华章数学原版精品系列

PEARSON

概率论基础教程

（英文版 · 第8版）

A First Course in

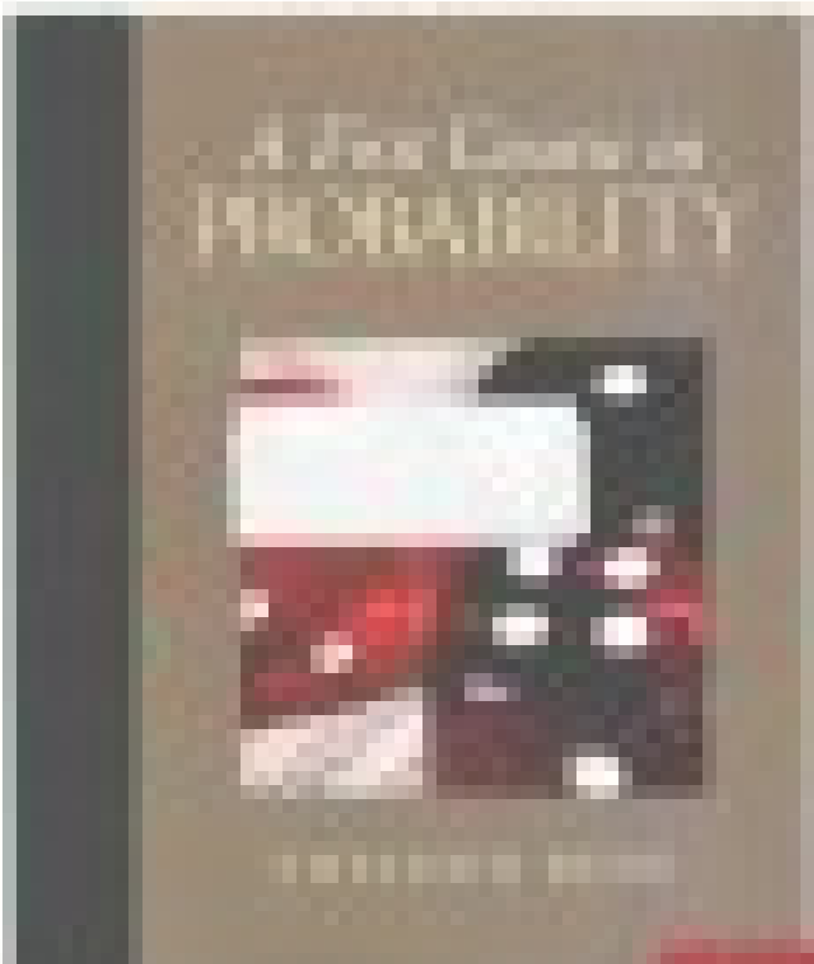
PROBABILITY

EIGHTH EDITION

SHELDON ROSS

（美）Sheldon M. Ross 著
南加州大学

概率论基础教程



华章数学原版精品系列

概率论基础教程

（英文版 · 第8版）

A First Course in Probability
(Eighth Edition)

（美）Sheldon M. Ross 著
南加州大学

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

概率论基础教程（英文版·第8版）/（美）罗斯（Ross, S. M.）著．—北京：机械工业出版社，2014.11

（华章数学原版精品系列）

书名原文：A First Course in Probability（Eighth Edition）

ISBN 978-7-111-48277-2

Ⅰ.概…　Ⅱ.罗…　Ⅲ.概率论－教材－英文　Ⅳ.O211

中国版本图书馆CIP数据核字（2014）第232859号

本书版权登记号：图字：01-2014-6190

出版发行：机械工业出版社（北京市西城区百万庄大街22号　邮政编码：100037）

责任编辑：明永玲　　责任校对：董纪丽

印　　刷：藁城市京瑞印刷有限公司　　版　　次：2014年11月第1版第1次印刷

开　　本：186mm×240mm　1/16　　印　　张：29.5

书　　号：ISBN 978-7-111-48277-2　　定　　价：79.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：（010）88378991　88361066　　投稿热线：（010）88379604

购书热线：（010）68326294　88379649　68995259　　读者信箱：hzjsj@hzbook.com

前　　言

“我们看到概率论实际上只是将常识归结为计算，它使我们能够用理性的头脑精确地评价凭某种直观感受到的、往往又不能解释清楚的见解……引人注意的是，概率论这门起源于对机会游戏进行思考的科学，早就应该成为人类知识中最重要的组成部分……生活中那些最重要的问题绝大部分其实只是概率论的问题.”著名的法国数学家和天文学家拉普拉斯侯爵（人称“法国的牛顿”）如是说. 尽管许多人认为，这位对概率论的发展作出过重大贡献的著名侯爵说话夸张了一些，但是概率论已经成为几乎所有的科学工作者、工程师、医务人员、法律工作者以及企业家们手中的基本工具已经成为事实. 实际上，有见识的人们不再问：“是这样？”而是问：“有多大的概率是这样？”

本书是概率论的入门教材，读者对象是数学、统计、工程和其他学科专业（包括计算机科学、生物学、社会科学和管理科学）的学生. 只需要读者具备初等微积分知识作为基础. 本书试图介绍概率论的数学理论，同时通过大量例子来展示这门学科的广泛的应用.

第 1 章阐述了组合分析的基本原理，它是计算概率的最有用的工具. 第 2 章介绍了概率论的公理体系，并且阐明如何应用这些公理计算各种概率. 第 3 章讨论概率论中极为重要的两个概念，即事件的条件概率和事件间的独立性. 通过一系列例子说明，当部分信息可利用时，条件概率就会发挥它的作用；即使在没有部分信息时，条件概率也可以使概率的计算变得容易. 利用“条件”计算概率这一极为重要的技巧还将出现在第 7 章，在那里我们用它来计算期望. 随机变量的概念在第 4 ～ 6 章引入. 第 4 章讨论离散型随机变量，第 5 章讨论连续型随机变量，第 6 章讨论随机变量的联合分布. 在第 4 章和第 5 章中讨论了两个重要概念，即随机变量的期望值和方差，并且对许多常见的随机变量，求出了相应的期望值和方差.

第 7 章进一步讨论了期望值的一些重要性质. 书中引入了许多例子，解释如何利用随机变量和的期望值等于随机变量期望值的和这一重要规律来计算随机变量的期望值. 本章中还有几节介绍条件期望（包括它在预测方面的应用）和矩母函数. 本章最后一节介绍了多元正态分布，同时给出了来自正态总体的样本均值和样本方差的联合分布的简单证明.

在第 8 章我们介绍了概率论的主要的理论结果. 特别地，我们证明了强大数定律和中心极限定理. 在强大数定律的证明中，我们假定了随机变量具有有限的四阶矩，因为在这种假定之下，证明非常简单. 在中心极限定理的证明中，我们假定了莱维连续性定理成立. 在本章中，我们还介绍了若干概率不等式，如马尔可夫不等式、切比雪夫不等式和切尔诺夫界. 在本章最后一节，我们给出用有相同期望值的泊松随机变量的相应概率去近似独立伯努利随机变量和的相关概率的误差界.

第 9 章阐述了一些额外的论题，如马尔可夫链、泊松过程以及信息编码理论初步，第 10 章介绍了统计模拟. 与以前的版本一样，在每章末给出了三组练习题，它们被指定为习题、理论习题和自检习题. 自检习题的完整解答在附录 B 给出，这部分练习题可以帮助学生检测他们对知识的掌握程度并为考试做准备.

新版变化

第 8 版继续对教材内容进行微调和优化，加入了很多新的习题和例子. 内容的选取不仅要适合学生的兴趣，还要有助于学生建立概率直觉. 为此，第 1 章例 5d 讨论了淘汰赛，第 7 章的例 4k 和例 5i 是多个赌徒破产问题的例子. 新版最主要的变化是随机变量和的期望等于随机变量期望的和这一重要规律，在第 4 章首次出现（而不是旧版的第 7 章）. 第 4 章还针对概率实验的样本空间有限时这一特殊情况，给出了这一规律的新的且初等的证明.

6.3 节介绍独立随机变量的和，这一节也有一些变化. 6.3.1 节是新增的一节，推导独立且具有相同

均匀分布的随机变量和的分布，并用所得到的结果证明了，具有（0，1）上均匀分布的独立随机变量，和大于 1 的那些随机变量的平均个数是 e．6.3.5 节也是新增的一节，推导具有独立几何分布但均值不同的随机变量和的分布．

致谢

Hossein Hamedani 仔细审阅了本教材，对此我深表感谢．同时我还要感谢下列人员对于这一版的改进提出宝贵的建议：Amir Ardestani（德黑兰理工大学），Joe Blitzstein（哈佛大学），Peter Nuesch（洛桑大学），Joseph Mitchell（纽约州立大学石溪分校），Alan Chambless（精算师），Robert Kriner, Israel David（本古里安大学），T. Lim（乔治・梅森大学），Wei Chen（罗格斯大学），D. Monrad（伊利诺伊大学），W. Rosenberger（乔治・梅森大学），E. Ionides（密歇根大学），J. Corvino（拉法叶学院），T. Seppalainen（威斯康星大学）．

最后，我要感谢下列对本书各个版本给出很多有价值的意见的人们．其中，对第 8 版的改进给出意见的审稿人，在其名字前面加了星号．

K. B. Athreya（爱荷华州立大学）
Richard Bass（康涅狄格大学）
Robert Bauer（伊利诺伊大学厄巴纳－尚佩恩分校）
Phillip Beckwith（密歇根科技大学）
Arthur Benjamin（哈维姆德学院）
Geoffrey Berresford（长岛大学）
Baidurya Bhattacharya（特拉华大学）
Howard Bird（圣克劳德州立大学）
Shahar Boneh（丹佛城市州立学院）
Jean Cadet（纽约州立大学石溪分校）
Steven Chiappari（圣克拉拉大学）
Nicolas Christou（加州大学洛杉矶分校）
James Clay（亚利桑那大学图森分校）
Francis Conlan（圣克拉拉大学）
* Justin Corvino（拉法叶学院）
Jay DeVore（圣路易斯－奥比斯波的加州技术大学）
Scott Emerson（华盛顿大学）
Thomas R. Fischer（德州农机大学）
Anant Godbole（密歇根科技大学）
Zakkhula Govindarajulu（肯塔基大学）
Richard Groeneveld（爱荷华州立大学）
Mike Hardy（麻省理工学院）
Bernard Harris（威斯康星大学）
Larry Harris（肯塔基大学）
David Heath（康奈尔大学）
Stephen Herschkorn（罗格斯大学）
Julia L.Higle（亚利桑那大学）
Mark Huber（杜克大学）
* Edward Ionides（密歇根大学）
Anastasia Ivanova（北卡罗来纳大学）
Hamid Jafarkhani（加州大学厄文分校）
Chuanshu Ji（北卡罗来纳大学 Chapel Hill 分校）
Robert Keener（密歇根大学）
Fred Leysieffer（佛罗里达州立大学）
Thomas Liggett（加州大学洛杉矶分校）
Helmut Mayer（佐治亚大学）
Bill McCormick（佐治亚大学）
Ian McKeague（佛罗里达州立大学）
R. Miller（斯坦福大学）
* Ditlev Monrad（伊利诺伊大学）
Robb J. Muirhead（密歇根大学）
Joe Naus（罗格斯大学）
Nhu Nguyen（新墨西哥州立大学）
Ellen O'Brien（乔治・梅森大学）
N.U. Prabhu（康奈尔大学）
Kathryn Prewitt（亚利桑那州立大学）
Jim Propp（威斯康星大学）
* William F. Rosenberger（乔治・梅森大学）
Myra Samuels（普度大学）
I. R. Savage（耶鲁大学）
Art Schwartz（密歇根大学安阿伯分校）
Therese Shelton（西南大学）
Malcolm Sherman（纽约州立大学奥尔巴尼分校）
Murad Taqqu（波士顿大学）
Eli Upfal（布朗大学）
Ed Wheeler（田纳西大学）
Allen Webster（布拉德利大学）

smross@usc.edu

Contents

Chapter 1

COMBINATORIAL ANALYSIS

Contents

1.1 Introduction

Here is a typical problem of interest involving probability: A communication system is to consist of n seemingly identical antennas that are to be lined up in a linear order. The resulting system will then be able to receive all incoming signals—and will be called *functional*—as long as no two consecutive antennas are defective. If it turns out that exactly m of the n antennas are defective, what is the probability that the resulting system will be functional? For instance, in the special case where $n = 4$ and $m = 2$, there are 6 possible system configurations, namely,

$$\begin{matrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{matrix}$$

where 1 means that the antenna is working and 0 that it is defective. Because the resulting system will be functional in the first 3 arrangements and not functional in the remaining 3, it seems reasonable to take $\frac{3}{6} = \frac{1}{2}$ as the desired probability. In the case of general n and m, we could compute the probability that the system is functional in a similar fashion. That is, we could count the number of configurations that result in the system's being functional and then divide by the total number of all possible configurations.

From the preceding discussion, we see that it would be useful to have an effective method for counting the number of ways that things can occur. In fact, many problems in probability theory can be solved simply by counting the number of different ways that a certain event can occur. The mathematical theory of counting is formally known as *combinatorial analysis*.

1.2 The Basic Principle of Counting

The basic principle of counting will be fundamental to all our work. Loosely put, it states that if one experiment can result in any of m possible outcomes and if another experiment can result in any of n possible outcomes, then there are mn possible outcomes of the two experiments.

> **The basic principle of counting**
>
> Suppose that two experiments are to be performed. Then if experiment 1 can result in any one of m possible outcomes and if, for each outcome of experiment 1, there are n possible outcomes of experiment 2, then together there are mn possible outcomes of the two experiments.

Proof of the Basic Principle: The basic principle may be proven by enumerating all the possible outcomes of the two experiments; that is,

$$\begin{array}{c}
(1,1),\ (1,2),\ \ldots,\ (1,n) \\
(2,1),\ (2,2),\ \ldots,\ (2,n) \\
\vdots \\
(m,1),\ (m,2),\ \ldots,\ (m,n)
\end{array}$$

where we say that the outcome is (i, j) if experiment 1 results in its ith possible outcome and experiment 2 then results in its jth possible outcome. Hence, the set of possible outcomes consists of m rows, each containing n elements. This proves the result.

Example 2a

A small community consists of 10 women, each of whom has 3 children. If one woman and one of her children are to be chosen as mother and child of the year, how many different choices are possible?

Solution By regarding the choice of the woman as the outcome of the first experiment and the subsequent choice of one of her children as the outcome of the second experiment, we see from the basic principle that there are $10 \times 3 = 30$ possible choices. ■

When there are more than two experiments to be performed, the basic principle can be generalized.

> **The generalized basic principle of counting**
>
> If r experiments that are to be performed are such that the first one may result in any of n_1 possible outcomes; and if, for each of these n_1 possible outcomes, there are n_2 possible outcomes of the second experiment; and if, for each of the possible outcomes of the first two experiments, there are n_3 possible outcomes of the third experiment; and if ..., then there is a total of $n_1 \cdot n_2 \cdots n_r$ possible outcomes of the r experiments.

Example 2b

A college planning committee consists of 3 freshmen, 4 sophomores, 5 juniors, and 2 seniors. A subcommittee of 4, consisting of 1 person from each class, is to be chosen. How many different subcommittees are possible?

Solution We may regard the choice of a subcommittee as the combined outcome of the four separate experiments of choosing a single representative from each of the classes. It then follows from the generalized version of the basic principle that there are $3 \times 4 \times 5 \times 2 = 120$ possible subcommittees. ■

Example 2c How many different 7-place license plates are possible if the first 3 places are to be occupied by letters and the final 4 by numbers?

Solution By the generalized version of the basic principle, the answer is $26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 175{,}760{,}000$. ■

Example 2d How many functions defined on n points are possible if each functional value is either 0 or 1?

Solution Let the points be $1, 2, \ldots, n$. Since $f(i)$ must be either 0 or 1 for each $i = 1, 2, \ldots, n$, it follows that there are 2^n possible functions. ■

Example 2e In Example 2c, how many license plates would be possible if repetition among letters or numbers were prohibited?

Solution In this case, there would be $26 \cdot 25 \cdot 24 \cdot 10 \cdot 9 \cdot 8 \cdot 7 = 78{,}624{,}000$ possible license plates. ■

1.3 Permutations

How many different ordered arrangements of the letters a, b, and c are possible? By direct enumeration we see that there are 6, namely, *abc, acb, bac, bca, cab*, and *cba*. Each arrangement is known as a *permutation*. Thus, there are 6 possible permutations of a set of 3 objects. This result could also have been obtained from the basic principle, since the first object in the permutation can be any of the 3, the second object in the permutation can then be chosen from any of the remaining 2, and the third object in the permutation is then the remaining 1. Thus, there are $3 \cdot 2 \cdot 1 = 6$ possible permutations.

Suppose now that we have n objects. Reasoning similar to that we have just used for the 3 letters then shows that there are

$$n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1 = n!$$

different permutations of the n objects.

Example 3a How many different batting orders are possible for a baseball team consisting of 9 players?

Solution There are $9! = 362{,}880$ possible batting orders. ■

Example 3b A class in probability theory consists of 6 men and 4 women. An examination is given, and the students are ranked according to their performance. Assume that no two students obtain the same score.

(a) How many different rankings are possible?

(b) If the men are ranked just among themselves and the women just among themselves, how many different rankings are possible?

Solution (a) Because each ranking corresponds to a particular ordered arrangement of the 10 people, the answer to this part is $10! = 3{,}628{,}800$.

(b) Since there are 6! possible rankings of the men among themselves and 4! possible rankings of the women among themselves, it follows from the basic principle that there are $(6!)(4!) = (720)(24) = 17{,}280$ possible rankings in this case. ■

Example 3c

Ms. Jones has 10 books that she is going to put on her bookshelf. Of these, 4 are mathematics books, 3 are chemistry books, 2 are history books, and 1 is a language book. Ms. Jones wants to arrange her books so that all the books dealing with the same subject are together on the shelf. How many different arrangements are possible?

Solution There are 4! 3! 2! 1! arrangements such that the mathematics books are first in line, then the chemistry books, then the history books, and then the language book. Similarly, for each possible ordering of the subjects, there are 4! 3! 2! 1! possible arrangements. Hence, as there are 4! possible orderings of the subjects, the desired answer is $4!\,4!\,3!\,2!\,1! = 6912$. ■

We shall now determine the number of permutations of a set of n objects when certain of the objects are indistinguishable from each other. To set this situation straight in our minds, consider the following example.

Example 3d

How many different letter arrangements can be formed from the letters *PEPPER*?

Solution We first note that there are 6! permutations of the letters $P_1E_1P_2P_3E_2R$ when the 3P's and the 2E's are distinguished from each other. However, consider any one of these permutations—for instance, $P_1P_2E_1P_3E_2R$. If we now permute the P's among themselves and the E's among themselves, then the resultant arrangement would still be of the form *PPEPER*. That is, all 3! 2! permutations

$$
\begin{array}{ll}
P_1P_2E_1P_3E_2R & P_1P_2E_2P_3E_1R \\
P_1P_3E_1P_2E_2R & P_1P_3E_2P_2E_1R \\
P_2P_1E_1P_3E_2R & P_2P_1E_2P_3E_1R \\
P_2P_3E_1P_1E_2R & P_2P_3E_2P_1E_1R \\
P_3P_1E_1P_2E_2R & P_3P_1E_2P_2E_1R \\
P_3P_2E_1P_1E_2R & P_3P_2E_2P_1E_1R
\end{array}
$$

are of the form *PPEPER*. Hence, there are $6!/(3!\,2!) = 60$ possible letter arrangements of the letters *PEPPER*. ■

In general, the same reasoning as that used in Example 3d shows that there are

$$\frac{n!}{n_1!\,n_2!\,\cdots\,n_r!}$$

different permutations of n objects, of which n_1 are alike, n_2 are alike, ..., n_r are alike.

Example 3e

A chess tournament has 10 competitors, of which 4 are Russian, 3 are from the United States, 2 are from Great Britain, and 1 is from Brazil. If the tournament result lists just the nationalities of the players in the order in which they placed, how many outcomes are possible?

Solution There are

$$\frac{10!}{4!\,3!\,2!\,1!} = 12{,}600$$

possible outcomes. ■

Example 3f How many different signals, each consisting of 9 flags hung in a line, can be made from a set of 4 white flags, 3 red flags, and 2 blue flags if all flags of the same color are identical?

Solution There are

$$\frac{9!}{4!\,3!\,2!} = 1260$$

different signals. ■

1.4 Combinations

We are often interested in determining the number of different groups of r objects that could be formed from a total of n objects. For instance, how many different groups of 3 could be selected from the 5 items A, B, C, D, and E? To answer this question, reason as follows: Since there are 5 ways to select the initial item, 4 ways to then select the next item, and 3 ways to select the final item, there are thus $5 \cdot 4 \cdot 3$ ways of selecting the group of 3 when the order in which the items are selected is relevant. However, since every group of 3—say, the group consisting of items A, B, and C—will be counted 6 times (that is, all of the permutations ABC, ACB, BAC, BCA, CAB, and CBA will be counted when the order of selection is relevant), it follows that the total number of groups that can be formed is

$$\frac{5 \cdot 4 \cdot 3}{3 \cdot 2 \cdot 1} = 10$$

In general, as $n(n-1)\cdots(n-r+1)$ represents the number of different ways that a group of r items could be selected from n items when the order of selection is relevant, and as each group of r items will be counted $r!$ times in this count, it follows that the number of different groups of r items that could be formed from a set of n items is

$$\frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{(n-r)!\,r!}$$

Notation and terminology

We define $\binom{n}{r}$, for $r \le n$, by

$$\binom{n}{r} = \frac{n!}{(n-r)!\,r!}$$

and say that $\binom{n}{r}$ represents the number of possible combinations of n objects taken r at a time.†

† By convention, 0! is defined to be 1. Thus, $\binom{n}{0} = \binom{n}{n} = 1$. We also take $\binom{n}{i}$ to be equal to 0 when either $i < 0$ or $i > n$.

Thus, $\binom{n}{r}$ represents the number of different groups of size r that could be selected from a set of n objects when the order of selection is not considered relevant.

Example 4a

A committee of 3 is to be formed from a group of 20 people. How many different committees are possible?

Solution There are $\binom{20}{3} = \frac{20 \cdot 19 \cdot 18}{3 \cdot 2 \cdot 1} = 1140$ possible committees. ■

Example 4b

From a group of 5 women and 7 men, how many different committees consisting of 2 women and 3 men can be formed? What if 2 of the men are feuding and refuse to serve on the committee together?

Solution As there are $\binom{5}{2}$ possible groups of 2 women, and $\binom{7}{3}$ possible groups of 3 men, it follows from the basic principle that there are $\binom{5}{2}\binom{7}{3} = \left(\frac{5 \cdot 4}{2 \cdot 1}\right)\frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 350$ possible committees consisting of 2 women and 3 men.

Now suppose that 2 of the men refuse to serve together. Because a total of $\binom{2}{2}\binom{5}{1} = 5$ out of the $\binom{7}{3} = 35$ possible groups of 3 men contain both of the feuding men, it follows that there are $35 - 5 = 30$ groups that do not contain both of the feuding men. Because there are still $\binom{5}{2} = 10$ ways to choose the 2 women, there are $30 \cdot 10 = 300$ possible committees in this case. ■

Example 4c

Consider a set of n antennas of which m are defective and $n - m$ are functional and assume that all of the defectives and all of the functionals are considered indistinguishable. How many linear orderings are there in which no two defectives are consecutive?

Solution Imagine that the $n - m$ functional antennas are lined up among themselves. Now, if no two defectives are to be consecutive, then the spaces between the functional antennas must each contain at most one defective antenna. That is, in the $n - m + 1$ possible positions—represented in Figure 1.1 by carets—between the $n - m$ functional antennas, we must select m of these in which to put the defective antennas. Hence, there are $\binom{n - m + 1}{m}$ possible orderings in which there is at least one functional antenna between any two defective ones. ■

$\wedge 1 \wedge 1 \wedge 1 \ldots \wedge 1 \wedge 1 \wedge$

1 = functional

$\wedge$ = place for at most one defective

Figure 1.1 No consecutive defectives.

A useful combinatorial identity is

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r} \qquad 1 \le r \le n \tag{4.1}$$

Equation (4.1) may be proved analytically or by the following combinatorial argument: Consider a group of n objects, and fix attention on some particular one of these objects—call it object 1. Now, there are $\binom{n-1}{r-1}$ groups of size r that contain object 1 (since each such group is formed by selecting $r - 1$ from the remaining $n - 1$ objects). Also, there are $\binom{n-1}{r}$ groups of size r that do not contain object 1. As there is a total of $\binom{n}{r}$ groups of size r, Equation (4.1) follows.

The values $\binom{n}{r}$ are often referred to as *binomial coefficients* because of their prominence in the binomial theorem.

The binomial theorem

$$(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \tag{4.2}$$

We shall present two proofs of the binomial theorem. The first is a proof by mathematical induction, and the second is a proof based on combinatorial considerations.

Proof of the Binomial Theorem by Induction: When $n = 1$, Equation (4.2) reduces to

$$x + y = \binom{1}{0} x^0 y^1 + \binom{1}{1} x^1 y^0 = y + x$$

Assume Equation (4.2) for $n - 1$. Now,

$$\begin{aligned}(x + y)^n &= (x + y)(x + y)^{n-1} \\ &= (x + y) \sum_{k=0}^{n-1} \binom{n-1}{k} x^k y^{n-1-k} \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} x^{k+1} y^{n-1-k} + \sum_{k=0}^{n-1} \binom{n-1}{k} x^k y^{n-k}\end{aligned}$$

Letting $i = k + 1$ in the first sum and $i = k$ in the second sum, we find that

$$\begin{aligned}(x + y)^n &= \sum_{i=1}^{n} \binom{n-1}{i-1} x^i y^{n-i} + \sum_{i=0}^{n-1} \binom{n-1}{i} x^i y^{n-i} \\ &= x^n + \sum_{i=1}^{n-1} \left[\binom{n-1}{i-1} + \binom{n-1}{i} \right] x^i y^{n-i} + y^n \\ &= x^n + \sum_{i=1}^{n-1} \binom{n}{i} x^i y^{n-i} + y^n \\ &= \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}\end{aligned}$$

where the next-to-last equality follows by Equation (4.1). By induction, the theorem is now proved.

Combinatorial Proof of the Binomial Theorem: Consider the product

$$(x_1 + y_1)(x_2 + y_2) \cdots (x_n + y_n)$$

Its expansion consists of the sum of 2^n terms, each term being the product of n factors. Furthermore, each of the 2^n terms in the sum will contain as a factor either x_i or y_i for each $i = 1, 2, \ldots, n$. For example,

$$(x_1 + y_1)(x_2 + y_2) = x_1x_2 + x_1y_2 + y_1x_2 + y_1y_2$$

Now, how many of the 2^n terms in the sum will have k of the x_i's and $(n - k)$ of the y_i's as factors? As each term consisting of k of the x_i's and $(n - k)$ of the y_i's corresponds to a choice of a group of k from the n values $x_1, x_2, \ldots, x_n$, there are $\binom{n}{k}$ such terms. Thus, letting $x_i = x, y_i = y, i = 1, \ldots, n$, we see that

$$(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Example 4d

Expand $(x + y)^3$.

Solution

$$\begin{aligned}(x + y)^3 &= \binom{3}{0} x^0 y^3 + \binom{3}{1} x^1 y^2 + \binom{3}{2} x^2 y^1 + \binom{3}{3} x^3 y^0 \\ &= y^3 + 3xy^2 + 3x^2y + x^3\end{aligned}$$

■

Example 4e

How many subsets are there of a set consisting of n elements?

Solution Since there are $\binom{n}{k}$ subsets of size k, the desired answer is

$$\sum_{k=0}^{n} \binom{n}{k} = (1 + 1)^n = 2^n$$

This result could also have been obtained by assigning either the number 0 or the number 1 to each element in the set. To each assignment of numbers, there corresponds, in a one-to-one fashion, a subset, namely, that subset consisting of all elements that were assigned the value 1. As there are 2^n possible assignments, the result follows.

Note that we have included the set consisting of 0 elements (that is, the null set) as a subset of the original set. Hence, the number of subsets that contain at least 1 element is $2^n - 1$. ■

1.5 Multinomial Coefficients

In this section, we consider the following problem: A set of n distinct items is to be divided into r distinct groups of respective sizes $n_1, n_2, \ldots, n_r$, where $\sum_{i=1}^{r} n_i = n$. How many different divisions are possible? To answer this question, we note that there are $\binom{n}{n_1}$ possible choices for the first group; for each choice of the first group, there are $\binom{n - n_1}{n_2}$ possible choices for the second group; for each choice of the first two groups, there are $\binom{n - n_1 - n_2}{n_3}$ possible choices for the third group; and so on. It then follows from the generalized version of the basic counting principle that there are

$$\begin{aligned}
&\binom{n}{n_1}\binom{n - n_1}{n_2}\cdots\binom{n - n_1 - n_2 - \cdots - n_{r-1}}{n_r} \\
&= \frac{n!}{(n - n_1)!\,n_1!}\frac{(n - n_1)!}{(n - n_1 - n_2)!\,n_2!}\cdots\frac{(n - n_1 - n_2 - \cdots - n_{r-1})!}{0!\,n_r!} \\
&= \frac{n!}{n_1!\,n_2!\cdots n_r!}
\end{aligned}$$

possible divisions.

Another way to see this result is to consider the n values $1, 1, \ldots, 1, 2, \ldots, 2, \ldots, r, \ldots, r$, where i appears n_i times, for $i = 1, \ldots, r$. Every permutation of these values corresponds to a division of the n items into the r groups in the following manner: Let the permutation $i_1, i_2, \ldots, i_n$ correspond to assigning item 1 to group i_1, item 2 to group i_2, and so on. For instance, if $n = 8$ and if $n_1 = 4$, $n_2 = 3$, and $n_3 = 1$, then the permutation $1, 1, 2, 3, 2, 1, 2, 1$ corresponds to assigning items $1, 2, 6, 8$ to the first group, items $3, 5, 7$ to the second group, and item 4 to the third group. Because every permutation yields a division of the items and every possible division results from some permutation, it follows that the number of divisions of n items into r distinct groups of sizes $n_1, n_2, \ldots, n_r$ is the same as the number of permutations of n items of which n_1 are alike, and n_2 are alike, ..., and n_r are alike, which was shown in Section 1.3 to equal $\frac{n!}{n_1!n_2!\cdots n_r!}$.