Introduction to Stochastic Processes
Professor Manjesh Hanawal
Industrial Engineering & Operations Research
Indian Institute of Technology, Bombay
Lecture No 46
Renewal Reward Theorem

So, we are just going to now cover one last topic, which is just with more generalization of, our elementary renewal theorem which is called Renewal Reward theorem. See like, in the setup when we consider elementary reward theorem. We kind of treated all cycles to have kind of similar value.

All cycles when I returned from state J 2 against state J, that is like kind of completing one cycle, again going from J to against state J that is completing one cycle. So, these are cycles. But what could happen that when you are looking at some applications, when you are going from this state to this state, again, state J to state J again in between you could be, accruing some reward or some cost.

So, for example, so in the battery case, when you and your battery went from full charge to discharge that completed one cycle, you may be interested in how much, my lamp burnt with that battery in that case. So, maybe let us say that your battery is powering something, for time being okay for time being, let us assume your battery is powering some electrical vehicle.

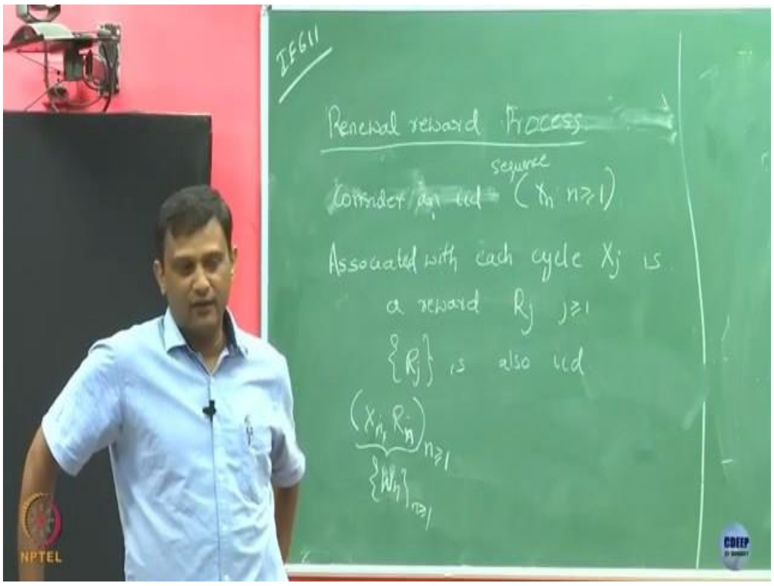
So, when your battery is powering some electrical vehicle, it may be moving some distance. So, when battery goes from full charge to zero charge, your vehicle might have covered some distance. The distance covered by your vehicle maybe kind of reward for you. Now, you want to basically measure for each and that could be random.

That could be again random because every time you charge your battery until it goes to zero. The amount of the distance that your vehicle travelled could be modelled as a stochastic variable. It is not always, it may not be deterministic because when you charge your battery till it goes zero.

That time it could have gone through different routes which could be having different level of traffics because of that it is not necessary that between, when from full charge to discharge, full discharge. It has to travel the same amount of distance. Maybe it travel, but that is a possibility that it could be stochastic also.

Now, in that case, you may be interested in what is the average reward I accumulated per cycle, okay. So, let us say the total in this example the total distance travelled, average total distance travelled. So, how we are going to characterize such things. So, such questions can be again answered based on a similar setup, but we have to bring in the notion of reward associated with each cycle.

(Reference Slide Time: 04:18)



So, let us try to formalize that, okay. Now, (())(4:33) simply now consider a renewal process, so what is that, now I am going to just consider an IID sequence Xn. I am not going to distinguish the first cycle and the rest of the cycles. So, it is just like all the cycles are same.

Now, associated with each cycle Xj, now I am going to call this as cycle one, life cycle one, what we call it as an Xj we called as lifetimes. One lifetime as I am going to call it as one cycle associated with each cycle Xj is a reward Rj, and the Rj is also iid. So, with each cycle J here is an associated reward Rj and this reward processes. I am also going to assume as iid. And now, so, combiningly I can write this process as Rii where, in each cycle there are two things.

One is the length of the cycle and associated reward but this reward may depend on the length of the cycle. So, Rn could be function of Xn, that is it could be dependent of Xn but it is independent of other cycles and other rewards. So, this Rn, it just is a function of Xn in that way. This is going to be again an iid process for each n.

So, now if you are going to treat this as one quantity, this quantity, let us say I call this as simply Yn which consists of Xn and Yn and if I treat this as a now let us call it Z, Z is also used, let us call w. This w is again an iid process where Wn corresponds to Xn and Yn.

Now, I am interested in accumulated reward. What I mean? As I said in the, in the electrical vehicle example, if I have run it for, run my vehicle for let us say till time T what is my accumulated reward till that time. So, in that, in the, in the vehicle example, electrical vehicle example Rn is the let us say total distance you travel, when your battery lasted for Xn duration.

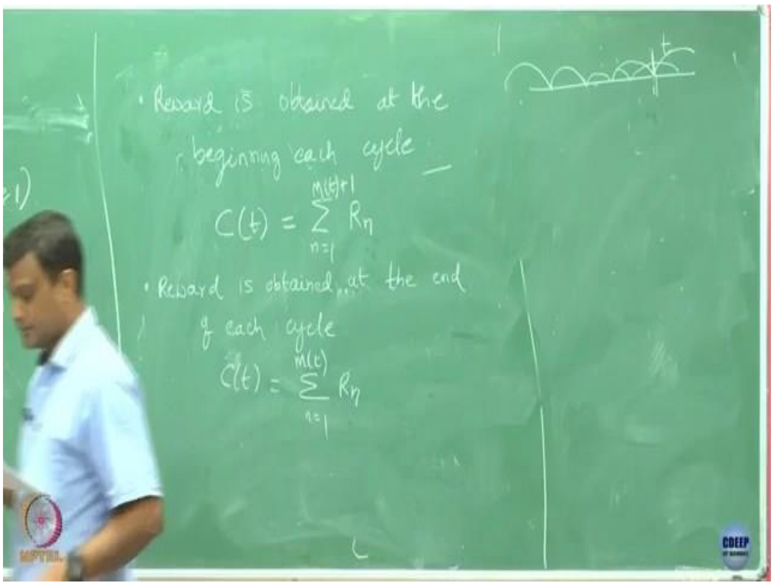
So, now how can, how to define my total reward for a time T. So, for example, let us say you have vehicle running and you want to define your reward after let us say T equals to 10 days. So, one natural way to do is, you just see in those 10 days how many cycles have completed, that is, how many times, have replaced your battery and then just add the rewards you got in each of the cycles.

And here you have to be a bit more specific because when you say 10th day, maybe the battery is just replaced yesterday and that battery is still running when you are looked at the 10th day.

So, one possibility is you can define that the reward till time 10th day is reward at accumulated till the completion of previous cycle or you could simply define that. When I say my reward, this reward is I obtained when I start the cycle or other possibility is when I end the cycle.

For example, in Electrical vehicle case when you completely charge your battery and put it in the vehicle. In the beginning you have not got anything. Only when you, let us say then you have battery complete then you look at how much distance you have covered after battery completely discharge then you can say this is the total reward I got in. In this case, you are defining your reward at the end of the cycle or there could be some cases where you could say that I got the reward as soon as this cycle started.

(Reference Slide Time: 10:45)



So, suppose you reward is obtained at the beginning of each cycle. So, then what you can do the total reward you accumulate till time T can be defined as summation of Rn, n equals to 1 to M of t. What is M of T indicates here? Number of cycles that have been completely covered, till times 0 to t right and it might be possible that the next cycle would have already started within the time T but when it has started you are already including that reward in the in this cycle, because you are getting this reward as soon as the cycle starts.

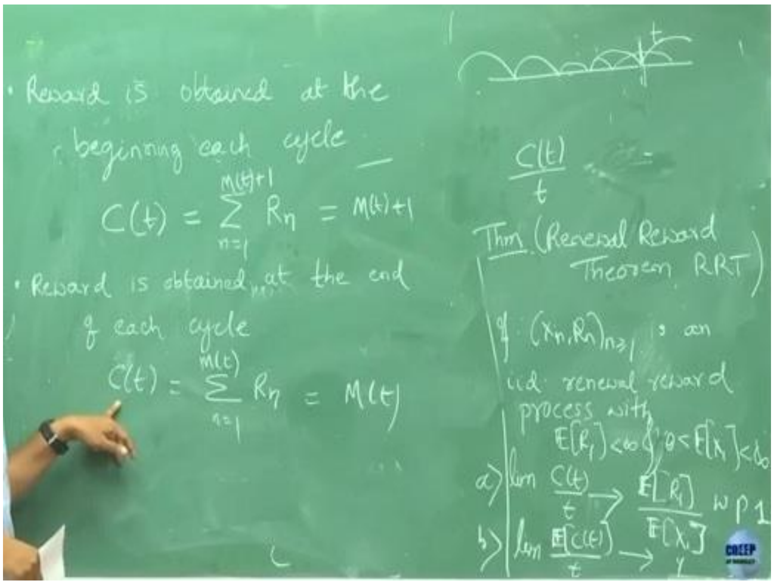
So, that is why you are going to include it here. Now, reward is obtained at the end of each cycle. So, if you are going to get the reward only at the end of the cycle. How you are going to define your total reward? Summation N equals to 1 to just Mt. Because you know that exactly Mt number of cycle have been completed till time T.

And that that is why I have to include only that many because I am getting reward only at the end of the cycles. That means I should only care about those cycles, which are completed within the interval 0t, whereas when I defined here when rewards I get at the beginning of the cycle when I defined C of t, I have to worry about all the cycles that have already begun within the M of 0 to T.

So, this is two possibilities. Depending on your applications, whether you can define it differently. So, let us take these two cases here. So, your reward has ended, let us say this is your time T, this is your time T, till this time T and let us say these are your cycles. One cycle went like this, another cycle went like this, another went and we just finished like this.

Now, we are saying beginning of this, now if it has ended this, the next one is actually started in the same instance. That gets included there. So, that is why that Mt plus 1 has to be here. So, okay, if it is this guy is ended here because of I have continuity the other guy also at the end of this, at T it has started. So, because this guy started right at T, the N plus 1th beginning of the N plus 1th is actually happening till T. So, that is what it has to be in this. So, this is, this is a boundary case. Most of the times what will happen? Maybe this guy will hand and that guy will go to the next one. So, in these things it is clear, only when this happens it is happening exactly at time. But we can, by our definition it has to be included here.

(Reference Slide Time: 15:37)



Now, I have defined total cost till time T and now I want to see what is my average cost. So, suppose I am interested in knowing how my cost per unit time changes. So, Ct is what? Ct is the accumulated reward time T and I am dividing it by T that means my average, I am basically looking at my average reward till time T.

So, you can imagine many applications where this will fit in, for example, if you have a, let us say you are running a factory where you have machine. Like machines needs lot of maintenance and you want them to be working because if they are stopping, that means you are stopping your production.

So, what you can do is okay, there are some critical machines in your entire operations which need to be really running. You do not want them to be down at any time. So, you can, what you can do is you see. I repair them and till the time they go back. This critical components, I am still operating when they go bad. Yeah, I am bad shape. But let us assume as soon as they go bad I am going to replace them with the new one and again my operation starts. But during the time they are operating, they are producing something and that production is going to be like a monetary benefit for you.

Let us say while they are operating, they were produce, let us say a 100 pack, 100 items that 100 items is like some benefit reward for you during the time of their operation. Again, when they go down and you replace them next, till the time they go bad again, they would have produced another set of items that is another set of rewards for you. Now, what you want to see is, what is the total items I produced, I am now interest. This is the total items I produce till time T and I am interested in average number of items I produce.

Now, so this is going to kind of give you a sense of based on your maintenance and operations, how the cycle of this operation of your critical component is changing and that is going to affect this. So, let us see. This is the average, so average reward we want to see. So, we want to analyse what is the limit, what is this value is going to be. See, often what happens, you might have seen that, while we are doing all these studies, we happily let T go to infinity, N go to infinity. Why is that? We know that is a kind of observe to do. Letting N go to infinity, "kya ho raha hain", T go to infinity.

So, these are basically because to get some intuition doing just analysis. If you want to find this, what is the value of this at a finite T, it is very hard. We can do that. But that needs more sophistication. This course is not for that. For that you need more machinery with whatever machinery we did you are already bored with that.

So, but with whatever machinery we have, we can only do such analysis, if you want to really understand, okay for every T "kya hota hain", I do not want what happens at N equals to infinity, T equals to infinity. I want it at some finite time. That is really what I am interest.

For that, the current tools, what we study in this course is not enough. You need to go much, much beyond that. So, that is what like with whatever we have and whatever we can say, we can only do this, what we call as limiting regime or asymptotic region. But it is not still bad because they still some kind of intuition what is happening.

So, okay, now comes this theorem Renewal Reward Theorem. So, it says that if, let us say Xn Rn is going to be, is an IID. So, this one we are going to call it as, when we have this, we call it as a renewal reward process. So, then there is a reward associated with the process we will call it renewal reward process is an IID, renewal reward process with equation of R1 is finite and expectation of (())(21:45) and finite.

Then it says t by t goes to, any guess? Yeah, good. You are a good guesser. So, expectation of R1 you said, or R1, okay divided by expectation of X1, "kyu"? When you guess do not ask "kyu". So, that is why we say yes or no, we ask prove or disprove. So, this is a probability 1 and again, this if you are going to look at, so if I take the expectation of this, what is the limit? It is same, c of t this is going to be the same. This is expectation sequence. So, what we are saying? The average reward goes to expectation of R1 divided by expectation of X1.

What is the expectation of R1? This is the reward per cycle expectation of X 1 is this is the expected length of 1 cycle, or basically this is expected length of each cycle because of the IID process. Now, whether this renewal reward theorem implies my earlier elementary renewal theorem?

So, suppose if you take this Rn equals 1, what is this going to be? M of T plus 1. If I Rn equals to 1, if I Rn equals to 1. This is simply an M of T. So, if you, I mean basically what we did is we kind of took this case. CT equals to Mt. CT equals to Mt is what the case we had gotten earlier, that means we earlier we said that in each cycle my reward is 1 we said. But that is not necessary like my reward could be something different.

This is why this is like a more generalization of the elementary renewal theorem we had. So, the generalization here is in each cycle the reward could be not just unit, it could be something else. And also it could be stochastic.

Now, this little result makes sense. That is the average reward is going to be, average reward divided by average cycle length does this makes sense? So, you are saying Ct is what? Total reward accumulated till time T and you are dividing that by total time. So, that is could be same as saying that okay, you just focus on one cycle and one and the reward in that cycle and just look at the average in that cycle.

If you look at it, this is also saying kind of average reward per unit time. The numerator, is that expected reward, denominator the expected time. So, in a sense this is expected reward per unit time and this is also asking that, but this is asking over a total time T. Now, because we are in this process, I could as well focus on one cycle and get the information from that one cycle okay.

Now, if you now want this, apply this kind of results to what I said your plant, manufacturing unit or something. Now, you have this kind of information, average. So, this is like a random quantity. The total reward you got, average total reward you got.

But now you kind of know what is your expected reward per cycle and what is the expected cycle length based on that you can directly get this and if you want more reward, what you could do is, you can, have to design your plant such that you are getting more reward per cycle or you have to design such that your cycle time is smaller.

So, this is all about this renewal process and renewal reward theorem. I wanted to say, so there are other aspects to this which is there in the book like you can read into that. For example, we have defined a process, aging process, and we have defined another process. What it is called. Yeah. We have time residual process and age process.

So, all we can do some analysis on both processes also, we can derive those process also. So, in the book, the properties of those processors are given. So, you should understand and this, those properties one can derive one can based on similar, similar ideas. What I have, we have discussed in the class, you can just read yourself and you will get, so we will stop here.