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:
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

<table>
<thead>
<tr>
<th></th>
<th>$Y = 1$</th>
<th>$Y = 2$</th>
<th>$Y = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T = -1$</td>
<td>.06</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>$T = 0$</td>
<td>.29</td>
<td>.14</td>
<td>.09</td>
</tr>
<tr>
<td>$T = 1$</td>
<td>.30</td>
<td>.08</td>
<td>.02</td>
</tr>
</tbody>
</table>

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

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?

<table>
<thead>
<tr>
<th>$Y = 1$</th>
<th>$Y = 2$</th>
<th>$Y = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T = -1$</td>
<td>.09</td>
<td>.04</td>
</tr>
<tr>
<td>$T = 0$</td>
<td>.45</td>
<td>.61</td>
</tr>
<tr>
<td>$T = 1$</td>
<td>.46</td>
<td>.35</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$Y = 1$</th>
<th>$Y = 2$</th>
<th>$Y = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T = -1$</td>
<td>.75</td>
<td>.125</td>
</tr>
<tr>
<td>$T = 0$</td>
<td>.56</td>
<td>.27</td>
</tr>
<tr>
<td>$T = 1$</td>
<td>.75</td>
<td>.2</td>
</tr>
</tbody>
</table>

0 ≤ $P(Y = y|T = t) ≤ 1$

(y, t) → number
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

<table>
<thead>
<tr>
<th></th>
<th>( Y = 1 )</th>
<th>( Y = 2 )</th>
<th>( Y = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T = -1 )</td>
<td>.06</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>( T = 0 )</td>
<td>.29</td>
<td>.14</td>
<td>.09</td>
</tr>
<tr>
<td>( T = 1 )</td>
<td>.30</td>
<td>.08</td>
<td>.02</td>
</tr>
</tbody>
</table>

### Conditional PMFs

**(A)** \( P(Y = y | T = t) \)

<table>
<thead>
<tr>
<th></th>
<th>( Y = 1 )</th>
<th>( Y = 2 )</th>
<th>( Y = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T = -1 )</td>
<td>.75</td>
<td>.125</td>
<td>.125</td>
</tr>
<tr>
<td>( T = 0 )</td>
<td>.56</td>
<td>.27</td>
<td>.17</td>
</tr>
<tr>
<td>( T = 1 )</td>
<td>.75</td>
<td>.2</td>
<td>.05</td>
</tr>
</tbody>
</table>

**(B)** \( P(T = t | Y = y) \)

<table>
<thead>
<tr>
<th></th>
<th>( Y = 1 )</th>
<th>( Y = 2 )</th>
<th>( Y = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T = -1 )</td>
<td>.09</td>
<td>.04</td>
<td>.08</td>
</tr>
<tr>
<td>( T = 0 )</td>
<td>.45</td>
<td>.61</td>
<td>.75</td>
</tr>
<tr>
<td>( T = 1 )</td>
<td>.46</td>
<td>.35</td>
<td>.17</td>
</tr>
</tbody>
</table>

\[
\frac{0.30}{0.06 + 0.29 + 0.30} = \frac{P(T=1, Y=3)}{P(T=1)}
\]

Conditional PMFs also sum to 1 conditioned on different events!
Extended to Amazon

Stainless Steel Mixing Bowls by Fixedline (Set of 5) Polished Mirror Finish Nesting Bowl, ¾ - 1.5 - 3 - 4.5 - 8 Quart - Cooking Supplies

**Customer Reviews:**

Average: 4.3 stars
Rated: 5 stars (4 reviews)

**Details:**

- BPA-free, non-toxic, dishwasher-safe stainless steel bowls
- Nesting design for easy storage
- Non-slipping bottoms for stability on countertops
- Mirror finish

**Price:** $24.95 & FREE Shipping on orders over $25 shipped by Amazon.

Get a 30-day trial of Prime. Details

Amazon.com Gift Card - $25

**Specifications:**

- 4.5 quart capacity
- Mirror finish
- Polished stainless steel

- **Customer Reviews:**
  - "These bowls are amazing! They're perfect for mixing and storing food. I love the nesting feature for easy storage." - by Customer Name
  - "Stainless steel bowls are a must-have in any kitchen. These are especially great because they're dishwasher-safe and nest together." - by Customer Name

**More Like This:**

- Stainless Steel Measuring Cups and Spoons Set
- Glass Measuring Cups and Spoons Set
- Measuring Spoons and Cups Set

**Related Products:**

- [Easy Mixing Spoons](product-url) - $9.95
- [3-Piece Mixing Bowl Set](product-url) - $24.95

**Customer who bought this item also bought:**

- [Stainless Steel Measuring Cups and Spoons Set](product-url) - $12.99
- [Acrylic Measuring Spoons](product-url) - $9.95

**P(bought item X | bought item Y)"
**Quick check**

**Number or function?**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>$P(X = 2</td>
</tr>
<tr>
<td>2.</td>
<td>$P(X = x</td>
</tr>
<tr>
<td>3.</td>
<td>$P(X = 2</td>
</tr>
<tr>
<td>4.</td>
<td>$P(X = x</td>
</tr>
</tbody>
</table>

**True or false?**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5.</td>
<td>$\sum_x P(X = x</td>
</tr>
<tr>
<td>6.</td>
<td>$\sum_y P(X = 2</td>
</tr>
<tr>
<td>7.</td>
<td>$\sum_x \sum_y P(X = x</td>
</tr>
<tr>
<td>8.</td>
<td>$\sum_x \left( \sum_y P(X = x</td>
</tr>
</tbody>
</table>
Quick check

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. $\sum_x \sum_y P(X = x|Y = y) = 1$  
   left to you to prove

8. $\sum_x \left( \sum_y \frac{P(X = x|Y = y)P(Y = y)}{\sum_x \sum_y P(x, y)} \right) = 1$  
   true
Web server requests, redux
Web server requests (Lecture: Independent RVs)

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

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 \( P(X = k | X + Y = n) \)?

\[
P(X = k | X + Y = n) = \frac{P(X = k, Y = n - k)}{P(X + Y = n)} = \frac{P(X = k)P(Y = n - k)}{P(X + Y = n)}
\]

\[
= \frac{e^{-\lambda_1} \lambda_1^k}{k!} \cdot \frac{e^{-\lambda_2} \lambda_2^{n-k}}{(n-k)!} \cdot \frac{n!}{e^{-(\lambda_1+\lambda_2)}(\lambda_1 + \lambda_2)^n} = \frac{n!}{k! (n-k)!} \cdot \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^n}
\]

\[
= \binom{n}{k} \left( \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^k \left( \frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{n-k} = X | X + Y \sim \text{Bin} \left( X + Y, \frac{\lambda_1}{\lambda_1 + \lambda_2} \right)
\]
Conditional Expectation
Conditional expectation

Recall 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

• 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 $E[S|D_2 = 6]$? 

$$E[S|D_2 = 6] = \sum_{x \leq 7} x P(S = x|D_2 = 6)$$

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

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

Intuitively: \hspace{1cm} 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

• Roll two 6-sided dice.
• Let roll 1 be $D_1$, roll 2 be $D_2$.
• Let $S =$ 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


\[ E[X|Y = y] = \sum_x x p_{X|Y}(x|y) \]
It’s been so long, our dice friends

• 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 $E[S|D_2 = 6]$?

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

   A. A function of $S$
   B. A function of $D_2$
   C. A number


   
   $E[S|D_2 = d_2] = E[D_1 + d_2|D_2 = d_2]$
   
   $= \sum_{d_1} (d_1 + d_2)P(D_1 = d_1|D_2 = d_2)$
   
   $= \sum_{d_1} d_1 P(D_1 = d_1) + d_2 \sum_{d_1} P(D_1 = d_1)$
   
   $= E[D_1] + d_2 = 3.5 + d_2$

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

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

\[ = \sum_y P(Y = y) \sum_x xP(X = x|Y = y) \]

\[ = \sum_y \left( \sum_x xP(X = x|Y = y)P(Y = y) \right) = \sum_y \left( \sum_x xP(X = x, Y = y) \right) \]

\[ = \sum_x \sum_y xP(X = x, Y = y) = \sum_x x \sum_y P(X = x, Y = y) \]

\[ = \sum_x xP(X = x) \]

\[ = E[X] \quad \text{...what?} \]
Another way to compute $E[X]$

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

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 $Y = y$

3. Compute a weighted sum (where weights are $P(Y = y)$)

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

Useful for analyzing recursive code!!
14: Conditional Expectation

Lisa Yan
May 6, 2020
Where are we now? A roadmap of CS109

**Monday**

Today: Statistics of multiple RVs!
- $\text{Var}(X + Y)$
- $E[X + Y]$
- $\text{Cov}(X, Y)$
- $\rho(X, Y)$

Last week: Joint distributions
- $p_{X,Y}(x, y)$

**Wednesday**

Conditional distributions
- $p_{X|Y}(x|y)$
- $E[X|Y]$ 

Time to kick it up a notch!

**Friday**

Modeling with Bayesian Networks
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. $E[X] \quad A.$
2. $E[X, Y] \quad E$
3. $E[X + Y] \quad A,$ expectation of a function of $X$ and $Y$
4. $E[X|Y] \quad B$
5. $E[X|Y = 6] \quad A$
6. $E[X = 1] \quad E$
7. $E[Y|X = x] \quad A,$ for a particular value of $X = x$

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
Conditional Expectation

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

```python
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 $Y = \text{return value of } \text{recurse()}$. What is $E[Y]$?

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

```python
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 $Y = \text{return value of } \text{recurse()}$. What is $E[Y]$?

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

$$E[Y] = \frac{1}{3}E[Y|X = 1]P(X = 1) + \frac{1}{3}E[Y|X = 2]P(X = 2) + \frac{1}{3}E[Y|X = 3]P(X = 3)$$

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

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

```python
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 $Y = \text{return value of recurse()}$. What is $E[Y]$?


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

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

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

(by yourself)
Analyzing recursive code

```python
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 $E[Y]$?


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

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

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

---

If $Y$ discrete

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

```python
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 $E[Y]$?

\[
\]

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

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

If $Y$ discrete

\[
E[X] = E[E[X|Y]] = \sum_y E[X|Y = y]P(Y = y)
\]
Analyzing recursive code

```python
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 $E[Y]$?

\[
\]

\[
\]

\[
E[Y] = \frac{3}{3} \times (1) + (5 + E[Y]) \times \frac{1}{3} + (7 + E[Y]) \times \frac{1}{3}
\]

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

\[
E[Y] = 15
\]

On your own: What is $\text{Var}(Y)$?
Interlude for jokes/announcements

too tired
Announcements

Problem Set 3
Due: Monday 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!

Independent RVs, defined another way

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

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

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

Note for conditional expectation, independent $X$ and $Y$ implies

$$E[X | Y = y] = \sum_x xp_{X|Y}(x | y) = \sum_x xp_X(x) = E[X]$$
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!
Random number of random variables

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

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

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

Indep $X, Y$

$E[X|Y = y] = E[X]$
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 $i./\text{day}$ $Y_i \sim \text{Poi}(8)$
- $X$ and all $Y_i$ are independent.

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

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

Alternate simpler problem: 100 ppl per day

$E[W] = E\left[\sum_{i=1}^{n} Y_i\right] = 100 \cdot E[Y_i]$ (linearity)
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 \( E[W] \)?

\[
E[W] = E \left[ \sum_{i=1}^{X} Y_i \right] = E \left[ \sum_{i=1}^{X} E \left[ Y_i \mid X \right] \right] \\
= E \left[ X E[Y_i] \right] \\
= E[Y_i] E[X] \quad \text{(scalar } E[Y_i]) \\
= 8 \cdot 50
\]
See you next time!

Have a GREAT day!
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 $P$(get best candidate)?

Fun fact:
• There is an $\alpha$-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 $k$

$$P_k(B) = \sum_{i=1}^{n} P_k(B|X = i)P(X = i) = \frac{1}{n} \sum_{i=1}^{n} P_k(B|X = i)$$ (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 $P$(get best candidate)?

Define:

$X = \text{position of best engineer candidate}$

$B = \text{event that you hire the best engineer}$

If $i \leq k : \ P_k(B|X = i) = 0$ \hspace{0.5cm} (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 $k+1, \ldots, i-1$.

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 \( k \):

\[
P_k(B) = \frac{1}{n} \sum_{i=k+1}^{n} \frac{k}{i-1} \approx \frac{k}{n} \int_{i=k+1}^{n} \frac{1}{i-1} \, di = \frac{k}{n} \ln(i-1) \bigg|_{i=k+1}^{n} = \frac{k}{n} \ln \left(1 - \frac{1}{n} \right) \approx \frac{k}{n} \ln \frac{n}{k}
\]

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

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

\[
\ln \frac{n}{k} = 1
\]

\[
k = \frac{n}{e}
\]

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