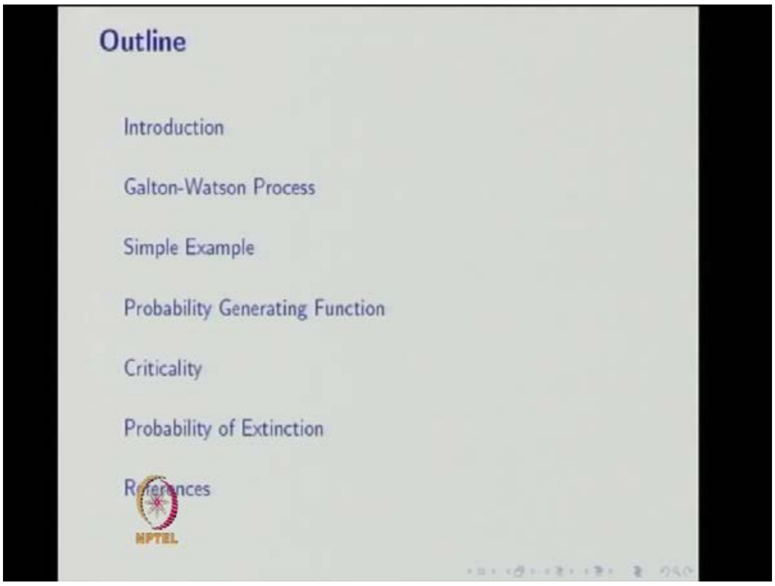
Stochastic Processes
Prof. Dr. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Module - 9
Branching Processes
Lecture - 1
Galton-Watson Process

This is stochastic processes, module nine, branching processes. In the module one we have discussed review of probability; module two, we have introduced stochastic process; module three, we have discussed stationary processes; module four, we have discussed discrete time Markov chain; module five, we have discussed continuous time Markov chain; module six, we have discussed the martingale; module seven; we have discussed about Brownian motion or wiener process; module eight, we have discussed renewal process.

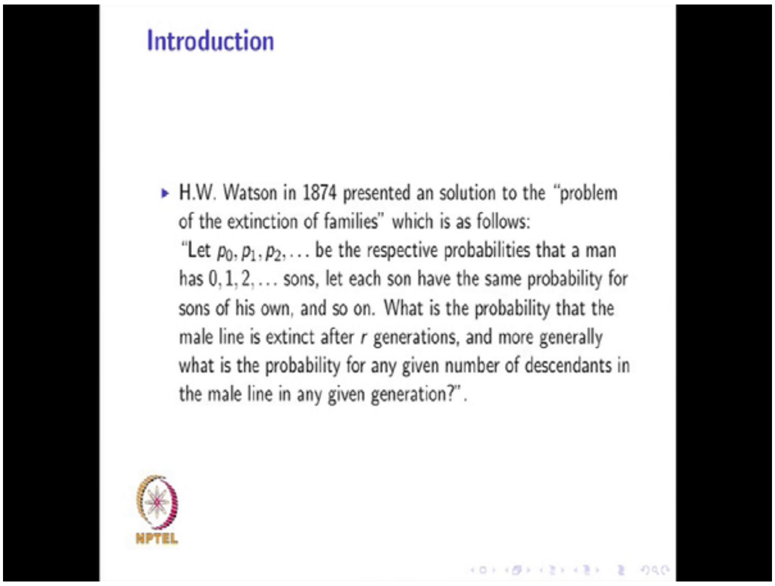
In this module we are going to discuss the branching processes. It covers two lectures, in these two lectures we are going to cover the definition and examples of branching processes. Two important branching processes, Galton-Watson branching process and that Markov branching process are going to be discussed. Important measures like perform probability of extinction, limit theorem, limit distribution and mean of branching process will be discussed.

(Refer Slide Time: 02:15)



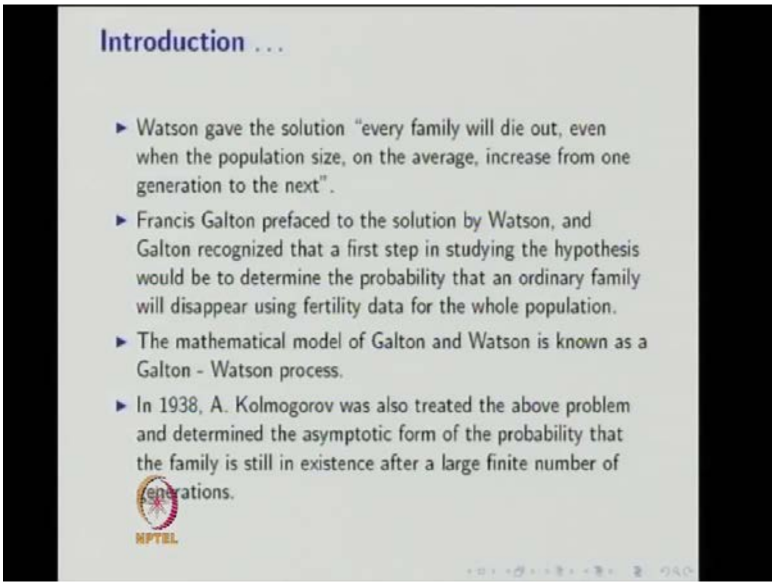
In the lecture 1 we are going to start with the introduction to branching processes, followed by that we are going to start the Galton-Watson process, few examples will be discussed. Then, we are going to discuss the probability generating function of Galton-Watson branching process. Then, we are going to discuss the criticality and probability of extinction.

(Refer Slide Time: 02:44)



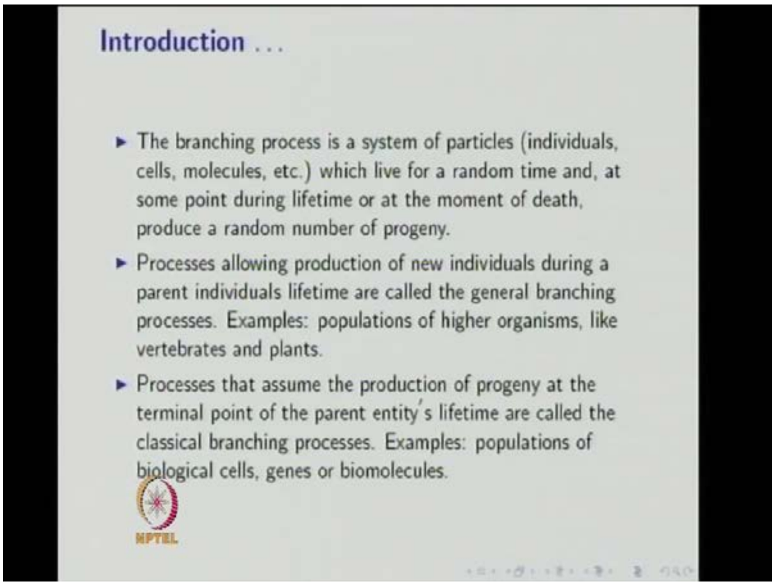
Watson in 1974 presented a solution to the problem of extinction of families, which is as follows. Let p naught, p 1, p 2 be the respective probabilities, that the man has 0, 1, 2 and so on, sons. Let each son have the same probability for sons of his own and so on. What is the probability, that the male line is extinct after r generations and more generally, what is the probability for any given number of descendants in the male line in any given generation?

(Refer Slide Time: 03:27)



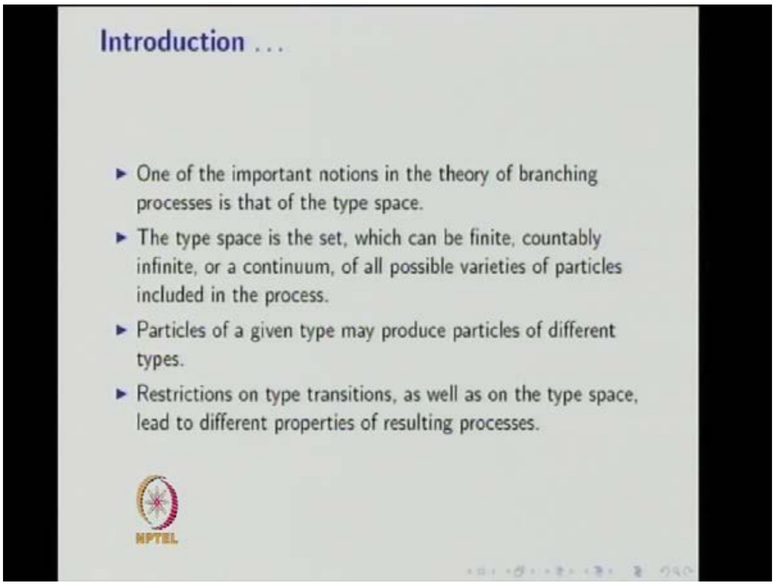
Watson gave the solution, every family will die out, even when the population size, on the average, increase from one generation to the next. Galton prefaced to the solution by Watson and Galton recognized that the first step in studying the hypothesis would be to determine the probability, that the ordinary family will disappear using fertility data for the whole population. The mathematical model of Galton and Watson is known as a Galton-Watson branching process. In 1938, Kolmogorov was also treated the above problem and determined the asymptotic form of the probability, that the family is still in existence after a large finite number of generations.

(Refer Slide Time: 04:32)



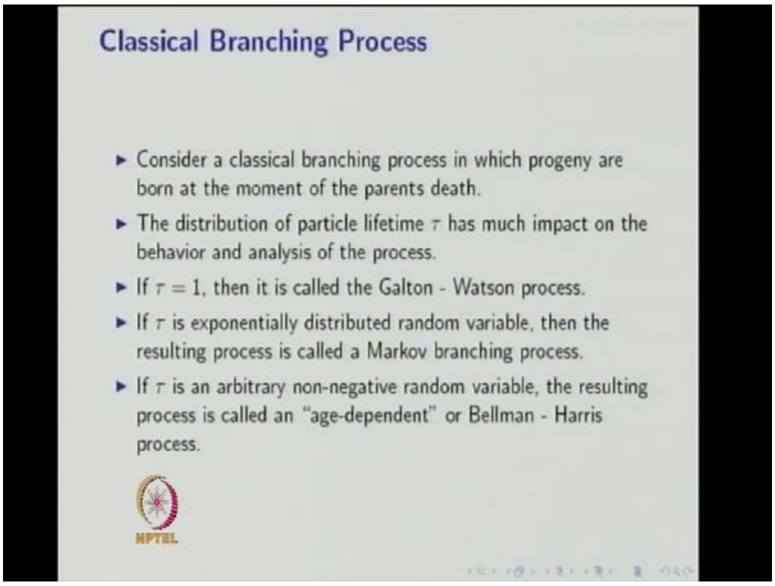
The branching process is a system of particles or individuals or cells or molecules, which live for a random time and, at some point during lifetime or at the moment of death, produce a, random number of progency, random number of progeny. Processes allowing production of new individuals during a parent individual's lifetime are called the general branching processes. Examples: populations of higher organisms like vertebrates and plants. Processes that assume production of progeny at the terminal point of the parent entity's lifetime are called the classical branching processes. Examples: population of biological cells, genes or biomolecules.

(Refer Slide Time: 05:47)



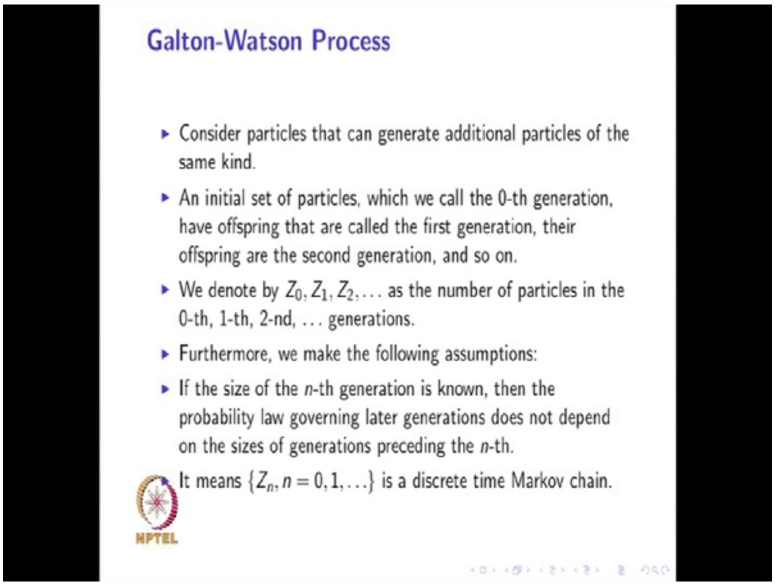
One of the important notation; one of the important notions in the theory of branching processes is that of the type space. The type space is the set, which can be finite, countably infinite or continuum, of all possible varieties of particles including in the process. Particles of a given type may produce particles of different types. Restrictions on the type transitions, as well as on the type space, lead to different properties of resulting processes.

(Refer Slide Time: 06:30)



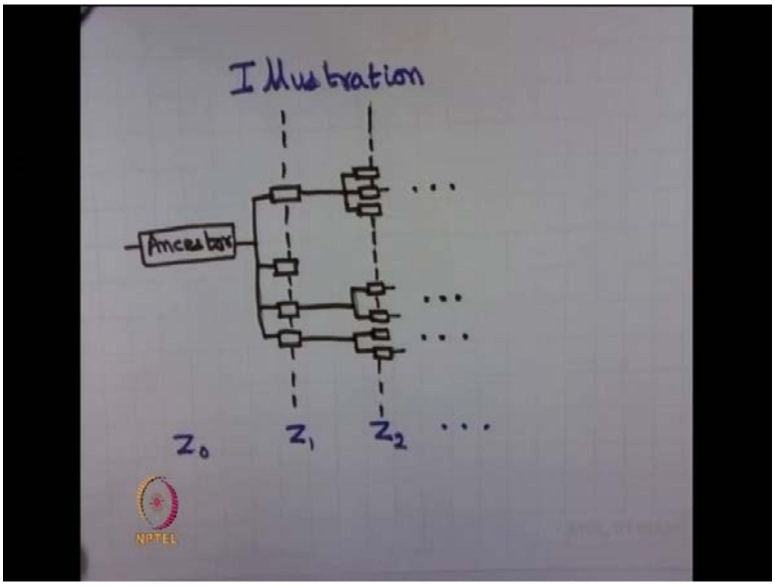
Now, we are going to discuss the classical branching processes. Consider a classical branching process in which progeny are born at the moment of parent's death. The distribution of particle lifetime tau has much impact on the behavior and analysis of the, of the price, of the process. If tau is equal to 1, then it is called Galton-Watson process. If tau is exponentially distributed random variable, then the resulting process is called Markov branching process. In this model we are going to discuss in detail Galton-Watson process and Markov branching process. If tau is arbitrary non-negative random variable, the resulting process is called age-dependent or Bellman-Harris process.

(Refer Slide Time: 07:36)



Now, we are going to discuss Galton-Watson process. In the next lecture, lecture 2, we are going to discuss Markov branching process. Consider particles, that generate, that can generate additional particles of the same kind. An initial set of particles, which we call the 0th generation having offspring, that are called 1st generation, their offsprings are the 2nd generation and so on. We denote by Z naught, Z 1, Z 2, as the number of particles in the 0th, 1st, 2nd generations. We see the illustration, this is Z naught, Z 1, Z 2 and so on.

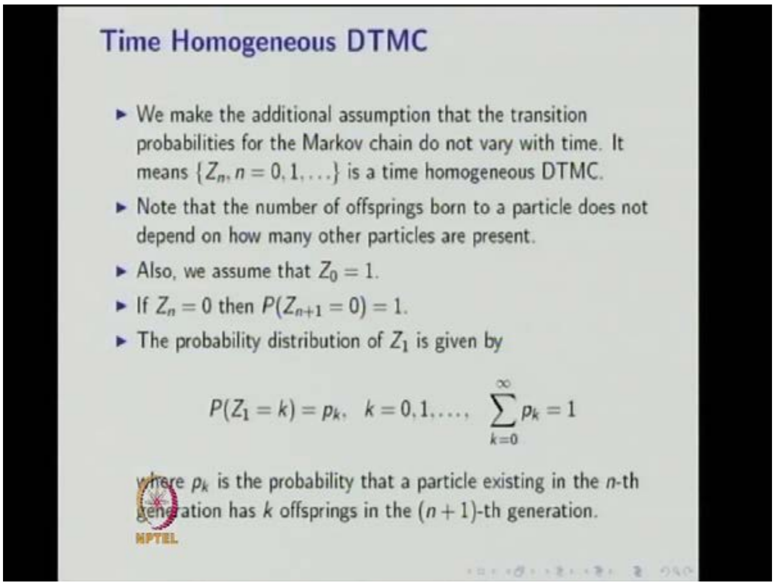
(Refer Slide Time: 08:42)



So, this is the illustration of the number of particles in the 1st generation, 0th generation, 1st generation, 2nd generation and so on.

Furthermore, we make the following assumptions. If the size of the nth generation is known, then the probability law governing later generations does not depend on the sizes of generations preceding the nth. It means the sequence of random variables Z n form a discrete time Markov chain.

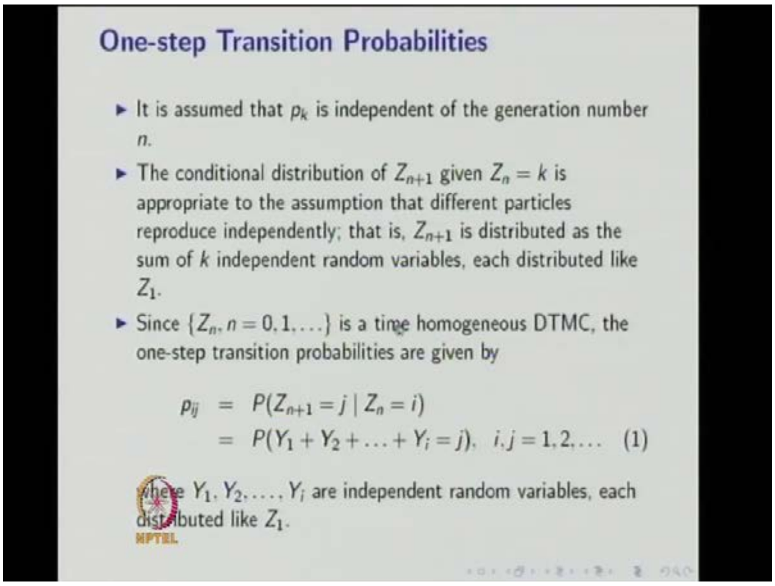
(Refer Slide Time: 09:20)



We make additional assumption that the transition probabilities for the Markov chain do not vary with time. It means, the sequence of random variables Z n is a time homogenous discrete time Markov chain. We have discussed the discrete time Markov chain in the module 4 in detail. Note that the number of offspring's born to a particle does not depend on how many other particles are present.

Also, we assume, that Z naught is equal to 1. In the illustration also we made it Z naught is equal to 1. Z 1 is a random number of particles; Z 2 is a random number of particles and so on. If Z n is equal to 0, then the probability of Z n plus 1 is equal to 0 is equal to 1. The probability distribution of Z 1 is given by the probability of Z 1 takes a value k, that is nothing but in, in notation p suffix k and the summation of p ks is equal to 1, where p k is the probability, that a particle existing in the nth generation has k offsprings in the n plus 1th generation. So, this is the probability mass function for the random variable Z 1.

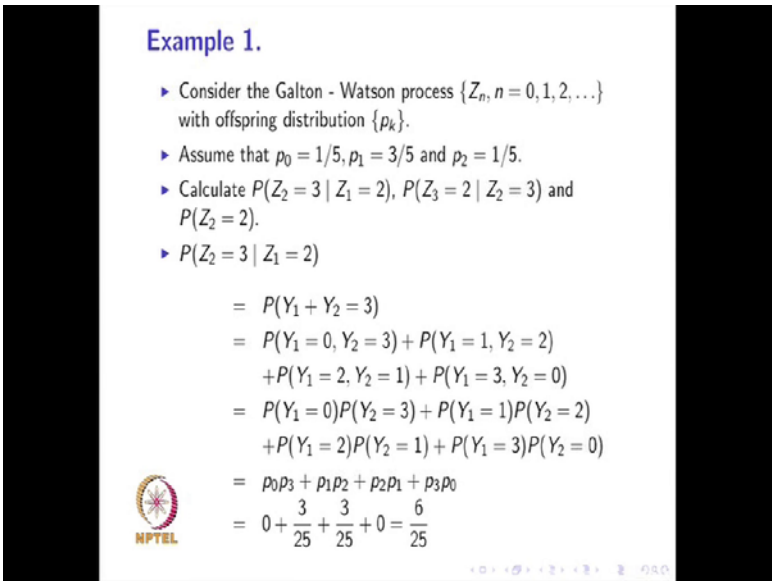
(Refer Slide Time: 11:04)



It is assumed, that p k is independent of the generation number n. The conditional distribution of Z n plus 1 given Z n is equal to k, is appropriate to the assumption, that the different particles are reproduce independent, that is Z n plus 1 is distributed as the sum of k independent random variables, each distributed like Z 1.

Since the sequence of random variables Z n is a time homogenous discrete time Markov chain, the one step transition probabilities are given by p suffix (i, j) that is nothing but the probability, that Z n plus 1 is equal to j, given Z n is equal to I, that is same as Y 1 plus Y 2, plus, and so on plus Y i, that takes a value j for (i, j) belonging to 1, 2 and so on, where Y i’s are i.i.d. random variables, each distributed like Z 1. So, the illustration for the Galton-Watson process is Z naught is equal to 1, Z 1 here it is 4 and Z 2 is 7 and so on.

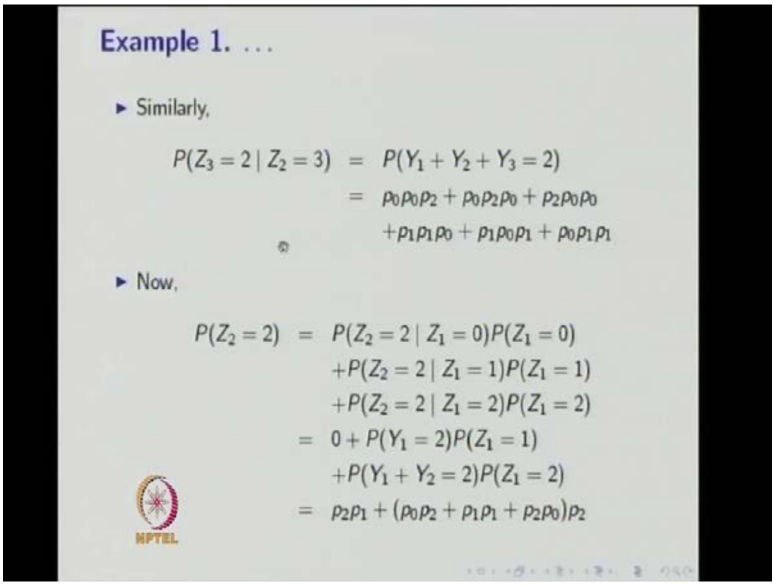
(Refer Slide Time: 12:51)



Now, let us consider a simple example. Consider the Galton-Watson process Z n with offspring distribution p k. Assume, that p naught is equal to 1 by 5, p 1 is equal to 3 by 5 and p 2 equal to 1 by 5. So, this is the probability mass function for the random variable Z 1. Our interest is to find out the conditional probability p. Probability of Z 2 is equal to 3, given Z 1 was 2 and also probability, that Z 2 is equal to 2, given Z 2 was 3 and also, probability of Z 2 is equal to 2. So, the conditional probability p Z 2 is equal to 3 given Z 1 is equal to 2, that is same as probability of Y 1 plus Y 2 equal to 3. That is possible, either Y 1 is equal to 0 or Y 2 equal to 3 or Y 1 is equal to 1, Y 2 is equal to 2 or Y 1 is equal to 2 or, and Y 2 is equal to 1 or Y 1 is equal to 3 and Y 2 is equal to 0.

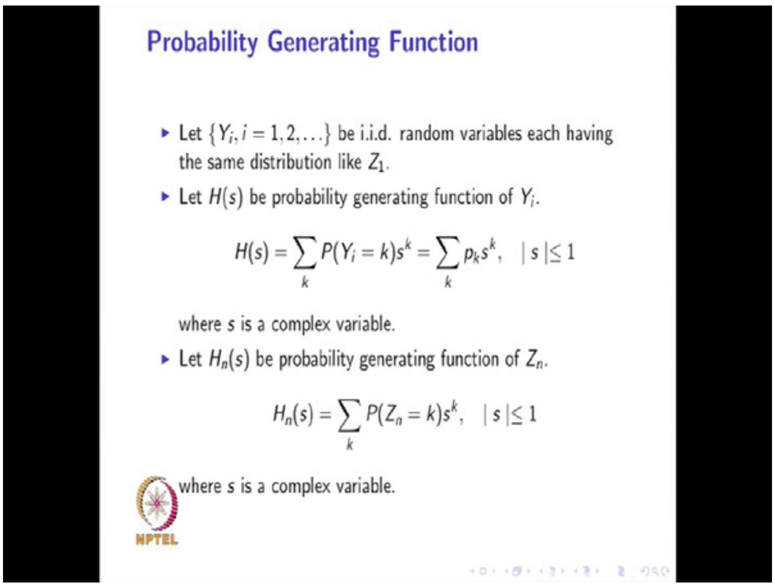
Since Y i's are i.i.d. random variables, you can write down this as a probability of Y 1 is equal to 0 into probability of Y 2 is equal to 3 and so on. Probability of Y 1 is equal to 0, that is nothing but p naught. Probability of Y 2 is equal to 3 that is nothing but p 3. Similarly, the second expression is p 1 into p 2; third one is p 2 into p 1; the last one is p 3 p naught. Since p 3 is equal to 0, the first term and the last term will be 0. So, you will get p 1 p 2 plus p 2 p 1 that is same as 6 by 25. So, this is the conditional probability of p of Z 2 is equal to 3 given Z 1 is equal to 2.

(Refer Slide Time: 15:09)



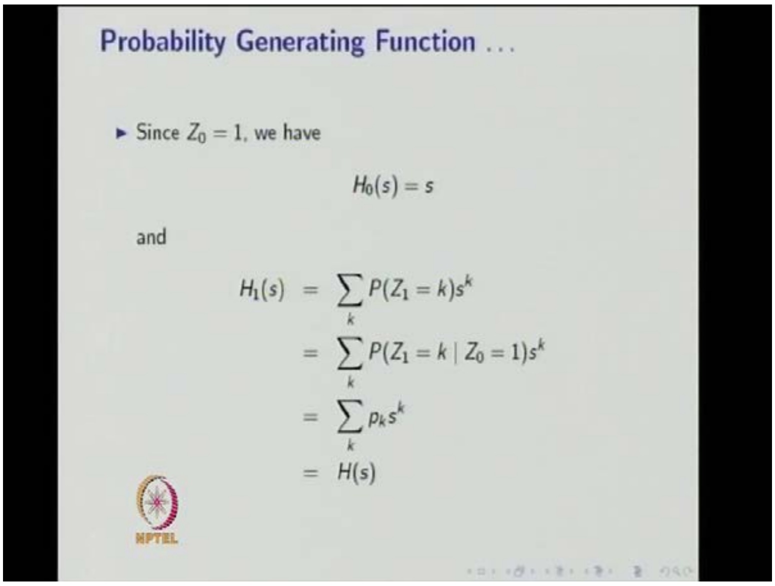
Similarly, one can find P of Z 3 is equal to 2 given Z 2 is equal to 3, that is possible when Y 1 plus Y 2 plus Y 3 is equal to 2. So, the probability of Y 1 plus Y 2 plus Y 3 is equal to 2; we have six possibilities. Substitute the value of p naught, p 1, p 2, you will get the numerical value of this conditional probability. Similarly, one can find the probability of Z 2 is equal to 2 also. Probability of Z 2 is equal to 2, same as probability of Z 2 is equal to 2 given Z 1 is equal to 0 multiplied by probability of Z 1 is equal to 0 plus the combination with the probability of Z 1 is equal to 1, probability of Z 1 is equal to 2. Substitute the values, then you will get the probability of Z 2 is equal to 2.

(Refer Slide Time: 16:12)



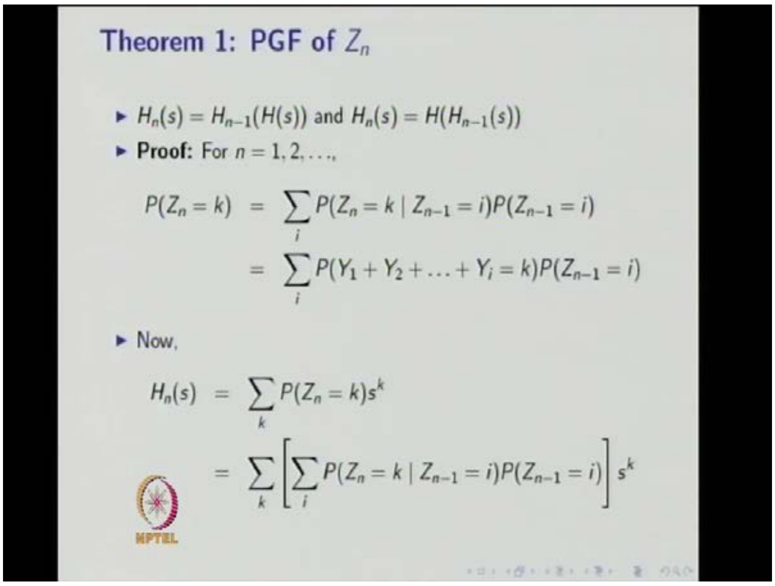
Now, we are going to discuss the probability generating function for the branching process. Let Y i's be i.i.d. random variables, each having the same distribution like Z 1. Let H of s be the probability generating function of Y i's; let H n of s be the probability generating function of Z n.

(Refer Slide Time: 16:47)



Since Z naught is equal to 1, you will get H naught of s is same as s. Now, our interest is to find out the probability generating function for Z 1 that is same as the probability generating function of Y i's. The probability generating function of Y i's is H of s, therefore H 1 of s is same as H of S.

(Refer Slide Time: 17:17)

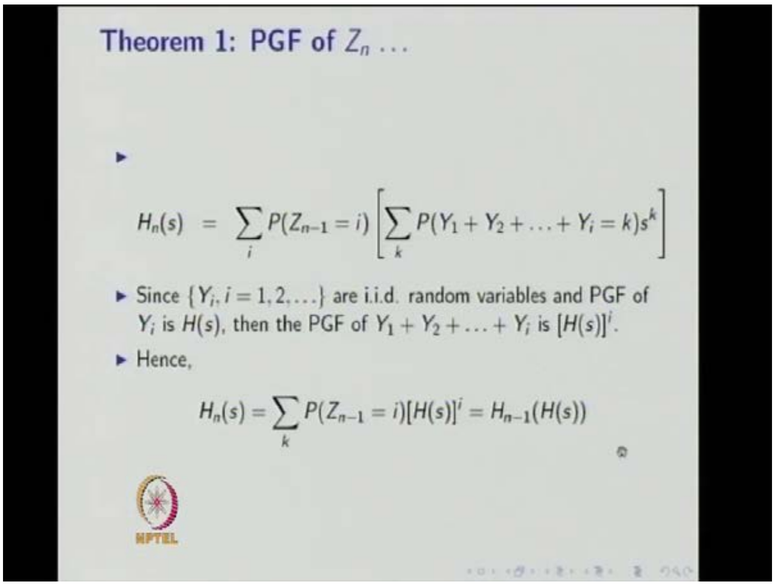


Our interest is to find out the probability generating function for Z i, Z n, where n is 1, 2 3 and so on. This we are going to give it as a theorem, H n of s. This is nothing but the probability generating function for the random variable Z n, is same as H of n minus 1 of H of s and H n of s also can be written in the form of H of H n minus 1 of s. Let us see the proof of this.

We know that probability of Z n is equal to k that is nothing but summation over i's. Probability of Z n is equal to k, given Z n minus 1 is equal to i multiplied by probability of Z n minus 1 is equal to i. We know, that this condition probability is nothing but probability of Y 1 plus Y 2 and so on plus Y i is equal to k multiplied by probability of Z n minus 1 is equal to i.

Now, we will go for finding out the probability generating function for the random variable Z n. Substitute probability of Z n is equal to k from above in this equation. Therefore, you will have summation of k, substitute the probability of Z n is equal to k, that is nothing but summation over i. Probability of a, conditional probability multiplied by probability of Z n minus 1 is equal to i multiplied by s power k.

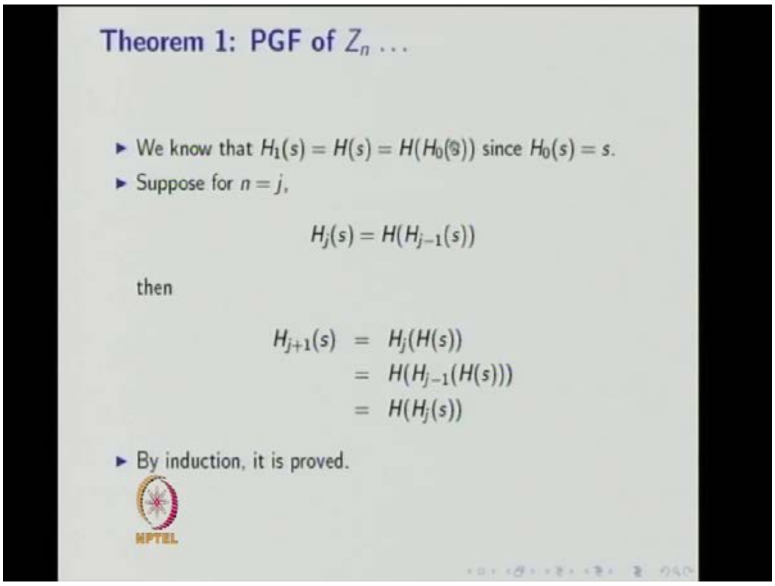
(Refer Slide Time: 19:09)



Substitute, the conditional probability is nothing but the probability of Y 1 plus Y 2 plus so on. Y i is equal to k into s power k. Since Y i's are i.i.d. random variables and the probability generating function of Y i's are nothing but H i, sorry, H of s. The probability generating function of sum of i random variables Y i's, that is nothing but H of s whole power i because of Y i's are i.i.d random variables and probability generating function of Y i's is equal to H of s. Therefore, the probability generating function of Y 1 plus Y 2 and so on plus Y i is H of s whole power i.

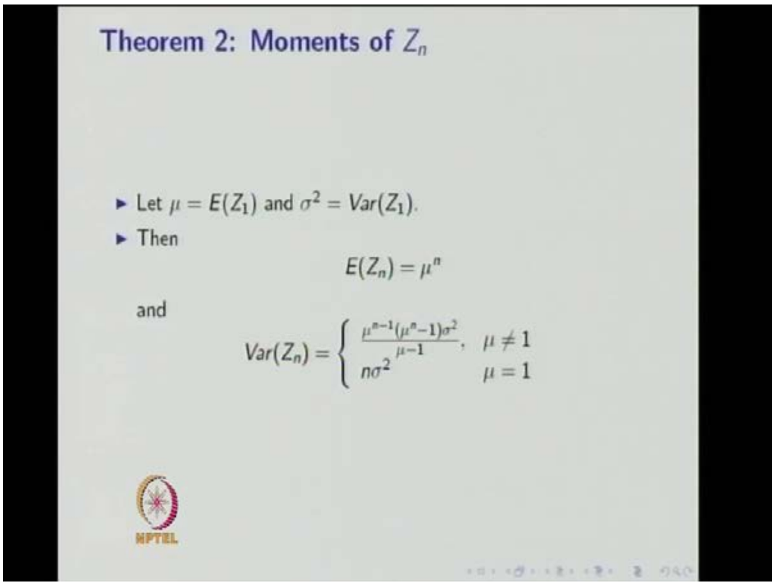
Therefore, substitute, this is nothing but the probability generating function of Y 1 plus Y 2 plus and so on plus Y i. Therefore, H n of s is nothing but summation over k H n of s is nothing but summation over i probability of Z n minus 1 is equal to i times H of s whole power I, replace this by H of s whole power i, therefore summation i probability of Z n minus 1 is equal to i H of s whole power i. So, that is nothing, but the probability generating function for the random variable Z n minus 1 with the replacement s by H of s. Therefore, H n of s is nothing, but H n minus 1 of H of s.

(Refer Slide Time: 21:16)



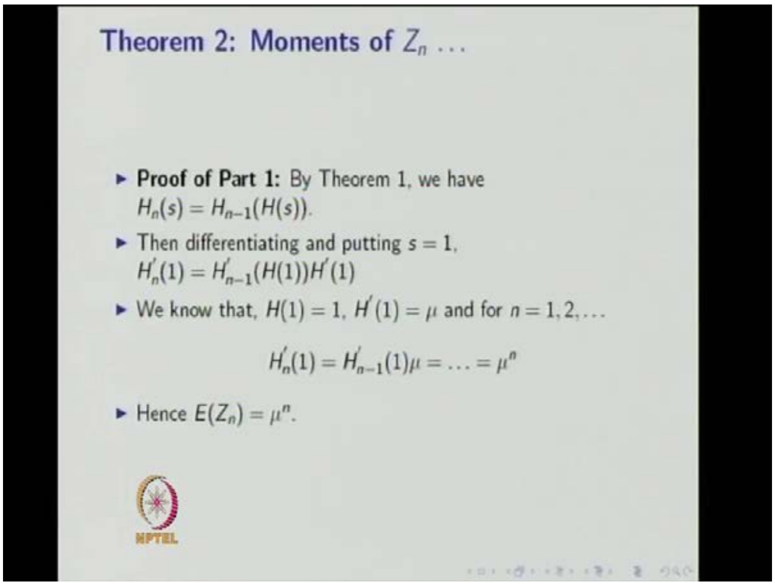
So, the first part is proved. H n of s is same as H n minus 1 of H of s. Now, we are going to prove the second part. We know, that H 1 of s is H of s, that is same as H of H naught of s because H naught of s is nothing, but s. Suppose this is true for n is equal to j, that means, H, H of j of s is H of H of j minus 1 time of s. Then, we can find out what is H of j plus 1 of s? That is nothing, but H j of H of s that is same as H of H j minus 1 times H of s that is same as H of H j of s. By induction it is proved.

(Refer Slide Time: 22:02)



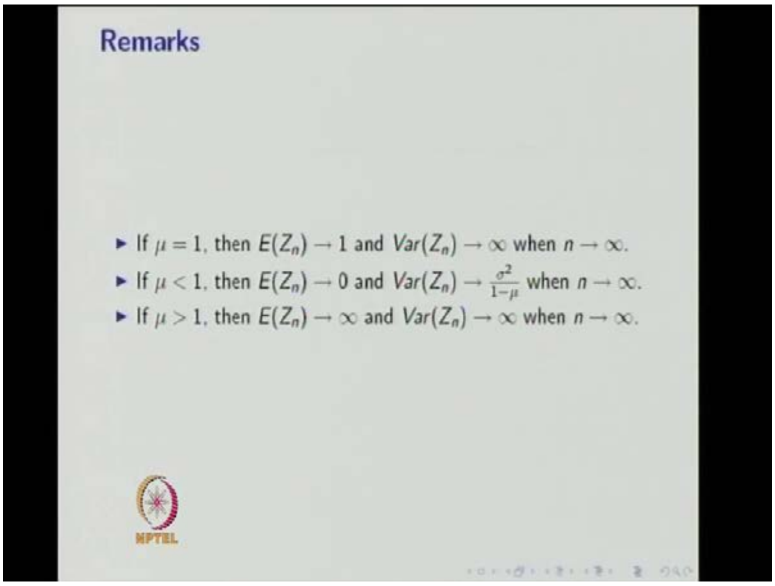
Now, we are finding the moments of Z n. Let mu is equal to expectation of Z 1 and sigma square is nothing but the variance of Z 1. Then, expectation Z of n, that is mu power n and the variance of Z n will be for mu equal to 1. It is n times sigma square for mu is not equal to 1. It will be mu power n minus 1 times mu n minus 1 times sigma square divided by mu minus 1.

(Refer Slide Time: 22:43)



We prove the part 1, that is, expectation of Z n is equal to mu power n. By theorem 1, you, we know that H n of s is H n minus 1 of H of s. Then, differentiating and putting s equal to 1 we will, we will get H n dash of 1 that is same as H n minus 1 dash of H of 1 into H dash of 1. We know, that H 1 H of 1 is 1 H dash of 1 is mu. Therefore, for n, n is equal to 1, 2 and so on. We will get H n dash of 1 is same as H n minus 1 dash of 1 into mu that is same as H n minus 2 dash of 1 into mu square and so on. Therefore, you will get mu power n because you know that H dash of 1 is mu. By recursively you will get, H n dash of 1 is mu power n. Therefore, expectation of Z n is nothing, but H n dash of 1 that is same as (( )).

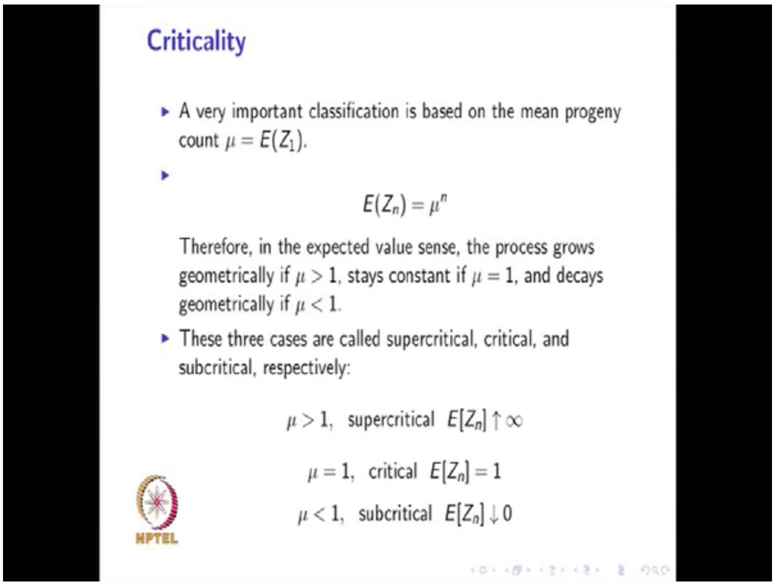
(Refer Slide Time: 24:04)



So, till now we have discussed the probability generating function of Z power n and also, we have discussed mean and variance of Z of n. As a remark, if mu is equal to 1, the expectation, expectation of Z n tends to 1, whereas variance of Z n tends to infinity as n tends to infinity. You can see it from this theorem, as mu tends to 1, expectation of Z n will tends to 1, whereas the variance Z n tends to infinity because that is same as n times sigma square. Therefore, as mu tends to 1, the variance of Z n tends to infinity as n tends to infinity.

Whereas if mu is less than 1, the expectation of Z n tends to 0 and variance of Z n will be sigma square divided by 1 minus mu as n tends to infinity. Similarly, if mu is greater than 1, then the expectation of Z n will tends to infinity and the variance of Z n tends to infinity as n tends to infinity. This also you see it from the theorem.

(Refer Slide Time: 25:31)

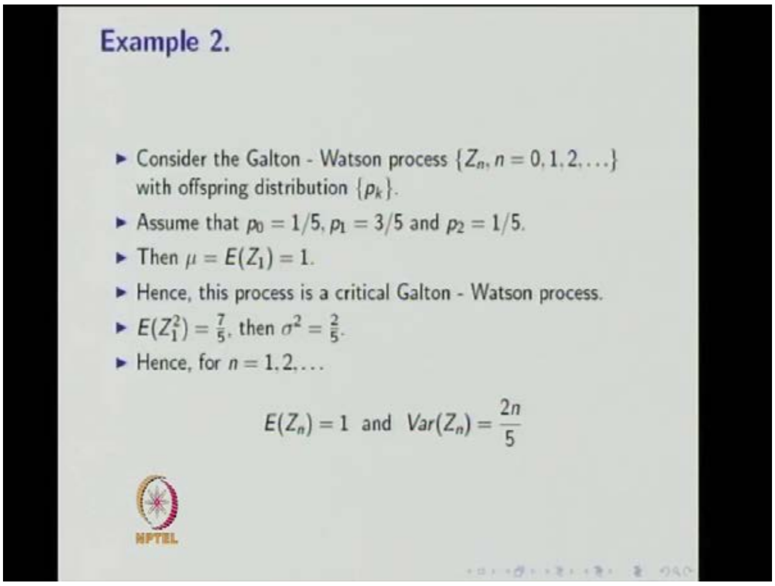


Now, we are going to discuss the criticality. A very important classification is based on mean, progency, progeny count mu is equal to expectation of Z 1; a very important classification is based on the mean progeny count mu is equal to expectation of Z 1.

You know, that expectation of Z n is equal to mu power n, just now we have proved it in the theorem 1. Therefore, in the expected value sense the process grows geometrically if mu is greater than 1, stays constant if mu is equal to 1 and decays geometrically if mu is less than 1. From the expectation of Z n is equal to mu power n, you can conclude if mu is greater than 1. The process grows geometrically if mu is equal to 1, then the process stays constant, whereas if mu is less than 1, the process decays geometrically. Thus three cases are called supercritical, critical and subcritical respectively.

That means, if mu is greater than 1, then the process is called supercritical, in this case the expectation of Z n tends to infinity. The mu is equal to 1, the process is called critical and the expectation of Z n is equal to 1 as n tends to infinity. When mu is less than 1, the process is called subcritical and expectation of Z n will tends to 0 as n tends to infinity.

(Refer Slide Time: 27:42)

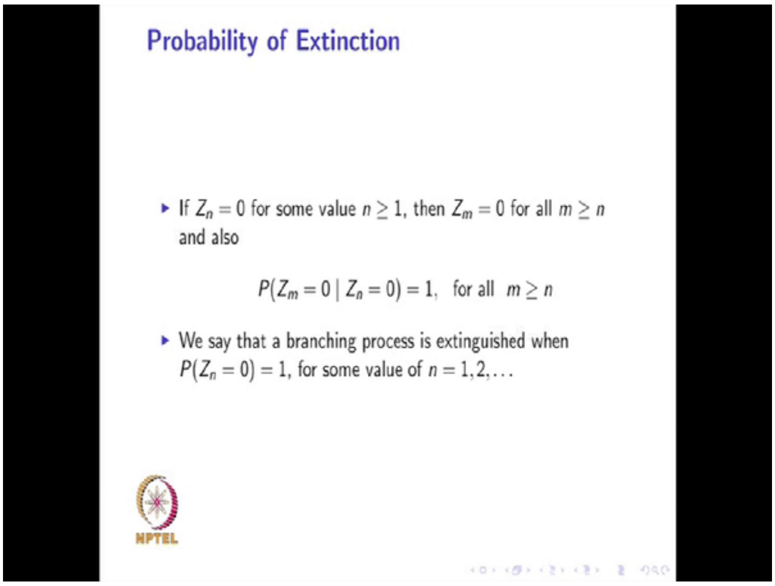


Now, we are going to consider the second example. Consider the Galton-Watson process Z n with offspring distribution p k, which is the same problem, example 1, with the assumption p naught is equal to 1 by 5, p 1 is equal to 3 by 5 and p 2 is equal to 1 by 5.

Now, you can find out the mean of Z 1, that is nothing but 1 because p naught is equal to 1 by 5, p 1 is equal to 3 by 5, p 2 is equal to 1 by 5, you will get mean of Z 1 will be 1. Hence this process is called critical Galton-Watson process because mu is equal to 1. We can find out the variance of Z 1 also, first we find out expectation of Z 1 square that will be 7 by 5. Hence variance equal to expectation of Z 1 square minus expectation of Z 1 whole square, that will be 2 by 5.

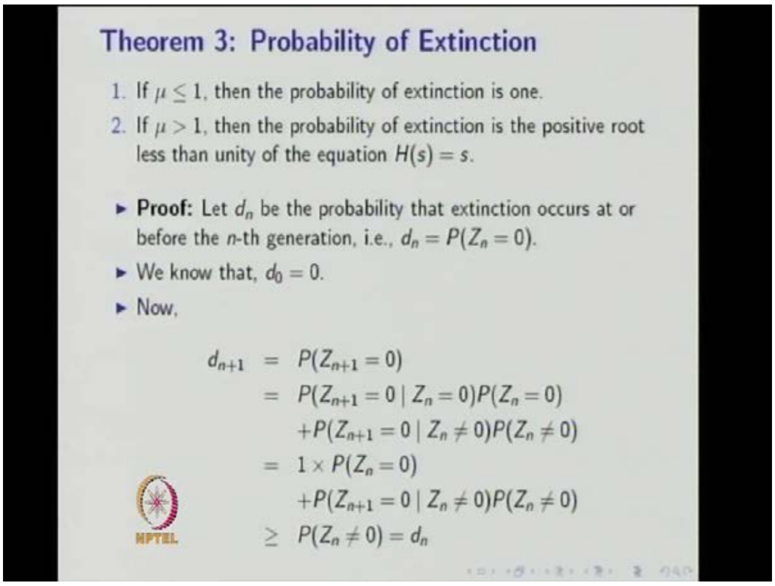
Since you know mu and sigma square using the theorem 1, you can find out the moments of Z n, first and second order moment expectation of Z n and variance of Z n. Since mu is equal to 1 the expectation of Z n will be 1, variance of Z n will be n times sigma square. In this problem the sigma square is 2 by 5, therefore mean of Z n is 1, variance of Z n is 2 n by 5.

(Refer Slide Time: 29:43)



Now, we are going to discuss the very important concept called probability of extinction. If Z n equal to 0 for some value for n is greater than or equal to 1, then Z m will be 0 for all m greater than or equal to n and also the conditional probability of Z m is equal to 0, given Z n is equal to 0, that will be 1 for all m greater than or equal to 1. We say, that the branching process is, extent, extinguished when probability of Z n is equal to 0 will be 1 for some value of n.

(Refer Slide Time: 30:30)



Now, we discuss the probability of extinction in theorem 3. If mu is less than or equal to 1, then the probability of extinction is 1. If mu greater than 1, then the probability of extinction is the positive root less than unity of the equation H of s is equal to s, where H of s is the probability generating function of Z 1.

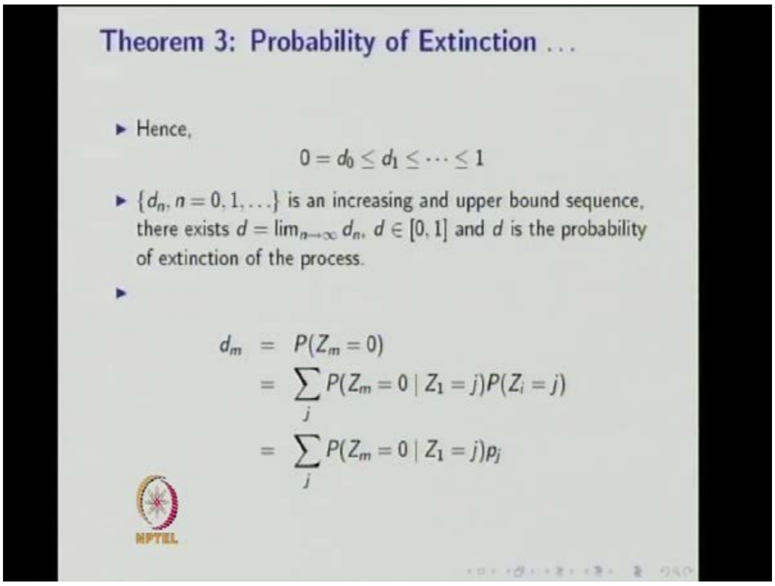
Here we consider a Galton-Watson branching process Z n with offspring distribution p k and p k forms the probability mass function for the random variable Z 1. If mu is less than or equal to 1, then the probability of extinction is 1. If mu is greater than 1, then the probability of extinction is the positive root less than unity of the equation H of s equal to s, that means, we have to solve the equation H of s is equal to s. From that you can get the probability of extinction where mu is the mean of the random variable Z 1.

Let us see the proof. Let d n be the probability, that extinction occurs at or before the nth generation. Hence, d n is nothing but the probability of Z n is equal to 0. We know, that d naught will be 0 because we made the assumption Z naught is equal to 1. Now, we will find out d n plus 1. d n plus 1 is nothing but, by the definition, it is d n plus 1, is nothing but probability of Z n plus 1 equal to 0.

We can write probability of Z n plus 1 equal to 0 using conditional probabilities, that is same as probability of Z n plus 1 is equal to 0, given Z n was 0 multiplied by probability of Z n is equal to 0 plus probability of Z n plus 1 is equal to 0 given Z n is not equal to 0 multiplied by probability of Z n is not equal to 0. We know, that probability of Z n plus 1 is equal to 0 given Z n is equal to 0, that is equal to 1. Therefore, this will be 1 times probability of Z n is equal to 0 plus probability of Z n plus 1 is equal to 0 given Z n is not equal to 0 multiplied by probability of Z n not equal to 0. Obviously, this will be greater than equal or equal to probability of Z n is not equal to 0 because we are adding some probability plus probability multiplied by probability of Z n is not equal to 0. Therefore, this quantity will be greater than or equal to probability of Z n not equal to 0.

This quantity, this quantity will be, this is same as 1 times probability of Z n is equal to 0 plus probability of Z n plus 1 equal to 0 given probability of Z n is not equal to 0 multiplied by probability of Z n is not equal to 0. Obviously, this quantity will be greater than or equal to 0, the second term. Therefore, the whole result will be greater than or equal to probability of Z n equal to 0. You know, that probability of Z n is equal to 0 is nothing but d n, therefore d n plus 1 will be greater than or equal to d n. This is true for n is equal to 1, 2 and so on.

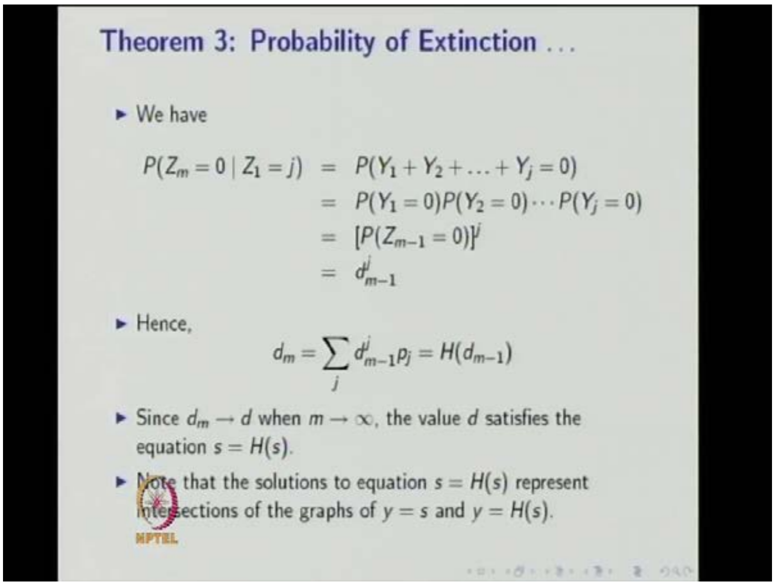
(Refer Slide Time: 34:59)



Hence, d naught will be 0. d naught is less than or equal to d 1, d 1 will be less than or equal to d 2 and so on. And since d i's are the probability of extinction, therefore that will be less than or equal to 1. Hence, d n is a sequence and upper bound sequence. d n is an increasing and upper bound sequence, there exists d, that is nothing but the limit n tends to infinity of d n and the d is belonging to the closed interval 0 to 1 and d is the probability of extinction of the process.

Now, we will find out what is the d? You know, that d m is nothing but probability of Z m equal to 0, that is nothing but summation over j probability of Z m is equal to 0 given Z 1 is equal to j multiplied by probability of Z 1 is equal to j. You know, that probability of Z 1 is equal to j is nothing but p j. Therefore, the d m is nothing but summation over j probability of Z m is equal to 0 given Z 1 is equal to g j multiplied by p j.

(Refer Slide Time: 36:27)

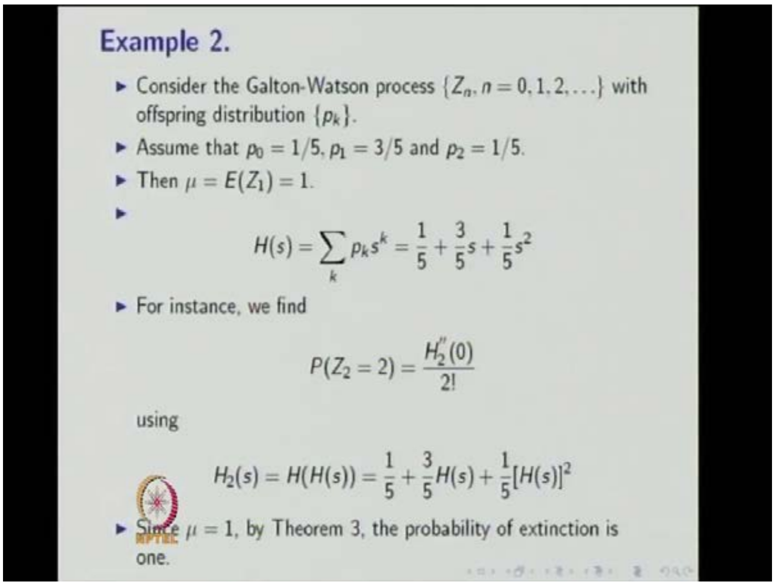


We know, that probability of Z m is equal to 0 given Z 1 is equal to j is nothing but probability of Y 1 plus Y 2 plus and so on. Y j is equal to 0 since Y i's are i.i.d. random variables, that is nothing but probability of Y 1 is equal to 0 Y 2 is equal to 0 and so on. Probability of Y j is equal to 0 that is same as the probability of Z m minus 1 equal to 0 whole power j, that is nothing but d suffix m minus 1 power j.

Hence, you can substitute this result in this equation. This conditional probability is nothing but d suffix m minus 1 power j. Therefore, d m will be summation over j, d of m minus 1 power j p j and that is nothing but the probability generating function of d m minus 1. So, hence we get d m is equal to H of d m minus 1. Since d m tends to d as m tends to infinity, the value d satisfies the equation s is equal to H of s because d m is equal to H of d m minus 1. Therefore, the value d satisfies the equation s is equal to H of s.

Note, that the solution to the equation s is equal to H of s represents intersections of the graphs of Y is equal s and Y is equal to H of s. So, the d will be probability of extinction and we are not going to discuss furthermore about how to solve s is equal to H of s and finding the probability of extinction.

(Refer Slide Time: 38:39)



We will consider the third example, Z n be the sequence of random variable, which is a Galton-Watson process with offspring distribution p k. Similar to the example 1 and 2, p naught is equal to 1 by 5, p 1 is equal to 3 by 5, p 2 is equal to 1 by 5. Already we got the mean of Z 1 that is equal to 1; therefore, this is the critical Galton-Watson process.

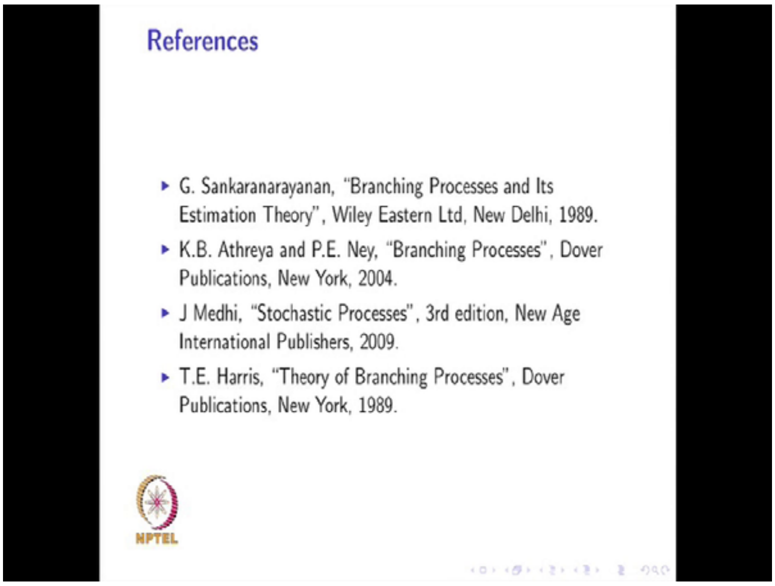
Now, you can find out the probability generating function of Z 1. Using this one can find the probability of Z is equal to or we can find the distribution of Z n also. For instant, we find probability of Z is equal to 2, for that we, you need the probability generating function of Z 2, that means, you should know what is the probability generating function of H 2 of s.

With the help of H 1 of s and H of s, one can find the H 2 of s that is what we have proved it in the theorem 2. H n, H n of s will be H n minus 1 of H of s. So, here we put n is equal to 2, find the probability generating function of Z 2 using the probability generating function of Z 2. We are finding the probability of Z 2 is equal to 2. So, probability of Z is equal to 2 will be H 2 double dash of 0 divided by 2 factorial.

So, first you find out H 2 of s that is nothing but H 1 of H of s that is same as H of H of s. So, replace s by H of s in the H of s expression, that is, 1 by 5 plus 3 by 5 times s plus 1 by 5 s square. So, replace s by H of s, you will get H 2 of s. Once you know H 2 of s, you differentiate twice, then substitute s is equal to 0, then divide by 2, you will get the probability of Z 2 is equal 2. Since mu is equal to 1, this is the critical Watson process.

By using theorem 3 we conclude the probability of extinction is 1. Theorem 3 says, if mu is less than or equal to 1, then the probability of extinction is 1. So, since in this problem the mu is equal to 1, hence by theorem 3 the probability of extinction is 1.

(Refer Slide Time: 41:57)



So, in this lecture we have covered the definition and the examples of a branching process, in particular we have discussed Galton-Watson discrete branching process. We have discussed probability generating function of Z n; we have discussed mean and variance of Z n. Also, we have seen three examples, through that we found the conditional probability, probability generating function mean and variance of Z n. Finally, we have discussed probability of extinction for the Galton-Watson process.

In the next lecture we are going to cover another important branching process, that is, Markov branching process, which is a continuous type branching process. In this lecture we have discussed discrete type branching process, that is, Galton-Watson process. In the next lecture we are going to cover continuous type branching process, that is, Markov branching process and some more important branching processes. Here are the references.