MTH 224, Spring 2024

Instructor: Bahman Angoshtari

Lecture 13

Section 4.3: conditional distribution of discrete random variable, conditional expectation, law of total expectation (tower property)

13.1. Discrete Conditional Distributions

- Recall conditional probability:

$$\mathbb{P}(E \mid F) = \frac{P(E \cap F)}{P(F)}.$$

Definition 13.1. For discrete random variable X and Y, we define the following conditional pmf:

$$p_{X|Y}(x|y) = p_{X|Y=y}(x) := \mathbb{P}(X = x \mid Y = y) = \frac{P(X = x \cap Y = y)}{P(Y = y)} = \frac{p_{X,Y}(x,y)}{p_y(y)}.$$

The random variable with the above pmf is denoted by $X|_{Y=y}$.

Example 13.2. Roll two fair dice. Let X be the largest number rolled, and Y the smallest number. Find pmf of $X|_{Y=2}$ and $X|_{Y=4}$.

Solution. The joint pmf is the following (we have previously calculated this joint pmf):

	X=1	X=2	X=3	X=4	X=5	X=6
Y=1	1/36	1/18	1/18	1/18	1/18	1/18
Y=2	0	1/36	1/18	1/18	1/18	1/18
Y=3	0	0	1/36	1/18	1/18	1/18
Y=4	0	0	0	1/36	1/18	1/18
Y=5	0	0	0	0	1/36	1/18
Y=6	0	0	0	0	0	1/36

or $p_{X,Y}(x,y) = \begin{cases} 0, & x < y \\ \frac{1}{36}, & x = y \\ \frac{1}{18}, & x > y \end{cases}.$

The marginal pmf values for Y are:

$$p_Y(2) = 4 \cdot \frac{1}{18} + \frac{1}{36} = \frac{1}{4}, \quad \text{and} \quad p_Y(4) = 2 \cdot \frac{1}{18} + \frac{1}{36} = \frac{5}{36}.$$

Therefore, the desired conditional pmfs are given by:

$X\|_{Y=y}$	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$	Total
$Y = 2$	0	$\frac{1/36}{1/4} = \frac{1}{9}$	$\frac{1/18}{1/4} = \frac{2}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	1
$Y = 4$	0	0	0	$\frac{1/36}{5/36} = \frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$	1

13.2. Discrete Conditional Expectation

DEFINITION 13.3. Let X, Y be random variables and $p_{X|Y}(x|y)$ be the conditional distribution of X given $Y = y$. We then define the conditional expectation of X given the event $Y = y$ as follows:

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

Note that $\mathbb{E}[X \mid Y = y]$ is a function of y, so, we can define a function $g(y)$ as follows:

$$g(y) := \mathbb{E}[X \mid Y = y].$$

We now define the conditional expectation of X given Y as the following **random variable**

$$\mathbb{E}[X|Y] := g(Y).$$

- In particular, if the discrete random variable Y has the range $I_Y = \{y_1, y_2, \ldots, y_N\}$, then

$$\mathbb{E}[X|Y] = \begin{cases} \mathbb{E}[X|Y = y_1]; & \text{if } Y = y_1, \\ \mathbb{E}[X|Y = y_2]; & \text{if } Y = y_2, \\ \vdots & \\ \mathbb{E}[X|Y = y_N]; & \text{if } Y = y_N. \end{cases}$$

REMARK 13.4. It is very important that you pay attention to the difference between $\mathbb{E}[X \mid Y = y]$ and $\mathbb{E}[X \mid Y]$. In particular, $\mathbb{E}[X \mid Y = y]$ is a number (more accurately, a function $g(y)$ of the value y). While, $\mathbb{E}[X \mid Y]$ is a random variable (more accurately, it is $g(Y)$).

EXAMPLE 13.5. Let X be the # of bedrooms and Y be the number of laptops in an apartment. The joint and marginal pmf of X and Y are as follows:

	$Y = 0$	$Y = 1$	$Y = 2$	p_X
$X = 0$	0.05	0.12	0.03	0.2
$X = 1$	0.07	0.1	0.08	0.25
$X = 2$	0.02	0.26	0.27	0.55
p_Y	0.14	0.48	0.38	1

Find: a) $\mathbb{E}[X|Y = 1]$,b) $\mathbb{E}[X|Y]$.

SOLUTION. a) First, we calculate the pmf $p_{X|Y}(x|1) = \frac{p_{X,Y}(x,1)}{p_Y(1)}$ for $x = 0, 1, 2$:

$$p_{X|Y}(0|1) = \frac{0.12}{0.48} = 0.25, \quad p_{X|Y}(1|1) = \frac{0.1}{0.48} \approx 0.21, \quad p_{X|Y}(2|1) = \frac{0.26}{0.48} \approx 0.54.$$

Then, we calculate:

$$\mathbb{E}[X|Y = 1] = 0 \times p_{X|Y}(0|1) + 1 \times p_{X|Y}(1|1) + 2 \times p_{X|Y}(2|1) = 0.21 + 2 \times 0.54 = 1.29.$$

b) We have that

$$\mathbb{E}[X|Y] = \begin{cases} \mathbb{E}[X|Y = 0]; & \text{if } Y = 0, \\ \mathbb{E}[X|Y = 1]; & \text{if } Y = 1, \\ \mathbb{E}[X|Y = 2]; & \text{if } Y = 2. \end{cases}$$

We have already calculated $\mathbb{E}[X|Y=1] = 1.29$. We need to calculate the other two conditional expectations:

$$\begin{aligned}\mathbb{E}[X|Y=0] &= 0 \times p_{X|Y}(0|0) + 1 \times p_{X|Y}(1|0) + 2 \times p_{X|Y}(2|0) \\ &= \frac{p_{XY}(1,0)}{p_Y(0)} + 2 \times \frac{p_{XY}(2,0)}{p_Y(0)} = \frac{0.07}{0.14} + 2 \times \frac{0.02}{0.14} = \frac{11}{14} \approx 0.786.\end{aligned}$$

$$\begin{aligned}\mathbb{E}[X|Y=2] &= 0 \times p_{X|Y}(0|2) + 1 \times p_{X|Y}(1|2) + 2 \times p_{X|Y}(2|2) \\ &= \frac{0.08}{0.38} + 2 \times \frac{0.27}{0.38} = \frac{31}{19} \approx 1.63.\end{aligned}$$

Finally,

$$\mathbb{E}[X|Y] = \begin{cases} 0.786; & \text{if } Y = 0, \\ 1.29; & \text{if } Y = 1, \\ 1.63; & \text{if } Y = 2. \end{cases}$$

13.3. Law of total expectation (tower property)

- The following rule, known as the **law of total expectation (LTE)** or the **tower property** of conditional expectation, is a fundamental property of the conditional expectation.

THEOREM 13.6. *We have* $\mathbb{E}\left[\mathbb{E}\left[X|Y\right]\right] = \mathbb{E}\left[X\right]$.

PROOF. Using the definition of conditional expectation:

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[X|Y\right]\right] &= \sum_y p_Y(y)\,\mathbb{E}\left[X|Y=y\right] = \sum_y p_Y(y)\left(\sum_x x p_{X|Y}(x|y)\right) \\ &= \sum_y p_Y(y)\left(\sum_x x\frac{p_{XY}(x,y)}{p_Y(y)}\right) = \sum_y\sum_x x p_Y(y)\frac{p_{XY}(x,y)}{p_Y(y)} \\ &= \sum_y\sum_x x p_{XY}(x,y) = \mathbb{E}[X].\end{aligned}$$

□

EXAMPLE 13.7. Let X be the # of bedrooms and Y be the number of laptops in an apartment. The joint and marginal pmf of X and Y are as follows:

	$Y=0$	$Y=1$	$Y=2$	p_X
$X=0$	0.05	0.12	0.03	0.2
$X=1$	0.07	0.1	0.08	0.25
$X=2$	0.02	0.26	0.27	0.55
p_Y	0.14	0.48	0.38	1

Calculate $\mathbb{E}[X]$ and $\mathbb{E}[\mathbb{E}[X|Y]]$.

SOLUTION. By LTE, $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X] = 1 \times 0.25 + 2 \times 0.55 = 1.35$. Let us also calculate $\mathbb{E}[\mathbb{E}[X|Y]]$ using the definition. In the last example of Lecture 17, we calculated:

$$\mathbb{E}[X|Y] = \begin{cases} 0.786; & \text{if } Y = 0, \\ 1.29; & \text{if } Y = 1, \\ 1.63; & \text{if } Y = 2. \end{cases}$$

Therefore,

$$\begin{aligned}\mathbb{E}[\mathbb{E}[X|Y]] &= \mathbb{E}[X|Y=0]p_Y(0) + \mathbb{E}[X|Y=1]p_Y(1) + \mathbb{E}[X|Y=2]p_Y(2) \\ &\approx 0.786 \times 0.14 + 1.29 \times 0.48 + 1.63 \times 0.38 \approx 1.35.\end{aligned}$$

So, we double checked that the tower property holds!

Example 13.8. Let $X \sim \mathrm{G}(p)$. Show that $\mathbb{E}[X] = \frac{1}{p}$.

Solution. Let $A =$ the first trial is success. Then, by the law of total expectation (LTE), we have

$$\mathbb{E}[X] = \mathbb{E}\big[\mathbb{E}[X|\mathbb{1}_A]\big] = \mathbb{E}[X|\mathbb{1}_A = 1]\mathbb{P}(\mathbb{1}_A = 1) + \mathbb{E}[X|\mathbb{1}_A = 0]\mathbb{P}(\mathbb{1}_A = 0).$$

Note that $\mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(A) = p$ and $\mathbb{P}(\mathbb{1}_A = 0) = \mathbb{P}(A^c) = 1 - p$. If A happens, then $X = 1$. So, $\mathbb{E}[X|\mathbb{1}_A = 1] = 1$. If A^c happens, then $X = 1 + Y$, with $Y \sim G(p)$. So, $\mathbb{E}[X|\mathbb{1}_A = 0] = \mathbb{E}[1 + Y] = 1 + \mathbb{E}[Y] = 1 + \mathbb{E}[X]$. Finally, we obtain that

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}[X|\mathbb{1}_A = 1]\mathbb{P}(A) + \mathbb{E}[X|\mathbb{1}_A = 0]\mathbb{P}(A^c). \\ &= 1 \times p + (1 + \mathbb{E}[X])(1 - p) \\ &= p + 1 - p + (1 - p)\mathbb{E}[X].\end{aligned}$$

We thus have that $\mathbb{E}[X] = (1 - p)\,\mathbb{E}(X) + 1$, which yields $\mathbb{E}[X] = \frac{1}{p}$.