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.

PROFESSOR: So by now you have seen pretty much every possible trick there is in basic probability theory, about how to calculate distributions, and so on. You have the basic tools to do pretty much anything. So what's coming after this?

Well, probability is useful for developing the science of inference, and this is a subject to which we're going to come back at the end of the semester. Another chapter, which is what we will be doing over the next few weeks, is to deal with phenomena that evolve in time. So so-called random processes or stochastic processes. So what is this about?

So in the real world, you don't just throw two random variables and go home. Rather the world goes on. So you generate the random variable, then you get more random variables, and things evolve in time. And random processes are supposed to be models that capture the evolution of random phenomena over time. So that's what we will be doing.

Now when we have evolution in time, mathematically speaking, you can use discrete time or continuous time. Of course, discrete time is easier. And that's where we're going to start from. And we're going to start from the easiest, simplest random process, which is the so-called Bernoulli process, which is nothing but just a sequence of coin flips. You keep flipping a coin and keep going forever. That's what the Bernoulli process is.

So in some sense it's something that you have already seen. But we're going to introduce a few additional ideas here that will be useful and relevant as we go along and we move on to continuous time processes. So we're going to define the Bernoulli process, talk about some basic properties that the process has, and derive a few formulas, and exploit the special structure that it has to do a few quite interesting things.

By the way, where does the word Bernoulli come from? Well the Bernoulli's were a family of mathematicians, Swiss mathematicians and scientists around the 1700s. There were so many of them that actually-- and some of them had the same first name-- historians even have difficulty of figuring out who exactly did what. But in any case, you can imagine that at the

dinner table they were probably flipping coins and doing Bernoulli trials. So maybe that was their pass-time.

OK. So what is the Bernoulli process? The Bernoulli process is nothing but a sequence of independent Bernoulli trials that you can think of as coin flips. So you can think the result of each trial being heads or tails. It's a little more convenient maybe to talk about successes and failures instead of heads or tails. Or if you wish numerical values, to use a 1 for a success and 0 for a failure. So the model is that each one of these trials has the same probability of success, p. And the other assumption is that these trials are statistically independent of each other.

So what could be some examples of Bernoulli trials? You buy a lottery ticket every week and you win or lose. Presumably, these are independent of each other. And if it's the same kind of lottery, the probability of winning should be the same during every week.

Maybe you want to model the financial markets. And a crude model could be that on any given day the Dow Jones is going to go up or down with a certain probability. Well that probability must be somewhere around 0.5, or so. This is a crude model of financial markets. You say, probably there is more into them. Life is not that simple. But actually it's a pretty reasonable model. It takes quite a bit of work to come up with more sophisticated models that can do better predictions than just pure heads and tails.

Now more interesting, perhaps to the examples we will be dealing with in this class-- a Bernoulli process is a good model for streams of arrivals of any kind to a facility. So it could be a bank, and you are sitting at the door of the bank. And at every second, you check whether a customer came in during that second or not. Or you can think about arrivals of jobs to a server. Or any other kind of requests to a service system.

So requests, or jobs, arrive at random times. You split the time into time slots. And during each time slot something comes or something does not come. And for many applications, it's a reasonable assumption to make that arrivals on any given slot are independent of arrivals in any other time slot. So each time slot can be viewed as a trial, where either something comes or doesn't come. And different trials are independent of each other.

Now there's two assumptions that we're making here. One is the independence assumption. The other is that this number, p, probability of success, is constant. Now if you think about the bank example, if you stand outside the bank at 9:30 in the morning, you'll see arrivals

happening at a certain rate. If you stand outside the bank at 12:00 noon, probably arrivals are more frequent. Which means that the given time slot has a higher probability of seeing an arrival around noon time.

This means that the assumption of a constant p is probably not correct in that setting, if you're talking about the whole day. So the probability of successes or arrivals in the morning is going to be smaller than what it would be at noon. But if you're talking about a time period, let's say 10:00 to 10:15, probably all slots have the same probability of seeing an arrival and it's a good approximation. So we're going to stick with the assumption that p is constant, doesn't change with time.

Now that we have our model what do we do with it? Well, we start talking about the statistical properties that it has. And here there's two slightly different perspectives of thinking about what a random process is. The simplest version is to think about the random process as being just a sequence of random variables.

We know what random variables are. We know what multiple random variables are. So it's just an experiment that has associated with it a bunch of random variables. So once you have random variables, what do you do instinctively? You talk about the distribution of these random variables. We already specified for the Bernoulli process that each Xi is a Bernoulli random variable, with probability of success equal to p.

That specifies the distribution of the random variable X, or Xt, for general time t. Then you can calculate expected values and variances, and so on. So the expected value is, with probability p, you get a 1. And with probability 1 - p, you get a 0. So the expected value is equal to p.

And then we have seen before a formula for the variance of the Bernoulli random variable, which is p times 1-p. So this way we basically now have all the statistical properties of the random variable Xt, and we have those properties for every t. Is this enough of a probabilistic description of a random process? Well, no. You need to know how the different random variables relate to each other.

If you're talking about a general random process, you would like to know things. For example, the joint distribution of X2, with X5, and X7. For example, that might be something that you're interested in. And the way you specify it is by giving the joint PMF of these random variables. And you have to do that for every collection, or any subset, of the random variables you are interested in.

So to have a complete description of a random processes, you need to specify for me all the possible joint distributions. And once you have all the possible joint distributions, then you can answer, in principle, any questions you might be interested in.

How did we get around this issue for the Bernoulli process? I didn't give you the joint distributions explicitly. But I gave them to you implicitly. And this is because I told you that the different random variables are independent of each other. So at least for the Bernoulli process, where we make the independence assumption, we know that this is going to be the product of the PMFs. And since I have told you what the individual PMFs are, this means that you automatically know all the joint PMFs. And we can go to business based on that.

All right. So this is one view of what a random process is, just a collection of random variables. There's another view that's a little more abstract, which is the following. The entire process is to be thought of as one long experiment. So we go back to the chapter one view of probabilistic models. So there must be a sample space involved. What is the sample space?

If I do my infinite, long experiment of flipping an infinite number of coins, a typical outcome of the experiment would be a sequence of 0's and 1's. So this could be one possible outcome of the experiment, just an infinite sequence of 0's and 1's. My sample space is the set of all possible outcomes of this kind. Here's another possible outcome, and so on. And essentially we're dealing with a sample space, which is the space of all sequences of 0's and 1's. And we're making some sort of probabilistic assumption about what may happen in that experiment.

So one particular sequence that we may be interested in is the sequence of obtaining all 1's. So this is the sequence that gives you 1's forever. Once you take the point of view that this is our sample space-- its the space of all infinite sequences-- you can start asking questions that have to do with infinite sequences. Such as the question, what's the probability of obtaining the infinite sequence that consists of all 1's? So what is this probability? Let's see how we could calculate it.

So the probability of obtaining all 1's is certainly less than or equal to the probability of obtaining 1's, just in the first 10 tosses. OK. This is asking for more things to happen than this. If this event is true, then this is also true. Therefore the probability of this is smaller than the probability of that. This event is contained in that event. This implies this. So we have this

inequality.

Now what's the probability of obtaining 1's in 10 trials? This is just p to the 10th because the trials are independent. Now of course there's no reason why I chose 10 here. The same argument goes through if I use an arbitrary number, k. And this has to be true for all k.

So this probability is less than p to the k, no matter what k I choose. Therefore, this must be less than or equal to the limit of this, as k goes to infinity. This is smaller than that for all k's. Let k go to infinity, take k arbitrarily large, this number is going to become arbitrarily small. It goes to 0. And that proves that the probability of an infinite sequence of 1's is equal to 0.

So take limits of both sides. It's going to be less than or equal to the limit-- I shouldn't take a limit here. The probability is less than or equal to the limit of p to the k, as k goes to infinity, which is 0. So this proves in a formal way that the sequence of all 1's has 0 probability.

If you have an infinite number of coin flips, what's the probability that all of the coin flips result in heads? The probability of this happening is equal to zero. So this particular sequence has 0 probability. Of course, I'm assuming here that p is less than 1, strictly less than 1.

Now the interesting thing is that if you look at any other infinite sequence, and you try to calculate the probability of that infinite sequence, you would get a product of (1-p) times 1, 1-p times 1, 1-p, times p times p, times 1-p and so on. You keep multiplying numbers that are less than 1.

Again, I'm making the assumption that p is between 0 and 1. So 1-p is less than 1, p is less than 1. You keep multiplying numbers less than 1. If you multiply infinitely many such numbers, the infinite product becomes 0. So any individual sequence in this sample space actually has 0 probability. And that is a little bit counter-intuitive perhaps. But the situation is more like the situation where we deal with continuous random variables.

So if you could draw a continuous random variable, every possible outcome has 0 probability. And that's fine. But all of the outcomes collectively still have positive probability. So the situation here is very much similar. So the space of infinite sequences of 0's and 1's, that sample space is very much like a continuous space.

If you want to push that analogy further, you could think of this as the expansion of a real number. Or the representation of a real number in binary. Take a real number, write it down in binary, you are going to get an infinite sequence of 0's and 1's. So you can think of each

possible outcome here essentially as a real number.

So the experiment of doing an infinite number of coin flips is sort of similar to the experiment of picking a real number at random. When you pick real numbers at random, any particular real number has 0 probability. So similarly here, any particular infinite sequence has 0 probability.

So if we were to push that analogy further, there would be a few interesting things we could do. But we will not push it further. This is just to give you an indication that things can get pretty subtle and interesting once you start talking about random processes that involve forever, over the infinite time horizon. So things get interesting even in this context of the simple Bernoulli process.

Just to give you a preview of what's coming further, today we're going to talk just about the Bernoulli process. And you should make sure before the next lecture-- I guess between the exam and the next lecture-- to understand everything we do today. Because next time we're going to do everything once more, but in continuous time. And in continuous time, things become more subtle and a little more difficult. But we are going to build on what we understand for the discrete time case.

Now both the Bernoulli process and its continuous time analog has a property that we call memorylessness, whatever happened in the past does not affect the future. Later on in this class we're going to talk about more general random processes, so-called Markov chains, in which there are certain dependences across time. That is, what has happened in the past will have some bearing on what may happen in the future.

So it's like having coin flips where the outcome of the next coin flip has some dependence on the previous coin flip. And that gives us a richer class of models. And once we get there, essentially we will have covered all possible models.

So for random processes that are practically useful and which you can manipulate, Markov chains are a pretty general class of models. And almost any real world phenomenon that evolves in time can be approximately modeled using Markov chains. So even though this is a first class in probability, we will get pretty far in that direction.

All right. So now let's start doing a few calculations and answer some questions about the Bernoulli process. So again, the best way to think in terms of models that correspond to the Bernoulli process is in terms of arrivals of jobs to a facility. And there's two types of questions

that you can ask. In a given amount of time, how many jobs arrived? Or conversely, for a given number of jobs, how much time did it take for them to arrive?

So we're going to deal with these two questions, starting with the first. For a given amount of time-- that is, for a given number of time periods-- how many arrivals have we had? How many of those Xi's happen to be 1's?

We fix the number of time slots-- let's say n time slots-- and you measure the number of successes. Well this is a very familiar random variable. The number of successes in n independent coin flips-- or in n independent trials-- is a binomial random variable. So we know its distribution is given by the binomial PMF, and it's just this, for k going from 0 up to n.

And we know everything by now about this random variable. We know its expected value is n times p. And we know the variance, which is n times p, times 1-p. So there's nothing new here. That's the easy part.

So now let's look at the opposite kind of question. Instead of fixing the time and asking how many arrivals, now let us fix the number of arrivals and ask how much time did it take. And let's start with the time until the first arrival.

So the process starts. We got our slots. And we see, perhaps, a sequence of 0's and then at some point we get a 1. The number of trials it took until we get a 1, we're going to call it T1. And it's the time of the first arrival. OK. What is the probability distribution of T1? What kind of random variable is it?

We've gone through this before. The event that the first arrival happens at time little t is the event that the first t-1 trials were failures, and the trial number t happens to be a success. So for the first success to happen at time slot number 5, it means that the first 4 slots had failures and the 5th slot had a success.

So the probability of this happening is the probability of having failures in the first t -1 trials, and having a success at trial number 1. And this is the formula for t equal 1,2, and so on. So we know what this distribution is. It's the so-called geometric distribution.

Let me jump this through this for a minute. In the past, we did calculate the expected value of the geometric distribution, and it's 1/p. Which means that if p is small, you expect to take a long time until the first success. And then there's a formula also for the variance of T1, which

we never formally derived in class, but it was in your textbook and it just happens to be this.

All right. So nothing new until this point. Now, let's talk about this property, the memorylessness property. We kind of touched on this property when we discussed-- when we did the derivation in class of the expected value of T1.

Now what is the memoryless property? It's essentially a consequence of independence. If I tell you the results of my coin flips up to a certain time, this, because of independence, doesn't give you any information about the coin flips after that time. So knowing that we had lots of 0's here does not change what I believe about the future coin flips, because the future coin flips are going to be just independent coin flips with a given probability, p, for obtaining tails.

So this is a statement that I made about a specific time. That is, you do coin flips until 12 o'clock. And then at 12 o'clock, you start watching. No matter what happens before 12 o'clock, after 12:00, what you're going to see is just a sequence of independent Bernoulli trials with the same probability, p. Whatever happened in the past is irrelevant.

Now instead of talking about the fixed time at which you start watching, let's think about a situation where your sister sits in the next room, flips the coins until she observes the first success, and then calls you inside. And you start watching after this time. What are you're going to see?

Well, you're going to see a coin flip with probability p of success. You're going to see another trial that has probability p as a success, and these are all independent of each other. So what you're going to see starting at that time is going to be just a sequence of independent Bernoulli trials, as if the process was starting at this time. How long it took for the first success to occur doesn't have any bearing on what is going to happen afterwards. What happens afterwards is still a sequence of independent coin flips.

And this story is actually even more general. So your sister watches the coin flips and at some point tells you, oh, something really interesting is happening here. I got this string of a hundred 1's in a row. Come and watch. Now when you go in there and you start watching, do you expect to see something unusual?

There were unusual things that happened before you were called in. Does this means that you're going to see unusual things afterwards? No. Afterwards, what you're going to see is, again, just a sequence of independent coin flips. The fact that some strange things happened

before doesn't have any bearing as to what is going to happen in the future.

So if the roulettes in the casino are properly made, the fact that there were 3 reds in a row doesn't affect the odds of whether in the next roll it's going to be a red or a black. So whatever happens in the past-- no matter how unusual it is-- at the time when you're called in, what's going to happen in the future is going to be just independent Bernoulli trials, with the same probability, p.

The only case where this story changes is if your sister has a little bit of foresight. So your sister can look ahead into the future and knows that the next 10 coin flips will be heads, and calls you before those 10 flips will happen. If she calls you in, then what are you going to see? You're not going to see independent Bernoulli trials, since she has psychic powers and she knows that the next ones would be 1's. She called you in and you will see a sequence of 1's. So it's no more independent Bernoulli trials.

So what's the subtle difference here? The future is independent from the past, provided that the time that you are called and asked to start watching is determined by someone who doesn't have any foresight, who cannot see the future. If you are called in, just on the basis of what has happened so far, then you don't have any information about the future.

And one special case is the picture here. You have your coin flips. Once you see a one that happens, once you see a success, you are called in. You are called in on the basis of what happened in the past, but without any foresight.

OK. And this subtle distinction is what's going to make our next example interesting and subtle. So here's the question. You buy a lottery ticket every day, so we have a Bernoulli process that's running in time. And you're interested in the length of the first string of losing days. What does that mean?

So suppose that a typical sequence of events could be this one. So what are we discussing here? We're looking at the first string of losing days, where losing days means 0's. So the string of losing days is this string here. Let's call the length of that string, L. We're interested in the random variable, which is the length of this interval. What kind of random variable is it?

OK. Here's one possible way you might think about the problem. OK. Starting from this time, and looking until this time here, what are we looking at? We're looking at the time, starting from here, until the first success. So the past doesn't matter. Starting from here we have coin

flips until the first success. The time until the first success in a Bernoulli process-- we just discussed that it's a geometric random variable.

So your first conjecture would be that this random variable here, which is 1 longer than the one we are interested in, that perhaps is a geometric random variable. And if this were so, then you could say that the random variable, L, is a geometric, minus 1. Can that be the correct answer?

A geometric random variable, what values does it take? It takes values 1, 2, 3, and so on. 1 minus a geometric would take values from 0, 1, 2, and so on. Can the random variable L be 0? No. The random variable L is the length of a string of losing days. So the shortest that L could be, would be just 1. If you get just one losing day and then you start winning, L would be equal to 1.

So L cannot be 0 by definition, which means that L + 1 cannot be 1, by definition. But if L +1 were geometric, it could be equal to 1. Therefore this random variable, L + 1, is not a geometric.

OK. Why is it not geometric? I started watching at this time. From this time until the first success, that should be a geometric random variable. Where's the catch? If I'm asked to start watching at this time, it's because my sister knows that the next one was a failure. This is the time where the string of failures starts. In order to know that they should start watching here, it's the same as if I'm told that the next one is a failure.

So to be asked to start watching at this time requires that someone looked in the future. And in that case, it's no longer true that these will be independent Bernoulli trials. In fact, they're not. If you start watching here, you're certain that the next one is a failure. The next one is not an independent Bernoulli trial. That's why the argument that would claim that this L + 1 is geometric would be incorrect.

So if this is not the correct answer, which is the correct answer? The correct answer goes as follows. Your sister is watching. Your sister sees the first failure, and then tells you, OK, the failures-- or losing days-- have started. Come in and watch. So you start to watching at this time. And you start watching until the first success comes. This will be a geometric random variable.

So from here to here, this will be geometric. So things happen. You are asked to start

watching. After you start watching, the future is just a sequence of independent Bernoulli trials. And the time until the first failure occurs, this is going to be a geometric random variable with parameter p. And then you notice that the interval of interest is exactly the same as the length of this interval. This starts one time step later, and ends one time step later. So conclusion is that L is actually geometric, with parameter p.

OK, it looks like I'm missing one slide. Can I cheat a little from here?

OK. So now that we dealt with the time until the first arrival, we can start talking about the time until the second arrival, and so on. How do we define these? After the first arrival happens, we're going to have a sequence of time slots with no arrivals, and then the next arrival is going to happen.

So we call this time that elapses-- or number of time slots after the first arrival until the next one-- we call it T2. This is the second inter-arrival time, that is, time between arrivals. Once this arrival has happened, then we wait and see how many more it takes until the third arrival. And we call this time here, T3.

We're interested in the time of the k-th arrival, which is going to be just the sum of the first k inter-arrival times. So for example, let's say Y3 is the time that the third arrival comes. Y3 is just the sum of T1, plus T2, plus T3. So we're interested in this random variable, Y3, and it's the sum of inter-arrival times. To understand what kind of random variable it is, I guess we should understand what kind of random variables these are going to be.

So what kind of random variable is T2? Your sister is doing her coin flips until a success is observed for the first time. Based on that information about what has happened so far, you are called into the room. And you start watching until a success is observed again.

So after you start watching, what you have is just a sequence of independent Bernoulli trials. So each one of these has probability p of being a success. The time it's going to take until the first success, this number, T2, is going to be again just another geometric random variable. It's as if the process just started. After you are called into the room, you have no foresight, you don't have any information about the future, other than the fact that these are going to be independent Bernoulli trials.

So T2 itself is going to be geometric with the same parameter p. And then you can continue the arguments and argue that T3 is also geometric with the same parameter p. Furthermore,

whatever happened, how long it took until you were called in, it doesn't change the statistics about what's going to happen in the future. So whatever happens in the future is independent from the past. So T1, T2, and T3 are independent random variables.

So conclusion is that the time until the third arrival is the sum of 3 independent geometric random variables, with the same parameter. And this is true more generally. The time until the k-th arrival is going to be the sum of k independent random variables. So in general, Yk is going to be T1 plus Tk, where the Ti's are geometric, with the same parameter p, and independent.

So now what's more natural than trying to find the distribution of the random variable Yk? How can we find it? So I fixed k for you. Let's say k is 100. I'm interested in how long it takes until 100 customers arrive. How can we find the distribution of Yk?

Well one way of doing it is to use this lovely convolution formula. Take a geometric, convolve it with another geometric, you get something. Take that something that you got, convolve it with a geometric once more, do this 99 times, and this gives you the distribution of Yk. So that's definitely doable, and it's extremely tedious.

Let's try to find the distribution of Yk using a shortcut. So the probability that Yk is equal to t. So we're trying to find the PMF of Yk. k has been fixed for us. And we want to calculate this probability for the various values of t, because this is going to give us the PMF of Yk. OK.

What is this event? What does it take for the k-th arrival to be at time t? For that to happen, we need two things. In the first t -1 slots, how many arrivals should we have gotten? k - 1. And then in the last slot, we get one more arrival, and that's the k-th one. So this is the probability that we have k - 1 arrivals in the time interval from 1 up to t. And then, an arrival at time t.

That's the only way that it can happen, that the k-th arrival happens at time t. We need to have an arrival at time t. And before that time, we need to have exactly k - 1 arrivals. Now this is an event that refers-- t-1. In the previous time slots we had exactly k -1 arrivals. And then at the last time slot we get one more arrival.

Now the interesting thing is that this event here has to do with what happened from time 1 up to time t -1. This event has to do with what happened at time t. Different time slots are independent of each other. So this event and that event are independent.

So this means that we can multiply their probabilities. So take the probability of this. What is

that? Well probability of having a certain number of arrivals in a certain number of time slots, these are just the binomial probabilities.

So this is, out of t - 1 slots, to get exactly k - 1 arrivals, p to the k-1, (1-p) to the t-1 - (k-1), this gives us t-k. And then we multiply with this probability, the probability of an arrival, at time t is equal to p. And so this is the formula for the PMF of the number-- of the time it takes until the k-th arrival happens.

Does it agree with the formula in your handout? Or its not there? It's not there. OK.

Yeah. OK. So that's the formula and it is true for what values of t? [INAUDIBLE]. It takes at least k time slots in order to get k arrivals, so this formula should be true for k larger than or equal to t. For t larger than or equal to k.

All right. So this gives us the PMF of the random variable Yk. Of course, we may also be interested in the mean and variance of Yk. But this is a lot easier. Since Yk is the sum of independent random variables, the expected value of Yk is going to be just k times the expected value of your typical t.

So the expected value of Yk is going to be just k times 1/p, which is the mean of the geometric. And similarly for the variance, it's going to be k times the variance of a geometric. So we have everything there is to know about the distribution of how long it takes until the first arrival comes.

OK. Finally, let's do a few more things about the Bernoulli process. It's interesting to talk about several processes at the time. So in the situation here of splitting a Bernoulli process is where you have arrivals that come to a server. And that's a picture of which slots get arrivals. But actually maybe you have two servers. And whenever an arrival comes to the system, you flip a coin and with some probability, q, you send it to one server. And with probability 1-q, you send it to another server.

So there is a single arrival stream, but two possible servers. And whenever there's an arrival, you either send it here or you send it there. And each time you decide where you send it by flipping an independent coin that has its own bias q. The coin flips that decide where do you send it are assumed to be independent from the arrival process itself.

So there's two coin flips that are happening. At each time slot, there's a coin flip that decides

whether you have an arrival in this process here, and that coin flip is with parameter p. And if you have something that arrives, you flip another coin with probabilities q, and 1-q, that decides whether you send it up there or you send it down there.

So what kind of arrival process does this server see? At any given time slot, there's probability p that there's an arrival here. And there's a further probability q that this arrival gets sent up there. So the probability that this server sees an arrival at any given time is p times q. So this process here is going to be a Bernoulli process, but with a different parameter, p times q. And this one down here, with the same argument, is going to be Bernoulli with parameter p times (1-q).

So by taking a Bernoulli stream of arrivals and splitting it into two, you get two separate Bernoulli processes. This is going to be a Bernoulli process, that's going to be a Bernoulli process. Well actually, I'm running a little too fast. What does it take to verify that it's a Bernoulli process? At each time slot, it's a 0 or 1. And it's going to be a 1, you're going to see an arrival with probability p times q.

What else do we need to verify, to be able to tell-- to say that it's a Bernoulli process? We need to make sure that whatever happens in this process, in different time slots, are statistically independent from each other. Is that property true? For example, what happens in this time slot whether you got an arrival or not, is it independent from what happened at that time slot?

The answer is yes for the following reason. What happens in this time slot has to do with the coin flip associated with the original process at this time, and the coin flip that decides where to send things. What happens at that time slot has to do with the coin flip here, and the additional coin flip that decides where to send it if something came.

Now all these coin flips are independent of each other. The coin flips that determine whether we have an arrival here is independent from the coin flips that determined whether we had an arrival there. And you can generalize this argument and conclude that, indeed, every time slot here is independent from any other time slot. And this does make it a Bernoulli process.

And the reason is that, in the original process, every time slot is independent from every other time slot. And the additional assumption that the coin flips that we're using to decide where to send things, these are also independent of each other. So we're using here the basic property that functions of independent things remain independent.

There's a converse picture of this. Instead of taking one stream and splitting it into two streams, you can do the opposite. You could start from two streams of arrivals. Let's say you have arrivals of men and you have arrivals of women, but you don't care about gender. And the only thing you record is whether, in a given time slot, you had an arrival or not.

Notice that here we may have an arrival of a man and the arrival of a woman. We just record it with a 1, by saying there was an arrival. So in the merged process, we're not keeping track of how many arrivals we had total. We just record whether there was an arrival or not an arrival.

So an arrival gets recorded here if, and only if, one or both of these streams had an arrival. So that we call a merging of two Bernoull-- of two processes, of two arrival processes. So let's make the assumption that this arrival process is independent from that arrival process.

So what happens at the typical slot here? I'm going to see an arrival, unless none of these had an arrival. So the probability of an arrival in a typical time slot is going to be 1 minus the probability of no arrival. And the event of no arrival corresponds to the first process having no arrival, and the second process having no arrival. So there's no arrival in the merged process if, and only if, there's no arrival in the first process and no arrival in the second process.

We're assuming that the two processes are independent and that's why we can multiply probabilities here. And then you can take this formula and it simplifies to $p + q$, minus p times q.

So each time slot of the merged process has a certain probability of seeing an arrival. Is the merged process a Bernoulli process? Yes, it is after you verify the additional property that different slots are independent of each other.

Why are they independent? What happens in this slot has to do with that slot, and that slot down here. These two slots-- so what happens here, has to do with what happens here and there. What happens in this slot has to do with whatever happened here and there.

Now, whatever happens here and there is independent from whatever happens here and there. Therefore, what happens here is independent from what happens there. So the independence property is preserved. The different slots of this merged process are independent of each other. So the merged process is itself a Bernoulli process.

So please digest these two pictures of merging and splitting, because we're going to revisit

them in continuous time where things are little subtler than that. OK. Good luck on the exam and see you in a week.