Probability, Statistics and Stochastic Processes

This book is divided into three parts;
Part I Probability, Part II Statistics and Part III Stochastic processes. The discussion ranges from basic concepts such as conditional probability and independence to more advanced concepts such as limit theorems and stochastic processes. Some research topics such as the rapid mixing property of reversible Markov chains are also discussed. In this way, the reader can successfully build a bridge from these basic statistical ideas to more advanced applications and research topics.

cupress.chula.ac.th

Probability, statistic...
ISBN 978-974-03-4025-6
9 789740 340256
C322
2240200 330.00 บาท

cupress
Chulalongkorn University Press
Phayathai Road, Pathumwan, Bangkok 10330
Tel. 0-2218-3269-70 Fax. 0-2218-3547
e-mail:cupress@chula.ac.th
www.cupress.chula.ac.th
Knowledge to All

Production of this printed book helps reducing global warming effect by 100% compensating the amount of greenhouse gases emissions
0.53 kg. CO_2 eq/book

Chulalongkorn University press

Probability, Statistics and Stochastic Processes

Seksan Kiatsupaibul

CUP. 2494

cupress

Probability, Statistics and Stochastic Processes

Seksan Kiatsupaibul

Probability, Statistics and Stochastic Processes

Probability, Statistics and Stochastic Processes

Seksan Kiatsupaibul

Chulalongkorn University press

2021

330.-

Seksan Kiatsupaibul
Probability, statistics and stochastic processes / Seksan Kiatsupaibul

1. Probability. 2. Statistics. 3. Stochastic processes.

519.2
ISBN 978-974-03-4025-6
CUP 2496

Knowledge to All
www.cupress.chula.ac.th

Copyright © 2021, by Chulalongkorn University Press
First Printing 2021 600 copies

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying recording, or otherwise, without prior written permission of the publisher.

Distributed by Chulalongkorn University Book Center
Phayathai Road, Pathumwan, Bangkok 10330, Thailand
Branch Sala Prakiew Tel. 0-2218-7000-3 Fax. 0-2255-4441
Siam Square Tel. 0-2218-9881-2 Fax. 0-2254-9495
Naresuan University (Phitsanulok) Tel. 0-5526-0162-4 Fax. 0-5526-0165
Suranaree University (Nakhonratchasrima) Tel. 0-4421-6131-4 Fax. 0-4421-6135
Burapha University (Chonburi) Tel. 0-3839-4855-9 Fax. 0-3839-3239
Chulachomklao Royal Military Academy (Nakhon-Nayok) Tel. 0-3739-3023 Fax. 0-3739-3023
University of Phayao Tel. 0-5446-6799-800 Fax. 0-5446-6798
Chamchuri Square (4th Floor) Tel. 0-2160-5300-1 Fax. 0-2160-5304
Huamark Tel. 0-2374-1378 Fax. 0-2374-1377
Rajamangala University of Technology Isan Tel.0-4423-0000 Fax.0-4423-3052
Call Center Tel. 0-2255-4433 http://chulabook.com and Networks
Wholesale Huamark (Intersection Huamark) Tel. 0-2374-1374-5 Fax. 0-2374-1377

Editorial Department: Chosita Watcharothai
Artwork: Seksan Kiatsupabul
Cover: Khomsan Kiatsupaibul
Printed by: Chulalongkorn University Press Tel. 0-2218-3549-50 Fax. 0-2218-3551

Preface

In the Department of Statistics at Chulalongkorn Business School, the interest of students are diverse. Students' concentrations range from statistical methodology and actuarial science to information technology and data science. However, there is one common background that they all need, which is a solid statistical knowledge. Therefore, it is mandatory for our students both at undergraduate and graduate levels to be well equipped with probability and statistical concepts. Nevertheless, traditional statistics textbooks that discuss only independent sequences seem to be inadequate for this group of students with diverse interests. Dependent sequences or stochastic processes are also very important in key subjects that students with specific concentrations are required to take. Therefore, this textbook is intended to cover basic non-measure theoretic statistical concepts of both independent sequences and stochastic processes that will prepare students well for their future careers in our undergraduate and graduate programs.

This book can be divided into three parts; Part I Probability, Part II Statistics and Part III Stochastic processes. Probability is covered by the first seven chapters. Statistics ranges from Chapter 8 to Chapter 11. Stochastic processes are discussed in the last three chapters, Chapter 12 to Chapter 14. The connections among the three parts are developed over the course of the discussions. Graphics are supplied where possible to facilitate understanding of the ideas. Exercises are provided not only for practice but also as an extension of the topics presented in each chapter. The discussion ranges from the most basic concepts such as conditional probability and independence to more advanced concepts such as limit theorems and stochastic processes. We even touch on some research topics such as the rapid mixing property of a reversible Markov chains at the end. In this way, I hope that the readers will successfully build a bridge from the basic statistical ideas to the applications and research topics that they engage in.

In Part I Probability, Chapter 1 defines a probability space, the mathematical foundation upon which all statistical objects are based. Chapter 2 elaborates to the concepts of conditional probability and independence. Chapter 3 defines random variables, probability distribution functions and summary measurement such as expectations and variances. Chapter 4 introduces basic probability distribution functions of both discrete and continuous types. Chapter 5 discusses joint random variables and independent random variables. Chapter 6 extends conditional probability to conditional expectation, preparing for a study on dependence sequences or stochastic processes in Part III Stochastic

Processes. In Chapter 7, we apply additional mathematical techniques to the concepts introduced in the first five chapters, resulting in some useful probability distributions that will be used in Part II Statistics.

In Part II Statistics, Chapter 8 extends the idea of independent random variables to a sequence of independent random variables. A natural consequence is the law of large numbers and the central limit theorem. A random sample is also defined, preparing for a discussion of statistical theory in subsequent chapters. Chapters 9 and 10 discuss basic statistical inference theory. Chapter 9 discusses statistical estimation and confidence intervals. Chapter 10 discusses basic hypothesis testing procedures. Simple linear regression and multiple linear regression are discussed in Chapter 11. Confidence intervals and hypothesis testing are applied to form statistical inferences for linear regression model parameters.

In Part III Stochastic Processes, Chapter 12 extends the idea of joint random variables to a sequence of joint random variables or a stochastic process. We limit our discussion to only discrete time and discrete state space stochastic processes. A simple random walk is introduced as one of the most basic discrete time discrete state space stochastic processes. Chapter 13 introduces Markov chains and their limit theorems. In Chapter 14, the discussion of the Markov chains is narrowed down to reversible chains. Advanced topics on the rapid mixing property of a reversible Markov chain is also discussed in this last chapter.

I would like to thank Professor Robert L. Smith for his inspiration over the years. My special gratitude also goes to Professor Anthony Hayter who helped review the book from when it was only an incomplete manuscript. I would also like to thank the anonymous reviewers whose comments significantly help improve the readability of the final version. Chulalongkorn Business School provides an excellent environment for academic pursuits, and because of that this book has been made possible.

Seksan Kiatsupaibul
January, 2021
Chulalongkorn Business School
Bangkok, Thailand

Contents

List of Figures

Chapter 1

Probability Space

1.1 Introduction

People usually make decisions in the face of uncertainty. Better understanding the nature of uncertainty helps people make better decisions. A scientific way of understanding uncertainty is through probability and statistics. This textbook presents basic concepts of probability, statistics and stochastic processes. The main objective is to provide basic tools for decision making under uncertainty.

Let us begin our investigation by considering an example of a coin tossing experiment. In this first example, a fair coin is tossed two times. Each time the coin may turn head or tail with equal chance. No other outcomes are considered. For example, we do not consider the outcome of the coin landing on its side. Note that the order of the two tosses is also of concern. Therefore, there are totally four possible outcomes, which are

1. head and head (HH),
2. head and tail (HT),
3. tail and head (TH),
4. tail and tail (TT).

Observe that HT and TH are considered two different outcomes since the order counts. This random experiment will guide us through most of the early concepts of probability.

1.2 Probability Space

A random experiment is an activity whose outcome is uncertain. To completely describe a random experiment, one requires three elements. The first element is the set of all possible outcomes. The second element is the events that we are interested in. The third element is the probability or the degree that an event would occur. These three elements combine to completely capture a random experiment. We call the combination of the three a *probability space*. We denote the probability space by a triplet notation $(\Omega, \mathcal{A}, \mathbb{P})$. In the remaining of this chapter, we consider each element in detail.

1.2.1 Sample Space

The first element of the probability space is the sample space, denoted by Ω. It is the set or the collection of all possible outcomes from the random experiment.

Definition 1.1 *The sample space Ω is the set of all possible outcomes from the random experiment.*

Example 1.1 *Consider a random experiment of one toss of a coin. There are two possible outcomes: head (H) or tail (T). Let Ω_1 be its sample space. Using set notations, we then write*

$$\Omega_1 = \{H, T\}$$

Example 1.2 *Consider a random experiment of two tosses of a coin. There are four possible outcomes: HH, HT, TH, TT.*

$$\Omega_2 = \{HH, HT, TH, TT\}$$

1.2.2 Events

An *event* is a collection of some outcomes.

Definition 1.2 *Given a sample space Ω, an event is a subset of Ω.*

It represents an event in a random experiment. *An event occurs when an outcome in it becomes the experiment result.*

Example 1.3 *Consider the example of two tosses of a coin where $\Omega_2 = \{HH, HT, TH, TT\}$. An example of an event is*

"the event that the experiment result is head followed by head."

We usually denote an event by a capital letter in the English alphabet. For example, if we denote the above event by E_1, then

E_1 = *the event that the result is head followed by head.*

This event can be written in set notation as

$$E_1 = \{HH\}.$$

Note that E_1 is a subset of Ω_2.

Consider another example:

"the event that there is exactly one head from the two tosses."

If we denote this event by E_2, then

$$E_2 = \{HT, TH\}.$$

E_2 is also a subset of Ω_2. There are two outcomes in this event. The event E_2 is said to occur if either of the two outcomes appears to be the result of the experiment.

Consider another example:

"the event that there is at least one head from the two tosses."

Denote this by E_3.

$$E_3 = \{HH, HT, TH\}.$$

Again E_3 is a subset of Ω_2. E_3 occurs if one of the three outcomes appears to be the result of the experiment.

From the above examples, observe that we can use either words or set notations to describe events. In some cases, set notations are more concised. In other cases, verbal expression are more efficient. In what follows, we switch between the two representations of events, employing the one that is most efficient for the problem at hand. Let us keep in mind that all events described in set notations have their verbal counterpart.

For a technical reason, in every random experiment, we always have two special events, namely *the impossible event* and *the sure event*. The impossible event is represented by the empty set, $\emptyset$. It is a convenient way to collectively combine all unwanted outcomes into a single notion. For example, an outcome of a coin landing on its side is considered impossible in our analysis. We then use $\emptyset$ to capture this notion. The sure event is represented by the sample space, Ω. Since every random experiment will end up in an outcome, the sample space considered as an event will always occur, making it the sure event.

We use the notation $\mathcal{A}$ to denote the collection of events that we aim at assigning probability. In a finite sample space, we usually let $\mathcal{A}$ be the set of all possible events or the power set of the sample space, denoted by 2^Ω.

Example 1.4 *In Example 1.1, the sample space Ω_1 consists of two outcomes. Let $\mathcal{A}_1$ denote the collection of all possible events from Ω_1. Then $\mathcal{A}_1$ consists of four events.*

$$\mathcal{A}_1 = \{\emptyset, \{H\}, \{T\}, \Omega_1\}.$$

In Example 1.2, the sample space Ω_2 consists of four outcomes. Let $\mathcal{A}_2$ denote the collection of all possible events from Ω_2. Then $\mathcal{A}_2$ consists of sixteen events.

$$\begin{aligned}\mathcal{A}_2 = \{&\emptyset, \{HH\}, \{HT\}, \{TH\}, \{TT\}, \{HH, HT\}, \{HH, TH\}, \{HH, TT\},\\ &\{HT, TH\}, \{HT, TT\}, \{TH, TT\}, \{HH, HT, TH\}\\ &\{HH, HT, TT\}, \{HH, TH, TT\}, \{HT, TH, TT\}, \Omega_2\}.\end{aligned}$$

In fact, the collection of events, also referred to as event sets, $\mathcal{A}$ does not need to contain all possible events. An event set $\mathcal{A}$ can be any collection of events that satisfies the following three properties.

(i) The sure event Ω and the impossible event $\emptyset$ are in $\mathcal{A}$.

(ii) If event A is in $\mathcal{A}$, then A^c is also in $\mathcal{A}$.

(iii) If $\{A_1, A_2, \ldots,\}$ is a countable sequence of events in $\mathcal{A}$, then $\cup_{i=1}^{\infty} A_i$ is in $\mathcal{A}$.

An event set $\mathcal{A}$ that satisfies the above three properties is called a σ-algebra or a σ-field. Note that, in property (iii), countable sequences include finite sequence. Note also that property (ii) and property (iii) imply that, if $\{A_1, A_2, \ldots\}$ is a sequence of events in $\mathcal{A}$, then the intersection $\cap_{i=1}^{\infty} A_i$ must be in $\mathcal{A}$. See Problem 1.3 in the Exercises. The power set of the sample space, 2^{Ω}, is always a σ-algebra. In what follows, for a finite sample space Ω, if not specified otherwise, the σ-algebra $\mathcal{A}$ is taken to be 2^{Ω}.

Example 1.5 *In Example 1.2, the sample space Ω_2 consists of four outcomes. The following three event sets, $\mathcal{A}_{2.0}, \mathcal{A}_{2.1}, \mathcal{A}_{2.2}$, are all σ-algebra.*

$$\mathcal{A}_{2.0} = \{\emptyset, \Omega_2\}.$$

$$\mathcal{A}_{2.1} = \{\emptyset, \{HH, HT\}, \{TT, TH\}, \Omega_2.\}$$

$$\begin{aligned}\mathcal{A}_{2.2} = \mathcal{A}_2 = \{&\emptyset, \{HH\}, \{HT\}, \{TH\}, \{TT\}, \{HH, HT\}, \{HH, TH\}, \\ &\{HH, TT\}, \{HT, TH\}, \{HT, TT\}, \{TH, TT\}, \{HH, HT, TH\} \\ &\{HH, HT, TT\}, \{HH, TH, TT\}, \{HT, TH, TT\}, \Omega_2\}.\end{aligned}$$

The first event set, $\mathcal{A}_{2.0}$, is a trivial σ-algebra. The third event set, $\mathcal{A}_{2.2}$, is the set of all possible events from Ω_2. The second event set, $\mathcal{A}_{2.1}$, can be regarded as the event set when we are interested in only the result of the first toss. Observe that, in $\mathcal{A}_{2.1}$, we put $\{HH, HT\}$ into the same event and does not differentiate between outcome HH and HT since both outcomes have H as the result from the first toss. We can construct more σ-algebras from Ω_2. See Problem 1.4 in Exercises.

Before moving on to the third element, let us define two technical terms, a *simple event* and *mutually exclusive events*. A simple event is an event that has only one outcome. When we say that an outcome occurs, we mean the simple event that contains only that specific outcome occurs.

Two events are called mutually exclusive if they cannot occur simultaneously. In other words, there are no common outcomes between the two events. In Example 1.3, E_1 and E_2 are mutually exclusive since the single outcome in E_1 is not in E_2, and no outcomes in E_2 appear in E_1. However, E_1 and E_3 are not mutually exclusive since they have HH as their common outcome. In addition, E_2 and E_3 are not mutually exclusive since both outcomes in E_2 appear in E_3. In set terminology, we say that two events, say A and

B, are mutually exclusive if their intersection are empty, i.e. $A \cap B = \emptyset$. For example, $E_1 \cap E_2 = \emptyset$, so E_1 and E_2 are mutually exclusive. (Recall that an intersection of two sets is the operation to create one new set by taking only the common members of the two.)

1.2.3 Probability

The third element of the probability space is the probability assigned to each event. The probability assigned to an event expresses the degree of chance at which that event would occur. For an event E the probablity of E is denoted by $\mathbb{P}(E)$. Probability cannot be arbitrary. It has to satisfy the following three rules

Definition 1.3 *Given the sample space Ω and the event set $\mathcal{A}$ that is a σ-algebra, the probability, denoted by $\mathbb{P}$, is a nonnegative real value function assigned to each event in $\mathcal{A}$, i.e., $\mathbb{P} : \mathcal{A} \to \Re$, such that:*

(i) $\mathbb{P}(\Omega) = 1$ *and* $\mathbb{P}(\emptyset) = 0$,

(ii) $0 \le \mathbb{P}(E) \le 1$ *for each event E in $\mathcal{A}$,*

(iii) For two mutually exclusive events E_1 and E_2 in $\mathcal{A}$,

$$\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$$

Rule (i) states that the probability of the sure event to occur is 1. On the other hand, the probability of the impossible event to occur is zero. It sets the upper bound and the lower bound for the probability. Rule (ii) states that the probabilities of all other events must lie between these two bounds. An event with probability 1 would surely occur. In contrast, an event with probability 0 has no chance to occur. The higher the probability, the more chance the event would occur.

In Rule (iii), $\cup$ denotes the union in set operations. Rule (iii) states that, if a big event is decomposed into two mutually exclusive smaller events, the probability of the big event is the sum of the probabilities of the two smaller components.

In fact, Rule (iii) can be extended to include a combination of more than two sets. In general, Rule (iii) can be replaced by the following Rule (iii').

(iii') $\mathbb{P}(\cup_{i=1}^{\infty} E_i) = \displaystyle\sum_{i=1}^{\infty} \mathbb{P}(E_i)$ where $E_i \cap E_j = \emptyset$ for $i \neq j$. (E_i and E_j are mutually exclusive for all i and j.)

The union is taken over $i = 1, 2, 3, \ldots$, as well as the summation.

Example 1.6 *Consider the random experiment of one toss of a coin. We usually assign probability to all of the events as follows.*

$$\mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(\Omega_1) = 1, \quad \mathbb{P}(\{H\}) = \frac{1}{2}, \quad \mathbb{P}(\{T\}) = \frac{1}{2}$$

However, we can also assign the probability differently as follows.

$$\mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(\Omega_1) = 1, \quad \mathbb{P}(\{H\}) = \frac{1}{3}, \quad \mathbb{P}(\{T\}) = \frac{2}{3}$$

Both assignments of the probabilities are legitimate. The first one models a fair coin. The second one models an unfair coin with more chance for the coin to turn tail. In the examples to follow, when we consider a fair coin, we take the first assignment.

Example 1.7 *Consider the example of two tosses of a coin. In Example 1.4, we have seen that there are 16 possible events. Observe that if both coins are fair,*

$$\mathbb{P}(\{HH\}) = \mathbb{P}(\{HT\}) = \mathbb{P}(\{TH\}) = \mathbb{P}(\{TT\}) = \frac{1}{4}$$

The probability of other bigger *events can be deduced from Rule (iii). For example, the probability of the event $\{HH, HT, TH\}$ can be deduced as follows.*

$$\begin{aligned}
\mathbb{P}(\{HH, HT, TH\}) &= \mathbb{P}(\{HH\} \cup \{HT\} \cup \{TH\}) \\
&= \mathbb{P}(\{HH\}) + \mathbb{P}(\{HT\}) + \mathbb{P}(\{TH\}) \\
&= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} \\
&= \frac{3}{4}
\end{aligned}$$

The second equality follows by applying Rule (iii) to those three simple events that are mutually exclusive.

One useful identity that we will repeatedly refer to is the probability of not occurring. Let A be an arbitrary event. Then the event that A does not occur is denoted by its complement, A^c. Observe that

$$\Omega = A \cup A^c,$$

and A and A^c cannot occur simultaneously, i.e.

$$A \cap A^c = \emptyset.$$

Therefore, by the first and the third rule of probability, respectively,

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c)$$

Hence, the probability that A would not occur is $1 - \mathbb{P}(A)$.

Theorem 1.1 *For an event A and its complement A^c,*

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A). \tag{1.1}$$

Another useful identity is the probability of a combined event.

Theorem 1.2 *Let A and B be two arbitrary events. Then*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B). \tag{1.2}$$

To see why (1.2) is true, we define event $A - B$ as the event that contains outcomes in A that is not in B, i.e.,

$$A - B = A \cap B^c.$$

We then observe that $A \cup B$ can be divided into three mutually exclusive events.

$$A \cup B = (A - B) \cup (A \cap B) \cup (B - A).$$

Therefore,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A - B) + \mathbb{P}(A \cap B) + \mathbb{P}(B - A).$$

But,

$$\begin{aligned}
&\mathbb{P}(A - B) + \mathbb{P}(A \cap B) + \mathbb{P}(B - A) \\
&= [\mathbb{P}(A - B) + \mathbb{P}(A \cap B)] + [\mathbb{P}(B - A) + \mathbb{P}(A \cap B)] - \mathbb{P}(A \cap B) \\
&= \mathbb{P}((A - B) \cup (A \cap B)) + \mathbb{P}((B - A) \cup (A \cap B)) - \mathbb{P}(A \cap B) \\
&\quad \text{since } (A - B),\ (A \cap B) \text{ and } (B - A) \text{ are mutually exclusive} \\
&= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).
\end{aligned}$$

Equation (1.2) follows.

Before leaving the definition of probability, note that, from the three rules of probability, if the sample space is a countable set, then the probability of all simple events must sum to one.

1.2.4 Probability Space

Definition 1.4 *A probability space is the combination of three components: sample space Ω, event sets $\mathcal{A}$ and probability $\mathbb{P}$. When referring to a probability space, one uses the triplet notation $(\Omega, \mathcal{A}, \mathbb{P})$.*

A probability space $(\Omega, \mathcal{A}, \mathbb{P})$ completely specifies a random experiment and set up a framework for further analysis. In most cases, the probability space is implicitly specified. For example, referring to Example 1.2, specifying that the random experiment involves two tosses of a fair coin implicitly specifies the probability space to be considered.

1.3 A Simple Case

A simple example of nontrivial probability space is a situation when each outcome is equally likely to occur. Consider the following example.

Example 1.8 *One hundred people are invited to a party. Upon arrival, each is told to drop his or her name card into a closed box. At the end of the party, a card is randomly selected from the box. The person whose name is on the selected card wins a gift. If you are invited, what is the probability that you win the gift? If you and your partner are joining this party, what is the probability that you or your partner win the gift? Finally, if your family (your partner, you and your son and daughter), what is the probability?*

To apply the probability space concept, assume for simplicity that we tagged each person with a number, instead of name, say 1 to 100. Consider the random selection process as the random experiment. Therefore, each outcome of this experiment is a guest that would be selected, and, with numbers replacing names, our sample space is then

$$\Omega = \{1, 2, \ldots, 100\}.$$

There are 100 outcomes altogether.

An event is a subset of Ω. For example, if you, alone, and your number is 21, the event E_1 of you winning the gift would be

$$E_1 = \{21\}$$

If you and your partner join the party with number 21 and 22, the event E_2 of winning would be

$$E_2 = \{21, 22\}$$

If the whole family join with running numbers 21 to 24, the event E_3 of the family winning would be

$$E_3 = \{21, 22, 23, 24\}$$

Since each person is equally likely to be picked, A_i, representing the event that person i is picked, would have the same probability of occurring.

$$A_i = \{i\}$$

and

$$\mathbb{P}(A_i) = \mathbb{P}(\{i\}) = \mathbb{P}(\{j\}) = \mathbb{P}(A_j), \quad \textit{for all } i, j$$

By the rules of probability

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(\cup_{i=1}^{100} A_i) = \sum_{i=1}^{100} \mathbb{P}(A_i) = 100\mathbb{P}(A_1)$$

The third equation results because A_i's cannot occur at the same time and applying the third rule of probability. The last equation comes because each event A_i has the same probability as A_1. As a result,

$$\mathbb{P}(A_1) = \frac{1}{100} = \mathbb{P}(A_i), \quad \textit{for all } i$$

Answering the questions, if you are at the party alone,

$$\mathbb{P}(\textit{you win the gift}) = \mathbb{P}(E_1) = \mathbb{P}(A_{21}) = \frac{1}{100}$$

If you and you partner are at the party,

$$\begin{aligned}\mathbb{P}(\textit{you or your partner win the gift}) &= \mathbb{P}(E_2) \\ &= \mathbb{P}(A_{21} \cup A_{22}) \\ &= \mathbb{P}(A_{21}) + \mathbb{P}(A_{22}) \\ &\quad \textit{by the third rule of probability} \\ &= \frac{1}{100} + \frac{1}{100} \\ &= \frac{2}{100}.\end{aligned}$$

Finally, if your whole family is in the party,

$$\mathbb{P}(\textit{your family wins the gift}) = \frac{4}{100}.$$

We can summarize the above example into a formula. For the case when each outcome is equally likely to occur, the probability that an event E would occur,

$$\mathbb{P}(E) = \frac{\text{Number of outcomes in } E}{\text{Number of outcomes in the sample space}} \tag{1.3}$$

We see that, in this simple case, it is crucial to be able to count the number of the outcomes. We assume that the readers have been exposed to basic counting techniques, such as computations of permutation and combination. Once we can count the number of the outcomes of a given event, the probability of that event is readily computed.

Simple as it may seem, this kind of probability space is a building block for many other important probability spaces in more complex problems.

Example 1.9 *An urn contains 20 balls, 12 of which are white. The rest are black. Randomly select 6 balls from the urn. What is the probability that 4 are white?*

Recall that the number of k-combinations from n outcomes can be expressed in a *formula of factorials as*

$$\text{Number of } k\text{-combinations from } n \text{ outcomes} = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Therefore, this random experiment of random ball selection has totally

$$\binom{20}{6} = \frac{20!}{14!6!} \quad \text{outcomes.}$$

Since there is no reason why any of them is more likely to occur, we can assume that each of them is equally likely to occur. Therefore, the event A that a combination of 6 balls has been picked has probability,

$$\mathbb{P}(A) = \frac{1}{\binom{20}{6}} = \frac{14!6!}{20!}$$

Let

$$E = \text{Event that 4 out of 6 are white.}$$

Observe that, to form an outcome in E, 4 out of 12 white balls and 2 out of 8 black balls must be chosen. Therefore, the total number of possible combinations to form outcomes in E is then

$$\binom{12}{4}\binom{8}{2} = \left(\frac{12!}{4!8!}\right)\left(\frac{8!}{2!6!}\right) \quad \text{outcomes.}$$

By (1.3),

$$\mathbb{P}(E) = \frac{\binom{12}{4}\binom{8}{2}}{\binom{20}{6}}$$

In general, for a group of N objects of two types, let there be K objects of the first type and $N-K$ objects of the second type. (Note that N and K are integer where $K < N$.) Select n objects out of the group randomly, where $n < N$ and $n < K$. Let

$$E = \text{Event that there are } k \text{ objects of the first type}$$

By (1.3),

$$\mathbb{P}(E) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}} \tag{1.4}$$

This type of probability is encountered frequently in practice and is called the *hypergeometric probability distribution*. More details on the hypergeomeric probability distribution will be discussed in Chapter 4.

In this section, we introduced an analysis of the case where each outcome is equally likely to occur. However, there are cases where an outcome is more or less likely to occur than one another. We have seen one of such examples in Example 1.6 when we toss an unfair coin. Those cases with *non-uniform* probabilities are sometimes more complicated to analyze. Nevertheless, the probabilities have to obey the three rules of probability introduced in Subsection 1.2.3, and Rule (iii) will be an important tool to help calculate the probability of a complex event.

Exercises

1.1 Show that the sample space of tossing a fair coin infinitely many times is equivalent to the set of real numbers in the interval $[0,1)$.

1.2 Prove De Morgan's law $(A \cup B)^c = A^c \cap B^c$ by using Venn diagram.

1.3 Let $\mathcal{A}$ be a σ-algebra. Show that property (ii) and (iii) of a σ-algebra imply the following property (iv).

(iv) If $\{A_1, A_2, \ldots\}$ is a sequence of events in $\mathcal{A}$, then the intersection $\cap_{i=1}^{\infty} A_i$ is in $\mathcal{A}$.

1.4 From Example 1.5, construct a σ-algebra from Ω_2 that contains eight events.

1.5 Consider a sample space Ω with three outcomes $\Omega = \{a, b, c\}$. Construct three σ-algebras from Ω denoted by $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ such that $\mathcal{A}_1 \neq \mathcal{A}_2 \neq \mathcal{A}_3$ and $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{A}_3$.

1.6 Let the sample space Ω be the real numbers between 0 and 1, including 0 but excluding 1, i.e. $\Omega = [0,1)$. Let $\mathcal{A}$ be the set of all non-empty open intervals. Show that $\mathcal{A}$ is not a σ-algebra.

1.7 Let A_1, A_2 and A_3 be events from a sample space Ω. Show that

$$\begin{aligned}\mathbb{P}(A_1 \cup A_2 \cup A_3) \;=\; & \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1 \cap A_2) - \\ & \mathbb{P}(A_1 \cap A_3) - \mathbb{P}(A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3).\end{aligned}$$

1.8 (Inclusion-exclusion Formula) Let $A_1, A_2, \ldots, A_n$ be events from a sample space Ω. Show that

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n}\left((-1)^{k+1} \sum_{1 \le i_1 \le \cdots \le i_k \le n} \mathbb{P}\left(A_{i_1} \cap \cdots \cap A_{i_k}\right)\right).$$

1.9 At a party, n men took off their hats and left them at the front desk when they entered the party. When they left the party, each man randomly picked up one hat. A match occurs when a man gets back his hat.

(a) Find the probability of no matches. (Use the inclusion-exclusion formula in Problem 1.8.)

(b) Show that when n goes to infinity, the probability of no matches approaches $1/e$.

1.10 An urn contains 20 balls, 6 are white, 6 are red and 8 are black. Randomly select 6 balls from the urn. What is the probability of getting 2 of each color?

1.11 A suitcase has a 3-digit combination lock. If we forgot the correct 3-digit code, what is the probability that a random 3-digit code would unlock it?

1.12 Each credit card contains a 16-digit credit card number plus 3-digit security code. To authorize a payment, one requires the complete 19-digit code. Suppose that there are one hundred million cards being active in the market. What is the probability that a random 19-digit code can authorize a payment?

1.13 In the English language, there are 71 three-letter words that have meaning. What is the probability that a random chosen three-letter word will have a meaning.

1.14 In the English language, there are 55 three-distinct-letter words that have meaning. What is the probability that a randomly chosen three-distinct-letter word will have a meaning.

1.15 Drawing two cards from a deck of regular 52 cards, what is the probability of obtaining a pair?

1.16 Drawing two cards from a deck of regular 52 cards, what is the probability that the second card is a queen?

1.17 Tossing two dice, what is the probability that the sum is seven?

1.18 In a soccer match, there are 22 players from the two teams plus one referee on the field. What is the probability that no one on the field has the same birthday (same day and same month)?

1.19 In a random group of n people, the probability that at least one pair has the same birthday is 0.99. What is the smallest number n that makes this happens?

1.20 In a city, families of one person, two persons, three persons and more-than-three persons constitute the same proportion.

(a) What is the probability that a randomly selected family has two or three persons?

(b) A family with children must have one or two parents. Thirty percent of all families have one child. One half of the three-person families have one child. Selecting a family randomly, what is the probability that this family has two persons or one child?