

## 6.041SC Probabilistic Systems Analysis and Applied Probability, Fall 2013 Transcript – Lecture 3

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

PROFESSOR: Let us start. So as always, we're to have a quick review of what we discussed last time. And then today we're going to introduce just one new concept, the notion of independence of two events. And we will play with that concept.

So what did we talk about last time? The idea is that we have an experiment, and the experiment has a sample space  $\omega$ . And then somebody comes and tells us you know the outcome of the experiment happens to lie inside this particular event  $B$ . Given this information, it kind of changes what we know about the situation. It tells us that the outcome is going to be somewhere inside here. So this is essentially our new sample space.

And now we need to we reassign probabilities to the various possible outcomes, because, for example, these outcomes, even if they had positive probability beforehand, now that we're told that  $B$  occurred, those outcomes out there are going to have zero probability. So we need to revise our probabilities. The new probabilities are called conditional probabilities, and they're defined this way.

The conditional probability that  $A$  occurs given that we're told that  $B$  occurred is calculated by this formula, which tells us the following-- out of the total probability that was initially assigned to the event  $B$ , what fraction of that probability is assigned to outcomes that also make  $A$  to happen? So out of the total probability assigned to  $B$ , we see what fraction of that total probability is assigned to those elements here that will also make  $A$  happen.

Conditional probabilities are left undefined if the denominator here is zero. An easy consequence of the definition is if we bring that term to the other side, then we can find the probability of two things happening by taking the probability that the first thing happens, and then, given that the first thing happened, the conditional probability that the second one happens.

Then we saw last time that we can divide and conquer in calculating probabilities of mildly complicated events by breaking it down into different scenarios. So event  $B$  can happen in two ways. It can happen either together with  $A$ , which is this probability, or it can happen together with  $A$  complement, which is this probability. So basically what we're saying that the total probability of  $B$  is the probability of this, which is  $A$  intersection  $B$ , plus the probability of that, which is  $A$  complement intersection  $B$ .

So these two facts here, multiplication rule and the total probability theorem, are basic tools that one uses to break down probability calculations into a simpler parts. So we find probabilities of two things happening by looking at each one at a time. And this is what we do to break up a situation with two different possible scenarios.

Then we also have the Bayes rule, which does the following. Given a model that has conditional probabilities of this kind, the Bayes rule allows us to calculate conditional probabilities in which the events appear in different order. You can think of these probabilities as describing a causal model of a certain situation, whereas these are the probabilities that you get after you do some inference based on the information that you have available.

Now the Bayes rule, we derived it, and it's a trivial half-line calculation. But it underlies lots and lots of useful things in the real world. We had the radar example last time. You can think of more complicated situations in which there's a bunch or lots of different hypotheses about the environment. Given any particular setting in the environment, you have a measuring device that can produce many different outcomes. And you observe the final outcome out of your measuring device, and you're trying to guess which particular branch occurred. That is, you're trying to guess the state of the world based on a particular measurement.

That's what inference is all about. So real world problems only differ from the simple example that we saw last time in that this kind of tree is a little more complicated. You might have infinitely many possible outcomes here and so on. So setting up the model may be more elaborate, but the basic calculation that's done based on the Bayes rule is essentially the same as the one that we saw.

Now something that we discuss is that sometimes we use conditional probabilities to describe models, and let's do this by looking at a model where we toss a coin three times. And how do we use conditional probabilities to describe the situation? So we have one experiment. But that one experiment consists of three consecutive coin tosses. So the possible outcomes, our sample space, consists of strings of length 3 that tell us whether we had heads, tails, and in what sequence. So three heads in a row is one particular outcome.

So what is the meaning of those labels in front of the branches? So this  $P$  here, of course, stands for the probability that the first toss resulted in heads. And let me use this notation to denote that the first was heads. I put an  $H$  in toss one.

How about the meaning of this probability here? Well the meaning of this probability is a conditional one. It's the conditional probability that the second toss resulted in heads, given that the first one resulted in heads. And similarly this label here corresponds to the probability that the third toss resulted in heads, given that the first one and the second one resulted in heads. So in this particular model that I wrote down here, those probabilities,  $P$ , of obtaining heads remain the same no matter what happened in the previous toss.

For example, even if the first toss was tails, we still have the same probability,  $P$ , that the second one is heads, given that the first one was tails. So we're assuming that no matter what happened in the first toss, the second toss will still have a conditional probability equal to  $P$ . So that conditional probability does not depend on what happened in the first toss. And we will see that this is a very special situation, and that's really the concept of independence that we are going to introduce shortly.

But before we get to independence, let's practice once more the three skills that we covered last time in this example. So first skill was multiplication rule. How do you find the probability of several things happening? That is the probability that we have tails followed by heads followed by tails. So here we're talking about this particular outcome here, tails followed by heads followed by tails. And the way we calculate such a probability is by multiplying conditional probabilities along the path that takes us to this outcome. And so these conditional probabilities are recorded here. So it's going to be  $(1 - P)$  times  $P$  times  $(1 - P)$ . So this is the multiplication rule.

Second question is how do we find the probability of a mildly complicated event? So the event of interest here that I wrote down is the probability that in the three tosses, we had a total of one head. Exactly one head. This is an event that can happen in multiple ways. It happens here. It happens here. And it also happens here. So we want to find the total probability of the event consisting of these three outcomes. What do we do? We just add the probabilities of each individual outcome. How do we find the probability of an individual outcome? Well, that's what we just did.

Now notice that this outcome has probability  $P$  times  $(1 - P)$  squared. That one should not be there. So where is it? Ah. It's this one.

OK, so the probability of this outcome is  $(1 - P)$  times  $P$  times  $(1 - P)$ , the same probability. And finally, this one is again  $(1 - P)$  squared times  $P$ .

So this event of one head can happen in three ways. And each one of those three ways has the same probability of occurring. And this is the answer.

And finally, the last thing that we learned how to do is to use the Bayes rule to calculate and make an inference. So somebody tells you that there was exactly one head in your three tosses. What is the probability that the first toss resulted in heads? OK, I guess you can guess the answer here if I tell you that there were three tosses. One of them was heads. Where was that head in the first, the second, or the third?

Well, by symmetry, they should all be equally likely. So there should be probably just  $1/3$  that that head occurred in the first toss. Let's check our intuition using the definitions. So the definition of conditional probability tells us the conditional probability is the probability of both things happening. First toss is heads, and we have exactly one head divided by the probability of one head.

What is the probability that the first toss is heads, and we have exactly one head? This is the same as the event heads, tails, tails. If I tell you that the first is heads, and there's only one head, it means that the others are tails. So this is the probability of heads, tails, tails divided by the probability of one head. And we know all of these quantities probability of heads, tails, tails is  $P$  times  $(1 - P)$  squared. Probability of one head is  $3$  times  $P$  times  $(1 - P)$  squared. So the final answer is  $1/3$ , which is what you should have guessed on intuitive grounds.

Very good. So we got our practice on the material that we did cover last time. Again, think. There's basically three basic skills that we are practicing and exercising here. In the problems, quizzes, and in the real life, you may have to apply those three skills in somewhat more complicated settings, but in the end that's what it boils down to usually.

Now let's focus on this special feature of this particular model that I discussed a little earlier. Think of the event heads in the second toss. Initially, the probability of heads in the second toss, you know, that it's  $P$ , the probability of success of your coin. If I tell you that the first toss resulted in heads, what's the probability that the second toss is heads? It's again  $P$ . If I tell you that the first toss was tails, what's the probability that the second toss is heads? It's again  $P$ . So whether I tell you the result of the first toss, or I don't tell you, it doesn't make any difference to you. You would always say the probability of heads in the second toss is going to  $P$ , no matter what happened in the first toss.

This is a special situation to which we're going to give a name, and we're going to call that property independence. Basically independence between two things stands for the fact that the first thing, whether it occurred or not, doesn't give you any information, does not cause you to change your beliefs about the second event. This is the intuition. Let's try to translate this into mathematics.

We have two events, and we're going to say that they're independent if your initial beliefs about  $B$  are not going to change if I tell you that  $A$  occurred. So you believe something how likely  $B$  is. Then somebody comes and tells you, you know,  $A$  has happened. Are you going to change your beliefs? No, I'm not going to change them. Whenever you are in such a situation, then you say that the two events are independent.

Intuitively, the fact that  $A$  occurred does not convey any information to you about the likelihood of event  $B$ . The information that  $A$  provides is not so useful, is not relevant.  $A$  has to do with something else. It's not useful for your guessing whether  $B$  is going to occur or not.

So we can take this as a first attempt into a definition of independence. Now remember that we have this property, the probability of two things happening is the probability of the first times the conditional probability of the second. If we have independence, this conditional probability is the same as the unconditional probability.

So if we have independence according to that definition, we get this property that you can find the probability of two things happening by just multiplying their individual probabilities. Probability of heads in the first toss is  $1/2$ . Probability of heads in the second toss is  $1/2$ . Probability of heads heads is  $1/4$ . That's what happens if your two tosses are independent of each other.

So this property here is a consequence of this definition, but it's actually nicer, better, simpler, cleaner, more beautiful to take this as our definition instead of that one. Are the two definitions equivalent? Well, they're almost the same, except for one thing. Conditional probabilities are only defined if you condition on an event that has positive probability.

So this definition would be limited to cases where event A has positive probability, whereas this definition is something that you can write down always. We will say that two events are independent if and only if their probability of happening simultaneously is equal to the product of their two individual probabilities.

And in particular, we can have events of zero probability. There's nothing wrong with that. If A has 0 probability, then  $A \cap B$  will also have zero probability, because it's an even smaller event. And so we're going to get zero is equal to zero. A corollary of what I just said, if an event A has zero probability, it's actually independent of any other event in our model, because we're going to get zero is equal to zero. And the definition is going to be satisfied.

This is a little bit harder to reconcile with the intuition we have about independence, but then again, it's part of the mathematical definition. So what I want you to retain is this notion that the independence is something that you can check formally using this definition, but also you can check intuitively by if, in some cases, you can reason that whatever happens and determines whether A is going to occur or not, has nothing absolutely to do with whatever happens and determines whether B is going to occur or not.

So if I'm doing a science experiment in this room, and it gets hit by some noise that's causes randomness. And then five years later, somebody somewhere else does the same science experiment somewhere else, it gets hit by other noise, you would usually say that these experiments are independent. So what events happen in one experiment are not going to change your beliefs about what might be happening in the other, because the sources of noise in these two experiments are completely unrelated. They have nothing to do with each other.

So if I flip a coin here today, and I flip a coin in my office tomorrow, one shouldn't affect the other. So the events that I get from these should be independent. So that's usually how independence arises. By having distinct physical phenomena that do not interact.

Sometimes you also get independence even though there is a physical interaction, but you just happen to have a numerical accident. A and B might be physically related very tightly, but a numerical accident happens and you get equality here, that's another case where we do get independence.

Now suppose that we have two events that are laid out like this. Are these two events independent or not? The picture kind of tells you that one is separate from the other. But separate has nothing to do with independent. In fact, these two events are as dependent as Siamese twins. Why is that?

If I tell you that A occurred, then you are certain that B did not occur. So information about the occurrence of A definitely affects your beliefs about the possible occurrence or non-occurrence of B. When the picture is like that, knowing that A occurred will change drastically my beliefs about B, because now I suddenly become certain that B did not occur.

So a picture like this is a case actually of extreme dependence. So don't confuse independence with disjointness. They're very different types of properties.

AUDIENCE: Question.

PROFESSOR: Yes?

AUDIENCE: So I understand the explanation, but the probability of A intersect B [INAUDIBLE] to zero, because they're disjoint.

PROFESSOR: Yes.

AUDIENCE: But then the product of probability A and probability B, one of them is going to be 1. [INAUDIBLE]

PROFESSOR: No, suppose that the probabilities are  $1/3$ ,  $1/4$ , and the rest is out there. You check the definition of independence. Probability of A intersection B is zero. Probability of A times the probability of B is  $1/12$ . The two are not equal. Therefore we do not have independence.

AUDIENCE: Right. So what's wrong with the intuition of the probability of A being 1, and the other one being 0? [INAUDIBLE].

PROFESSOR: No. The probability of A given B is equal to 0. Probability of A is equal to  $1/3$ . So again, these two are different. So we had some initial beliefs about A, but as soon as we are told that B occurred, our beliefs about A changed. And so since our beliefs changed, that means that B conveys information about A.

AUDIENCE: So can you not draw independent [INAUDIBLE] on a Venn diagram?

PROFESSOR: I can't hear you.

AUDIENCE: Can you draw independence on a Venn diagram?

PROFESSOR: No, the Venn diagram is never enough to decide independence. So the typical picture in which you're going to have independence would be one event this way, and another event this way. You need to take the probability of this times the probability of that, and check that, numerically, it's equal to the probability of this intersection. So it's more than a Venn diagram. Numbers need to come out right.

Now we did say some time ago that conditional probabilities are just like ordinary probabilities, and whatever we do in probability theory can also be done in conditional universes. Talking about conditional probabilities. So since we have a notion of independence, then there should be also a notion of conditional independence. So independence was defined by the probability that A intersection B is equal to the probability of A times the probability of B.

What would be a reasonable definition of conditional independence? Conditional independence would mean that this same property could be true, but in a conditional universe where we are told that the certain event happens. So if we're told that the event C has happened, then were transported in a conditional universe where the only thing that matters are conditional

probabilities. And this is just the same plain, previous definition of independence, but applied in a conditional universe.

So this is the definition of conditional independence. So it's independence, but with reference to the conditional probabilities. And intuitively it has, again, the same meaning, that in the conditional world, if I tell you that A occurred, then that doesn't change your beliefs about B.

So suppose you had a picture like this. And somebody told you that events A and B are independent unconditionally. Then somebody comes and tells you that event C actually has occurred, so we now live in this new universe. In this new universe, is the independence of A and B going to be preserved or not? Are A and B independent in this new universe?

The answer is no, because in the new universe, whatever is left of event A is this piece. Whatever is left of event B is this piece. And these two pieces are disjoint. So we are back in a situation of this kind. So in the conditional universe, A and B are disjoint. And therefore, generically, they're not going to be independent.

What's the moral of this example? Having independence in the original model does not imply independence in a conditional model.

The opposite is also possible. And let's illustrate by another example. So I have two coins, and both of them are badly biased. One coin is much biased in favor of heads. The other coin is much biased in favor of tails. So the probabilities being 90%.

Let's consider independent flips of coin A. This is the relevant model. This is a model of two independent flips of the first coin. There's going to be two flips, and each one has probability 0.9 of being heads. So that's a model that describes coin A. You can think of this as a conditional model which is a model of the coin flips conditioned on the fact that they have chosen coin A.

Alternatively we could be dealing with coin B. In a conditional world where we chose coin B and flip it twice, this is the relevant model. The probability of two heads, for example, is the probability of heads the first time, heads the second time, and each one is 0.1.

Now I'm building this into a bigger experiment in which I first start by choosing one of the two coins at random. So I have these two coins. I blindly pick one of them. And then I start flipping them.

So the question now is, are the coin flips, or the coin tosses, are they independent of each other? If we just stay inside this sub-model here, are the coin flips independent? They are independent, because the probability of heads in the second toss is the same, 0.9, no matter what happened in the first toss. So the conditional probabilities of what happens in the second toss are not affected by the outcome of the first toss. So the second toss and the first toss are independent. So here we're just dealing with plain, independent coin flips.

Similarly the coin flips within this sub-model are also independent. Now the question is, if we look at the big model as just one probability model, instead of looking at the conditional sub-

models, are the coin flips independent of each other? Does the outcome of a few coin flips give you information about subsequent coin flips?

Well if I observe ten heads in a row-- So instead of two coin flips, now let's think of doing more of them so that the tree gets expanded.

So let's start with this. I don't know which coin it is. What's the probability that the 11th coin toss is going to be heads? There's complete symmetry here, so the answer could not be anything other than  $1/2$ . So let's justify it, why is it  $1/2$ ?

Well, the probability that the 11th toss is heads, how can that outcome happen? It can happen in two ways. You can choose coin A, which happens with probability  $1/2$ . And having chosen coin A, there's probability 0.9 that it results in that you get heads in the 11th toss. Or you can choose coin B. And if it's coin B when you flip it, there's probably 0.1 that you have heads. So the final answer is  $1/2$ .

So each one of the coins is biased, but they're biased in different ways. If I don't know which coin it is, their two biases kind of cancel out, and the probability of obtaining heads is just in the middle, then it's  $1/2$ .

Now if someone tells you that the first ten tosses were heads, is that going to change your beliefs about the 11th toss? Here's how a reasonable person would think about it.

If it's coin B the probability of obtaining 10 heads in a row is negligible. It's going to be 0.1 to the 10th. If it's coin A. The probability of 10 heads in a row is a more reasonable number. It's 0.9 to the 10th. So this event is a lot more likely to occur with coin A, rather than coin B.

The plausible explanation of having seen ten heads in a row is that I actually chose coin A. When you see ten heads in a row, you are pretty certain that it's coin A that we're dealing with. And once you're pretty certain that it's coin A that we're dealing with, what's the probability that the next toss is heads? It's going to be 0.9.

So essentially here I'm doing an inference calculation. Given this information, I'm making an inference about which coin I'm dealing with. I become pretty certain that it's coin A, and given that it's coin A, this probability is going to be 0.9. And I'm putting an approximate sign here, because the inference that I did is approximate. I'm pretty certain it's coin A. I'm not 100% certain that it's coin A.

But in any case what happens here is that the unconditional probability is different from the conditional probability. This information here makes me change my beliefs about the 11th toss. And this means that the 11th toss is dependent on the previous tosses. So the coin tosses have now become dependent. What is the physical link that causes this dependence? Well, the physical link is the choice of the coin. By choosing a particular coin, I'm introducing a pattern in the future coin tosses. And that pattern is what causes dependence.

OK, so I've been playing a little bit too loose with the language here, because we defined the concept of independence of two events. But here I have been referring to independent coin tosses, where I'm thinking about many coin tosses, like 10 or 11 of them.

So to be proper, I should have defined for you also the notion of independence of multiple events, not just two. We don't want to just say coin toss one is independent from coin toss two. We want to be able to say something like, these 10 then coin tosses are all independent of each other. Intuitively what that means should be the same thing-- that information about some of the coin tosses doesn't change your beliefs about the remaining coin tosses. How do we translate that into a mathematical definition?

Well, an ugly attempt would be to impose requirements such as this. Think of  $A_1$  being the event that the first flip was heads.  $A_2$  is the event of that the second flip was heads.  $A_3$ , the third flip, was heads, and so on.

Here is an event whose occurrence is not determined by the first three coin flips. And here's an event whose occurrence or not is determined by the fifth and sixth coin flip. If we think physically that all those coin flips have nothing to do with each other, information about the fifth and sixth coin flip are not going to change what we expect from the first three. So the probability of this event, the conditional probability, should be the same as the unconditional probability. And we would like a relation of this kind to be true, no matter what kind of formula you write down, as long as the events that show up here are different from the events that show up there.

OK. That's sort of an ugly definition. The mathematical definition that actually does the job, and leads to all the formulas of this kind, is the following. We're going to say that the collection of events are independent if we can find the probability of their joint occurrence by just multiplying probabilities. And that will be true even if you look at sub-collections of these events.

Let's make that more precise. If we have three events, the definition tells us that the three events are independent if the following are true. Probability  $A_1$  and  $A_2$  and  $A_3$ , you can calculate this probability by multiplying individual probabilities. But the same is true even if you take fewer events. Just a few indices out of the indices that we have available. So we also require  $P(A_1 \text{ intersection } A_2)$  is  $P(A_1)$  times  $P(A_2)$ . And similarly for the other possibilities of choosing the indices.

OK, so independence, mathematical definition, requires that calculating probabilities of any intersection of the events we have in our hands, that calculation can be done by just multiplying individual probabilities. And this has to apply to the case where we consider all of the events in our hands or just sub-collections of those events.

Now these relations just by themselves are called pairwise independence. So this relation, for example, tells us that  $A_1$  is independent from  $A_2$ . This tells us that  $A_2$  is independent from  $A_3$ . This will tell us that  $A_1$  is independent from  $A_3$ . But independence of all the events together actually requires a little more. One more equality that has to do with all three events being considered at the same time.

And this extra equality is not redundant. It actually does make a difference. Independence and pairwise independence are different things. So let's illustrate the situation with an example. Suppose we have two coin flips. The coin tosses are independent, so the bias is  $1/2$ , so all possible outcomes have a probability of  $1/2$  times  $1/2$ , which is  $1/4$ .

And let's consider now a bunch of different events. One event is that the first toss is heads. This is this blue set here. Another event is the second toss is heads. And this is this black event here.

OK. Are these two events independent? If you check it mathematically, yes. Probability of A is  $1/2$ . Probability of B is  $1/2$ . Probability of A times probability of B is  $1/4$ , which is the same as the probability of A intersection B, which is this set. So we have just checked mathematically that A and B are independent.

Now let's consider a third event which is that the first and second toss give the same result. I'll use a different color. First and second toss to give the same result. This is the event that we obtain heads, heads or tails, tails. So this is the probability of C. What's the probability of C? Well, C is made up of two outcomes, each one of which has probability  $1/4$ , so the probability of C is  $1/2$ . What is the probability of C intersection A? C intersection A is just this one outcome, and has probability  $1/4$ .

What's the probability of A intersection B intersection C? The three events intersect just this outcome, so this probability is also  $1/4$ .

OK. What's the probability of C given A and B?

If A has occurred, and B has occurred, you are certain that this outcome here happened. If the first toss is H and the second toss is H, then you're certain of the first and second toss gave the same result. So the conditional probability of C given A and B is equal to 1.

So do we have independence in this example? We don't. C, that we obtain the same result in the first and the second toss, has probability  $1/2$ . Half of the possible outcomes give us two coin flips with the same result-- heads, heads or tails, tails. So the probability of C is  $1/2$ .

But if I tell you that the events A and B both occurred, then you're certain that C occurred. If I tell you that we had heads and heads, then you're certain the outcomes were the same. So the conditional probability is different from the unconditional probability. So by combining these two relations together, we get that the three events are not independent.

But are they pairwise independent? Is A independent from B? Yes, because probability of A times probability of B is  $1/4$ , which is probability of A intersection B. Is C independent from A? Well, the probability of C and A is  $1/4$ . The probability of C is  $1/2$ . The probability of A is  $1/2$ . So it checks.  $1/4$  is equal to  $1/2$  and  $1/2$ , so event C and event A are independent.

Knowing that the first toss was heads does not change your beliefs about whether the two tosses are going to have the same outcome or not. Knowing that the first was heads, well, the second is equally likely to be heads or tails. So event C has just the same probability, again,  $1/2$ , to occur.

To put it the opposite way, if I tell you that the two results were the same-- so it's either heads, heads or tails, tails-- what does that tell you about the first toss? Is it heads, or is it tails? Well, it doesn't tell you anything. It could be either over the two, so the probability of heads in the first toss is equal to  $1/2$ , and telling you C occurred does not change anything.

So this is an example that illustrates the case where we have three events in which we check that pairwise independence holds for any combination of two of these events. We have the probability of their intersection is equal to the product of their probabilities. On the other hand, the three events taken all together are not independent. A doesn't tell me anything useful, whether C is going to occur or not. B doesn't tell me anything useful. But if I tell you that both A and B occurred, the two of them together tell me something useful about C. Namely, they tell me that C certainly has occurred.

Very good. So independence is this somewhat subtle concept. Once you grasp the intuition of what it really means, then things perhaps fall in place. But it's a concept where it's easy to get some misunderstanding. So just take some time to digest.

So to lighten things up, I'm going to spend the remaining four minutes talking about the very nice, simple problem that involves conditional probabilities and the like. So here's the problem, formulated exactly as it shows up in various textbooks. And the formulation says the following.

Well, consider one of those anachronistic places where they still have kings or queens, and where actually boys take precedence over girls. So if there is a boy-- if the royal family has a boy, then he will become the king even if he has an older sister who might be the queen.

So we have one of those royal families. That royal family had two children, and we know that there is a king. There is a king, which means that at least one of the two children was a boy. Otherwise we wouldn't have a king. What is the probability that the king's sibling is female?

OK. I guess we need to make some assumptions about genetics. Let's assume that every child is a boy or a girl with probability  $1/2$ , and that different children, what they are is independent from what the other children were. So every childbirth is basically a coin flip.

OK, so if you take that, you say, well, the king is a child. His sibling is another child. Children are independent of each other. So the probability that the sibling is a girl is  $1/2$ . That's the naive answer. Now let's try to do it formally.

Let's set up a model of the experiment. The royal family had two children, as we're told, so there's four outcomes-- boy boy, boy girl, girl boy, and girl girl. Now, we are told that there is a king, which means what? This outcome here did not happen. It is not possible. There are three outcomes that remain possible. So this is our conditional sample space given that there is king.

What are the probabilities for the original model? Well with the model that we assume that every child is a boy or a girl independently with probability  $1/2$ , then the four outcomes would be equally likely, and they're like this. These are the original probabilities. But once we are told that

this outcome did not happen, because we have a king, then we are transported to the smaller sample space.

In this sample space, what's the probability that the sibling is a girl? Well the sibling is a girl in two out of the three outcomes. So the probability that the sibling is a girl is actually  $2/3$ . So that's supposed to be the right answer. Maybe a little counter-intuitive.

So you can play smart and say, oh I understand such problems better than you, here is a trick problem and here's why the answer is  $2/3$ . But actually I'm not fully justified in saying that the answer is  $2/3$ . I made lots of hidden assumptions when I put this model down, which I didn't yet state. So to reverse engineer this answer, let's actually think what's the probability model for which this would have been the right answer. And here's the probability model.

The royal family-- the royal parents decided to have exactly two children. They went and had them. It turned out that at least one was a boy and became a king. Under this scenario-- that they decide to have exactly two children-- then this is the big sample space. It turned out that one was a boy. That eliminates this outcome. And then this picture is correct and this is the right answer.

But there's hidden assumptions being there. How about if the royal family had followed the following strategy? We're going to have children until we get a boy, so that we get a king, and then we'll stop. OK, given they have two children, what's the probability that the sibling is a girl?

It's 1. The reason that they had two children was because the first was a girl, so they had to have a second. So assumptions about reproductive practices actually need to come in, and they're going to affect the decisions. Or, if it's one of those ancient kingdoms where a king would always make sure to strangle any of his brothers, then the probability that the sibling is a girl is actually 1 again, and so on.

So it means that one needs to be careful when you start with loosely worded problems to make sure exactly what it means and what assumptions you're making. All right, see you next week.

MIT OpenCourseWare  
<http://ocw.mit.edu>

6.041SC Probabilistic Systems Analysis and Applied Probability  
Fall 2013

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.