Universitext

Leonid B. Koralov
Yakov G. Sinai

Theory of Probability and Random Processes

Second Edition

概率论和随机过程 第2版

Springer

世界图书出版公司
www.wpcbj.com.cn

Leonid B. Koralov   Yakov G. Sinai

Theory of Probability and Random Processes

Second Edition

Springer

图书在版编目（CIP）数据

概率论和随机过程：第2版＝Theory of Probability and Random Processes：英文/（美）凯罗勒夫著．—影印本．—北京：世界图书出版公司北京公司，2012.3
ISBN 978－7－5100－4410－6

Ⅰ．①概…　Ⅱ．①凯…　Ⅲ．①概率论—高等学校—教材—英文
②随机过程—高等学校—教材—英文　Ⅳ．①O21

中国版本图书馆CIP数据核字（2012）第030523号

书　　名： Theory of Probability and Random Processes 2nd ed.
作　　者： Leonid B. Koralov，Yakov G. Sinai
中 译 名： 概率论和随机过程 第2版
责任编辑： 高蓉　刘慧

出 版 者： 世界图书出版公司北京公司
印 刷 者： 三河市国英印务有限公司
发　　行： 世界图书出版公司北京公司（北京朝内大街137号　100010）
联系电话： 010－64021602，010－64015659
电子信箱： kjb@wpcbj.com.cn

开　　本： 24开
印　　张： 15.5
版　　次： 2012年06月
版权登记： 图字：01－2012－1249

书　　号： 978－7－5100－4410－6/O·934　　**定　　价：** 55.00元

Leonid B. Koralov
Yakov G. Sinai
Department of Mathematics
Princeton University
Princeton, NJ 08544, USA

e-mail: koralov@math.princeton.edu
sinai@math.princeton.edu

Mathematics Subject Classification (2000): 60xx

Library of Congress Control Number: 2007931050

ISBN: 978-3-540-25484-3 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springer.com

Preface

This book is primarily based on a one-year course that has been taught for a number of years at Princeton University to advanced undergraduate and graduate students. During the last year a similar course has also been taught at the University of Maryland.

We would like to express our thanks to Ms. Sophie Lucas and Prof. Rafael Herrera who read the manuscript and suggested many corrections. We are particularly grateful to Prof. Boris Gurevich for making many important suggestions on both the mathematical content and style.

While writing this book, L. Koralov was supported by a National Science Foundation grant (DMS-0405152). Y. Sinai was supported by a National Science Foundation grant (DMS-0600996).

Leonid Koralov
Yakov Sinai

Contents

Part I Probability Theory

Part I

Probability Theory

1

Random Variables and Their Distributions

1.1 Spaces of Elementary Outcomes, σ-Algebras, and Measures

The first object encountered in probability theory is the space of elementary outcomes. It is simply a non-empty set, usually denoted by Ω, whose elements $\omega \in \Omega$ are called elementary outcomes. Here are several simple examples.

Example. Take a finite set $X = \{x^1, ..., x^r\}$ and the set Ω consisting of sequences $\omega = (\omega_1, ..., \omega_n)$ of length $n \geq 1$, where $\omega_i \in X$ for each $1 \leq i \leq n$. In applications, ω is a result of n statistical experiments, while ω_i is the result of the i-th experiment. It is clear that $|\Omega| = r^n$, where $|\Omega|$ denotes the number of elements in the finite set Ω. If $X = \{0, 1\}$, then each ω is a sequence of length n made of zeros and ones. Such a space Ω can be used to model the result of n consecutive tosses of a coin. If $X = \{1, 2, 3, 4, 5, 6\}$, then Ω can be viewed as the space of outcomes for n rolls of a die.

Example. A generalization of the previous example can be obtained as follows. Let X be a finite or countable set, and I be a finite set. Then $\Omega = X^I$ is the space of all functions from I to X.

If $X = \{0, 1\}$ and $I \subset \mathbb{Z}^d$ is a finite set, then each $\omega \in \Omega$ is a configuration of zeros and ones on a bounded subset of d-dimensional lattice. Such spaces appear in statistical physics, percolation theory, etc.

Example. Consider a lottery game where one tries to guess n distinct numbers and the order in which they will appear out of a pool of r numbers (with $n \leq r$). In order to model this game, define $X = \{1, ..., r\}$. Let Ω consist of sequences $\omega = (\omega_1, ..., \omega_n)$ of length n such that $\omega_i \in X, \omega_i \neq \omega_j$ for $i \neq j$, and $X = \{1, ..., r\}$. It is easy to show that $|\Omega| = r!/(r - n)!$.

Later in this section we shall define the notion of a probability measure, or simply probability. It is a function which ascribes real numbers between zero and one to certain (but not necessarily all!) subsets $A \subseteq \Omega$. If Ω is interpreted as the space of possible outcomes of an experiment, then the probability of A may be interpreted as the likelihood that the outcome of the experiment belongs to A. Before we introduce the notion of probability we need to discuss the classes of sets on which it will be defined.

Definition 1.1. *A collection $\mathcal{G}$ of subsets of Ω is called an algebra if it has the following three properties.*

1. $\Omega \in \mathcal{G}$.
2. $C \in \mathcal{G}$ *implies that* $\Omega \backslash C \in \mathcal{G}$.
3. $C_1, ..., C_n \in \mathcal{G}$ *implies that* $\bigcup_{i=1}^n C_i \in \mathcal{G}$.

Example. Given a set of elementary outcomes Ω, let $\underline{\mathcal{G}}$ contain two elements: the empty set and the entire set Ω, that is $\underline{\mathcal{G}} = \{\emptyset, \Omega\}$. Define $\overline{\mathcal{G}}$ as the collection of all the subsets of Ω. It is clear that both $\underline{\mathcal{G}}$ and $\overline{\mathcal{G}}$ satisfy the definition of an algebra. Let us show that if Ω is finite, then the algebra $\overline{\mathcal{G}}$ contains $2^{|\Omega|}$ elements.

Take any $C \subseteq \Omega$ and introduce the function $\chi_C(\omega)$ on Ω:

$$\chi_C(\omega) = \begin{cases} 1 & \text{if } \omega \in C, \\ 0 & \text{otherwise,} \end{cases}$$

which is called the indicator of C. It is clear that any function on Ω taking values zero and one is an indicator function of some set and determines this set uniquely. Namely, the set consists of those ω, where the function is equal to one. The number of distinct functions from Ω to the set $\{0, 1\}$ is equal to $2^{|\Omega|}$.

Lemma 1.2. *Let Ω be a space of elementary outcomes, and $\mathcal{G}$ be an algebra. Then*

1. *The empty set is an element of* $\mathcal{G}$.
2. *If* $C_1, ..., C_n \in \mathcal{G}$, *then* $\bigcap_{i=1}^n C_i \in \mathcal{G}$.
3. *If* $C_1, C_2 \in \mathcal{G}$, *then* $C_1 \setminus C_2 \in \mathcal{G}$.

Proof. Take $C = \Omega \in \mathcal{G}$ and apply the second property of Definition 1.1 to obtain that $\emptyset \in \mathcal{G}$. To prove the second statement, we note that

$$\Omega \backslash \bigcap_{i=1}^n C_i = \bigcup_{i=1}^n (\Omega \backslash C_i) \in \mathcal{G}.$$

Consequently, $\bigcap_{i=1}^n C_i \in \mathcal{G}$. For the third statement, we write

$$C_1 \setminus C_2 = \Omega \setminus ((\Omega \setminus C_1) \cup C_2) \in \mathcal{G}.$$

□

Lemma 1.3. *If an algebra $\mathcal{G}$ is finite, then there exist non-empty sets $B_1, ..., B_m \in \mathcal{G}$ such that*

1. *$B_i \cap B_j = \emptyset$ if $i \neq j$.*
2. *$\Omega = \bigcup_{i=1}^m B_m$.*
3. *For any set $C \in \mathcal{G}$ there is a set $I \subseteq \{1, ..., m\}$ such that $C = \bigcup_{i \in I} B_i$ (with the convention that $C = \emptyset$ if $I = \emptyset$).*

Remark 1.4. The collection of sets $B_i, i = 1, ..., m$, defines a partition of Ω. Thus, finite algebras are generated by finite partitions.

Remark 1.5. Any finite algebra $\mathcal{G}$ has 2^m elements for some integer $m \in \mathbb{N}$. Indeed, by Lemma 1.3, there is a one-to-one correspondence between $\mathcal{G}$ and the collection of subsets of the set $\{1, ..., m\}$.

Proof of Lemma 1.3. Let us number all the elements of $\mathcal{G}$ in an arbitrary way:

$$\mathcal{G} = \{C_1, ..., C_s\}.$$

For any set $C \in \mathcal{G}$, let

$$C^1 = C, \quad C^{-1} = \Omega \backslash C.$$

Consider a sequence $b = (b_1, ..., b_s)$ such that each b_i is either $+1$ or -1 and set

$$B^b = \bigcap_{i=1}^s C_i^{b_i}.$$

From the definition of an algebra and Lemma 1.2 it follows that $B^b \in \mathcal{G}$. Furthermore, since

$$C_i = \bigcup_{b: b_i = 1} B^b,$$

any element C_i of $\mathcal{G}$ can be obtained as a union of some of the B^b. If $b' \neq b''$, then $B^{b'} \cap B^{b''} = \emptyset$. Indeed, $b' \neq b''$ means that $b_i{}' \neq b_i{}''$ for some i, say $b_i{}' = 1, b_i{}'' = -1$. In the expression for $B^{b'}$ we find $C_i^1 = C_i$, so $B^{b'} \subseteq C_i$. In the expression for $B^{b''}$ we find $C_i^{-1} = \Omega \backslash C_i$, so $B^{b''} \subseteq \Omega \backslash C_i$. Therefore, all B^b are pair-wise disjoint. We can now take as B_i those B^b which are not empty. □

Definition 1.6. *A collection $\mathcal{F}$ of subsets of Ω is called a σ-algebra if $\mathcal{F}$ is an algebra which is closed under countable unions, that is $C_i \in \mathcal{F}$, $i \geq 1$, implies that $\bigcup_{i=1}^\infty C_i \in \mathcal{F}$. The elements of $\mathcal{F}$ are called measurable sets, or events.*

As above, the simplest examples of a σ-algebra are the trivial σ-algebra, $\underline{\mathcal{F}} = \{\emptyset, \Omega\}$, and the σ-algebra $\overline{\mathcal{F}}$ which consists of all the subsets of Ω.

Definition 1.7. *A measurable space is a pair $(\Omega, \mathcal{F})$, where Ω is a space of elementary outcomes and $\mathcal{F}$ is a σ-algebra of subsets of Ω.*

Remark 1.8. A space of elementary outcomes is said to be discrete if it has a finite or countable number of elements. Whenever we consider a measurable space $(\Omega, \mathcal{F})$ with a discrete space Ω, we shall assume that $\mathcal{F}$ consists of all the subsets of Ω.

The following lemma can be proved in the same way as Lemma 1.2.

Lemma 1.9. *Let $(\Omega, \mathcal{F})$ be a measurable space. If $C_i \in \mathcal{F}$, $i \geq 1$, then $\bigcap_{i=1}^{\infty} C_i \in \mathcal{F}$.*

It may seem that there is little difference between the concepts of an algebra and a σ-algebra. However, such an appearance is deceptive. As we shall see, any interesting theory (such as measure theory or probability theory) requires the notion of a σ-algebra.

Definition 1.10. *Let $(\Omega, \mathcal{F})$ be a measurable space. A function $\xi : \Omega \to \mathbb{R}$ is said to be $\mathcal{F}$-measurable (or simply measurable) if $\{\omega : a \leq \xi(\omega) < b\} \in \mathcal{F}$ for each $a, b \in \mathbb{R}$.*

Below we shall see that linear combinations and products of measurable functions are again measurable functions. If Ω is discrete, then any real-valued function on Ω is a measurable, since $\mathcal{F}$ contains all the subsets of Ω.

In order to understand the concept of measurability better, consider the case where $\mathcal{F}$ is finite. Lemma 1.3 implies that $\mathcal{F}$ corresponds to a finite partition of Ω into subsets $B_1, ..., B_m$, and each $C \in \mathcal{F}$ is a union of some of the B_i.

Theorem 1.11. *If ξ is $\mathcal{F}$-measurable, then it takes a constant value on each element of the partition B_i, $1 \leq i \leq m$.*

Proof. Suppose that ξ takes at least two values, a and b, with $a < b$ on the set B_j for some $1 \leq j \leq m$. The set $\{\omega : a \leq \xi(w) < (a+b)/2\}$ must contain at least one point from B_j, yet it does not contain the entire set B_j. Thus it can not be represented as a union of some of the B_i, which contradicts the $\mathcal{F}$-measurability of the set. □

Definition 1.12. *Let $(\Omega, \mathcal{F})$ be a measurable space. A function $\mu : \mathcal{F} \to [0, \infty)$ is called a finite non-negative measure if*

$$\mu\left(\bigcup_{i=1}^{\infty} C_i\right) = \sum_{i=1}^{\infty} \mu(C_i)$$

whenever $C_i \in \mathcal{F}$, $i \geq 1$, are such that $C_i \cap C_j = \emptyset$ for $i \neq j$.

The property expressed in Definition 1.12 is called the countable additivity (or the σ-additivity) of the measure.

Remark 1.13. Most often we shall omit the words finite and non-negative, and simply refer to μ as a measure. Thus, a measure is a σ-additive function on $\mathcal{F}$ with values in $\mathbb{R}^+$. In contrast, σ-finite and signed measures, to be introduced in Chapter 3, take values in $\mathbb{R}^+ \cup \{+\infty\}$ and $\mathbb{R}$, respectively.

Definition 1.14. *Let g be a binary function on Ω with values* 1 (*true*) *and* 0 (*false*). *It is said that g is true almost everywhere if there is an event C with $\mu(C) = \mu(\Omega)$ such that $g(\omega) = 1$ for all $\omega \in C$.*

Definition 1.15. *A measure* P *on a measurable space $(\Omega, \mathcal{F})$ is called a probability measure or a probability distribution if* $\mathrm{P}(\Omega) = 1$.

Definition 1.16. *A probability space is a triplet $(\Omega, \mathcal{F}, \mathrm{P})$, where $(\Omega, \mathcal{F})$ is a measurable space and* P *is a probability measure. If $C \in \mathcal{F}$, then the number* $\mathrm{P}(C)$ *is called the probability of C.*

Definition 1.17. *A measurable function defined on a probability space is called a random variable.*

Remark 1.18. When P is a probability measure, the term "almost surely" is often used instead of "almost everywhere".

Remark 1.19. Let us replace the σ-additivity condition in Definition 1.12 by the following: if $C_i \in \mathcal{F}$ for $1 \leq i \leq n$, where n is finite, and $C_i \cap C_j = \emptyset$ for $i \neq j$, then

$$\mu(\bigcup_{i=1}^{n} C_i) = \sum_{i=1}^{n} \mu(C_i) .$$

This condition leads to the notion of a finitely additive function, instead of a measure. Notice that finite additivity implies superadditivity for infinite sequences of sets. Namely,

$$\mu(\bigcup_{i=1}^{\infty} C_i) \geq \sum_{i=1}^{\infty} \mu(C_i)$$

if the sets C_i are disjoint. Indeed, otherwise we could find a sufficiently large n such that

$$\mu(\bigcup_{i=1}^{\infty} C_i) < \sum_{i=1}^{n} \mu(C_i) ,$$

which would violate the finite additivity.

Let Ω be discrete. Then $\mathrm{p}(\omega) = \mathrm{P}(\{\omega\})$ is the probability of the elementary outcome ω. It follows from the definition of the probability measure that

1. $\mathrm{p}(\omega) \geq 0$.
2. $\sum_{\omega \in \Omega} \mathrm{p}(\omega) = 1$.