

Stochastic Processes

Module 9: Branching Processes

Lecture 1: Galton - Watson Process

Dr. S. Dharmaraja

Department of Mathematics
Indian Institute of Technology Delhi
New Delhi, India
<http://web.iitd.ac.in/~dharmar>

◀ ◯ ▶ ⏪ ⏩ 🔍

Video Course on
Stochastic Processes -1

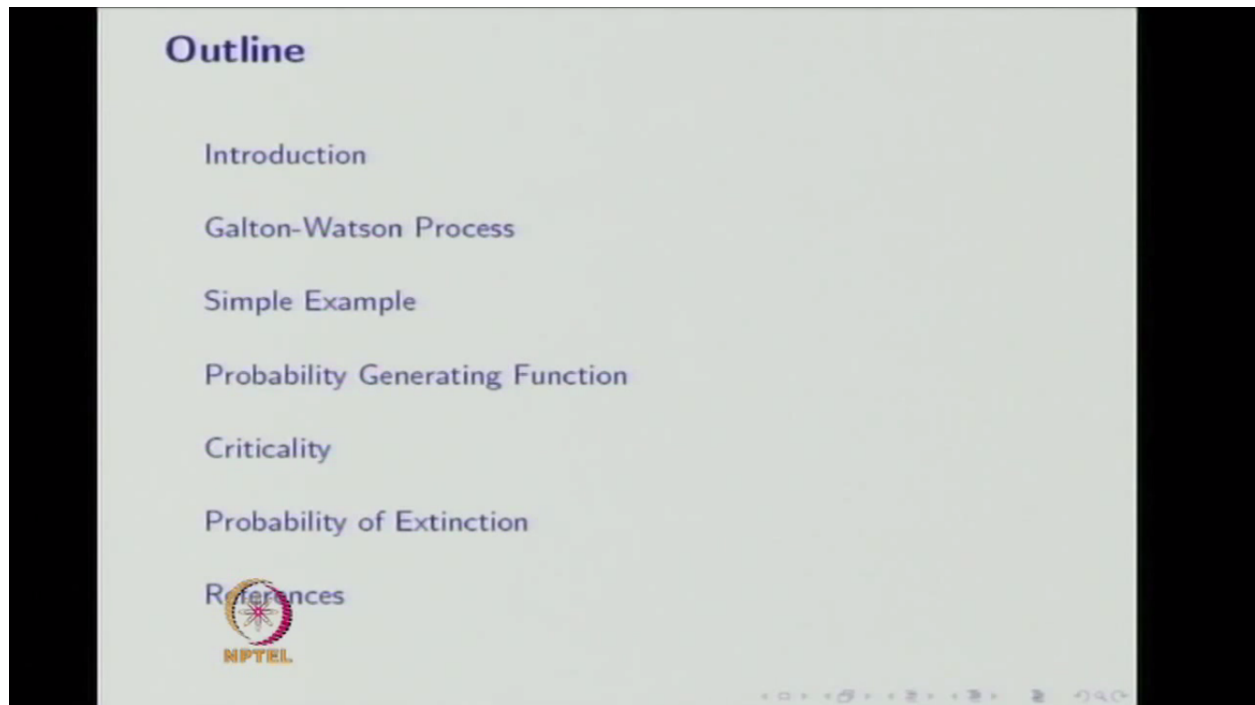
By

Dr. S Dharmaraja
Department of Mathematics, IIT Delhi

Module #9
Branching Processes

Lecture # 1
Galton-Watson Process

This is a stochastic processes module 9, branching processes. In the module 1 we have discussed a review of probability. Model 2, we have introduced a stochastic process. Model 3, we have discussed stationary processes. Model 4 we have discussed a discrete time Markov chain. Model 5 we have discussed a continuous time Markov chain. In the model 6 we have discussed the Martingale. Model 7 we have discussed about Brownian motion or Wiener process. Model 8 we have discussed renewal process. In this model 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 Markov branching process are going to be discussed.

Important measures like probability of extension, limit theorem, limit distribution and mean of branching process will be discussed. In the lecture 1 we are going to start with the interaction 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.

Introduction

- ▶ H.W. Watson in 1874 presented an solution to the “problem of the extinction of families” which is as follows:
“Let p_0, p_1, p_2, \dots be the respective probabilities that a man has 0, 1, 2, \dots 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?”.



Watson in 1874 presented a solution to the problem of extinction of families which is as follows”
Let p_0, p_1, p_2, \dots be the respective probabilities that the man has 0, 1, 2, so on sons that each some 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. 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 preface to the solution by Watson and Galton recognized that a first step in studying hypothesis would be to determine the probability that the ordinary family will disappear using fertility data for the whole population.

Introduction . . .

- ▶ Watson gave the solution “every family will die out, even when the population size, on the average, increase from one generation to the next”.
- ▶ Francis Galton prefaced to the solution by Watson, and Galton recognized that a first step in studying the hypothesis would be to determine the probability that an ordinary family will disappear using fertility data for the whole population.
- ▶ The mathematical model of Galton and Watson is known as a Galton - Watson process.
- ▶ In 1938, A. 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



The mathematical model of Galton and Watson is known as Galton-Watson branching process. In 1938 Kolmogorov was also treated the above problem and determined asymptotic form of the probability that the family is still in existence after a large finite number of generations. The branching process is a system of particles or individuals or cells or molecules which leave for a random time and at some point during lifetime or at the moment of death produce a 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 entities lifetime are called the classical branching processes. Examples populations of biological cells. Genes or biomolecules.

Introduction . . .

- ▶ 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 a continuum, of all possible varieties of particles included in the process.
- ▶ Particles of a given type may produce particles of different types.
- ▶ Restrictions on type transitions, as well as on the type space, lead to different properties of resulting processes.



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 a different types. Restrictions on the type transitions as well as on the type space lead to different properties of resulting processes.

Classical Branching Process

- ▶ Consider a classical branching process in which progeny are born at the moment of the parents death.
- ▶ The distribution of particle lifetime τ has much impact on the behavior and analysis of the process.
- ▶ If $\tau = 1$, then it is called the Galton - Watson process.
- ▶ If τ is exponentially distributed random variable, then the resulting process is called a Markov branching process.
- ▶ If τ is an arbitrary non-negative random variable, the resulting process is called an "age-dependent" or Bellman - Harris process.



Galton-Watson Process

- ▶ Consider particles that can generate additional particles of the same kind.
- ▶ An initial set of particles, which we call the 0-th generation, have offspring that are called the first generation, their offspring are the second generation, and so on.
- ▶ We denote by Z_0, Z_1, Z_2, \dots as the number of particles in the 0-th, 1-th, 2-nd, ... generations.
- ▶ Furthermore, we make the following assumptions:
 - ▶ If the size of the n -th generation is known, then the probability law governing later generations does not depend on the sizes of generations preceding the n -th.



It means $\{Z_n, n = 0, 1, \dots\}$ is a discrete time Markov chain.

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

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 0-th generation having offspring that are called first generation, their offsprings are the second generation and so on. We denote by Z naught, Z_1, Z_2 , as the number of particles in the 0-th, first, second generations. Please see the illustration. This is that Z naught, Z_1, Z_2 , and so on. So this is an illustration of the number of particles in the first generation 0-th generation, first generation, second generation, and so on.

Furthermore, we make the following assumptions. If the size of the n th generation is known then the probability law governing later generations does not depend on the sizes of generations preceding the n th. It means the sequence of random variables Z_n form a discrete time Markov chain. We make an additional assumption that the transition probabilities for the Markov chain do not vary with the time. It means the sequence of random variables Z_n is a time homogeneous discrete time Markov chain.

Time Homogeneous DTMC

09:39 / 12:42

- ▶ We make the additional assumption that the transition probabilities for the Markov chain do not vary with time. It means $\{Z_n, n = 0, 1, \dots\}$ is a time homogeneous DTMC.
- ▶ Note that the number of offsprings born to a particle does not depend on how many other particles are present.
- ▶ Also, we assume that $Z_0 = 1$.
- ▶ If $Z_n = 0$ then $P(Z_{n+1} = 0) = 1$.
- ▶ The probability distribution of Z_1 is given by

$$P(Z_1 = k) = p_k, \quad k = 0, 1, \dots, \quad \sum_{k=0}^{\infty} p_k = 1$$

where p_k is the probability that a particle existing in the n -th generation has k offsprings in the $(n + 1)$ -th generation.

We have discussed the discrete time Markov chain in the model four in detail. Note that the number off springs born to a particle does not depend on how many other particles are present. Also we assume that Z_0 is equal to 1.

In the illustration also we made it Z_0 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+1} 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 notation P_{k} and the summation of P_k 's is equal to 1. That P_k is the probability that a particle existing in the n th generation has K offsprings in the $n + 1$ -th generation. So this is the probability mass function for the random variable Z_1 .

One-step Transition Probabilities

- ▶ It is assumed that p_k is independent of the generation number n .
- ▶ The conditional distribution of Z_{n+1} given $Z_n = k$ is appropriate to the assumption that different particles reproduce independently; that is, Z_{n+1} is distributed as the sum of k independent random variables, each distributed like Z_1 .
- ▶ Since $\{Z_n, n = 0, 1, \dots\}$ is a time homogeneous DTMC, the one-step transition probabilities are given by

$$\begin{aligned} p_{ij} &= P(Z_{n+1} = j \mid Z_n = i) \\ &= P(Y_1 + Y_2 + \dots + Y_i = j), \quad i, j = 1, 2, \dots \quad (1) \end{aligned}$$



where Y_1, Y_2, \dots, Y_i are independent random variables, each distributed like Z_1 .

It is assumed that p_k is independent of the generation number n . The conditional distribution of $Z_n + 1$ given $Z_n = k$ is appropriate to the assumption that the different particles reproduce independently. That is $Z_n + 1$ 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 homogeneous discrete time Markov chain the one step transition probabilities are given by P_{ij} that is nothing but the probability that $Z_n + 1$ is equal to j given Z_n is equal to i . That is same as $Y_1 + Y_2 + \dots + Y_i$ that takes a value j for i, j belonging to $1, 2$ and so on where Y_i 's are iid random variables each distributed like Z_1 .

So the illustration for the Galton-Watson process is a $Z_n = 1, Z_1 = 4$ and $Z_2 = 7$ and so on.