PROFESSOR: So now we start on a unit of about a half a dozen lectures on probability theory which most students have been exposed to, to some degree, in high school. We'll be taking a more thorough and theoretical look at the subject in our six lectures but, before we begin, let's give a little pitch for the significance of it. There's been extensive debate among the faculty that probability theory belongs right up there with physics and chemistry and math as something that should be a fundamental requirement for all students to know. It plays an absolutely fundamental role in the hard sciences, and the social sciences, and in engineering that pervades all those subjects. And it's hard to imagine somebody legitimately being called fully-educated if they don't understand the basics of probability theory.

Historically, probability theory starts off in a somewhat disreputable way in the 17th and early 18th centuries with the analysis of gambling, but then it goes on to be the basis for the insurance industry and underwriting, predicting life expectancies, so that you could understand what kind of premiums to charge. And then it goes on to allow the interpretation of noisy data with errors in it and the degree to which it confirms scientific and social science hypotheses.

But true to the historical basis, let's begin with an example from gambling that illustrates the first idea of probability and then we're going to be working up to a methodology for inventing probability models, called the tree model. So let's begin with an example from poker and I'd like to ask a question. If I deal a hand of five cards in poker, what's the probability of getting exactly two jacks? So there are 13 ranks and there are four kinds of jacks-- space, hearts, diamonds, clubs-- what's the probability that, among my five cards, I'm going to get two of them? Well, that's really a counting problem because I'm going to ask, first of all, how many possible five-card hands are there?

We can think of these as the outcomes of a random experiment of just picking five cards. And there are 52 choose 5 five-card hands in a 52-card deck. Then, there are 4 choose 2 ways of picking the suits for the two jacks that we have and so the total number of hands that have two jacks is simply 4 choose 2 times 52 minus 4, the remaining 48 cards, choose the remaining 3 cards in the five-card hand. And then what we would say is that the probability of two jacks is basically the number of hands with two jacks divided by the total number of hands. Turns out to be about 0.04 and, under this interpretation, basically, what we're thinking of probability as telling us is, what fraction of the time do I get what I want? What's the fraction of the time that I quote, "win" , if winning consists of getting a pair of jacks and, by symmetry and the fact that we think of one hand is as likely to come up as another, this fraction of hands that equal two jacks, it makes sense to think of that as that's the probability that we'll get that hand. If we think of all the hands as being equally likely, we yank 1 out of the deck, the fraction of time that we would expect to get two jacks is this number. About 0.04.

So, the general setup of probability, the first idea based on this illustration with a pair of jacks, is that, abstractly, we have some random experiment that's capable of producing outcomes. These are mathematical black boxes called outcomes. Now, a certain set of the outcomes, we will think of as an event that we're interested in whether or not it happens. We could think of it as the event of getting two jacks or the event of winning some game. Then we define the probability of an event as simply the fraction of the outcomes in the event divided by the total number of outcomes. Among all the outcomes, what fraction of outcomes are in the event? And we define that to be the probability of the event. That's the first naive idea about probability theory and it applies to a lot of cases, but not always.

So now, let's begin with an example which illustrates why this first idea needs to be refined and it doesn't really give us the kind of theory of probability that we'd like. So let's turn to a game that was really famous in the 1970s. An enormously popular TV game hosted by a man named Monty Hall. The actual name of the TV show was called Let's Make a Deal, but we'll refer to it as the Monty Hall game, and the way that this Let's Make A Deal show worked was, roughly, that there were three doors. This is an actual picture of the stage set. Door 1, door 2, door 3. And by the way, this game show still has a fan base. There's a website for it that you can look at. Even 40 years later, people are still caught up in the dynamics of the game.

So there are these three doors and the idea is that behind the doors, they're going to have a prize behind one of them and some kind of booby prize, often a goat held by a beautiful woman holding a goat on a leash just to keep things visually interesting, and that's what you got if you lost. And contestants were going to somehow or other pick a door and hope that the prize was behind it. There's a picture of the staff. There's Monty Hall and the woman who was his assistant, Carol Merrill. Her job was to pick doors to open and show them to contestants to see what was behind them.

OK. So here are the rules for the Monty Hall game. The actual quiz show had more flexible rules but, for simplicity, we're going to define a simple, precise, and fixed set of rules. The rules are that, behind the three doors, two of the doors are going to have goats and one of the doors is going to have a prize behind it. Often the prize is something like an automobile. Something really desirable. So we can assume that the staff, on purpose, will place the price at random behind the three doors because they don't want anybody to have a guess that some doors are more likely than others to have the prize and they're not allowed to cheat. That is, once they've decided which door is going to have the price, it's supposed to stay there throughout the game. They can't move it in response to which door that the contestants pick. That would be cheating.

OK. Next, the contestant is given an opportunity to pick one of the doors. They're all closed and it's hard to understand how the contestant would make a choice, but if the contestant was worried about the staff trying to outguess him on where to put the goat and where to put the prize, the contestant should just pick all the doors with equally likelihood. Then he can't be beaten by their trying to outguess him. He can only be beaten by if they cheated him by moving the goat after he picked or moving the prize after he picked. At this point, once the contestant has picked a door-- let's say he picks door 2-- then Monty instructs Carol to open a door with a goat behind it. So he's going to choose an unpicked door. If the contestant has picked door 2, that means that door 1 and door 3 are unpicked doors, and Monty tells Carol, open either door 1 or door 3, whichever one-- or perhaps both-- have a goat behind them.

And so Carol is going to open one of those doors and show a goat and everybody knows that they're going to see a goat because that's the way the game works. And then at this point, when the contestant has seen that there's a door that has a goat behind it and they're sitting on a picked door and there's another unopened door that hasn't been picked, the contestant's job is to decide whether to stick with the door that they originally picked or switch to the other unopened door. So if they picked door 2 and Carol opened door 3, they could stick with door 2 or they could switch to the closed door 1 and hope that maybe 1 has the price behind it. Those are the rules of the game.

Now, the game got a lot of prominence in a magazine column written by a woman named Marilyn Vos Savant. The name of the magazine column was called Ask Marilyn and she advertises herself as having the highest recorded IQ of all time, some IQ of 200, and so she runs a popular science and math column with various kinds of puzzles. And she took up the analysis of the Monty Hall statistics and came to a conclusion and the conclusion caused a firestorm of response. Letters from all sorts of readers, even quite sophisticated PhD Mathematicians who were arguing with her conclusion about the way the game worked and the probability of winning according to how the contested behaved.

The debate basically came down to these two positions. Position 1 said that sticking and switching were equally good. It really didn't matter what the contestant did, whether they stuck with the door that they originally picked or switched to the unpicked door after the third door had been opened and that their likelihood of finding the prize was the same. And the other argument, very emphatically, said switching is much better. You should really switch no matter what. And how can we resolve this question?

Well, the general method that we're proposing for dealing with problems like this where we're really trying to figure out, what is the probability model? Is to draw a tree that shows, step-by-step, the progress of the process or experiment that's going to yield a random output and try to assign probabilities to each of the branches of the tree as you go and use that as a guide for how to assign probabilities to outcomes. So let's begin, first of all, by finding out what the outcomes are, and we're going to be analyzing the switch strategy. So, just for definiteness, let's suppose that the contestant adopts the strategy that they pick a door, Carol opens a door that shows a goat, and they're going to switch to the non-goat closed door that they did not originally pick. They're going to switch to the other door that they can switch to and we're going to ask about, what are the outcomes and consequences of winning or losing if you adopt that strategy?

Well, the tree of possibilities goes like this. The first step in this process that we've described is that the staff picks a prize location, a door to put the prize behind, and so there are three possibilities. They could put the prize behind door 1, door 2, and door 3. OK Well, let's examine the possibility that they put the prize behind door 1. So the next stage is they pick a door and if the prize is behind one and they pick a door, again, there are three possible doors that the contestant might pick. The contestant has no idea where the price is and so the contestant can choose either door 1 or door 2 or door 3. At that point, the third event in this random process, or experiment, is that Carol opens a door that has a goat behind it.

So let's examine those possibilities. So, one possibility is that the prize is behind one and the contestant picks door one, initially. Well that means that Carol can open either door 2 or door 3 in that circumstance because both of them have goats behind them. On the other hand, if the prize is at 1 and the contestant picks door 2, the two closed doors have-- one has the prize, 1, and the other doesn't have the prize, 3. Carol has to open door three. Likewise, if the contestant picks door 3 when the prize is behind door 1, Carol has to open door 2. Here she's got a two-way branch. She can choose to open either of the two goat doors, 2 or 3. Here there's only one unopened door with a goat, she's got to open 3 there, too. OK.

And that describes the outcomes of the experiment. That's the process of the experiment and these nodes at the end, these leaves of the tree, describe the final outcomes on this branch. Now, if you look at the classification of these outcomes according to winning and losing, well, we're looking at the switch strategy. So if the price was behind 1 and the contestant picked door 1 initially, then their strategy is to switch and they're going to switch away from the prize door. So whichever door Carol opened to reveal the goat, 2 or 3, the contestant is going to switch to the other one and they're going to lose. So both of these outcomes count as losses for the contestant.

On the other hand, if the prize was behind door 1 and the contestant picked door 2, then Carol opens the non-prize door, 3, and the contestant switches from 2. The only choice they have is to switch to 1, the prize door. They win. And this other case is symmetric. And that summarizes the wins and losses in this branch of the tree. Now, of course, the rest of the tree is symmetric so we don't need to talk it through again. This is just simply the case where the prize is behind 2. The contestant has the same choices and [? Marilyn ?] has the same choices of which unopened door to choose and likewise if the prize is behind 3.

So if we look at this tree, the tree is telling us that this is an experiment which we think of as having twelve outcomes, four in each of these major branches. So there are twelve outcomes of this random experiment, of which, six are losses and six are wins for the contestant and so we discover that there's six wins and six losses. Now, the way that this game works, if you think about it-- if the switching strategy wins, that means that the sticking strategy would have lost because if switching wins, it meant that you switched to the door that had the prize and so if you hadn't switched, you must have been at a door that didn't have the prize and likewise. If switching loses, then you must have switched from the door with the prize to a door that didn't have the prize-- switching-- and that means if you'd stuck, you would have won.

So what we can say is that really analyzing the switch strategy enables us to analyze the stick strategy simultaneously because you win by sticking if and only if you lose by switching. Now this simplification doesn't hold when there's more than three doors, and that's another exercise, but for now, it's telling us that if we analyze the switch strategy, we also understand the stick strategy. And of course, that means that if you use the stick strategy then the six wins become losses and the six losses become wins and, again, there are six ways to lose and six ways to win. So the first false conclusion from this is by reasoning about it as though they were poker hands, and the false conclusion says, look, sticking and switching win with the same number of outcomes and lose with the same number of outcomes. So it really doesn't matter whether you stick or switch because the probability of winning, in both cases, is half the outcome. 6 out of 12. The probability doesn't matter. It makes no difference whether you stick or switch. And that's wrong, and we will see why soon.

The other false argument is that we think about what happens after Carol has opened a door. So, where are we? The contestant has picked a door, has no idea where the goat or the prize is. Carol opens the door and shows the contestant a goat. What's left? Well, there's two closed doors left. One is the door with the prize and the other is the door without the price that has a goat behind it and, by symmetry of the doors, the contestant has no idea what's behind the door that he picked or the remaining unopened door. They're equally likely to contain the prize and so the argument is, again, that whether you stick or switch between those two doors that haven't yet been opened, it doesn't really matter and so, again, the stick strategy and the switch strategy each win with the same 50-50 probability. And that's wrong, too.

What's wrong? Well, let's go back and look at this tree a little bit more carefully to understand what's going on. And the first thing to notice about the tree is that the structure of the tree leading to the leaves is not the same. Here's a leaf that has degree [? 2. ?] One way to get in and only one way out and here's a leaf that has degree 3. One way in and two ways out, if we think of going from the root to the leaf. And so it's not clear that these branches, these leaves, should be treated the same way. Well let's think about it more carefully, about-- how are we going to assign probabilities to the various steps of the experiment?

Well, what we're going to assume, for simplicity, is that the staff chooses a door at random to place the prize. So that means that each of these branches occurs with probability 1/3. 1/3 of the time, they put the prize behind door 1, 1/3 behind door 2, and 1/3 behind door 3. OK. Let's continue exploring the branch where they put the prize behind door 1. At that point, the contestant is going to pick a door and they can pick either door 1, 2, or 3 and, absent any knowledge and also to be sure that they can't be outguessed by the staff realizing that they mostly prefer door 1. So if they're going to switch, they'll put the prize behind door 1 to fool the contestant. The contestant's protection is, pick a door at random. Choose door 1 1/3 of the time, and door 2 1/3 of the time, and door 3 1/3 of the time in a completely unpredictable way. And so the contestants is going to choose each of those possible doors as their first choice with probability 1/3.

Now what happens next? Well, the next thing that happens is that Carol opens a door. Now this is the case where Carol has a choice. The prize is behind one and the contestant happened to pick door 1. That means doors 2 and 3 both have goats and, again, for simplicity, let's assume the Carol, when she has a choice-- she can open either door 2 or door 3, here-- does them with equal probability. So we're going to assign probability 1/2 to her opening door 2 when she has the choice between 2 or 3 and probability 1/2 that she'll open door 3 and, by the way, we saw that those were losing outcomes for the contestant.

But here, things are a little different. If the prize is behind door 1 and the contestant has chosen door 2, Carol has no choice but to open the only other unchosen door with the goat behind, namely, door 3. So we could say that this choice, really, is probability 1 and I got a little bit ahead of myself here but, having filled in the probabilities on these edges, what we figured out is that the probability of this topmost branch of losing is we said, well, 1/3 of the time you go here and 1/3 of that third you go here and 1/2 of that time you go to this vertex. So it's 1/3 of 1/3 and 1/2 of that, or a weight of 1/18 and, by symmetry, this gets weight 1/18. But this way, 1/3 of the time, the prize is behind door 1. 1/3 of the time, the contestant picks door 2 and after that, Carol is was forced to open door 3. So this branch occurs with certainty, as with probability 1, which means that we wind up at this leaf 1/3 of 1/3 of the time for sure, and its weight is 1/9. And of course, by symmetry, the similar weights get assigned to the winning and the losing.

So what we've concluded is that, although there are six wins, the weight of the wins is 6/9 because they're each worth 1/9 of the time and that winning will occur 2/3 of the time. Likewise, there are six losses but they each only occur 1/18 of the time and so we lose 1/3 third of the time by the switch strategy. The summary, then, is that the probability of winning if you switch is 2/3 and, by the remark that you win with switching if and only if you lose with sticking, it follows that you lose by sticking 2/3 of the time. And so sticking is really a bad strategy and switching is the dominant way to go.

Now, in class, we back up this theoretical analysis. It's very logical but the question is, is it true? And you can do statistical experiments and have students pick doors and goats and prizes and, sure enough, it turns out that roughly 2/3 of the time, and closer and closer to 2/3 the more times you play the game, the switching strategy wins 2/3 of the time. So, the second key idea in probability theory is that the outcomes may have different probabilities. They may have different weights. Unlike the poker hand case, when we look more closely at a random experiment with different outcomes, we will agree that, for various kinds of reasons of symmetry or logic and so on, that it make sense to assign different probability weights to the different outcomes. It's not the case that the outcomes have uniform probability, that they're all equally likely.

So, to summarize, what happens, especially-- this example illustrates the confusion about of probability theory that was engendered to even some serious experts-- but, in general, intuition is very important, as in any subject, but it's also dangerous in probability theory. Particularly, for beginners who aren't experienced about some of these traps that you can fall into and so our proposal is that you be very wary of intuitive arguments. They're valuable but you need another disciplined way to check them, and we propose that you stick with what we call the four-part method when you're trying to devise a probability model for some random experiment. So, the steps are, first, that you try to identify the outcomes of the random experiment and this is where the tree structure comes up. If you try to model, step-by-step at each stage of the tree, what the possible sub-steps are in the overall process that yields the random outcome, that's where the tree comes in as we illustrated with Monty Hall.

The next thing to do is, among the outcomes, identify the ones that you consider to be of the winning events or the winning outcomes or the outcomes in the event that you are concerned about whether or not it happens. Getting two jacks, picking the door with the prize. So you need to identify the target event whose probability you're interested in. We could call it the winning event, the probability of winning. The third key step is to try to use the tree and logic of it to assign probabilities to the outcomes and the fourth step, then, is, simply, to compute the probability of the event which you do in a very straightforward way by basically adding up the probabilities of each of the outcomes in the event. That is the four-step method. Now, this Monty Hall tree that we came up with was very literal and wildly, unnecessarily complicated. So let's take another look at that and a simpler argument that will lead us to the same conclusion about how the Monty Hall game works, and we'll do that in the next video.