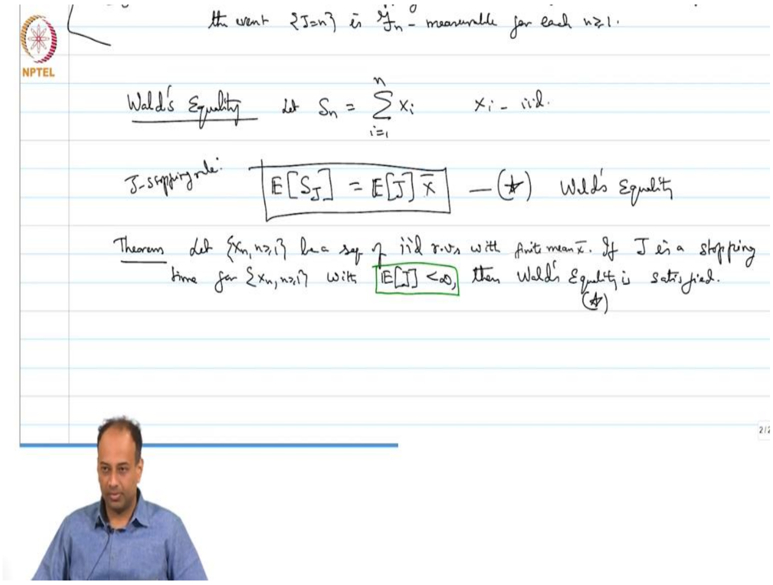
Stochastic Modeling and the Theory of Queues
Prof. Krishna Jagannathan
Department of Electrical Engineering
Indian Institute of Technology - Madras

Module - 4
Lecture - 30
Wald's Equality

(Refer Slide Time: 00:16)



Wald's equality: So, it relates; so, let S n is equal to sum over i equals 1 to n X i. So, these are my sum total winnings until time n, where n is something deterministic. Wald's equality says, what is your expected winning at your stopping time? So, you have decided to stop at some point as a stopping rule. So, J is the stopping rule. What is your expected total winning? So, sum of X 1 + X 2 + ... till X J.

J is where you decided to stop, but this J is now a function of, it is potentially a function of all these X, the X's you have seen. It is not a function of what is going to come. So, you stopped based on some stopping rule. What is the expected amount of winnings you have? This, it relates to your X bar which is the average of X, expected value of X and the expected value of the stopping time, stopping rule.

So, it says something very non-surprising, which says that it is equal to expected J times X bar. And so, I should say this properly, right? Let me call this some star. Theorem: So, this is for IID random variables, X i. So, this star is called Wald's equality. Let X n be a sequence of

IID random variables with finite mean X bar. If J is a stopping time for X n with expectation of J finite, then Wald's equality is satisfied; which means that this guy holds.

So, Wald's equality is star; star is satisfied. So, the key conditions are that J should be a stopping rule for these X n's. And this is very important; we will see; expected J should be finite. So, not only am I saying that you should stop with probability 1; well, if you do not even stop with probability 1, it is not even a stopping rule, it is a defective stopping rule. So, in addition to being a legitimate stopping rule, meaning that you stop with probability 1; so, J is finite with probability 1; in addition to that, you need expected J to be finite.

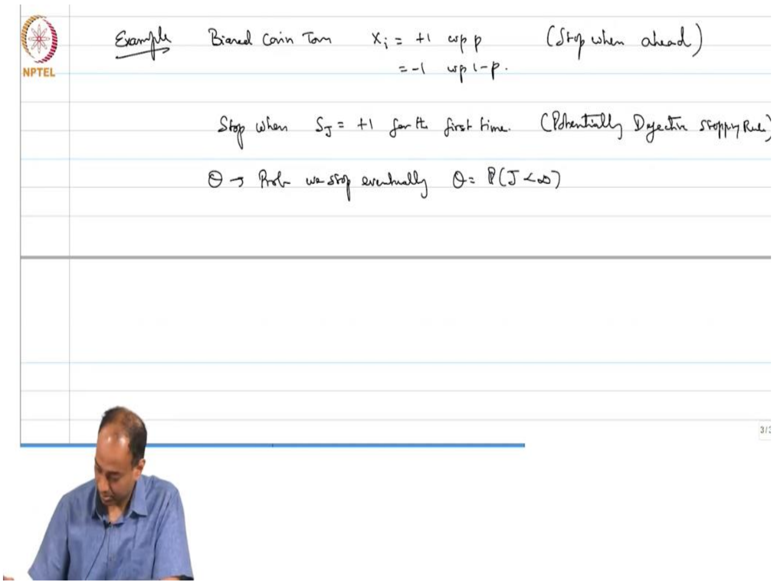
If the expected J is infinite, Wald's equality may not hold; it is a necessary for this theorem to hold. So, I hope you know that a random variable could be finite with probability 1, and the expected value still could be infinite. That is not allowed here. We are asking for more than J just being a legitimate stopping rule. This is what the statement is, okay? We will prove this of course. I just want to remark that; this sort of a thing, you have already seen when this J; see, if the number of terms you are summing is a random variable, some n which is independent of the X i's, then expected S n is equal to expected n times expected X.

This you have already seen in basic probability. That you already know. You can just take iterated expectations and prove it. Here, this is a; a similar thing holds also for stopping rules. What is the issue here? That n or here what we call J is not independent of the X n's; is dependent on the X n's, but it is dependent in a stopping rule fashion. So, essentially, it says that, if you have a stopping rule, your expected winning is simply, you are not really beating the system.

You are just saying that it is, your average winning in each play times the expected number of plays. That is all this is saying. X bar is your average winning in each play; expected J is the number of plays; so, that is your total expected winning. That is what this is saying. So, this would not be true for example, if you say that, I am going to stop based on whether my next few returns are going to be bad or going to be good.

Then you can gain more or lose more depending on what you do. Then, this kind of a relationship will not be true. So, before we prove Wald's equality, let me give you one example.

(Refer Slide Time: 06:35)



Let us say that you have a biased coin tossing or a biased random walk. So, X i = +1 with probability p, and -1 with probability 1 - p; and you are looking at, the stopping rule is this, stop when S J equals +1 for the first time. So, here is a gambler who could win with probability p, 1 rupee, or lose with probability 1 - p, 1 rupee. And the first time he gains a rupee, he is happy. He will stop the first time he gains a rupee.

So, he could stop in the first trial itself. If X 1 turns out to be +1, he will happily go back. And if S 1 turns out to be -1, then I hope that I will build up to 0 and then go to +1, and then I will stop. So, this is the stopping rule. Moment I am ahead; so, this is called stop when ahead. **"Professor - student conversation starts"** Good question. So, the question is, is this even a legitimate stopping rule? Well, it satisfies the stopping rule condition that you do not look ahead, right?

You look at all your X i's; if the sum of your X i's so far is +1, the first time you get to +1, you stop. So, you are not looking ahead. That much is clear. But is it finite with probability 1? That is not clear. It could be defective. So, this stopping rule could be potentially defective. So, expectation of what will be finite? Expectation of S J may be finite. See, S J is +1, right? Whenever you stop, S J is +1. I do not know that, right?

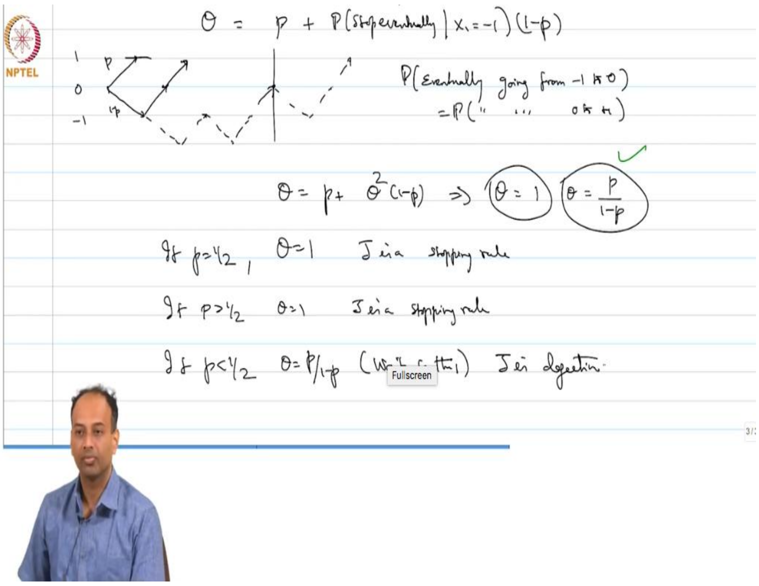
It is not, right? J is not, it is not so simple. See, you are not waiting for the first success. See, the question is saying that, is J simply your geometric random variable where you are waiting for the first +1? I am not waiting for the first +1, I am waiting for my first gain of 1 rupee.

See, there is no use if I lose 2 rupees and gain a rupee; I do not stop. I should be ahead by 1 rupee. I start with 0, let us say; I do not have any money. The moment I gain 1 rupee I leave.

I can borrow, let us say; I do not have any money. If I lose money, I lose money; it is not like you are actually paying. So, it is not a geometric random variable, not at all. So, it is much more complicated. **"Professor - student conversation ends"** So, you can look at this, you can easily analyse this using a Markov chain, but we have not done Markov chains. So, we will use a trick, we will use some recursive trick here.

So, let me actually; first let me argue this out more intuitively. See, I stop if my first X i is 1; I win 1 rupee, I am out. So, if theta is the probability that we stop eventually; I guess, so, then theta is simply the probability that J is finite. So, if theta is equal to 1, it is a stopping rule. If theta is something less than 1, it is not a stopping rule, it is a defective stopping rule. So, let me argue this out. It will turn out that; we will see how this theta behaves. This theta is the probability that we stop eventually.

(Refer Slide Time: 11:21)



So, using the total probability rule, I can write, the probability that I stop eventually as the probability that I stop in the first trial itself. So, what is the probability that I stop in the first trial, given that I got +1? The probability that I stop in the first trial given that I got +1 is 1, times the probability of getting a +1. So, that will be p, plus probably that I stop eventually, given that I got -1 times probability that I got -1; which is simply; got it.

So, theta is a probability that I stop eventually. So, I am splitting this into 2 disjoint event. I stop; so, X 1 has to be +1 or -1. Given that this X 1 is 1, I stop eventually. If X 1 is -1, what is the probability that I stop eventually? I write like this. So, it is like this, right? So, you can look at this as a random walk. So, if this is 0; if I hit +1, I stop; finished. So, this has probability p.

If I get a -1 with probability 1 - p, then I have to, from -1, I have to eventually find my way to +1. See, I do not have to go; I can go like that or I can do a lot of things now. So, basically, I have to go eventually from -1 to 0, and then go from 0 to 1. So, -1 to 0, I may not go like this, I may do; right; then, I may go like that. All these are possible. Once I get to 0, again I have to get to +1.

See, but the key issue here is, because these X i's are IID, the probability of going eventually from -1 to 0 is the same as eventually going from 0 to 1. See, if you eventually go like this from -1 to 0, those exact sample paths will also go from 0 to 1, if you start at 1. And once you reach 0, again you have to eventually get to 1. So, my point is that, whatever happens here and whatever happens here, are basically statistically identical.

Going from -1 to 0 and 0 to 1 are statistically identical, because of the IID nature of these random variables. So, bottom line is that, probability that you stop eventually given you ended up negative, is like going from -1 to 0 and then going from 0 to 1. And the probability of that; so, what is the probability of eventually going from 0 to 1? It is the original probability, right? It is theta. Probability of eventually going from 0 to 1 is theta.

So, similarly, the probability of eventually going from -1 to 0 is also theta. So, I can write this, probability of going from -1 to 0 is equal to probability of eventually going from 0 to +1. This you will agree. It does not matter where I start, right? You just have to take 1 step, and that is your theta. So, for probability of stopping eventually, given that you start at -1, the probability that you eventually end up at +1 is simply theta times theta.

So, I can write, theta is equal to p plus theta square times 1 - p. And because they are independent; you go from -1 to 0 and then independently go from 0 to 1; and these are the same probability. You solve this, what do you get? This implies theta equal to; see, theta is equal to 1 is clearly a solution; it is a quadratic in theta, right? Theta equal to 1 is clearly a solution, and theta equal to p over 1 - p is another solution, right?

So, the question is, why is the probability of stopping eventually given X 1 equal to -1? So, if X 1 is -1, I am at -1, right? I want to go from -1 to 1 eventually. In order to go from -1 to 1, I should eventually go from -1 to 0 and then eventually go from 0 to 1. Probability of eventually going from 0 to 1 is my original theta. And because of the IID nature, probability of going from -1 to 0 is also statistically the same; it is theta.

And these are independent, so I multiply. So, what is the issue here? So, if p is equal to half, then theta equal 1 is the only solution. So, if p is equal to half, so, the cases are as follows: So, if p equal to half, then theta equal to 1, is the only solution. So, J is a stopping rule. If p is greater than half, then what happens? Theta equal to 1 is the only solution. So, then, J is a stopping rule; because, p greater than half, the second solution is not even a probability; you can just throw it away.

So, if p less than half, we seem to have 2 solutions, which are valid probabilities 1 and p over 1 - p. So, if p is like one-third, then p by 1 - p will be half. So, the question is, then, is theta equal to 1 correct or theta equal to p over 1 - p correct? It turns out; this requires some Markov chains; that this is the correct solution. If p is less than half, then we can show that theta equal to p over 1 - p is the correct solution. Wait for this.

We can show this; we can use birthrate Markov chains to show this. Then, J is defective. So, you stop only with probability p by 1 - p, if p is less than half; there is a positive probability that you do not stop. Now, we are almost out of time; but we can use Wald if you can establish what is expectation of J, is it finite or infinite? Then, we can go ahead and use Wald. So, we will continue this in the next class. Next class, we will also prove Wald's theorem, Wald's equality. Okay? Thank you. So, we will continue this tomorrow.