工业和信息化普通高等教育“十二五”规划教材

概率论与数理统计及其应用

李昌兴 ◎ 主编
张素梅　林椹尠　赵美霞　邢务强 ◎ 编著
于海波 ◎ 主审

Probability theory and mathematical statistics with applications

人民邮电出版社
POSTS & TELECOM PRESS

工业和信息化普通高等教育“十二五”规划教材

概率论与数理统计及其应用

李昌兴 ◎ 主编
张素梅　林榵勣　赵美霞　邢务强 ◎ 编著
于海波 ◎ 主审

Probability theory and mathematical statistics with applications

人民邮电出版社
北京

图书在版编目（C I P）数据

概率论与数理统计及其应用 / 李昌兴主编 ；张素梅等编著. -- 北京 ：人民邮电出版社，2012.8
ISBN 978-7-115-27889-0

Ⅰ. ①概… Ⅱ. ①李… ②张… Ⅲ. ①概率论②数理统计 Ⅳ. ①O21

中国版本图书馆CIP数据核字(2012)第058481号

内 容 提 要

本书根据教育部最新颁布的概率论与数理统计教学基本要求，结合作者多年的教学实践编写而成，书中内容包括：随机事件与概率、随机变量及其分布、多维随机变量及其分布、随机变量的数字特征、大数定律和中心极限定理、数理统计的基本概念、参数估计、假设检验、回归分析、随机过程的基本知识、马尔可夫链、平稳随机过程．全书重点着眼于介绍概率论、数理统计、随机过程中的基本概念、基本原理和基本方法，强调直观性，加强可读性，突出基本思想，注重实际应用．

本书可供高等院校工科和其他非数学类专业的学生使用，也可供各类专业技术人员参考．

概率论与数理统计及其应用

◆ 主　　编　李昌兴
　 编　　著　张素梅　林椹尠　赵美霞　邢务强
　 主　　审　于海波
　 责任编辑　贾　楠

◆ 人民邮电出版社出版发行　　北京市崇文区夕照寺街 14 号
　 邮编　100061　　电子邮件　315@ptpress.com.cn
　 网址　http://www.ptpress.com.cn
　 中国铁道出版社印刷厂印刷

◆ 开本：787×1092　1/16
　 印张：17.25　　　　2012 年 8 月第 1 版
　 字数：412 千字　　　2012 年 8 月北京第 1 次印刷

ISBN 978-7-115-27889-0

定价：36.00 元

读者服务热线：(010)67170985　印装质量热线：(010)67129223
反盗版热线：(010)67171154

前　言

概率论与数理统计是研究随机现象统计规律性的一门数学课程，也是普通高等院校本科各专业普遍开设的一门公共基础课程．目前，我国高等教育已由“精英教育”转向“大众化教育”，为了适应新的教学形势、学生基础和教学特点，我们编写了《概率论与数理统计及其应用》这部教材．

在本书编写的过程中，我们认真贯彻落实教育部“高等教育面向21世纪教学内容和课程体系改革计划”的要求和精神，严格执行教育部数学与统计学教学指导委员会2009年修订的工科和管理类“本科数学基础课程（概率论与数理统计）教学基本要求”，消化吸收国内外同类教材的精华，融入近年来教学改革的新理念和新成果，坚持边实践、边总结，力求教材建设与时俱进，精益求精．本书着重突出以下几个特色．

（1）整合教学内容，重组知识结构．对现行教材的内容进行重构，力争使相近内容相对集中，知识模块由大到小、先易后难，做到结构严谨、逻辑清晰、叙述清楚．

（2）重视问题驱动，激活思考探索．在引入基本概念、基本定理之前，提出一些富有启发性、触及数学本质的问题．围绕问题的产生、发展与解决，展开内容讨论，以引导学生进行思考与探索，重现数学知识的返现过程．

（3）确保基本要求，降低知识难度．在保证基本概念的理解、基本方法的使用、基本技能训练的前提下，适当降低知识难度．在叙述论证上，条理清晰，重点突出，难点分散，删除令人费解的繁杂证明，淡化严格抽象的理论推导，代之以直观简要的说明或例证，使学生易于抓住理论推导的基本思路与精神实质．简化解题技巧训练，侧重于基本能力的培养与提高．

（4）注重数学思想，突出实际应用．揭示各章节内容之间矛盾的发生、冲突与转化，在对立中求得统一．例题与习题的配备贴近生活、贴近实际、贴近应用．如女士品茶问题、出行路线的选择问题、录取者的最低分数线问题、求职面试决策问题等．对于这些问题的讨论，有利于激发学生的学习兴趣，有利于提高学生的综合素质．为加强培养学生的创新能力，部分章节增设适量的半开放性和开放性的实例．

本书分三个部分．概率论部分（第1章至第5章）作为基础知识，为读者提供必要的理论基础；数理统计部分（第6章至第9章）主要讲述了参数估计和假设检验，简要介绍了回归分析；随机过程部分（第10章至第12章）在阐述清楚基本知识的基础上，主要讨论了马

尔可夫链和平稳随机过程．数理统计和随机过程两部分是相互独立的，可根据专业需要选用．其中第 3 章、第 9 章由张素梅编写，第 4 章由林椹彬编写，第 5 章、第 6 章由赵美霞编写，第 7 章、第 8 章由邢务强编写，其余各章由李昌兴编写，最后由李昌兴统纂定稿，由于海波担任主审．

本书可作为高等院校各工科、理科（非数学专业）及管理类各专业概率论与数理统计课程的教材，也可供各类有关技术人员参考．

在本书的编写过程中，我们参阅大量国内同类教材及相关辅导书，得到了有益的启迪和教益，谨向有关作者表示谢意！

本书虽然经过我们深思熟虑和反复推敲，但难免一疏，欢迎广大读者批评指正，使本书在教学实践中不断完善．

编　者

2012 年 3 月

目录

第 1 章 随机事件与概率

概率论与数理统计是研究随机现象统计规律性的一门数学学科. 为了揭示随机现象的内在规律，建立严密的逻辑体系，本章主要介绍随机事件、概率的定义、古典概型与几何概型、条件概率以及事件的独立性等内容.

1.1 随机试验、样本空间

在自然界和人类的社会生活中常常会出现各种各样的现象，例如，一枚硬币向上抛起后必然落地；每天早晨太阳从东方升起；在相同大气压与温度的条件下，气罐内的分子对罐壁的压力是常数. 这类现象的共同特点是在可以控制条件一定时，观测到的结果也一定，而且可以预测，这类现象称为**确定性现象**. 另一类现象则不然，例如，用同一门炮向同一目标射击，各次弹着点不尽相同，在一次射击之前无法预测弹着点的确切位置. 再如，抛一枚硬币，着地时可能出现正面向上，也可能出现反面向上，而在每次抛掷之前，无法确定正面向上还是反面向上，呈现出不确定性. 但人们经过长期实践并深入研究之后，发现这类现象在大量重复试验或观察下，它的结果呈现出某种规律性. 例如，同一门炮向同一目标射击的弹着点按照一定的规律分布，多次重复抛一枚硬币得到正面向上的次数大致占到抛掷总次数的一半等. 这种在个别观测中其结果呈现出不确定性，而在大量重复观测中其结果又呈现出规律性的现象称为**随机现象**，这种规律称之为**统计规律**.

1.1.1 随机试验

为叙述方便，我们通常把对某一事物的某一特征的一次观察、测量或进行一次科学实验等统称为一种试验. 概率论与数理统计研究的是随机现象的统计规律性，需要进行各种各样的试验.

一般地，如果一个试验满足下列条件：

（1）每次试验的可能结果不止一个，并且在试验之前能明确试验的所有可能结果；

（2）进行一次试验之前不能预知哪一个结果会出现.

称这样的试验为**随机试验**，用 E 来表示. 如果随机试验在相同的条件下可以重复进行，则称为**可重复的随机试验**，否则，称为**不可重复的随机试验**.

下面举一些随机试验的例子.

E_1：抛一枚硬币，观察正面 H、反面 T 出现的情况.

E_2：将一枚硬币连续抛 2 次，观察正面 H、反面 T 出现的情况.

E_3：将一枚硬币连续抛 2 次，观察正面 H 出现的次数.

E_4：抛一颗骰子，观察出现的点数.

E_5：记录某城市 119 防灾指挥中心一昼夜接到用户的呼叫次数.

E_6：在一批电子元件中任意抽取一只，测试它的寿命.

E_7：某人计划去西双版纳旅游，观察在预定的一天能否安全抵达目的地.

E_8：观察某场篮球比赛的输赢.

E_9：观察某年国民生产总值的增长率.

试验 E_1～E_6 都是可重复的随机试验，而 E_7～E_9 均是不可重复的随机试验. 可重复的随机试验已得到广泛深入的研究，有一套成熟的理论和方法. 但随着科学技术的进步和社会经济的发展，特别是现代管理和决策分析的需要，不可重复的随机试验的研究引起了人们的广泛关注. 但本书除了个别章节以外，只讨论可重复的随机试验，因此，在不引起混淆的情况下，以后把可重复的随机试验也简称为随机试验或试验.

1.1.2 样本空间

对于随机试验，尽管在每次试验之前不能预知试验的结果，但试验的所有可能结果组成的集合是已知的，我们把随机试验 E 的所有可能结果组成的集合称为 E 的**样本空间**，记为 S. 样本空间的元素，即 E 的每个结果，称为**样本点**.

下面写出上述试验 E_i $(i=1,2,3,4,5,6)$ 的样本空间 S_i：

$S_1=\{H,T\}$；

$S_2=\{HH,HT,TH,TT\}$；

$S_3=\{0,1,2\}$；

$S_4=\{1,2,3,4,5,6\}$；

$S_5=\{0,1,2,3,\cdots\}$；

$S_6=\{t\,|\,t\geqslant 0\}$.

应该注意的是：试验 E_2 和 E_3 的过程都是将一枚硬币连续抛两次，但是由于试验的目的不同，所以样本空间 S_2 和 S_3 截然不同，这说明试验的目的决定着试验所对应的样本空间.

1.2 随机事件

1.2.1 随机事件

在研究随机试验时，人们不仅关心试验的单个样本点，而且常常对于试验的某些样本点所组成的集合更感兴趣. 例如，若规定某种电子元件的寿命小于 1 500h 的为次品，那么在 E_6 中我们关心电子元件的寿命是否有 $t\geqslant 1500\,(\text{h})$，满足这一条件的样本点组成 S_6 的一个子集 $A=\{t\,|\,t\geqslant 1500\}$，并称 A 是试验 E_6 的一个随机事件. 显然，当且仅当子集 A 中的一个样本点出现时，有 $t\geqslant 1500\,(\text{h})$，即电子元件为合格品. 如某次测试电子元件的寿命是 1 650h，便认为 A 在这次试验中发生了.

一般地，称试验 E 的样本空间 S 的子集为 E 的**随机事件**，简称为**事件**[①]. 在一次试验中，当且仅当这一子集中的一个样本点出现时，称这一**事件发生**. 随机事件常用大写字母 A,B,C 等来表示.

下面举出一些随机事件的例子.

例 1.1 在 E_2 中，事件"两次抛硬币均出现同一面"，即

$$A_1=\{HH,TT\}$$

事件"两次抛掷中至少出现一次正面"，即

$$A_2=\{HH,HT,TH\}$$

在 E_4 中事件"出奇数点"，即

$$A_3=\{1,3,5\}$$

在 E_6 中，事件"寿命不超过 1 700h"，即

$$A_4=\{t\mid 0\leqslant t\leqslant 1\,700\}$$

特别地，由一个样本点组成的单点集，称为**基本事件**. 例如，试验 E_1 中有 2 个基本事件 $\{H\}$ 和 $\{T\}$，试验 E_3 中有 3 个基本事件 $\{0\},\{1\},\{2\}$.

样本空间 S 包含所有的样本点，它是 S 自身的子集，在每次试验中都发生，S 称为**必然事件**. 空集 $\varnothing$ 不包含任何样本点，它也作为样本空间的子集，而它在每次试验中都不发生，称为**不可能事件**.

1.2.2 事件间的关系与运算

在一个样本空间中，可以有许多随机事件，希望通过对较简单的事件的了解，去掌握复杂的事件. 为此，需要研究事件之间的关系与运算.

由于事件是一个集合，因此，事件之间的关系与运算自然按照集合论中集合间的关系与运算来处理.下面这些关系和运算的提法是根据集合间的关系与运算，以及"事件发生"的含义给出的.

设试验 E 的样本空间为 S，而 $A,B,A_k(k=1,2,\cdots)$ 是 S 的子集.

（1）事件的包含 如果事件 A 发生必然导致事件 B 发生，则称事件 B 包含事件 A，记为 $B\supset A$ 或者 $A\subset B$.

例如，在试验 E_6 中，事件 A 表示"电子元件寿命不超过 1 300 小时"，事件 B 表示"电子元件寿命不超过 1 400 小时"，易见 $A\subset B$（或 $B\supset A$）.

（2）事件的相等 如果事件 $A\subset B$ 且事件 $B\subset A$，则称事件 A 与事件 B 相等，记为 $A=B$.

（3）和事件 事件 A 与 B 至少有一个发生的事件，称为事件 A 与 B 的和事件，记为 $A\cup B$

[①] 严格地说，事件是指 S 中的满足某些条件的子集. 当 S 是由有限个元素或由可列个元素组成时，每个子集都可作为一个事件. 若 S 是由不可列的无限个元素组成时，某些子集必须排除在外. 幸运的是，这种不容许的子集在实际问题中几乎不会遇到. 今后，每当我们讲到一个事件时，都是假定它是允许考虑的那种子集. 读者如有兴趣，可参考其他资料.

（或 $A+B$）.

例如，在试验 E_2 中，事件 A 表示“两次都出现正面”，事件 B 表示“两次都出现反面”，则和事件 $A\cup B$ 表示“两次出现同一面”.

类似地，称 $\bigcup\limits_{k=1}^{n} A_k$（或 $\sum\limits_{k=1}^{n} A_k$）为 n 个事件 $A_1, A_2, \cdots, A_n$ 的和事件，称 $\bigcup\limits_{k=1}^{+\infty} A_k$（或 $\sum\limits_{k=1}^{+\infty} A_k$）为可列个[①]事件 $A_1, A_2, \cdots$ 的和事件.

（4）积事件 事件 A 与 B 同时发生的事件，称为事件 A 与 B 的积事件，记为 $A\cap B$（或 AB）.

例如，某种圆柱形零件的长度与外径都合格才是合格的，事件 A 表示“长度合格”，B 表示“外径合格”，那么积事件 $A\cap B$ 表示“产品合格”.

类似地，称 $\bigcap\limits_{k=1}^{n} A_k$ 为 n 个事件 $A_1, A_2, \cdots, A_n$ 的积事件，称 $\bigcap\limits_{k=1}^{+\infty} A_k$ 为可列个事件 $A_1, A_2, \cdots$ 的积事件.

（5）差事件 事件 A 发生而 B 不发生的事件，称为事件 A 与 B 的差事件，记为 $A-B$.

例如，某输油管道长 10 km，事件 A 表示“前 5 km 正常工作”，事件 B 表示“后 5 km 正常工作”，那么差事件 $A-B$ 表示“前 5 km 正常工作，而后 5 km 非正常工作”.

（6）互不相容事件 如果事件 A 与 B 不能同时发生，则称事件 A 与 B 互不相容（或**互斥**）.

如果 n 个事件 $A_1, A_2, \cdots, A_n$ 中任意两个事件是互不相容的，则称这 n 个事件 $A_1, A_2, \cdots, A_n$ 两两互不相容.

如在试验 E_6 中，记事件 A 表示“电子元件寿命不超过 2 000 小时”，事件 B 表示“电子元件寿命至少为 3 000 小时”，那么事件 A 与 B 互不相容. 如果事件 C 表示“电子元件寿命超过 2 000 小时，少于 3 000 小时”，那么 A、B、C 两两互不相容.

（7）对立事件 如果事件 A 与 B 必有一个发生，且仅有一个发生，则称事件 A 与 B 互为对立事件. 事件 A 的对立事件记为 $\overline{A}$.

例如，某公司经过一年的市场运作，事件 A 表示“该公司年底结算赢利”，那么事件 $\overline{A}$ 表示“公司年底结算不赢利”.

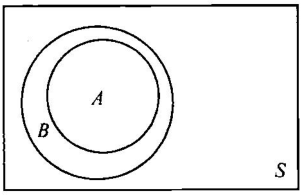
（8）完备事件组 如果事件 $A_1, A_2, \cdots, A_n$ 两两互不相容，即 $A_iA_j=\varnothing\,(i\neq j, i, j=1,2,\cdots,n)$，且 $\bigcup\limits_{k=1}^{n} A_k=S$，则称 $A_1, A_2, \cdots, A_n$ 是样本空间 S 的一个完备事件组或样本空间 S 的一个**划分**.

若 $A_1, A_2, \cdots, A_n$ 是一个完备事件组，那么对于每次试验，事件 $A_1, A_2, \cdots, A_n$ 中有一个且只有一个发生.

例如，在 E_4 中，E_4 的一组事件 $A_1=\{1,2,3\}$，$A_2=\{4,5\}$，$A_3=\{6\}$ 是 S 的一个划分，而事件 $B_1=\{1,2,3\}$，$B_2=\{3,4,5\}$，$B_3=\{6\}$ 不是 S 的划分.

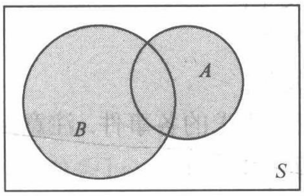
用图 1-1 至图 1-6 可直观表示上述事件间的关系及运算. 例如，在图 1-1 中长方形表示样本空间 S，圆 A 与圆 B 分别表示事件 A 与事件 B，事件 B 包含事件 A. 又如，在图 1-2 中长方形表示样本空间 S，圆 A 与圆 B 分别表示事件 A 与事件 B，而阴影部分表示和事件 $A\cup B$.

[①] 可列个的含义是指无限多个元素可以按照某种次序排成一列，例如，自然数有可列无限多个.



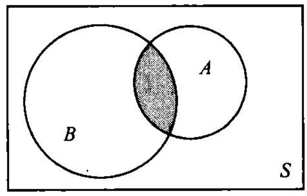
$A\subset B$

图 1-1



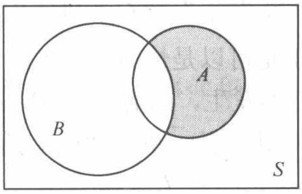
$A\cup B$

图 1-2



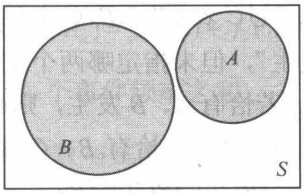
$A\cap B$

图 1-3



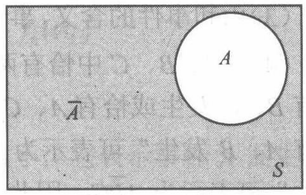
$A-B$

图 1-4



$A\cap B=\Phi$

图 1-5



$A\cup\overline{A}=S$　　$A\cap\overline{A}=\Phi$

图 1-6

在进行事件的运算时，经常要用到下列定律. 设 A, B, C 为事件，则有

交换律：$A\bigcup B=B\bigcup A$；

$A\bigcap B=B\bigcap A$.

结合律：$(A\bigcup B)\bigcup C=A\bigcup(B\bigcup C)$；

$(A\bigcap B)\bigcap C=A\bigcap(B\bigcap C)$.

分配律：$(A\bigcup B)\bigcap C=(A\bigcap C)\bigcup(B\bigcap C)$；

$(A\bigcap B)\bigcup C=(A\bigcup C)\bigcap(B\bigcup C)$.

德·摩根[①]**律**：$\overline{A\bigcup B}=\overline{A}\bigcap\overline{B}$；

$\overline{A\bigcap B}=\overline{A}\bigcup\overline{B}$.

例 1.2　设试验 E 为抛一枚骰子观察出现的点数，样本空间 $S=\{1,2,3,4,5,6\}$. 事件 $A=\{2,4,6\}$，$B=\{1,3\}$，$C=\{3,4\}$，求 $A\bigcap B$，$B\bigcup C$，$A\bigcup(B\bigcap C)$，$\overline{A\bigcup B}$，$C-A$.

解　$A\bigcap B=\{2,4,6\}\bigcap\{1,3\}=\varnothing$

$B\bigcup C=\{1,3\}\bigcup\{3,4\}=\{1,3,4\}$

$A\bigcup(B\bigcap C)=\{2,4,6\}\bigcup(\{1,3\}\bigcap\{3,4\})=\{2,4,6\}\bigcup\{3\}=\{2,3,4,6\}$

$\overline{A\bigcup B}=\overline{\{1,2,3,4,6\}}=\{5\}$

$C-A=\{3,4\}-\{2,4,6\}=\{3\}$

例 1.3　设 A、B、C 为三事件，用 A、B、C 的运算关系表示下列各事件：

（1）A 发生，B 与 C 不发生；

（2）A、B、C 都不发生；

（3）A、B、C 中至少有一个发生；

（4）A、B、C 中恰有两个发生；

---

① 德·摩根（De Morgan，1806—1871），19 世纪有相当的影响力的数学家，主要在分析学、代数学、数学史及逻辑学等方面做出重要的贡献.

（5）A、B、C 中至少有两个发生；

（6）A、B、C 中不多于一个发生；

（7）A、B、C 中至多有两个发生.

解 以 $G_i(i=1,2,3,4,5,6,7)$ 表示所求的各事件. 注意到一个事件不发生，则它的对立事件一定发生. 如事件 A 不发生，则 $\overline{A}$ 一定发生.

（1）“事件 A 发生，B 与 C 不发生”意味着“A、$\overline{B}$、$\overline{C}$ 三个事件同时发生”，即有 $G_1 = A\overline{B}\overline{C}$.

（2）“A、B、C 都不发生”也就是“$\overline{A}$、$\overline{B}$、$\overline{C}$ 都发生”，即 $G_2 = \overline{A}\,\overline{B}\,\overline{C}$.

（3）由和事件的含义，事件 $A\cup B\cup C$ 表示“A、B、C 中至少有一个发生”，故 $G_3 = A\cup B\cup C$.

（4）“A、B、C 中恰有两个发生”，但未指定哪两个发生，于是可以是恰有 A、B 发生或恰有 B、C 发生或恰有 A、C 发生. 若恰有 A、B 发生，则必有 C 不发生，因而“A、B、C 中恰有 A、B 发生”可表示为 $AB\overline{C}$. 类似地，“恰有 B、C 发生”可表示为 $\overline{A}BC$；“恰有 A、C 发生”可表示为 $A\overline{B}C$. 因此，$G_4 = AB\overline{C}\cup\overline{A}BC\cup A\overline{B}C$.

（5）“A、B、C 中至少有两个发生”等价于“A、B 同时发生”或“A、C 同时发生”或“B、C 同时发生”，即事件 AB、BC、CA 中至少有一个发生，即 $G_5 = AB\cup BC\cup CA$.

（6）“A、B、C 中不多于一个发生”表示“A、B、C 中至少有两个不发生”，也就是“$\overline{A}\,\overline{B}$、$\overline{B}\,\overline{C}$、$\overline{A}\,\overline{C}$ 中至少有一个发生”，即 $G_6 = \overline{A}\,\overline{B}\cup\overline{A}\,\overline{C}\cup\overline{B}\,\overline{C}$.

又“A、B、C 中不多于一个发生”的对立事件是“A、B、C 中至少有两个发生”，所以也可以把“A、B、C 中不多于一个发生”表示为“A、B、C 都不发生或者恰有一个发生”，故又有 $G_6 = \overline{A}\,\overline{B}\,\overline{C}\cup A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C$.

（7）“A、B、C 中至多有两个发生”等价于“A、B、C 中至少有一个不发生”，即 $G_7 = \overline{A}\cup\overline{B}\cup\overline{C}$.

我们也可把“A、B、C 中至多有两个发生”分解成“事件 A、B、C 三个事件不同时发生”、“A、B、C 三个事件恰好有一个发生”、“A、B、C 三个事件恰好有两个发生”的和事件，即 $G_7 = \overline{A}\,\overline{B}\,\overline{C}\cup\overline{A}\,\overline{B}C\cup\overline{A}B\overline{C}\cup A\overline{B}\,\overline{C}\cup AB\overline{C}\cup\overline{A}BC\cup A\overline{B}C$.

此外，“A、B、C 中至多有两个发生”的对立事件为“A、B、C 三个事件都发生”，所以 $G_7 = \overline{ABC}$.

1.3 随机事件的概率

除了必然事件和不可能事件外，任意一个事件在一次试验中可能发生，也可能不发生. 我们常常希望知道这些事件在一次试验中发生的可能性的大小. 例如，商业保险机构为了获得最大利润，就必须研究意外事件发生可能性的大小. 为了确定水坝的高度，就要知道河流在造水坝地段每年最大洪水达到某一高度这一事件发生的可能性的大小. 我们希望找到一个恰当的数字刻画事件在一次试验中发生的可能性的大小. 为此，我们先引入频率的概念.

1.3.1 事件的频率

定义 1.1 在相同的条件下将试验重复进行 n 次，在这 n 次试验中，事件 A 发生了 n_A 次，

n_A 称为事件 A 在这 n 次试验中发生的**频数**，而比值 $f_n(A)=\frac{n_A}{n}$ 称为事件 A 在这 n 次试验中发生的**频率**.

根据定义，易知频率具有如下性质：

（1）对任意的事件 A，有 $0\leqslant f_n(A)\leqslant 1$；

（2）对于必然事件 S，$f_n(S)=1$；

（3）若 $A_1,A_2,\cdots,A_k$ 是两两互不相容的事件，则

$$f_n(A_1\cup A_2\cup\cdots\cup A_k)=f_n(A_1)+f_n(A_2)+\cdots+f_n(A_k)$$

即两两互不相容事件和的频率等于每个事件频率之和.

由于事件 A 的频率是它发生的次数与试验次数之比 $\frac{n_A}{n}$ 的大小，它表示事件 A 发生的频繁程度. 频率越大，事件 A 发生就越频繁，这就意味着 A 在一次试验中发生的可能性越大. 因此，直观的想法就是用事件 A 的频率表示事件 A 在一次试验中发生的可能性的大小，但是否可行呢？我们先考察下面的例子.

例 1.4　抛一枚质地均匀的硬币试验. 我们将一枚硬币抛掷 10 次、100 次、1 000 次，各做 10 遍. 统计数据如表 1-1 所示（n 表示抛硬币的次数，n_H 表示出现正面的次数，$f_n(H)$ 表示出现正面的频率）.

表 1-1

实验序号	n=10		n=100		n=1 000	
	n_H	$f_n(H)$	n_H	$f_n(H)$	n_H	$f_n(H)$
1	4	0.4	58	0.44	503	0.503
2	5	0.5	48	0.48	496	0.496
3	3	0.3	55	0.55	482	0.482
4	6	0.6	52	0.52	518	0.518
5	7	0.7	46	0.46	498	0.498
6	6	0.6	42	0.42	516	0.516
7	5	0.5	51	0.51	495	0.495
8	2	0.2	57	0.57	527	0.527
9	9	0.9	45	0.45	476	0.476
10	5	0.5	53	0.53	503	0.503

这种试验历史上有人做过，其数据如表 1-2 所示.

表 1-2

试　验　者	n	n_H	$f_n(H)$	$\lvert f_n(H)-0.5\rvert$
德・摩根	2 048	1 061	0.518 1	0.018 1
蒲丰[①]	4 040	2 048	0.506 9	0.006 9

① 蒲丰（Buffon, 1707—1788），法国数学家、自然科学家，几何概率的开创者.

续表

试验者	n	n_H	$f_n(H)$	$\|f_n(H)-0.5\|$
威廉·费勒[1]	10 000	4 979	0.497 9	0.002 1
皮尔逊[2]	12 000	6 019	0.501 6	0.001 6
皮尔逊	24 000	12 012	0.500 5	0.000 5
维尼	30 000	14 994	0.499 8	0.000 2

从上述数据可以看出，抛硬币的次数 n 较小时，出现正面的频率 $f_n(H)$ 在 0 与 1 之间波动，而且波动幅度较大. 但随着 n 的增大，频率 $f_n(H)$ 呈现出稳定性，即当 n 逐渐增大时，$f_n(H)$ 总在 0.5 的附近摆动，而摆动的幅度越来越小，也就是 $f_n(H)$ 逐渐稳定于 0.5.

例 1.5 考查英语中特定字母出现的频率. 当观察字母个数 n（试验次数）较小时，频率具有较大幅度的随机波动. 但当 n 增大时，频率呈现出稳定性. 表 1-3 就是一份英文字母的统计表[3].

表 1-3

字母	频率	字母	频率	字母	频率	字母	频率
A	0.078 8	B	0.015 6	C	0.026 8	D	0.038 9
E	0.126 8	F	0.025 6	G	0.018 7	H	0.057 3
I	0.070 7	J	0.001 0	K	0.006 0	L	0.039 4
M	0.024 4	N	0.070 6	O	0.077 6	P	0.018 6
Q	0.000 9	R	0.059 4	S	0.063 4	T	0.098 7
U	0.028 0	V	0.010 2	W	0.021 4	X	0.001 6
Y	0.020 2	Z	0.000 6				

大量的实验证实：当重复试验次数 n 逐渐增大时，频率 $f_n(A)$ 呈现出稳定性，逐渐稳定于某个常数. 这种“频率稳定性”即通常所说的统计规律性，它揭示了隐藏在随机现象中的规律性. 我们让试验重复大量的次数，计算 $f_n(A)$，以它刻画事件 A 发生的可能性的大小是相对合适的.

但是，在实际问题中，我们不可能也没有必要对每一个事件做大量的试验，从中求得事件的频率，用以刻画事件发生可能性的大小. 同时，为了理论研究的需要，我们从频率的稳定性和频率的性质得到启发，给出刻画事件发生可能性大小的概率的定义.

1.3.2 事件的概率

定义 1.2 设 E 为随机试验，S 是它的样本空间. 对于 E 中的每一个事件赋予一个实数，记为 $P(A)$，称为事件 A 的概率. 如果集合函数 $P(\cdot)$ 满足下列条件：

（1）**非负性**：对 E 中的每一个事件 A，有 $P(A)\geqslant 0$；

（2）**规范性**：对于必然事件，有 $P(S)=1$；

---

[1] 威廉·费勒（William Feller，1907—1970），克罗地亚裔美国数学家，20 世纪最伟大的概率学家之一.

[2] 皮尔逊（Pearson,1857—1936），英国数学家、哲学家，现代统计学的创始人之一，被尊称为统计学之父.

[3] Dewey. G. 统计了约 438 023 个英语单词中各字母出现的频率.

（3）**可列可加性**：设 $A_1, A_2, \cdots, A_n, \cdots$ 是两两互不相容的事件，即对于 $A_iA_j = \varnothing$，$i \neq j$，$i, j = 1, 2, 3, \cdots$，有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots) = P(A_1) + P(A_2) + \cdots + P(A_n) + \cdots$$

第 5 章中的伯努利[1]大数定理将阐明：当试验次数 n 充分大时，事件 A 的频率 $f_n(A)$ 在一定意义下接近于事件 A 的概率 $P(A)$. 基于这一事实，我们有理由用概率 $P(A)$ 来度量 A 在一次试验中发生可能性的大小.

由概率的定义可以推出概率具有以下性质.

性质 1　$P(\varnothing) = 0$

证　令 $A_n = \varnothing\ (n = 1, 2, 3, \cdots)$，则 $\bigcup\limits_{n=1}^{+\infty} A_n = \varnothing$，且 $A_iA_j = \varnothing$，$i \neq j$，$i, j = 1, 2, 3, \cdots$，由概率的可列可加性得

$$P(\varnothing) = P(\bigcup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} P(A_n) = \sum_{n=1}^{+\infty} P(\varnothing)$$

再由概率的非负性可知，$P(\varnothing) \geqslant 0$，故由上式知 $P(\varnothing) = 0$.

性质 2（可加性）　若 $A_1, A_2, \cdots, A_n$ 是两两互不相容的事件，则有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

证　令 $A_k = \varnothing\ (k = n+1, n+2, n+3, \cdots)$，则有 $A_iA_j = \varnothing$，$i \neq j$，$i, j = 1, 2, 3, \cdots$，由概率的可列可加性得

$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots \cup A_n) &= P(\bigcup_{k=1}^{+\infty} A_k) = \sum_{k=1}^{+\infty} P(A_k) = \sum_{k=1}^{n} P(A_k) + \sum_{k=n+1}^{+\infty} P(A_k) \\ &= P(A_1) + P(A_2) + \cdots + P(A_n) \end{aligned}$$

性质 2 得证.

性质 3　设 A、B 是两个事件，且 $A \subset B$，则有 $P(B - A) = P(B) - P(A)$，$P(A) \leqslant P(B)$.

证　由 $A \subset B$（参见图 1-1）得 $B = A \cup (B - A)$，且 $A \cap (B - A) = \varnothing$，再由性质 2 得

$$P(B) = P(A) + P(B - A)$$

即

$$P(B - A) = P(B) - P(A)$$

又由概率的非负性知 $P(B - A) \geqslant 0$，从而 $P(A) \leqslant P(B)$.

性质 4　对任意一事件 A，有 $P(A) \leqslant 1$.

证　因为 $A \subset S$，由性质 3 得

$$P(A) \leqslant P(S) = 1$$

性质 5　对任意一事件 A，有 $P(\overline{A}) = 1 - P(A)$.

证　因为 $A \cup \overline{A} = S$，$A \cap \overline{A} = \varnothing$，由性质 2 得

$$1 = P(S) = P(A \cup \overline{A}) = P(A) + P(\overline{A})$$

[1] 伯努利（Jacob Bernoulli，1654—1705），瑞士数学家，极坐标和大数定律的创始人.