PROFESSOR: We just saw some random variables come up in the bigger number game. And we're going to be talking now about random variables, just formally what they are and their definition of independence for random variables. But let's begin by looking at the informal idea.

Again, a random variable is a number that's produced by a random process. So a typical example that comes up where you get a random variable is you've got some system that you're watching and you're going to time it to see when the next crash comes, if it crashes. So assuming that this is unpredictable that it happens in some random way, then the number of hours from the present until the next time the system crashes is a number that's produced by this random process of whether the system works or not.

Number of faulty pixels in a monitor. When you're building the monitors and delivering them to the actual computer manufacturers, there's a certain probability that some of the millions of pixels in the monitor are going to be faulty. And you could think of that number of pixels is also produced from an unpredictable randomness in the manufacturing process.

One that really is modeled in physics as random is when you have a Geiger counter, you're measuring alpha particles. The number of alpha particles that are detected by a given Geiger counter in a second is believed to be a random number. There's a distribution that it has but the number of alpha particles is not always the same from second to second, and so it's a random variable.

And finally, we'll look at the standard abstract example of flipping coins. And if I flip coins then the number of heads in a given number of flips-- let's say I flip a coin n times. The number of heads will be another rather standard random variable.

OK what is abstractly a random variable? Oops, I'm getting ahead of myself again. Let's look at that example of three fair coins. So each coin has a probability of being heads that's a half and tails being a half. I'm going to flip the three of them.

And I'm going to assume that they're distinguishable. So there's a first coin, a second coin, and a third coin. Or alternatively you could think of flipping the same coin three times.

So the number of heads is a number that comes out of this random process of flipping the three coins. So it's a number that's either from 0 to 3. There could be no heads or all heads.

So it is a basic example of a random variable where you're producing this integer based on how the coins flip. Another one is simply a [? 0-1 ?] valued random variable where it signals 1 if all 3 coins match in what they come up with, and 0 if they don't match.

Now once I have these random variables defined, one of the things that's a convenient use of random variables is to use them to define various kinds of events. So the event that C equals 1, that's an event that-- it's a set of outcomes where the count is 1 and it has a certain probability. This is the event of exactly 1 head. There are 3 possible outcomes among the 8 outcomes of heads and tails with 3 coins. So it has probability 3/8.

I could also just talk about the outcome that C is greater than or equal to 1. Well C is greater than or equal to 1 when there is at least 1 head. Or put another way, the only time that C is not greater than or equal to 1 is when you have all tails. So there's a 7/8 chance, 7 out of 8 outcomes involve 1 or more heads. So the probability that C greater than or equal to 1 is 7/8.

Here's a weirder one. I can use the two variables C and M to define an event. What's the probability that C times M is greater than 0? Well since C and M are both non-negative variables, the probability that their product is greater than 0 is equal to the probability that each of them is greater than 0.

OK, what does it mean that M is greater than 0 and C is greater than 0? Well it says there's at least 1 head-- that's what C greater than 0 means. And M greater than 0 means all the coins match. This is an obscure way of describing the event all heads, and it has a course probability 1/8.

Now we come to the formal definition. So formally, a random variable is simply a function that maps outcomes in the sample space to numbers. We think of the outcomes in the sample space as the results of a random experiment. They are an outcome and they have a probability.

And when the outcome is translated into a real number that you think of as being produced as a result of that outcome, that's what the random variable does. So formally, a random variable is not a variable. Or it's a function that maps the sample space to the real numbers. And it's got to be total, by the way. It's a total function.

Usually this would be a real valued random variable. Usually it's the real numbers. Might be a subset of the real numbers like the integer valued random variables. Occasionally we'll use complex valued random variables. Actually, that happens in physics a good deal in quantum mechanics, but not for our purposes. We're just going to mean real value from now on when we talk about random variables.

So abstractly or intuitively what the random variable is doing really is it just packaging together in one object R, the random variable, a whole bunch of events that are defined by the value that R takes. So for every possible real number, if I look at the event that R is equal to a, that's an interesting event. And it's one of the basic events that R puts together. And if you knew the answer to all of these R equals a's, then you really know a lot about R.

And with this understanding that R is a package of events of the form R is equal to a, then a lot of the event properties carry right over to random variables directly. That's why this little topic of introducing random variables is also about independence because the definition of independence carries right over. Namely, a bunch of random variables are mutually independent if the events that they define are all mutually independent. So if and only if the events that are-- each event defined by R1 and R2 and through Rn, that set of events are mutually independent no matter what the values are chosen that we decide to look at for R1 and R2 through Rn.

And of course there's an alternative way we could always express independent events in terms of products instead of conditional probabilities. So we could say-- or instead of invoking the idea of mutual independence we could say explicitly where it comes from as an equation. It means that the probability that R1 is equal to a1 and R2 is equal to a1 and Rn is equal to an is equal to the product of the probabilities-- of the individual probabilities-- that R1 is a1 times the probability of R2 is a2. And the definition then of mutual independence of the random variables R1 through n, Rn holds is that this equation it holds for all possible values, little a1 through little an.

So let's just practice. Are the variables C, which is the count of the number of heads when you flip three coins, and M, [? the 0-1 ?] valued random variable that tells you whether there's a match, are they independent? Well certainly not, because there's definitely a positive probability that the count will be 1 that you'll get at least a head.

And there's a positive probability that they all will match. It's the probability of a quarter. So the product of those 2 is positive, but of course the probability that you match and you'll have exactly 1 head is 0 because if you have exactly 1 head you must have 2 tails and there's no match. So without thinking very hard about what the probabilities are we can immediately see that the product is not equal to the probability of the conjunction or the and, and therefore they're not independent.

Well here's one that's a little bit more interesting. In order to explain it I've got to set up the idea of an indicator variable, which itself is a very important concept. So if I have an event A, I can package A into a random variable. Just like the match random variable was really packaging the event that the coins matched into a [? 0-1 ?] valued variable, I'm going to define the indicator variable for any event A to be 1 if A occurs and 0 if A does not occur.

So now I'm able to capture everything that matters about event A by the random variable IA. If I have IA I know what A is, and if I have A I know what IA is. And it means that really I can think of events as special cases of random variables.

Now when you do this you need a sanity check. Because remember we've defined independence of random variables one way. I mean it's a concept of independence that holds for random variables. We have another concept of independence that holds for events.

Now the definition for random variable was motivated by the definition for events but it's a different definition of independence of different kinds of objects. Now if this correspondence between events and indicator variables is going to make sense and not confuse us it should be the case that two events are independent if and only if their indicator variables are independent. That is, IA and IB are independent if and only if the events A and B are independent.

And this is a lovely little exercise. It's like a three-line proof for you to verify. I'm not going bother to do it on the slide because it's good practice. So this would be a moment to stop and verify that using the two definitions of independence, the definition of what it means for IA and IB to be independent as random variables and comparing that to the definition of what it means for A and B to be independent as events, they match.

If we look at the event of an odd number of heads we can ask now whether the event M, which is the indicator variable for a match-- the random variable M-- and the indicator variable IO are dependent or not. Now both of these depend on all the three coins. IO is looking at all 3 coins to see if there are an odd number of heads, M is looking at all 3 coins to see if they're all heads or all tails. And it's not clear with all that common basis for returning what value they have. It's not immediately obvious that they're independent, but as a matter of fact they are.

And again this is absolutely something that you should check out. If you don't stop the video now to work it out, you should definitely do it afterward. It's an important little exercise and it's easy to check. All you have to do is check that the probabilities of the event of odd number of heads in the event all match are independent as events. Or you could use the random variable definition and check that these two random variables were independent by checking 4 equations because this can have values 0 and 1. And this can have value 0 and 1.

Remember with independent events we had the idea that if A was independent of B it really meant that A was independent of everything about B. In particular it was independent of the complement of B as well. And a similar property holds for random variables.

So intuitively if R is independent of S then R is really independent of any information at all that you have about S. And that can be made more precise that R is independent of any information about S by saying pick an arbitrary function that maps R to R, total function. So what I can do is think of f as giving me some information about the value of S. So if R is independent of S then in fact R is independent of f of S, any transformation of S by a fixed non-random function.

And of course the notion of k-way independence carries right over from the event case. If I have k random-- if I have a bunch of random variables, a large number much more than k, they're k-way independent if every set of k of them are mutually independent. And of course as with events we use the 2-way case to call them pairwise independent.

Again, we saw an example of this in terms of events already but we can rephrase it now in terms of indicator variables. If we let Hi be the indicator variable for a head on a flip i-- of the i flip of a coin-- where i ranges from 1 through k, if we have k coins and Hi is the indicator variable for how coin I came out, whether or not there's a head, now O can be nicely expressed. The notion that there's an odd number of heads is simply the mod 2 sum of the Hi's. And this by the way, is a trick that we'll be using regularly that events now can be defined rather nicely in terms of doing operations on the arithmetic values of indicator variables.

So O is nothing but the mod 2 sum of the values of the indicator variables Hi from 1 to k. And what we saw when we were working with their event version is that any k of these events are independent. I've got k plus 1. There's k Hi's and there's O, which makes the k plus 1-- k plus first. [AUDIO OUT] And the reason why any k of them were independent was discussed in the previous slide when we were looking at the events of there being an odd number of heads and a head coming up on the i flip.

The reason why pairwise independence gets singled out is that we'll see that for a bunch of major applications this pairwise independence is sufficient and rather than verifying mutual independence. It's harder to check mutual independence. You've got a lot more equations to check.

And in fact it often doesn't hold in circumstances where pairwise does hold. So this is good to know. We'll be making use of it in an application later when we look at sampling and the law of large numbers.