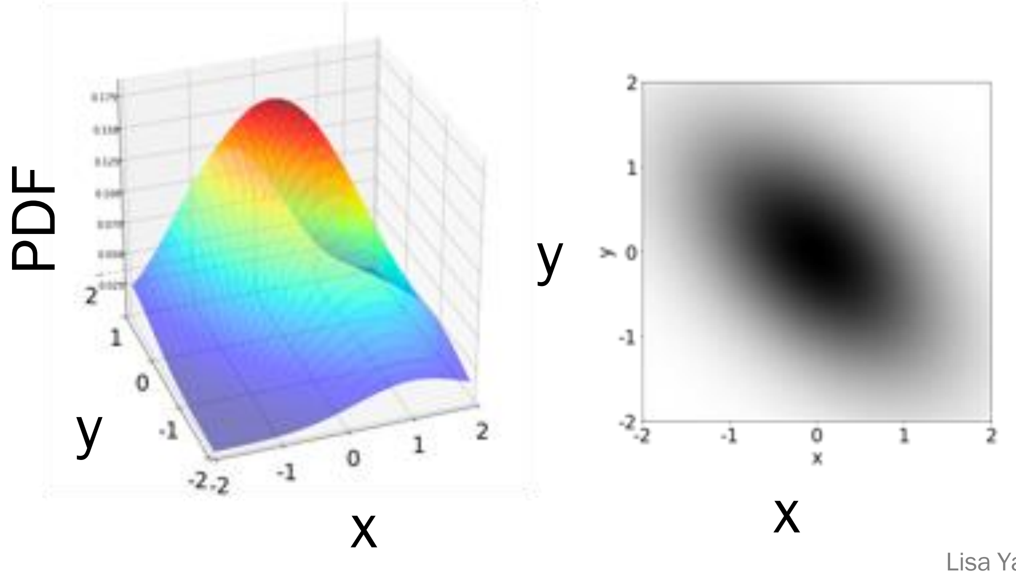
1.



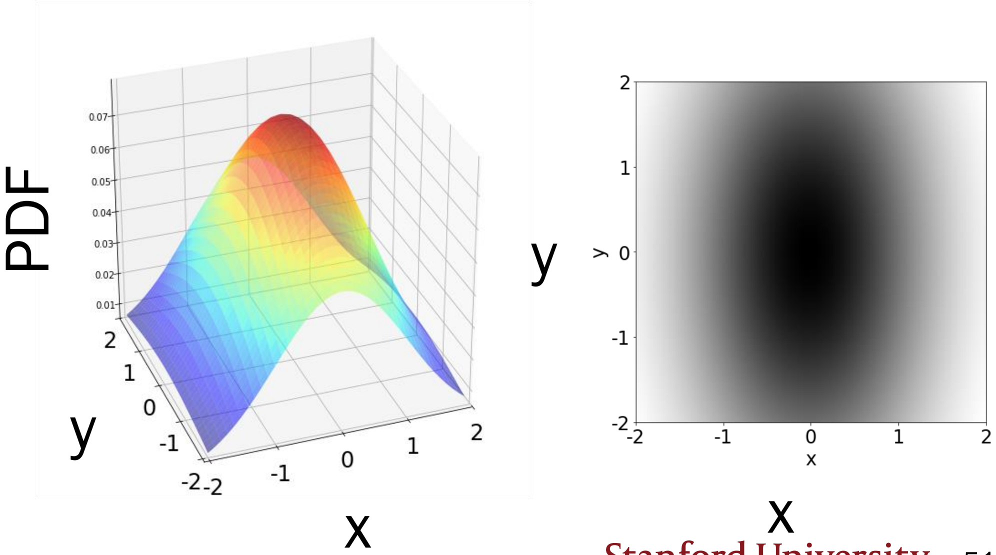
2.



3.



4.



Probabilities from joint CDFs

Recall for a single RV X with CDF F_X:

CDF: $P(X \leq x) = F_X(x)$

$$P(a < X \leq b) = F_X(b) - F(a)$$

For two RVs X and Y with joint CDF $F_{X,Y}$:

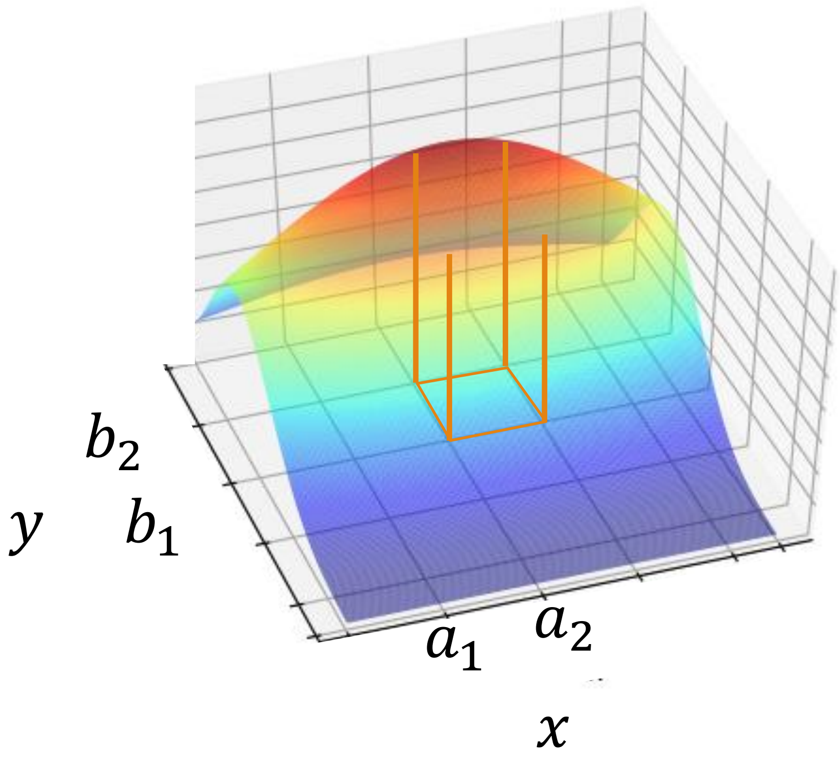
Joint CDF: $P(X \leq x, Y \leq y) = F_{X,Y}(x, y)$

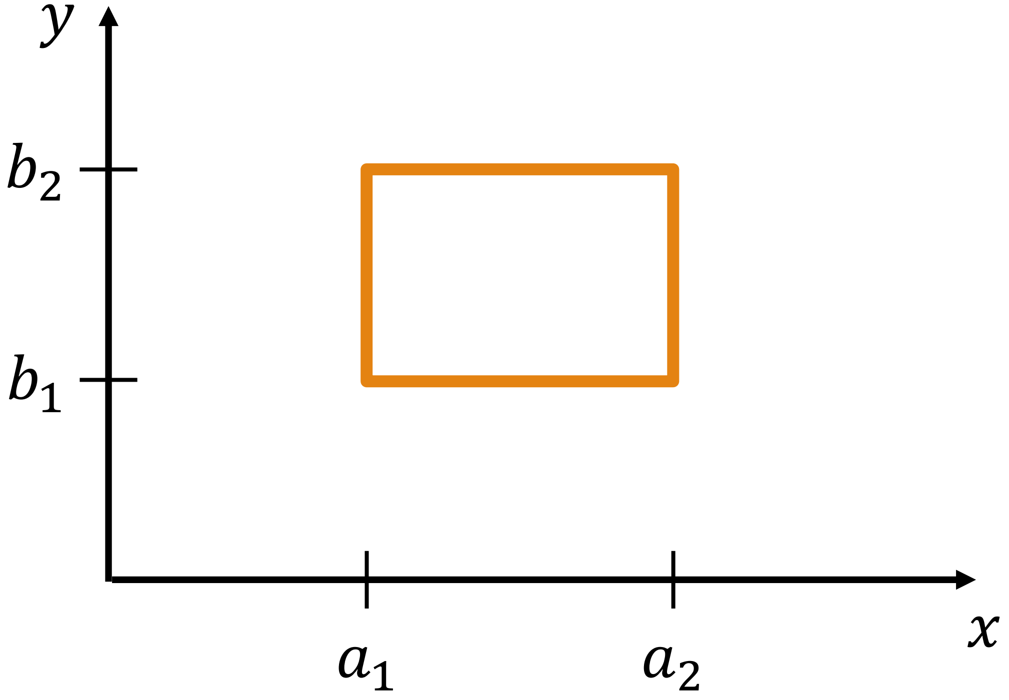
$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) =$$
$$F_{X,Y}(a_{2,}b_2) - F_{X,Y}(a_{1,}b_2) - F_{X,Y}(a_{2,}b_1) + F_{X,Y}(a_{1,}b_1)$$

Note strict inequalities; these properties hold for both discrete and continuous RVs.

Probabilities from joint CDFs

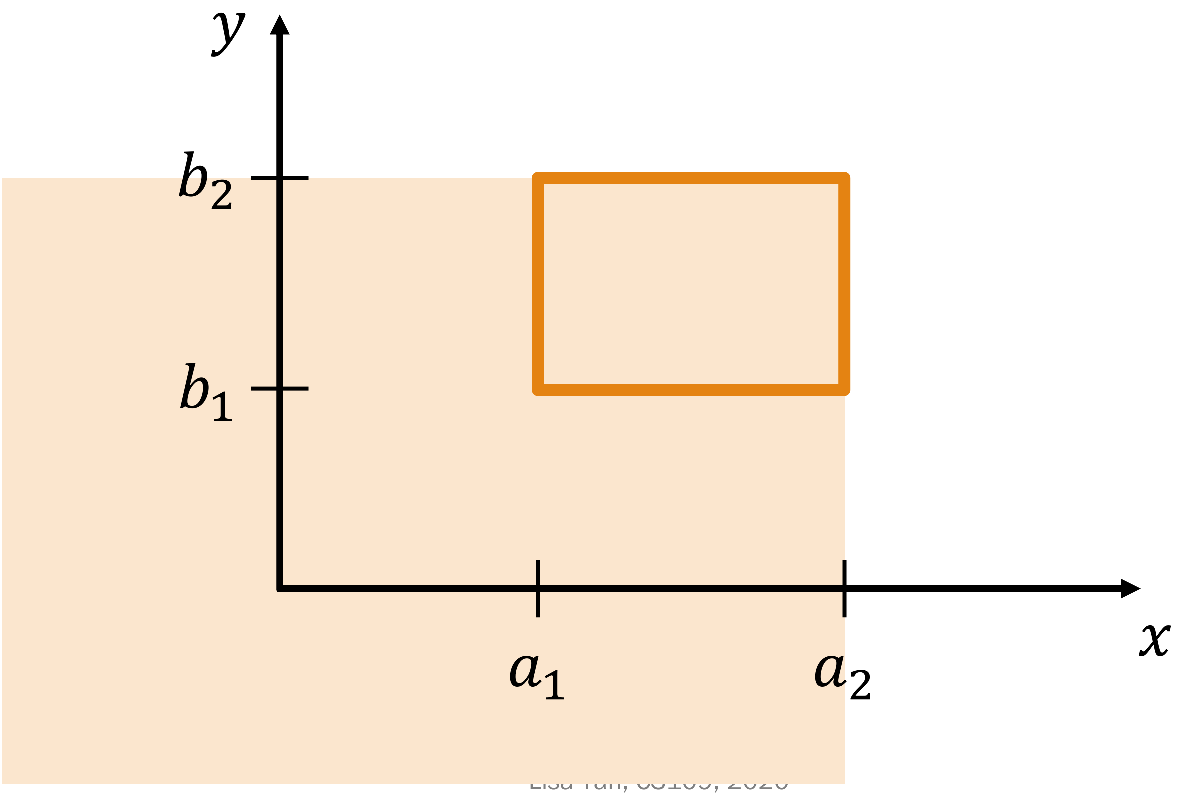
$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) =$$
$$F_{X,Y}(a_2, b_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(a_2, b_1) + F_{X,Y}(a_1, b_1)$$





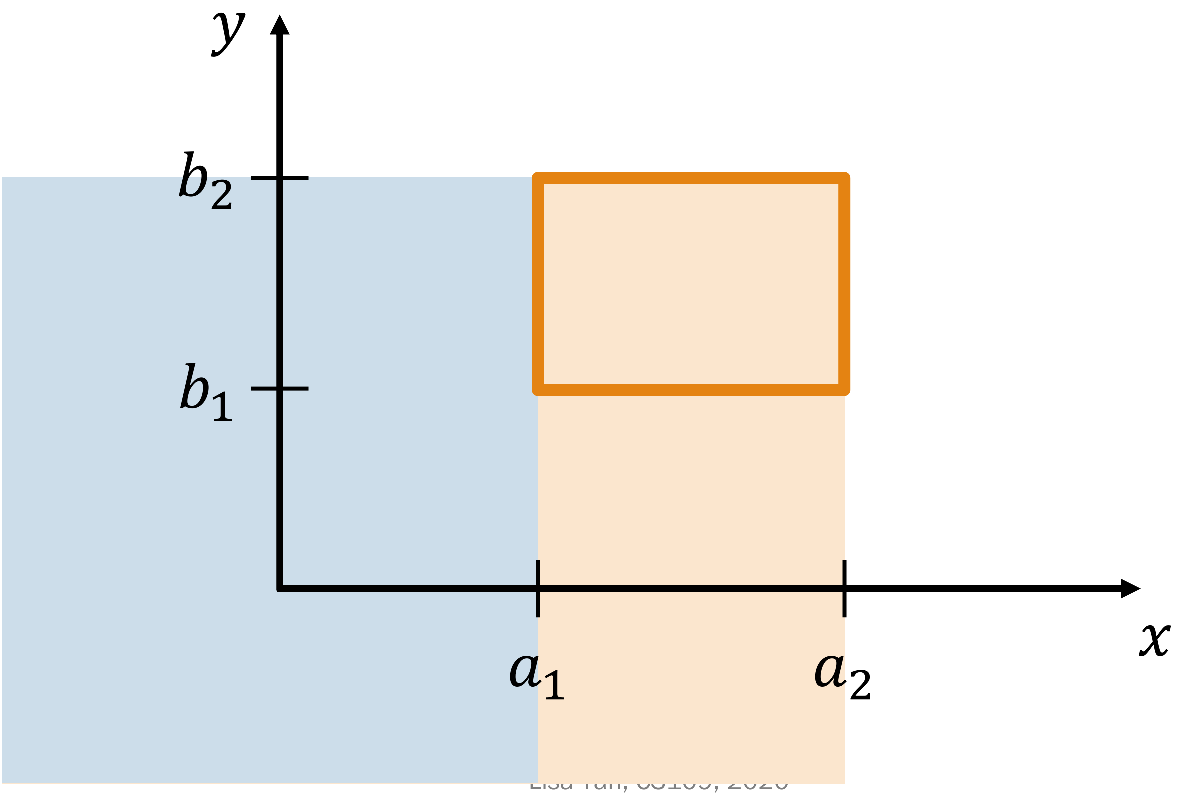
Probabilities from joint CDFs

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



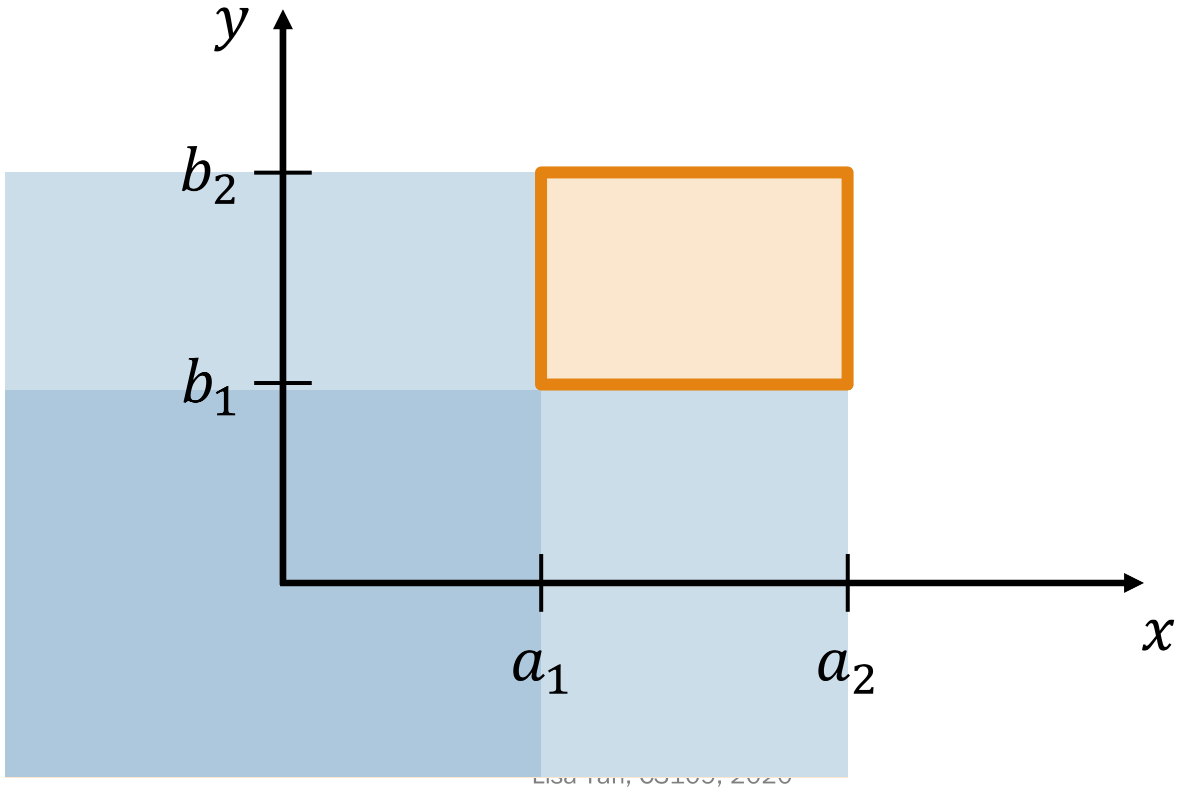
Probabilities from joint CDFs

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



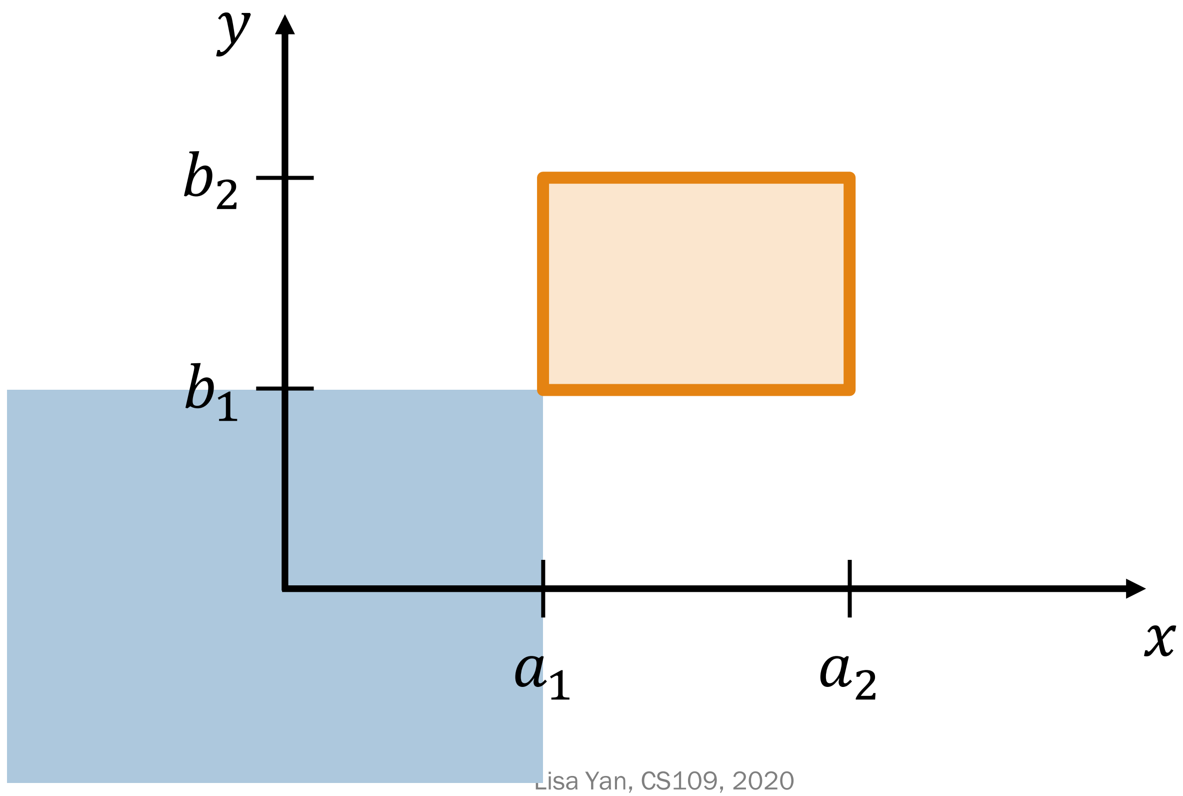
Probabilities from joint CDFs

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



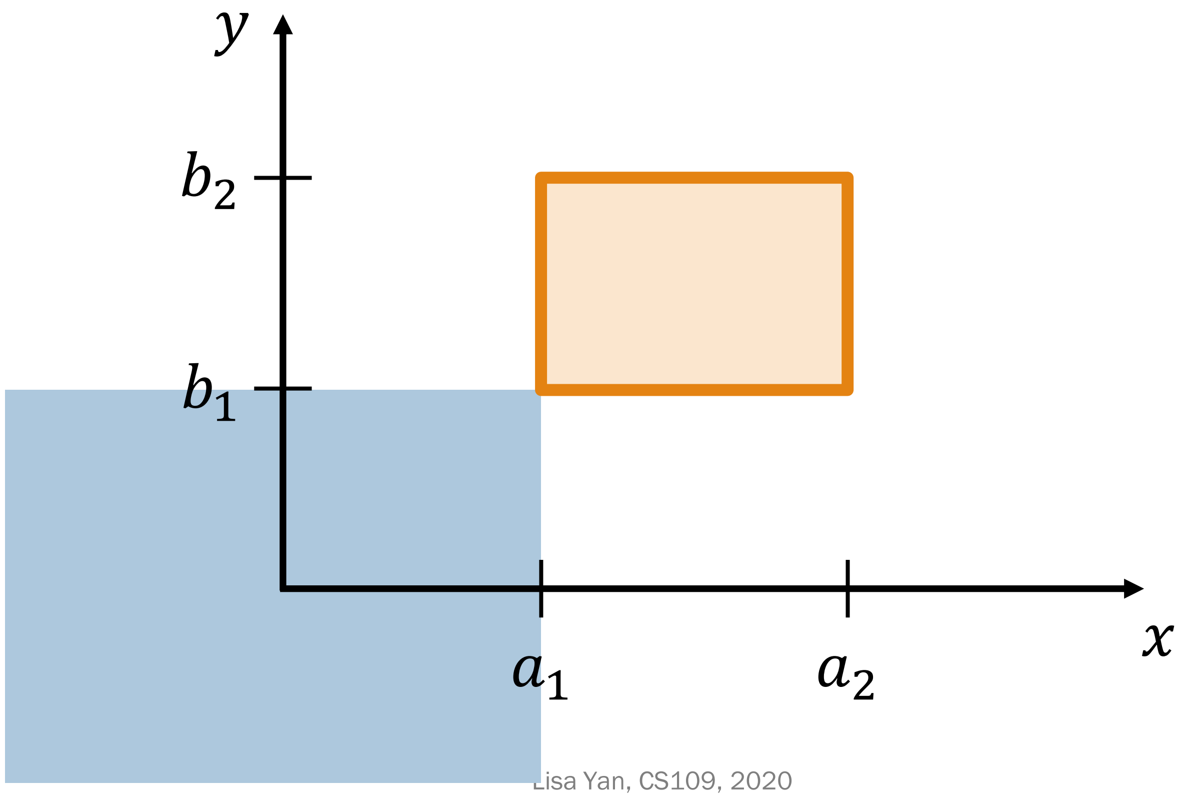
Probabilities from joint CDFs

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



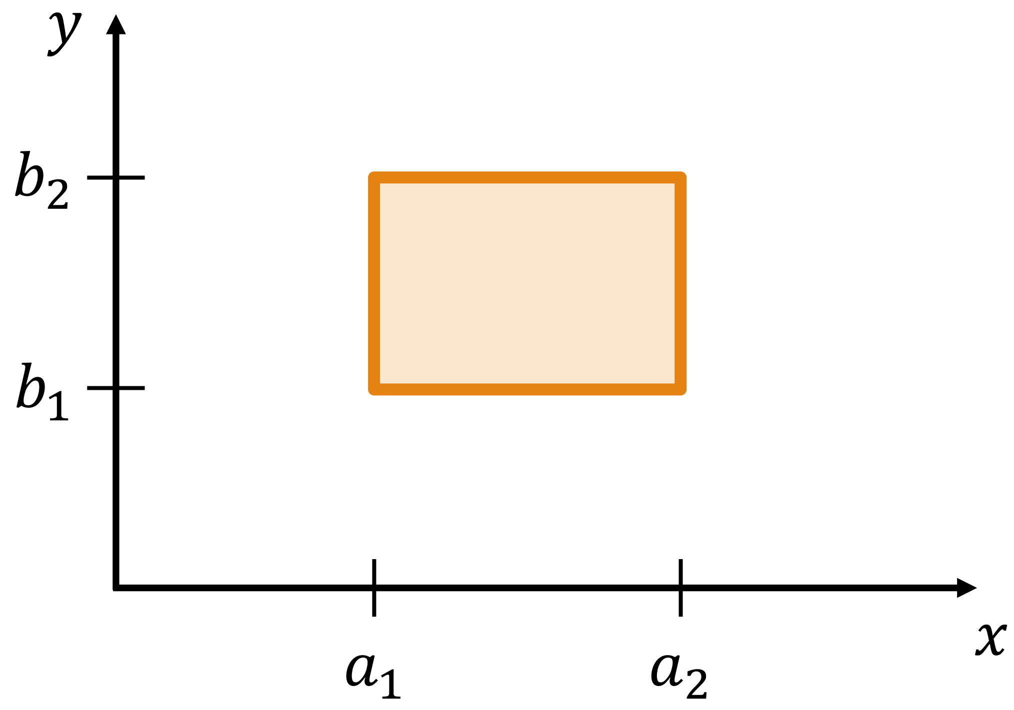
Probabilities from joint CDFs

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



Probabilities from joint CDFs

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



Probability with Instagram!

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

(for next time)

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 blur

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

In a Gaussian blur, for every pixel:

- Weight each pixel by the probability that X and Y are both within the pixel bounds
- The weighting function is a Bivariate Gaussian (Normal) standard deviation parameter σ

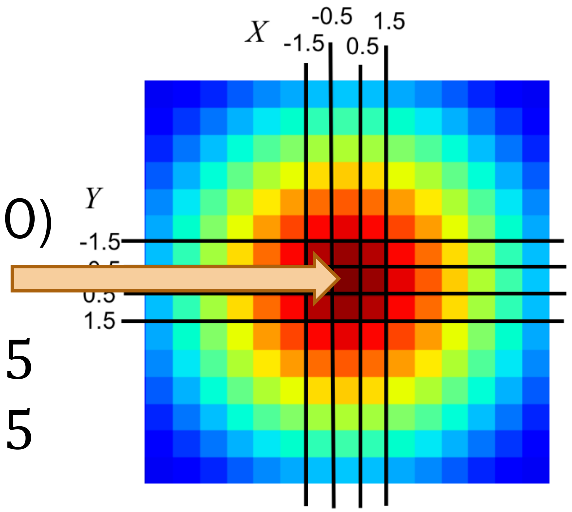
Weight matrix:

Center pixel: (0, 0)

Pixel bounds:

$-0.5 < x \le 0.5$

$-0.5 < y \le 0.5$



Gaussian blurring with $\sigma = 3$:

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

→ Independent $X \sim \mathcal{N}(0, 3^2), Y \sim \mathcal{N}(0, 3^2)$

→ Joint CDF: $F_{X,Y}(x,y) = \Phi\left(\frac{x}{3}\right)\Phi\left(\frac{y}{3}\right)$

What is the weight of the center pixel?

$P(-0.5 < X \le 0.5, -0.5 < Y \le 0.5) =$

$= 0.206$

Next time:

More Cont. Joint and Central Limit Theorem

Extra

1. Integral practice

Let X and Y be two continuous random variables with joint PDF:

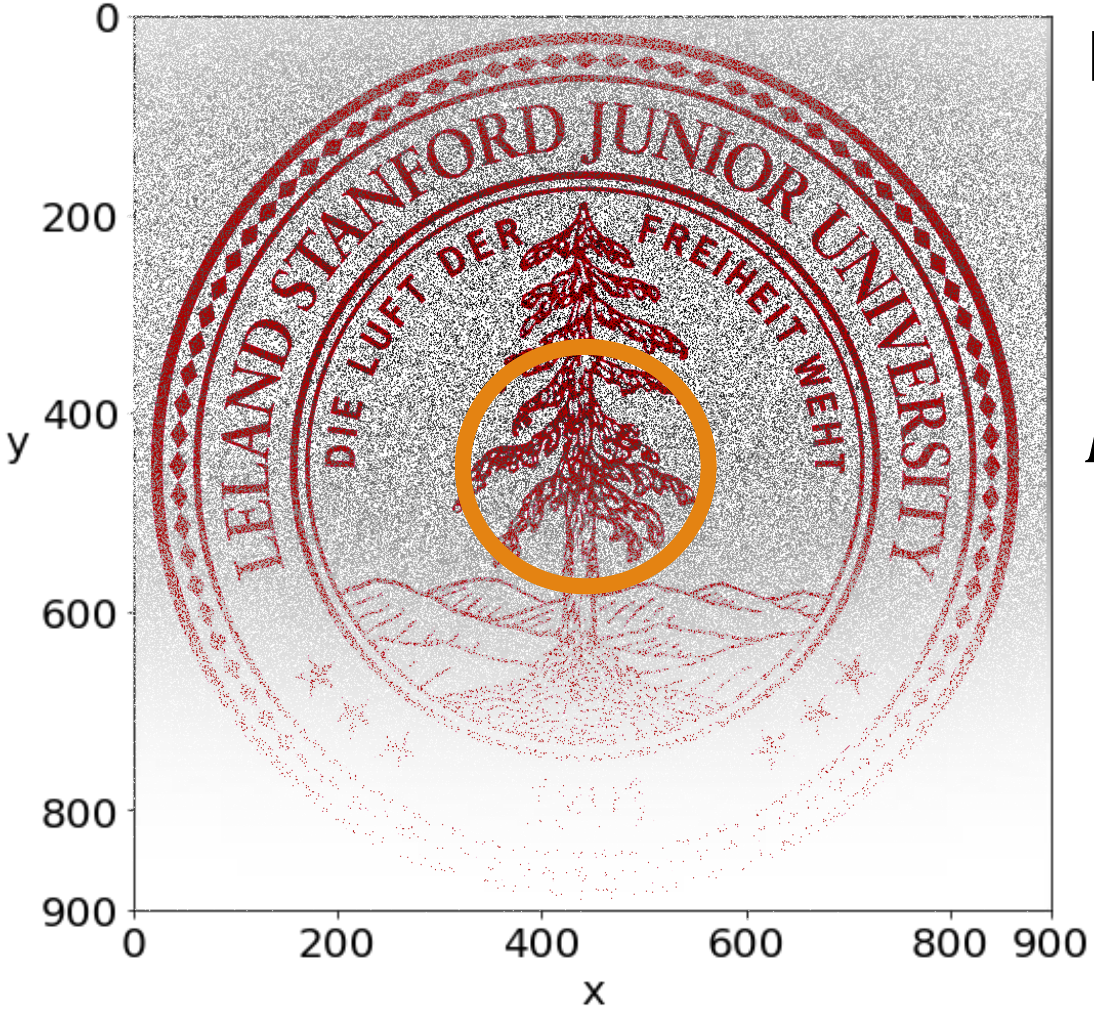
$$f(x,y) = \begin{cases} 4xy & 0 \le x, y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

What is $P(X \le Y)$?

$$P(X \le Y) = \iint_{\substack{x \le y, \\ 0 \le x, y \le 1}} 4xy\, dx\, dy = \int_{y=0}^{1} \int_{x \le y} 4xy\, dx\, dy = \int_{y=0}^{1} \int_{x=0}^{y} 4xy\, dx\, dy$$

$$= \int_{y=0}^{1} 4y \left[\frac{x^2}{2}\right]_0^y dy = \int_{y=0}^{1} 2y^3 dy = \left[\frac{2}{4} y^4\right]_0^1 = \frac{1}{2}$$

2. How do you integrate over a circle?



P(dart hits within $r = 10$ pixels of center)?

$$P(x^2 + y^2 \le 10^2) = \iint_{x^2 + y^2 \le 10^2} f_{X,Y}(x, y)\, dy\, dx$$

Let's try an example that doesn't involve integrating a Normal RV ☺

2. Imperfection on Disk

You have a disk surface, a circle of radius R. Suppose you have a single point imperfection uniformly distributed on the disk.

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi R^2} & x^2 + y^2 \le R^2 \\ 0 & \text{otherwise} \end{cases}$$

What are the marginal distributions of X and Y? Are X and Y independent?

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \frac{1}{\pi R^2} \int_{x^2+y^2 \le R^2} dy \qquad \text{where } -R \le x \le R$$

$$= \frac{1}{\pi R^2} \int_{y=-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy \qquad = \frac{2\sqrt{R^2 - x^2}}{\pi R^2}$$

$$f_Y(y) = \frac{2\sqrt{R^2 - y}}{\pi R^2}$$ where $-R \le y \le R$, by symmetry

No, X and Y are **dependent.**

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