

PRINCIPLES OF BEHAVIORAL ECONOMICS

Prof. Sujata Kar

Department of Management Studies

IIT Roorkee

Week 19

Lecture 19

Welcome back to the course on principles of behavioral economics. This is lecture 19, and I am going to talk about the fundamentals of probability theory. Basically, we are beginning with a new topic on decisions under risk and uncertainty. So, whenever there are risks and uncertainty, we need to bring in the concept of probability, and prospect theory also uses probability. So, we are going to cover the basics of Bayesian analysis and prospect theory.

In order to understand them better, we begin with the fundamentals of probability theory. So, in this module, we are going to first introduce some basic definitions and rules. We begin with very basic definitions like outcome space. The outcome space is the set of all possible individual outcomes. For example, if I flip a coin, then heads and tails,

both of them together constitute the outcome space; these are the possible outcomes. In a similar fashion, if I roll a die, then the outcome space will have six components: 1, 2, 3, 4, 5, 6. That is, these are the six possible outcomes. So, this is the outcome space when a die is rolled once. Now, what is an outcome? Understandably, an outcome is a subset of the outcome space.

So here in this case, H, which is head or tail when a coin is flipped, are individual outcomes. In a similar fashion, 1, 2, 3, so these are individual outcomes. On the other hand, if a dice is rolled, alternative outcomes, the way it can be 1, 2, 3, 4, we can also have outcomes like 1, 2, 3 together for a number less than or equal to 3. So if we are interested in an outcome where we are looking at the possibility of a number turning up which is between 1 and 3, both inclusive,

then this would be the outcome space. In a similar fashion, if we are concerned about outcomes, that are only even numbers, then our outcome space would be 2, 4, 6. If we are interested in odd number outcomes, then our outcome space would be 1, 3, 5, and so on. Sorry, these are outcome spaces, and these are outcomes. So 2, 4, 6 could be an outcome;

1, 3, 5 is another outcome. Now, talking about Probability, how do we define probability? The probability function is a function which is written like Pr.

At times, it is also written as simply P, that assigns a real number to each outcome. And that real number is basically probability. The probability of an outcome A is the number Pr(A) assigned to the outcome A by the probability function Pr(.). So, we at times need to talk about probability functions, but otherwise, the probability values could at times be obvious also.

If we take an example, the probability of rolling a 6 is denoted by Pr(6). So, when I am rolling a fair die, the probability of a 6 coming up is denoted by Pr(6). The range of probability is basically the probability of any outcome A is a number between 0 and 1 both inclusive. So, probability actually begins from 0.

So, something which cannot happen is given a probability of 0, and something which is certain to happen is given a probability of 1. Any other things which are possible take values between 0 and 1. So, a certain occurrence is given a probability of 1. Something which cannot happen is given a probability of 0. So, for example, what is the probability that a fish will have two legs?

That is, we can assign a probability of 0 to it. Similarly, if I am asked, 'What is the probability that you are going to sleep for at least 5 hours tonight?' I would say, 'Yeah, it is one.' I'm going to sleep for at least five hours tonight. Or, I generally sleep for at least five hours every night.

So, I assign a probability of one. This is a certain thing. What is the probability that I'm going to get my salary credited to my salary account at the end of every month? I would assign a probability of 1. So, this is how something which is very much certain—of course, there are things which can suddenly change.

Setting aside those exigencies, emergencies, extremely uncertain situations, we can assign probability 1 to most of the events about which we are pretty much certain. Anything in between—something which we are uncertain about—we would be assigning probability or the rate of possibility between 0 and 1. There is something called the Equiprobability rule, which is: if there are n equally probable individual outcomes and we are denoting them by A1, A2, A3, and An, then the probability of any one individual outcome Ai is equal to 1/n. That is, the probability of Ai is equal to 1/n. We would write it like this.

$$\text{Pr}(A) = 1/n$$

So, for example, if a fair dice is rolled, the probability of getting any number between 1 and 6, both inclusive, is $1/6$. It's a fair dice, so when I throw it, there is an equal probability of any number turning up, and as a result, each individual number has a probability of $1/6$. I would write, or we would write, probability 1 equals probability 2, and so on, up to probability 6, all equal to $1/6$. Two outcomes, A and B, are called mutually exclusive if only one of them can happen at a time. For example, in the case of rolling a dice, we cannot have 1 and 2 appearing simultaneously. Only one number can come up.

So each one of these numbers, one to six, is mutually exclusive. Similarly, when we flip a coin, heads and tails are mutually exclusive outcomes. When a fair coin is flipped, since only one of them can occur, it's not possible for both heads and tails to come up simultaneously. So mutually exclusive events are those where, specifically, if one happens, the other cannot happen. Now we apply something called the OR rule for mutually exclusive outcomes. If two outcomes, A and B, are mutually exclusive, then the probability of A or B equals the probability of A plus the probability of B.

So we would write it like this: probability of A or B equals the probability of A plus the probability of B. If you want to know the probability of rolling a 1 or 2 when you roll a fair six-sided die, then this would be given as follows. So the probability of 1 is $\text{Pr}(1)$, and the probability of 2 is $\text{Pr}(2)$. The probability is $1/3$. I'm just summing up $1/6$ and $1/6$.

$$\begin{aligned}\text{Pr}(A \vee B) &= \text{Pr}(A) + \text{Pr}(B) \\ &= \text{Pr}(1) + \text{Pr}(2) = 1/6 + 1/6 = 1/3\end{aligned}$$

So this is to be remembered: when the events are mutually exclusive, then the probability of one happening or the other happening can be added together to arrive at the probability of one or the other taking place. Then we talk about the 'everything' rule. The probability of the entire outcome space is equal to one. This simply means, if I take the rolling dice example, the outcome space is 1, 2, 3, 4, 5, 6, and the probability of all of them happening.

Of course, they cannot happen together, so I need to apply the OR rule, that is, the probability of 1 or 2 or 3 or 4 or 5 or 6, so them together—the probability of all of them happening. I by, I mean, this or this or this, not together, of course, is, again, we sum them up. So, probability 1, probability 2, probability 3, probability 4, 5, and 6, all of them summed up together would be—I am basically summing up 1 upon 6, 6 times, right? And that is why I have $1/6$ multiplied by 6, which gives us 1.

$$\Pr(1,2,3,4,5,6) = \Pr(1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6) = \Pr(1) + \Pr(2) + \Pr(3) + \Pr(4) + \Pr(5) + \Pr(6) = 1/6 \times 6 = 1.$$

There is also something called the 'not' rule, right?

The probability that some outcome A will not occur is equal to 1 minus the probability that it does occur. For example, we would first of all write it not like this. So, probability not A is 1 minus probability A. For instance, if I think that today there is a 20% possibility of raining. So if A is rain, then the probability of A, according to my opinion, is 20%, that is, 0.2.

The probability of no rain would be 1 minus 0.2, which equals 0.8. So, there is an 80% chance that it won't rain today. Suppose you want to know the probability of rolling anything but a 6 when rolling a 6-sided die. Then what? Again, the simple rule is that probability not 6, is 1 minus probability 6, which is 5/6. Two outcomes, A and B, are independent if the occurrence of one does not affect the probability of the occurrence of the other.

$$\Pr(-A) = 1 - \Pr(A)$$

$$\Pr(-6) = 1 - \Pr(6) = 1 - 1/6 = 5/6$$

For example, the first time you flip a coin, you get heads. The second time you flip a coin, you get tails. So you can see that the coin, of course, does not remember what it produced on the first trial, and then it doesn't decide what is going to happen in the second trial. So whenever we flip a coin, the thing that comes up is actually independent of each other. First heads do not necessarily or do not at all imply anything

that the next one will be either heads or tails. Both of them have equal probability. So, that is why we say that these two events or outcomes are independent of each other. Alternatively, you can understand that as I took the example of today being a rainy day; rainy days and sunny days are not independent of each other. If it is going to rain, then it cannot be sunny, or if it is sunny, then it cannot rain.

So, they are not independent. But there could be other instances when we can take examples of independent events. For example, whether we are going to have lunch or not does not depend on whether it is rainy or sunny. So regardless of whether it is rainy or sunny, we can have lunch. We definitely have lunch.

So that's why or we might not have lunch. If we do not have lunch, then that could be completely for different reasons. So at least for me say weather does not impact the possibility of me having a lunch. So if the possibility of me having a lunch is say 90% and the possibility of the day being rainy is 20% I would say that these two events are at least independent for me. One is not impacting the occurrence or the prevalence of the other.

Now we apply AND rule to independent events. If two outcomes A and B are independent, then the probability of A and B equals the probability of A multiplied by the probability of B. So, if I take my rain and lunch example, I said that the probability that I am going to have lunch is 0.9 and probability that it is going to rain is 0.2. So, what is the probability that I am going to have lunch and today being a rainy day is 0.9 multiplied by 0.2. We get 0.18.

Something like this we studied during the discussion on the conjunction fallacy as described by Kahneman and Tversky. This is the same rule that applies here. Now, we take another example. Generally, we would write it as the probability of A and B equals the probability of A multiplied by the probability of B. So, you can see it here. Suppose you flip a fair coin twice; what is the probability of getting two heads?

If the outcomes are denoted by H1 and H2, this is the head from the first trial, this is the head from the second trial. Then, the probability of H1 and H2 is the probability that you get a head in the first trial and the probability you get a head in the second trial. So, in each trial, the probability is $1/2$, and as a result, the probability of getting two heads is $1/4$. The probability that something happens given that some other thing happens is called conditional probability.

So, now we are going to discuss conditional probability, which simply implies that the occurrence of one event is conditional upon the occurrence of some other event. We write the probability of A given C, or the probability of A conditional upon C, as probability A then the straight line C. Suppose H means the patient has a headache, and T means the patient has a brain tumor. So, probability H/T implies that the patient has a headache given that he has a brain tumor, and probability T/H says that the patient has a brain tumor given that he has a headache. So, you can see that the two things are not the same.

And the two things do not necessarily have the same probability. It's possible that the probability of H given T is 0.9. It's possible that the probability of T given H is 0.6. So, they neither imply the same thing nor have to have the same numbers associated with them. They are completely different things.

But this is how we write the conditional probability. Now we take an example of conditional probability. Suppose from a pack of 52 cards, you draw two cards without replacement. So once one card is taken out, then you take out another card. What is the probability of drawing two aces?

So I just hope that all of you are familiar with the types of cards that are there in a pack of cards. So this is an ace of spades. the two outcomes are not of course independent. Why? Because there are only four aces in a pack of cards. So once you draw one card and if it turns out to be an ace, then that is going to impact the possibility of the second card being an ace. so we first ask what is the probability that the first card is an ace the answer is $4/52$ there are 52 cards and there are four aces so

$4/52$ is the probability that the first card will be an ace. $4/52$ equals to $1/13$. Next, we ask, what is the probability that the second card is an ace? Given that the first card was an ace. So if the first card was an ace, then you understand that there will be, first of all, only 51 cards left because I have already taken out one card.

And then there will be 3 aces left because 1 ace has already come. So, because there are 51 cards left in the deck and only 3 of them are aces, the probability is $3/51$ for the second card to be an ace given that the first was an ace. If the first one was not an ace, then it would be $4/51$. Now we apply the AND rule to obtain $4/52$ multiplied by $3/51$. So what is happening? This is happening: the first card is an ace, and then the second card is also an ace given that the first one was an ace, right?

So this AND rule is applied here to arrive at a number like $1/221$. In the previous slide, we calculated $P(A \text{ and } B)$. Now, if $P(A \text{ and } B)$ are two outcomes, then formally the conditional probability formula for the probability of A given B is written as the probability of A and B divided by the probability of B. Now, if we continue with the cards example, let A denote the second card is an ace and B denote the first card is an ace. So here, actually, we are trying to find out the probability of the second card being an ace while the first card was an ace.

We know that the probability of drawing two aces without replacement is $1/221$, which we calculated in the previous slide. We also know that the probability that the first card is an ace is $1/13$ because, as we discussed, there are four aces in a pack of 52. So, $4/52$ makes it $1/13$, right? Now, if I plug in these two values into this expression, then what I obtain is that $1/221$ divided by $1/13$, which is approximately equal to 0.0588, which is also approximately equal to $3/51$. Now, if you remember, the probability of the second card

being an ace, while the first card was also an ace, was calculated to be exactly $3/51$ in the previous slide.

So, this just shows that this formula works—that what is the probability of something occurring, given that something else has occurred previously. So, you would write it as the probability of both of them happening together, like probability A and B, divided by the probability of the first thing happening, that is probability B. This is the formula. Now, we talk about the general AND rule. The general AND rule says the probability of A given B multiplied by the probability of B is equal to the probability of B given A multiplied by the probability of A. This is very simple to prove because we already know this from the previous slide. Right? This is basically the conditional probability formula that we just learned in the previous slide. Now probability A and B is equal to probability A given B multiplied by probability B which is simply taking this to the other side so from this expression we will be deriving this one now logically probability A and B should be equal to probability B and A, where probability B and A would be written as this. So now you can see that if these two things are the same, then this and this should also be the same. So it is just written here: the probability of A given B multiplied by probability B is equal to the probability of B given A multiplied by probability A.

$$\Pr(A \& B) = \Pr(A) \times \Pr(B)$$

So this is called the general AND rule. Now, let us take an example of the general AND rule. Suppose you draw one card from a well-shuffled pack, and A means you draw an ace, and D means you draw a diamond. So it follows that probability of drawing an ace conditional upon probability of drawing a diamond multiplied by probability of drawing a diamond is equal to you draw a diamond conditional upon ace. multiplied by the probability of drawing an ace.

So you can check that this is true by plugging in the numbers. What are the numbers? So if you draw an ace, then that is because—see, if you draw a diamond previously or draw an ace, it is simply the probability of $4/52$. If you are drawing it for the first time, then it is $1/13$ multiplied by $13/52$, which is the probability of drawing a diamond. Diamonds—there are 13 diamonds in a pack of cards.

So, this is $13/52$. And similarly, if you are drawing a diamond, then this is $1/4$ multiplied by drawing an ace, which is $4/52$. That gives us $1/52$. So one trick here is that I'm just drawing one card here. So basically, drawing the first card is actually not conditional upon whether you are drawing diamonds or anything, because you're not drawing two cards.

So the probability of A given D is as good as the probability of A. Similarly, the probability of D given A is as good as the probability of D, because you are drawing a single card from the pack. And it is not conditional upon what you drew previously, because this is the first and only time you are drawing the card. Anyway, what it shows is that this formula holds—this formula holds good in this example. Now, we talk about independence conditions.

If two outcomes A and B are independent, then the following three claims are equivalent. So, this is just an elaboration of the independence condition. Just as in the previous example, I said that if you are drawing only one time, it is not conditional upon anything. So, the probability of drawing an ace is not conditional upon anything, and that is why the probability of A given D is equal to the probability of A. In a similar fashion, this is written in a generalized format.

1. $\Pr(A/B) = \Pr(A)$

2. $\Pr(B/A) = \Pr(B)$

The probability of A given B is equal to the probability of A. Similarly, the probability of B given A is equal to the probability of B. As I said, the probability of drawing a diamond given A is equal to the probability of a diamond because you are not drawing an ace. This is the only one card that you are taking out. And the probability of A and B equals the probability of A multiplied by the probability of B. This is, of course, something we discussed previously. This is valid only for independent events.

Next, we talk about the total probability rule. Conditional probabilities can also be used to compute unconditional probabilities. Generally, if I say that the probability of a head appearing from a fair coin is an unconditional probability. In this sense, it is not contingent upon something else happening, right? So, here conditional probabilities can also be used to compute unconditional probabilities.

We will see what the formula for that is. Suppose you are running a frisbee factory and you want to know the probability that one of your frisbees is defective. We need some more information. So this is the information. You have two machines producing frisbees.

A new one, denoted by B, produces 800 frisbees per day. And you have an old machine, denoted by not B, that produces 200 frisbees per day. Thus, the probability that a randomly selected frisbee from your factory was produced by machine B is 800 divided by 1000. Of these, 800 are produced by machine B and 200 by not B, the old machine. And 800 divided by 1000 makes it 0.8.

A randomly selected frisbee produced by machine B has a probability of 0.8. The probability that it was produced by machine not B would similarly be 200 divided by 1000, or we can write it as 1 minus the probability of B, which equals 0.2. Here we are applying the NOT rule that we learned in the beginning. Among the frisbees produced by the new machine, 1% are defective. Among those produced by the old one, 2% are defective.

The probability that a randomly selected frisbee produced by machine B is defective is written like this. So, it is defective provided that it was produced by machine B. What is the probability? It is 0.01 simply because we have mentioned it here clearly. Similarly, the probability that a randomly selected frisbee produced by machine not B is defective is 0.02 because this is 2 percent clearly mentioned here.

So, we write it that the probability of one randomly selected frisbee turning out to be defective provided that it was produced by machine B is 0.02. Now, what is the probability that a randomly selected frisbee from your factory is defective? So, someone coming from outside, not knowing how many machines or anything that you have, would just pick up a frisbee to inspect. What is the probability that it will turn out to be a defective one? The probability that a frisbee is defective, not conditional upon anything, is denoted by $\Pr(D)$, the probability of a frisbee being defective.

So, it equals the following probability: that the frisbee is produced by machine B and turns out to be defective or the frisbee is produced by machine not B and turns out to be defective. So, we cannot apply the AND rule here because just one frisbee is picked up; it has to be either produced by machine B or produced by not B, that is, the old machine. So, we need to apply the OR rule here. So, applying the OR rule, we obtain probability D equals what we write.

That it was produced in machine B and defective, or produced in machine not B and defective, which is equal to simply replacing this OR sign with a plus. It is the probability of defective and produced by machine B, again the probability of defective produced not by machine B. And we can always do this because these are mutually exclusive. Only one is picked up. It has to be from here or here.

So they are mutually exclusive events or outcomes. Now applying the general AND rule twice, that is one for D and B and another for D and not B, what we observe is that this is expanded like this. This is expanded like this. And now we plug in the values. What are the values?

Probability D given B, that is, it is produced by machine B and then it turns out to be defective, is 0.01, and any frisbee randomly picked up is produced by machine B is 0.8. Similarly, it is defective and produced by machine given that it was produced by machine not B is 0.02, and any randomly picked up frisbee produced by machine not B is 0.2. By multiplying and adding, we arrive at a number of 0.012, which is 1.2 percent. So the probability that a randomly picked up frisbee will turn out to be defective is 0.012, or a 1.2 percent chance that it is going to be defective. These calculations illustrate the rule of total probability, which can formally be stated as.

$$\begin{aligned} \Pr(D) &= \Pr(D \& B) + \Pr(D \& \neg B) \\ &= \Pr(D / B) \times \Pr(B) + \Pr(D / \neg B) \times \Pr(\neg B) \end{aligned}$$

$$\Pr(D) = 0.01 \times 0.8 + 0.02 \times 0.2 = 0.012 = 1.2\%$$

$\Pr(D)$ equals this expression plus this expression, right? So, these are, you know, directly taken from above. We take another example. You are scheduled to take the test required to be a certified scuba diver and very much hope you will pass. So, the probability of passing the test is given by P.

Not the probability, but the event that you are going to pass is given by P. The test can be easy, denoted by E, or not easy, denoted by not E. The probability that it is easy is 60%. So, the probability that it is going to be easy is 0.6. Obviously, the probability it is not going to be easy is 0.4. If it is easy, you estimate that the probability of passing is 90%. So, the probability you will pass given that it is easy is 0.9.

What else? If it is hard, you estimate that the probability is 50%. So, the probability of passing given that it is not easy is 0.5, right? So, what is the probability that you pass? How do we calculate it?

Simply, we use that formula—the total probability rule—applied here. So, this is how we write it, and we plug in the value P given E. You can see 0.9, probability E is 0.6, probability P not given not E is 0.5, and probability not E is 0.4. By multiplying and summing, we arrive at a number of 0.74, which tells us this is the probability of passing the test. Then, we finally talk about Bayes' rule or Bayes' theorem. Going back to the frisbee factory example, suppose you pick up one of the frisbees produced in your factory and find that it is defective.

Now, the question is: What is the probability that the defective frisbee was produced by the new machine? So far, we asked what is the probability that one randomly picked-up frisbee will be defective, regardless of which machine produced it. Here, we are asking what is the probability that a defective frisbee will be produced by your new machine.

Sorry, a defective frisbee will be produced by your new machine. Probability that the frisbee turns out to be defective and it is produced by—or rather, I would say probability that one defective frisbee is produced by machine B. So, we know that probability D and B can be written like this. This is what we learned because that has come from this formula—the conditional probability formula. The conditional probability formula says that probability D given B is equal to probability D and B divided by probability B.

So by simply changing sides, I can write it like this and calculate the values. It is defective and produced by machine B is 0.01 because 1% is defective, and it is produced by B is 0.8. So this is the formula of probability D and B, and in a similar fashion, D and not B would be 2% defective produced by machine not B. And 1 frisbee is produced by machine not B is 0.2, so 0.004. Now Bayes' rule or Bayes' theorem gives us probability B given D is equal to this, and the lower probability D, the lower expression, can be further expanded like this.

$$\Pr(B / D) = \frac{\Pr(D / B) \times \Pr(B)}{\Pr(D)}$$

$$= \frac{\Pr(D / B) \times \Pr(B)}{\Pr(D/b) \times \Pr(B) + \Pr(D/ -B) \times \Pr(-B)} = \frac{0.008}{0.012} = \frac{2}{3}$$

Which we obtain from the total probability rule. And then, plugging in the values, we will arrive at a number like 2/3. What does it say? This means that in 2 cases out of 3, that is, 2 defective frisbees or 3 defective frisbees if picked up, then 2 would turn out to be produced by machine B. A defective frisbee comes from the new machine.

So, 2 cases out of 3 defective frisbees come from the new machine. This may be surprising since the new machine has a lower rate of defective frisbees than the old one. However, it is explained by the fact that the new machine also produces far more frisbees than the old one. So, the new machine produces 800 frisbees per day, and 1% of 800 frisbees turn out to be 8 defective frisbees per day. While the old machine produces only 200 frisbees per day, and 2 percent of 200 is only 4.

So, as a result, you can see that overall the new machine produces a larger amount or a larger number of defective frisbees, even though it produces a larger total number of frisbees. So, in this, I come to the conclusion of this module. We have discussed basic

probability rules. Now, we will continue with Bayesian updating, followed by further discussions on risk attitude. Everywhere, we are going to apply these basic probability functions or rules in the ongoing topics specifically. These are the references that were used.

Thank you.