Lecture 3

Conditional Probabilities

Example 1. There are 5 women and 10 men in a room. Three of the women and 9 of the men are employed. You select a person at random from the room, all people being equally likely to be chosen. Clearly, Ω is the collection of all 15 people, and

$$\mathrm{P}\{\text{male}\} = \frac{2}{3}, \quad \mathrm{P}\{\text{female}\} = \frac{1}{3}, \quad \mathrm{P}\{\text{employed}\} = \frac{4}{5}.$$

Also,

$$\mathrm{P}\{\text{male and employed}\} = \frac{8}{15}, \quad \mathrm{P}\{\text{female and employed}\} = \frac{4}{15}.$$

Someone has looked at the result of the sample, and tells us that the person sampled is employed. Let $\mathrm{P}(\text{female} \mid \text{employed})$ denote the conditional probability of "female" given this piece of information. Then,

$$\mathrm{P}(\text{female} \mid \text{employed}) = \frac{|\text{female among employed}|}{|\text{employed}|} = \frac{3}{12} = \frac{1}{4}.$$

Definition 1. If A and B are events and $\mathrm{P}(B) > 0$, then the *conditional probability of A given B* is

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}.$$

For the previous example, this amounts to writing

$$\mathrm{P}(\text{Female} \mid \text{employed}) = \frac{|\text{female and employed}|/|\Omega|}{|\text{employed}|/|\Omega|} = \frac{1}{4}.$$

Example 2 (Multiplication rule). A useful formulation of the conditional probability formula: $\mathrm{P}(A \cap B) = \mathrm{P}(B)\mathrm{P}(A \mid B)$. And an induction argument can be used that shows that for all events $A_1, \ldots, A_n$,

$$\mathrm{P}(A_1 \cap \cdots \cap A_n) = \mathrm{P}(A_1)\mathrm{P}(A_2 \mid A_1)\mathrm{P}(A_3 \mid A_1 \cap A_2) \cdots \mathrm{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1}).$$

Example 3. If we deal two cards fairly from a standard deck, the probability of $K_1 \cap K_2$ [$K_j = \{\text{King on the } j \text{ draw}\}$] is

$$\mathrm{P}(K_1 \cap K_2) = \mathrm{P}(K_1)\mathrm{P}(K_2 \mid K_1) = \frac{4}{52} \times \frac{3}{51}.$$

This agrees with direct counting: $|K_1 \cap K_2| = 4 \times 3$, whereas $|\Omega| = 52 \times 51$.

Similarly,

$$\mathrm{P}(K_1 \cap K_2 \cap K_3) = \frac{4}{52} \times \frac{3}{51} \times \frac{2}{50}.$$

Or for that matter,

$$\mathrm{P}(K_1 \cap K_2 \cap K_3 \cap K_4) = \frac{4}{52} \times \frac{3}{51} \times \frac{2}{50} \times \frac{1}{49}.$$

Example 4 (Probability of a flush, p. 61 of Pitman). A five-card hand is dealt at random from a deck of 52 cards. [More specifically, every time a card is dealt, all remaining cards in the deck are equally likely.] What is P(flush), where a flush occurs when all cards have the same suit? Because probability of a disjoint union is the sum of the probabilities,

$$\mathrm{P}(\text{flush}) = \mathrm{P}(\text{spade flush}) + \mathrm{P}(\text{heart flush}) + \mathrm{P}(\text{diamond flush}) + \mathrm{P}(\text{club flush}).$$

All four flush probabilities on the right-hand side are the same [think about relabeling the cards]. Therefore,

$$\begin{aligned}
\mathrm{P}(\text{flush}) &= 4\mathrm{P}(\text{diamond flush}) \\
&= 4 \times \left(\frac{13}{52} \times \frac{12}{51} \times \frac{11}{50} \times \frac{10}{49} \times \frac{9}{48}\right) \\
&= \frac{12}{51} \times \frac{11}{50} \times \frac{10}{49} \times \frac{9}{48} \\
&\approx 0.001980792316927.
\end{aligned}$$

Theorem 1 (Law of total probability). *For all events A and B,*

$$\mathrm{P}(A) = \mathrm{P}(A \cap B) + \mathrm{P}(A \cap B^c).$$

If, in addition, $0 < \mathrm{P}(B) < 1$, then

$$\mathrm{P}(A) = \mathrm{P}(A \mid B)\mathrm{P}(B) + \mathrm{P}(A \mid B^c)\mathrm{P}(B^c).$$

Proof. For the first statement, note that $A = (A \cap B) \cup (A \cap B^c)$ is a disjoint union. For the second, write $\mathrm{P}(A \cap B) = \mathrm{P}(A \mid B)\mathrm{P}(B)$ and $\mathrm{P}(A \cap B^c) = \mathrm{P}(A \mid B^c)\mathrm{P}(B^c)$. □

Example 5 (The Monte Hall problem)**.** Three doors: behind one is a nice prize; behind the other two lie goats. You choose a door at random. The host (Monte Hall) opens another door, and gives you the option of changing your choice to the remaining unopened door. Should you take his offer?

The answer is "yes." Indeed, if W denotes the event that you win, then under the "not switch" model, we have

$$\mathrm{P}(W) = \frac{1}{3}. \tag{4}$$

Under the "switch model,"

$$\mathrm{P}(W) = \mathrm{P}(W \mid R)\mathrm{P}(R) + \mathrm{P}(W \mid R^c)\mathrm{P}(R^c),$$

where R denotes the event that you had it right in your first guess. Now $\mathrm{P}(R) = 1/3$, but because you are going to switch, $\mathrm{P}(W \mid R) = 0$ and $\mathrm{P}(W \mid R^c) = 1$. Therefore,

$$\mathrm{P}(W) = \frac{2}{3}.$$

Compare this with (4) to see that you should always switch. What are my assumptions on Ω? This is an important issue, as can be seen from reading the nice discussion (p. 75) of the text.

Example 6. There are three types of people: poor (π), middle-income (μ), and rich (ρ). 40% of all π, 45% of μ, and 60% of ρ are over 25 years old (Θ). Find $\mathrm{P}(\Theta)$. The result of Theorem 1 gets replaced with

$$\begin{aligned}
\mathrm{P}(\Theta) &= \mathrm{P}(\Theta \cap \pi) + \mathrm{P}(\Theta \cap \mu) + \mathrm{P}(\Theta \cap \rho) \\
&= \mathrm{P}(\Theta \mid \pi)\mathrm{P}(\pi) + \mathrm{P}(\Theta \mid \mu)\mathrm{P}(\mu) + \mathrm{P}(\Theta \mid \rho)\mathrm{P}(\rho) \\
&= 0.4\mathrm{P}(\pi) + 0.45\mathrm{P}(\mu) + 0.6\mathrm{P}(\rho).
\end{aligned}$$

If we knew $\mathrm{P}(\pi)$ and $\mathrm{P}(\mu)$, then we could solve. For example, suppose $\mathrm{P}(\pi) = 0.1$ and $\mathrm{P}(\mu) = 0.3$. Then $\mathrm{P}(\rho) = 0.6$ (why?), and

$$\mathrm{P}(\Theta) = (0.4 \times 0.1) + (0.45 \times 0.3) + (0.6 \times 0.6) = 0.535.$$

Bayes's Theorem

The following question arises from time to time: Suppose A and B are two events of positive probability. If we know $\mathrm{P}(B \mid A)$ then what is $\mathrm{P}(A \mid B)$? We can proceed as follows:

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{\mathrm{P}(B \mid A)\mathrm{P}(A)}{\mathrm{P}(B)}.$$

If we know only the conditional probabilities, then we can write $\mathrm{P}(B)$, in turn, using Theorem 1, and obtain

Theorem 2 (Bayes's Formula)**.** *If A, A^c and B are events of positive probability, then*

$$\mathrm{P}(A \mid B) = \frac{\mathrm{P}(B \mid A)\mathrm{P}(A)}{\mathrm{P}(B \mid A)\mathrm{P}(A) + \mathrm{P}(B \mid A^c)\mathrm{P}(A^c)}.$$

Example 7. There are two coins on a table. The first tosses heads with probability 1/2, whereas the second tosses heads with probability 1/3. You select one at random and toss it. What are the chances that you toss heads?

Question: What is Ω?

Question: Someone tells you that the end result of this game was heads. What are the odds that it was the first coin that was chosen?

Let C denote the event that you selected the first coin. Let H denote the event that you tossed heads. We know: $\mathrm{P}(C) = 1/2$, $\mathrm{P}(H \mid C) = 1/2$, and $\mathrm{P}(H \mid C^c) = 1/3$. By Bayes's formula,

$$\begin{aligned} \mathrm{P}(C \mid H) &= \frac{\mathrm{P}(H \mid C)\mathrm{P}(C)}{\mathrm{P}(H \mid C)\mathrm{P}(C) + \mathrm{P}(H \mid C^c)\mathrm{P}(C^c)} \\ &= \frac{\frac{1}{2} \times \frac{1}{2}}{\left(\frac{1}{2} \times \frac{1}{2}\right) + \left(\frac{1}{3} \times \frac{1}{2}\right)} \\ &= \frac{3}{5}. \end{aligned}$$

Bayes' rule has a natural generalization. Suppose $A_1, \ldots, A_n$ and B are events such that all of the A_j's are disjoint [$A_i \cap A_j = \varnothing$ if $i \neq j$] and at least one of the A_j's necessarily occurs [$A_1 \cup \cdots \cup A_n = \Omega$. For instance, $A_1 = A$ and $A_2 = A^c$ with $n = 2$.]. Then,

$$\mathrm{P}(A_j \mid B) = \frac{\mathrm{P}(B \mid A_j)\mathrm{P}(A_j)}{\mathrm{P}(B \mid A_1)\mathrm{P}(A_1) + \cdots + \mathrm{P}(B_n \mid A_n)\mathrm{P}(A_n)}.$$

When $n = 2$ and $A_1 = A$ and $A_2 = A^c$, then this is the former Bayes's theorem. The general formula follows from similar arguments as the $n = 2$ case.