Lecture 9

Andrei Antonenko

Feb 11, 2005

1 Conditioning on Event

Assume we are given a probability model, and we are told that a certain event A occurred. Therefore, we can use the conditional probabilities instead of original. One can talk about conditional PMFs which provide the probabilities of the possible values of a random variable, conditioned on the occurrence of some event. This idea is developed in this section.

The **conditional PMF** of a random variable X, conditioned on a particular event A with $P(A) > 0$, is defined by

$$p_{X|A}(x) = P(X = x|A) = \frac{P(\{X = x\} \cap A)}{P(A)}. \tag{1}$$

Since events $\{X = x\} \cap A$ are disjoint for different values of A, we have

$$P(A) = \sum_x P(\{X = x\} \cap A),$$

and therefore

$$\sum_x p_{X|A}(x) = 1.$$

Example 1.1. Consider a roll of a fair 6-sided die. Assume we know that the number obtained was even: $A = \{2, 4, 6\}$. If X is a random variable, which is equal to a number tossed, we can write the conditional PMF of X:

$$\begin{aligned} p_{X|A}(x) &= P(X = x| \text{ roll is even}) \\ &= \frac{P(X = x \text{ and roll is even})}{P(\text{roll is even})} \\ &= \begin{cases} 1/3, & \text{if } x = 2, 4, 6 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

2 Conditioning on another random variable

Assume we are conducting an experiment, and there are several (two) random variables associated with this experiment, X and Y. We might be given some information about the value of one of them, say we know that $Y = y$. This information might change our beliefs about X, and affect its distribution. Formally, we will define a **conditional PMF of X given Y**, as the conditional probability that $X = x$ given that $Y = y$:

$$p_{X|Y}(x|y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)}. \tag{2}$$

For any fixed value of $Y = y$, we can consider $p_{X|Y}(x|y)$ as a function of x. This function is a valid PMF for X. It assigns nonnegative values to all possible x, and all of them add up to 1.

The conditional PMF is often convenient to use to calculate joint PMF:

$$p_{X,Y}(x, y) = p_{X|Y}(x|y)p_Y(y). \tag{3}$$

Example 2.1. Consider 4 rolls of a fair 6-sided die. Let X be the number of 1's, and Y be the number of 2's. We will find the joint probability mass function of X and Y.

Since Y is a number of 2's, and the probability of getting 2 is 1/6, Y has a binomial distribution with parameters $n = 4$ and $p = 1/6$:

$$p_Y(y) = \binom{4}{y}\left(\frac{1}{6}\right)^{y}\left(\frac{5}{6}\right)^{4-y}.$$

Now let's fix the value $Y = y$ - number of 2's obtained in 4 rolls. For 1's we have $4 - y$ rolls left, and the probability of getting 1 is 1/5, since now we have only 5 possibilities: 1,3,4,5, and 6. Therefore, given the value $Y = y$, X has a binomial distribution with parameters $n = 4 - y$ and $p = 1/5$:

$$p_{X|Y}(x|y) = \binom{4-y}{x}\left(\frac{1}{5}\right)^{x}\left(\frac{4}{5}\right)^{4-y-x}.$$

Multiplying the two expressions above, we get the joint distribution of X and Y:

$$p_{X,Y}(x, y) = p_Y(y)p_{X|Y}(x|y) = \binom{4}{y}\left(\frac{1}{6}\right)^{y}\left(\frac{5}{6}\right)^{4-y}\binom{4-y}{x}\left(\frac{1}{5}\right)^{x}\left(\frac{4}{5}\right)^{4-y-x}.$$

The marginal distribution of X was defined as

$$p_X(x) = \sum_y p_{X,Y}(x, y).$$

If we substitute the expression for $p_{X,Y}(x, y)$, we get:

$$p_X(x) = \sum_y p_{X|Y}(x|y)p_Y(y), \tag{4}$$

which is the equivalent of total probability theorem for random variables.

Example 2.2. Assume X is a travel time of the message, and Y is a length of message. Y can be equal to 100 with probability 5/6 and to 10,000 with probability 1/6 (short messages and long messages), therefore

$$p_Y(y) = \begin{cases} 5/6, & \text{if } y = 100 \\ 1/6, & \text{if } y = 10,000 \end{cases}$$

Now assume that the transmission time of the message depends on Y and equal to $(Y/10,000)$ with probability $1/2$, $(Y/1,000)$ with probability $1/3$, and $(Y/100)$ with probability $1/6$ (network may be busy, and in this case the transmission time is longer). Therefore, assuming the message is short ($Y = 100$), we have:

$$p_{X|Y}(x|100) = \begin{cases} 1/2, & \text{if } x = 100/10,000 = 1/100 \text{ sec} \\ 1/3, & \text{if } x = 100/1,000 = 1/10 \text{ sec} \\ 1/6, & \text{if } x = 100/100 = 1 \text{ sec} \end{cases}$$

Now, assuming the message is long ($Y = 10,000$), we get:

$$p_{X|Y}(x|10,000) = \begin{cases} 1/2, & \text{if } x = 10,000/10,000 = 1 \text{ sec} \\ 1/3, & \text{if } x = 10,000/1,000 = 10 \text{ sec} \\ 1/6, & \text{if } x = 10,000/100 = 100 \text{ sec} \end{cases}$$

Now we can compute the (unconditional) distribution of the transmission time of the message:

$$\begin{aligned}
p_X\left(\frac{1}{100}\right) &= p_{X|Y}\left(\frac{1}{100}|Y = 100\right) p_Y(100) + p_{X|Y}\left(\frac{1}{100}|Y = 10,000\right) p_Y(10,000) \\
&= \frac{1}{2}\cdot\frac{5}{6} + 0\cdot\frac{1}{6} = \frac{1}{2}\cdot\frac{5}{6} \\
p_X\left(\frac{1}{10}\right) &= p_{X|Y}\left(\frac{1}{10}|Y = 100\right) p_Y(100) + p_{X|Y}\left(\frac{1}{10}|Y = 10,000\right) p_Y(10,000) \\
&= \frac{1}{3}\cdot\frac{5}{6} + 0\cdot\frac{1}{6} = \frac{1}{3}\cdot\frac{5}{6} \\
p_X(1) &= p_{X|Y}(1|Y = 100)p_Y(100) + p_{X|Y}(1|Y = 10,000)p_Y(10,000) \\
&= \frac{1}{6}\cdot\frac{5}{6} + \frac{1}{2}\cdot\frac{1}{6} \\
p_X(10) &= p_{X|Y}(10|Y = 100)p_Y(100) + p_{X|Y}(10|Y = 10,000)p_Y(10,000) \\
&= 0\cdot\frac{5}{6} + \frac{1}{3}\cdot\frac{1}{6} = \frac{1}{3}\cdot\frac{1}{6} \\
p_X(100) &= p_{X|Y}(100|Y = 100)p_Y(100) + p_{X|Y}(100|Y = 10,000)p_Y(10,000) \\
&= 0\cdot\frac{5}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{6}\cdot\frac{1}{6}
\end{aligned}$$

3 Conditional expectation

Assuming we have information that the event A happened. In this case, the expectation of X can be substituted with the **conditional expectation**, conditioned on the occurrence of the event A:

$$\mathbf{E}[X|A] = \sum_x x p_{X|A}(x|A). \tag{5}$$

For example, in case of a roll of a 6-sided die, we have:

$$\mathbf{E}[X] = \frac{1}{6}\cdot 1 + \frac{1}{6}\cdot 2 + \frac{1}{6}\cdot 3 + \frac{1}{6}\cdot 4 + \frac{1}{6}\cdot 5 + \frac{1}{6}\cdot 6 = 3.5.$$

Assume we know that the number rolled was even (event A). In this case, using PMF from the example 1.1, we have conditional expectation:

$$\mathbf{E}[X|A] = \frac{1}{3} \cdot 2 + \frac{1}{3} \cdot 4 + \frac{1}{3} \cdot 6 = 4.$$

If instead of the event A we are given, that the value of the other random variable Y is y: $Y = y$, we can define conditional expectation as

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

We will see some properties of the conditional expectation in the next lecture.