PROFESSOR: So last time we started talking about random processes. A random process is a random experiment that evolves over time. And conceptually, it's important to realize that it's a single probabilistic experiment that has many stages. Actually, it has an infinite number of stages. And we discussed the simplest random process there is, the Bernoulli process, which is nothing but the sequence of Bernoulli trials-- an infinite sequence of Bernoulli trials. For example, flipping a coin over and over.

Once we understand what's going on with that process, then what we want is to move into a continuous time version of the Bernoulli process. And this is what we will call the Poisson process. And for the Poisson process, we're going to do exactly the same things that we did for the Bernoulli process. That is, talk about the number of arrivals during a given time period, and talk also about the time between consecutive arrivals, and for the distribution of inter-arrival times.

So let's start with a quick review of what we discussed last time. First, a note about language. If you think of coin tosses, we then talk about heads and tails. If you think of these as a sequence of trials, you can talk about successes and failures. The language that we will be using will be more the language of arrivals. That is, if in a given slot you have a success, you say that something arrived. If you have a failure, nothing arrived. And that language is a little more convenient and more natural, especially when we talk about continuous time-- to talk about arrivals instead of successes.

But in any case, for the Bernoulli process let's keep, for a little bit, the language of successes. Whereas working in discrete time, we have time slots. During each time slot, we have an independent Bernoulli trial. There is probability p of having a success. Different slots are independent of each other. And this probability p is the same for any given time slot.

So for this process we will discuss the one random variable of interest, which is the following. If we have n time slots, or n trials, how many arrivals will there be? Or how many successes will there be? Well, this is just given by the binomial PMF. Number of successes in n trials is a random variable that has a binomial PMF, and we know what this is.

Then we talked about inter-arrival times. The time until the first arrival happens has a geometric distribution. And we have seen that from some time ago. Now if you start thinking about the time until k arrivals happen, and we denote that by Yk, this is the time until the first arrival happens. And then after the first arrival happens, you have to wait some time until the second arrival happens, and so on. And then the time from the (k -1)th arrival, until arrival number k.

The important thing to realize here is that because the process has a memorylessness property, once the first arrival comes, it's as if we're starting from scratch and we will be flipping our coins until the next arrival comes. So the time it will take until the next arrival comes will also be a geometric random variable. And because different slots are independent, whatever happens after the first arrival is independent from whatever happened before. So T1 and T2 will be independent random variables. And similarly, all the way up to Tk.

So the time until the k-th arrival is a sum of independent geometric random variables, with the same parameter p. And we saw last time that we can find the probability distribution of Yk. The probability that Yk takes a value of t is equal to-- there's this combinatorial factor here, and then you get p to the k, (1-p) to the (t-k), and this formula is true for t equal to k, k+1, and so on. And this distribution has a name. It's called the Pascal PMF.

So this is all there is to know about the Bernoulli process. One important comment is to realize what exactly this memorylessness property is saying. So I discussed it a little bit last time. Let me reiterate it. So we have a Bernoulli process, which is a sequence of Bernoulli trials. And these are (0,1) random variables that keep going on forever.

So someone is watching this movie of Bernoulli trials B_t. And at some point, they say they think, or something interesting has happened, why don't you come in and start watching? So at some time t, they tell you to come in and start watching. So what you will see once you come in will be this future trials.

So actually what you will see is a random process, whose first random variable is going to be the first one that you see, B_(t +1). The second one is going to be this, and so on. So this is the process that's seen by the person who's asked to come in and start watching at that time. And the claim is that this process is itself a Bernoulli process, provided that the person who calls you into the room does not look into the future. The person who calls you into the room decides to call you in only on the basis of what they have seen so far.

So for example, who calls you into the room might have a rule that says, as soon as I see a sequence of 3 heads, I ask the other person to come in. So if they use that particular rule, it means that when you're called in, the previous 3 were heads. But this doesn't give you any information about the future. And so the future ones will be just independent Bernoulli trials.

If on the other hand, the person who calls you in has seen the movie before and they use a rule, such as, for example, I call you in just before 3 heads show up for the first time. So the person calls you in based on knowledge that these two would be three heads. If they have such foresight-- if they can look into the future-- then X1, X2, X3, they're certain to be three heads, so they do not correspond to random independent Bernoulli trials.

So to rephrase this, the process is memoryless. It does not matter what has happened in the past. And that's true even if you are called into the room and start watching at a random time, as long as that random time is determined in a causal way on the basis of what has happened so far.

So you are called into the room in a causal manner, just based on what's happened so far. What you're going to see starting from that time will still be a sequence of independent Bernoulli trials. And this is the argument that we used here, essentially, to argue that this T2 is an independent random variable from T1.

So a person is watching the movie, sees the first success. And on the basis of what they have seen-- they have just seen the first success-- they ask you to come in. You come in. What you're going to see is a sequence of Bernoulli trials. And you wait this long until the next success comes in. What you see is a Bernoulli process, as if the process was just starting right now. And that convinces us that this should be a geometric random variable of the same kind as this one, as independent from what happened before.

All right. So this is pretty much all there is to know about the Bernoulli process. Plus the two things that we did at the end of the last lecture where we merge two independent Bernoulli processes, we get a Bernoulli process. If we have a Bernoulli process and we split it by flipping a coin and sending things one way or the other, then we get two separate Bernoulli processes. And we see that all of these carry over to the continuous time. And our task for today is basically to work these continuous time variations.

So the Poisson process is a continuous time version of the Bernoulli process. Here's the motivation for considering it a Bernoulli process. So you have that person whose job is to sit outside the door of a bank. And they have this long sheet, and for every one second slot, they mark an X if a person came in, or they mark something else if no one came in during that slot.

Now the bank manager is a really scientifically trained person and wants very accurate results. So they tell you, don't use one second slots, use milliseconds slots. So you have all those slots and you keep filling if someone arrived or not during that slot. Well then you come up with an idea. Why use millisecond slots and keep putting crosses or zero's into each slot? It's much simpler if I just record the exact times when people came in. So time is continuous.

I don't keep doing something at every time slot. But instead of the time axis, I mark the times at which customers arrive. So there's no real need for slots. The only information that you want is when did we have arrivals of people. And we want to now model a process of this kind happening in continuous time, that has the same flavor, however, as the Bernoulli process. So that's the model we want to develop.

OK. So what are the properties that we're going to have? First, we're going to assume that intervals over the same length behave probabilistically in an identical fashion. So what does that mean? Think of an interval of some given length. During the interval of that length, there's going to be a random number of arrivals. And that random number of arrivals is going to have a probability distribution. So that probability distribution-- let's denote it by this notation.

We fix t, we fix the duration. So this is fixed. And we look at the different k's. The probability of having 0 arrivals, the probability of 1 arrival, the probability of 2 arrivals, and so on. So this thing is essentially a PMF. So it should have the property that the sum over all k's of this P_(k, tau) should be equal to 1.

Now, hidden inside this notation is an assumption of time homogeneity. That is, this probability distribution for the number of arrivals only depends on the length of the interval, but not the exact location of the interval on the time axis.

That is, if I take an interval of length tau, and I ask about the number of arrivals in this interval. And I take another interval of length tau, and I ask about the number of arrivals during that interval. Number of arrivals here, and number of arrivals there have the same probability distribution, which is denoted this way.

So the statistical behavior of arrivals here is the same as the statistical behavioral of arrivals there. What's the relation with the Bernoulli process? It's very much like the assumption-- the Bernoulli process-- that in different slots, we have the same probability of success. Every slot looks probabilistically as any other slot.

So similarly here, any interval of length tau looks probabilistically as any other interval of length tau. And the number of arrivals during that interval is a random variable described by these probabilities. Number of arrivals here is a random variable described by these same probabilities. So that's our first assumption.

Then what else? In the Bernoulli process we had the assumption that different time slots were independent of each other. Here we do not have time slots, but we can still think in a similar way and impose the following assumption, that these joint time intervals are statistically independent. What does that mean?

Does a random number of arrivals during this interval, and the random number of arrivals during this interval, and the random number of arrivals during this interval-- so these are three different random variables-- these three random variables are independent of each other. How many arrivals we got here is independent from how many arrivals we got there.

So this is similar to saying that different time slots were independent. That's what we did in discrete time. The continuous time analog is this independence assumption. So for example, in particular, number of arrivals here is independent from the number of arrivals there. So these are two basic assumptions about the process.

Now in order to write down a formula, eventually, about this probability distribution-- which is our next objective, we would like to say something specific about this distribution of number of arrivals-- we need to add a little more structure into the problem.

And we're going to make the following assumption. If we look at the time interval of length delta-- and delta now is supposed to be a small number, so a picture like this-- during a very small time interval, there is a probability that we get exactly one arrival, which is lambda times delta. Delta is the length of the interval and lambda is a proportionality factor, which is sort of the intensity of the arrival process.

Bigger lambda means that a little interval is more likely to get an arrival. So there's a probability lambda times delta of 1 arrival. The remaining probability goes to 0 arrivals. And when delta is small, the probability of 2 arrivals can be approximated by 0. So this is a description of what happens during a small, tiny slot.

Now this is something that's supposed to be true in some limiting sense, when delta is very small. So the exact version of this statement would be that this is an equality, plus order of delta squared terms. So this is an approximate equality. And what approximation means is that in the limit of small deltas, the dominant terms-- the constant and the first order term are given by this.

Now when delta is very small, second order terms in delta do not matter. They are small compared to first order terms. So we ignore this. So you can either think in terms of an exact relation, which is the probabilities are given by this, plus delta squared terms. Or if you want to be a little more loose, you just write here, as an approximate equality. And the understanding is that this equality holds-- approximately becomes more and more correct as delta goes to 0.

So another version of that statement would be that if you take the limit as delta goes to 0, of p, the probability of having 1 arrival in an interval of length delta, divided by delta, this is equal to lambda. So that would be one version of an exact statement of what we are assuming here.

So this lambda, we call it the arrival rate, or the intensity of the process. And clearly, if you double lambda, then a little interval is likely -- you expect to get -- the probability of obtaining an arrival during that interval has doubled. So in some sense we have twice as intense arrival process.

If you look at the number of arrivals during delta interval, what is the expected value of that random variable? Well with probability lambda delta we get 1 arrival. And with the remaining probability, we get 0 arrivals. So it's just lambda times delta. So expected number of arrivals during a little interval is lambda times delta. So expected number of arrivals is proportional to lambda, and that's again why we call lambda the arrival rate.

If you send delta to the denominator in this equality, it tells you that lambda is the expected number of arrivals per unit time. So the arrival rate is expected number of arrivals per unit time. And again, that justifies why we call lambda the intensity of this process.

All right. So where are we now? For the Bernoulli process, the number of arrivals during a given interval of length n had the PMF that we knew it was the binomial PMF. What is the formula for the corresponding PMF for the continuous time process? Somehow we would like to use our assumptions and come up with the formula for this quantity.

So this tells us about the distribution of number of arrivals during an interval of some general length. We have made assumptions about the number of arrivals during an interval of small length. An interval of big length is composed of many intervals of small length, so maybe this is the way to go. Take a big interval, and split it into many intervals of small length.

So we have here our time axis. And we have an interval of length tau. And I'm going to split it into lots of little intervals of length delta. So how many intervals are we going to have? The number of intervals is going to be the total time, divided by delta.

Now what happens during each one of these little intervals? As long as the intervals are small, what you have is that during an interval, you're going to have either 0 or 1 arrival. The probability of more than 1 arrival during a little interval is negligible.

So with this picture, you have essentially a Bernoulli process that consists of so many trials. And during each one of those trials, we have a probability of success, which is lambda times delta.

Different little intervals here are independent of each other. That's one of our assumptions, that these joint time intervals are independent. So approximately, what we have is a Bernoulli process. We have independence. We have the number of slots of interest. And during each one of the slots we have a certain probability of success.

So if we think of this as another good approximation of the Poisson process-- with the approximation becoming more and more accurate as delta goes to 0 -- what we should do would be to take the formula for the PMF of number of arrivals in a Bernoulli process, and then take the limit as delta goes to 0.

So in the Bernoulli process, the probability of k arrivals is n choose k, and then you have p to the k. Now in our case, we have here lambda times delta, delta is tau over n. Delta is tau over n, so p is lambda times tau divided by n. So here's our p -- Lambda tau over n -- to the power k, and then times one minus this-- this is our one minus p-- to the power n-k.

So this is the exact formula for the Bernoulli process. For the Poisson process, what we do is we take that formula and we let delta go to 0. As delta goes to 0, n goes to infinity. So that's the limit that we're taking.

On the other hand, this expression lambda times tau-- lambda times tau, what is it going to be? Lambda times tau is equal to n times p. n times p, is that what I want? No, let's see. Lambda tau is np. Yeah. So lambda tau is np.

All right. So we have this relation, lambda tau equals np. These two numbers being equal kind of makes sense. np is the expected number of successes you're going to get in the Bernoulli process. Lambda tau-- since lambda is the arrival rate and you have a total time of tau, lambda tau you can think of it as the number of expected arrivals in the Bernoulli process.

We're doing a Bernoulli approximation to the Poisson process. We take the formula for the Bernoulli, and now take the limit as n goes to infinity. Now lambda tau over n is equal to p, so it's clear what this term is going to give us. This is just p to the power k.

It will actually take a little more work than that. Now I'm not going to do the algebra, but I'm just telling you that one can take the limit in this formula here, as n goes to infinity. And that will give you another formula, the final formula for the Poisson PMF.

One thing to notice is that here you have something like 1 minus a constant over n, to the power n. And you may recall from calculus a formula of this kind, that this converges to e to the minus c. If you remember that formula from calculus, then you will expect that here, in the limit, you are going to get something like an e to the minus lambda tau. So indeed, we will get such a term.

There is some work that needs to be done to find the limit of this expression, times that expression. The algebra is not hard, it's in the text. Let's not spend more time doing this. But let me just give you the formula of what comes at the end. And the formula that comes at the end is of this form.

So what matters here is not so much the specific algebra that you will do to go from this formula to that one. It's kind of straightforward. What's important is the idea that the Poisson process, by definition, can be approximated by a Bernoulli process in which we have a very large number of slots-- n goes to infinity. Whereas we have a very small probability of success during each time slot. So a large number of slots, but tiny probability of success during each slot. And we take the limit as the slots become smaller and smaller.

So with this approximation we end up with this particular formula. And this is the so-called Poisson PMF. Now this function P here -- has two arguments. The important thing to realize is that when you think of this as a PMF, you fix t to tau. And for a fixed tau, now this is a PMF. As I said before, the sum over k has to be equal to 1. So for a given tau, these probabilities add up to 1. The formula is moderately messy, but not too messy. One can work with it without too much pain.

And what's the mean and variance of this PMF? Well what's the expected number of arrivals? If you think of this Bernoulli analogy, we know that the expected number of arrivals in the Bernoulli process is n times p. In the approximation that we're using in these procedure, n times p is the same as lambda tau. And that's why we get lambda tau to be the expected number of arrivals. Here I'm using t instead of tau. The expected number of arrivals is lambda t.

So if you double the time, you expect to get twice as many arrivals. If you double the arrival rate, you expect to get twice as many arrivals. How about the formula for the variance? The variance of the Bernoulli process is np, times one minus p.

What does this go to in the limit? In the limit that we're taking, as delta goes to zero, then p also goes to zero. The probability of success in any given slot goes to zero. So this term becomes insignificant. So this becomes n times p, which is again lambda t, or lambda tau.

So the variance, instead of having this more complicated formula of the variance is the Bernoulli process, here it gets simplified and it's lambda t. So interestingly, the variance in the Poisson process is exactly the same as the expected value. So you can look at this as just some interesting coincidence.

So now we're going to take this formula and see how to use it. First we're going to do a completely trivial, straightforward example. So 15 years ago when that example was made, email was coming at a rate of five messages per hour. I wish that was the case today. And now emails that are coming in, let's say during the day-- the arrival rates of emails are probably different in different times of the day. But if you fix a time slot, let's say 1:00 to 2:00 in the afternoon, there's probably a constant rate. And email arrivals are reasonably well modeled by a Poisson process.

Speaking of modeling, it's not just email arrivals. Whenever arrivals happen in a completely random way, without any additional structure, the Poisson process is a good model of these arrivals. So the times at which car accidents will happen, that's a Poisson processes. If you have a very, very weak light source that's shooting out photons, just one at a time, the times at which these photons will go out is well modeled again by a Poisson process. So it's completely random.

Or if you have a radioactive material where one atom at a time changes at random times. So it's a very slow radioactive decay. The time at which these alpha particles, or whatever we get emitted, again is going to be described by a Poisson process. So if you have arrivals, or emissions, that happen at completely random times, and once in a while you get an arrival or an event, then the Poisson process is a very good model for these events.

So back to emails. Get them at a rate of five messages per day, per hour. In 30 minutes this is half an hour. So what we have is that lambda t, total number of arrivals is-- the expected number of arrivals is-- lambda is five, t is one-half, if we talk about hours. So lambda t is two to the 0.5.

The probability of no new messages is the probability of zero, in time interval of length t, which, in our case, is one-half. And then we look back into the formula from the previous slide, and the probability of zero arrivals is lambda t to the power zero, divided by zero factorial, and then an e to the lambda t. And you plug in the numbers that we have. Lambda t to the zero power is one. Zero factorial is one. So we're left with e to the minus 2.5. And that number is 0.08.

Similarly, you can ask for the probability that you get exactly one message in half an hour. And that would be-- the probability of one message in one-half an hour-- is going to be lambda t to the first power, divided by 1 factorial, e to the minus lambda t, which-- as we now get the extra lambda t factor-- is going to be 2.5, e to the minus 2.5. And the numerical answer is 0.20. So this is how you use the PMF formula for the Poisson distribution that we had in the previous slide.

All right. So this was all about the distribution of the number of arrivals. What else did we do last time? Last time we also talked about the time it takes until the k-th arrival. OK. So let's try to figure out something about this particular distribution. We can derive the distribution of the time of the k-th arrival by using the exact same argument as we did last time.

So now the time of the k-th arrival is a continuous random variable. So it has a PDF. Since we are in continuous time, arrivals can happen at any time. So Yk is a continuous random variable. But now let's think of a time interval of length little delta. And use our usual interpretation of PDFs. The PDF of a random variable evaluated at a certain time times delta, this is the probability that the Yk falls in this little interval.

So as I've said before, this is the best way of thinking about PDFs. PDFs give you probabilities of little intervals. So now let's try to calculate this probability. For the k-th arrival to happen inside this little interval, we need two things. We need an arrival to happen in this interval, and we need k minus one arrivals to happen during that interval.

OK. You'll tell me, but it's possible that we might have the k minus one arrival happen here, and the k-th arrival to happen here. In principle, that's possible. But in the limit, when we take delta very small, the probability of having two arrivals in the same little slot is negligible. So assuming that no two arrivals can happen in the same mini slot, then for the k-th one to happen here, we must have k minus one during this interval.

Now because we have assumed that these joint intervals are independent of each other, this breaks down into the probability that we have exactly k minus one arrivals, during the interval from zero to t, times the probability of exactly one arrival during that little interval, which is lambda delta.

We do have a formula for this from the previous slide, which is lambda t, to the k minus 1, over k minus one factorial, times e to minus lambda t. And then lambda times delta. Did I miss something?

Yeah, OK. All right. And now you cancel this delta with that delta. And that gives us a formula for the PDF of the time until the k-th arrival. This PDF, of course, depends on the number k. The first arrival is going to happen somewhere in this range of time. So this is the PDF that it has.

The second arrival, of course, is going to happen later. And the PDF is this. So it's more likely to happen around these times. The third arrival has this PDF, so it's more likely to happen around those times.

And if you were to take k equal to 100, you might get a PDF-- it's extremely unlikely that the k-th arrival happens in the beginning, and it might happen somewhere down there, far into the future. So depending on which particular arrival we're talking about, it has a different probability distribution. The time of the 100th arrival, of course, is expected to be a lot larger than the time of the first arrival.

Incidentally, the time of the first arrival has a PDF whose form is quite simple. If you let k equal to one here, this term disappears. That term becomes a one. You're left with just lambda, e to the minus lambda. And you recognize it, it's the exponential distribution. So the time until the first arrival in a Poisson process is an exponential distribution.

What was the time of the first arrival in the Bernoulli process? It was a geometric distribution. Well, not coincidentally, these two look quite a bit like the other. A geometric distribution has this kind of shape. The exponential distribution has that kind of shape. The geometric is just a discrete version of the exponential. In the Bernoulli case, we are in discrete time. We have a PMF for the time of the first arrival, which is geometric.

In the Poisson case, what we get is the limit of the geometric as you let those lines become closer and closer, which gives you the exponential distribution. Now the Poisson process shares all the memorylessness properties of the Bernoulli process. And the way one can argue is just in terms of this picture.

Since the Poisson process is the limit of Bernoulli processes, whatever qualitative processes you have in the Bernoulli process remain valid for the Poisson process. In particular we have this memorylessness property. You let the Poisson process run for some time, and then you start watching it. What ever happened in the past has no bearing about the future.

Starting from right now, what's going to happen in the future is described again by a Poisson process, in the sense that during every little slot of length delta, there's going to be a probability of lambda delta of having an arrival. And that probably lambda delta is the same-- is always lambda delta-- no matter what happened in the past of the process.

And in particular, we could use this argument to say that the time until the k-th arrival is the time that it takes for the first arrival to happen. OK, let me do it for k equal to two. And then after the first arrival happens, you wait a certain amount of time until the second arrival happens.

Now once the first arrival happened, that's in the past. You start watching. From now on you have mini slots of length delta, each one having a probability of success lambda delta. It's as if we started the Poisson process from scratch. So starting from that time, the time until the next arrival is going to be again an exponential distribution, which doesn't care about what happened in the past, how long it took you for the first arrival.

So these two random variables are going to be independent and exponential, with the same parameter lambda. So among other things, what we have done here is we have essentially derived the PDF of the sum of k independent exponentials. The time of the k-th arrival is the sum of k inter-arrival times. The inter-arrival times are all independent of each other because of memorylessness. And they all have the same exponential distribution.

And by the way, this gives you a way to simulate the Poisson process. If you wanted to simulate it on your computer, you would have one option to break time into tiny, tiny slots. And for every tiny slot, use your random number generator to decide whether there was an arrival or not. To get it very accurate, you would have to use tiny, tiny slots. So that would be a lot of computation.

The more clever way of simulating the Poisson process is you use your random number generator to generate a sample from an exponential distribution and call that your first arrival time. Then go back to the random number generator, generate another independent sample, again from the same exponential distribution. That's the time between the first and the second arrival, and you keep going that way.

So as a sort of a quick summary, this is the big picture. This table doesn't tell you anything new. But it's good to have it as a reference, and to look at it, and to make sure you understand what all the different boxes are. Basically the Bernoulli process runs in discrete time. The Poisson process runs in continuous time.

There's an analogy of arrival rates, p per trial, or intensity per unit time. We did derive, or sketched the derivation for the PMF of the number of arrivals. And the Poisson distribution, which is the distribution that we get, this Pk of t. Pk and t is the limit of the binomial when we take the limit in this particular way, as delta goes to zero, and n goes to infinity.

The geometric becomes an exponential in the limit. And the distribution of the time of the k-th arrival-- we had a closed form formula last time for the Bernoulli process. We got the closed form formula this time for the Poisson process. And we actually used exactly the same argument to get these two closed form formulas.

All right. So now let's talk about adding or merging Poisson processes. And there's two statements that we can make here. One has to do with adding Poisson random variables, just random variables. There's another statement about adding Poisson processes. And the second is a bigger statement than the first. But this is a warm up. Let's work with the first statement.

So the claim is that the sum of independent Poisson random variables is Poisson. OK. So suppose that we have a Poisson process with rate-- just for simplicity-- lambda one. And I take the interval from zero to two. And that take then the interval from two until five. The number of arrivals during this interval-- let's call it n from zero to two-- is going to be a Poisson random variable, with parameter, or with mean, two.

The number of arrivals during this interval is n from time two until five. This is again a Poisson random variable with mean equal to three, because the arrival rate is 1 and the duration of the interval is three. These two random variables are independent. They obey the Poisson distribution that we derived before. If you add them, what you get is the number of arrivals during the interval from zero to five.

Now what kind of distribution does this random variable have? Well this is the number of arrivals over an interval of a certain length in a Poisson process. Therefore, this is also Poisson with mean five.

Because for the Poisson process we know that this number of arrivals is Poisson, this is Poisson, but also the number of overall arrivals is also Poisson. This establishes that the sum of a Poisson plus a Poisson random variable gives us another Poisson random variable. So adding Poisson random variables gives us a Poisson random variable.

But now I'm going to make a more general statement that it's not just number of arrivals during a fixed time interval-- it's not just numbers of arrivals for given time intervals-- but rather if you take two different Poisson processes and add them up, the process itself is Poisson in the sense that this process is going to satisfy all the assumptions of a Poisson process.

So the story is that you have a red bulb that flashes at random times at the rate of lambda one. It's a Poisson process. You have an independent process where a green bulb flashes at random times. And you happen to be color blind, so you just see when something is flashing. So these two are assumed to be independent Poisson processes. What can we say about the process that you observe?

So in the processes that you observe, if you take a typical time interval of length little delta, what can happen during that little time interval? The red process may have something flashing. So red flashes. Or the red does not. And for the other bulb, the green bulb, there's two possibilities. The green one flashes. And the other possibility is that the green does not.

OK. So there's four possibilities about what can happen during a little slot. The probability that the red one flashes and the green one flashes, what is this probability? It's lambda one delta that the first one flashes, and lambda two delta that the second one does. I'm multiplying probabilities here because I'm making the assumption that the two processes are independent.

OK. Now the probability that the red one flashes is lambda one delta. But the green one doesn't is one, minus lambda two delta. Here the probability would be that the red one does not, times the probability that the green one does. And then here we have the probability that none of them flash, which is whatever is left. But it's one minus lambda one delta, times one minus lambda two delta.

Now we're thinking about delta as small. So think of the case where delta goes to zero, but in a way that we keep the first order terms. We keep the delta terms, but we throw away the delta squared terms. Delta squared terms are much smaller than the delta terms when delta becomes small. If we do that-- if we only keep the order of delta terms-- this term effectively disappears. This is delta squared. So we make it zero. So the probability of having simultaneously a red and a green flash during a little interval is negligible.

What do we get here? Lambda delta times one survives, but this times that doesn't. So we can throw that away. So the approximation that we get is lambda one delta. Similarly here, this goes away. We're left with a lambda two delta. And this is whatever remains, whatever is left.

So what do we have? That there is a probability of seeing a flash, either a red or a green, which is lambda one delta, plus lambda two delta. So if we take a little interval of length delta here, it's going to see an arrival with probability approximately lambda one, plus lambda two, delta.

So every slot in this merged process has an arrival probability with a rate which is the sum of the rates of these two processes. So this is one part of the definition of the Poisson process. There's a few more things that one would need to verify. Namely, that intervals of the same length have the same probability distribution and that different slots are independent of each other.

This can be argued by starting from here because different intervals in this process are independent from each other. Different intervals here are independent from each other. It's not hard to argue that different intervals in the merged process will also be independent of each other.

So the conclusion that comes at the end is that this process is a Poisson process, with a total rate which is equal to the sum of the rate of the two processes. And now if I tell you that an arrival happened in the merged process at a certain time, how likely is it that it came from here? How likely is it?

We go to this picture. Given that an arrival occurred-- which is the event that this or that happened-- what is the probability that it came from the first process, the red one? Well it's the probability of this divided by the probability of this, times that.

Given that this event occurred, you want to find the conditional probability of that sub event. So we're asking the question, out of the total probability of these two, what fraction of that probability is assigned here? And this is lambda one delta, after we ignore the other terms.

This is lambda two delta. So that fraction is going to be lambda one, over lambda one plus lambda two. What does this tell you? If lambda one and lambda two are equal, given that I saw an arrival here, it's equally likely to be red or green. But if the reds have a much higher arrival rate, when I see an arrival here, it's more likely this number will be large. So it's more likely to have come from the red process.

OK so we'll continue with this story and do some applications next time.