The Monty Hall Problem

The Monty Hall problem has become too much of a classic, it is taught so often students sometimes get tired of hearing about it. It became well known in 1990 when Marilyn von Savant, an author who had a puzzle column in Parade magazine, wrote about the problem. Thousands of irate and misguided men wrote to incorrectly point out errors in what she had written, as such it The Monty Hall problem teaches us about the behaviour of irate men as well as statistics. It even features in Ian McEwan's novel Sweet Tooth. As such we won't look at the solution to the problem, instead I will describe the problem and you can work it out yourselves, one way to approach the solution is to use conditional probability, the subject of this set of notes.

The Monty Hall problem is based on a game show 'Let's make a deal', hosted by Monty Hall, who died at the age of 96 in 2017. In the problem, a kind of idealization of what really went on in the show, a contestant who does well gets the chance to try for the big prize, a car. They are presented with three doors and are asked to choose one: the car is behind one door, behind the other two are goats. The contestant, foolishly in my view, wants the car. They pick a door. Monty Hall then opens one of the two remaining doors to reveal a goat. He gives the contestant the chance to change doors. Should they change?

Conditional probability

One of the main characters in the Ian Banks novel *Canal Dreams* is afraid of flying; people try to reason her out of it, pointing out that being in a plane crash is very unlikely. She points out in return that it is even less likely if you aren't on a plane. We often want to model random or unknown events that are related to each other; for example, on a rare dry day in Summer in Galway, where I grew up, it was more likely to rain the next day if you could see the hills of Clare clearly across the bay. This sort of situation is modelled using conditional probabilities.

Consider the sample space made up of pairs of points:

$$X = \{(r, c), (r, u), (d, c), (d, u)\} \tag{1}$$

The outcome (r, c) corresponds to rain and a clear view of the hills, (r, u) to rain and no clear view, (d, c) to dry and a clear view and (d, u) to dry and no clear view. Let's imagine we have probabilities for each of these outcomes:

	(r, c)	(r, u)	(d, c)	(d, u)
p	1/2	1/4	1/16	3/16

Thus, the probability that it is going to rain and the hills are clear is 1/2; the probability that the hills are unclear and it is going to stay dry is 3/16.

It is useful to also define some events, for example R for rain

$$R = \{(r, c), (r, u)\} \tag{2}$$

and C for clear

$$C = \{(r, c), (d, c)\} \tag{3}$$

Equally well $\bar{R} = \{(d, c), (d, u)\}$ is the set of all points not in R, it is called the **complement** of R and is given by

$$\bar{R} = X \setminus R \tag{4}$$

where '\' is **set minus**. This event corresponds to a dry day. It is easy to work out

$$P(R) = 3/4 \tag{5}$$

and

$$P(C) = 9/16 \tag{6}$$

Now we can also work out probabilities for the intersections, for example:

$$P(R \cap C) = 1/2 \tag{7}$$

This is the probability of R and C both happening, in the small example we are looking at here, this is just one outcome (r, c), but, in general, the intersection might have more points in it. Now, notice that by definition

$$C = (R \cap C) \cup (\bar{R} \cap C) \tag{8}$$

and, since $(R \cap C) \cap (\bar{R} \cap C) = \emptyset$ then

$$P(C) = P(R \cap C) + P(\bar{R} \cap C) \tag{9}$$

Hence, the probability of C is the probability of C and R plus the probability of C and not R, or, in words; the probability that it is clear is the probability that it is clear and raining plus the probability that it is clear and not raining.

Now we define

$$P(R|C) = \frac{P(R \cap C)}{P(C)} \tag{10}$$

This is the probability of R given C; if we know it is clear how likely is it to be raining. The denominator $P(C)$ is needed because the probabilities must add up to one and $P(C) = P(R \cap C) + P(\bar{R} \cap C)$. Thus

$$P(R|C) = \frac{1/2}{9/16} = \frac{8}{9} \approx 0.89 \tag{11}$$

In other words, if we can see the hills of Clare, there is a 89% chance it will rain tomorrow.

More generally, if we have two events A and B then the conditional probability is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{12}$$

Here is a quick example. In the card game 21 you get given two cards and you add up their values; you then decide whether to take more cards, with the aim of being no more than and but as close to 21 as possible. In 21 picture cards count as ten; you can count an ace as one or 11; for simplicity lets pretend an ace has to be one and we are concentrating on the original hand of two cards. Now imagine you are given two cards; if the first card is a king what is the chance your hand is bigger the 17 in value? Let B be the set of hands that has a king as the first card and A is the set of hands worth more than 17. First lets calculate $P(B)$. There are 52×51 hands in total, and 4×51 of them have a king as the first card so

$$P(B) = \frac{4}{52} = \frac{1}{13} \tag{13}$$

Now, lets count the number of hands with a king as the first card, and a total value over 17. Given that the first hand is a king for the total to be over 17 the second card is an 8, 9, 10 or picture card. This might seem like there are $6 \times 4 = 24$ cards to go along with the king, but it isn't, since the first card is a king there are only 23 other cards to go along with it to make 18 or higher. Thus

$$P(A \cap B) = \frac{4 \times 23}{51 \times 52} \tag{14}$$

where the first four corresponds to the choice of king. Thus

$$P(A|B) = \frac{23}{51} \approx 0.45 \tag{15}$$

In fact, conditional probability is very frequently calculated in situations like the first one, where the sample space is made of pairs. This make sense, it is like there are two things to measure and you use conditional probability to see how one relates to the other. Here is another example; imagine thinking about the journey times from Temple Meades to the top of Park Street. Consider a table of the describing the frequency different modes of transport are used and different times of travel are measured.

	car	walk	cycle
< 20	0.025	0.1	0.2
$[20, 40]$	0.125	0.35	0.05
> 40	0.1	0.05	0

Using the obvious notation we can work out the **marginal** distributions: for vehicles $P(\text{car}) = P(\text{cycle}) = 0.25$ whereas $P(\text{walk}) = 0.5$, for times $P(< 20) = 0.325$, $P([20, 40]) = 0.525$ and $P(> 40) = 0.15$. Now, say someone rushes in to your meeting and says they are sorry but it took them over 40 minutes to get from Temple Meades, what is the chance they travelled by car:

$$P(\text{car} \mid > 40) = \frac{P(\text{car} \cap > 40)}{P(> 40)} = \frac{0.1}{0.15} = \frac{2}{3} \tag{16}$$

Looking at the bottom row of the table you can see how it all works; 'car' makes up two thirds of the values on that row; the denominator just serves to normalize the row so that it is a probability.

Summary

- The **conditional probability** of event R given C:

$$P(R|C) = \frac{P(R \cap C)}{P(C)} \tag{17}$$

  This is the probability of getting an outcome in event R if we know the outcome is in event C.

Example question

This is known as the 'second sibling' problem and like the Monty Hall problem it was popularized by Marilyn Vos Savant:

> A shopkeeper says she has two new baby beagles to show you, but she doesn't know whether they're male, female, or a pair. You tell her that you want only a male, and she telephones the fellow who's giving them a bath. "Is at least one a male?" she asks him. "Yes!" she informs you with a smile. What is the probability that the other one is a male?

solution So we are assuming that males and females are equally likely and that the sex of the two pups is independent; the set of outcomes, using the obvious notation, is $\mathcal{X} = \{(m,m),(m,f),(f,m),(f,f)\}$ with each event having probability 1/4. The event that at least one is male is $\mathcal{M} = \{(m,m),(m,f),(f,m)\}$. This means that $P(\mathcal{M}) = 3/4$. Lets call the event that the second dog is male $\mathcal{S}$ so $\mathcal{M} \cap \mathcal{S} = \{(m,m)\}$ and

$$P(\mathcal{M} \cap \mathcal{S}) = \frac{1}{4} \tag{18}$$

and hence

$$P(\mathcal{S}|\mathcal{M}) = \frac{1}{3} \tag{19}$$

and so the probability the second dog is male is one in three.