14: Conditional Expectation

Jerry Cain
February 9, 2024

Lecture Discussion on Ed
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 Year**
- 1: Freshmen and Sophomores
- 2: Juniors and Seniors
- 3: Graduate Students and SCPD

**Mood Mood**
- −1: 😕
- 0: 😐
- 1: 😊

**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>

Joint PMFs sum to 1.

\[
\sum_{t} \sum_{y} p(Y=y, T=t) = 1
\]
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>$T = -1$</th>
<th>$Y = 1$</th>
<th>$Y = 2$</th>
<th>$Y = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y = 1$</td>
<td>.09</td>
<td>.04</td>
<td>.08</td>
</tr>
<tr>
<td>$Y = 2$</td>
<td>.45</td>
<td>.61</td>
<td>.75</td>
</tr>
<tr>
<td>$Y = 3$</td>
<td>.46</td>
<td>.35</td>
<td>.17</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>$T = 1$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y = 1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y = 2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y = 3$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each column sums to 1... what does that mean?
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></th>
<th>(Y = 1)</th>
<th>(Y = 2)</th>
<th>(Y = 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T = -1)</td>
<td>(0.06)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>(T = 0)</td>
<td>(0.29)</td>
<td>(0.14)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>(T = 1)</td>
<td>(0.30)</td>
<td>(0.08)</td>
<td>(0.02)</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>(0.09)</td>
<td>(0.04)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>(T = 0)</td>
<td>(0.45)</td>
<td>(0.61)</td>
<td>(0.75)</td>
</tr>
<tr>
<td>(T = 1)</td>
<td>(0.46)</td>
<td>(0.35)</td>
<td>(0.17)</td>
</tr>
</tbody>
</table>

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

<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>(0.75)</td>
<td>(0.125)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>(T = 0)</td>
<td>(0.56)</td>
<td>(0.27)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>(T = 1)</td>
<td>(0.75)</td>
<td>(0.2)</td>
<td>(0.05)</td>
</tr>
</tbody>
</table>

Conditional PMFs also sum to 1 conditioned on different events!
Quick check

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

Number or function?

1. \( P(X = 2|Y = 5) \)
   number

2. \( P(X = x|Y = 5) = \frac{P_X(x, 5)}{P_Y(5)} \)
   1-D function on \( x \)

3. \( P(X = 2|Y = y) = \frac{P_{XY}(x, y)}{P_Y(y)} \)
   1-D function on \( y \)

4. \( P(X = x|Y = y) = \frac{P_{XY}(x, y)}{P_Y(y)} \)
   2-D function on \( x \) and \( y \)

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

\[ P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \]
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 $S =$ value of $D_1 + D_2$.

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

$$E[S|D_2 = 6] = \sum_{x=7}^{12} xp(S = x|D_2 = 6)$$

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

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

Intuitively: $6 + E[D_1] = 6 + 3.5 = 9.5$
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 (in, like, three slides)

\[ \text{cliffhanger!} \]
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 $E[S|D_2 = 6]$?

   \[
   \frac{57}{6} = 9.5
   \]

2. What is $E[S|D_2]$?
   
   A. A function of $S$
   
   B. A function of $D_2$
   
   C. A number

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 $E[S|D_2 = 6]$?
   \[
   \frac{57}{6} = 9.5
   \]

2. What is $E[S|D_2]$?
   
   **A.** A function of $S$
   
   **B.** A function of $D_2$
   
   **C.** A number


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

\[
= \sum_{d_1} d_1P(D_1 = d_1) + d_2 \sum_{d_1} P(D_1 = d_1) = E[D_1] + d_2 = 3.5 + d_2
\]

\[
E[S|D_2] = 3.5 + D_2
\]
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:

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

3. Law of total expectation:

\[ E[X] = E[E[X \mid Y]] \quad \text{what?} \]

*This inner expectation is a random variable in Y*
Proof of Law of Total Expectation

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

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

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

\[ = E[X] \]
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 $Y$
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.
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 \( E[Y] \)?
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 $E[Y]$?

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


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

When $X = 1$, return 3.
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 $E[Y]$?

\[
\]

What is $E[Y|X = 1]$?

B. $E[5 + Y] = 5 + E[Y]$
C. $5 + E[Y|X = 2]$
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. $E[5 + Y] = 5 + E[Y]$
- C. $5 + E[Y|X = 2]$
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()`.

$$E[Y|X = 3] = E[7 + Y]$$
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 $\text{recurse}()$. What is $E[Y]$?

If $Y$ discrete

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


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

$$E[Y] = 15$$

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

$$E[X|Y = y] = \sum x p_{X|Y}(x|y) = \sum x p_X(x) = E[X]$$
Random number of random variables

Suppose you have a website: zerothworldproblems.com. Let:

- $X$ = # of people per day who visit your site. $X \sim \text{Bin}(100, 0.5)$
- $Y_i$ = # of minutes spent per day 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 $E[W]$?

\[ E[X|Y = y] = E[X] \]
Random number of random variables

Suppose you have a website: zerothworldproblems.com. Let:

- \( X = \# \) of people per day who visit your site. \( X \sim \text{Bin}(100, 0.5) \)
- \( Y_i = \# \) of minutes spent per day 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 \( E[W] \)?

\[
E[W] = E\left[ \sum_{i=1}^{X} Y_i \right] = E\left[ \sum_{i=1}^{X} E[Y_i | X] \right] \\
= E\left[ X E[Y_i] \right] \\
= E[Y_i] E[X] \quad \text{(scalar } E[Y_i]) \\
= 8 \cdot 50
\]
Breaking News
Before You Go! CS109 Challenge

Do something cool and creative with probability!

Grand Prize:
- All exams replaced with 100%

All Serious Entries:
- Extra credit (between %0.5 and 2% added to overall average)

Optional Proposal:  Sun. 02/25, 11:59pm
Due:                 Sat. 03/09, 11:59pm