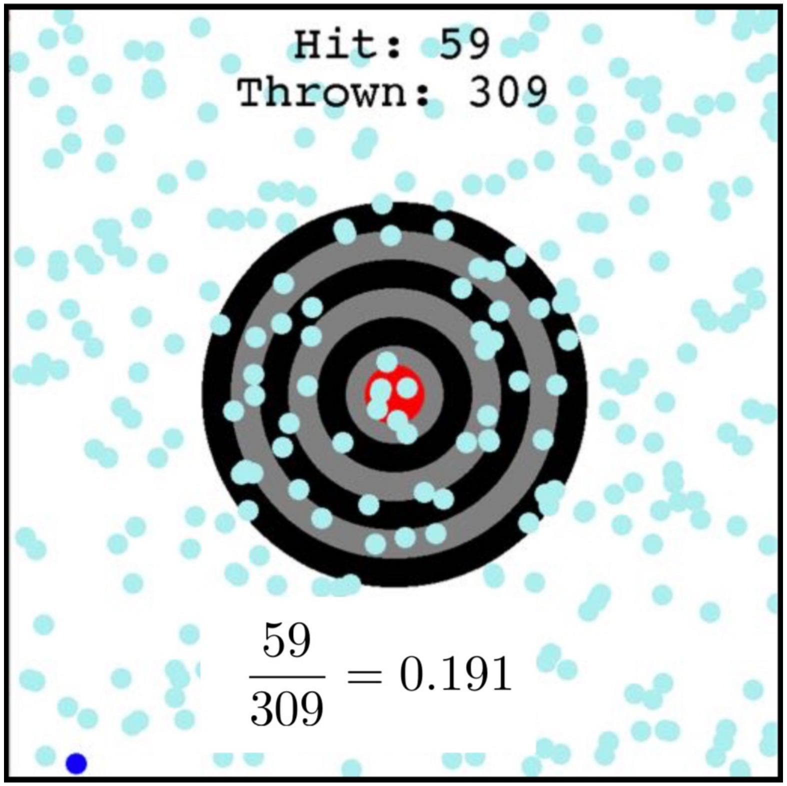
# 16: Continuous Joint Distributions

Jerry Cain

May 2, 2022

# 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.

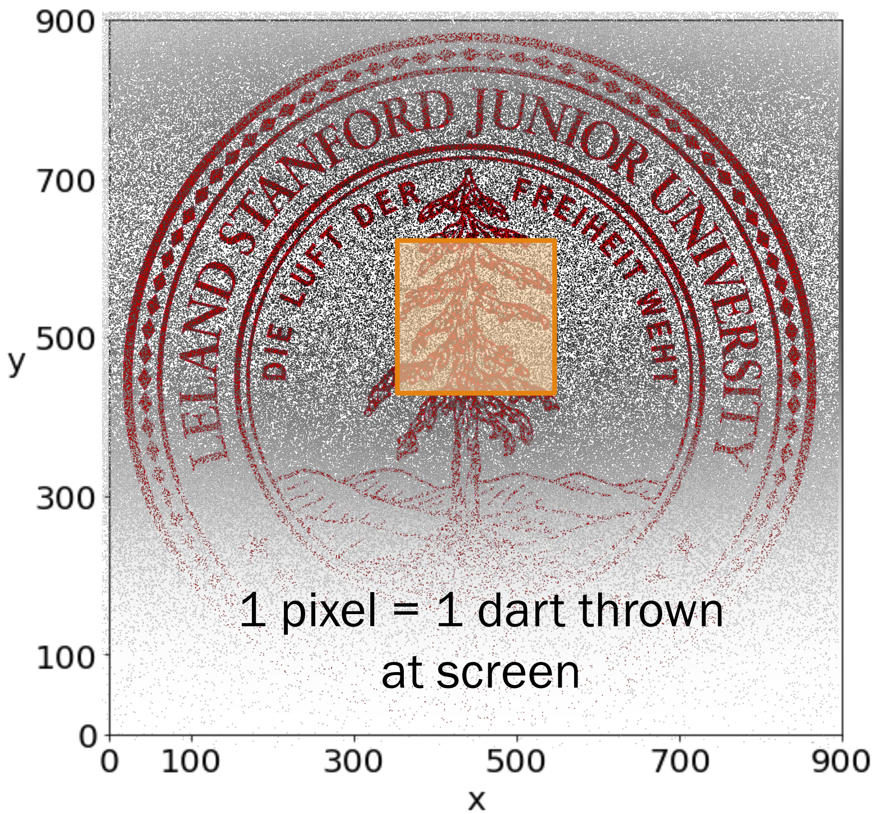
If we throw another dart according to the same distribution, what is

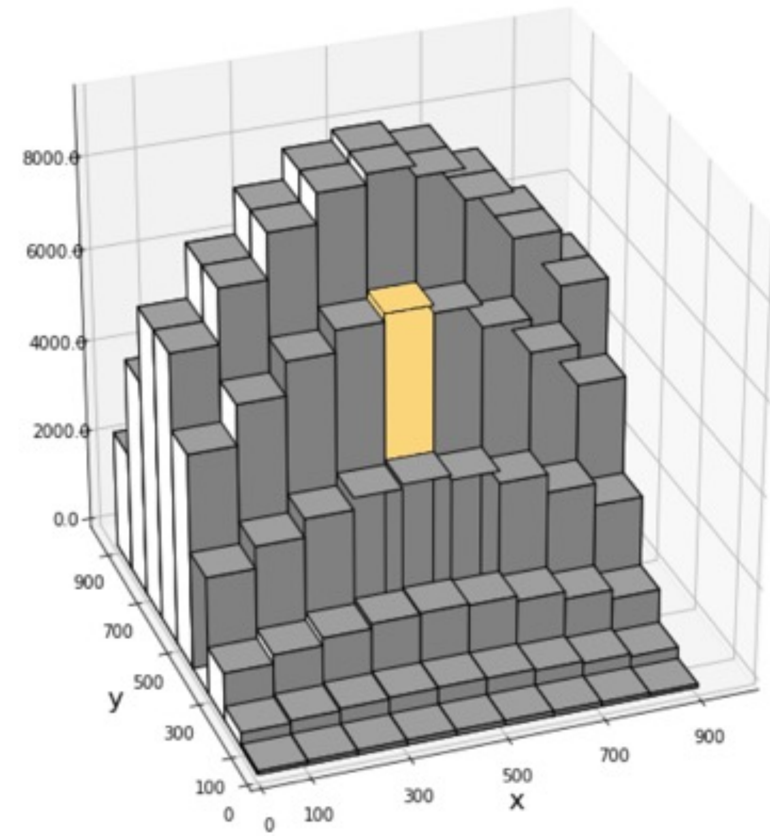
P(dart hits within $r$ pixels of center)?

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)?

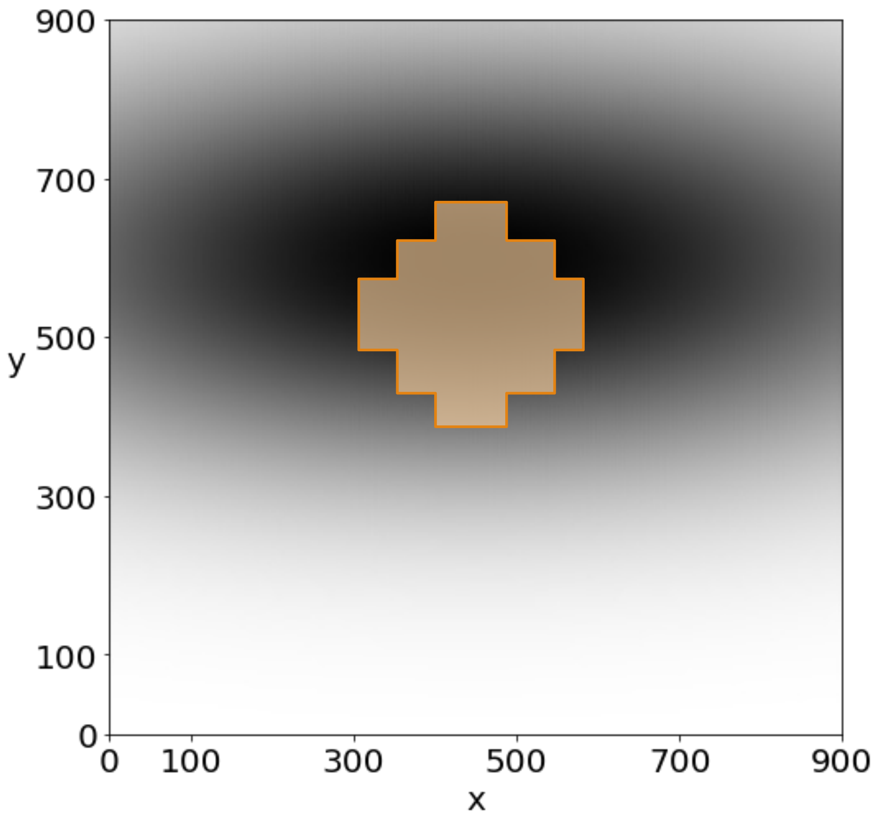


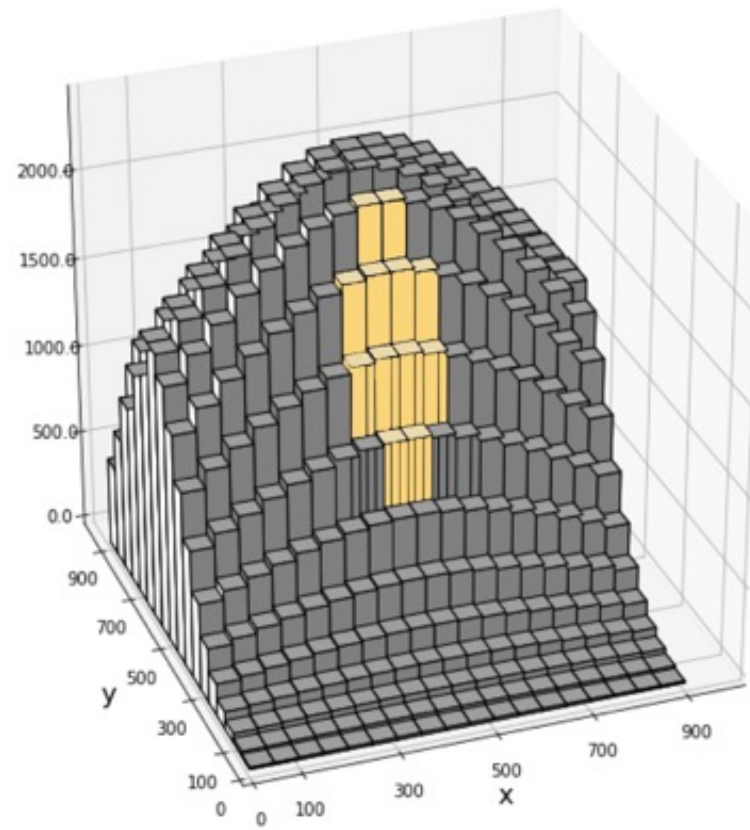




Possible dart counts (in 100x100 boxes)

# CS109 logo with darts

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

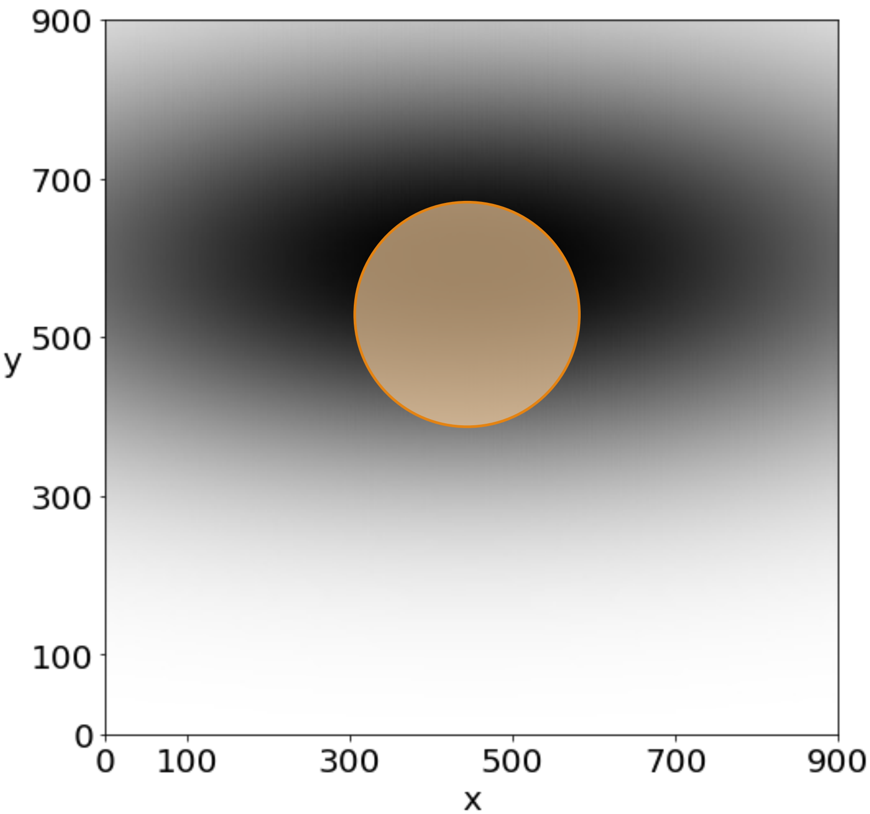


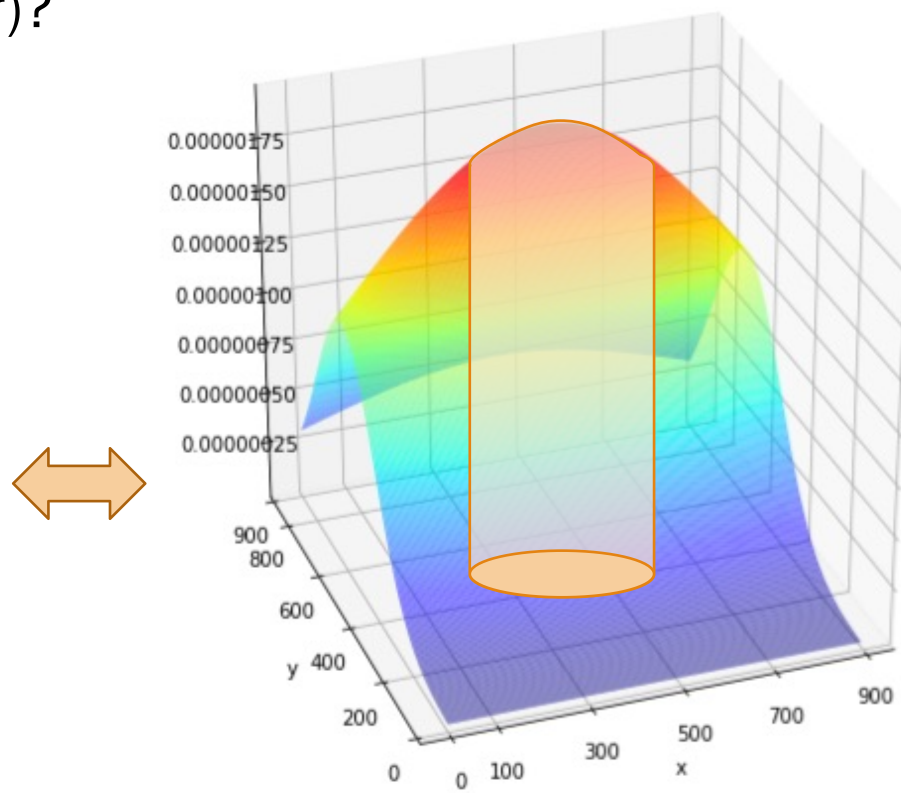




Possible dart counts (in 50x50 boxes)

# 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:

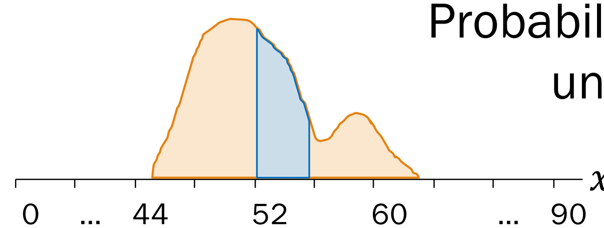
$$P\left(a_1 \leq X \leq a_{2,}\ \ b_1 \leq Y \leq b_2\right) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{X,Y}(x,y) dy\, dx$$

# From one continuous RV to jointly continuous RVs

Single continuous RV $X$

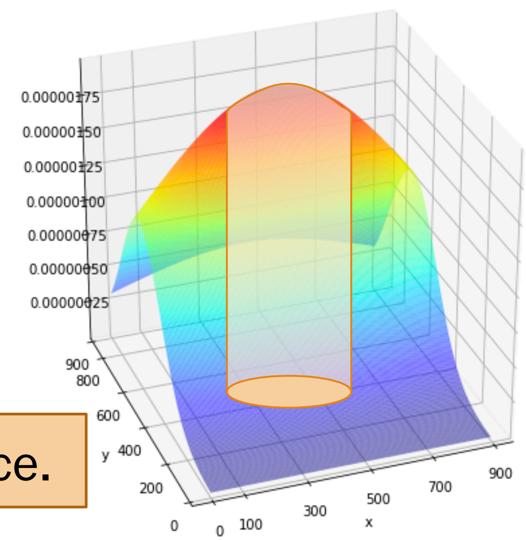
- PDF $f_X$ such that $\int_{-\infty}^{\infty} f_X(x)dx = 1$
- Integrate to get probabilities

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



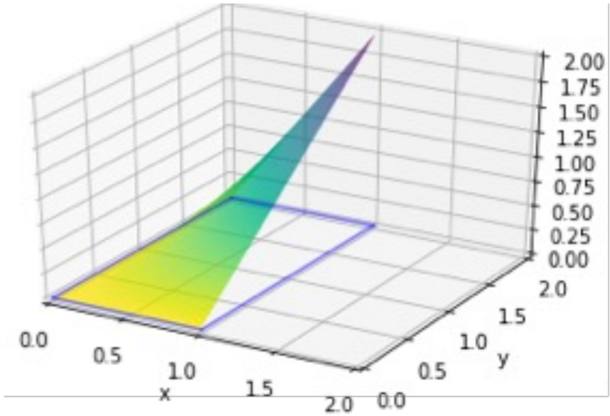

Probability for jointly continuous RVs is **volume** under a surface.

# 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 $Y$?



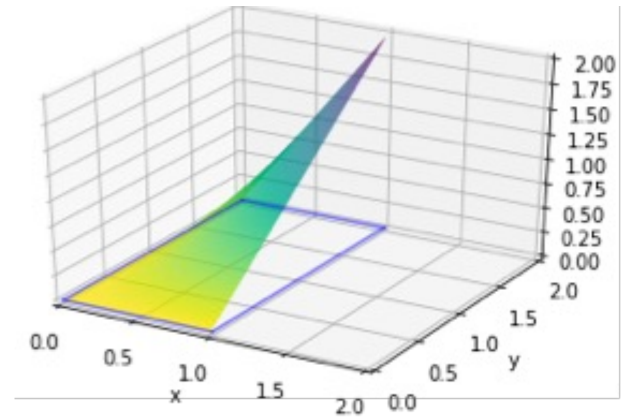

Write down the definite double integral that must integrate to 1:

# Double integrals without tears

Let $X$ and $Y$ be two continuous random variables.

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

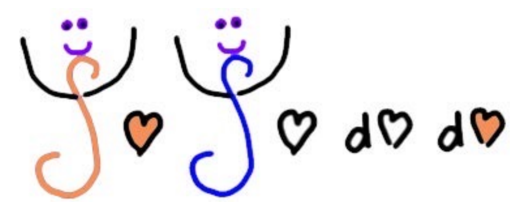
Is $g(x, y) = xy$ a valid joint PDF over $X$ and $Y$?




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

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

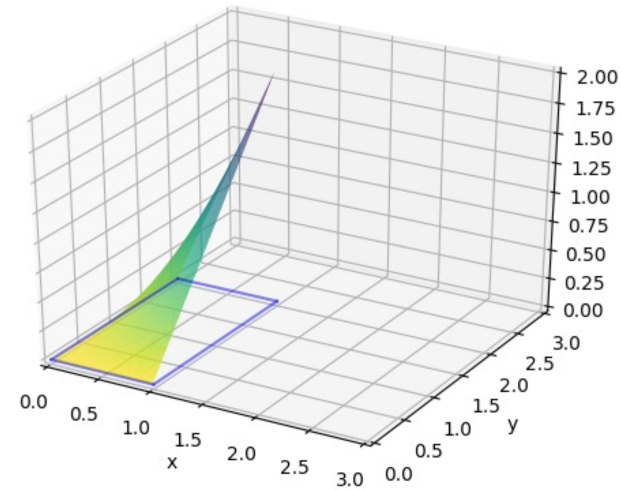
(used in next slide)



# Double integrals without tears

Let $X$ and $Y$ be two continuous random variables.

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




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:

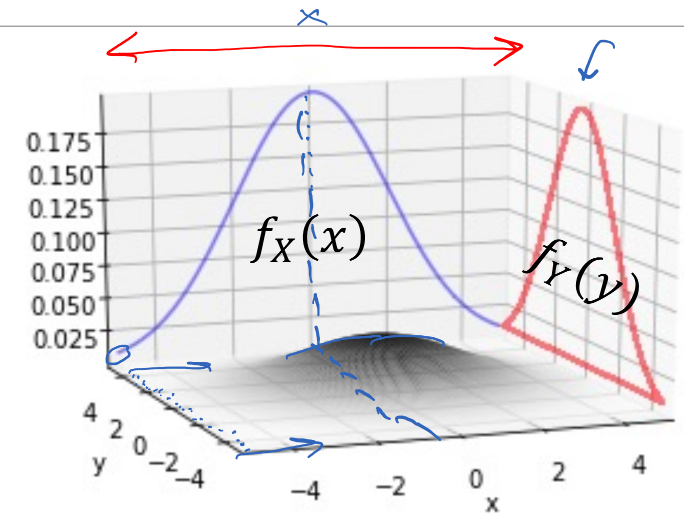
$$\int_{y=0}^{2} y\,\frac{1}{2}dy = \left[\frac{y^2}{4}\right]_{y=0}^{2} = 1 - 0 = 1$$

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:

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

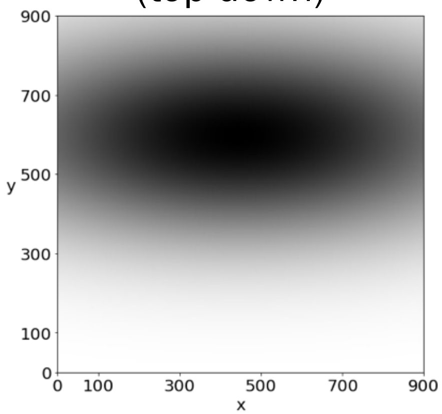



The marginal density functions (**marginal PDFs**) are therefore:

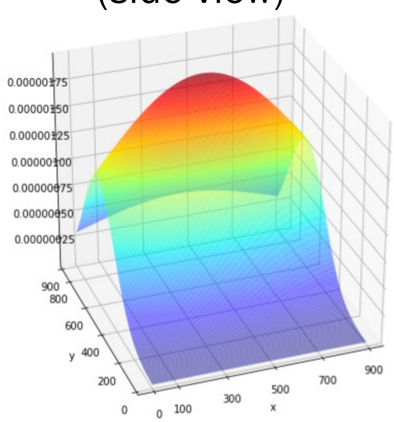
$$f_X(a) = \int_{-\infty}^{\infty} f_{X,Y}(a,y)\,dy \qquad f_Y(b) = \int_{-\infty}^{\infty} f_{X,Y}(x,b)\,dx$$

# Back to darts!

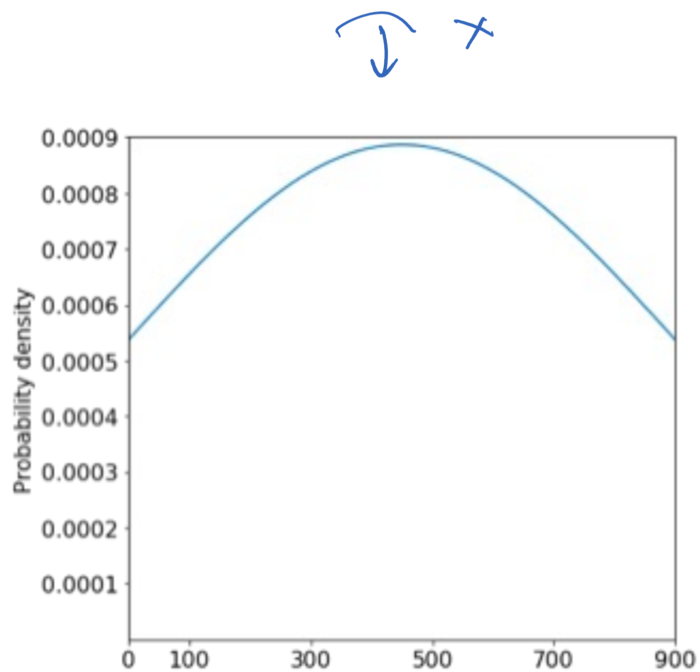
(top-down)

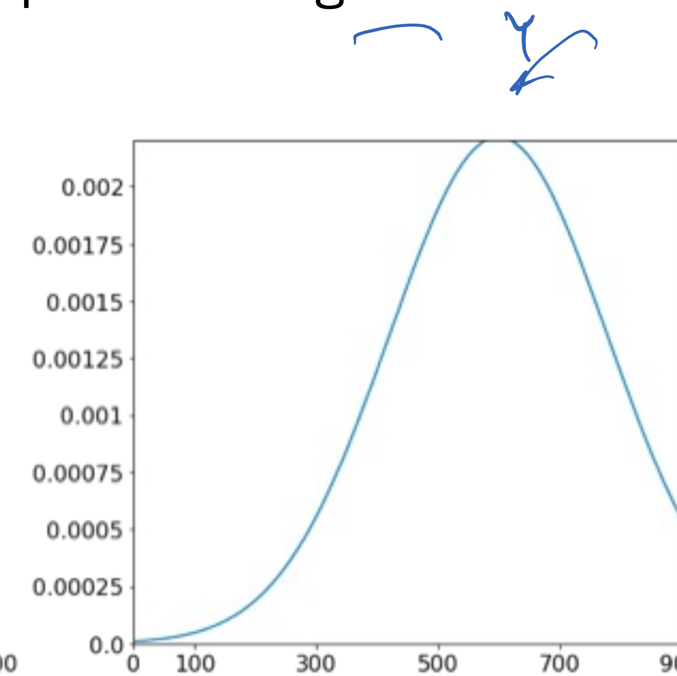


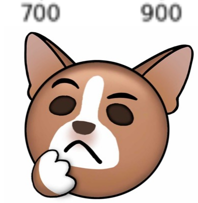

(side view)



Match $X$ and $Y$ to their respective marginal PDFs:

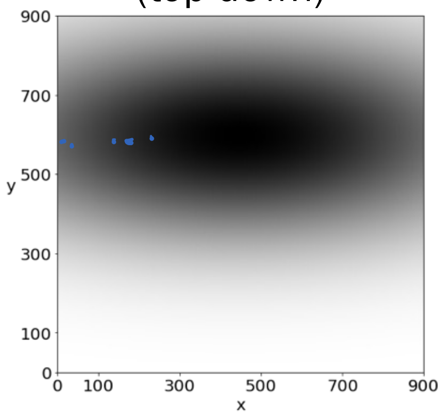


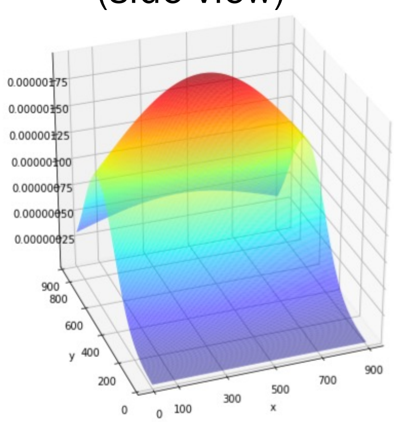



# Back to darts!

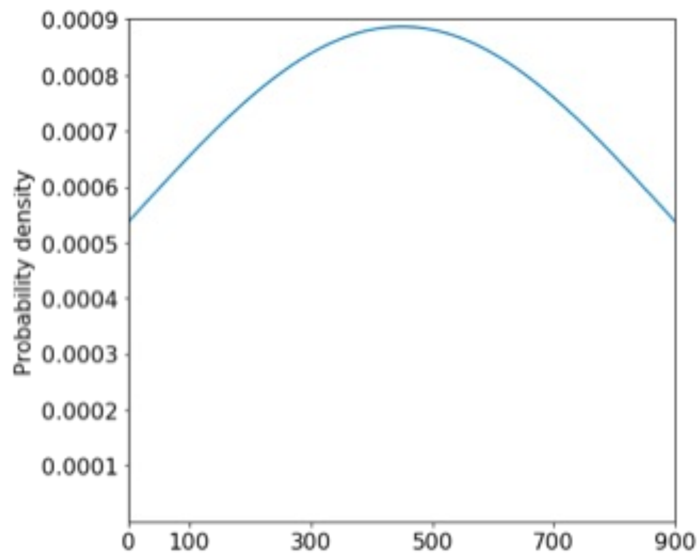
(top-down)



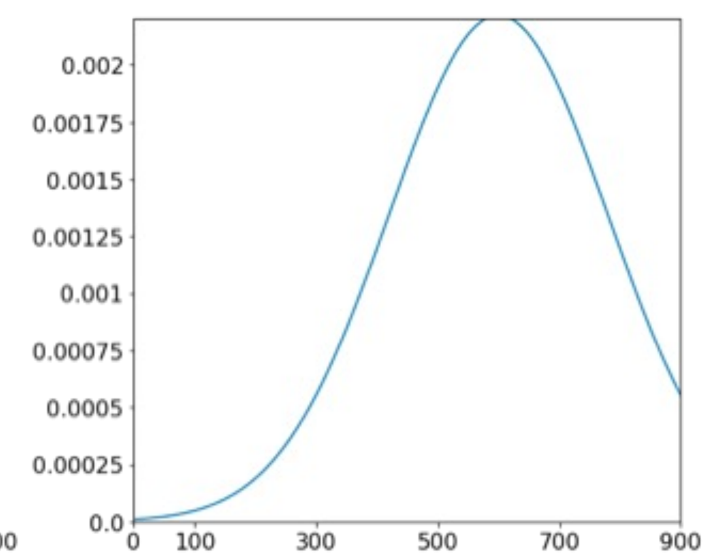

(side view)




Match $X$ and $Y$ to their respective marginal PDFs:




pixel x




pixel y

# 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 $Y$:

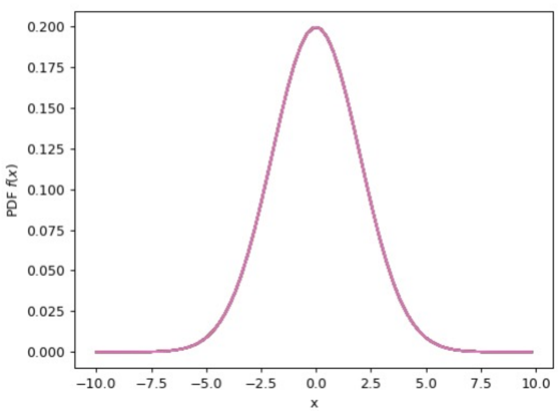
$$F_{X,Y}(a, b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x, y) dy\, dx$$

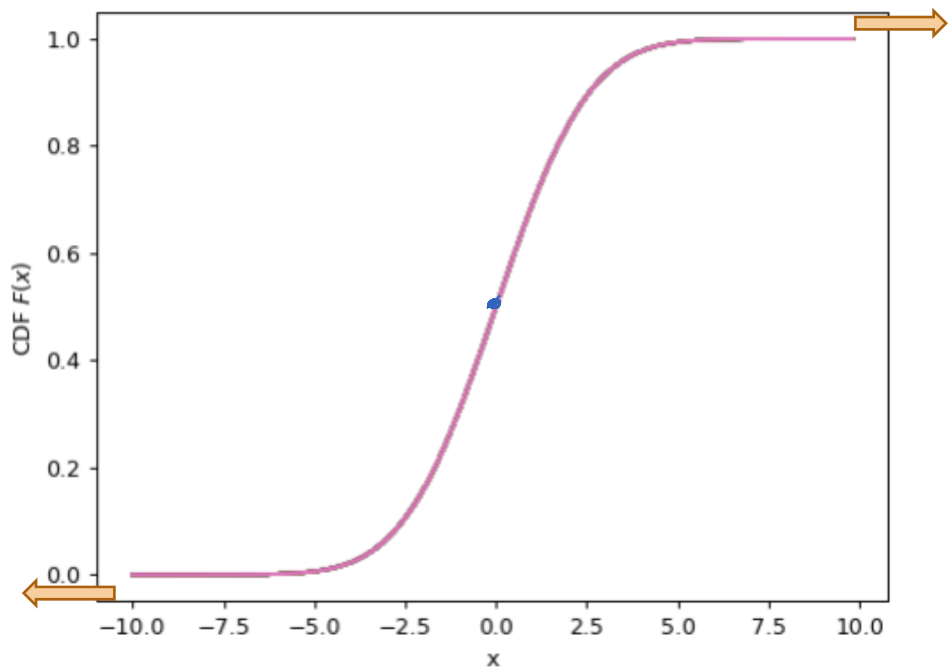
$$f_{X,Y}(a, b) = \frac{\partial^2}{\partial a\, \partial b} F_{X,Y}(a, b)$$

# Single variable CDF, graphically

Review

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



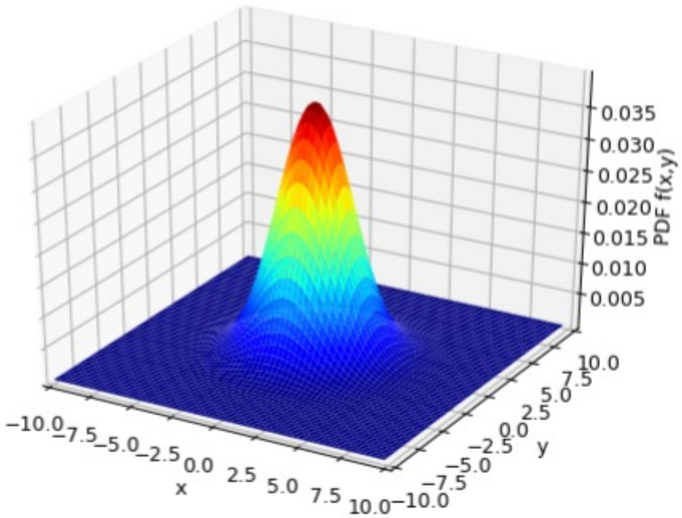

$$\lim_{x \to -\infty} F_X(x) = 0$$

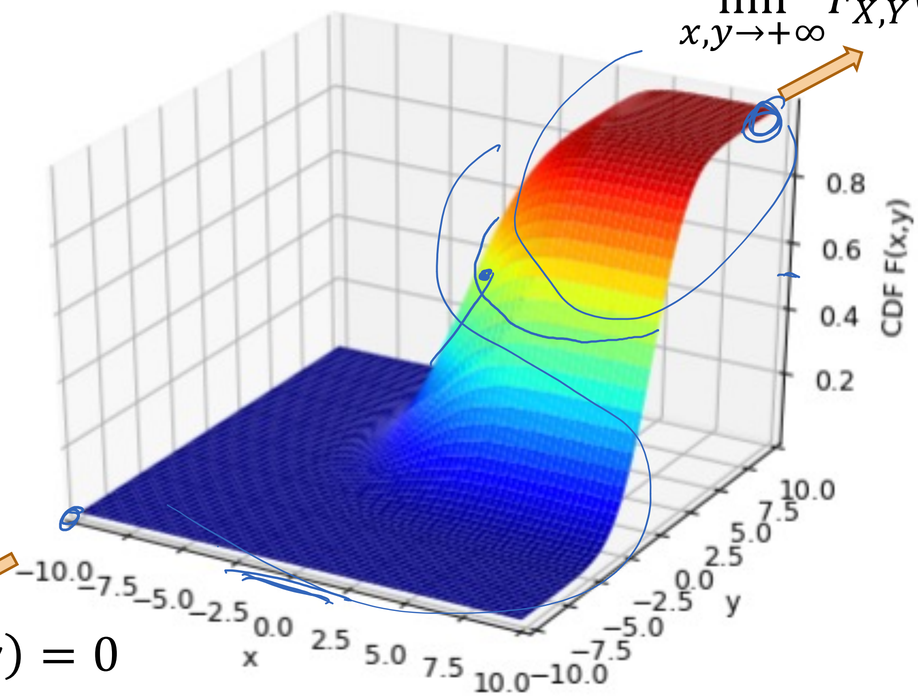
$f_X(x)$

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

# 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:

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

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:

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

must be separable

# 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: $g(x) = 3e^{-3x}$ is valid PDF, $h(y) = 2e^{-2y}$

4x vs y

✅ 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.

# More pop quiz! (still kidding)

$X$ 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 $E[X + Y]$?

# More pop quiz! (still kidding)

$X$ and $Y$ have the following joint PDF:

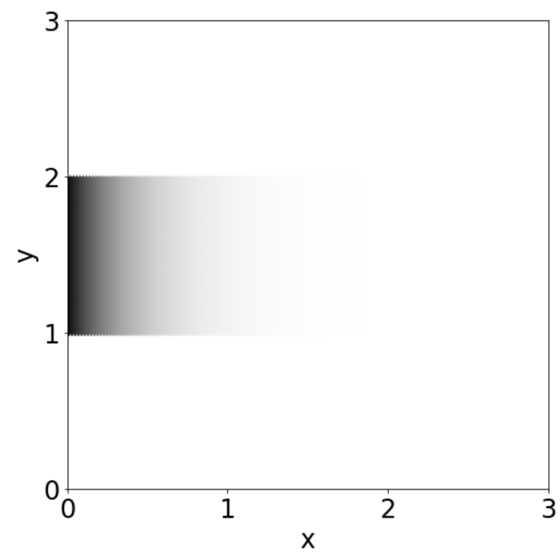
$$f_{X,Y}(x,y) = 3e^{-3x}$$
where $0 < x < \infty, 1 < y < 2$

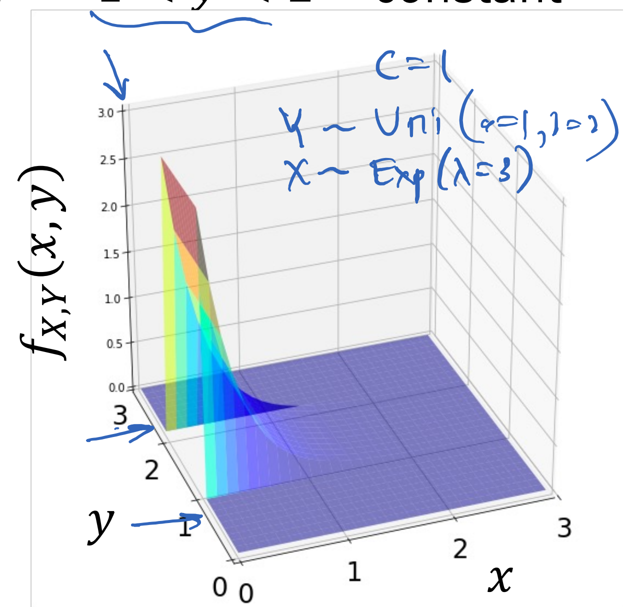
1. Are $X$ and $Y$ independent? ✅

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

$C$ is a constant

2. What is the marginal PDF of $X$? Of $Y$?
3. What is $E[X+Y]$?



# More pop quiz! (still kidding)

$X$ 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

$C = 1$

$X \sim Exp(\lambda = 3)$
$Y \sim Uni(a=1, b=2)$

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

$g(x) = 3e^{-3x}$
$h(y) = 1$

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

Approach 1: LOTUS $E[X+Y] = \int_0^\infty \int_1^2 (x+y)\, dy\, dx$

Approach 2: $E[X+Y] = E[X] + E[Y]$ ($\lambda = 3$) $= 1/3 + 3/2 = 11/6$

# 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:

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

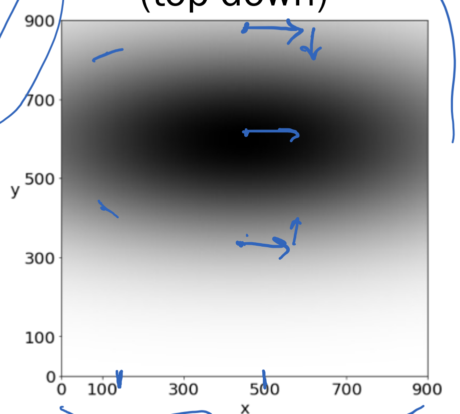
- Vector $\boldsymbol{X} = (X_1, X_2)$
- Mean vector $\boldsymbol{\mu} = (\mu_1, \mu_2)$, Covariance matrix: $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$

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

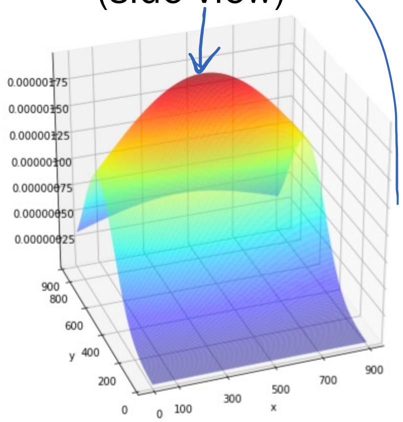
We will focus on understanding the **shape** of a bivariate Normal RV.

# Back to darts

(top-down)




(side view)



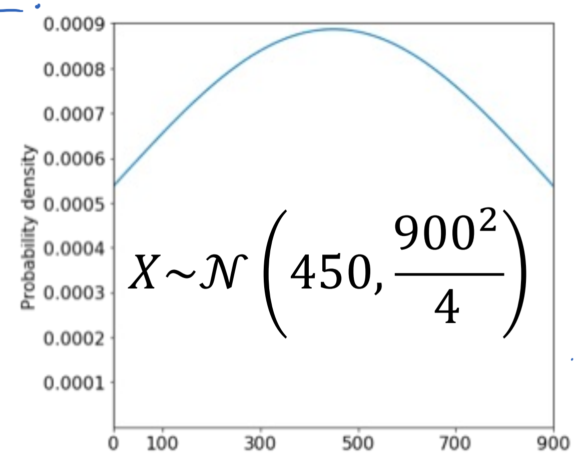
These darts were actually thrown according to a bivariate normal distribution:

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

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

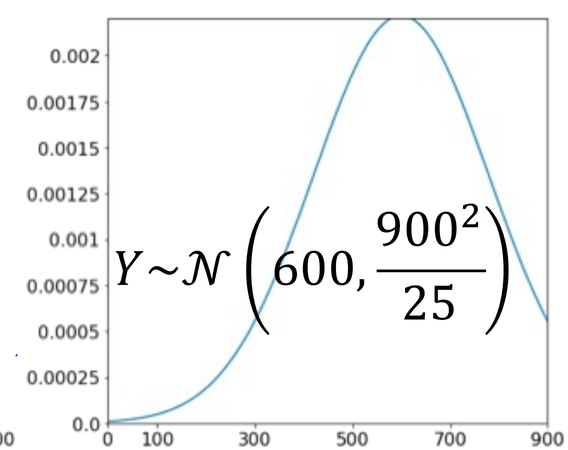
$$\boldsymbol{\Sigma} = \begin{bmatrix} 900^2/4 & 0 \\ 0 & 900^2/25 \end{bmatrix}$$

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

$f(x,y) = g(x)h(y)$

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: $\rho\sigma_1\sigma_2 = 0$)

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

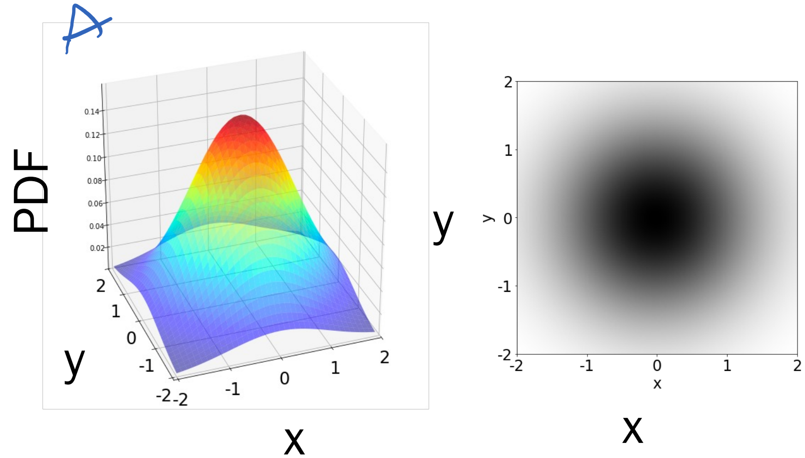
$X_1$ and $X_2$ are **independent** with marginal distributions $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$

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

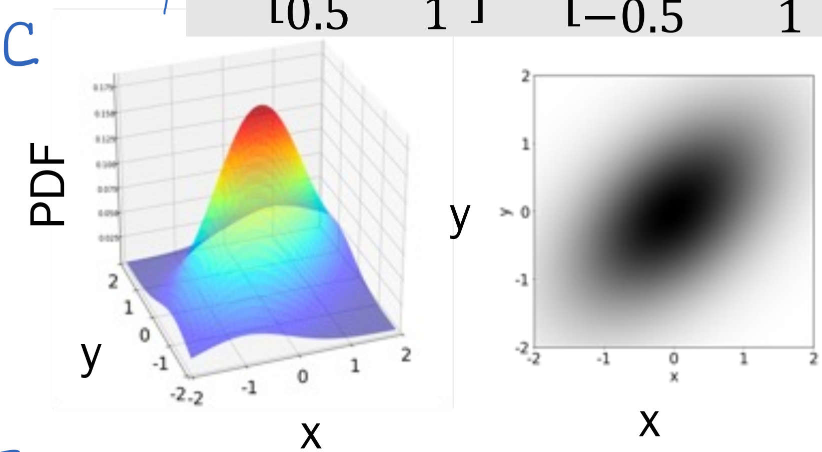
A. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ B. $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

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

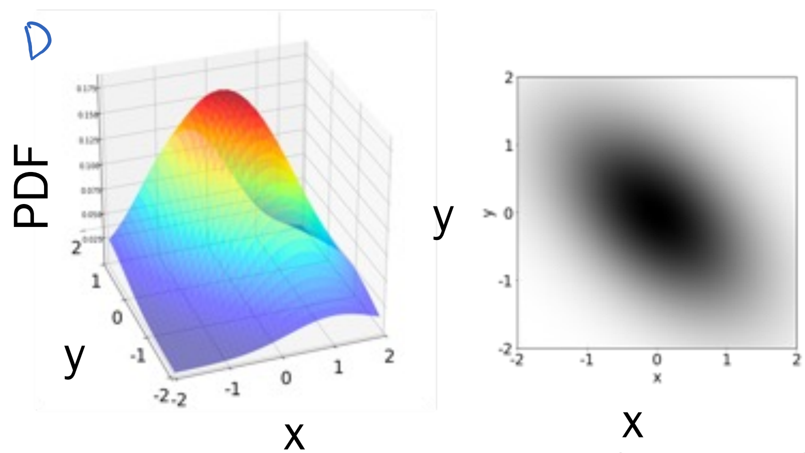
1.



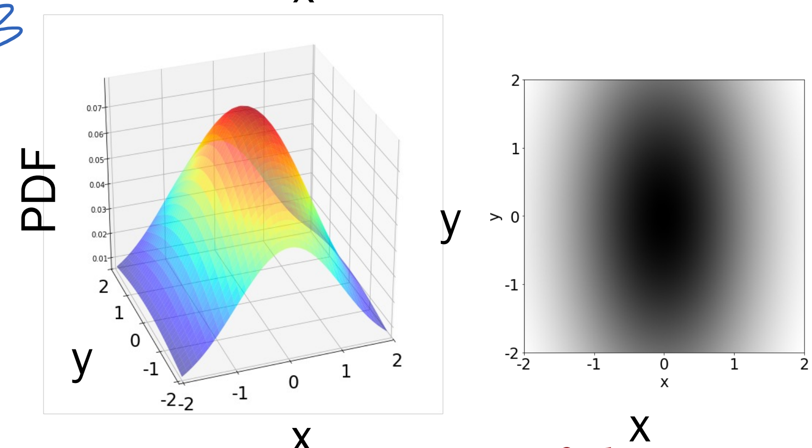

2.




3.




4.

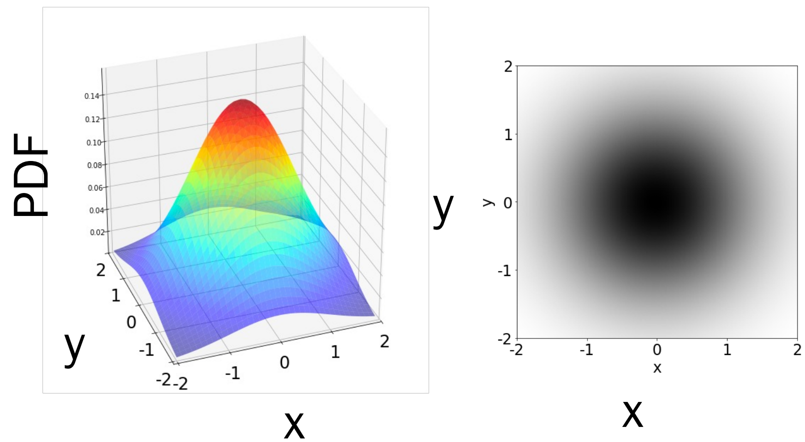


# $(X, Y)$ Matching (all have $\boldsymbol{\mu} = (0, 0)$)

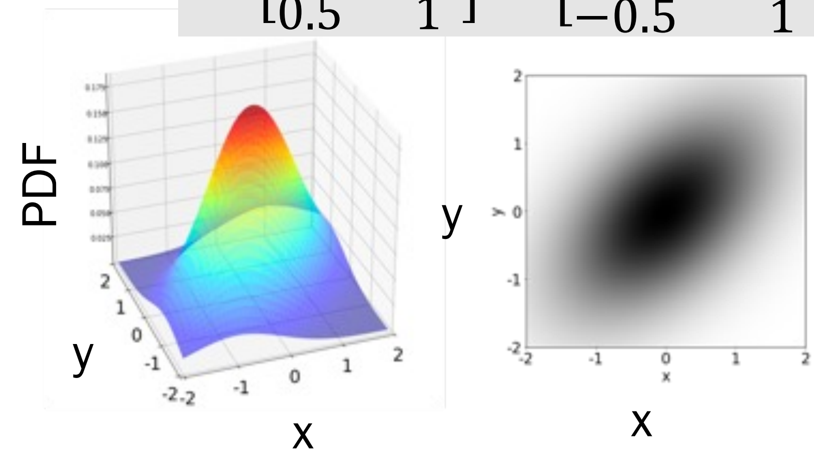
A. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ B. $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

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

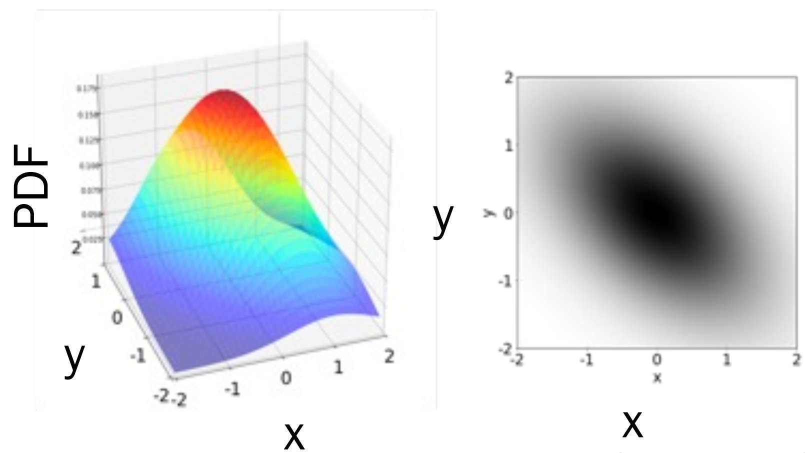
1.



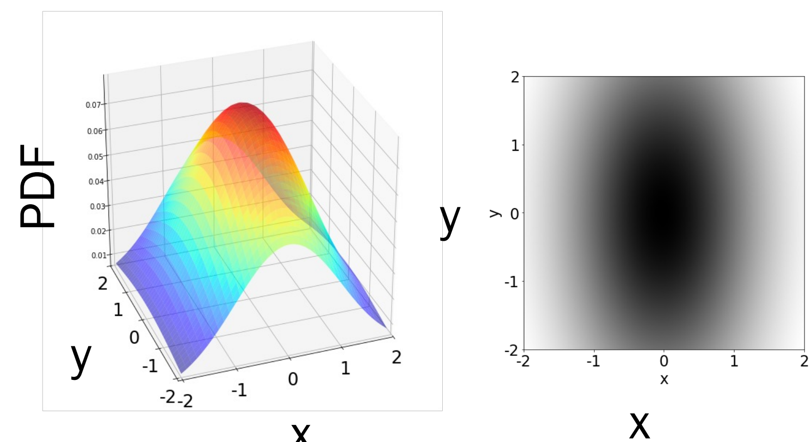

2.



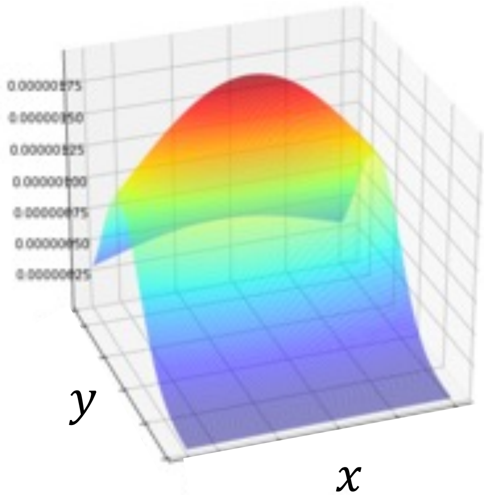

3.




4.



# Why are joint PDFs useful?




Independence
2-D support
Joint PDF
Joint CDF
Marginal PDF
(next time) Conditional PDF

- How 2 continuous RVs vary with each other

$P(X < Y)$,
$\text{Cov}(X,Y), \rho(X,Y)$

- How continuous RV is distributed given evidence (next time)

Given $Y = y$, the distribution of $X$

- How a continuous RV can be decomposed into 2 RVs (or vice versa)

Distribution of $Z = X + Y$
(which is a 1-D RV!)

# 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. Sum of indep. Binomials
B. (approximate) Sum of indep. Poissons
C. (approximate) Sum of indep. Normals
D. None/other

# 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 $P$(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. Sum of indep. Binomials
B. (approximate) Sum of indep. Poissons
C. (approximate) Sum of indep. Normals
D. None/other

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

$X \sim N(20, 18)$
$Y \sim N(40, 24)$
$X + Y \sim ??(60, 42)$ N

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

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

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

   $\approx 0.8023$

# Linear transforms vs. independence

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = X + X$. What is the distribution of $Y$?

- Are both approaches valid?

**Independent RVs approach**

Let $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$
be independent.
Then $Y = X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

**Linear transform approach**

Let $X \sim \mathcal{N}(\mu, \sigma^2)$.
If $Y = aX + b$,
then $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

# Linear transforms vs. independence

Let $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = X + X$. What is the distribution of $Y$?

- Are both approaches valid?

**Independent RVs approach** ❌

Let $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$
be independent.
Then $Y = X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

$Y = X + X$

$X + X \sim \mathcal{N}(\mu + \mu, \sigma^2 + \sigma^2)$?

$Y \sim \mathcal{N}(2\mu, 2\sigma^2)$?

$X$ is NOT independent of $X$!

**Linear transform approach** ✅

Let $X \sim \mathcal{N}(\mu, \sigma^2)$.
If $Y = aX + b$,
then $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

$Y = 2X$

$Y \sim \mathcal{N}(2\mu, 4\sigma^2)$

For independent $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2), X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$,
$aX_1 + bX_2 + c \sim \mathcal{N}(a\mu_1 + b\mu_2 + c, a^2\sigma_1^2 + b^2\sigma_2^2)$