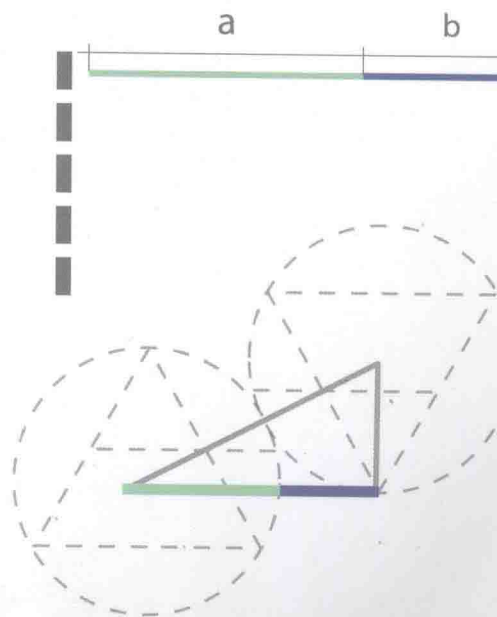


普通高等教育“十三五”规划教材

概率论 与数理统计

韩 明 编著



上海财经大学出版社
SHANGHAI UNIVERSITY OF FINANCE & ECONOMICS PRESS

普通高等教育“十二五”规划教材

概率论与数理统计

韩 明 编著



上海财经大学出版社

内容提要

全书共有九章组成,第1~第5章是概率论部分,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,大数定律及中心极限定理;第6~第9章是数理统计部分,内容包括:数理统计的基本概念,参数估计,假设检验,回归分析.并附录有“数学建模及大学生数学建模竞赛简介”和“概率论与数理统计实验简介”等.

本书既有继承国内相关教材的传统部分,又有汲取国外相关教材中流行的直观、灵活的风格.本书图文并茂,可读性强,着重讲解基本概念、统计思想,强调理论与方法的应用,并把数学实验与数学建模的思想方法融入教材中.

本书可供高等院校工科类、理科类(非数学专业)、经济管理类有关专业用作教材,也可供广大自学者参考.

图书在版编目(CIP)数据

概率论与数理统计 / 韩明编著. — 上海: 上海财经大学出版社, 2016. 8

(普通高等教育“十三五”规划教材)

ISBN 978-7-5642-2520-9/F. 2520

I. ①概… II. ①韩… III. ①概率论—高等学校—教材
②数理统计—高等学校—教材 IV. ①O21

中国版本图书馆 CIP 数据核字(2016)第 175050 号

□ 责任编辑 袁 敏

□ 封面设计 杨雪婷

GAILULUN YU SHULI TONGJI

概率论与数理统计

韩 明 编著

上海财经大学出版社出版发行
(上海市武东路 321 号乙 邮编 200434)

网 址: <http://www.sufep.com>

电子邮箱: webmaster@sufep.com

全国新华书店经销

同济大学印刷厂印刷

上海叶大印务发展有限公司装订

2016 年 8 月第 1 版 2016 年 8 月第 1 次印刷

787 mm×1 092 mm 1/16 13.25 印张 322 千字

印数: 0001—3000 定价: 28.00 元

前 言

本书是根据“教育部数学与统计学教学指导委员会”新修订的“工科类、经济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”，参考了近些年国内外出版的有关教材，并结合作者多年来的教学实践经验、教学研究成果和体会等编写而成的。

为贯彻落实 2010 年颁布的《国家中长期教育改革和发展规划纲要(2010—2020 年)》，教育部于 2011 年实施了《高等学校本科教学质量与教学改革工程》，2012 年又出台了《关于全面提高高等教育质量的若干意见》。这为新时期我国高等教育提出了战略部署，也为高等教育教学改革明确了方向。在 2013 年 9 月，中国高等教育学会发出了《关于组织申报“普通高等学校应用型本科教育教材规划”选题立项的通知》。本书作者经过多年来的教学实践，也深感一部适合我国应用型本科院校(非数学类专业)的《概率论及数理统计》教材的重要性。

“概率论与数理统计”是高等院校本科生各专业普遍开设的一门公共基础课程，它是研究随机现象的一门数学课程。阅读本书要求读者具备《高等数学》与《线性代数》的基本知识。本书可供高等院校工科类、理科类(非数学类专业)、经济管理类等有关专业用作教材(其中“*”部分是供选学的)，也可供广大自学者参考。

把数学实验(Mathematical Experiment)和数学建模(Mathematical Modelling)的思想方法融入大学数学主干课程，是当前大学数学教学改革的一个重要方向，也是在计算机普及的今天给大学数学教学改革提出的一个必须考虑的问题。本书借助数学实验把教材中一些不好理解的问题(如与极限有关的几个定理)通过可视化等形式展现给读者，使读者便于理解相关内容。关于数学建模方面的内容，见附录 A“数学建模及大学生数学建模竞赛简介”。本书把数学软件(MATLAB)与教材内容紧密结合，教材中的一些计算、画图的相关程序写在了附录 B 中。关于与“概率论与数理统计”课程有关的数学实验，见本书的附录 B“概率论与数理统计实验简介”。

本书有如下一些特色：

- (1) 在不违背科学性的前提下，尽量通俗；
- (2) 注重可读性，力争图文并茂。全书共有图 48 幅，相信它们会对读者理解相关内容有所帮助；
- (3) 着重讲解基本概念、统计思想，强调理论与方法的应用，适当淡化理论推导；
- (4) 全书精选例题 150 个，其中很多例题贴近日常生活、反映科技进步和社会发展，具有时代气息；
- (5) 习题按节设立，这样可以使习题具有针对性，全书共有习题 243 道，并在书后附有“习题参考答案”。

(6) 把数学实验与数学建模的思想方法融入教材中.

感谢王家宝教授的指导和鼓励. 本书的编写得到了宁波工程学院理学院的支持, 在此表示感谢.

虽然作者努力想把本书写成为一本既有新意又便于教学的教材, 但由于水平所限, 书中难免还有一些疏漏甚至错误, 恳请专家和读者批评指正.

韩 明

2016年8月

目 录

前言	1
第 1 章 随机事件与概率	1
1.1 随机事件及其运算	1
1.1.1 随机现象与统计规律性	1
1.1.2 随机试验与样本空间	1
1.1.3 随机事件、事件间的关系与运算	2
习题 1.1	5
1.2 事件的概率及其性质	5
1.2.1 频率与概率的统计定义	5
1.2.2 古典概型	6
1.2.3 几何概率	8
1.2.4 概率的公理化定义	9
习题 1.2	11
1.3 条件概率	12
1.3.1 条件概率与乘法公式	12
1.3.2 全概率公式与贝叶斯公式	13
习题 1.3	16
1.4 事件的独立性与伯努利概型	16
1.4.1 事件的独立性	16
1.4.2 伯努利概型	18
习题 1.4	19
本章附录 “概率论”发展简史	19
第 2 章 随机变量及其分布	22
2.1 随机变量的概念与离散型随机变量	22
2.1.1 随机变量的概念	22
2.1.2 离散型随机变量及其分布律	23



2.1.3 常见的离散型随机变量	24
习题 2.1	28
2.2 随机变量的分布函数	29
2.2.1 分布函数的定义	29
2.2.2 分布函数的性质	30
习题 2.2	31
2.3 连续型随机变量及其概率密度	32
2.3.1 连续型随机变量	32
2.3.2 常见的连续型随机变量	34
习题 2.3	40
2.4 随机变量函数的分布	41
2.4.1 离散型随机变量函数的分布	41
2.4.2 连续型随机变量函数的分布	41
习题 2.4	43
第3章 多维随机变量及其分布	45
3.1 二维随机变量及其分布	45
3.1.1 二维随机变量的定义、分布函数	45
3.1.2 二维离散型随机变量	46
3.1.3 二维连续型随机变量	47
习题 3.1	49
3.2 边缘分布	50
3.2.1 边缘分布律	50
3.2.2 边缘密度函数	52
习题 3.2	54
3.3 随机变量的独立性	55
习题 3.3	57
3.4 两个随机变量函数的分布	58
3.4.1 $Z=X+Y$ 的分布	58
3.4.2 $M=\max(X, Y)$ 和 $N=\min(X, Y)$ 的分布	60
习题 3.4	62
第4章 随机变量的数字特征	64
4.1 数学期望	64



4.1.1 数学期望的定义	64
4.1.2 随机变量函数的数学期望	66
4.1.3 数学期望的性质	69
习题 4.1	70
4.2 方差	71
4.2.1 方差的定义	71
4.2.2 方差的性质	72
4.2.3 常见分布的方差	73
习题 4.2	76
4.3 协方差、相关系数与矩	77
4.3.1 协方差与相关系数	77
4.3.2 独立性与不相关性	80
4.3.3 矩、协方差矩阵	82
习题 4.3	82
 第5章 大数定律及中心极限定理	 84
5.1 大数定律	84
5.1.1 切比雪夫不等式	84
5.1.2 三个大数定律	85
习题 5.1	89
5.2 中心极限定理	89
5.2.1 独立同分布中心极限定理	90
5.2.2 棣莫弗—拉普拉斯中心极限定理	91
习题 5.2	94
 第6章 数理统计的基本概念	 95
6.1 几个基本概念	95
6.1.1 总体与样本	95
6.1.2 直方图	97
6.1.3 统计量与样本矩	98
习题 6.1	100
6.2 三个重要分布与抽样定理	101
6.2.1 三个重要分布	101

6.2.2 正态总体下的抽样定理	106
习题 6.2	109
本章附录 “数理统计”发展简史	110
第7章 参数估计	113
7.1 点估计	113
7.1.1 矩估计法	113
7.1.2 极大似然估计法	115
习题 7.1	118
7.2 估计量的评选标准	119
7.2.1 无偏性	120
7.2.2 有效性与一致性	121
习题 7.2	122
7.3 区间估计	122
7.3.1 区间估计的定义	122
7.3.2 单个正态总体均值与方差的置信区间	125
7.3.3 两个正态总体均值之差与方差之比的置信区间	126
习题 7.3	129
第8章 假设检验	131
8.1 假设检验的基本思想与步骤	131
8.1.1 假设检验的基本思想	131
8.1.2 两类错误与假设检验的步骤	133
* 8.1.3 检验的 p -值	135
习题 8.1	136
8.2 单个正态总体均值与方差的检验	137
8.2.1 单个总体 $N(\mu, \sigma^2)$ 均值 μ 的检验	137
8.2.2 置信区间与假设检验的关系	138
8.2.3 单个总体 $N(\mu, \sigma^2)$ 方差 σ^2 的检验	139
习题 8.2	141
8.3 两个正态总体均值与方差的检验	141
8.3.1 两个正态总体均值之差的检验	141
8.3.2 两个正态总体方差之比的检验	143



习题 8.3	144
* 8.4 分布拟合检验	145
习题 8.4	148
* 第9章 回归分析	150
9.1 一元线性回归	150
9.1.1 基本概念	150
9.1.2 回归系数的最小二乘估计	151
9.1.3 回归方程的显著性检验	154
9.1.4 一元线性回归方程的预测	158
习题 9.1	159
9.2 可线性化的回归方程	160
习题 9.2	161
附录	163
附录 A 数学建模及大学生数学建模竞赛简介	163
附录 B 概率论与数理统计实验简介	168
附录 C 概率论与数理统计附表	174
习题参考答案	191
参考文献	200

第 1 章 随机事件与概率

在考虑一个事件是否会发生的时候,人们常关心该事件发生可能性的大小,就像用尺子来测量物体的长度一样,我们用概率来测量一个事件发生可能性的大小.

1990 年诺贝尔经济学奖的三位得主之一是马科维茨(Markowitz),他获奖的主要原因是提出了投资组合选择(portfolio selection)理论,他把投资组合的价格视为随机变量,用它的均值来衡量收益,用它的方差来度量风险(被称为“均值一方差分析理论”),该理论后来被誉为“华尔街的第一次革命”(注:随机变量、均值、方差是本课程后面将要介绍的内容).

《统计与真理——怎样运用偶然性》(C. R. Rao,美国宾西法尼亚州立大学教授,2002 年美国总统科学奖获得者)的扉页上写有这样一段话:

在终极的分析中,一切知识都是历史;
在抽象的意义下,一切科学都是数学;
在理性的基础上,所有的判断都是统计学.

1.1 随机事件及其运算

1.1.1 随机现象与统计规律性

在自然界与人类社会活动中,人们观察到的现象是多种多样的,但归结起来它们大体上可以分为两类:一类是确定性现象,另一类是随机现象.例如,向上抛一个石子必然下落;同性电荷必然相互排斥.这类在一定条件下必然发生的现象,称为**确定性现象**(或**必然现象**).

在相同条件下抛一枚硬币,其结果可能是正面朝上,也可能是反面朝上,在抛掷之前无法预知抛掷的结果,结果呈现出不确定性;但多次重复抛同一枚硬币,得到正面朝上与反面朝上两个结果大致各占一半,结果呈现出规律性.在大量重复试验中,其结果所呈现出的规律性,称为**统计规律性**.这类在个别试验中其结果呈现出不确定性,在大量重复试验中其结果呈现出规律性的现象,称为**随机现象**(或**偶然现象**).值得注意的是,确定性现象在一定条件下其结果只有一个,而随机现象其结果不止一个.

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,其理论与方法的应用非常广泛,几乎遍及所有科学技术领域、工农业生产、国民经济以及我们的日常生活.

1.1.2 随机试验与样本空间

我们遇到过各种试验,包括各种科学试验.在这里我们把试验作广义理解,对某一事物的某一特征的观察,也认为是一种试验.为了研究随机现象的统计规律性,我们需要进行各种试验.

如果一个试验同时满足下列三个条件:

- (1) 可以在相同的条件下重复地进行(简称“可重复性”);
- (2) 每次试验的可能结果不止一个,并且能事先明确试验的所有可能结果(简称“唯一性”);
- (3) 进行一次试验之前不能确定哪一个结果会出现(简称“不确定性”).

则称这样的试验为随机试验,有时把随机试验简称为试验(experiment),用 E 来表示.我们是通过随机试验来研究随机现象的.

值得注意的是,随机试验要求试验在相同的条件下可以重复.当然也有很多随机现象是不能重复的,例如某场足球赛的输赢是不能重复的,某些经济现象(如经济增长率等)也是不能重复的.概率论与数理统计主要研究能大量重复的随机现象,但也十分注意研究不能重复的随机现象.

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,用 Ω 来表示.样本空间 Ω 中的元素,即试验 E 的每个结果,称为样本点,用 ω 来表示.

【例 1.1.1】 以下是七个随机试验,请写出它们的样本空间.

E_1 :抛一枚硬币,用 H (head)表示正面朝上,用 T (tail)表示反面朝上,观察正面和反面出现的情况;

E_2 :将一枚硬币抛掷三次,观察正面 H 和反面 T 出现的情况;

E_3 :将一枚硬币抛掷三次,观察正面出现的次数;

E_4 :抛一颗骰子,观察出现的点数;

E_5 :记录某城市 114 电话号码查询台一昼夜接到的呼叫次数;

E_6 :在一批灯泡中任意抽取一只,测试它的寿命;

E_7 :向平面区域 $D=\{(x, y) | x^2+y^2 \leq 1\}$ 内随机投掷一点,观察落点的坐标(假设该落点一定落在 D 内).

解:上面七个随机试验 $E_1, E_2, \dots, E_7$ 的样本空间分别为:

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

$$\Omega_2 = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\};$$

$$\Omega_3 = \{0, 1, 2, 3\};$$

$$\Omega_4 = \{1, 2, 3, 4, 5, 6\};$$

$$\Omega_5 = \{0, 1, 2, 3, \dots\};$$

$$\Omega_6 = \{t | t \geq 0\};$$

$$\Omega_7 = \{(x, y) | x^2 + y^2 \leq 1\}.$$

应该注意的是,样本空间中的元素是由试验的目的所确定的.例如,在例 1.1.1 中, E_2 和 E_3 同是将一枚硬币抛掷三次,由于试验的目的不同,样本空间中的元素也不同.

1.1.3 随机事件、事件间的关系与运算

在进行随机试验时,人们常常关心满足某种条件的那些样本点组成的集合,即“随机试验的某些样本点组成的集合”(亦即样本空间的子集).例如,若规定某种灯泡的寿命小于 1 000 h 为次品,则我们在例 1.1.1 的 E_6 中关心是否有 $t \geq 1\,000$ h,满足这个条件的样本点组成样本空间 Ω_6 的一个子集 $\{t | t \geq 1\,000\}$.

称试验 E 的样本空间 Ω 的子集为 E 的随机事件(或“随机试验的某些样本点组成的集

合”),简称事件(event).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.随机事件一般用大写字母 A, B, C 等来表示.

【例 1.1.2】在例 1.1.1 中,看几个事件的例子.对于 E_2 ,事件“第一次出现 H ”,即 $A_1 = \{HHH, HHT, HTH, HTT\}$;事件“三次出现同一面”,即 $A_2 = \{HHH, TTT\}$.对于 E_4 ,事件“出现偶数点”,即 $A_3 = \{2, 4, 6\}$.

由一个样本点组成的单点集,称为基本事件.例如,在例 1.1.1 的 E_1 中,有两个基本事件 $\{H\}$ 和 $\{T\}$;在 E_3 中,有 4 个基本事件 $\{0\}, \{1\}, \{2\}, \{3\}$.

样本空间 Ω 包含所有样本点,它是自身的子集,在每次试验中它总是发生的,称为必然事件.

空集 $\emptyset$ 不包含任何样本点,它也作为样本空间的子集,它在每次试验中都不发生,称为不可能事件.

事件是一个集合,所以事件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.下面这些关系与运算的提法,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 Ω ,而 $A, B, A_i (i=1, 2, \dots)$ 是 Ω 的子集.

(1) 若 $A \subset B$,则称事件 B 包含事件 A ,这指的是事件 A 发生必然导致事件 B 发生.

若 $A \subset B$ 且 $A \supset B$,则称事件 A 与事件 B 相等,记为 $A=B$.

(2) 事件 $A \cup B = \{x | x \in A \text{ 或 } x \in B\}$ 称为事件 A 与事件 B 的和事件(或事件的并).当且仅当 A, B 中至少有一个事件发生时,事件 $A \cup B$ 发生.

类似地,称 $\bigcup_{i=1}^n A_i$ 为 n 个事件 $A_1, A_2, \dots, A_n$ 的和事件,称 $\bigcup_{i=1}^{+\infty} A_i$ 为可列个事件 $A_1, A_2, \dots$ 的和事件.

(3) 事件 $A \cap B = \{x | x \in A \text{ 且 } x \in B\}$ 称为事件 A 与事件 B 的积事件(或事件的交).当且仅当 A, B 同时发生时,事件 $A \cap B$ 发生.事件 A 与事件 B 的积事件,简记作 AB .

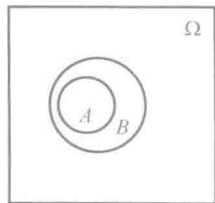
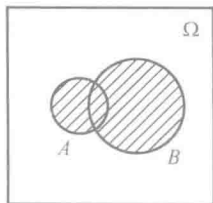
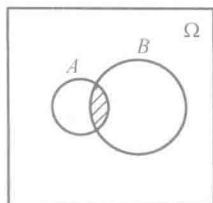
类似地,称 $\bigcap_{i=1}^n A_i$ 为 n 个事件 $A_1, A_2, \dots, A_n$ 的积事件,称 $\bigcap_{i=1}^{+\infty} A_i$ 为可列个事件 $A_1, A_2, \dots$ 的积事件.

(4) 事件 $A - B = \{x | x \in A \text{ 且 } x \notin B\}$ 称为事件 A 与事件 B 的差事件(或事件的差).当且仅当 A 发生而 B 不发生时,事件 $A - B$ 发生.

(5) 若 $A \cup B = \Omega$ 且 $A \cap B = \emptyset$,则称事件 A 与事件 B 互为对立事件(或逆事件).记事件 A 的对立事件为 $\bar{A}$, $\bar{A} = \Omega - A$.

(6) 若 $A \cap B = \emptyset$,则称事件 A 与事件 B 互不相容(或互斥).这指的是事件 A 与事件 B 不能同时发生.显然,同一个试验中各个基本事件是两两互不相容的.

我们可以用维恩(Venn)图来表示上述事件间的关系,如图 1-1~图 1-6 所示.

图 1-1 $A \subset B$ 图 1-2 $A \cup B$ 图 1-3 $A \cap B$

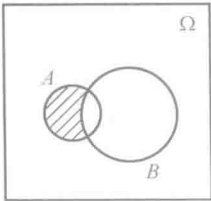


图 1-4 $A - B$

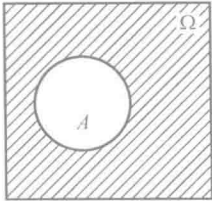


图 1-5 $\bar{A}$

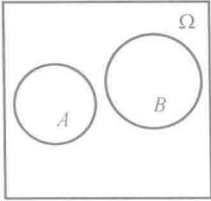


图 1-6 A 与 B 互不相容

在进行事件的运算时,经常要用到下述定律. 设 $A, B, C, A_i (i=1, 2, \dots, n)$ 为事件, 则有

交换律: $A \cup B = B \cup A, A \cap B = B \cap A$.

结合律: $A \cup (B \cup C) = (A \cup B) \cup C, A \cap (B \cap C) = (A \cap B) \cap C$.

分配律: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C), A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

对偶律: $\overline{A \cup B} = \bar{A} \cap \bar{B}, \overline{A \cap B} = \bar{A} \cup \bar{B}; \overline{\bigcup_{i=1}^n A_i} = \bigcap_{i=1}^n \bar{A}_i, \overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \bar{A}_i$.

在集合论、概率论中符号与意义的对照,见表 1-1.

表 1-1 在集合论、概率论中符号与意义的对照

符号	集合论	概率论
Ω	全集	样本空间, 必然事件
$\emptyset$	空集	不可能事件
$\omega (\in \Omega)$	元素	样本点
$\{\omega\}$	单点集	基本事件
$A (\subset \Omega)$	子集 A	事件 A
$A \subset B$	集合 B 包含集合 A	事件 B 包含事件 A
$A = B$	集合 A 与 B 相等	事件 A 与 B 相等
$A \cup B$	集合 A 与 B 的并集	事件 A 与 B 的和事件
$A \cap B$	集合 A 与 B 的交集	事件 A 与 B 的积事件
$\bar{A}$	集合 A 的余集	事件 A 的对立事件
$A - B$	集合 A 与 B 的差集	事件 A 与 B 的差事件
$A \cap B = \emptyset$	集合 A 与 B 没有公共元素	事件 A 与 B 互不相容

【例 1.1.3】 考察学生在一次数学考试中的成绩(括号中的区间表示成绩所处的范围), 记 $A = \text{“优秀”}([90, 100])$, $B = \text{“良好”}([80, 90))$, $C = \text{“中等”}([70, 80))$, $D = \text{“及格”}([60, 70))$, $E = \text{“未通过”}([0, 60))$, $F = \text{“通过”}([60, 100])$, 则 A, B, C, D, E 为两两互不相容事件; E 与 F 互为对立事件, 即 $\bar{E} = F$; $F = A \cup B \cup C \cup D$.

【例 1.1.4】 对于例 1.1.2 中的 $A_1 = \{HHH, HHT, HTH, HTT\}$, $A_2 = \{HHH, TTT\}$, 求 $A_1 \cup A_2, A_1 \cap A_2, A_1 - A_2, \bar{A}_1 \cup \bar{A}_2$.

解: 根据例 1.1.1 知样本空间为 $\Omega_2 = \{HHH, HHT, HTH, THH, HTT, THT,$



$TTT\}$, 则 $A_1 \cup A_2 = \{HHH, HHT, HTH, HTT, TTT\}$, $A_1 \cap A_2 = \{HHH\}$, $A_1 - A_2 = \{HHT, HTH, HTT\}$, $\overline{A_1 \cup A_2} = \Omega - A_1 \cup A_2 = \{THT, TTH, THH\}$.

习 题 1.1

1. 以下是五个随机试验, 请写出它们的样本空间:
 - (1) 抛一枚硬币, 用 H 表示正面朝上, 用 T 表示反面朝上, 观察正面和反面出现的情况;
 - (2) 将一枚硬币抛掷两次, 观察正面 H 、反面 T 出现的情况;
 - (3) 将一枚硬币抛掷两次, 观察正面出现的次数;
 - (4) 在单位圆内任意取一点, 记录它的(直角)坐标;
 - (5) 掷两颗骰子, 观察其点数.
2. 袋中装有编号为 1, 2, 3, 4, 5 的五个相同的球. 若从中任取三个球, 请写出这个随机试验的样本空间, 并计算基本事件总数.
3. 设 A, B, C 表示三个随机事件, 用 A, B, C 的运算关系表示下列各事件: (1) A, B, C 都发生; (2) A, B, C 都不发生; (3) A, B, C 中恰好有两个发生.
4. 一名射手向某个目标射击三次, 事件 A_i 表示射手第 i 次射击时击中目标 ($i=1, 2, 3$), 试用文字叙述下列事件: (1) $\bar{A}_1 \cup \bar{A}_2$; (2) $A_1 \cup A_2 \cup A_3$; (3) $\bar{A}_1 A_2$; (4) $A_2 \cup \bar{A}_3$.
5. 一位工人生产四个零件, 以事件 A_i 表示他生产的第 i 个零件是不合格品, $i=1, 2, 3, 4$. 请用诸 A_i 表示如下事件: (1) 全是合格品; (2) 全是不合格品; (3) 至少有一个零件是不合格品; (4) 恰好有一个零件是不合格品.
6. 请叙述下述事件的对立事件: (1) A = “掷两枚硬币, 皆为正面”; (2) B = “射击三次, 皆命中目标”; (3) C = “加工四个产品, 至少有一个正品”.
7. 下列说法是否正确, 为什么? (1) 若 $A \cup B = \Omega$, 则 A, B 互为对立事件; (2) 若 $ABC = \emptyset$, 则 A, B, C 两两互不相容.
8. 在分别标有 1, 2, 3, 4, 5, 6, 7, 8 的八张卡片中任取一张, 设事件 A 为“抽得一张标号不大于 4 的卡片”; 事件 B 为“抽得一张标号为偶数的卡片”; 事件 C 为“抽得一张标号为奇数的卡片”. 请用样本点表示如下事件: $A \cup B, AB, \bar{B}, A - B, B - A, BC, \overline{B \cup C}, (A \cup B)C$.

1.2 事件的概率及其性质

我们观察一个随机试验的各种事件, 一般来说, 有些事件出现的可能性大些, 有些事件出现的可能性小些, 我们需要一个刻画事件发生可能性大小的数量指标. 在实际问题中, 经常需要对随机事件发生的可能性大小进行定量计算, 而“概率”的概念正是源于这种需要而产生的.

1.2.1 频率与概率的统计定义

定义 1.2.1 在相同条件下, 进行了 n 次试验, 在这 n 次试验中, 事件 A 发生的次数 n_A 称为事件 A 发生的**频数**, 比值 $\frac{n_A}{n}$ 称为事件 A 发生的**频率**(frequency), 并记作 $f_n(A)$.

根据定义 1.2.1, 易知频率具有下述基本性质:

- (1) 对于任意事件 A , 有 $0 \leq f_n(A) \leq 1$;

- (2) 对于必然事件 Ω , $f_n(\Omega)=1$;
- (3) 对于两两互不相容的事件 $A_1, A_2, \cdots, A_k$, 有

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

由于事件 A 的频率是它发生的次数与试验次数之比 $\frac{n_A}{n}$, 其大小表示事件 A 发生的频繁程度, 因此, 直观的想法是用事件 A 的频率表示事件 A 在一次试验中发生的可能性的. 但这是否可行呢? 我们先看下面的例子.

【例 1.2.1】 抛一枚质地均匀硬币的试验, 历史上有人做过. 设 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, 得到如表 1-2 所示的数据.

表 1-2 抛硬币试验

试验者	n	n_H	$f_n(H)$	$ f_n(H)-0.5 $
德摩根	2 048	1 061	0.518 1	0.018 1
蒲丰	4 040	2 048	0.506 9	0.006 9
费勒	10 000	4 979	0.497 9	0.002 1
皮尔逊	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

从表 1-2 中的数据可以看出, 抛硬币的次数 n 较小时, 出现正面的频率 $f_n(H)$ 在 0 与 1 之间波动相对较大. 但随着 n 的增大, $f_n(H)$ 呈现出稳定性, 即当 n 逐渐增大时, $f_n(H)$ 总在 0.5 附近徘徊, 而逐渐稳定于 0.5.

例 1.2.1 说明, 随机事件在大量重复试验中其结果呈现出某种规律性, 而频率的稳定性正是这种规律性的表现.

定义 1.2.2(概率的统计定义) 在大量重复试验中, 若事件 A 发生的频率稳定地在某一个常数 p 附近摆动, 则称该常数 p 为事件 A 发生的**概率**(probability), 记作 $P(A)$, 即 $P(A)=p$.

应该指出, 频率是变动的, 而概率(频率的稳定值)则是常数. 频率提供了概率的一个可供想象的具体值, 并且在试验重复次数较大时, 可用频率作为概率的近似值, 这一点是频率最有价值的地方. 在日常生活中, 我们经常说的产品合格率、彩票中奖率等其实都是频率.

在足球比赛中, 人们很关心罚点球命中的可能性大小. 有人曾对 1930~1988 年世界各地的 53 274 场重大足球比赛作了统计, 在判罚的 15 382 个点球中有 11 172 个命中. 由此可得罚点球命中概率的近似值为 $\frac{11\,172}{15\,382}=0.726\,3$.

1.2.2 古典概型

在例 1.1.1 中的 E_1 和 E_4 , 它们具有两个共同特点:

- (1) 试验的样本空间只包含有限个元素;

(2) 试验中的每一个基本事件发生的可能性相同.

具有以上两个特点的试验,称为**古典型试验**.

定义 1.2.3(概率的古典型定义) 设随机试验 E 为古典型试验,它的样本空间为 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$,事件 A 包含 k 个基本事件,则事件 A 的概率为

$$P(A) = \frac{k}{n} \quad (1.2.1)$$

这里 k = 事件 A 包含的基本事件数, $n = \Omega$ 中基本事件的总数.

称满足定义 1.2.3 的概率模型为**古典概型**.显然,在古典概型中基本事件发生的概率都相等,因此古典概型又称为**等可能概型**.古典概型在概率论的产生和发展过程中是最早且最常用到的一种概率模型.

【例 1.2.2】 将一枚硬币抛掷三次.(1)设事件 A_1 为“恰有一次出现正面”,求 $P(A_1)$;
(2)设事件 A_2 为“至少有一次出现正面”,求 $P(A_2)$.

解:(1)“将一枚硬币抛掷三次”这个试验的样本空间为 $\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$,而 $A_1 = \{HTT, THT, TTH\}$.这是古典概型问题,根据式(1.2.1),得 $P(A_1) = \frac{3}{8}$.

(2) 由于 $A_2 = \{HHH, HHT, HTH, THH, HTT, THT, TTH\}$,所以根据式(1.2.1),得 $P(A_2) = \frac{7}{8}$.

【例 1.2.3】 将 n 只球随机地放入 $N(N \geq n)$ 个盒子中去,试求每个盒子至多有一只球的概率(设盒子的容量不限).

解:将 n 只球随机地放入 $N(N \geq n)$ 个盒子中去,每种放法是一个基本事件.易知,这是古典概型问题.由于每一只球都可以放入 N 个盒子中的任意一个,故共有 $N \times N \times \dots \times N = N^n$ 种不同的放法.而每个盒子至多放有一只球,共有 $N \times (N-1) \times \dots \times [N-(n-1)]$ 种不同的放法.根据式(1.2.1),得所求的概率为

$$\frac{N \times (N-1) \times \dots \times [N-(n-1)]}{N^n}.$$

有许多问题和本例有相同的数学模型.例如生日问题,假设每个人的生日在一年 365 天中的任意一天是等可能的,即等于 $\frac{1}{365}$,那么随机选取 $n(n \leq 365)$ 个人,根据例 1.2.3 的结果(取 $N=365$),则他(她)们的生日各不相同的概率为

$$\frac{365 \times 364 \times \dots \times [365 - (n-1)]}{365^n}.$$

则 n 个人中至少有两人生日相同的概率为

$$p_n = 1 - \frac{365 \times 364 \times \dots \times [365 - (n-1)]}{365^n}.$$

对 $n=10, 20, 30, 40, 50, 60, 70, 80$,计算结果如下表(MATLAB 程序见附录 B 的例