Conditional probability (Cont)

Math 477

September 27, 2014

1 $P(\cdot|F)$ is a probability

Conditioning on F means we treat F as our new sample space. This has a rigorous mathematical interpretation as followed:

$$\begin{aligned} 0 \leq P(E|F) &\leq 1 \\ P(S|F) &= 1 \end{aligned}$$

If $E_1, E_2, \cdots, E_n$ are mutually exclusive then

$$P(E_1 \cup E_2 \cup \cdots \cup E_n|F) = \sum_{i=1}^{n} P(E_i|F).$$

The 2nd equality can be replaced with $P(F|F) = 1$ to emphasize the fact that F now is our universe. A different way to express this idea is if we denote $Q(E) = P(.|F)$ then Q is a probability measure on the same sample space.

1.1 Conditional probability of a conditional probability

Having Q as a probabiltiy measure, we can define $Q(E|G)$. The question of course is how is $Q(E|G)$ related to the original probability P? The answer is as followed:

$$Q(E|G) = P(E|FG),$$

which you can check by definition:

$$
\begin{aligned}
Q(E|G) &= \frac{Q(EG)}{Q(G)} = \frac{P(EG|F)}{P(G|F)} \\
&= \frac{P(EGF)/P(F)}{P(GF)/P(F)} = \frac{P(EGF)}{P(GF)} = P(E|FG).
\end{aligned}
$$

The interpretation of the above equation is that observing F happened, our new sample space becomes F. Now suppose addtionally we observe G also happened. Then the grand total effect is as if we observe both F and G happened, which of course makes sense.

1.2 Application

The observation $P(\cdot|F)$ is a probability can lead us to extension of rules we have so far. For example, we have the following law of total probability, *conditioned on F*:

$$P(E_1|F) = P(E_1|E_2F)P(E_2|F) + P(E_1|E_2^cF)P(E_2^c|F).$$

The following example demonstrates its use.

Example 1.1. *Suppose the probability that during any given year, an accident-prone person will have an accident within next year is .4, while the probability that a non-accident-prone person will have an accident next year is .2. We also know that 30 % of the population is accident-prone. If a policy holder has an accident in the first year, what is the probability that he will have an accident in his second year?*

Ans: Note that this question asks you to consider probability conditioned on two different events: whether a person is accident-prone and his having an accident in the first year. So let E_1 be the event that the person has an accident in the first year, E_2 the event he has an accident in the 2nd year, F his being an accident-prone person. Then

$$P(E_2|E_1) = P(E_2|E_1F)P(F|E_1) + P(E_2|E_1F^c)P(F^c|E_1).$$

We have found

$$P(E_1) = P(E_1|F)P(F) + P(E_1|F^c)P(F^c) = (.4)(.3) + (.2)(.7) = .26.$$

Thus

$$\begin{aligned} P(F|E_1) &= P(E_1|F)P(F)/P(E_1) = (.4)(.3)/(.26) = \frac{6}{13}. \\ P(F^c|E_1) &= \frac{7}{13} \\ P(E_2|E_1F) &= P(E_2|F) = .4 \\ P(E_2|E_1F^c) &= P(E_2|F^c) = .2. \end{aligned}$$

Thus from the above formula, $P(E_2|E_1) = (.4)\frac{6}{13} + (.2)\frac{7}{13} \approx .29$.

1.3 Conditionally independent events

In the above example, we have *implicitly assumed* that the events E_2 and E_1 are independent given F (and also independent given F^c). More specifically, two events E_1, E_2 are conditionally independent given F if

$$P(E_1|E_2F) = P(E_1|F),$$

or equivalently,

$$P(E_1E_2|F) = P(E_1|F)P(E_2|F).$$

2 The hat problem revisited

We can approach the hat problem yet another way. Let F be the event that no one selects their correct hat, E_i be the event the ith person selecting his correct hat. Then

$$P(F) = P(F|E_1)P(E_1) + P(F|E_1^c)P(E_1^c).$$

Now $P(F|E_1) = 0$, so we need to compute $P(F|E_1^c)$ (we already know how to compute $P(E_1^c)$. Apply the rule again we have

$$P(F|E_1^c) = P(F|E_2E_1^c)P(E_2|E_1^c) + P(F|E_2^cE_1^c)P(E_2^c|E_1^c) = P(F|E_2^cE_1^c)P(E_2^c|E_1^c),$$

by a similar reasoning. How to compute $P(E_2^c|E_1^c)$? We need to use the definition.

$$P(E_2^cE_1^c) = P(E_2^cE_1^c)/P(E_1^c) = (1 - P(E_1 \cup E_2))/P(E_1^c).$$

Note that the denominator $P(E_1^c)$ will cancel with the $P(E_1^c)$ in the first equation. Similarly, we have

$$P(F|E_2^cE_1^c) = P(F|E_3^cE_2^cE_1^c)P(E_3^c|E_2^cE_1^c),$$

and you'll quickly discover that we have cancellation between $P(E_2^c E_1^c)$ in the denominator here and the above equation as well. Eventually we'll discover that

$$P(F) = P(E_1^c E_2^c \cdots E_{10}^c),$$

a tautology ! So there seems to be no gain at first glance. However, I would argue there are two benefits from this approach. First it is the "natural" approach we would use for this problem, thinking what is the probability if the first person gets his right hat or wrong hat etc. Second, this approach naturally suggests to look at the complementary problem, and along the way we are forced to compute $P(E_1 \cup E_2 \cup E_3 \cdots)$, which as you remember was the way to solve the problem. So actually we would arrive at the answer in a long way, but this approach would point to us the right direction to attack the problem.

3 The Monty Hall problem

Example 3.1. *There is a game show, where you're presented with 3 closed doors. Behind two doors are two goats (which are worthless) and behind a third door is a prize you want (for example, a car). You're first asked to pick a door.* **Without opening the door you picked**, *the showman will open another door, showing you the goat behind it (thus pointing out to you a wrong door). Now you're given two options: to stay with your choice or to switch the door. The question is should you switch?*

Ans: Before giving the answer, let's go over two wrong (but potentially hard to recognize) answers to the question. The first one is it doesn't matter, since the probability of you winning the prize is 1/3, whether you switch or not. The 2nd one is it doesn't matter, since the probability of you winning the prize is 1/2, whether you switch or not.

The reason the first one is wrong is because the answer failed to recognize the sample space has changed. BEFORE the wrong door is showed to you, the probability of picking the right one is 1/3. But AFTER the wrong door is showed, an event has happened, thus your sample space is change, and you're into a conditional probability scenario. Thus your probability might have changed.

The reason the 2nd one is wrong is also because the answer failed to recognize the sample space has changed in a correct way. It is better than the first one because it has recognized you're left with 2 doors, not 3 as before. However it failed to recognize that you've also CHOSEN a door, which is also an event that has happened that needs to be taken into account.

So how to properly answer this question? The answer can be done by counting. We count how many ways you can win if switch, given that you choose some door and the wrong door has been showed to you. Note that there are 3 ways you can choose a door, 2 wrong and 1 right. If you choose the right door, then if you switch after being showed a wrong door you will choose a wrong door. If you choose the wrong door, then if you switch after being showed a wrong door you will choose the right door. Since these events are equally likely, the probability you win if switch, conditioning on choosing some door and being showed a wrong door is 2/3. Thus you should switch.

Mathematically how do we formulate this problem? Let's assume you always switch your choice. Let F be the event that you win, E_1 be the event that the wrong door is showed to you, E_2 being the event that you chose the wrong door at the first try. We're computing $P(F|E_1)$ because the wrong door is always showed to you.

Then

$$P(F|E_1) = P(F|E_1E_2)P(E_2|E_1) + P(F|E_1E_2^c)P(E_2^c|E_1).$$

Now $P(E_2|E_1) = P(E_2) = 2/3$ and $P(E_2^c|E_1) = P(E_2^c) = 1/3$ since it is reasonable to assume that the event of choosing the door is independent from the event the wrong door is showed to you.

$P(F|E_1E_2) = 1$ and $P(F|E_1E_2^c) = 0$. Thus $P(F|E_1) = 2/3$ agreeing with our above reasoning.