

高 等 学 校 教 材

工程数学

概率统计简明教程

第二版

同济大学数学系 编

 高等教育出版社
HIGHER EDUCATION PRESS

高等学校教材

工程数学

概率统计简明教程

Gailü Tongji Jianming Jiaocheng

第二版

同济大学数学系 编



高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

内容提要

本书是普通高等教育“十五”国家级规划教材《工程数学 概率统计简明教程》的第二版,保持了第一版的特色,内容简练,直观性和可读性强,且富有思想内涵,特别适合少学时“概率论与数理统计”课程的教学需要。

本书在广泛征求读者意见的基础上对第一版教材做了适当修订,补充了若干有典型意义的例子、增加了一定数量的习题。本书内容包括随机事件、事件的概率、条件概率与事件的独立性、随机变量及其分布、二维随机变量及其分布、随机变量的函数及其分布、随机变量的数字特征、统计量和抽样分布、点估计、区间估计、假设检验、一元线性回归等。每章末有小结,书末附有辅助材料——统计软件 Excel简介。

本书可供高等院校工科各专业及其他非数学类专业学生使用,也适用于多层次办学的“概率论与数理统计”课程的教学需要。

图书在版编目(CIP)数据

工程数学. 概率统计简明教程/ 同济大学数学系编.
--2 版. --北京:高等教育出版社,2012. 6
ISBN 978 - 7 - 04 - 035198 - 9

I. ①工… II. ①同… III. ①工程数学 - 高等学校 - 教材②概率论 - 高等学校 - 教材③数理统计 - 高等学校 - 教材 IV. ①TB11

中国版本图书馆 CIP 数据核字(2012)第 108050 号

策划编辑 王 强 责任编辑 蒋 青 封面设计 赵 阳 版式设计 于 婕
插图绘制 黄建英 责任校对 刘春萍 责任印制 尤 静

出版发行	高等教育出版社	网 址	http://www.hep.edu.cn
社 址	北京市西城区德外大街 4 号		http://www.hep.com.cn
邮政编码	100120	网上订购	http://www.landraco.com
印 刷	三河市华润印刷有限公司		http://www.landraco.com.cn
开 本	787mm×960mm 1/16	版 次	2003 年 7 月第 1 版
印 张	16.5		2012 年 6 月第 2 版
字 数	300 千字	印 次	2012 年 6 月第 1 次印刷
购书热线	010-58581118	定 价	30.80 元(含光盘)
咨询电话	400-810-0598		

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换
版权所有 侵权必究
物 料 号 35198-00

第二版前言

本书第一版于2003年出版。八年的教学实践,使作者团队积累了丰富的教学经验,同时我们听取和吸收广大读者的宝贵意见,在此基础上对第一版进行了修订。

本书的再版以提高教材质量,方便读者使用为宗旨,除修改第一版中存在的疏漏和不当之处外,主要作了以下几个方面的修订:

一、将第八、九章合并为一章;增加了一元线性回归作为新的一章;添加了辅助材料——统计软件 Excel 简介,使教材更富实用性。

二、充实了不少应用实例,特别是带有实际数据的实例,以增加教材的可读性;习题训练是学习本课程的重要环节,本次修订较大范围地扩充了习题的数量,其中大部分是基本训练题,还有一部分则是正文内容的补充。

三、在各章末添加了小结,以说明该章主要概念的背景和含义、重要方法的概率意义和统计思想,同时也指明需要掌握的重点,起到提纲挈领的作用。

四、增加了部分加“*”的内容,例如少数章节添加了附录,对正文中涉及的少数几个理论结果给出了证明。加“*”的内容不作为对学生的基本要求,而是供对此有兴趣或有志于进一步深造的读者参考。

五、书后为读者提供了一个光盘,内容有课堂教学的辅助 PPT 课件、使用 Excel 软件求解的应用实例演示等。

在本书的修订过程中,同济大学使用过第一版教材的老师,上海财经大学的孙燕老师和扬州大学的高峻老师提供了许多宝贵的意见;本书的修订得到了高等教育出版社的王强和蒋青两位编辑的热诚帮助和大力支持,他们为本书的顺利出版付出了大量心血,他们的工作细致周到、精益求精,对提高本书质量起了重要作用。我们一并表示衷心的感谢。

通过本次修订,期待本书质量有所提高,但难免还会有不少错漏和不妥之处,敬请各位同行和广大师生批评指正。

同济大学 柴根象、蒋凤瑛、杨筱萼

2011年12月

第一版前言

概率论与数理统计,在我国高校的绝大部分工科、理科专业及管理类专业,都是一门重要的基础课程。这不仅是因为它在各个领域中的应用广泛性,而且从人才素质的全面培养来说,这门课程也是不可或缺的。例如,进入 21 世纪之后,人们可以通过各种媒体获得越来越多的统计信息,它们传递着政府部门的重要政策取向,没有良好的数理统计知识就不可能很好地把握这些统计信息的特性,并善加运用。

本书着眼于介绍概率论和数理统计中的基本概念、基本原理和基本方法,它们都是初步的,但又是基本的。强调直观性和应用背景,注重可读性,突出基本思想是本书的特点。期望能对后续课程的学习以及进一步深造有所裨益,能对随机思维能力的增强和统计素质的培养有所裨益。

本书的概率论部分是根据同济大学数学教研室主编的《高等数学》(1978 年版)的第十四章改编的,数理统计部分则完全是新编的。这里顺便提及的是,早在 1982 年我系叶润修同志已做过改编的尝试,并由高等教育出版社出版了《概率论》一书。该书在使用中很受读者欢迎,但出版近 20 年来,一直未作修改,由于叶润修同志已故世,无法对该书进行修订,这不能不说是一种遗憾。本书的出版在某种意义上也可以说是弥补这一不足。

本书的部分内容打上 * 号,一般可以不读。作为一本教材,本书在选材及编排上,充分考虑到能适应不同层次的需要,有较大的灵活性。我们建议:若只选概率论部分,大约需 36 学时,而欲使用全部内容,需 54 学时;打 * 号的内容可供工科研究生和攻读 MBA 的读者参考。

本书的编写分工如下:第 1—3 章、8—12 章由柴根象同志执笔,第 4—6 章由蒋凤瑛同志执笔,第 7 章由梁汉营同志执笔,习题及解答由蒋凤瑛和杨筱茵同志执笔,最后由柴根象同志统稿、定稿。

本书的出版得到高等教育出版社徐刚、张忠月两位同志的大力支持;天津大学齐植兰同志仔细地审阅了本书的初稿,提出了许多宝贵的意见,这对提高本书质量起了重要作用;在本书的酝酿过程中,我系的郭镜明、徐建平同志做了大量的协调工作,推动了本书的写作。此外,我们的研究生孙燕、吴月琴为本书手稿的打印付出了辛勤的劳动,特在此一并表示由衷的谢意。

由于作者学识和阅历所限,书中不当和疏漏之处在所难免,敬请各位同行和读者不吝赐教。

编者

2003 年 1 月

目 录

第一章 随机事件	1
第一节 样本空间和随机事件	1
第二节 事件关系和运算	3
小结	4
习题一	5
第二章 事件的概率	7
第一节 概率的概念	7
第二节 古典概型	8
第三节 几何概型	11
第四节 概率的公理化定义	12
附录	14
小结	16
习题二	17
第三章 条件概率与事件的独立性	19
第一节 条件概率	19
第二节 全概率公式	21
第三节 贝叶斯公式	23
第四节 事件的独立性	25
第五节 伯努利试验和二项概率	29
第六节 主观概率	30
小结	31
习题三	31
第四章 随机变量及其分布	35
第一节 随机变量及分布函数	35
第二节 离散型随机变量	38
第三节 连续型随机变量	44
小结	52
习题四	53
第五章 二维随机变量及其分布	56
第一节 二维随机变量及分布函数	56
第二节 二维离散型随机变量	57
第三节 二维连续型随机变量	59

第四节	边缘分布	61
第五节	随机变量的独立性	65
*第六节	条件分布	67
小结	70
习题五	71
第六章	随机变量的函数及其分布	74
第一节	一维随机变量的函数及其分布	74
第二节	多元随机变量的函数的分布	79
小结	85
习题六	86
第七章	随机变量的数字特征	89
第一节	数学期望与中位数	89
第二节	方差和标准差	97
第三节	协方差和相关系数	99
*第四节	切比雪夫不等式及大数律	103
第五节	中心极限定理	105
小结	107
习题七	109
第八章	统计量和抽样分布	113
第一节	统计与统计学	113
第二节	统计量	118
第三节	抽样分布	121
附录	127
小结	128
习题八	128
第九章	点估计	130
第一节	点估计问题	130
第二节	估计方法	132
第三节	点估计的优良性	137
小结	140
习题九	140
第十章	区间估计	143
第一节	置信区间	143
第二节	正态总体下的置信区间	145
*第三节	抽样推断	148
小结	151
习题十	152
第十一章	假设检验	154

第一节	检验的基本原理	154
第二节	显著性水平检验法与正态总体检验	158
第三节	拟合优度检验	163
小结	168
习题十一	169
第十二章	一元线性回归	173
第一节	若干基本概念	173
第二节	一元线性回归的检验和置信推断	179
第三节	预测	186
附录	190
小结	191
习题十二	192
辅助材料	统计软件 Excel 简介	196
附表一	泊松分布表	220
附表二	标准正态分布表	223
附表三	χ^2 分布表	225
附表四	t 分布表	227
附表五	p 值表	229
部分习题参考答案	233
参考书目	252

第一章 随机事件

第一节 样本空间和随机事件

在科学研究和社会生活中,常常要在—组给定条件下进行实验或观察,例如在—定的大气压下观察对水加热,随着温度升高会发生什么现象;又如—in 闹市区的某个街口,在—个给定时间段内观察交通堵塞现象,等等,统称实验和观察为试验.在各种试验中,就试验相伴的现象的特点,又区别出—种称作随机试验的试验,如前面所举的交通堵塞试验,事先无法预知是否堵塞以及堵塞次数是多少.也就是说试验将要出现什么结果是随机的;而对于水加热这一例来说,如观察在—个标准大气压下水加热到 $100\text{ }^{\circ}\text{C}$ 会发生什么结果,其答案是预先就可以说出来的,因此没有什么随机性可言.

一般地,称具有以下两个特点的试验为随机试验:(1) 试验的所有可能结果是已知的或者是可以确定的;(2) 每次试验究竟将会发生什么结果是事先无法预知的.依此定义,上面提到的“水加热”不是随机试验,而“堵车”是随机试验.再来看—些例子:

例 1 投掷—枚均匀骰子,观察朝上面的点数,则可能结果可以是出现 1 点,2 点, \dots ,6 点中的—个.

例 2 在—批量很大的同型号产品中,混有比例为 p 的次品.从中—件接—件地随机抽取 n 次,每次抽后不—放回(简称不—放回抽取),观察抽到的 n 件产品中的次品数,则可能结果可以是次品数为 0 件,1 件, \dots , n 件中的—个.

例 3 对上海证券交易所每个交易日的综合收盘指数进行观察,则可能结果可以是 0 点到 10 000 点中的任何—个点数.

我们注意到,这三个例子都有上面提到的特点(1)、(2),因此都是随机试验,而且此外还有第三个特点,即试验可在相同条件下重复.应该说对大多数随机试验都具有第三个特点,然而也有不少例外,如

例 4 观察某地明天的天气是下雨还是晴天.

例 5 某人计划去某地旅游,观察在预定的一天能否安全抵达目的地.

很明显,这两例都是随机试验,但除非时间能够倒转,它们都是不可重复的—次性试验.从历史上看,可重复试验已经得到广泛深入的研究,有—套成熟的理论和方法.但随着社会经济的发展,特别是现代经营管理和决策分析的需要,

不可重复的随机试验的研究已引起人们的关注. 只是因篇幅限制, 本书除了个别章节外, 只研究可重复的随机试验.

对于随机试验, 我们关心的是相伴的随机现象. 为研究方便起见, 我们称: 在随机试验中, 对某些现象或某种情况的陈述为随机事件, 或简称事件. 对于指定的一次试验, 一个特定的事件可能发生, 也可能不发生, 这就是事件的随机性. 如在例 1 中, 我们关注“出现点数不大于 4”这一事件, 当试验出现结果 3 点时, 该事件发生; 而当结果出现 5 点时, 该事件不发生. 要判定一个事件是否在一次试验中发生, 必须当该次试验有了结果以后才能知晓.

称试验的每一个可能结果为样本点, 用 ω 表示. 它是一个最为基本的元素, 如例 1 中, 有 6 个样本点, 它们分别是出现 1 点到 6 点这样六个可能结果; 例 2 中有 $n+1$ 个样本点, 它们分别是: 次品数为 0 件, 1 件, \dots , n 件这样 $n+1$ 个可能结果. 又称样本点全体为样本空间, 记之为 Ω .

例 6 试给出下述随机试验的样本空间:

E_1 : 在某交通路口的某个时段, 观察机动车的流量,

E_2 : 向一个直径为 50 cm 的靶子射击, 观察弹着点的位置,

E_3 : 从含有两件次品 a_1, a_2 和三件正品 b_1, b_2, b_3 的产品中, 任取两件, 观察产品可能出现的情况.

解 试验 E_1 的可能结果为经过该路口的机动车辆数, 可以为 $0, 1, 2, \dots$. 因而

$$\Omega_1 = \{0, 1, 2, 3, \dots\}.$$

对于 E_2 , 设弹着点 ω 的坐标为 (x, y) , 则按题意应满足 $x^2 + y^2 \leq 25^2$, 因而 E_2 的样本空间为

$$\Omega_2 = \{(x, y) \mid x^2 + y^2 \leq 25^2\}.$$

E_3 的样本空间为

$$\Omega_3 = \{(a_1, a_2), (a_1, b_1), (a_1, b_2), (a_1, b_3), (a_2, b_1), \\ (a_2, b_2), (a_2, b_3), (b_1, b_2), (b_1, b_3), (b_2, b_3)\}.$$

显然, 每次试验有且只有一个含在样本空间中的试验结果发生, 事件是由试验的某些可能结果构成, 因此事件是样本空间的子集, 通常用大写字母 A, B, C, \dots 等记之. 如例 1 中, 记 $\omega_j =$ “出现点数 j ”, $j=1, 2, \dots, 6$ 为 6 个样本点, 若事件 $A = \{\text{出现点数不大于 } 4\}$, $B = \{\text{出现偶数点}\}$, 则 $\Omega = \{\omega_1, \dots, \omega_6\}$, $A = \{\omega_1, \dots, \omega_4\}$, $B = \{\omega_2, \omega_4, \omega_6\}$. 又在例 2 中, 记 $\omega_j =$ “抽到的 n 件产品中恰有 j 件次品”, $j=0, 1, \dots, n$, 若事件 $C = \{\text{次品件数不少于 } 3\}$, 则 $\Omega = \{\omega_0, \omega_1, \dots, \omega_n\}$, $C = \{\omega_3, \dots, \omega_n\}$.

依事件的定义, 样本空间 Ω 本身也是事件, 它包含了所有可能的试验结果, 因此不论在哪一次试验它都发生, 称之为必然事件. 而不含任何样本点的空集

(记之为 \emptyset),也是样本空间的子集,它在任何一次试验中都不会发生,称之为不可能事件,如例1中{掷出点数为7点}是不可能事件.

必然事件和不可能事件是随机事件的特例,尽管它们本身已无随机性可言,但在概率论中起着重要作用.

第二节 事件关系和运算

实际问题中遇到的随机事件往往是比较复杂的,在求解相关问题时,其关键的一步是将较复杂的事件分解成较简单事件的“组合”.如

例7 有两门火炮同时向一架飞机射击,考察事件 $A = \{\text{击落飞机}\}$.依常识,“击落飞机”等价于“击中驾驶员”或者“同时击中两个发动机”,因此 A 是一个较复杂的事件.如记 $B_i = \{\text{击中第}i\text{个发动机}\}, i = 1, 2, C = \{\text{击中驾驶员}\}$,相对 A 而言, B_1, B_2 及 C 都较 A 为简单.我们的问题是如何建立 A 与 B_1, B_2, C 之间的联系.

下面先讨论事件之间的关系,如果事件 A 发生必导致事件 B 发生,则称 A 蕴涵了 B ,或者说 B 包含了 A ,记为 $A \subset B$.

若 A, B 互相蕴涵,即 $A \subset B, B \subset A$ 同时成立,则称 A 与 B 相等,记作 $A = B$.

例8(例1续) 若 $A = \{\text{出现2点}\}, B = \{\text{出现偶数点}\}, C = \{\text{出现2或4或6点}\}$,则 $A \subset B, B \subset C$.

若事件 A, B 不能在一次试验中同时发生,则称 A, B 互斥或互不相容.依定义,两个事件互斥,当且仅当它们不含公共的样本点.互斥事件的一个重要特例是互为对立事件或补事件,即对任一事件 A ,称 $B = \{A \text{ 不发生}\}$ 为 A 的对立事件,或 A 的补事件,且记 $B = \bar{A}$,易知 $\bar{\bar{A}} = A$,因此当 B 为 A 的补事件时, A 也为 B 的补事件.有时也称 A, B 互补.

例9(例3续) 若 $A = \{\text{收盘指数在1500点以下}\}, B = \{\text{收盘指数在1500点或以上}\}$,则 $B = \bar{A}$.

我们也可将互斥关系推广到多个事件,称事件 A_1, A_2, \dots, A_n 是两两互斥的,如果对任意 $1 \leq i < j \leq n, A_i$ 与 A_j 是互斥的.

至此,我们已经建立了事件之间的四种关系.现在,再讨论事件之间的运算.

事件 A, B 的并事件是这样的事件: A 发生或 B 发生;等价地, A, B 至少发生一个,记之为 $A \cup B$.我们可表述为

$$A \cup B = \{A, B \text{ 至少发生一个}\},$$

图1.1(a)是 $A \cup B$ 的维恩(Venn)图,非常形象直观.

事件 A, B 的交事件是这样的事件: A, B 同时发生,记之为 $A \cap B$ 或 AB .也可以直接表述为

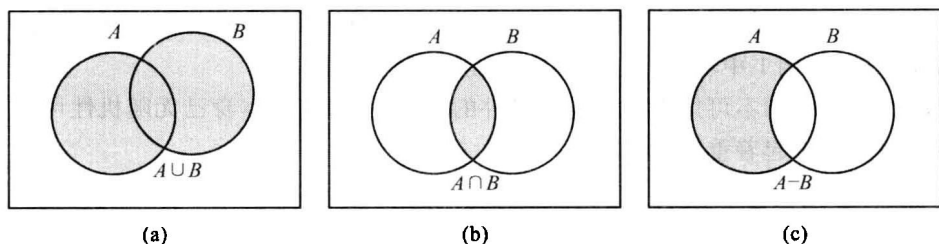


图 1.1

$$A \cap B = \{A, B \text{ 同时发生}\},$$

其维恩图见图 1.1(b).

运用事件的“补”关系及“交”运算可导出第三种运算,即事件之差:

$$A - B \equiv A\bar{B} = \{A \text{ 发生}, B \text{ 不发生}\}.$$

其维恩图见图 1.1(c).

例 10 利用事件的运算,可将事件之间的“互斥”、“互补”关系表述如下:

A, B 互斥,当且仅当 $A \cap B = \emptyset$;

A, B 互补,当且仅当 $A \cup B = \Omega, A \cap B = \emptyset$.

例 11(例 7 续) 依例 7 的记号,击落飞机这一事件 A 可分解为

$$A = B_1 B_2 \cup C.$$

如同数的四则运算有运算规则,事件的运算也遵循一定规则. 以下 A, B, C 为任意三个事件.

- (1) 交换律 $A \cup B = B \cup A, A \cap B = B \cap A$;
- (2) 结合律 $A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C),$
 $A \cap B \cap C = (A \cap B) \cap C = A \cap (B \cap C)$;
- (3) 分配律 $(A \cup B) \cap C = AC \cup BC$;
- (4) 对偶律 $\overline{A \cup B} = \bar{A} \cap \bar{B}, \overline{A \cap B} = \bar{A} \cup \bar{B}.$

例 12(例 7 续)

$$\begin{aligned} \bar{A} &= \{\text{未击落飞机}\} \\ &= (\bar{B}_1 \cup \bar{B}_2) \cap \bar{C} \\ &= (\bar{B}_1 \cap \bar{C}) \cup (\bar{B}_2 \cap \bar{C}) \\ &= \{\text{两个发动机至少有一未被击中且未击中驾驶员}\}. \end{aligned}$$

小 结

随机试验的每个可能的结果称之为样本点,而试验的所有可能结果的集合

称之为样本空间. 事件是试验中某些现象或某些情况的陈述, 它可以用样本空间的某个子集来描述. 事件的随机性表现在: 对指定的一次试验, 一个特定的事件可能发生, 也可能不发生. 事件之间的关系有: 包含(作为其特例相等)和互斥(作为其特例互补); 事件的并表示诸事件至少发生一个, 而事件的交则是诸事件同时发生, 事件的补则是该事件不发生. 事件运算的对偶律是很有用的, 需善加运用.

习 题 一

1. 用集合的形式写出下列随机试验的样本空间与随机事件 A :
 - (1) 掷两枚均匀骰子, 观察朝上面的点数, 事件 A 表示“点数之和为 7”;
 - (2) 记录某电话总机一分钟内接到的呼唤次数, 事件 A 表示“一分钟内呼唤次数不超过 3 次”;
 - (3) 从一批灯泡中随机抽取一只, 测试它的寿命, 事件 A 表示“寿命在 2 000 到 2 500 小时之间”.
2. 投掷三枚大小相同的均匀硬币, 观察它们出现的面.
 - (1) 试写出该试验的样本空间;
 - (2) 试写出下列事件所包含的样本点: $A = \{\text{至少出现一个正面}\}$, $B = \{\text{出现一正、二反}\}$, $C = \{\text{出现不多于一个正面}\}$;
 - (3) 如记 $A_i = \{\text{第 } i \text{ 枚硬币出现正面}\}$, $i = 1, 2, 3$, 试用 A_1, A_2, A_3 表示事件 A, B, C .
3. 袋中有 10 个球, 分别编有号码 1 ~ 10, 从中任取 1 球, 设 $A = \{\text{取得球的号码是偶数}\}$, $B = \{\text{取得球的号码是奇数}\}$, $C = \{\text{取得球的号码小于 5}\}$, 问下列运算表示什么事件:
 - (1) $A \cup B$; (2) AB ; (3) AC ; (4) \overline{AC} ; (5) $\overline{A} \overline{C}$; (6) $\overline{B \cup C}$; (7) $A - C$.
4. 在区间 $[0, 2]$ 上任取一数, 记 $A = \left\{x \mid \frac{1}{2} < x \leq 1\right\}$, $B = \left\{x \mid \frac{1}{4} \leq x \leq \frac{3}{2}\right\}$, 求下列事件的表达式:
 - (1) $A \cup B$; (2) \overline{AB} ; (3) $A \overline{B}$; (4) $A \cup \overline{B}$.
5. 用事件 A, B, C 的运算关系表示下列事件:
 - (1) A 出现, B, C 都不出现;
 - (2) A, B 都出现, C 不出现;
 - (3) 所有三个事件都出现;
 - (4) 三个事件中至少有一个出现;
 - (5) 三个事件都不出现;
 - (6) 不多于一个事件出现;
 - (7) 不多于两个事件出现;
 - (8) 三个事件中至少有两个出现.
6. 一批产品中有合格品和废品, 从中有放回地抽取三件产品, 设 A_i 表示事件“第 i 次抽

到废品”，试用 A_i 的运算表示下列各个事件：

- (1) 第一次、第二次中至少有一次抽到废品；
- (2) 只有第一次抽到废品；
- (3) 三次都抽到废品；
- (4) 至少有一次抽到合格品；
- (5) 只有两次抽到废品。

7. 接连进行三次射击，设 $A_i = \{\text{第 } i \text{ 次射击命中}\} (i=1, 2, 3)$ ，试用 A_1, A_2, A_3 表示下述事件：

- (1) $A = \{\text{前两次至少有一次击中目标}\}$ ；
- (2) $B = \{\text{三次射击恰好命中两次}\}$ ；
- (3) $C = \{\text{三次射击至少命中两次}\}$ ；
- (4) $D = \{\text{三次射击都未命中}\}$ 。

8. 盒中放有 a 个白球 b 个黑球，从中有放回地抽取 r 次（每次抽一个，记录其颜色，然后放回盒中，再进行下一次抽取），记 $A_i = \{\text{第 } i \text{ 次抽到白球}\} (i=1, 2, \dots, r)$ ，试用 $\{A_i\}$ 表示下述事件：

- (1) $A = \{\text{首个白球出现在第 } k \text{ 次}\}$ ；
- (2) $B = \{\text{抽到的 } r \text{ 个球同色}\}$ ，

其中 $1 \leq k \leq r$ 。

9. 试说明什么情况下，下列事件的关系式成立：

- (1) $ABC = A$ ；
- (2) $A \cup B \cup C = A$ 。

第二章 事件的概率

第一节 概率的概念

人们常会谈论一批产品的次品率是多少,或者某射手在一定条件下击中目标的命中率是多少.这表明,在日常生活中,人们已经形成一种共识:尽管随机事件有随机性,但在一次试验中发生的可能性大小是客观存在的,且是可以度量的.具体来说,如记事件 A_1 为“在一批产品中随机抽取一件是次品”,事件 A_2 为“某射手击中目标”,则 A_1, A_2 都是相应试验下的随机事件, A_1 发生的可能性大小就是次品率,而 A_2 发生的可能性大小就是命中率.例如次品率为 5%,意味着平均抽 100 件产品,其中有 5 件为次品,当然具体操作时,有时多于 5 件,有时不到 5 件,但平均为 5% 正是事件随机性中所蕴含的规律性.

我们称在随机试验中,事件 A 发生的可能性大小为事件 A 的概率,记之为 $P(A)$.

历史上,曾有许多学者做过大量的试验,例如蒲丰、皮尔逊等人先后做过掷一枚均匀硬币的试验,观察“正面朝上”这一事件(记为 A) 在 n 次试验中出现的次数,前者投掷 $n = 4\ 040$ 次, A 出现 2 048 次;后者投掷 $n = 24\ 000$ 次, A 出现 12 012 次.因此 A 出现的频率 $\left(= \frac{A \text{ 出现的次数}}{\text{试验总次数}} \right)$ 分别为 0.506 9 和 0.500 5,而且他们发现,随着试验次数增大,事件 A 出现的频率总是围绕 0.5 上下波动,且越来越接近 0.5.

概率的统计定义正是综合了大量的类似以上的试验所揭示的随机现象的规律性,定义事件的概率为频率的稳定值,依此定义,在上述中事件 A 发生的概率即为 0.5.

概率的统计定义非常直观,但在理论上的不严密也是很明显的,其实际用处是概率的近似计算,即当试验次数 n 充分大时,事件的概率可用它的频率近似.

由频率的性质可知概率满足:

(i) $0 \leq P(A) \leq 1$;

(ii) $P(\Omega) = 1$;

(iii) 若 A_1, A_2, \dots, A_n 互斥,则 $P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k)$.

这里顺便指出的是, 概率的统计定义并未为概率的计算提供任何具体规则. 在相当长的时间内, 对于具体的试验模型, 如何确定相关事件的概率, 一直是人们所关注的问题.

第二节 古典概型

本节开始介绍在概率论发展早期受到关注的两类试验模型, 其一是本节要介绍的古典概型, 其二是几何概型, 几何概型将放在下一节介绍.

若我们的试验有如下特征:

- (i) 试验的可能结果只有有限个;
- (ii) 各个可能结果出现是等可能的,

则称此试验为古典概型.

由有限性, 不妨设试验一共有 n 个可能结果, 也就是说样本点总数为 n , 而所考察的事件 A 含有其中的 k 个 (也称为有利于 A 的样本点数), 则事件 A 发生的概率为

$$P(A) = \frac{k}{n} = \frac{\text{有利于 } A \text{ 的样本点数}}{\text{样本点总数}}. \quad (1)$$

公式(1)是古典概型概率的计算公式, 具体操作涉及样本点的计数. 当涉及研究对象比较复杂时, 这种计数并非一目了然, 需要熟悉以下的基本计数原理:

设有 m 个试验, 第 1 个试验有 n_1 种可能结果, 对于第 i ($2 \leq i \leq m$) 次试验, 前 $i-1$ 个试验的每一种可能结果, 都使第 i 个试验有 n_i 种可能结果, 则 m 个试验一共有 $n_1 \times n_2 \times \cdots \times n_m$ 种可能结果. 此外, 本章末尾的附录, 介绍了排列、组合有关知识, 其熟练应用对求复杂的计数是有益的.

例 1 设有批量为 100 的同型号产品, 其中次品有 30 件. 现按以下两种方式随机抽取 2 件产品: (a) 有放回抽取, 即先任意抽取一件, 观察后放回, 再从中任取一件; (b) 不放回抽取, 即先任意抽取一件, 观察后不放入, 从剩下的产品中再任取一件. 试分别按这两种抽样方式求:

- (1) 两件都是次品的概率;
- (2) 第 1 件是次品, 第 2 件是正品的概率.

解 易知本题的试验为古典概型. 记

$$A = \{\text{两件都是次品}\}, B = \{\text{第 1 件是次品, 第 2 件是正品}\}.$$

先考虑有放回情形: 在两次抽取中每次抽取都有 100 种可能结果, 因此依计数原理样本点总数为 $n = 100^2 = 10\,000$. 事件 A 发生, 指每次是从 30 件次品中抽取的, 即每次抽取有 30 种可能结果, 因而有利于 A 的样本点数 $k = 30^2 = 900$. 于是

$$P(A) = \frac{30^2}{100^2} = 0.09.$$

同理,事件 B 发生,必须第 1 次取自 30 件次品,第 2 件取自 70 件正品,因此有利于事件 B 的样本点数为 30×70 ,所以

$$P(B) = \frac{30 \times 70}{100^2} = 0.21.$$

再考虑不放回抽取情形,此时第 1 次抽取仍然有 100 种可能结果,但第 2 次抽取只有 99 种可能结果,因而样本点总数 $n = 100 \times 99$. 同理有利于 A 的样本点数为 30×29 ,有利于 B 的为 30×70 . 因此

$$P(A) = \frac{30 \times 29}{100 \times 99} = \frac{29}{330} \approx 0.088,$$

$$P(B) = \frac{30 \times 70}{100 \times 99} = \frac{7}{33} \approx 0.21.$$

例 2 某城市电话号码升位后为八位数,且第一位为 6 或 8. 求:

- (1) 随机抽取的一个电话号码为不重复的八位数的概率;
- (2) 随机抽取的电话号码末位数是 8 的概率.

解 分别记问题(1)、(2)的事件为 A 及 B . 注意到除第一位外,其余位数可取自 0 到 9 这 10 个数中任意一个,因此有 10 种可能结果. 又第一位数只能填 6 和 8,因此只有 2 种可能结果,由此样本点总数 $n = 2 \times 10^7$.

事件 A 中的号码要求不重复,因此容易得到有利于 A 的样本点数为 $2 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3$,而有利于 B 的样本点数为 $2 \times 10^6 \times 1$,于是求得

$$P(A) = \frac{2 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3}{2 \times 10^7} = 0.01814,$$

$$P(B) = \frac{2 \times 10^6}{2 \times 10^7} = 0.1.$$

例 3(女士品茶问题) 一位常饮牛奶加茶的女士称:她能从一杯冲好的饮料中辨别出先放茶还是先放牛奶. 并且她在 10 次试验中都正确地辨别出来,问该女士的说法是否可信?

解 假设该女士说法不可信,即假定该女士纯粹是猜测. 则在此假设下每次试验的两个可能结果:牛奶+茶或茶+牛奶,是等可能的,适用古典概型. 10 次试验一共有 2^{10} 个等可能结果. 如记事件 $A = \{$ 在 10 次试验中都能正确指出放置牛奶和茶的先后次序 $\}$,则在全部 2^{10} 个样本点中 A 只含其中的一个,因而

$$P(A) = \frac{1}{2^{10}} = 0.0009766.$$

这是一个非常小的概率. 人们在日常生活中遵循一个称之为“实际推断原理”的准则:一个小概率事件在一次试验中实际是不会发生的. 依此原理 A 实际不发