

Stochastic Processes

NPTEL - II Web Course

N. Selvaraju

Department of Mathematics
Indian Institute of Technology Guwahati

&

S. Dharmaraja

Department of Mathematics
Indian Institute of Technology Delhi

Modulewise Contents

1	Probability Essentials	1
1	Introduction	1
1	Probability Spaces	1
2	Conditional Probability and Independence	3
2	Random Variables and Probability Distributions	7
1	Random Variables and Distribution Functions	7
2	Probability Mass and Density Functions	10
3	Functions of Random Variables	14
3	Random Vectors and Expectations	16
1	Random Vectors	16
2	Expectations	20
3	Generating Functions	24
4	Sequences of Random Variables	26
1	Sequences of Random Variables	26
2	Limit Theorems	28
2	Introduction to Stochastic Processes	1
1	What are Stochastic Processes?	1
2	Classes of Stochastic Processes	8
1	Independent Process	8

MODULEWISE CONTENTS

iv

2	Bernoulli Process	8
3	Processes with Independent Increments	9
4	Processes with Stationary Increments	9
5	Markov Processes	9
6	Poisson Process	10
7	Simple Random Walks	11
8	Second-Order Processes	11
9	Stationary Processes	11
10	Gaussian Processes	12
11	Martingales	12
12	Brownian Motion	13
3	Discrete-Time Markov Chains	1
1	Introduction and Examples	1
1	Markov Chain	2
2	State Transition Diagram	7
2	Classification of States - I	8
1	Irreducible Chains	8
2	Recurrence and Transience	9
3	Classification of States - II	14
1	Periodicity	14
2	Ergodicity	14
4	Stability of Markov chains	18
1	Limiting Distribution	18
2	Stationary Distribution	20
3	Finite Markov Chains	21
4	Markov chain with countable state space	23

MODULEWISE CONTENTS

v

5	Reducible Markov Chains	25
1	Finite chains with single closed class	25
2	MC with a single class of positive recurrent states	25
3	Absorbing MCs	26
6	Reversed and Time-Reversible Markov Chains	28
4	Continuous-Time Markov Chains	1
1	Definition and Kolmogorov Equations	1
1	Introduction	1
2	Definition	1
3	Chapman-Kolmogorov equations for time homogeneous CTMC	3
2	Limiting and Stationary Distributions	8
1	Limiting distribution	9
2	Stationary distribution	10
3	Some notes	12
3	Poisson Processes - I	14
1	Definition of Poisson Process	16
2	Properties of Poisson processes	17
4	Poisson Processes - II	21
1	Examples	21
2	Non-homogeneous Poisson process	22
3	Compound Poisson process	23
4	Compensated Poisson Process	24
5	Birth-Death Process	25
1	Definition	25
2	Finite BDP	27
3	Pure birth process	27

MODULEWISE CONTENTS

4	Pure death process	28
6	<i>M/M/1</i> Queueing Model	30
1	Queueing systems	30
2	<i>M/M/1</i> Queueing model	32
3	Little's law	35
7	Simple Markovian Queueing Models	38
1	<i>M/M/c</i> Queueing model	38
2	<i>M/M/1/N</i> Queueing model	40
3	<i>M/M/c/K</i> Queueing model	42
4	<i>M/M/c/c</i> loss system	43
5	<i>M/M/∞</i> self service system	43
5	Martingales	1
1	Filtrations and Conditional Expectations	1
1	X_t is \mathcal{F}_t -measurable	1
2	Filtrations	2
3	Adaptability	4
4	Conditional Expectations	4
2	Conditional Expectations	7
1	Rules on Conditional Expectations	7
2	Conditional Expectation Given σ -field	9
3	Generated σ -fields	11
1	σ -field generated by a random vector \mathbf{Y}	11
2	Properties of Conditional Expectations given σ -field	14
4	Martingales, Sub-martingales and Super-martingales	16
1	Martingales	16
2	Doob's Martingale Process	18

MODULEWISE CONTENTS

vii

3	Some Results	18
5	Stopping Times and Inequalities	21
1	Stopping Times	21
2	Some Results	23
6	Convergence Theorems	25
1	Optional Sampling Theorem	25
2	Martingale Convergence Theorem	26
3	Martingale in Continuous Time	27
6	Brownian Motion	1
1	Brownian Motion	1
2	Properties of Brownian Motions-I	6
3	Properties of Brownian Motion-II	13
4	Processes Derived from Brownian Motion	18
5	Stochastic Differential Equations and Itô Integrals	23
1	Itô Calculus	23
2	Itô Integrals	24
3	Stochastic Differential Equations	30
6	Itô's Formula	32
7	SDEs and their Applications in Finance	37
7	Renewal Processes	1
1	Renewal Function and Renewal Equation	1
1	Introduction	1
2	Distribution of $N(t)$	3
3	Renewal Function	4
4	Renewal Equation	5
5	Age, Excess and Spread at Time t	7

MODULEWISE CONTENTS

viii

2	Generalized Renewal Processes and Limit Theorems	9
1	Renewal Reward Process	9
2	Markov Reward Model	10
3	Alternative Renewal Process	15
4	Delayed Renewal process	15
3	Markov Renewal and Markov Regenerative Processes	17
1	Semi-Markov Process	17
2	Regenerative Process	20
4	Non-Markovian Queues	25
1	<i>M/G/1</i> Queueing Model	25
5	Non Markovian Queues (contd.)	32
6	Non Markovian Queues (contd.)	37
1	<i>GI/M/1</i> Queue	37
8	Branching Processes	1
1	Galton - Watson Process	1
2	Properties of GW Process	8
3	Markov Branching Process	12
1	Markov Branching Process	12
9	Stationary Processes	1
1	Stationary Processes	1
1	Introduction	1
2	Important Definitions	1
3	Strict Sense and Wide Sense Stationary Process	3
4	Examples of Stationary Process	4
2	Some Special Stationary Processes	7
1	Time Series	7

MODULEWISE CONTENTS

ix

2	Pure Random Process	7
3	Moving Average Process	8
4	Autoregressive Process	9

Module 1: Probability Essentials

Lecture 1 Introduction

In our everyday life, we encounter many systems that we feel should see improvements. The system that we observe may be the planes hovering around airports waiting to land while others waiting to take off, or may be the price movements of a particular stock in a stock exchange, or may be speed of the internet in your home, or may be production of some items in a manufacturing company. In order to improve the behaviour of such systems, one needs to analyse them first by abstracting their essential features. These are typical examples of systems whose behaviour vary in time in a random manner and interest in studying them has never reduced and will never be. The mathematical models that abstract the essential features of such systems are known as stochastic processes or random processes.

In this module, we will briefly review the probability essentials that form backbone to the study of stochastic processes. Those who are familiar with the material covered in this module may either skip this module entirely or gloss over the module quickly to revise the concepts and get familiar with the notations used throughout the course.

1 Probability Spaces

Probability theory is concerned with modelling of a phenomenon that behave in an unpredictable manner. In probabilistic modelling, the first basic concept is that of a ‘random experiment’, like observing outcome of tossing a coin or throwing a dice or observing life length of device etc. In all such examples, we do not know what is going to be the outcome. They are thus known as random experiments which essentially means that we can only enumerate all possible outcomes of the experiment but are not sure which one of them actually will happen. We abstract all possible outcomes of a random experiment by a set Ω and call it the *sample space* of the experiment. The elements of Ω are called sample points or elementary events.

We are often interested in the study of set of sample points, rather than a single sample point only. These are called as events of the experiment and the events satisfy some consistency requirements among themselves. In its full generality, they

form a collection known as a σ -algebra (or σ -field) over Ω which is defined as below.

Definition 1.1. A nonempty collection \mathcal{F} of subsets of Ω is called a σ -algebra (over Ω) if \mathcal{F} is closed under the operations of countable unions and complementations, i.e.,

1. $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$ whenever $A_n \in \mathcal{F}, n \geq 1$, and
2. $A^c \in \mathcal{F}$ whenever $A \in \mathcal{F}$.

The pair (Ω, \mathcal{F}) is called a measurable space.

Example 1.1. The largest σ -field of subsets of a fixed set Ω is the collection of all subsets of Ω , i.e., the power set P^Ω . The smallest σ -field consists of the two sets \emptyset and Ω . These σ -fields are sometimes called as trivial σ -fields.

Example 1.2. Let Ω be an infinite set, and let \mathcal{F} be the collection of all finite subsets of Ω . Then \mathcal{F} does not contain Ω and is not closed under complementation, and so is not a σ -algebra (or a σ -field) on Ω .

In the definition given above, if we replace the ‘countable union’ part by a ‘finite union’, then the collection is called as a field. It can be seen easily that each σ -field on Ω is a field on Ω , since, for example, the union of the finite sequence A_1, A_2, \dots, A_n is the same as the union of the infinite sequence $A_1, A_2, \dots, A_n, A_n, A_n, \dots$. Similarly, every finite field (a field with a finite number of elements) is a σ -field too. But, in general, a field may not be a σ -field.

Example 1.3. Let A be a nonempty proper subset of Ω , and let $\mathcal{F} = \{\emptyset, \Omega, A, A^c\}$. Then \mathcal{F} is the smallest σ -field containing A . For if \mathcal{G} is a σ -field and $A \in \mathcal{G}$, then by definition of a σ -field, Ω, \emptyset , and A^c belong to \mathcal{G} , hence $\mathcal{F} \subseteq \mathcal{G}$. But \mathcal{F} is a σ -field, for if we form complements or unions of sets in \mathcal{F} , we obtain sets in \mathcal{F} . Thus \mathcal{F} is a σ -field that is included in any σ -field containing A , and the result follows.

Exercise 1.1. Let A, B be nonempty proper subsets of Ω . Determine the smallest σ -field containing A and B . How many elements are there in it? Extending the idea, describe explicitly the smallest σ -field containing a finite number of subsets A_1, A_2, \dots, A_n of Ω .

Generalizing the idea of the previous example, one can talk about the *smallest* σ -field containing a class of sets. If \mathcal{C} is a class of subsets of a set Ω , the smallest σ -field containing the sets of \mathcal{C} will be written as $\sigma(\mathcal{C})$, and will be called the *minimal*

σ -field over \mathcal{C} or the σ -field generated by \mathcal{C} . And, \mathcal{C} will be called a generator for the σ -field $\sigma(\mathcal{C})$.

Example 1.4. (*Borel σ -field*) In probability theory, the σ -field of interest is what is known as the Borel σ -field (especially in the case when $\Omega = \mathbb{R}$). Denoted as \mathcal{B} (or $\mathcal{B}(\mathbb{R})$), the Borel σ -field over \mathbb{R} is the σ -field generated by the class of all intervals of the form (a, b) , $a, b \in \mathbb{R}$. A Borel set is an element of the Borel σ -field.

Note that the class of open intervals is only one of the many generators of the Borel σ -field. And, though practically all the subsets of \mathbb{R} that we encounter are Borel sets, there exists non-Borel sets. But we will not worry about that.

The next idea concerns the measurement i.e. the probability of the events, which we define below in an axiomatic way.

Definition 1.2. A function $P : \mathcal{F} \rightarrow \mathbb{R}$ is called a probability measure (or simply probability) if

1. $0 \leq P(A) \leq 1, \forall A \in \mathcal{F}$,
2. $P(\Omega) = 1$, and
3. $A_1, A_2, \dots \in \mathcal{F}$ are mutually exclusive, then $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

The triplet (Ω, \mathcal{F}, P) is called a probability space.

If you are familiar with measure theory, you can realize immediately that the probability measure is a measure of total mass one. You should keep in mind the fact that there is always an underlying probability space in every probabilistic modelling problem. Sometimes it can be described easily and sometimes it may not be.

Example 1.5. In the case of throwing a die twice, $\Omega = \{(i, j) : i, j = 1, 2, \dots, 6\}$, where i is the outcome of the first throw and j the outcome of the second throw. If we take $\mathcal{F} = P^\Omega$ then we get a particular probability space with probability measure defined by say $P(A) = \frac{|A|}{|\Omega|}$, $A \in \mathcal{F}$ where $|A|$ is cardinality of set A .

If we take $\mathcal{F} = \{(1, 1)\}, \{(1, 1)\}^c, \Omega, \emptyset\}$, then we get a different probability space.

2 Conditional Probability and Independence

Consider families with two children. What is the probability that a family has two boys, given that it has one boy? The answer however is not $\frac{1}{2}$. Here, $\Omega =$

$\{BB, BG, GG, GB\}$ and let $H = \{BB, BG, GB\}$ be the event that the family has at least one boy. Under the condition that family has one boy, H can be considered as the new sample space and all outcomes are considered to be equally likely then conditional probability of event A is $P(A | H) = \frac{1}{3}$, where A is the event that the family has two boys, and in that case $P(A | H) = \frac{P(A \cap H)}{P(H)} = \frac{1}{3}$. In general, we have the following definition.

Definition 1.3. *Given a probability space (Ω, \mathcal{F}, P) , let $B \in \mathcal{F}$ be some fixed event such that $P(B) > 0$, then the conditional probability of $A \in \mathcal{F}$ given B (denoted by $P(A | B)$) is equal to $\frac{P(A \cap B)}{P(B)}$.*

Theorem 1.1. *Given (Ω, \mathcal{F}, P) , consider a fixed $B \in \mathcal{F}$ such that $P(B) > 0$. Then the function $P(. | B)$ behaves like an ordinary probability measure, that is it satisfies all the three axioms.*

Proof. 1. $P(A | B) \geq 0$ for any $A \in \mathcal{F}$

2. $P(\Omega | B) = 1$

3. If A_i for $i = 1, 2, \dots$ is a sequence of mutually exclusive (disjoint) events then

$$P(\cup_{i=1}^{\infty} A_i | B) = \sum_{i=1}^{\infty} P(A_i | B).$$

This is because

$$P(\cup_{i=1}^{\infty} A_i | B) = \frac{P(\cup_{i=1}^{\infty} (A_i \cap B))}{P(B)} = \sum_{i=1}^{\infty} \frac{P(A_i \cap B)}{P(B)} = \sum_{i=1}^{\infty} P(A_i | B).$$

Also the original probability measure P on \mathcal{F} actually is of the same form where $B = \Omega$.

□

We have the following theorem which is known as the theorem of total probability.

Theorem 1.2. *Let $A_i \in \mathcal{F}$ for $i = 1, 2, \dots$ be a sequence of mutually exclusive and exhaustive events such that $P(A_i) > 0$ for all i . If $B \in \mathcal{F}$ be any other event then*

$$P(B) = \sum_{i=1}^{\infty} P(B | A_i)P(A_i).$$

The proof of the above theorem is simple on noting that $P(B) = P(B \cap \Omega) = P(B \cap (\cup_{i=1}^{\infty} A_i)) = P(\cup_{i=1}^{\infty} (A_i \cap B)) = \sum_{i=1}^{\infty} P(A_i \cap B)$, since the A_i 's are mutually exclusive. Hence, $P(B) = \sum_{i=1}^{\infty} P(B | A_i)P(A_i)$.

A very useful result is the following theorem.

Theorem 1.3. (*Bayes theorem*) Let $A_i \in \mathcal{F}$ for $i = 1, 2, \dots$ be a sequence of mutually exclusive and exhaustive events such that $P(A_i) > 0$ for all i . If $B \in \mathcal{F}$, $P(B) > 0$ be any other event then

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{\sum_{i=1}^{\infty} P(B | A_i)P(A_i)}.$$

The theorem follows from the fact that $P(A_i | B) = \frac{P(A_i \cap B)}{P(B)}$, and then using total probability result.

Example 1.6. In a bolt factory machines M_1, M_2 and M_3 manufacture respectively 25%, 35% and 40% of the total production. Of their output 5, 4, 2 percent respectively are defective bolts. A bolt is drawn from a day's production and found to be defective. What is the probability that it was manufactured by machine M_3 ?

Solution: Let A_1, A_2 and A_3 represent the events that a bolt selected at random is manufactured by the machines M_1, M_2 and M_3 respectively, and let B denote the event of its being defective. It is given that $P(A_1) = 0.25, P(A_2) = 0.35, P(A_3) = 0.40, P(B|A_1) = 0.05, P(B|A_2) = 0.04, P(B|A_3) = 0.02$. Therefore, the probability that a defective bolt selected at random is manufactured by machine M_3 is computed using Bayes' rule as

$$P(A_3|B) = \frac{P(A_3)P(B|A_3)}{\sum_{i=1}^3 P(A_i)P(B|A_i)} = \frac{(0.40)(0.02)}{(0.25)(0.05) + (0.35)(0.04) + (0.40)(0.02)} = \frac{16}{69}.$$

Definition 1.4. Given a probability space (Ω, \mathcal{F}, P) , two events A and B are said to be statistically independent or stochastically independent or (simply) independent if

$$P(A \cap B) = P(A)P(B).$$

It can be easily shown that the above is the same as saying that if $P(B) > 0$ then A and B are independent if $P(A | B) = P(A)$ and if $P(B^c) > 0$ then A and B are independent if $P(A | B^c) = P(A)$.

A is said to be independent of itself if $P(A)$ is either 0 or 1.

Three events A_1, A_2, A_3 are said to be independent if they are pairwise independent and also $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$.

The following is an example which illustrates that pairwise independence need not imply independence of the events if the number of events is more than 2.

Example 1.7. Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ be a sample and \mathcal{F} be the power set of Ω . Let all the outcomes in the sample space be equally probable, let $A_1 = \{\omega_1, \omega_2\}$, $A_2 = \{\omega_2, \omega_3\}$, $A_3 = \{\omega_3, \omega_4\}$, then $P(A_1 \cap A_2) = P(A_1)P(A_2) = \frac{1}{2}$, $P(A_2 \cap A_3) = P(A_2)P(A_3) = \frac{1}{2}$, $P(A_1 \cap A_3) = P(A_1)P(A_3) = \frac{1}{2}$ and $P(A_1 \cap A_2 \cap A_3) = P(\emptyset) = 0$. But $P(A_1)P(A_2)P(A_3) = \frac{1}{8}$ and hence these three events are not statistically independent although they are pairwise independent.

Another question which needs to be addressed is that if $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$, then does it imply that the events are pairwise independent? Again, the answer is no. To see how, repeat the above experiment with Ω having 8 elements such that A_1, A_2, A_3 have four elements each and $A_1 \cap A_2 \cap A_3$ has one element.

Similarly, if we have n events say A_1, A_2, \dots, A_n then the A_i 's are said to be independent if they satisfy the following equalities.

1. $P(A_i \cap A_j) = P(A_i)P(A_j)$ for all $i < j$, there are n_{C_2} such pairs (i, j) .
2. $P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)$ for all $i < j < k$, there are n_{C_3} such triplets (i, j, k) .
- ⋮
3. $P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2) \dots P(A_n)$

Thus, in total one has to check $2^n - n - 1$ above identities.