普通高等教育“十二五”规划教材

Probability and Stochastic Processes

张丽华 周 清 主编

北京邮电大学出版社
www.buptpress.com

普通高等教育"十二五"规划教材

Probability and Stochastic Processes

张丽华 主编

北京邮电大学出版社
·北京·

内 容 简 介

本书系统地介绍概率论与随机过程的基本概念、基本方法、基本理论以及应用。本书分为八章。前四章介绍概率论的一般知识及应用，后四章介绍随机过程的一般知识及应用。

该教材注重概念之间的联系和背景介绍，强调知识的应用，而且本书所有内容是自包含的，讲述浅显易懂，便于自学。

该教材供非数学专业应应用型本科理工类一学期（64-72学时）学习使用。

图书在版编目(CIP)数据

概率论与随机过程 = Probability and StochasticProcesses / 张丽华主编. -- 北京 : 北京邮电大学出版社, 2015.12

ISBN 978-7-5635-4537-7

Ⅰ.①概… Ⅱ.①张… Ⅲ.①概率论—高等学校—教材②随机过程—高等学校—教材 Ⅳ.①O21

中国版本图书馆CIP数据核字(2015)第228387号

书　　名：概率论与随机过程 = Probability and StochasticProcesses
著作责任者：张丽华　主编
责 任 编 辑：张珊珊
出 版 发 行：北京邮电大学出版社
社　　址：北京市海淀区西土城路10号(邮编:100876)
发 行 部：电话:010-62282185　传真:010-62283578
E-mail：publish@bupt.edu.cn
经　　销：各地新华书店
印　　刷：
开　　本：787 mm×1 092 mm　1/16
印　　张：14.75
字　　数：362千字
版　　次：2015年12月第1版　2015年12月第1次印刷

ISBN 978-7-5635-4537-7　　定　价：30.00元

· 如有印装质量问题，请与北京邮电大学出版社发行部联系 ·

2015年12月第1次印刷

Preface

Probability theory provides a precise mathematical description of chance of random events. As a follow-up course to probability theory, stochastic processes concern sequences of events governed by probabilistic laws. In an effort to quantify, understand, and predict the effects of randomness, the mathematical theory of probability and stochastic processes has been developed. They are really fundamental importance to nearly all scientists, engineers, medical practitioners, jurists, industrialists and so on. Like Pierre Simon Laplace said "The most important questions of life are, for the most part, really only problems of probability".

This book is the introduction to the basic concepts and knowledge about probability and stochastic processes, which serves for one-semester course. The main intended audience for this book is undergraduate students in pure and applied sciences, especially those in engineering. What we endeavored to do in this book are:

1. to present a systematic introductory account of several principal areas in probability and stochastic processes step by step.

2. to emphasize two basic methods, the distribution function and the numerical characteristics, through the whole book to aim.

3. to pay great attention to the connection between concepts.

4. to provide a lot of examples and applications of probability and stochastic processes.

5. to put introduction in each chapter and each section to reader.

The aim we have done as above is to make readers study easily and to stimulate students interest in learning, moreover, to broaden readers horizons and to strengthen the logical and problem-solving ability.

The book is organized as follows.

Chapter 1 introduces basic ideas of probability theory, the axioms of probability theory and shows how they can be applied to compute various probabilities of interest. In this Chapter we also deal with the extremely important subjects of conditional probability and independence of events.

In Chapter 2 we mainly deals with the random variable in one dimension. The important concepts, distribution, the expected value and the variance of a random variable, are introduced here. This chapter also presents such probability inequalities as Markovs inequality, Chebyshevs inequality, and gives some applications of these inequalities.

Bivariate random vectors, jointly distributed random variables, are dealt with in Chapter 3, additional properties of the expected value are also considered in this chapter. The relationships between random variables are concentrated here. The final section introduces the multivariate distribution.

Chapter 4 presents the major theoretical results of probability theory. In particular, we prove the law of large numbers and the central limit theorem. We emphasize the two basic methods, distribution and moments, are still works for random variable sequences. But the results about those are always regarded as the realm of stochastic processes.

Probability provides models for analyzing random or unpredictable outcomes. The main new ingredient in stochastic processes, which is introduced in Chapter 5, is the explicit role of time. According to the dependence relations among the random variables, there have different kinds of stochastic processes, which will be partly introduced in this book.

Chapter 6 gives the definitions and properties of two kinds of stationary processes: strict-sense stationary processes(SSS) and wide-sense stationary processes(WSS). The ergodic theorem and spectral representation of a WSS are introduced here. The analysis of linear systems with random processes as inputs is given in the end.

In Chapter 7, we consider the finite Markov chains, which represent the simplest generalization of independent processes and have the property that the future is independent of the past given the present. The properties, characterization, and the long-term behavior are described here.

In Chapter 8 we are concerned with a type of stochastic process having independent increment. In particular, Poisson processes and Wiener processes are introduced here.

The examples and exercises have been chosen to illustrate the subject, to help the student solve various kinds of problems and for their entertainment value.

Bupt
Zhang Lihua
Zhou Qing

Contents

Chapter 1

Events and Their Probabilities

The most important questions of life are, for the most part, really only problems of probability.

— Pierre Simon Laplace

What is probability? How to understand it? We try to explain the two problems in this chapter. First, we review briefly the history of probability. Then we give some basic concepts which are necessary to define probability. Based on these concepts, we give three different interpretations: classical probability, geometric probability and frequency. The axiomatic definition of probability is shown to unify various interpretations about probability. If some event occurs, then the probability of original event which we considered before may be changed. We discuss this kind of probability, which is called conditional probability. At last, from the definition of conditional probability, we introduce the independence of events, which is important in many probability models.

1.1 The History of Probability

Probability theory is the branch of mathematics concerned with probability, the analysis of random phenomena.

The mathematical theory of probability originates from analyzing games of chance. Gerolamo Cardano (1501–1576)'s book about games of chance, Liber de ludo aleae (in English, Book on Games of Chance), written around 1564, but not published until 1663, is believed to be the first systematic treatment of probability. Cardano used the game of throwing dice to understand the basic concepts of probability.

One of the famous probability problems is the problem of points. The problem concerns a game of chance with two players who have equal chances of winning each round. The players contribute equally to a prize pot, and agree in advance that the first player who have won a certain number of rounds will collect the entire prize. Now suppose that the game is interrupted by external circumstances before either player has achieved victory. How does one then divide the pot fairly? It is tacitly understood that the division should depend somehow on the number of rounds won by each player, and a player who is close to winning should get a larger part of the pot. But the problem is not merely one of calculation, it also includes deciding what a "fair" division should mean in the first place.

Cardano analyzed the problem but did not get the right answer. So it is generally believed that the modern probability theory was started by the French mathematicians Blaise Pascal (1623–1662) and Pierre Fermat (1601–1665) when they succeeded in deriving the problem of points. In their works, the concepts of the probability of a stochastic event and the expected or mean value of a random variable can be found. Although their investigations were concerned with problems connected with games of chance, the importance of these new concepts was clear to them, as Christiaan Huygens (1629–1695) points out in the first printed probability text, On Calculations in Games of Chance (1657), "The reader will note that we are dealing not only with games, but also that the foundations of a very interesting and profound theory are being laid here." Later, Jacob Bernoulli (1654–1705), Abraham De Moivre (1667–1754), Thomas Bayes (1702–1761), Marquis Pierre Simon Laplace (1749–1827), Johann Carl Friedrich Gauss (1777–1855), and Simeon Denis Poisson (1781–1840) contributed significantly to the development of probability theory. The notable contributors from the Russian school include Pafnuty Lvovich Chebyshev (1821–1894), and his students Andrey Andreyevich Markov (1856–1922) and Aleksandr Mikhailovich Lyapunov (1857–1918) with important works dealing with the law of large numbers.

The deductive theory based on the axiomatic definition of probability that is popular today, which is mainly attributed to Andrei Nikolaevich Kolmogorov (1903–1987), who in the 1930s along with Paul Lévy (1886–1971) found a close connection between the theory of probability and the mathematical theory of sets and functions of a real variable. Although Émile Borel (1871–1956) had arrived at these ideas earlier, putting probability theory on this modern frame work is mainly due to the early 20th century mathematicians. As is the case with all parts of mathematics, probability theory is constructed by means of the axiomatic method. So we regard probability theory as a part of mathematics.

The theory of probability has been developed steadily and applied widely in diverse fields of study since the seventeenth century. One of the most striking features of the present day is the steadily increasing use of the ideas of probability theory in a wide variety of scientific fields, involving matters as remote and different as the prediction by geneticists of the relative frequency with which various characteristics occur in groups of individuals, the calculation by telephone engineers of the density of telephone traffic, the maintenance by industrial engineers of manufactured products at a certain standard of quality, the transmission (by engineers concerned with the design of communications and automatic control systems) of signals in the presence of noise, and the study by physicists of thermal noise in electric circuits and the

Brownian motion of particles immersed in a liquid or gas.

More details about the history of probability can be found in references [1, 2].

1.2 Experiment, Sample Space and Random Event

Various kinds of chance are well-known to everyone of us from our everyday experience: the outcome of a coin-toss or die-roll, the length of time spent waiting in line, how meteorological phenomena will proceed. In all such situations, we are unable to predict the outcome of an "experiment" or the future course of a process. So our recognition and quantitative characterization of real-word phenomena cannot be complete (indeed being impossible in many situations) unless we account for the great role played by randomness.

A random phenomenon is a situation in which we know what outcomes could happen, but we do not know which particular outcome did or will happen. "Random" in statistics is not a synonym for "haphazard" but a description of a kind of order that emerges only in the long run. So in order to find the essence behind the "Random" phenomena, we need do enough experiments. Usually, such experiments are called random experiment. To describe the outcomes of the random experiments, we need the concepts of sample space and random events.

1.2.1 Basic Definitions

Random experiment usually has the following three characteristics.

(i) Repeatability: it can be repeated under the same conditions.

(ii) Predictability: it has more than one outcome and we know all possible outcomes before the experiment.

(iii) Uncertainty: the outcome of the experiment will not be known in advance.

In fact, our efforts are focused on the random experiments which can be repeated under the same conditions. But we also use the probability method to study some subjects which can not be repeated in application, such as lifetime etc. In this book, we shall abbreviate "random experiment" to experiment and denote it by E. Let us see some examples.

E_1 : Determination of the sex of a newborn child.

E_2 : Roll a die and observe which number appears.

E_3 : Flip two coins and observe the outcomes.

E_4 : Roll two dice and observe the outcomes.

E_5 : Observe call times for a call center.

E_6 : Measure the lifetime of cars.

Now a question is how to record the experiment data after we finish the experiment. Each possible outcome is called an ***sample point*** or ***elementary event*** and is denoted by s or

ω. The set of all possible outcomes of an random experiment (if we could know all possible outcomes) is known as the ***sample space*** of the random experiment and is denoted by S or Ω. In this book, we use the notations ω and Ω.

In the following, we give examples about sample space corresponding to $E_1 \sim E_6$.

Example 1.2.1. For E_1, $\Omega_1 = \{g, b\}$, where the outcome g means that the child is a girl and b that it is a boy.

Example 1.2.2. For E_2, $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$, where the outcome i means that i appeared on the die, $i = 1, 2, 3, 4, 5, 6$.

Example 1.2.3. For E_3, $\Omega_3 = \{(H, H), (H, T), (T, H), (T, T)\}$, where H means head and T means tail.

Example 1.2.4. For E_4,

$$\Omega_4 = \begin{pmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{pmatrix}$$

where the outcome (i, j) is said to occur if i appears on the first die and j appears on the second die, $i, j = 1, 2, \cdots, 6$.

Example 1.2.5. For E_5, $\Omega_5 = 0, 1, 2, \cdots$, where the outcome i is the number of call times, $i = 0, 1, \cdots$.

Example 1.2.6. For E_6, $\Omega_6 = [0, \infty)$, where the outcome t is the lifetime of a car, $0 \leqslant t < \infty$.

Any subset of the sample space S is known as an ***random event*** or ***event***. Events are usually denoted by capital letters $A, B, C, \cdots$. We say that the event A occurs when the outcome of the experiment lies in A. Those events must occur in the experiment are called the ***inevitable events***. We always denote inevitable event by S or Ω (in this book, we use Ω). Sample space Ω is an inevitable event. For Example 1.2.2, $A = \{1, 2, 3, 4, 5, 6, 7, 8\}$ is an inevitable event. Those could not happen anytime are said to be ***impossible events***. We usually denote impossible events by $\varnothing$.

In Example 1.2.1, if $A = \{g\}$, then A is the event that newborn child is a girl.

In Example 1.2.2, if $A = \{2, 4, 6\}$, then A would be the event that an even number appears on the roll.

In Example 1.2.3, if $A = \{(H, H), (H, T)\}$, then A is the event that a head appears on the first coin.

In Example 1.2.4, if $A = \{(1, 3), (2, 2), (3, 1)\}$, then A is the event that the sum of the dice equals four.

In Example 1.2.5, if $A = (2, 5)$, then A is the event that the number of the call times for a call center is 3 or 4.

In Example 1.2.6, if $A = (2, 5)$, then A is the event that the car lasts between two and five years.

1.2.2 Events as Sets

Since random events are subsets of sample space, the relationships and operations between random events could be described in term of set theory.

Let Ω be the sample space of the random experiment E and $A, B, A_i (i = 1, 2, \cdots)$ be the random events of E.

1. It is said that an event A is ***contained in*** another event B if every outcome that belongs to the subset defining the event A also belongs to the subset defining the event B. The relation between two events is expressed symbolically by $\boldsymbol{A \subset B}$. That means ***if event A occurs, then B occurs***. The relation $A \subset B$ is also expressed by saying that A is a subset of B. Equivalently, if $A \subset B$, we may say that B *contains* A and may write $B \supset A$. In the experiment E_2, suppose that $A = \{2, 4, 6\}$ and $C = \{2, 3, 4, 5, 6\}$, it follows that $A \subset C$. It should be noted that $A \subset C$ for every event A.

If two events A and B are so related that $A \subset B$ and $B \subset A$, it follows that A and B must contain exactly the same points. In other words, $A = B$.

2. ***The union*** of the event A and the event B, denoted by $\boldsymbol{A \cup B}$, consists of all outcomes that lie in either A or B. That is, the event $A \cup B$ will occur if ***either A or B occurs***. In the experiment E_1, suppose that $A = \{H\}$ and $B = \{T\}$, then $A \cup B = \{H, T\} = S$.

We also define unions of more than two events in a similar manner.

$$\bigcup_{i=1}^{n} A_i \triangleq A_1 \cup A_2 \cup \cdots \cup A_n$$

is called the union of events $A_1, A_2, \cdots, A_n$ and defined to be the event that consists of all outcomes that are in A_i for at least one value of $i = 1, 2, \cdots, n$.

$$\bigcup_{i=1}^{\infty} A_i \triangleq A_1 \cup A_2 \cup \cdots$$

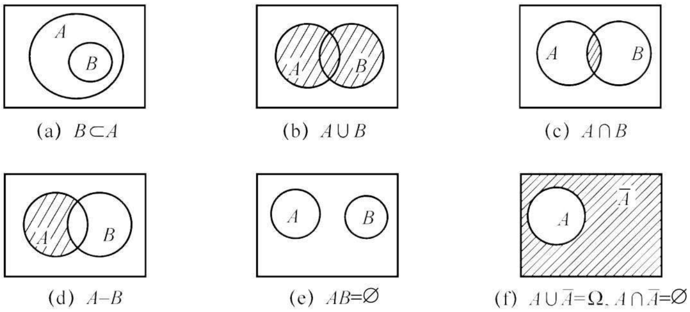

Figure 1.1: The relationships between events A and B

is called the union of events $A_1, A_2, \cdots$.

3. ***The intersection*** of the event A and the event B, denoted by $\boldsymbol{A \cap B}$ or AB, consists of all outcomes which are both in A and B. That is, the event AB will occur only if ***both A and B occur***. In the experiment E_3, if $A = \{(H, H), (T, H), (T, T)\}$ and $B = \{(T, H), (T, T)\}$, then

$$B \subset A, AB = B = \{(T, H), (T, T)\}$$

and thus AB would occur if the outcome is either (T, H) or (T, T).

Similarly, we can define the intersection of events $A_1, A_2, \cdots, A_n$ by the event

$$\bigcap_{i=1}^{n} A_i \triangleq A_1 \cap A_2 \cap \cdots \cap A_n$$

which consists of those outcomes that are in all of the events $A_i, i = 1, 2, \cdots, n$. And

$$\bigcap_{i=1}^{\infty} A_i \triangleq A_1 \cap A_2 \cap \cdots$$

is called the intersection of events $A_1, A_2, \cdots$.

4. ***The difference*** of events A and B, denoted by $\boldsymbol{A - B}$, consists of the outcomes that are in event A but not in event B. That means the event $A - B$ will occur if ***A occurs but B does not occur***. In the experiment E_5, if $A = (2, 12)$ and $B = (4, 13]$, then $A - B = (2, 4]$.

5. Events A and B are ***disjoint or mutually exclusive*** if $\boldsymbol{AB} = \varnothing$, which means AB would not consists of any outcome and hence could not occur. That is, ***A and B can not happen at the same time***.

More generally, the events $A_1, A_2, \cdots$ are said to be ***pairwise disjoint*** or mutually exclusive if $A_i A_j = \varnothing$ whenever $i \neq j$.

6. Event B is said to be the ***complement*** of event A with respect to Ω if $A \cup B = \Omega$ and $AB = \varnothing$. That is, B will occur if and only if ***A does not occur***. B is usually denoted by $\overline{A}$ or A^c. In the experiment E_4, if $A = \{(1,3),(2,2),(3,1)\}$, then $\overline{A}$ will occur if the sum of the dice does not equal four.

Since events could be regarded as sets, we can describe their relationships by Venn diagram, see Fig. 1.1. The jargons in probability theory are different from those in set theory. Here we list their connections and differences in Table1.1.

Table 1.1: The jargons in set theory and probability theory

Typical notation	Set jargon	Probability jargon
Ω	Collection of objects	Sample space
ω	Member of Ω	Elementary event, outcome
A	Subset of Ω	Events that some outcome in A occurs
A or $\overline{A}$	Complement of A	Event that no outcome in A occurs
$A \cap B$	Intersection	Both A and B
$A \cup B$	Union	Either A or B or both
$A - B$	Difference	A, but no B
$A \subseteq B$	Inclusion	If A, then B
$\varnothing$	Empty set	Impossible event
Ω	Whole space	Certain set

Let A, B and C be the random events of experiment E. The operations of the events will satisfy the following rules:

(i) Commutatively $\quad A \cup B = B \cup A,\ AB = BA.$

(ii) Associatively $\quad A \cup (B \cup C) = (A \cup B) \cup C = A \cup B \cup C.$

$A(BC) = (AB)C = ABC.$

(iii) Distributively $\quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$

$A(B \cup C) = (AB) \cup (AC).$

(iv) De Morgan's law $\quad \overline{A \cup B} = \overline{A} \cap \overline{B},\ \overline{A \cap B} = \overline{A} \cup \overline{B}.$

More generally, $\overline{(\bigcup_i A_i)} = \bigcap \overline{A_i},\ \overline{(\bigcap_i A_i)} = \bigcup \overline{A_i}.$

In addition, there are some common properties, such as

(1) $\overline{\overline{A}} = A$.

(2) $A \cup B \supset A$ and $A \cup B \supset B$. In particular, if $A \subset B$, then $A \cup B = B$.

(3) $A \cap B \subset A$ and $A \cap B \subset B$. In particular, if $A \subset B$, then $A \cap B = A$.

(4) $A - B = A - AB = A\overline{B}$.

(5) $A \cup B = A \cup \overline{A}B$.

Example 1.2.7. Consider the sample space Ω consists of all positive integers less than 10, i.e., $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Let A be the event consisting of all even numbers and B be the event consisting of numbers divisible by 3. Find $\overline{A}, A \cup B, AB$ and $\overline{A}B$.

Solution. We have $A = \{2, 4, 6, 8\}$ and $B = \{3, 6, 9\}$. Thus

$\overline{A} = \{1, 3, 5, 7, 9\}$, $A \cup B = \{2, 3, 4, 6, 8, 9\}$, $AB = \{6\}$ and $\overline{A}B = \{3, 9\}$. ■

Example 1.2.8. Suppose that A, B and C are three events and $D = \{$at least one of the three events will occur$\}$. Try to describe the event D by events A, B and C.

Solution. We describe the event D by three different ways:

(i) directed method: $D = A \cup B \cup C$,

(ii) decomposition method: $D = A\overline{B}\,\overline{C} \cup \overline{A}B\overline{C} \cup \overline{A}\,\overline{B}C \cup \overline{A}BC \cup A\overline{B}C \cup AB\overline{C} \cup ABC$,

(iii) inverse method: $D = \overline{\overline{D}} = \overline{\{A,\ B \text{ and } C \text{ will not occur}\}} = \overline{\overline{A}\,\overline{B}\,\overline{C}}$. ■

1.3 Probabilities Defined on Events

The events in random experiment may occur or not. The problem of discussing the chance they occur is probability problem. In this section, we give three different interpretations about probability: classical probability, geometric probability and frequency. We also calculate many concrete probabilities of specific events in different cases.

1.3.1 Classical Probability

The classical definition or interpretation of probability is identified with the works of Jacob Bernoulli and Pierre-Simon Laplace. It is based on the concept of equally likely outcomes. For example, when a fair coin is tossed, there are two possible outcomes: a head or a tail. If we assume that the sum of these outcomes' probabilities is 1, then both the probability of a head and the probability of a tail must be 1/2 respectively. More generally, if the outcome of some experiment is one of n different outcomes, and if these n outcomes are equally likely to occur, then the probability of each outcome is $1/n$.

Generally, a random experiment E is ***classical*** if

(i) E contains only different limited basic events, that is, $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$. We call this kind of sample space ***simple space***, and