JOHN TSITSIKLIS: Today we're going to finish our discussion of the Poisson process. We're going to see a few of its properties, do a few interesting problems, some more interesting than others. So go through a few examples and then we're going to talk about some quite strange things that happen with the Poisson process.

So the first thing is to remember what the Poisson processes is. It's a model, let's say, of arrivals of customers that are, in some sense, quote unquote, completely random, that is a customer can arrive at any point in time. All points in time are equally likely. And different points in time are sort of independent of other points in time. So the fact that I got an arrival now doesn't tell me anything about whether there's going to be an arrival at some other time.

In some sense, it's a continuous time version of the Bernoulli process. So the best way to think about the Poisson process is that we divide time into extremely tiny slots. And in each time slot, there's an independent possibility of having an arrival. Different time slots are independent of each other. On the other hand, when the slot is tiny, the probability for obtaining an arrival during that tiny slot is itself going to be tiny.

So we capture these properties into a formal definition what the Poisson process is. We have a probability mass function for the number of arrivals, k, during an interval of a given length. So this is the sort of basic description of the distribution of the number of arrivals. So tau is fixed. And k is the parameter. So when we add over all k's, the sum of these probabilities has to be equal to 1.

There's a time homogeneity assumption, which is hidden in this, namely, the only thing that matters is the duration of the time interval, not where the time interval sits on the real axis. Then we have an independence assumption. Intervals that are disjoint are statistically independent from each other. So any information you give me about arrivals during this time interval doesn't change my beliefs about what's going to happen during another time interval. So this is a generalization of the idea that we had in Bernoulli processes that different time slots are independent of each other.

And then to specify this function, the distribution of the number of arrivals, we sort of go in stages. We first specify this function for the case where the time interval is very small. And I'm telling you what those probabilities will be. And based on these then, we do some calculations and to find the formula for the distribution of the number of arrivals for intervals of a general duration. So for a small duration, delta, the probability of obtaining 1 arrival is lambda delta. The remaining probability is assigned to the event that we get to no arrivals during that interval.

The probability of obtaining more than 1 arrival in a tiny interval is essentially 0. And when we say essentially, it's means modular, terms that of order delta squared. And when delta is very small, anything which is delta squared can be ignored. So up to delta squared terms, that's what happened during a little interval.

Now if we know the probability distribution for the number of arrivals in a little interval. We can use this to get the distribution for the number of arrivals over several intervals. How do we do that? The big interval is composed of many little intervals. Each little interval is independent from any other little interval, so is it is as if we have a sequence of Bernoulli trials. Each Bernoulli trial is associated with a little interval and has a small probability of obtaining a success or an arrival during that mini-slot.

On the other hand, when delta is small, and you take a big interval and chop it up, you get a large number of little intervals. So what we essentially have here is a Bernoulli process, in which is the number of trials is huge but the probability of success during any given trial is tiny. The average number of trials ends up being proportional to the length of the interval. If you have twice as large an interval, it's as if you're having twice as many over these mini-trials, so the expected number of arrivals will increase proportionately.

There's also this parameter lambda, which we interpret as expected number of arrivals per unit time. And it comes in those probabilities here. When you double lambda, this means that a little interval is twice as likely to get an arrival. So you would expect to get twice as many arrivals as well. That's why the expected number of arrivals during an interval of length tau also scales proportional to this parameter lambda. Somewhat unexpectedly, it turns out that the variance of the number of arrivals is also the same as the mean. This is a peculiarity that happens in the Poisson process.

So this is one way of thinking about Poisson process, in terms of little intervals, each one of which has a tiny probability of success. And we think of the distribution associated with that process as being described by this particular PMF. So this is the PMF for the number of arrivals during an interval of a fixed duration, tau. It's a PMF that extends all over the entire range of non-negative integers.

So the number of arrivals you can get during an interval for certain length can be anything. You can get as many arrivals as you want. Of course the probability of getting a zillion arrivals is going to be tiny. But in principle, this is possible. And that's because an interval, even if it's a fixed length, consists of an infinite number of mini-slots in some sense. You can divide, chop it up, into as many mini-slots as you want. So in principle, it's possible that every mini-slot gets an arrival. In principle, it's possible to get an arbitrarily large number of arrivals.

So this particular formula here is not very intuitive when you look at it. But it's a legitimate PMF. And it's called the Poisson PMF. It's the PMF that describes the number of arrivals. So that's one way of thinking about the Poisson process, where the basic object of interest would be this PMF and you try to work with it.

There's another way of thinking about what happens in the Poisson process. And this has to do with letting things evolve in time. You start at time 0. There's going to be a time at which the first arrival occurs, and call that time T1. This time turns out to have an exponential distribution with parameter lambda. Once you get an arrival, it's as if the process starts fresh.

The best way to understand why this is the case is by thinking in terms of the analogy with the Bernoulli process. If you believe that statement for the Bernoulli process, since this is a limiting case, it should also be true. So starting from this time, we're going to wait a random amount of time until we get the second arrival This random amount of time, let's call it T2. This time, T2 is also going to have an exponential distribution with the same parameter, lambda. And these two are going to be independent of each other. OK?

So the Poisson process has all the same memorylessness properties that the Bernoulli process has. What's another way of thinking of this property? So think of a process where you have a light bulb. The time at the light bulb burns out, you can model it by an exponential random variable. And suppose that they tell you that so far, we're are sitting at some time, T. And I tell you that the light bulb has not yet burned out. What does this tell you about the future of the light bulb? Is the fact that they didn't burn out, so far, is it good news or is it bad news? Would you rather keep this light bulb that has worked for t times steps and is still OK? Or would you rather use a new light bulb that starts new at that point in time?

Because of the memorylessness property, the past of that light bulb doesn't matter. So the future of this light bulb is statistically the same as the future of a new light bulb. For both of them, the time until they burn out is going to be described an exponential distribution. So one way that people described the situation is to say that used is exactly as good as a new. So a used on is no worse than a new one. A used one is no better than a new one. So a used light bulb that hasn't yet burnt out is exactly as good as a new light bulb. So that's another way of thinking about the memorylessness that we have in the Poisson process.

Back to this picture. The time until the second arrival is the sum of two independent exponential random variables. So, in principle, you can use the convolution formula to find the distribution of T1 plus T2, and that would be what we call Y2, the time until the second arrival. But there's also a direct way of obtaining to the distribution of Y2, and this is the calculation that we did last time on the blackboard. And actually, we did it more generally. We found the time until the case arrival occurs. It has a closed form formula, which is called the Erlang distribution with k degrees of freedom.

So let's see what's going on here. It's a distribution Of what kind? It's a continuous distribution. It's a probability density function. This is because the time is a continuous random variable. Time is continuous. Arrivals can happen at any time. So we're talking about the PDF. This k is just the parameter of the distribution. We're talking about the k-th arrival, so k is a fixed number. Lambda is another parameter of the distribution, which is the arrival rate So it's a PDF over the Y's, whereas lambda and k are parameters of the distribution. OK.

So this was what we knew from last time. Just to get some practice, let us do a problem that's not too difficult, but just to see how we use the various formulas that we have. So Poisson was a mathematician, but Poisson also means fish in French. So Poisson goes fishing. And let's assume that fish are caught according to a Poisson process.

That's not too bad an assumption. At any given point in time, you have a little probability that a fish would be caught. And whether you catch one now is sort of independent about whether at some later time a fish will be caught or not. So let's just make this assumption. And suppose that the rules of the game are that you-- Fish are being called it the certain rate of 0.6 per hour. You fish for 2 hours, no matter what. And then there are two possibilities. If I have caught a fish, I stop and go home. So if some fish have been caught, so there's at least 1 arrival during this interval, I go home. Or if nothing has being caught, I continue fishing until I catch something. And then I go home. So that's the description of what is going to happen.

And now let's starts asking questions of all sorts. What is the probability that I'm going to be fishing for more than 2 hours? I will be fishing for more than 2 hours, if and only if no fish were caught during those 2 hours, in which case, I will have to continue. Therefore, this is just this quantity. The probability of catching 2 fish in-- of catching 0 fish in the next 2 hours, and according to the formula that we have, this is going to be e to the minus lambda times how much time we have.

There's another way of thinking about this. The probability that I fish for more than 2 hours is the probability that the first catch happens after time 2, which would be the integral from 2 to infinity of the density of the first arrival time. And that density is an exponential. So you do the integral of an exponential, and, of course, you would get the same answer. OK. That's easy.

So what's the probability of fishing for more than 2 but less than 5 hours? What does it take for this to happen? For this to happen, we need to catch 0 fish from time 0 to 2 and catch the first fish sometime between 2 and 5. So if you-- one way of thinking about what's happening here might be to say that there's a Poisson process that keeps going on forever. But as soon as I catch the first fish, instead of continuing fishing and obtaining those other fish I just go home right now.

Now the fact that I go home before time 5 means that, if I were to stay until time 5, I would have caught at least 1 fish. I might have caught more than 1. So the event of interest here is that the first catch happens between times 2 and 5. So one way of calculating this quantity would be-- Its the probability that the first catch happens between times 2 and 5. Another way to deal with it is to say, this is the probability that I caught 0 fish in the first 2 hours and then the probability that I catch at least 1 fish during the next 3 hours.

This. What is this? The probability of 0 fish in the next 3 hours is the probability of 0 fish during this time. 1 minus this is the probability of catching at least 1 fish, of having at least 1 arrival, between times 2 and 5. If there's at least 1 arrival between times 2 and 5, then I would have gone home by time 5. So both of these, if you plug-in numbers and all that, of course, are going to give you the same answer.

Now next, what's the probability that I catch at least 2 fish? In which scenario are we? Under this scenario, I go home when I catch my first fish. So in order to catch at least 2 fish, it must be in this case. So this is the same as the event that I catch at least 2 fish during the first 2 time steps. So it's going to be the probability from 2 to infinity, the probability that I catch 2 fish, or that I catch 3 fish, or I catch more than that.

So it's this quantity. k is the number of fish that I catch. At least 2, so k goes from 2 to infinity. These are the probabilities of catching a number k of fish during this interval. And if you want a simpler form without an infinite sum, this would be 1 minus the probability of catching 0 fish, minus the probability of catching 1 fish, during a time interval of length 2. Another way to think of it. I'm going to catch 2 fish, at least 2 fish, if and only if the second fish caught in this process happens before time 2. So that's another way of thinking about the same event. So it's going to be the probability that the random variable Y2, the arrival time over the second fish, is less than or equal to 2. OK.

The next one is a little trickier. Here we need to do a little bit of divide and conquer. Overall, in this expedition, what the expected number of fish to be caught? One way to think about it is to try to use the total expectations theorem. And think of expected number of fish, given this scenario, or expected number of fish, given this scenario. That's a little more complicated than the way I'm going to do it.

The way I'm going to do is to think as follows-- Expected number of fish is the expected number of fish caught between times 0 and 2 plus expected number of fish caught after time 2. So what's the expected number caught between time 0 and 2? This is lambda t. So lambda is 0.6 times 2. This is the expected number of fish that are caught between times 0 and 2.

Now let's think about the expected number of fish caught afterwards. How many fish are being caught afterwards? Well it depends on the scenario. If we're in this scenario, we've gone home and we catch 0. If we're in this scenario, then we continue fishing until we catch one. So the expected number of fish to be caught after time 2 is going to be the probability of this scenario times 1. And the probability of that scenario is the probability that they call it's 0 fish during the first 2 time steps times 1, which is the number of fish I'm going to catch if I continue.

The expected total fishing time we can calculate exactly the same way. I'm jumping to the last one. My total fishing time has a period of 2 time steps. I'm going to fish for 2 time steps no matter what. And then if I caught 0 fish, which happens with this probability, my expected time is going to be the expected time from here onwards, which is the expected value of this geometric random variable with parameter lambda. So the expected time is 1 over lambda. And in our case this, is 1/0.6.

Finally, if I tell you that I have been fishing for 4 hours and nothing has been caught so far, how much do you expect this quantity to be? Here is the story that, again, that for the Poisson process used is as good as new. The process does not have any memory. Given what happens in the past doesn't matter for the future. It's as if the process starts new at this point in time. So this one is going to be, again, the same exponentially distributed random variable with the same parameter lambda.

So expected time until an arrival comes is an exponential distribut -- has an exponential distribution with parameter lambda, no matter what has happened in the past. Starting from now and looking into the future, it's as if the process has just started. So it's going to be 1 over lambda, which is 1/0.6. OK.

Now our next example is going to be a little more complicated or subtle. But before we get to the example, let's refresh our memory about what we discussed last time about merging Poisson independent Poisson processes. Instead of drawing the picture that way, another way we could draw it could be this. We have a Poisson process with rate lambda1, and a Poisson process with rate lambda2. They have, each one of these, have their arrivals. And then we form the merged process. And the merged process records an arrival whenever there's an arrival in either of the two processes.

This process in that process are assumed to be independent of each other. Now different times in this process and that process are independent of each other. So what happens in these two time intervals is independent from what happens in these two time intervals. These two time intervals to determine what happens here. These two time intervals determine what happens there. So because these are independent from these, this means that this is also independent from that. So the independence assumption is satisfied for the merged process.

And the merged process turns out to be a Poisson process. And if you want to find the arrival rate for that process, you argue as follows. During a little interval of length delta, we have probability lambda1 delta of having an arrival in this process. We have probability lambda2 delta of an arrival in this process, plus second order terms in delta, which we're ignoring. And then you do the calculation and you find that in this process, you're going to have an arrival probability, which is lambda1 plus lambda2, again ignoring second order in delta-- terms that are second order in delta. So the merged process is a Poisson process whose arrival rate is the sum of the arrival rates of the individual processes.

And the calculation we did at the end of the last lecture-- If I tell you that the new arrival happened here, where did that arrival come from? Did it come from here or from there? If the lambda1 is equal to lambda2, then by symmetry you would say that it's equally likely to have come from here or to come from there. But if this lambda is much bigger than that lambda, the fact that they saw an arrival is more likely to have come from there. And the formula that captures this is the following. This is the probability that my arrival has come from this particular stream rather than that particular stream.

So when an arrival comes and you ask, what is the origin of that arrival? It's as if I'm flipping a coin with these odds. And depending on outcome of that coin, I'm going to tell you came from there or it came from there. So the origin of an arrival is either this stream or that stream. And this is the probability that the origin of the arrival is that one. Now if we look at 2 different arrivals, and we ask about their origins-- So let's think about the origin of this arrival and compare it with the origin that arrival.

The origin of this arrival is random. It could be right be either this or that. And this is the relevant probability. The origin of that arrival is random. It could be either here or is there, and again, with the same relevant probability. Question. The origin of this arrival, is it dependent or independent from the origin that arrival? And here's how the argument goes. Separate times are independent. Whatever has happened in the process during this set of times is independent from whatever happened in the process during that set of times. Because different times have nothing to do with each other, the origin of this, of an arrival here, has nothing to do with the origin of an arrival there. So the origins of different arrivals are also independent random variables.

So if I tell you that-- yeah. OK. So it as if that each time that you have an arrival in the merge process, it's as if you're flipping a coin to determine where did that arrival came from and these coins are independent of each other. OK. OK.

Now we're going to use this-- what we know about merged processes to solve the problem that would be harder to do, if you were not using ideas from Poisson processes. So the formulation of the problem has nothing to do with the Poisson process. The formulation is the following. We have 3 light-bulbs. And each light bulb is independent and is going to die out at the time that's exponentially distributed. So 3 light bulbs. They start their lives and then at some point they die or burn out. So let's think of this as X, this as Y, and this as Z.

And we're interested in the time until the last light-bulb burns out. So we're interested in the maximum of the 3 random variables, X, Y, and Z. And in particular, we want to find the expected value of this maximum. OK.

So you can do derived distribution, use the expected value rule, anything you want. You can get this answer using the tools that you already have in your hands. But now let us see how we can connect to this picture with a Poisson picture and come up with the answer in a very simple way. What is an exponential random variable? An exponential random variable is the first act in the long play that involves a whole Poisson process. So an exponential random variable is the first act of a Poisson movie. Same thing here. You can think of this random variable as being part of some Poisson process that has been running. So it's part of this bigger picture.

We're still interested in the maximum of the 3. The other arrivals are not going to affect our answers. It's just, conceptually speaking, we can think of the exponential random variable as being embedded in a bigger Poisson picture. So we have 3 Poisson process that are running in parallel. Let us split the expected time until the last burnout into pieces, which is time until the first burnout, time from the first until the second, and time from the second until the third. And find the expected values of each one of these pieces.

What can we say about the expected value of this? This is the first arrival out of all of these 3 Poisson processes. It's the first event that happens when you look at all of these processes simultaneously. So 3 Poisson processes running in parallel. We're interested in the time until one of them, any one of them, gets in arrival. Rephrase. We merged the 3 Poisson processes, and we ask for the time until we observe an arrival in the merged process.

When 1 of the 3 gets an arrival for the first time, the merged process gets its first arrival. So what's the expected value of this time until the first burnout? It's going to be the expected value of a Poisson random variable. So the first burnout is going to have an expected value, which is-- OK. It's a Poisson process. The merged process of the 3 has a collective arrival rate, which is 3 times lambda.

So this is the parameter over the exponential distribution that describes the time until the first arrival in the merged process. And the expected value of this random variable is 1 over that. When you have an exponential random variable with parameter lambda, the expected value of that random variable is 1 over lambda. Here we're talking about the first arrival time in a process with rate 3 lambda. The expected time until the first arrival is 1 over (3 lambda). Alright.

So at this time, this bulb, this arrival happened, this bulb has been burned. So we don't care about that bulb anymore. We start at this time, and we look forward. This bulb has been burned. So let's just look forward from now on. What have we got? We have two bulbs that are burning. We have a Poisson process that's the bigger picture of what could happen to that light bulb, if we were to keep replacing it. Another Poisson process. These two processes are, again, independent.

From this time until that time, how long does it take? It's the time until either this process records an arrival or that process records and arrival. That's the same as the time that the merged process of these two records an arrival. So we're talking about the expected time until the first arrival in a merged process. The merged process is Poisson. It's Poisson with rate 2 lambda. So that extra time is going to take-- the expected value is going to be 1 over the (rate of that Poisson process). So 1 over (2 lambda) is the expected value of this random variable.

So at this point, this bulb now is also burned. So we start looking from this time on. That part of the picture disappears. Starting from this time, what's the expected value until that remaining light-bulb burns out? Well, as we said before, in a Poisson process or with exponential random variables, we have memorylessness. A used bulb is as good as a new one. So it's as if we're starting from scratch here. So this is going to be an exponential random variable with parameter lambda. And the expected value of it is going to be 1 over lambda.

So the beauty of approaching this problem in this particular way is, of course, that we manage to do everything without any calculus at all, without striking an integral, without trying to calculate expectations in any form. Most of the non-trivial problems that you encounter in the Poisson world basically involve tricks of these kind. You have a question and you try to rephrase it, trying to think in terms of what might happen in the Poisson setting, use memorylessness, use merging, et cetera, et cetera.

Now we talked about merging. It turns out that the splitting of Poisson processes also works in a nice way. The story here is exactly the same as for the Bernoulli process. So I'm having a Poisson process. And each time, with some rate lambda, and each time that an arrival comes, I'm going to send it to that stream and the record an arrival here with some probability P. And I'm going to send it to the other stream with some probability 1 minus P. So either of this will happen or that will happen, depending on the outcome of the coin flip that I do. Each time that then arrival occurs, I flip a coin and I decide whether to record it here or there. This is called splitting a Poisson process into two pieces.

What kind of process do we get here? If you look at the little interval for length delta, what's the probability that this little interval gets an arrival? It's the probability that this one gets an arrival, which is lambda delta times the probability that after I get an arrival my coin flip came out to be that way, so that it sends me there. So this means that this little interval is going to have probability lambda delta P. Or maybe more suggestively, I should write it as lambda P times delta.

So every little interval has a probability of an arrival proportional to delta. The proportionality factor is lambda P. So lambda P is the rate of that process. And then you go through the mental exercise that you went through for the Bernoulli process to argue that a different intervals here are independent and so on. And that completes checking that this process is going to be a Poisson process.

So when you split a Poisson process by doing independent coin flips each time that something happens, the processes that you get is again a Poisson process, but of course with a reduced rate. So instead of the word splitting, sometimes people also use the words thinning-out. That is, out of the arrivals that came, you keep a few but throw away a few. OK.

So now the last topic over this lecture is a quite curious phenomenon that goes under the name of random incidents. So here's the story. Buses have been running on Mass Ave. from time immemorial. And the bus company that runs the buses claims that they come as a Poisson process with some rate, let's say, of 4 buses per hour. So that the expected time between bus arrivals is going to be 15 minutes. OK. Alright.

So people have been complaining that they have been showing up there. They think the buses are taking too long. So you are asked to investigate. Is the company-- Does it operate according to its promises or not. So you send an undercover agent to go and check the interarrival times of the buses. Are they 15 minutes? Or are they longer?

So you put your dark glasses and you show up at the bus stop at some random time. And you go and ask the guy in the falafel truck, how long has it been since the last arrival? So of course that guy works for the FBI, right? So they tell you, well, it's been, let's say, 12 minutes since the last bus arrival. And then you say, "Oh, 12 minutes. Average time is 15. So a bus should be coming any time now."

Is that correct? No, you wouldn't think that way. It's a Poisson process. It doesn't matter how long it has been since the last bus arrival. So you don't go through that fallacy. Instead of predicting how long it's going to be, you just sit down there and wait and measure the time. And you find that this is, let's say, 11 minutes. And you go to your boss and report, "Well, it took-- I went there and the time from the previous bus to the next one was 23 minutes. It's more than the 15 that they said."

So go and do that again. You go day after day. You keep these statistics of the length of this interval. And you tell your boss it's a lot more than 15. It tends to be more like 30 or so. So the bus company is cheating us. Does the bus company really run Poisson buses at the rate that they have promised? Well let's analyze the situation here and figure out what the length of this interval should be, on the average.

The naive argument is that this interval is an interarrival time. And interarrival times, on the average, are 15 minutes, if the company runs indeed Poisson processes with these interarrival times. But actually the situation is a little more subtle because this is not a typical interarrival interval. This interarrival interval consists of two pieces. Let's call them T1 and T1 prime. What can you tell me about those two random variables? What kind of random variable is T1? Starting from this time, with the Poisson process, the past doesn't matter. It's the time until an arrival happens. So T1 is going to be an exponential random variable with parameter lambda.

So in particular, the expected value of T1 is going to be 15 by itself. How about the random variable T1 prime. What kind of random variable is it? This is like the first arrival in a Poisson process that runs backwards in time. What kind of process is a Poisson process running backwards in time? Let's think of coin flips. Suppose you have a movie of coin flips. And for some accident, that fascinating movie, you happen to watch it backwards. Will it look any different statistically? No. It's going to be just the sequence of random coin flips.

So a Bernoulli process that's runs in reverse time is statistically identical to a Bernoulli process in forward time. The Poisson process is a limit of the Bernoulli. So, same story with the Poisson process. If you run it backwards in time it looks the same. So looking backwards in time, this is a Poisson process. And T1 prime is the time until the first arrival in this backward process.

So T1 prime is also going to be an exponential random variable with the same parameter, lambda. And the expected value of T1 prime is 15. Conclusion is that the expected length of this interval is going to be 30 minutes. And the fact that this agent found the average to be something like 30 does not contradict the claims of the bus company that they're running Poisson buses with a rate of lambda equal to 4. OK.

So maybe the company can this way-- they can defend themselves in court. But there's something puzzling here. How long is the interarrival time? Is it 15? Or is it 30? On the average. The issue is what do we mean by a typical interarrival time. When we say typical, we mean some kind of average. But average over what? And here's two different ways of thinking about averages. You number the buses. And you have bus number 100. You have bus number 101, bus number 102, bus number 110, and so on.

One way of thinking about averages is that you pick a bus number at random. I pick, let's say, that bus, all buses being sort of equally likely to be picked. And I measure this interarrival time. So for a typical bus. Then, starting from here until there, the expected time has to be 1 over lambda, for the Poisson process.

But what we did in this experiment was something different. We didn't pick a bus at random. We picked a time at random. And if the picture is, let's say, this way, I'm much more likely to pick this interval and therefore this interarrival time, rather than that interval. Because, this interval corresponds to very few times. So if I'm picking a time at random and, in some sense, let's say, uniform, so that all times are equally likely, I'm much more likely to fall inside a big interval rather than a small interval.

So a person who shows up at the bus stop at a random time. They're selecting an interval in a biased way, with the bias favor of longer intervals. And that's why what they observe is a random variable that has a larger expected value then the ordinary expected value.

So the subtlety here is to realize that we're talking between two different kinds of experiments. Picking a bus number at random verses picking an interval at random with a bias in favor of longer intervals. Lots of paradoxes that one can cook up using Poisson processes and random processes in general often have to do with the story of this kind.

The phenomenon that we had in this particular example also shows up in general, whenever you have other kinds of arrival processes. So the Poisson process is the simplest arrival process there is, where the interarrival times are exponential random variables. There's a larger class of models. They're called renewal processes, in which, again, we have a sequence of successive arrivals, interarrival times are identically distributed and independent, but they may come from a general distribution.

So to make the same point of the previous example but in a much simpler setting, suppose that bus interarrival times are either 5 or 10 minutes apart. So you get some intervals that are of length 5. You get some that are of length 10. And suppose that these are equally likely. So we have -- not exactly -- In the long run, we have as many 5 minute intervals as we have 10 minute intervals.

So the average interarrival time is 7 and 1/2. But if a person shows up at a random time, what are they going to see? Do we have as many 5s as 10s? But every 10 covers twice as much space. So if I show up at a random time, I have probability 2/3 falling inside an interval of duration 10. And I have one 1/3 probability of falling inside an interval of duration 5. That's because, out of the whole real line, 2/3 of it is covered by intervals of length 10, just because they're longer. 1/3 is covered by the smaller intervals.

Now if I fall inside an interval of length 10 and I measure the length of the interval that I fell into, that's going to be 10. But if I fall inside an interval of length 5 and I measure how long it is, I'm going to get a 5. And that, of course, is going to be different than 7.5. OK. And which number should be bigger? It's the second number that's bigger because this one is biased in favor of the longer intervals. So that's, again, another illustration of the different results that you get when you have this random incidence phenomenon.

So the bottom line, again, is that if you talk about a typical interarrival time, one must be very precise in specifying what we mean typical. So typical means sort of random. But to use the word random, you must specify very precisely what is the random experiment that you are using. And if you're not careful, you can get into apparent puzzles, such as the following. Suppose somebody tells you the average family size is 4, but the average person lives in a family of size 6. Is that compatible? Family size is 4 on the average, but typical people live, on the average, in families of size 6. Well yes. There's no contradiction here.

We're talking about two different experiments. In one experiment, I pick a family at random, and I tell you the average family is 4. In another experiment, I pick a person at random and I tell you that this person, on the average, will be in their family of size 6. And what is the catch here? That if I pick a person at random, large families are more likely to be picked. So there's a bias in favor of large families.

Or if you want to survey, let's say, are trains crowded in your city? Or are buses crowded? One choice is to pick a bus at random and inspect how crowded it is. Another choice is to pick a typical person and ask them, "Did you ride the bus today? Was it's crowded?" Well suppose that in this city there's one bus that's extremely crowded and all the other buses are completely empty. If you ask a person. "Was your bus crowded?" They will tell you, "Yes, my bus was crowded." There's no witness from the empty buses to testify in their favor. So by sampling people instead of sampling buses, you're going to get different result.

And in the process industry, if your job is to inspect and check cookies, you will be faced with a big dilemma. Do you want to find out how many chocolate chips there are on a typical cookie? Are you going to interview cookies or are you going to interview chocolate chips and ask them how many other chips where there on your cookie? And you're going to get different answers in these cases. So moral is, one has to be very precise on how you formulate the sampling procedure that you have. And you'll get different answers.