

14: Conditional Distributions

Lisa Yan

October 23, 2019

Sum of independent random variables

$$X \sim \text{Bin}(n_1, p), Y \sim \text{Bin}(n_2, p)$$

X, Y independent



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

$$X \sim \text{Poi}(\lambda_1), Y \sim \text{Poi}(\lambda_2)$$

X, Y independent



$$X + Y \sim \text{Poi}(\lambda_1 + \lambda_2)$$

$$X \sim \mathcal{N}(\mu_1, \sigma_1^2)$$

$$Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

X, Y independent



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

Note: these also hold in the **general case** (≥ 2 variables)

Quick questions

1. X and Y have the following joint PDF:
Are X and Y independent?

$$f_{X,Y}(x,y) = \frac{8}{3}x^3y$$

where $0 < x < 1, 1 < y < 2$

2. Let $X \sim \text{Bin}(30, 0.01)$ and $Y \sim \text{Bin}(50, 0.02)$ be independent RVs. Can we use sum of independent Poisson RVs to approximate $P(X + Y = 1)$?



Quick questions

1. X and Y have the following joint PDF:
Are X and Y independent?

$$f_{X,Y}(x,y) = \frac{8}{3}x^3y$$

where $0 < x < 1, 1 < y < 2$

Separable functions $g(x) = C_1x^3$
 $h(y) = C_2y$, where C_1, C_2 are constants

2. Let $A \sim \text{Bin}(30, 0.01)$ and $B \sim \text{Bin}(50, 0.02)$ be independent RVs. Can we use sum of independent Poisson RVs to approximate $P(A + B = 2)$?



Quick questions

2. Let $A \sim \text{Bin}(30, 0.01)$ and $B \sim \text{Bin}(50, 0.02)$ be independent RVs. Can we use sum of independent Poisson RVs to approximate $P(A + B = 2)$?

Sol 1: Approximate as sum of Poissons

$$A \approx X \sim \text{Poi}(\lambda_1 = 30 \cdot 0.01 = 0.3)$$

$$B \approx Y \sim \text{Poi}(\lambda_2 = 50 \cdot 0.02 = 1)$$

$$P(A + B = 2) \approx P(X + Y = 2) = \frac{\lambda^2}{2!} e^{-\lambda}$$

$$X + Y \sim \text{Poi}(\lambda = \lambda_1 + \lambda_2 = 1.3) \approx 0.2302$$

Sol 2: No approximation

$$P(A + B = 2) = \sum_{k=0}^2 P(A = k)P(B = 2 - k)$$

$$= \sum_{k=0}^2 \binom{30}{k} 0.01^k (0.99)^{30-k} \binom{50}{2-k} 0.02^{2-k} 0.98^{48+k} \approx 0.2327$$



Today's plan

→ Sum of two Uniform independent RVs

Expectation of sum of two RVs

midterm content up to here

Discrete conditional distributions

Ratio of probabilities interlude

Continuous conditional distributions

Recall 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

Dance, Dance, Convolution Extreme

Recall 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



Sum of independent Uniforms

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

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent random variables.

What is the distribution of $X + Y$, f_{X+Y} ?

$$f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} \underbrace{f_X(k) f_Y(\alpha - k)} dk$$

$f_X(k) f_Y(\alpha - k) = 1$ when: (select one)

- A. always
- B. $0 \leq k \leq 1$
- C. $0 \leq \alpha - k \leq 1$
- D. $0 \leq \alpha \leq 2$
- E. Other



Sum of independent Uniforms

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

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent random variables.

What is the distribution of $X + Y$, f_{X+Y} ?

$$f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} \underbrace{f_X(k) f_Y(\alpha - k)} dk$$

$f_X(k) f_Y(\alpha - k) = 1$ when: (select one)

- A. always
- B. $0 \leq k \leq 1$
- C. $0 \leq \alpha - k \leq 1$
- D. $0 \leq \alpha \leq 2$
- E. Other



Sum of independent Uniforms

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

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent random variables.

What is the distribution of $X + Y$, f_{X+Y} ?

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

$f_X(k) f_Y(\alpha - k) = 1$ when:

$$0 \leq \alpha \leq 2 \text{ and}$$

$$0 \leq k \leq 1 \text{ and}$$

$$0 \leq \alpha - k \leq 1 \Rightarrow \alpha - 1 \leq k \leq \alpha$$

The precise integration
bounds on k depend on α .

What are the bounds on k when:

1. $\alpha = 1/2$?

2. $\alpha = 3/2$?



Sum of independent Uniforms

$$\begin{array}{l} X \text{ and } Y \\ \text{independent} \\ + \text{ continuous} \end{array} f_{X+Y}(\alpha) = \int_{-\infty}^{\infty} f_X(x)f_Y(\alpha - x) dx$$

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent random variables.

What is the distribution of $X + Y$, f_{X+Y} ?

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

$f_X(k)f_Y(\alpha - k) = 1$ when:

$$0 \leq \alpha \leq 2 \text{ and}$$

$$0 \leq k \leq 1 \text{ and}$$

$$0 \leq \alpha - k \leq 1 \Rightarrow \alpha - 1 \leq k \leq \alpha$$

The precise integration
bounds on k depend on α .

What are the bounds on k when:

1. $\alpha = 1/2$?

$$0 \leq k \leq \alpha$$

$$\int_{k=0}^{\alpha} 1dk = \alpha = 1/2$$

2. $\alpha = 3/2$?

$$\alpha - 1 \leq k \leq 1$$

$$\int_{k=\alpha-1}^1 1dk = 2 - \alpha = 1/2$$



Sum of independent Uniforms

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

Let $X \sim \text{Uni}(0,1)$ and $Y \sim \text{Uni}(0,1)$ be independent random variables.

What is the distribution of $X + Y$, f_{X+Y} ?

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

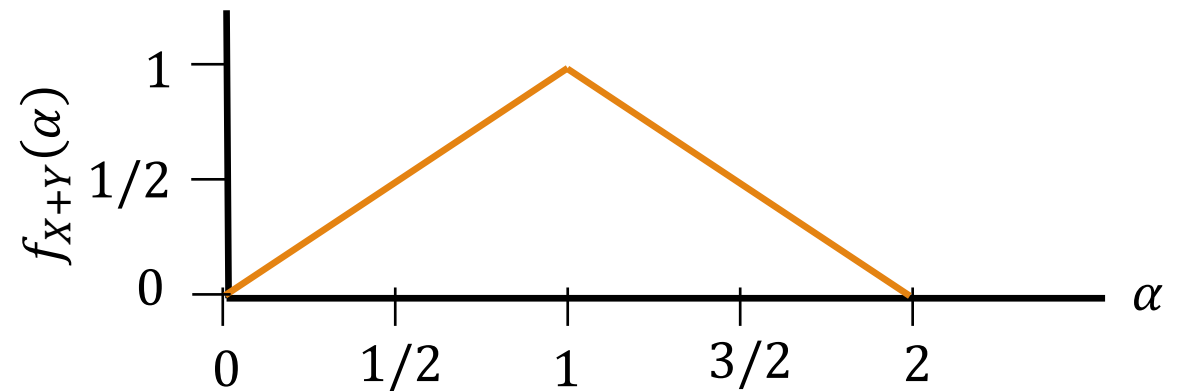
$f_X(k)f_Y(\alpha - k) = 1$ when:

$$0 \leq \alpha \leq 2 \text{ and}$$

$$0 \leq k \leq 1 \text{ and}$$

$$0 \leq \alpha - k \leq 1 \Rightarrow \alpha - 1 \leq k \leq \alpha$$

The precise integration
bounds on k depend on α .



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



whew

Today's plan

Sum of two Uniform independent RVs

→ Expectation of sum of two RVs

Discrete conditional distributions

Continuous conditional distributions

Properties of Expectation, extended to two RVs

1. Linearity:

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$

2. Expectation of a sum = sum of expectation:

$$E[X + Y] = E[X] + E[Y]$$

(we've seen this;
we'll prove this next)

3. Unconscious statistician:

$$E[g(X, Y)] = \sum_x \sum_y g(x, y) p_{X, Y}(x, y)$$

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X, Y}(x, y) dx dy$$

Proof of expectation of a sum of RVs

$$E[X + Y] = E[X] + E[Y]$$

$$E[X + Y] = \sum_x \sum_y (x + y)p_{X,Y}(x, y)$$

LOTUS,
 $g(X, Y) = X + Y$

$$= \sum_x \sum_y xp_{X,Y}(x, y) + \sum_x \sum_y yp_{X,Y}(x, y)$$

$$= \sum_x x \sum_y p_{X,Y}(x, y) + \sum_y y \sum_x p_{X,Y}(x, y)$$

Linearity of summations
(cont. case: linearity of integrals)

$$= \sum_x xp_X(x) + \sum_y yp_Y(y)$$

Marginal PMFs for X and Y

$$= E[X] + E[Y]$$



Even if the **joint distribution is unknown**, you can calculate the **expectation of sum as sum of expectations**.

Example: $E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E[X_i]$ despite dependent trials X_i

$$X \sim \text{Bin}(n, p) \quad E[X] = np$$

$$X = \sum_{i=1}^n X_i \quad \begin{array}{l} \text{Let } X_i = i\text{th trial is heads} \\ X_i \sim \text{Ber}(p), E[X_i] = p \end{array}$$



$$E[X] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n p = np$$

Expectations of common RVs

$$X \sim \text{Bin}(n, p) \quad E[X] = np$$

$$X = \sum_{i=1}^n X_i \quad \begin{array}{l} \text{Let } X_i = i\text{th trial is heads} \\ X_i \sim \text{Ber}(p), E[X_i] = p \end{array} \quad \Rightarrow \quad E[X] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n p = np$$

$$Y \sim \text{NegBin}(r, p) \quad E[Y] = \frac{r}{p}$$

Suppose:

$$Y = \sum_{i=1}^? Y_i$$

How should we define Y_i ?

- A. $Y_i = i\text{th trial is heads. } Y_i \sim \text{Ber}(p), i = 1, \dots, n$
- B. $Y_i = \# \text{ trials to get } i\text{th success (after } (i - 1)\text{th success)}$
 $Y_i \sim \text{Geo}(p), i = 1, \dots, r$
- C. $Y_i = \# \text{ successes in } n \text{ trials}$
 $Y_i \sim \text{Bin}(n, p), i = 1, \dots, r, \text{ we look for } P(Y_i = 1)$



Expectations of common RVs

$$X \sim \text{Bin}(n, p) \quad E[X] = np$$

$$X = \sum_{i=1}^n X_i \quad \begin{array}{l} \text{Let } X_i = i\text{th trial is heads} \\ X_i \sim \text{Ber}(p), E[X_i] = p \end{array} \quad \Rightarrow \quad E[X] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n p = np$$

$$Y \sim \text{NegBin}(r, p) \quad E[Y] = \frac{r}{p}$$

Suppose:

$$Y = \sum_{i=1}^? Y_i$$

How should we define Y_i ?

- A. $Y_i = i\text{th trial is heads. } Y_i \sim \text{Ber}(p), i = 1, \dots, n$
- B.** $Y_i = \# \text{ trials to get } i\text{th success (after } (i - 1)\text{th success)}$
 $Y_i \sim \text{Geo}(p), i = 1, \dots, r$
- C. $Y_i = \# \text{ successes in } n \text{ trials}$
 $Y_i \sim \text{Bin}(n, p), i = 1, \dots, r, \text{ we look for } P(Y_i = 1)$



Expectations of common RVs

$$X \sim \text{Bin}(n, p) \quad E[X] = np$$

$$X = \sum_{i=1}^n X_i \quad \begin{array}{l} \text{Let } X_i = i\text{th trial is heads} \\ X_i \sim \text{Ber}(p), E[X_i] = p \end{array} \quad \Rightarrow \quad E[X] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n p = np$$

$$Y \sim \text{NegBin}(r, p) \quad E[Y] = \frac{r}{p}$$

$$Y = \sum_{i=1}^r Y_i \quad \begin{array}{l} \text{Let } Y_i = \# \text{ trials to get } i\text{th} \\ \text{success (after} \\ (i-1)\text{th success)} \\ Y_i \sim \text{Geo}(p), E[Y_i] = \frac{1}{p} \end{array} \quad \Rightarrow \quad E[Y] = E\left[\sum_{i=1}^r Y_i\right] = \sum_{i=1}^r E[Y_i] = \sum_{i=1}^r \frac{1}{p} = \frac{r}{p}$$



(beginning of non-midterm content)

Break for jokes/ announcements

Announcements

Midterm study tips

Easy to do:

Charts/equations

(essential,
but not hard)

Harder to do:

Glean common strategies

(**top priority:**
reflect and
form links)

from practice exams/section handouts/psets

No matter what:

Start early. **Take breaks.** Stay hydrated. Sleep.

Concept checks

Week 5's: **Wednesday 10/30 1pm**

Includes mid-quarter feedback

Today's plan

Sum of two Uniform independent RVs

Expectation of sum of two RVs

→ Discrete conditional distributions

Continuous conditional distributions

CS109 roadmap

Multiple events:

intersection

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

conditional probability

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

independence

$$P(EF) = P(E)P(F)$$

Joint (**Multivariate**) distributions

joint PMF/PDF

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

conditional distributions?

today

independent RVs

sum of independent RVs

Model ALL the things!



Discrete conditional distributions

Recall the definition of the conditional probability of event E given event F :

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

For discrete random variables X and Y , the **conditional PMF** of X given Y is

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

Different notation,
same idea:

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

Quick check

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

Number or function?

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

number

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

1-D function

3. $P(X = 2|Y = y)$

1-D function

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

2-D function

True or false?

5. $\sum_x P(X = x|Y = 5) = 1$ true

6. $\sum_y P(X = 2|Y = y) = 1$ false

7. $\sum_x \left(\sum_y P(X = x|Y = y)P(Y = y) \right) = 1$

true



Discrete probabilities of CS109

CS109 Acquaintance form
(Fall 2019)

Each student responds with (major X ,
year Y , bool pokemon master M)



$M = 1$ (yes) or $M = 0$ (no)



Discrete probabilities of CS109

Each student responds with (major X , year Y , bool pokemon master M):

Joint PMF of X, Y, M

		CS $X = 1$	SymSys/ MCS/EE $X = 2$	Other Eng/Sci/ /Math $X = 3$	Hum/ SocSci/ Ling $X = 4$	Double major $X = 5$	Undec. $X = 6$
$M = 0$	$Y = 1$.006	.000	.000	.000	.000	.000
	$Y = 2$.155	.069	.034	.006	.023	.029
	$Y = 3$.092	.063	.023	.006	.006	.000
	$Y = 4$.017	.029	.011	.006	.000	.000
	$Y \geq 5$.029	.006	.011	.006	.000	.000
	$M = 1$	$Y = 1$.000	.000	.000	.000	.000
$Y = 2$.126	.040	.017	.017	.000	.017
$Y = 3$.046	.040	.006	.011	.000	.006
$Y = 4$.006	.006	.000	.000	.000	.000
$Y \geq 5$.006	.000	.017	.011	.000	.000

$$P(Y \geq 5, X = 1, M = 1)$$

Joint PMF of X, M

	CS $X = 1$	SymSys/ MCS/EE $X = 2$	Other Eng/Sci/ Math $X = 3$	Hum/ SocSci/ Ling $X = 4$	Double major $X = 5$	Undec. $X = 6$
$M = 0$.299	.167	.080	.023	.029	.029
$M = 1$.184	.086	.040	.040	.000	.023

$$P(X = 1, M = 1) = 0.18$$



Discrete probabilities of CS109

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

The below tables are conditional probability tables for the conditional PMFs

$P(M = m|X = x)$ and $P(X = x|M = m)$.

1. Which table is which?
2. Fill in the missing probability.

Joint PMF of X, M

	CS $X = 1$	Other SymSys/ MCS/EE $X = 2$	Eng/Sci/ Math $X = 3$	Hum/ SocSci/ Ling $X = 4$	Double major $X = 5$	Undec. $X = 6$
$M = 0$.299	.167	.080	.023	.029	.029
$M = 1$.184	.086	.040	.040	.000	.023

$$P(X = 1, M = 1) = 0.18$$

	CS $X = 1$	Other SymSys/ MCS/EE $X = 2$	Eng/Sci/ Math $X = 3$	Hum/ SocSci/ Ling $X = 4$	Double major $X = 5$	Undec. $X = 6$
$M = 0$.619	.659	.667	.364	1.000	.556
$M = 1$.381	.341	.333	.636	.000	.444

	CS $X = 1$	Other SymSys/ MCS/EE $X = 2$	Eng/Sci/ Math $X = 3$	Hum/ SocSci/ Ling $X = 4$	Double major $X = 5$	Undec. $X = 6$
$M = 0$.477	.266	.128	.037	.046	.046
$M = 1$.231	.108	.108	.000	.062



Discrete probabilities of CS109

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

The below tables are conditional probability tables for the conditional PMFs

$P(M = m|X = x)$ and $P(X = x|M = m)$.

1. Which table is which?
2. Fill in the missing probability.

Joint PMF of X, M

	CS $X = 1$	SymSys/ MCS/EE $X = 2$	Other Eng/Sci/ Math $X = 3$	Hum/ SocSci/ Ling $X = 4$	Double major $X = 5$	Undec. $X = 6$
$M = 0$	0.30	0.17	0.08	0.02	0.03	0.03
$M = 1$	0.18	0.09	0.04	0.04	0.00	0.02

$$P(X = 1, M = 1) = 0.18$$

Conditional PMF $P(M = m|X = x)$

	CS $X = 1$	SymSys/ MCS/EE $X = 2$	Other Eng/Sci/ Math $X = 3$	Hum/ SocSci/ Ling $X = 4$	Double major $X = 5$	Undec. $X = 6$
$M = 0$.619	.659	.667	.364	1.000	.556
$M = 1$.381	.341	.333	.636	.000	.444

Conditional PMF $P(X = x|M = m)$

	CS $X = 1$	SymSys/ MCS/EE $X = 2$	Other Eng/Sci/ Math $X = 3$	Hum/ SocSci/ Ling $X = 4$	Double major $X = 5$	Undec. $X = 6$
$M = 0$.477	.266	.128	.037	.046	.046
$M = 1$.492	.231	.108	.108	.000	.062



Be clear about which probabilities should sum to one.



Extended to Amazon



Roll over image to zoom in

FINEDINE
Stainless Steel Mixing Bowls by Finedine (Set of 6) Polished Mirror Finish Nesting Bowl, 3/4 - 1.5-3 - 4-5 - 8 Quart - Cooking Supplies

★★★★★ 2,366 customer reviews | 75 answered questions

Amazon's Choice for "stainless steel mixing bowls"

Price: **\$24.99 & FREE Shipping** on orders over \$25 shipped by Amazon. [Details](#)

Get \$40 off instantly: Pay \$0.00 upon approval for the Amazon.com Store Card.

✓prime | Try Fast, Free Shipping

- With graduating sizes of 3/4, 1.5, 3, 4, 5 and 8 quart, the bowl set allows users to be well equipped for serving fruit salads, marinating for the grill, and adding last ingredients for dessert.
- Stainless steel bowls with commercial grade metal that can be used as both baking mixing bowls and serving bowls. These metal bowls won't stain or absorb odors and resist rust for years of durability.
- An easy to grip rounded-lip on the stainless steel bowl set makes handling easier while a generous wide rim allows contents to flow evenly when pouring; flat base stabilizes the silver bowls making mixing all the easier.
- A space saving stackable design helps de-clutter kitchen cupboards while the attractive polished mirror finish on the large mixing bowls adds a luxurious aesthetic.
- This incredible stainless steel mixing bowl set is refrigerator, freezer, and dishwasher safe for quick and easy meal prep and clean up. They'd also make a great gift!

[Compare with similar items](#)

[Used & new \(7\) from \\$20.62 & FREE shipping on orders over \\$25.00. Details](#)

[Report incorrect product information.](#)

Packaging may reveal contents. Choose **Conceal Package** at checkout.

KELIWA
Easy home baking
[Shop now](#)


 Keliwa 12 Cup Silicone Muffin - Cupcake Baking Pan / Silicone Mold Non - Stick ...
 ★★★★★ 1,803
 \$9.99 ✓prime

Ad feedback

Customers who



ExcelSteel Stainless Steel Colanders, Set of 3
 ★★★★★ 301
 \$15.83 ✓prime



1Easylife 18/8 Stainless Steel Measuring Spoons, Set of 6 for Measuring Dry and Liquid Ingredients
 ★★★★★ 1,854
 #1 Best Seller in Measuring Spoons
 \$9.99 ✓prime



New Star Foodservice 42917 Stainless Steel 4pcs Measuring Cups and Spoons Combo Set
 ★★★★★ 1,042
 #1 Best Seller in Specialty Spoons
 \$9.95 ✓prime



Rubbermaid Easy Find Lids Food Storage Containers, Racer Red, 42-Piece Set 1880801
 ★★★★★ 10,319
 \$19.99 ✓prime



Mlusco 5 Piece Silicone Cooking Utensil Set with Natural Acacia Hard Wood Handle
 ★★★★★ 461
 \$20.99 ✓prime



Bellemain Micro-perforated Stainless Steel 5-quart Colander-Dishwasher Safe
 ★★★★★ 2,797
 #1 Best Seller in Colanders
 \$19.95 ✓prime



AmazonBasics 6-Piece Nonstick Bakeware Set
 ★★★★★ 67
 \$19.99 ✓prime



HOMWE Kitchen Cutting Board (3-Piece Set) | Juice Grooves w/ Easy-Grip Handles | BPA-Free,...
 ★★★★★ 240
 \$14.97 ✓prime

P(bought item X | bought item Y)

Today's plan

Sum of two Uniform independent RVs

Expectation of sum of two RVs

Discrete conditional distributions

→ Ratio of probabilities interlude

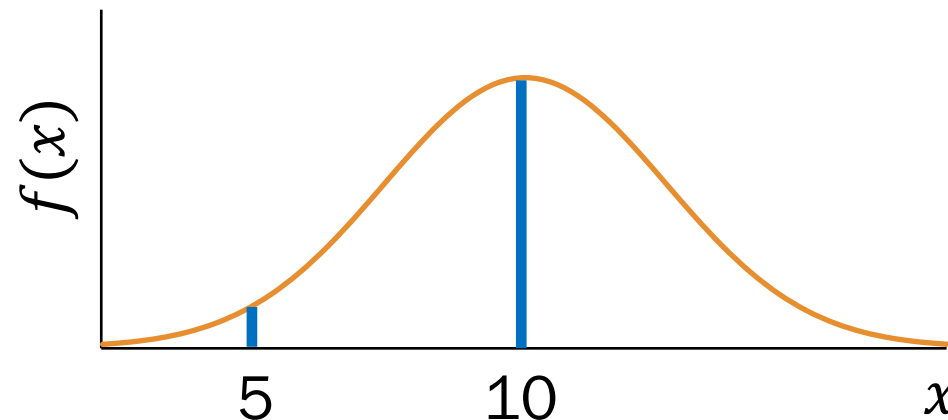
Continuous conditional distributions

Relative probabilities of continuous random variables

Let X = time to finish problem set 2.

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. $0/0 = \text{undefined}$
- B. $\frac{f(10)}{f(5)}$
- C. stay healthy

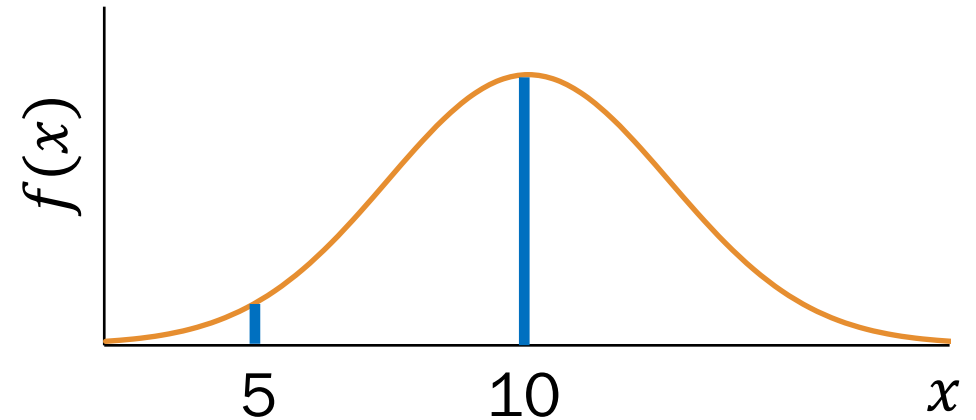


Relative probabilities of continuous random variables

Let X = time to finish problem set 2.

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. $0/0 = \text{undefined}$
- B. $\frac{f(10)}{f(5)}$
- C. stay healthy

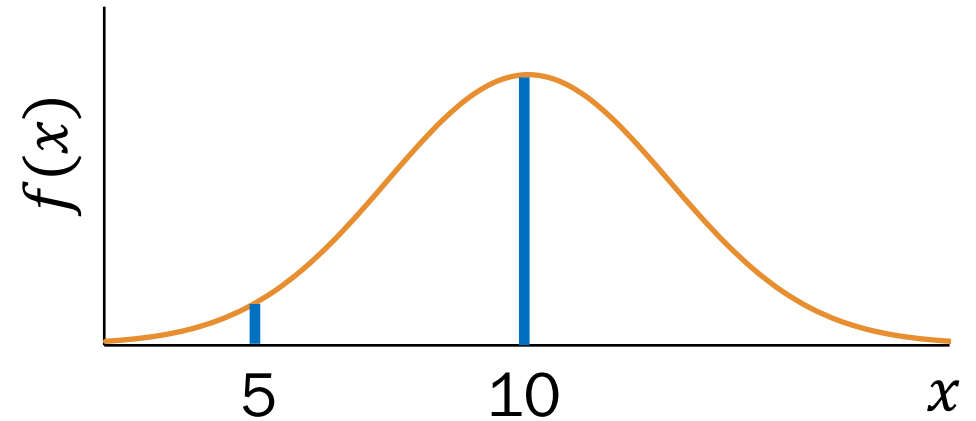


Relative probabilities of continuous random variables

Let X = time to finish problem set 2.

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

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



$$\frac{P(X = 10)}{P(X = 5)} = \frac{f(10)}{f(5)} \rightarrow P(X = a) = P\left(a - \frac{\varepsilon}{2} \leq X \leq a + \frac{\varepsilon}{2}\right) = \int_{a - \frac{\varepsilon}{2}}^{a + \frac{\varepsilon}{2}} f(x) dx \approx \varepsilon f(a)$$

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

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

👉 Ratios of PDFs are meaningful

Today's plan

Sum of two Uniform independent RVs

Expectation of sum of two RVs

Discrete conditional distributions

Ratio of probabilities interlude

→ Continuous conditional distributions

Continuous conditional distributions

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

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

Intuition: $P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \iff f_{X|Y}(x|y)\varepsilon_x = \frac{f_{X,Y}(x, y)\varepsilon_x\varepsilon_y}{f_Y(y)\varepsilon_y}$

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

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

Bayes' Theorem with Continuous RVs

For continuous RVs X and Y ,

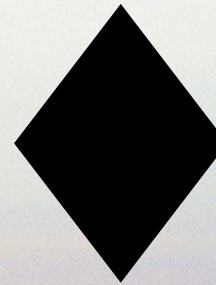
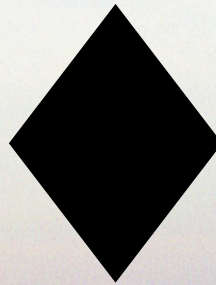
$$f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y)f_Y(y)}{f_X(x)}$$

Intuition:

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



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



This example is hard!
Don't feel discouraged!

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).

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:

$$f_{X,Y|D}(x, y|d) = \frac{\overset{\text{likelihood}}{\text{(of evidence)}} f_{D|X,Y}(d|x, y) \overset{\text{prior}}{\text{belief}} f_{X,Y}(x, y)}{\underset{\text{normalization constant}}{f_D(d)}}$$

Tracking in 2-D space

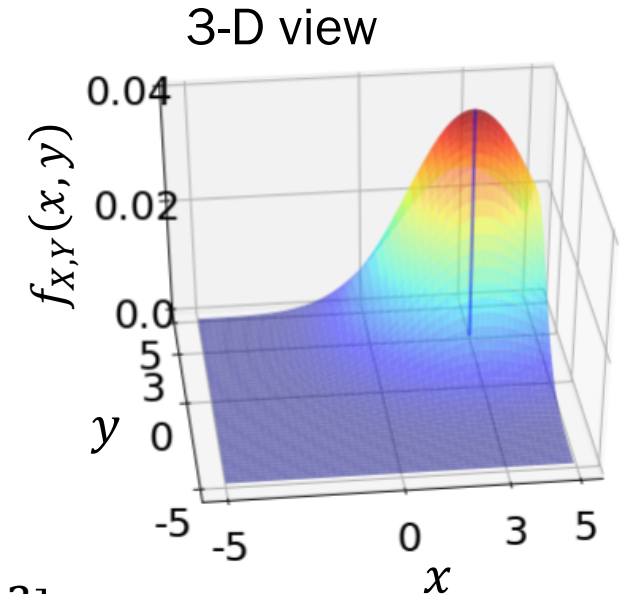
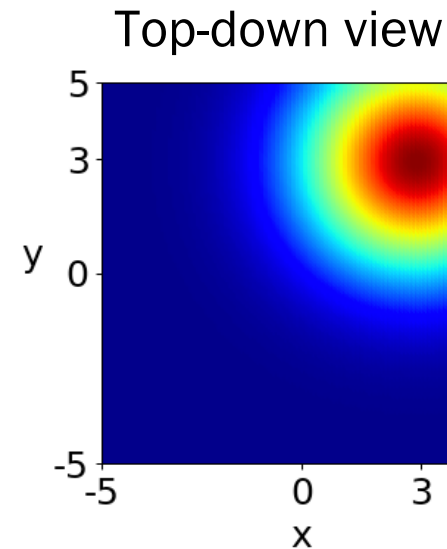
- You have a **prior belief** about the 2-D location of an object, (X, Y) .
- You observe a noisy distance measurement, $D = 4$.
- What is your updated (posterior) belief of the 2-D location of the object after observing the measurement?

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

Suppose the prior distribution is a symmetric bivariate normal distribution:

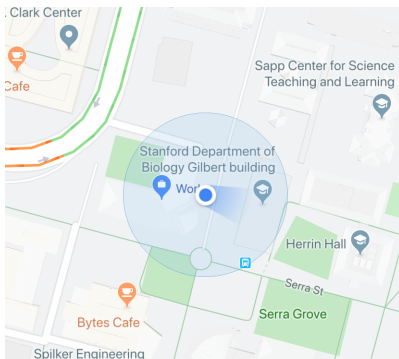
$$f_{X,Y}(x, y) = \frac{1}{2\pi 2^2} e^{-\frac{[(x-3)^2 + (y-3)^2]}{2(2^2)}} = K_1 \cdot e^{-\frac{[(x-3)^2 + (y-3)^2]}{8}}$$

normalizing constant



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?



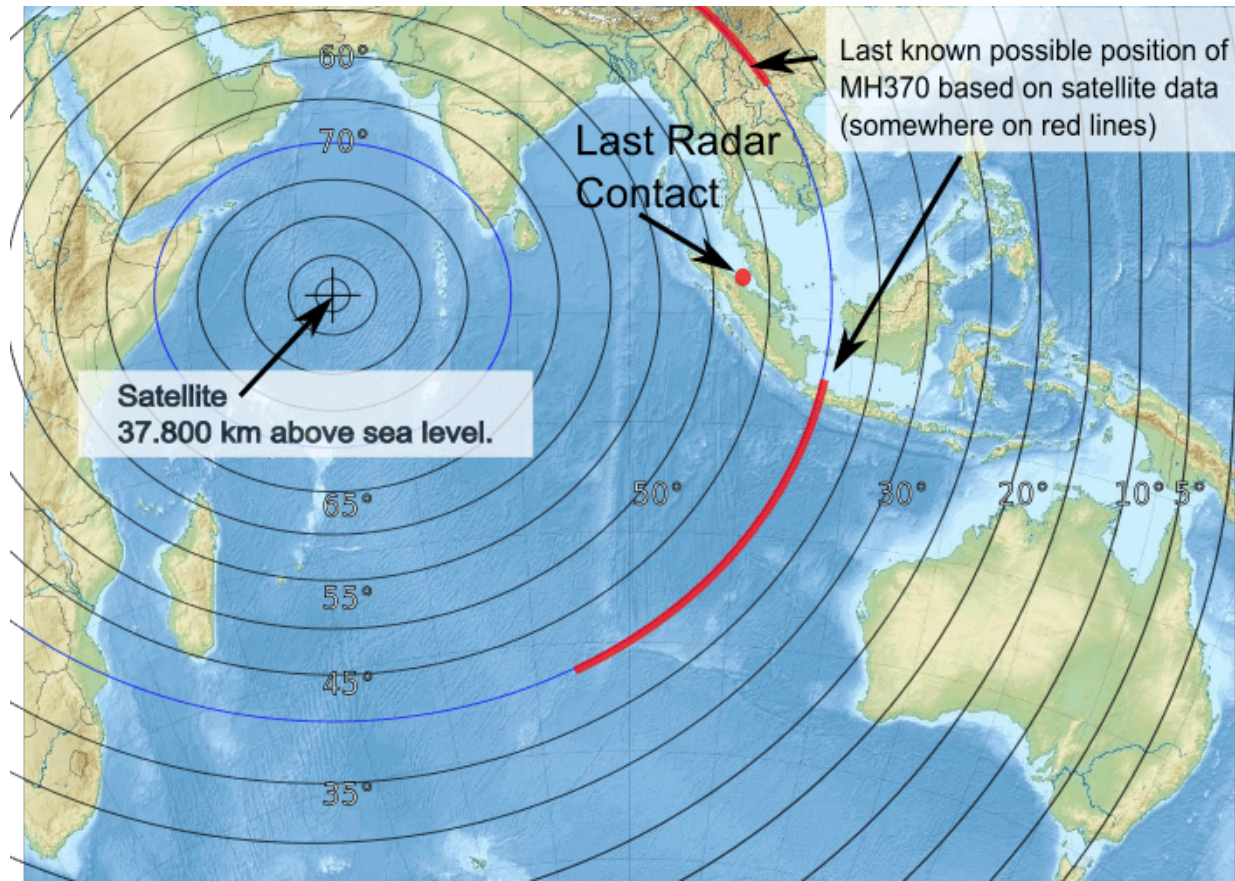
Let D = distance from the satellite (radially).

Suppose you knew your actual position: (x, y) .

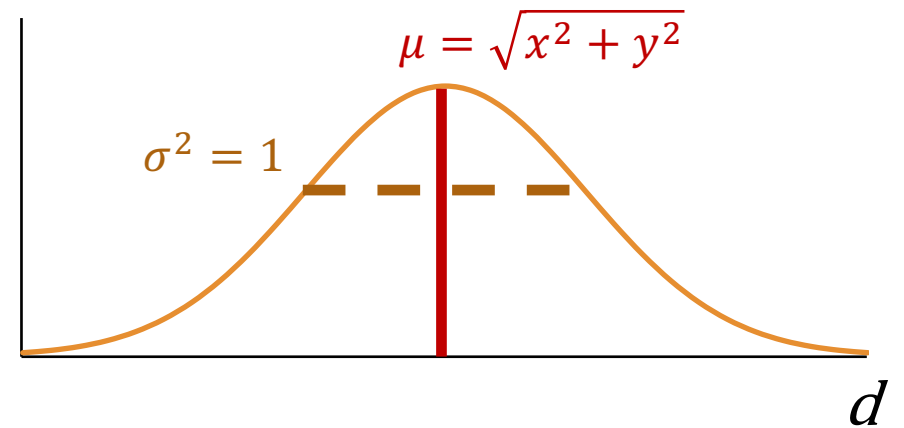
- D is still noisy! Suppose noise is unit variance: $\sigma^2 = 1$
- On average, D is your actual position: $\mu = \sqrt{x^2 + y^2}$

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

Tracking in 2-D space



probability density

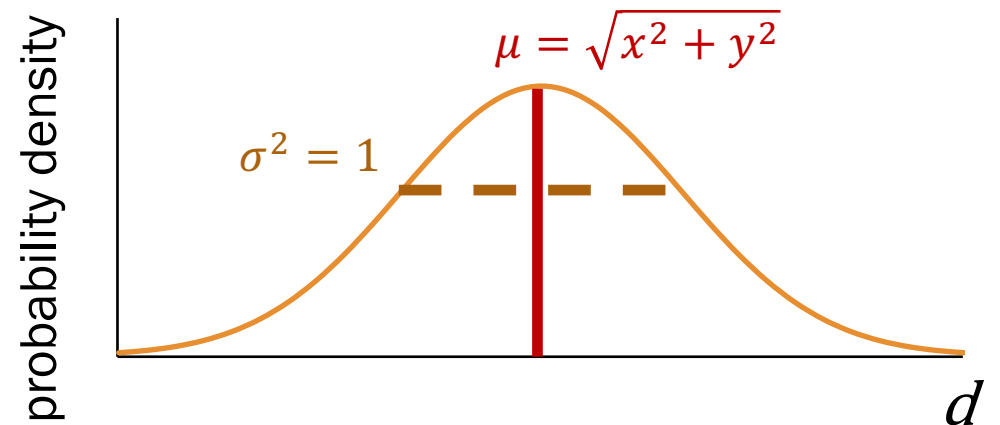


Distance measurement of a ping is normal with respect to the true location.

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?

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



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

Distance measurement of a ping is normal with respect to the true location.

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?

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

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

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

substitute
 μ and σ^2

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{(d-\sqrt{x^2+y^2})^2}{2}} = K_2 \cdot e^{-\frac{(d-\sqrt{x^2+y^2})^2}{2}}$$

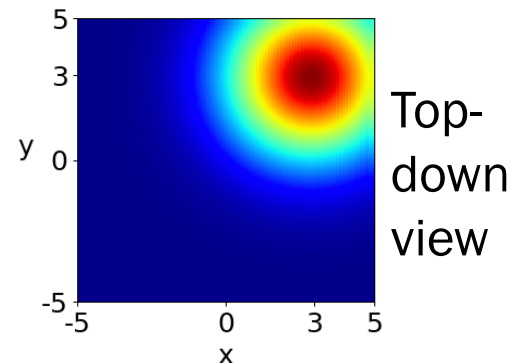
normalizing constant

Deep breath

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?

Prior belief

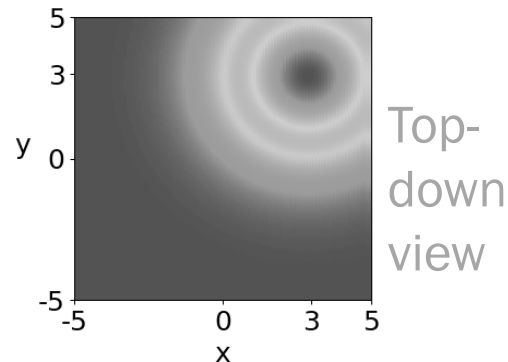


$$f_{X,Y}(x, y) = K_1 \cdot e^{-\frac{[(x-3)^2 + (y-3)^2]}{8}}$$

Tracking in 2-D space

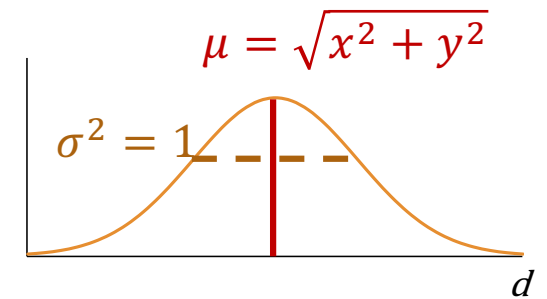
- You have a prior belief about the 2-D location of an object, (X, Y) .
- You observe a **noisy distance measurement**, $D = 4$.
- What is your updated (posterior) belief of the 2-D location of the object after observing the measurement?

Prior belief



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

Observation likelihood

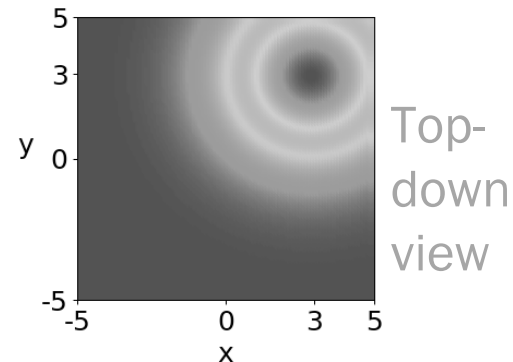


$$f_{D|X,Y}(d|x, y) = K_2 \cdot e^{-\frac{(d - \sqrt{x^2 + y^2})^2}{2}}$$

Tracking in 2-D space

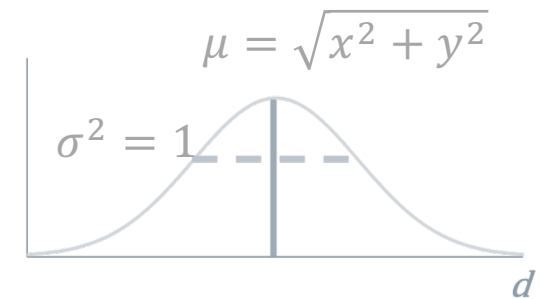
- You have a prior belief about the 2-D location of an object, (X, Y) .
- You observe a noisy distance measurement, $D = 4$.
- What is your **updated (posterior) belief** of the 2-D location of the object after observing the measurement?

Prior belief



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

Observation likelihood



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

Posterior belief

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

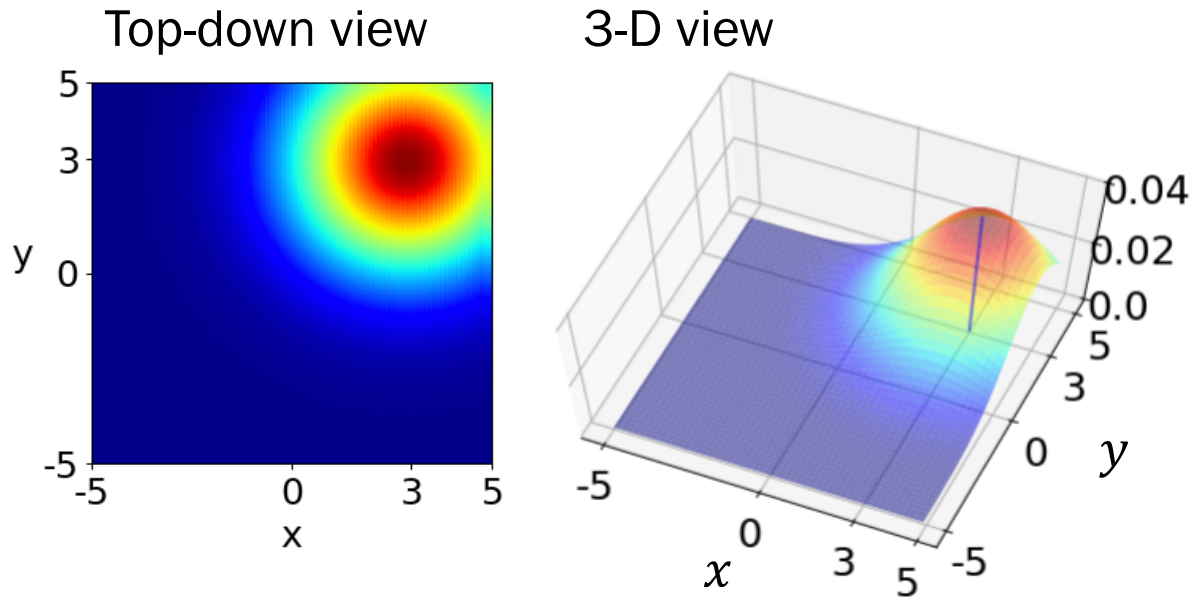
Tracking in 2-D space

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

$$\begin{aligned} f_{X,Y|D}(X = x, Y = y | D = 4) &= \frac{\overset{\text{likelihood of } D = 4}{f_{D|X,Y}(D = 4 | X = x, Y = y)} \overset{\text{prior belief}}{f_{X,Y}(x, y)}}{f(D = 4)} \quad \text{Bayes' Theorem} \\ &= \frac{K_2 \cdot e^{-\frac{(4 - \sqrt{x^2 + y^2})^2}{2}} \cdot K_1 \cdot e^{-\frac{[(x-3)^2 + (y-3)^2]}{8}}}{f(D = 4)} \\ &= \frac{K_3 \cdot e^{-\left[\frac{(4 - \sqrt{x^2 + y^2})^2}{2} + \frac{[(x-3)^2 + (y-3)^2]}{8}\right]}}{f(D = 4)} \\ &= K_4 \cdot e^{-\left[\frac{(4 - \sqrt{x^2 + y^2})^2}{2} + \frac{[(x-3)^2 + (y-3)^2]}{8}\right]} \quad \text{For your notes...} \end{aligned}$$

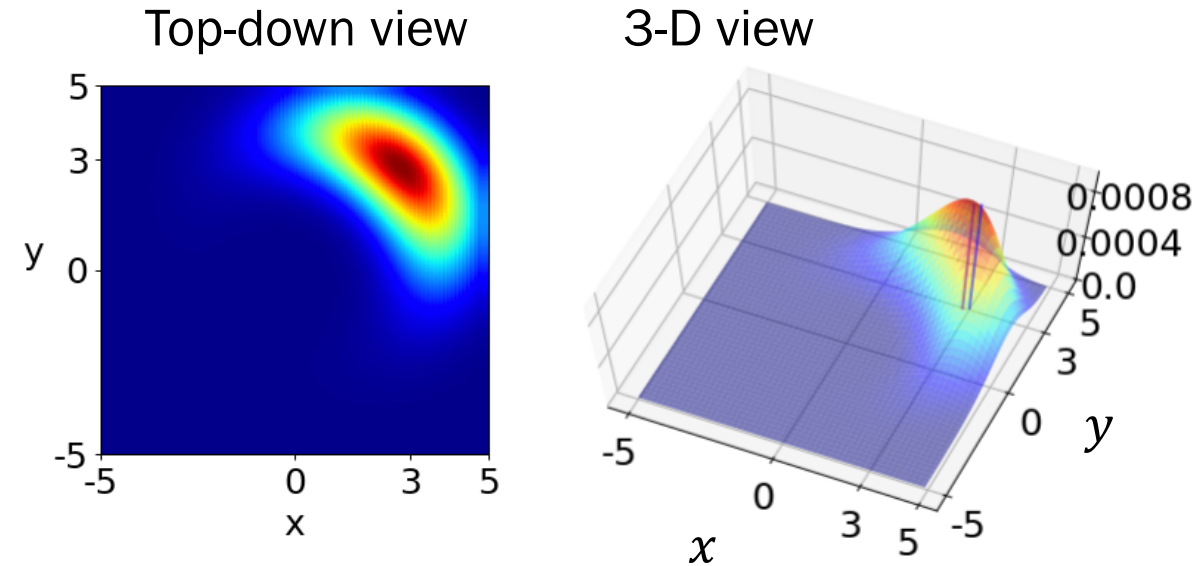
Tracking in 2-D space: Posterior belief

Prior belief



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

Posterior belief



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

Good job today

