CIS 2033, Spring 2017

Instructor: David Dobor

April 22, 2017

PMFs and conditional PMFs have come up on multiple occasions throughout this course. Recall that any PMF has an associated expectation. And so conditional PMFs also have associated expectations, which we call conditional expectations. We have already seen many examples of conditional expectations for the case where we condition on an event, A. The case where we condition on random variables is exactly the same, as we see in the following examples.

The Law of Total Expectation

As a refresher, let's start with an example that should feel familiar. [1]

[1] The example is taken from Wikipedia and is slightly modified to fit our notation.

EXAMPLE 1. Two factories supply light bulbs to the market. Factory A's bulbs work for an average of 5000 hours, whereas factory B's bulbs work for an average of 4000 hours. It is known that factory A supplies 60% of the total bulbs available. What is the expected length of time that a purchased bulb will work for?

LET X denote the random variable that measures the lifetime of a bulb. We are interested in $\mathbf{E}[X]$, the expected lifetime of the bulb. Let A denote the event that a bulb came from factory A and let B stand for the event that it came from factory B. We are given that:

- $\mathbf{P}(A) = 6/10$
- $\mathbf{P}(B) = 4/10$
- $\mathbf{E}[X \mid A] = 5000$
- $\mathbf{E}[X \mid B] = 4000$

The total expectaion theorem says that:

$$\begin{aligned}\mathbf{E}[X] &= \mathbf{E}[X \mid A] \cdot \mathbf{P}(A) + \mathbf{E}[X \mid B] \cdot \mathbf{P}(B) \\ &= 5000 \cdot \frac{6}{10} + 4000 \cdot \frac{4}{10} \\ &= 4600.\end{aligned}$$

Note that here we condition X on events A and B. The case where we condition on a random variable is exactly the same. We let the event A be the event that some random variable Y takes on a specific value. We then calculate the expectation using the relevant conditional probabilities, those that are given by the conditional PMF.

EXAMPLE 2. A miner is trapped in a mine containing 3 doors. The first door leads to a tunnel that will take him to safety after 3 hours of travel. The second door leads to a tunnel that will take him to safety after 5 hours of travel. The third door leads to a tunnel that will take him to safety after 7 hours of travel. If we assume that the miner is equally likely to choose any one of the doors, what is the expected length of time until he reaches safety?

LET Y denote the door that the miner chooses. Imagine that the three doors are numbered: $1, 2, 3$. Note that Y is a random variable, and the event that the miner chooses the first door is $Y = 1$; the event that he chooses the second door is $Y = 2$, and so on. Let X denote the amount of time (in hours) until the miner reaches safety,

This problem should feel no different from the one we just saw in Example 1. The law of total probability, rewritten in the "random variable notation", tells us that

$$\begin{aligned}\mathbf{E}[X] &= \mathbf{E}[X \mid Y = 1] \cdot p_Y(1) + \mathbf{E}[X \mid Y = 2] \cdot p_Y(2) + \mathbf{E}[X \mid Y = 3] \cdot p_Y(3) \\ &= 3 \cdot \frac{1}{3} + 5 \cdot \frac{1}{3} + 7 \cdot \frac{1}{3} \\ &= 5.\end{aligned}$$

So, on the average, it takes the miner 5 hours to reach safety.

IN SHORT, informally, the story with the total expectation theorem is as follows. The expectation of any random variable X can be found by taking the conditional expectations of X under each one the several different scenarios, $A_1, A_2, \ldots, A_n$, and weighing them according to the probabilities of the different scenarios:

$$\mathbf{E}[X] = \mathbf{P}(A_1)\ \mathbf{E}[X \mid A_1] + \ldots + \mathbf{P}(A_n)\ \mathbf{E}[X \mid A_n]$$

And if we let the event that Y takes on a specific value be one of the different scenarios, the total expectation theorem can be written as follows:

$$\mathbf{E}[X] = \sum_y p_Y(y)\ \mathbf{E}[X \mid Y = y] \qquad (1)$$

Here again each term in the sum is the probability of a given scenario times the expected value of X under that particular scenario.

The Law of Total Expectation for Continuous Random Variables

There is an obvious analog of the total expectation theorem in the case where X and Y are continuous. The only difference is that, in the expression (1) above, we replace the PMF, $p_Y(y)$, with the PDF, $f_Y(y)$, and we integrate instead of summing over all the possible values of y:

$$\mathbf{E}[X] = \int_{-\infty}^{\infty} f_Y(y)\ \mathbf{E}[X \mid Y = y]\ dy$$

The interpretation is that we consider all possibilities for Y. Under each possibility of Y we find the expected value of X. And then we weigh those different possibilities according to the corresponding values of the density. So we're taking a weighted average of these conditional expectations to obtain the overall expectation of the random variable X.

EXAMPLE 3. Suppose Y is uniform over $[1, 2]$. Suppose further that given $Y = y$, X is an exponential random variable with parameter $\lambda = y$, so we can write

$$X \mid Y = y \quad \sim \quad \text{Exponential}(y), \quad \text{or that} \quad f_{X|Y}(x \mid y) = ye^{-yx}$$
$$\text{whenever } x \geq 0 \text{ and } 0 \text{ otherwise.}$$

Now recall that the expected value of the exponentially distributed random variable with parameter λ is $1/\lambda$ and so we have that

$$\mathbf{E}[X \mid Y = y] = \frac{1}{y}.$$

Applying the total expectation theorem, we have

$$\begin{aligned}
\mathbf{E}[X] &= \int_{-\infty}^{\infty} \mathbf{E}[X \mid Y = y]\ f_Y(y)\ dy \\
&= \int_{1}^{2} \mathbf{E}[X \mid Y = y] \cdot 1\ dy \\
&= \int_{1}^{2} \frac{1}{y}\ dy \\
&= \ln 2.
\end{aligned}$$