OMCIT 592 Module 08 Self-Paced 03 (instructor Val Tannen)

Reference to this self-paced segment in seg.08.04

This is a segment that contains material meant to be learned *at your own pace*. We are trying to assist you in this endeavor by organizing the material in a manner similar to the way it is outlined in the recorded segments, however with one additional suggestion.

When you see the following marker:

we suggest that you stop and make sure you thoroughly understood the material presented so far before you proceed further.

Details of Monty Hall

You have just learned about independence. It turns out that **lack of independence** can lead to hard to understand situations. Specifically, in the lecture segment ”Sketching the Monty Hall problem” you learned about an interesting but puzzling problem in which independence fails at a crucial stage. We now analyze the problem in more detail.

The Monty Hall Problem (again). On the show Lets Make a Deal, hosted by Monty Hall, there are three doors. There is a prize behind one of the doors and goats (?!) behind the other two. The contestant chooses a door. Then the host opens a different door behind which there is always a goat. The contestant is then given a choice to either switch doors or to stay put. The contestant wins the prize if and only if the contestant chooses the door with the prize behind it. Is it to the contestants benefit to switch doors?

Answer. Let us start by thinking about an approach to solving this problem because there is a lot to be learned from the way we solve the problem itself, including a motivation for understanding conditional probability later.

Let A, B, C be the three doors. We should think about the random process here as developing in stages: prize placement, contestant's (lets call her Ann) choice, and Monty's action. One key question is how these three stages interact with each other. Another key question regards the likelihoods of the various choices that Ann, Monty, etc., make.

Thus its clear that we must rely on some **prior likelihood assumptions**, some explicit and some implicit in the text of the problem:

1. We will assume that the prize is placed by game staff behind one of the three doors with equal likelihood.

2. Ann does not know where the prize is, hence she chooses her door **independently** of that. Moreover we will assume that she chooses one of the three doors with **equal** likelihood.

3. Monty must never reveal the prize, so his action does **depend** on where the prize was placed and on which door Ann opens (this is the crucial failure of independence). Moreover, if it turns out that Monty has a choice among two doors, we will assume that he opens one of them with **equal** likelihood.

Given these prior likelihood assumptions, we proceed with our solution methodically.

Step 1: determine the sample space, i.e., who are the outcomes Each stage here has a source of randomness: (1) prize door choice, (2) Ann's door choice, and (3) Monty's door choice. Thus we can say that each stage has its own outcome, which is one of three doors. We will label the doors A, B and C.

A good rule of thumb in such a situation is to define the global outcome as a sequence (of length three, in this case) of stage outcomes for the various sources of randomness. For instance if the prize is placed behind A, Ann chooses A and Monty opens B then the outcome is AAB. It would seem that we might have $3 \times 3 \times 3 = 27$ outcomes but the constraints of the process rule some of the outcomes out, because Monty will never open the door that hides the prize. This leaves 12 outcomes, and you can see all of them in Figure 1. Lets use the notation

$$\mathcal{M} = \{AAB, AAC, ABC, ACB, BAC, BBA, BBC, BCA, CAB, CBA, CCA, CCB\}$$

for the sample space. The diagram shown in the figure is called a tree of (all) possibilities. Each edge is labeled with the name of the door that was opened at the corresponding stage. The outcomes are on the leaves of the tree.

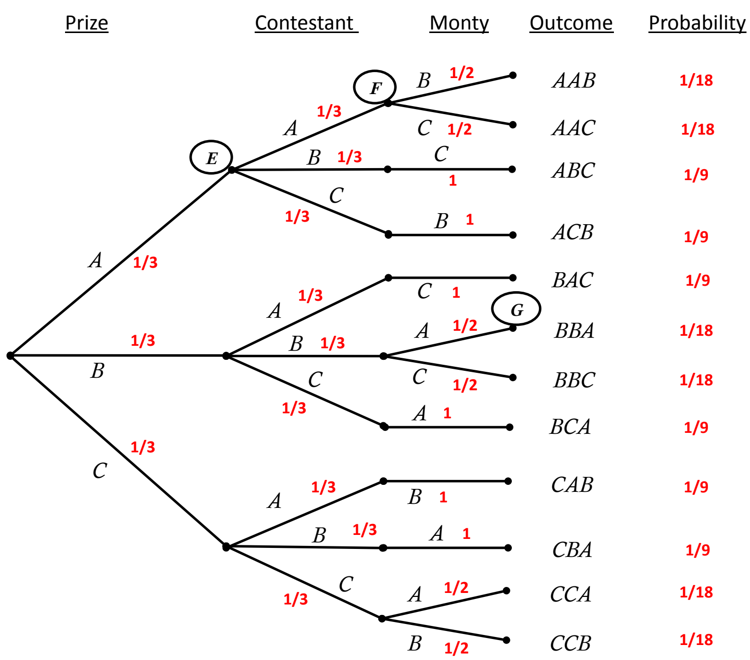

Figure 1: Tree of possibilities diagram for the Monty Hall Problem

Step 2: determine the events of interest We are interested in the outcomes in which Ann wins if she switches door choices. They are ABC, ACB, BAC, BCA, CAB and CBA. The set of these outcomes, is the event of interest here. In general, the ability to capture the events of interest is a sanity check on the modeling done in Step 1.

Step 3: determine the probabilities of the outcomes Here we begin by assigning local probabilities to the edges of the tree of possibilities, see Figure 1. In doing this we use the equal likelihood assumptions, hence the $\frac{1}{3}$ or $\frac{1}{2}$, or even the probability 1 when something happens for sure.

What is the meaning of these local probabilities and how do they relate to the probabilities of the events in the space $\mathcal{M}$? The nodes of the tree correspond to the events formed by the outcomes on the subordinate leaves. For instance, E in the figure corresponds to the event $\{AAB, AAC, ABC, ACB\}$ whose meaning is "prize is behind A". Moreover, F in the figure corresponds to the event $\{AAB, AAC\}$ whose meaning is "prize is behind A and Ann chooses A". Finally, the leaf labeled with AAB corresponds to the event whose meaning is "prize is behind A and Ann chooses A and Monty opens B".

Next is a slightly magical move: to compute the probability of an event corresponding to a node G (see figure) we multiply the edge probabilities along the branch of the tree from root to G: $(\frac{1}{3})(\frac{1}{3})(\frac{1}{2}) = \frac{1}{18}$. Why? It turns out that this is a consequence of understanding the concept of conditional probability and its **chain rule** as we shall see in a future segment. This gives us the outcome probabilities in Figure 1.

Here is a sanity check. The event "Ann chooses door A" is $\{AAB, AAC, BAC, CAB\}$ and its probability is $\frac{1}{18} + \frac{1}{18} + \frac{1}{9} + \frac{1}{9} = \frac{1}{3}$. We can check that this event is independent of events like prize is behind door A. While we can say that the 1/3's on the first stage edges do correspond to events in $\mathcal{M}$, this is not the case for second and third stage edges, which are conditional on events that have already happened. For example, 1/2 is the probability that Monty opens door B given that the prize is behind A and Ann chose A also. Note that we are also in agreement with the definition of of independence: the probability of F is $(\frac{1}{3})(\frac{1}{3})$ and we have noted that prize placement and Anns choice are independent.

Step 4: compute the probability of the events of interest This is straightforward now:

$$Pr[\{ABC, ACB, BAC, BCA, CAB, CBA\}] = 6 \cdot \frac{1}{9} = \frac{2}{3}$$

Therefore, Ann reasons that the probability that she is in a situation in which it is in her interest to switch is $(\frac{2}{3}) > (\frac{1}{2})$ and she switches doors.