

14: Conditional Expectation

Lisa Yan

May 6, 2020

Quick slide reference

3	Conditional distributions	14a_conditional_distributions
11	Web server requests, redux	14b_web_servers
14	Conditional expectation	14c_cond_expectation
20	Law of Total Expectation	14d_law_of_total_expectation
24	Exercises	LIVE

Discrete conditional distributions

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:

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

Discrete probabilities of CS109

Each student responds with:

Year Y

- 1: Frosh/Soph
- 2: Jr/Sr
- 3: Co-term/grad/NDO

Timezone T (12pm California time in my timezone is):

- -1: AM
- 0: noon
- 1: PM

	<u>Joint PMF</u>		
	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$.06	.01	.01
$T = 0$.29	.14	.09
$T = 1$.30	.08	.02

$$P(Y = 3, T = 1)$$

Joint PMFs sum to 1.

Discrete probabilities of CS109

The below are **conditional probability tables** for conditional PMFs

(A) $P(Y = y|T = t)$ and (B) $P(T = t|Y = y)$.

1. Which is which?
2. What's the missing probability?

	Joint PMF		
	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$.06	.01	.01
$T = 0$.29	.14	.09
$T = 1$.30	.08	.02

	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$.09	.04	.08
$T = 0$.45	.61	.75
$T = 1$.46	.35	.17

	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$.75	.125	?
$T = 0$.56	.27	.17
$T = 1$.75	.2	.05



Discrete probabilities of CS109

The below are **conditional probability tables** for conditional PMFs

(A) $P(Y = y|T = t)$ and (B) $P(T = t|Y = y)$.

- Which is which?
- What's the missing probability?

	Joint PMF		
	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$.06	.01	.01
$T = 0$.29	.14	.09
$T = 1$.30	.08	.02

(B) $P(T = t|Y = y)$

	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$.09	.04	.08
$T = 0$.45	.61	.75
$T = 1$.46	.35	.17

$$.30 / (.06 + .29 + .30)$$

(A) $P(Y = y|T = t)$

	$Y = 1$	$Y = 2$	$Y = 3$
$T = -1$.75	.125	.125
$T = 0$.56	.27	.17
$T = 1$.75	.2	.05

1-.75-.125

Conditional PMFs also sum to 1 conditioned on different events!

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

Ad feedback



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

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)

Quick check

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

Number or function?

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

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

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

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

True or false?

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

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

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

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



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. $\sum_x \sum_y P(X = x|Y = y) = 1$ false

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

Web server requests, redux

Let $N = \#$ of requests to a web server per day. Suppose $N \sim \text{Poi}(\lambda)$.

- Each request independently comes from a human (prob. p), or bot ($1 - p$).
- Let X be $\#$ of human requests/day, and Y be $\#$ of bot requests/day.

Are X and Y independent? What are their marginal PMFs?

Our approach:

- Yes, independent Poisson random variables:

$$X \sim \text{Poi}(\lambda p), Y \sim \text{Poi}(\lambda(1 - p))$$

- Two big parts of our derivation:

- $P(X = n, Y = m) = P(X = n | N = n + m)P(N = n + m)$
- $X | N = n + m \sim \text{Bin}(n + m, p)$

A conditional distribution, $X | N!$

Web server requests, redux

(Note: this is a different problem setup from the previous slide)

Consider the number of requests to a web server per day.

- Let $X = \#$ requests from humans/day. $X \sim \text{Poi}(\lambda_1)$
- Let $Y = \#$ requests from bots/day. $Y \sim \text{Poi}(\lambda_2)$
- X and Y are independent. $\rightarrow X + Y \sim \text{Poi}(\lambda_1 + \lambda_2)$

What is $P(X = k | X + Y = n)$?

$$\begin{aligned} P(X = k | X + Y = n) &= \frac{P(X = k, Y = n - k)}{P(X + Y = n)} = \frac{P(X = k)P(Y = n - k)}{P(X + Y = n)} \quad (X, Y \text{ indep.}) \\ &= \frac{e^{-\lambda_1} \lambda_1^k}{k!} \cdot \frac{e^{-\lambda_2} \lambda_2^{n-k}}{(n-k)!} \cdot \frac{n!}{e^{-(\lambda_1 + \lambda_2)} (\lambda_1 + \lambda_2)^n} = \frac{n!}{k! (n-k)!} \cdot \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^n} \\ &= \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-k} \end{aligned}$$

$X | X + Y \sim \text{Bin} \left(X + Y, \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)$

Conditional Expectation

Conditional expectation

Recall the the conditional PMF of X given $Y = y$:

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

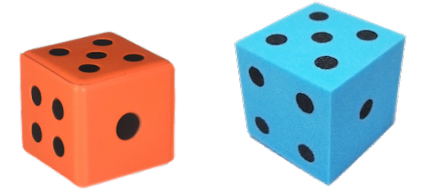
The **conditional expectation** of X given $Y = y$ is

$$E[X|Y = y] = \sum_x xP(X = x|Y = y) = \sum_x xp_{X|Y}(x|y)$$

It's been so long, our dice friends

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

- Roll two 6-sided dice.
- Let roll 1 be D_1 , roll 2 be D_2 .
- Let $S = \text{value of } D_1 + D_2$.



1. What is $E[S|D_2 = 6]$? $E[S|D_2 = 6] = \sum_{x=7}^{12} x P(S = x|D_2 = 6)$

$$= \left(\frac{1}{6}\right) (7 + 8 + 9 + 10 + 11 + 12)$$
$$= \frac{57}{6} = 9.5$$

Intuitively: $6 + E[D_1] = 6 + 3.5 = 9.5$

Let's prove this!

Properties of conditional expectation

1. LOTUS:

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

2. Linearity of conditional expectation:

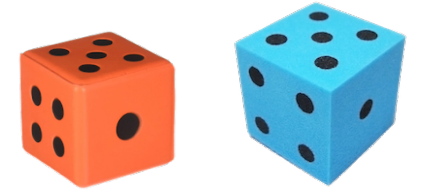
$$E\left[\sum_{i=1}^n X_i | Y = y\right] = \sum_{i=1}^n E[X_i | Y = y]$$

3. Law of total expectation (next time)

It's been so long, our dice friends

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

- Roll two 6-sided dice.
- Let roll 1 be D_1 , roll 2 be D_2 .
- Let $S = \text{value of } D_1 + D_2$.



1. What is $E[S|D_2 = 6]$?

$$\frac{57}{6} = 9.5$$

2. What is $E[S|D_2]$?

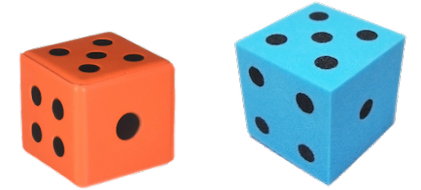
- A. A function of S
 - B. A function of D_2
 - C. A number
3. Give an expression for $E[S|D_2]$.



It's been so long, our dice friends

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

- Roll two 6-sided dice.
- Let roll 1 be D_1 , roll 2 be D_2 .
- Let $S = \text{value of } D_1 + D_2$.



1. What is $E[S|D_2 = 6]$?

$$\frac{57}{6} = 9.5$$

2. What is $E[S|D_2]$?

- A. A function of S
 - B.** A function of D_2
 - C. A number
3. Give an expression for $E[S|D_2]$.

$$E[S|D_2 = d_2] = E[D_1 + d_2|D_2 = d_2]$$

$$= \sum_{d_1} (d_1 + d_2) P(D_1 = d_1 | D_2 = d_2)$$

$$= \sum_{d_1} d_1 P(D_1 = d_1) + d_2 \sum_{d_1} P(D_1 = d_1)$$

($D_1 = d_1, D_2 = d_2$
independent events)

$$= E[D_1] + d_2 = 3.5 + d_2$$

$$E[S|D_2] = 3.5 + D_2$$

Law of Total Expectation

Properties of conditional expectation

1. LOTUS:

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

2. Linearity of conditional expectation:

$$E\left[\sum_{i=1}^n X_i | Y = y\right] = \sum_{i=1}^n E[X_i | Y = y]$$

3. Law of total expectation:

$$E[X] = E[E[X|Y]] \quad \text{what?!}$$

Proof of Law of Total Expectation

$$E[X] = E[E[X|Y]]$$

$$\begin{aligned} E[E[X|Y]] &= E[g(Y)] = \sum_y P(Y = y)E[X|Y = y] && \text{(LOTUS, } g(Y) = E[X|Y]) \\ &= \sum_y P(Y = y) \sum_x xP(X = x|Y = y) && \text{(def of conditional expectation)} \\ &= \sum_y \left(\sum_x xP(X = x|Y = y)P(Y = y) \right) = \sum_y \left(\sum_x xP(X = x, Y = y) \right) && \text{(chain rule)} \\ &= \sum_x \sum_y xP(X = x, Y = y) = \sum_x x \sum_y P(X = x, Y = y) && \text{(switch order of summations)} \\ &= \sum_x xP(X = x) && \text{(marginalization)} \\ &= E[X] \quad \dots\text{what?} \end{aligned}$$

Another way to compute $E[X]$

$$E[X] = E[E[X|Y]]$$

$$E[E[X|Y]] = \sum_y P(Y = y)E[X|Y = y] = E[X]$$

If we only have a conditional PMF of X on some discrete variable Y , we can compute $E[X]$ as follows:

1. Compute expectation of X given some value of $Y = y$
2. Repeat step 1 for all values of Y
3. Compute a weighted sum (where weights are $P(Y = y)$)

```
def recurse():  
    if (random.random() < 0.5):  
        return 3  
    else: return (2 + recurse())
```

Useful for analyzing recursive code!!

(live)

14: Conditional Expectation

Lisa Yan

May 6, 2020

Where are we now? A roadmap of CS109

Today: Statistics of multiple RVs!

$$\text{Var}(X + Y)$$

$$E[X + Y]$$

$$\text{Cov}(X, Y)$$

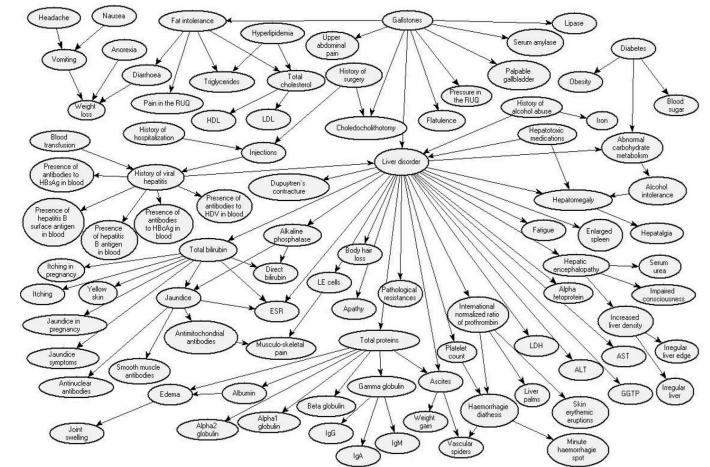
$$\rho(X, Y)$$

Last week: Joint distributions

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



Friday: Modeling with Bayesian Networks



Wednesday:
Conditional distributions

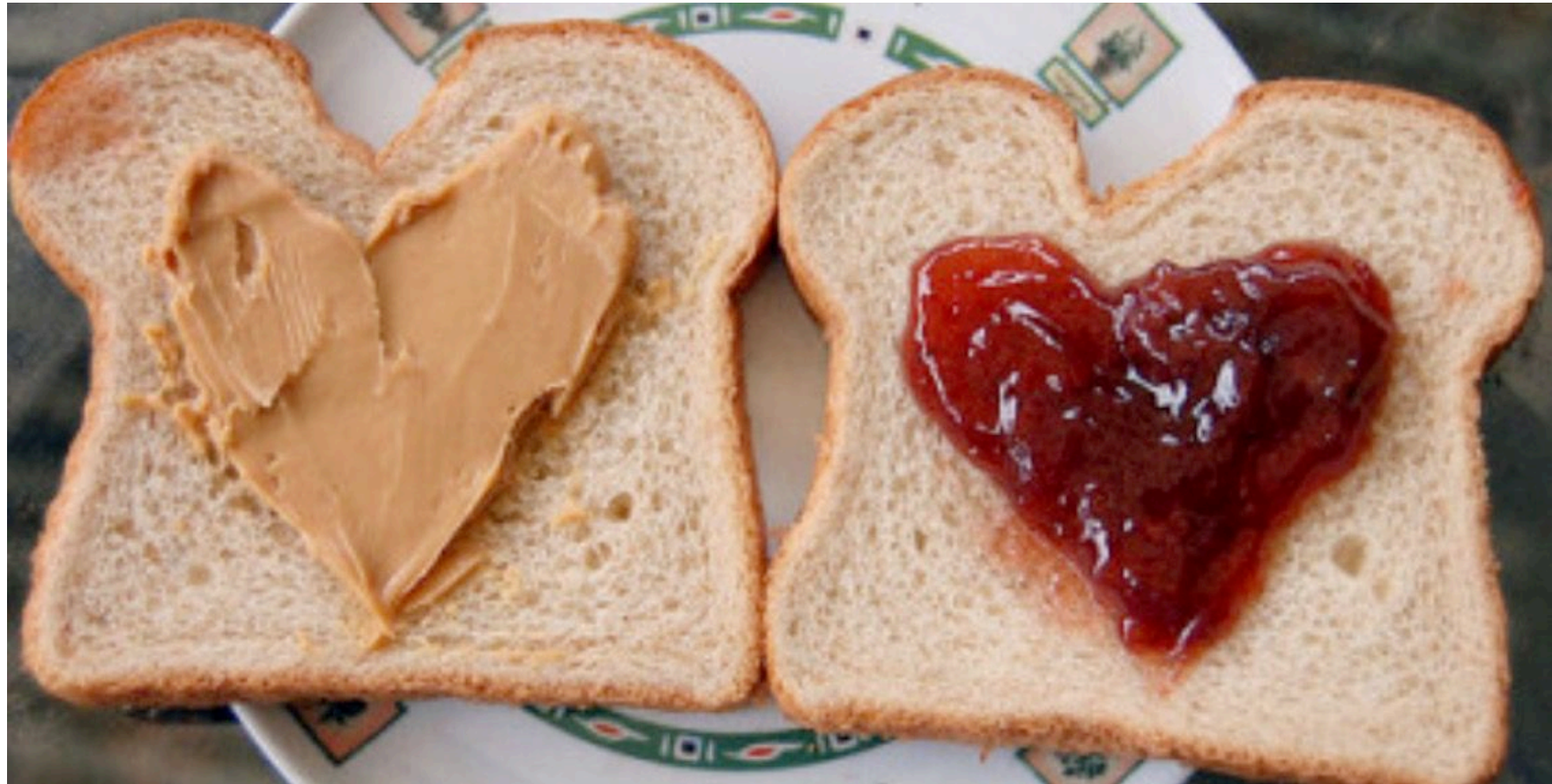
$$p_{X|Y}(x|y)$$

$$E[X|Y]$$

Time to kick it up a notch!



Conditional Expectation



Conditional Distributions

Expectation

Breakout Rooms

Check out the question on the next slide (Slide 28). Post any clarifications here!

<https://us.edstem.org/courses/109/discussion/54694>

Breakout rooms: 4 min. Introduce yourself!



Quick check

1. $E[X]$
2. $E[X, Y]$
3. $E[X + Y]$
4. $E[X|Y]$
5. $E[X|Y = 6]$
6. $E[X = 1]$
- 7.* $E[Y|X = x]$

- A. value
- B. random variable, function of Y
- C. random variable, function of X
- D. function of X and Y
- E. doesn't make sense



Quick check

1. $E[X]$
2. $E[X, Y]$
3. $E[X + Y]$
4. $E[X|Y]$
5. $E[X|Y = 6]$
6. $E[X = 1]$
- 7.* $E[Y|X = x]$

- A. value
- B. random variable, function of Y
- C. random variable, function of X
- D. function of X and Y
- E. doesn't make sense

The conditional expectation of X given $Y = y$ is

$$E[X|Y = y] = \sum_x xP(X = x|Y = y) = \sum_x xp_{X|Y}(x|y)$$

- Interpret: $E[X|Y]$ is a random variable that takes on the value $E[X|Y = y]$ with probability $P(Y = y)$

The **Law of Total Expectation** states that

$$E[E[X|Y]] = \sum_y E[X|Y = y]P(Y = y) = E[X]$$

- Apply: $E[X]$ can be calculated as the expectation of $E[X|Y]$

Analyzing recursive code

$$E[X] = E[E[X|Y]] = \sum_y E[X|Y = y]P(Y = y)$$

```
def recurse():  
    # equally likely values 1,2,3  
    x = np.random.choice([1,2,3])  
    if (x == 1): return 3  
    elif (x == 2): return (5 + recurse())  
    else: return (7 + recurse())
```

Let Y = return value of `recurse()`.
What is $E[Y]$?

Analyzing recursive code

$$E[X] = E[E[X|Y]] = \sum_y E[X|Y = y]P(Y = y)$$

```
def recurse():  
    # equally likely values 1,2,3  
    x = np.random.choice([1,2,3])  
    if (x == 1): return 3  
    elif (x == 2): return (5 + recurse())  
    else: return (7 + recurse())
```

Let Y = return value of `recurse()`.
What is $E[Y]$?

$$E[Y] = E[Y|X = 1]P(X = 1) + E[Y|X = 2]P(X = 2) + E[Y|X = 3]P(X = 3)$$



$$E[Y|X = 1] = 3$$

When $X = 1$, return 3.

Think

Slide 34 has a question to go over by yourself.

Post any clarifications here!

<https://us.edstem.org/courses/109/discussion/54694>

Think by yourself: 2 min



Analyzing recursive code

$$E[X] = E[E[X|Y]] = \sum_y E[X|Y = y]P(Y = y) \quad \text{If } Y \text{ discrete}$$

```
def recurse():  
    # equally likely values 1,2,3  
    x = np.random.choice([1,2,3])  
    if (x == 1): return 3  
    elif (x == 2): return (5 + recurse())  
    else: return (7 + recurse())
```

Let Y = return value of `recurse()`.
What is $E[Y]$?

$$E[Y] = E[Y|X = 1]P(X = 1) + E[Y|X = 2]P(X = 2) + E[Y|X = 3]P(X = 3)$$

$$E[Y|X = 1] = 3$$

What is $E[Y|X = 2]$?

- A. $E[5] + Y$
- B. $E[Y + 5] = 5 + E[Y]$
- C. $5 + E[Y|X = 2]$



(by yourself)

Analyzing recursive code


$$E[X] = E[E[X|Y]] = \sum_y E[X|Y = y]P(Y = y) \quad \text{If } Y \text{ discrete}$$

```
def recurse():  
    # equally likely values 1,2,3  
    x = np.random.choice([1,2,3])  
    if (x == 1): return 3  
    elif (x == 2): return (5 + recurse())  
    else: return (7 + recurse())
```

Let Y = return value of `recurse()`.
What is $E[Y]$?

$$E[Y] = E[Y|X = 1]P(X = 1) + E[Y|X = 2]P(X = 2) + E[Y|X = 3]P(X = 3)$$


 $E[Y|X = 1] = 3$


When $X = 2$, return 5 +
a future return value of `recurse()`.

What is $E[Y|X = 2]$?

- A. $E[5] + Y$
- B. $E[Y + 5] = 5 + E[Y]$
- C. $5 + E[Y|X = 2]$

Analyzing recursive code

$$E[X] = E[E[X|Y]] = \sum_y E[X|Y = y]P(Y = y) \quad \text{If } Y \text{ discrete}$$

```
def recurse():  
    # equally likely values 1,2,3  
    x = np.random.choice([1,2,3])  
    if (x == 1): return 3  
    elif (x == 2): return (5 + recurse())  
    else: return (7 + recurse())
```

Let Y = return value of `recurse()`.
What is $E[Y]$?

$$E[Y] = E[Y|X = 1]P(X = 1) + E[Y|X = 2]P(X = 2) + E[Y|X = 3]P(X = 3)$$

$$E[Y|X = 1] = 3$$

$$E[Y|X = 2] = E[5 + Y]$$

When $X = 3$, return
7 + a future return value
of `recurse()`.

$$E[Y|X = 3] = E[7 + Y]$$

Analyzing recursive code

$$E[X] = E[E[X|Y]] = \sum_y E[X|Y = y]P(Y = y) \quad \text{If } Y \text{ discrete}$$

```
def recurse():  
    # equally likely values 1,2,3  
    x = np.random.choice([1,2,3])  
    if (x == 1): return 3  
    elif (x == 2): return (5 + recurse())  
    else: return (7 + recurse())
```

Let Y = return value of `recurse()`.
What is $E[Y]$?

$$E[Y] = E[Y|X = 1]P(X = 1) + E[Y|X = 2]P(X = 2) + E[Y|X = 3]P(X = 3)$$

$$E[Y|X = 1] = 3 \quad E[Y|X = 2] = E[5 + Y] \quad E[Y|X = 3] = E[7 + Y]$$

$$E[Y] = 3(1/3) + (5 + E[Y])(1/3) + (7 + E[Y])(1/3)$$

$$E[Y] = (1/3)(15 + 2E[Y]) = 5 + (2/3)E[Y]$$

$$E[Y] = 15$$

On your own: What is $\text{Var}(Y)$?

Interlude for jokes/announcements

Announcements

Problem Set 3

Due: Wednesday 5/8 10am

Covers: Up to and including Lecture 11

Interesting probability news

U.S. Recession Model at 100% Confirms Downturn Is Already Here

“Bloomberg Economics created a model last year to determine America’s recession odds.”

- I encourage you to read through and understand the parameters used to define this model!

100%

Chance of Recession Within 12 Months

[CS109 Current Events Spreadsheet](#)

<https://www.bloomberg.com/graphics/us-economic-recession-tracker/>

Independent RVs, defined another way

If X and Y are independent discrete random variables, then $\forall x, y$:

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

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

Note for conditional expectation, independent X and Y implies

$$E[X|Y = y] = \sum_x xp_{X|Y}(x|y) = \sum_x xp_X(x) = E[X]$$

Breakout Rooms

Check out the question on the next slide (Slide 43). Post any clarifications here!

<https://us.edstem.org/courses/109/discussion/54694>

Breakout rooms: 4 min. Introduce yourself!



$$E[X|Y = y] = E[X]$$

Random number of random variables

Say you have a website: BestJokesEver.com. Let:

- $X = \#$ of people per day who visit your site. $X \sim \text{Bin}(100, 0.5)$
- $Y_i = \#$ of minutes spent by visitor i . $Y_i \sim \text{Poi}(8)$
- X and all Y_i are independent.

The time spent by all visitors per day is $W = \sum_{i=1}^X Y_i$. What is $E[W]$?



Random number of random variables

Say you have a website: BestJokesEver.com. Let:

- $X = \#$ of people per day who visit your site. $X \sim \text{Bin}(100, 0.5)$
- $Y_i = \#$ of minutes spent by visitor i . $Y_i \sim \text{Poi}(8)$
- X and all Y_i are independent.

The time spent by all visitors per day is $W = \sum_{i=1}^X Y_i$. What is $E[W]$?

$$E[W] = E \left[\sum_{i=1}^X Y_i \right] = E \left[E \left[\sum_{i=1}^X Y_i \mid X \right] \right]$$

$$= E[XE[Y_i]]$$

$$= E[Y_i]E[X] \quad (\text{scalar } E[Y_i])$$

$$= 8 \cdot 50$$

Suppose $X = x$.

$$E \left[\sum_{i=1}^x Y_i \mid X = x \right] = \sum_{i=1}^x E[Y_i \mid X = x] \quad (\text{linearity})$$

$$= \sum_{i=1}^x E[Y_i] \quad (\text{independence})$$

$$= xE[Y_i]$$

See you next time!

Have a GREAT day!

(no video)

Extra

Hiring software engineers

Your company has only one job opening for a software engineer.

- n candidates interview, in order ($n!$ orderings equally likely)
- Must decide hire/no hire *immediately* after each interview

Strategy: **1.** Interview k (of n) candidates and reject all k
2. Accept the next candidate better than all of first k candidates.

What is your target k that maximizes $P(\text{get best candidate})$?

Fun fact:

- There is an α -to-1 factor difference in productivity b/t the “best” and “average” software engineer.
- Steve jobs said $\alpha=25$, Mark Zuckerberg claims $\alpha=100$

Hiring software engineers

Your company has only one job opening for a software engineer.

- n candidates interview, in order ($n!$ orderings equally likely)
- Must decide hire/no hire *immediately* after each interview

Strategy: 1. Interview k (of n) candidates and reject all k
2. Accept the next candidate better than all of first k candidates.

What is your target k that maximizes $P(\text{get best candidate})$?

Define: X = position of best engineer candidate (1, 2, ..., n)

B = event that you hire the best engineer

Want to maximize for k : $P_k(B)$ = probability of B when using strategy for a given k

$$P_k(B) = \sum_{i=1}^n P_k(B|X=i)P(X=i) = \frac{1}{n} \sum_{i=1}^n P_k(B|X=i) \quad (\text{law of total probability})$$

Hiring software engineers

Your company has only one job opening for a software engineer.

Strategy:

1. Interview k (of n) candidates and reject all k
2. Accept the next candidate better than all of first k candidates.

What is your target k that maximizes $P(\text{get best candidate})$?

Define: X = position of best engineer candidate
 B = event that you hire the best engineer

If $i \leq k$: $P_k(B|X = i) = 0$ (we fired best candidate already)

Else:

We must not hire prior to the i -th candidate.

$$P_k(B|X = i) = \frac{k}{i-1}$$

→ We must have fired the best of the $i-1$ first candidates.

→ The best of the $i-1$ needs to be our comparison point for positions $k+1, \dots, i-1$.

→ The best of the $i-1$ needs to be one of our first k comparison/auto-fire

$$P_k(B) = \frac{1}{n} \sum_{i=1}^n P_k(B|X = i) = \frac{1}{n} \sum_{i=k+1}^n \frac{k}{i-1} \leftarrow \text{Want to maximize over } k$$

Hiring software engineers

Your company has only one job opening for a software engineer.

- Strategy:
1. Interview k (of n) candidates and reject all k
 2. Accept the next candidate better than all of first k candidates.

What is your target k that maximizes $P(\text{get best candidate})$?



Want to maximize over k :

$$P_k(B) = \frac{1}{n} \sum_{i=k+1}^n \frac{k}{i-1} \approx \frac{k}{n} \int_{i=k+1}^n \frac{1}{i-1} di = \frac{k}{n} \ln(i-1) \Big|_{i=k+1}^n = \frac{k}{n} \ln \frac{n-1}{k} \approx \frac{k}{n} \ln \frac{n}{k}$$

Maximize by differentiating w.r.t k , set to 0, solve for k :

$$\frac{d}{dk} \left(\frac{k}{n} \ln \frac{n}{k} \right) = \frac{1}{n} \ln \frac{n}{k} + \frac{k}{n} \cdot \frac{k}{n} \cdot \frac{-n}{k^2} = 0$$

$$\ln \frac{n}{k} = 1$$

$$k = \frac{n}{e}$$

1. Interview $\frac{n}{e}$ candidates
2. Pick best based on strategy
3. $P_k(B) \approx 1/e \approx 0.368$