

高等院校精品课系列教材

高等院校经济数学系列教材



概率论

P r o b a b i l i t y

何其祥 主编

高等院校精品课系列教材
高等院校经济数学系列教材

概 率 论

何其祥 主编

 上海财经大学出版社

图书在版编目(CIP)数据

概率论/何其祥主编. —上海:上海财经大学出版社,2016.2

高等院校精品课系列教材

高等院校经济数学系列教材

ISBN 978-7-5642-2359-5/F·2359

I. ①概… II. ①何… III. ①概率论-高等学校-教材 IV. ①O21

中国版本图书馆 CIP 数据核字(2016)第 024063 号

- ☐ 责任编辑 刘光本
- ☐ 责编电邮 lgb55@126.com
- ☐ 责编电话 021-65904890
- ☐ 封面设计 钱宇辰

GAI LÜ LUN

概 率 论

何其祥 主编

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

上海景条印刷有限公司装订

2016 年 2 月第 1 版 2016 年 2 月第 1 次印刷

787mm×1092mm 1/16 10 印张 256 千字

(习题集 9 印张 230 千字)

印数:0 001—4 000 定价:41.00 元

(本教材赠习题集, 请向售书单位索取)

前 言

概率论是研究随机现象数量规律的学科,是建立和分析带有随机因素的数学模型的基础,在经济管理、工程技术、自然科学等众多领域有着广泛的应用,尤其在近几十年中,概率统计的思想和方法被成功地用于经济、金融等领域。概率论是经济管理类学生重要的基础课程。

概率论(包括国内作者的著作和国外介绍过来的译本)已有相当的数量,其中不乏优秀教材,其内容是比较固定的。随着我国高等教育事业逐步普及化,尤其是独立学院的不断涌现,适合独立学院学生层次的教材还不多见。独立学院学生基础相对薄弱,且数学基础课程的课时数被不断压缩已是普遍现象。考虑到这些因素,根据教育部高等学校大学数学课程教学指导委员会制定的《大学数学课程教学基本要求》,我们研究和借鉴了国内外众多优秀教材的结构和内容安排,并考虑到经管类学生的需求和特点,编写了这本教材。全书共五章,由何其祥主编:第一章由何其祥编写,第二章由晁海舟编写,第三章由邹晓光、沈炳良编写,第四章由刘小锋、罗平编写,第五章由胡志明编写。全书力求用通俗的语言和方法介绍这一课程的主要思想、概念,同时保持其系统性、严谨性。

本书可读性强,既可以作为独立学院或相当层次的学生作为教学用书,又可以用作其他专业学生的参考书。为满足不同层次学生的需求,本书内容分为必学部分和选学部分(加*部分),习题分为(A)、(B)两组,(B)组略难。

在概率论课程建设和本书的编写、出版过程中,得到了上海财经大学浙江学院领导的关心以及教务处的大力支持,并得到了上海财经大学出版社的协助,一并表示感谢。

由于编者水平有限,书中难免有不妥之处和错误,恳请广大读者和同行批评指正。

何其祥

2016年2月

目 录

前言	1
第一章 随机事件与概率	1
§ 1.1 随机现象与样本空间	1
一、随机现象	1
二、样本空间	2
§ 1.2 随机事件与频率稳定性	3
一、随机事件	3
二、事件之间的关系与运算	3
三、频率与概率	6
§ 1.3 随机事件的概率	7
一、古典概型	7
二、几何概率	11
三、概率的公理化定义与性质	14
§ 1.4 条件概率、全概率公式、贝叶斯公式	16
一、条件概率	16
二、全概率公式	19
三、贝叶斯公式	20
§ 1.5 事件独立性	23
一、两个事件的独立性	23
二、多个事件的独立性	24
三、贝努利概型	27
习题一	28
第二章 随机变量及其分布	33
§ 2.1 随机变量与分布函数	33
一、随机变量的概念	33
二、随机变量的分布函数	34
§ 2.2 离散型随机变量	37

一、离散型随机变量的概率分布	37
二、离散型随机变量的分布函数	39
三、常用离散型随机变量的分布	41
§ 2.3 连续型随机变量	46
一、连续型随机变量的概率密度函数	46
二、连续型随机变量的分布函数	48
三、常用连续型随机变量的分布	49
§ 2.4 随机变量函数的分布	56
一、离散型随机变量函数的分布	56
二、连续型随机变量函数的分布	57
习题二	59
第三章 随机向量及其分布	64
§ 3.1 二维随机向量及其联合分布函数	64
一、随机向量的概念	64
二、随机向量的联合分布函数	65
三、随机向量的边际分布函数	66
§ 3.2 二维离散型随机向量	66
一、二维离散型随机向量的联合概率分布	66
二、二维离散型随机向量的边际概率分布	69
* 三、二维离散型随机向量的条件概率分布	71
§ 3.3 二维连续型随机向量	72
一、二维连续型随机向量的联合密度函数	72
二、二维连续型随机向量的边际密度函数	77
三、条件密度函数	78
四、两种常用的二维连续型随机向量的分布	78
§ 3.4 随机变量的独立性	81
一、随机变量独立性的定义	81
二、离散型随机向量独立的等价命题	81
三、连续型随机向量独立的等价命题	84
§ 3.5 二维随机向量函数的分布	86
一、二维离散型随机向量函数的分布	86
二、二维连续型随机向量函数的分布	88
三、可加性	92
习题三	96
第四章 随机变量的数字特征	102
§ 4.1 数学期望	102
一、数学期望的概念	102
二、常用离散型分布的数学期望	103

三、常用连续型分布的数学期望	104
四、随机变量函数的数学期望	105
五、二维随机向量函数的数学期望	106
六、数学期望的性质	107
§ 4.2 方差	107
一、方差的概念	108
二、常用离散型分布的方差	108
三、常用连续型分布的方差	110
四、方差的性质	111
* 五、切比雪夫不等式	111
§ 4.3 协方差和相关系数	112
一、协方差	112
二、相关系数	114
§ 4.4 随机变量的矩和 n 维随机向量	117
一、随机变量的原点矩和中心矩	117
* 二、 n 维随机向量的概念	117
* 三、 n 维随机向量的协方差矩阵	117
习题四	118
第五章 大数定律与中心极限定理	121
§ 5.1 大数定律	121
§ 5.2 中心极限定理	123
一、独立同分布场合的中心极限定理	123
二、二项分布的极限	124
习题五	127
习题参考答案	130
附录一 二项分布表	142
附录二 泊松分布表	147
附录三 标准正态分布表	149

第一章

随机事件与概率

概率论是研究随机现象的数量规律的学科,是统计学的理论基础.事件和概率是概率论中最基本的两个概念.在这一章中,我们将以深入浅出的方式介绍这些概念,主要内容包括:随机现象与随机试验、随机事件及其运算、古典概型和几何概率、概率的公理化定义、条件概率、全概率公式和贝叶斯公式、事件的独立性和贝努利概型等.

§ 1.1 随机现象与样本空间

一、随机现象

当我们观察自然界和人类社会中各种事物的变化规律时,会发现两种不同类型的现象:一种我们称为**决定性现象**,它在一定条件下必然出现某个结果.例如,在没有外力作用下,做匀速直线运动的物体必然继续做匀速直线运动;太阳必然从东方升起;任何一种生物总要经历生长、发育、衰老直至死亡等阶段;等等.这种在一定条件下必然发生的事情称为**必然事件**;反之,那种在一定条件下必然不会发生的事情称为**不可能事件**.例如,“在一个大气压下,没有加热到 100°C 的水沸腾”是不可能的.

必然事件和不可能事件,虽然表现形式有所不同,但两者的本质是一样的.必然事件的反面就是不可能事件,而不可能事件的反面就是必然事件.必然事件和不可能事件组成决定性现象,它广泛存在于自然现象和社会现象中,概率论以外的数学学科研究的就是决定性现象的数量规律.

除了决定性现象,在自然界和社会现象中还存在着与它有着本质区别的另一种现象.例如,今天无法准确地确定明天的最高或最低气温;金融领域事先无法断言将来某时刻某证交所的指数;经过某交叉路口时事先不能确定交通信号灯的颜色;同一条生产线上用同样的工艺生产出来的灯泡寿命长短也呈现出偶然性,这是由于生产过程中乃至对灯泡寿命试验的过程中种种偶然性的条件差异使每只灯泡的寿命不能事先确定;概率论中最经典的例子要数向上掷一枚硬币,结果可能是正面或反面朝上,事先无法断定.这些现象的共同特点是:在基本条件不变的情况下,一系列的试验或观察会得到各种不同的结果.换言之,就某一次的试验或观察而言,它可能会出现这种结果,也可能出现那种结果,事先无法确定,呈现出一种偶然性,这种现象称为**随机现象**(random phenomenon).概率论的研究对象就是这种随机现象的数量规律.

那么,随机现象是否普遍呢?回答是肯定的.世界上的万事万物都是相互依赖、相互影响的,某一次的观察或试验,其结果如何往往受到许多偶然因素的影响,表现形式往往是偶然的

或“随机的”，因而随机现象是普遍存在的，从而概率论的研究不但具有理论上的意义，而且具有广泛的应用价值。

事实上，概率论和以它为基础的数理统计作为现代数学的重要组成部分，由于它们在研究方法上的鲜明特色，在自然科学、社会科学、工程技术、军事科学等各个领域有着广泛的应用，特别是最近几十年中，概率论与数理统计的思想和方法在经济、管理、金融、保险等方面得到了大量、深入的应用，使这些领域的研究方法和研究范围取得了实质性的进步并得到了蓬勃的发展。正因为如此，概率论与数理统计已成为财经类专业最重要的数学基础课之一。

二、样本空间

对于随机现象，我们感兴趣的是它的结果，因此必须对它进行观察或试验，这种对随机现象的某一特征的试验或观察称为**随机试验**，简称**试验** (trial)。具体来说，称一个试验为随机试验，必须满足下述条件：

- (1) 试验可以在相同的条件下重复进行；
- (2) 试验之前能确定所有可能的结果；
- (3) 试验之前不能确定将会出现所有可能结果中的哪一个。

例如， T_1 ：向上掷一枚硬币，观察正反面出现的情况；

T_2 ：检查某企业应收账款余额；

T_3 ：观察某市过去一年发生的交通事故数。

我们把随机试验的每一个可能结果称为**样本点** (sample point)，用 ω 表示；样本点全体组成的集合称为**样本空间** (sample space)，用 Ω 表示。要认识一个随机试验，首先必须弄清楚它可能出现的各种结果，因此确定样本空间是研究随机现象的第一步。下面我们来看几个例子。

例 1-1 向上掷一枚骰子，观察朝上一面的点数，所有可能的结果为 $\Omega = \{1, 2, 3, 4, 5, 6\}$ 。如果将六个面分别涂上红、黄、蓝、绿、白、黑六种颜色，则 $\Omega = \{\text{红、黄、蓝、绿、白、黑}\}$ ，这两个样本空间表面上不同，但本质上是一致的，它们可以统一抽象地记为 $\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}$ 。

一般地，像上面只有有限个样本点的样本空间可以表示为 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ 。

样本空间的这种抽象表示，实质上是抓住了随机现象的本质，使那些表面上不同而本质上一致的各种随机现象能用统一的模型来表示。例如，只包含两个样本点的样本空间，既能作为掷硬币出现正面、反面的模型，又能用于人寿保险中“生存”与“死亡”的模型，以及描述天气时“下雨”与“不下雨”的模型，或者某公共汽车站“有人排队”与“无人排队”的模型等。

例 1-2 观察某“110”报警台在 $[0, t]$ 内来到的报警电话数，其结果显然为一非负整数，但很难确定报警数的上界，因此样本空间 $\Omega = \{0, 1, 2, \dots\}$ ，这一样本空间包含有无穷多个样本点，但它们可以按某种次序排列出来，这时我们称它有**可列个**样本点，其一般形式为 $\Omega = \{\omega_1, \omega_2, \dots\}$ 。

例 1-3 测量某一零件的长度，考察其测量结果与真正长度的误差，样本空间 Ω 可取作 $[-M, M]$ ，其中 M 为最大正误差，如果无法确定这一最大值，将 Ω 取作 $(-\infty, +\infty)$ 也无妨。这里的样本空间 $[-M, M]$ 也含有无穷多个样本点，但它无法像例 1-2 中的样本空间那样将样本点一一排出，我们称这样的样本空间所包含的样本点有**不可列个**。

例 1-4 为评价某学校小学生身体的生长发育状况，需要同时测量小学生的身高、体重和胸围，在这一随机试验中，任一样本点是一个有序数组 (x, y, z) ，其中 x, y, z 分别表示被测

量小学生的身高、体重和胸围,因此样本空间为 $\Omega = \{(x, y, z) | 0 < x \leq a, 0 < y \leq b, 0 < z \leq c\}$, 这里的 a, b, c 分别表示该校学生身高、体重和胸围的最大值. 这一样本空间所包含的样本点显然也有不可列个,但例 1-3 是一维的,而这里是高维的.

从以上这些例子不难发现,随着所讨论随机试验的不同,相应的样本空间可能很简单,也可能较复杂. 在今后的讨论中,我们一般假定样本空间是预先给定的,这种必要的抽象使我们能更好地抓住随机现象的本质,得到的结果也能得到广泛的应用.

§ 1.2 随机事件与频率稳定性

一、随机事件

对于随机现象,我们除了关心随机试验中某一结果是否会出现,往往还会关心试验的结果是否具有某一特征,称试验结果具有某一特征为**随机事件**,简称**事件**(event). 习惯上,用 A, B, C 等大写拉丁字母表示随机事件.

例如在例 1-1 中,样本空间 $\Omega = \{1, 2, 3, 4, 5, 6\}$, 如果以 A 表示“得到的为奇数”,显然 $A = \{1, 3, 5\}$, B 表示“得到的点数不大于 4”,则 $B = \{1, 2, 3, 4\}$. 这里的 A, B 均有一定的特征,为随机事件,它们都是由 Ω 中的若干个样本点所构成的. 当然,像“出现的点数为 1”“出现的点数为 5”等也是随机事件,它们只包含单个样本点. 又如在考察测量误差的例 1-3 中,样本空间 $\Omega = [-M, M]$, 如果以 A 表示“测量结果与真正长度的误差绝对值不大于 0.5 个单位”,则 $A = [-0.5, 0.5]$.

由此可见,随机事件是由若干个样本点组成的集合,或者说是样本空间的某个子集. 称某个随机事件 A 发生,当且仅当该事件所包含的某个样本点出现.

由于随机事件是样本空间的子集,所以样本空间 Ω 本身也可以看作一个事件. 由于在任一次试验中总有 Ω 中的某一样本点出现,也就是说 Ω 总发生,所以 Ω 就是必然事件. 其次,不包含有任何样本点的空集 $\emptyset$ 也可以作为一个事件,由于在任一次试验中总不可能有 $\emptyset$ 中的某一样本点出现,即 $\emptyset$ 总不发生,因此 $\emptyset$ 就是不可能事件. 当然,必然事件和不可能事件均属决定性现象,但我们宁可把它们作为随机现象的两个极端来处理,这样既是必需的也是方便的.

二、事件之间的关系与运算

给定一个样本空间,显然可以定义不止一个随机事件,分析这些事件之间的相互关系,不仅有助于认识事物的本质,而且可以通过对简单事件规律的研究去推算复杂事件的规律. 下面我们介绍事件之间的相互关系和运算. 在下面的叙述中,如果没有特别声明,均认为样本空间 Ω 已经给定,而 $A, B, A_i (i=1, 2, \dots)$ 等均为事件.

1. 事件之间的关系

(1) 包含关系

若事件 A 的每一个样本点都属于事件 B , 则称 A 包含于 B , 或称 B 包含 A , 记为 $A \subset B$ 或 $B \supset A$, 这时事件 A 的发生必然导致事件 B 的发生.

在例 1-2 中,若以 A 表示事件“ $[0, t]$ 内来到的报警电话不少于 1 500 次”,而 B 表示事件“ $[0, t]$ 内来到的报警电话数不少于 1 200 次”,那么 $A = \{1\,500, 1\,501, \dots\}$,

$B = \{1\ 200, 1\ 201, \dots\}$, 容易看出 A 的每一个样本点都属于事件 B , 即 $A \subset B$, 此时 A 的发生必然导致事件 B 的发生. 显然, 对任一事件 A , 总有 $\emptyset \subset A \subset \Omega$.

(2) 相等事件

对事件 A 与 B , 如果同时成立 $A \subset B$ 和 $B \subset A$, 则称事件 A 与事件 B 是等价的或 A 等于 B , 记为 $A = B$. 实际上, 此时 A 与 B 表示同一个事件, 它们所包含的样本点完全相同.

(3) 不相容事件

若在任何一次试验中, 事件 A 与 B 都不可能同时发生, 则称事件 A 与 B 互不相容, 也称 A 与 B 为互斥事件. 类似地, 若 n 个事件 $A_1, A_2, \dots, A_n$ 两两互不相容, 则称这 n 个事件互不相容. 更进一步, 若可列个事件 $A_1, A_2, \dots, A_n, \dots$ 两两互不相容, 则称 $A_1, A_2, \dots, A_n, \dots$ 互不相容.

2. 事件之间的运算

(1) 交(积)

由同时属于事件 A 与事件 B 的样本点组成的集合称为事件 A 与 B 的交或积, 记为 $A \cap B$ 或 AB , 事件 AB 表示事件 A 与事件 B 同时发生.

两个事件的交运算可以容易地推广到 n 个事件以至于可列个事件的场合. 事件 $A_1, A_2, \dots, A_n$ 的交事件记为 $A_1 \cap A_2 \cap \dots \cap A_n$ 或 $\bigcap_{i=1}^n A_i$, 它由同时属于 $A_1, A_2, \dots, A_n$ 的样本点组成, 表示 $A_1, A_2, \dots, A_n$ 同时发生. 在可列个事件的场合, 我们定义 $\bigcap_{i=1}^{\infty} A_i \triangleq \lim_{n \rightarrow \infty} \bigcap_{i=1}^n A_i$.

显然, 事件 A 与 B 互不相容即 $AB = \emptyset$. n 个事件 $A_1, A_2, \dots, A_n$ 互不相容, 即 $A_i A_j = \emptyset$, $i \neq j, i, j = 1, 2, \dots, n$. 可列个事件 $A_1, A_2, \dots, A_n, \dots$ 互不相容, 即 $A_i A_j = \emptyset, i \neq j, i, j = 1, 2, \dots$.

(2) 并

由至少属于事件 A 与 B 中的一个的样本点组成的集合称为事件 A 与 B 的并, 记为 $A \cup B$, 事件 $A \cup B$ 表示事件 A 与 B 中至少有一个发生.

如果事件 A 与 B 互不相容, 则称它们的并为和, 记为 $A + B$.

两个事件的并运算可以容易地推广到 n 个事件以至于可列个事件的场合. 事件 $A_1, A_2, \dots, A_n$ 的并事件记为 $A_1 \cup A_2 \cup \dots \cup A_n$ 或 $\bigcup_{i=1}^n A_i$, 它由至少属于 $A_1, A_2, \dots, A_n$ 中的一个的样本点组成, 表示 $A_1, A_2, \dots, A_n$ 中至少有一个发生. 同样, 如果事件 $A_1, A_2, \dots, A_n$ 互不相容, 则称它们的并为和, 记为 $A_1 + A_2 + \dots + A_n$ 或 $\sum_{i=1}^n A_i$.

对于可列个事件的场合, 我们定义 $\bigcup_{i=1}^{\infty} A_i \triangleq \lim_{n \rightarrow \infty} \bigcup_{i=1}^n A_i$. 进一步, 若这可列个事件互不相容, 则称它们的并为和, 记为 $\sum_{i=1}^{\infty} A_i$.

(3) 逆

对于事件 A , 由所有不属于 A 的样本点组成的集合称为事件 A 的逆事件或对立事件, 记为 $\bar{A}$, $\bar{A}$ 表示事件 A 不发生. 例如, 若 A 表示事件“ $[0, t]$ 内来到的报警电话数为偶数或 0”, 则 $\bar{A}$ 表示“ $[0, t]$ 内来到的报警电话数为奇数”. 显然, 若 $\bar{A}$ 是 A 的对立事件, 则 A 也是 $\bar{A}$ 的对立事件, 即 $\overline{\bar{A}} = A$. 必然事件和不可能事件互为对立事件.

(4) 差

由属于事件 A 而不属于 B 的样本点组成的集合称为事件 A 与 B 的差, 记为 $A-B$, 事件 $A-B$ 表示事件 A 发生而事件 B 不发生, 显然 $A-B=A\bar{B}=A-AB$.

在进行事件的运算时, 我们做如下顺序的约定: 首先进行逆的运算, 再进行交的运算, 最后进行并或差的运算.

由于事件是通过集合来定义的, 所以上面介绍的事件之间的关系与运算, 同相应集合之间的关系与运算非常相似. 因此, 我们可以借助集合论的知识和方法来帮助理解事件之间的关系与运算, 比如可以用下面直观的维恩(Venn)图来描述上面介绍的这些关系与运算. 另一方面, 应学会用概率论的观点来解释这些关系与运算(见图 1-1).

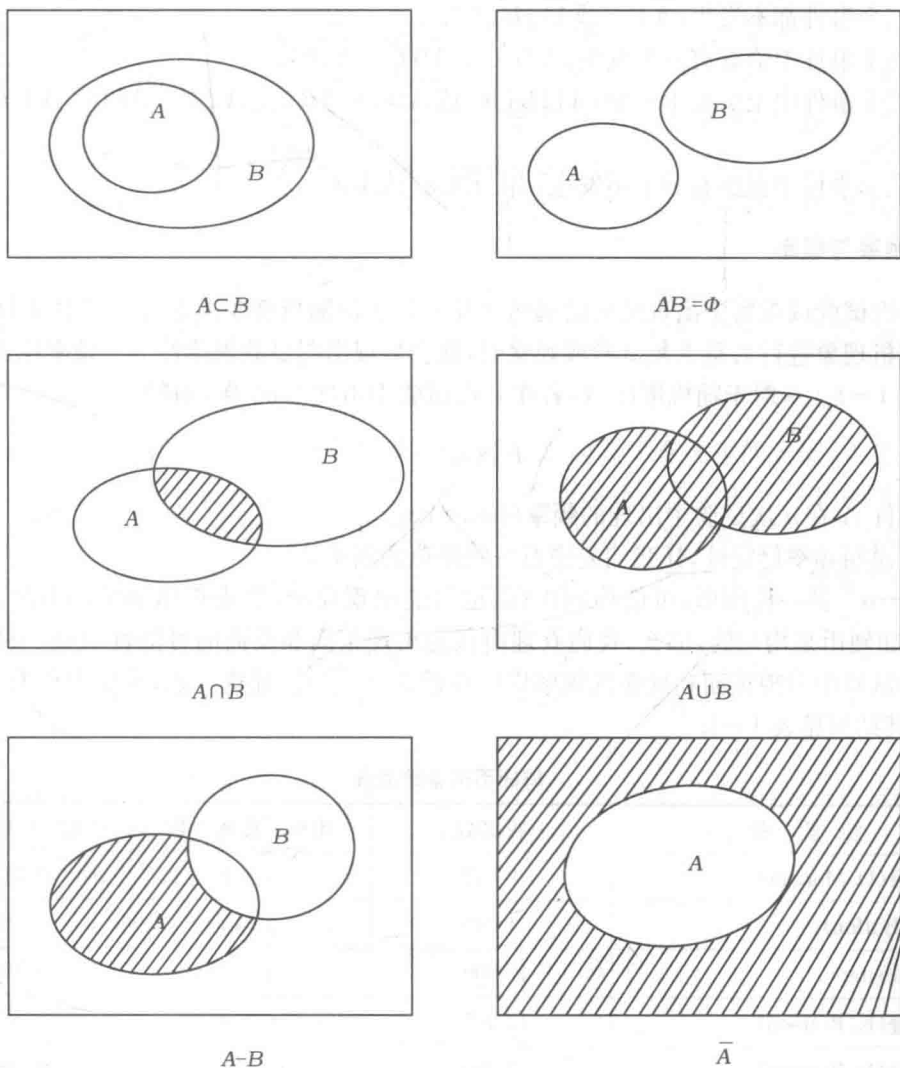


图 1-1 事件之间的关系与运算

对于事件之间的运算, 成立以下法则:

- (1) 交换律: $A \cup B = B \cup A$, $A \cap B = B \cap A$;
- (2) 结合律: $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$;

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

(4)德莫根(De Morgan)定理: $\overline{A \cup B} = \overline{A} \cap \overline{B}$, $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

上述运算法则可以推广到多个事件甚至可列个事件. 比如对于德莫根定理, 我们有

$$\begin{aligned} \overline{\bigcup_{i=1}^n A_i} &= \bigcap_{i=1}^n \overline{A_i}; & \overline{\bigcap_{i=1}^n A_i} &= \bigcup_{i=1}^n \overline{A_i}; \\ \overline{\bigcup_{i=1}^{\infty} A_i} &= \bigcap_{i=1}^{\infty} \overline{A_i}; & \overline{\bigcap_{i=1}^{\infty} A_i} &= \bigcup_{i=1}^{\infty} \overline{A_i}. \end{aligned}$$

例 1-5 设 A, B, C 为三个事件, 利用它们表示下列事件:

(1) A 发生而 B, C 都不发生: $A \overline{B} \overline{C}$ 或 $A - B - C$ 或 $A - (B \cup C)$;

(2) 三个事件都不发生: $\overline{A} \overline{B} \overline{C}$ 或 $\overline{A \cup B \cup C}$;

(3) 三个事件中恰好有一个发生: $A \overline{B} \overline{C} + \overline{A} B \overline{C} + \overline{A} \overline{B} C$;

(4) 三个事件中至少发生一个: $A \cup B \cup C$ 或 $ABC + AB \overline{C} + A \overline{B} C + \overline{A} BC + A \overline{B} \overline{C} + \overline{A} B \overline{C} + \overline{A} \overline{B} C$;

(5) 三个事件中至少有一个不发生: $\overline{A \cup B \cup C}$ 或 $\overline{ABC}$.

三、频率与概率

某一次试验或观察中出现结果的偶然性并不等于说随机现象是杂乱无章和无规律的, 当我们对随机现象进行大量重复试验或观察时, 就会呈现出明显的规律性——频率稳定性.

定义 1-1 对于随机事件 A , 若在 n 次试验中出现了 μ_n 次, 则称

$$F_N(A) = \frac{\mu_n}{n}$$

为随机事件 A 在 n 次试验中出现的频率(frequency).

为了说明频率稳定性, 让我们先来看一些经典的例子.

例 1-6 抛一枚硬币, 可能出现正面, 也可能出现反面, 事先作出确定的判断是不可能的. 但假如硬币是均匀的, 那么, 我们有理由认为出现正面和反面的可能性应该一样, 即在大量重复的试验中出现正面和反面的频率均应接近 50%. 为验证这一点, 历史上曾有不少人做过试验, 其结果见表 1-1.

表 1-1

掷硬币试验数据表

实 验 者	投掷数	出现正面的次数	出现正面的频率
德莫根(D. Morgan)	2 048	1 061	0.518 1
蒲丰(Buffon)	4 040	2 048	0.506 9
费勒(Feller)	10 000	4 979	0.497 9
皮尔逊(K. Pearson)	12 000	6 019	0.501 6
皮尔逊(K. Pearson)	24 000	12 012	0.500 5

例 1-7 任意指定一本英文书, 翻到任意一页, 任意指定该页中的某一行及该行中的任意一个位置, 记录所得到的结果, 大量重复这一试验, 会发现 26 个字母和其他字符(包括标点符号和空格)的使用频率相当稳定. 表 1-2 是经过大量试验后得出的.

表 1-2

字符试验数据表

字符	其他	E	T	O	A	N	I	R	S
频率	0.2	0.105	0.072	0.065 4	0.063	0.059	0.055	0.054	0.052
字符	H	D	L	C	F	U	M	D	
频率	0.047	0.035	0.029	0.023 4	0.022 5	0.022 5	0.021	0.017 5	
字符	W	G	B	U	K	X	J	Q	Z
频率	0.012	0.011	0.010 5	0.008	0.003	0.002	0.001	0.001	0.001

这样的例子还可以举出很多,例如多次测量同一零件的长度,其结果不一定相同,但当测量次数不断增加时,测量结果的算术平均会越来越接近于某一常数;衣服总在同样的部位破损;“水滴石穿”揭示的也是同样的道理. 这些例子表明这样一个事实:虽然在一次试验或观察中某一个随机事件 A 是否发生是偶然的,但当试验次数 n 很大时,事件 A 出现的频率总在某个固定常数附近摆动,而且一般来说, n 越大,摆动的幅度越小,这一规律称为**频率稳定性**(frequency stability).

频率稳定性是由观察对象的固有属性所决定的. 抛掷一枚硬币,可能出现正面或反面,假如硬币是均匀的,这一属性决定了当你多次抛掷硬币时,出现正面的次数总是接近总抛次数的 $\frac{1}{2}$;至于英文字母被使用的频率,虽然各个作家的写作风格各不相同,出版社也各有自己的排版方式,但字母被使用的情况总是受着文字结构、语法结构等固有规律的支配,正是这些规律决定了英文字母使用频率的稳定性. 因此,随机事件的频率稳定性表明了一个随机事件发生的可能性大小,是随机事件本身固有的客观属性,因此可以对它进行度量. 对于随机事件 A ,我们用 $P(A)$ 来刻画随机事件 A 发生的可能性大小,称 $P(A)$ 为**事件 A 发生的概率**(probability). 习惯上,称这一定义为**概率的统计定义**. 显然, $P(A)$ 即为频率稳定性中的稳定值.

对于随机现象,仅讨论它可能出现哪些结果,价值不大,而在讨论各种可能结果的同时,指出各种结果出现的可能性的**大小**才有意义. 概率概念的引进,使我们可以对随机现象做出定量的研究.

从上述可见,频率和概率既非同一概念,又有十分密切的联系. 比如,当试验次数 n 无限增大时,频率和概率之间应有某种极限关系,这正是概率论中的一大课题——**极限理论**的雏形,关于它的严格讨论将在第五章中给出.

§ 1.3 随机事件的概率

一、古典概型

1. 模型

概率论的一个基本任务,就是计算各种随机事件发生的概率. 在 § 1.2 中,我们将概率描述性地定义成事件发生的可能性大小,或者说频率的稳定值,但在许多实际问题中,为了求得某一事件的概率而进行大量重复的试验,有的是不经济的,有的是不可能的,所以我们无法根据这样的定义来计算各种事件发生的概率.

在这一节中,我们先讨论一类简单的随机试验.这类随机试验具有以下两个特征:

(1)试验的全部可能结果只有有限个,或者说只有有限个样本点,比如 n 个,即

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\};$$

(2)每个样本点 $\omega_i (i=1, 2, \dots, n)$ 出现的可能性即发生的概率相同,

$$P(\omega_1) = P(\omega_2) = \dots = P(\omega_n) = \frac{1}{n}.$$

具有这两个特征的随机现象的概率模型称为**古典概型**(classical probability model).

这类随机现象是概率论发展初期人们的主要研究对象,许多最初的概念和结果也都是对它做出的.甚至到了现在,古典概型在概率论中仍有一定的地位,这一方面是因为它简单而且直观,对它的讨论有助于理解概率论中的许多基本概念;另一方面,许多实际问题都可以概括为这一模型,古典概型有着较广泛的应用.

对于古典概型,它的样本空间为 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$,每个样本点出现的概率均为 $\frac{1}{n}$,若 A 为任一事件,由于它的每一个样本点的出现都可导致 A 的发生,而且每次试验不可能有两个样本点同时出现,因此它的概率 $P(A)$ 可看作 A 中各个样本点的概率之和.例如,若 $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_m}\}$, 则

$$\begin{aligned} P(A) &= P(\omega_{i_1}) + P(\omega_{i_2}) + \dots + P(\omega_{i_m}) \\ &= \underbrace{\frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n}}_m \\ &= \frac{m}{n}. \end{aligned}$$

所以在古典概型中,事件 A 的概率是一个分数,分母为样本点总数 n ,而分子是 A 所包含的样本点个数 m ,即

$$P(A) = \frac{m}{n} = \frac{A \text{ 包含的样本点数}}{\text{样本点总数}}.$$

法国数学家拉普拉斯(Laplace)在 1812 年把上式作为概率的一般定义,但因为这一定义只适合于古典概型,因此我们称它为**概率的古典定义**.

2. 排列与组合的基本公式

利用上述古典定义计算有关事件概率的过程中,经常会用到排列与组合的如下内容:

(1)乘法原理

设完成某件事须经过 m 个步骤,其中完成第 i 个步骤有 $n_i (i=1, 2, \dots, m)$ 种方法,则完成该事情共有 $n_1 \cdot n_2 \cdot \dots \cdot n_m$ 种方法.

(2)加法原理

设完成某件事有 m 种方式,而第 i 种方式有 $n_i (i=1, 2, \dots, m)$ 种方法,则完成该事情共有 $n_1 + n_2 + \dots + n_m$ 种方法.

(3)排列

(i)从 n 个不同的元素中有放回地任意取出 k 个,按照一定的次序进行排列,不同的排列总数为

$$n^k = \underbrace{n \cdot n \cdot \dots \cdot n}_k,$$

称为有重复的排列.

(ii) 从 n 个不同的元素中不放回地任意取出 k 个, 按照一定的次序进行排列, 不同的排列数为

$$P_n^k = n(n-1)(n-2)\cdots(n-k+1),$$

称这样的排列为选排列.

(iii) 在(ii)中, 若 $k=n$, 称其为全排列:

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

(4) 组合

从 n 个不同的元素中不放回地任意取出 k 个组成一组(此时不考虑元素之间的顺序), 称为组合, 不同的组合数有

$$C_n^k = \frac{P_n^k}{k!} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}.$$

3. 概率的计算

例 1-8 袋子中装有 10 个球, 其中 4 个黑球, 6 个白球, 大小、形状相同, 求下列事件的概率:

- (1) 从袋子中任取一球为白球;
- (2) 从袋子中任取两个球, 白球和黑球各一个;
- (3) 从袋子中任取 4 个球, 白球和黑球各两个.

解 (1) 将从 10 个球中任取一个的所有可能结果作为样本空间, 样本点总数 = 10, 以 A 记取得白球的事件, A 的样本点数 $m_A = 6$, 所以

$$P(A) = \frac{6}{10} = 0.6.$$

(2) 将从 10 个球中任取两个的所有可能结果作为样本空间, 所有可能的取法有 C_{10}^2 种, 以 B 表示“白球和黑球各一个”的事件, B 的样本点数 $m_B = C_4^1 \cdot C_6^1$, 所以

$$P(B) = \frac{C_4^1 \cdot C_6^1}{C_{10}^2} = \frac{8}{15}.$$

(3) 将从 10 个球中任取 4 个的所有可能结果作为样本空间, 所有可能的取法有 C_{10}^4 种, 以 C 表示“白球和黑球各两个”的事件, C 的样本点数 $m_C = C_4^2 \cdot C_6^2$, 因此

$$P(C) = \frac{C_4^2 \cdot C_6^2}{C_{10}^4} = \frac{3}{7}.$$

例 1-9 从 $0, 1, \dots, 9$ 中有放回地取 4 个数, 并按先后次序排列, 求下列事件的概率:

- (1) A_1 : 四个数字组成一四位数;
- (2) A_2 : 四个数字组成一四位偶数;
- (3) A_3 : 四个数字中 0 恰好出现两次.

解 从 $0, 1, \dots, 9$ 中有放回地取 4 个数的所有排列结果取作样本空间, 样本点总数显然为 10^4 .

(1) 四个数字组成一四位数, 只需第一位数字不为 0, 后三位数字可以任意, 因此 A_1 的样本点数为 $C_9^1 \cdot 10^3$, 所以

$$P(A_1) = \frac{C_9^1 \cdot 10^3}{10^4} = 0.9.$$

(2) 四个数字组成四位偶数, 第一位数字不能为 0, 且末位数字为 0 或偶数, 其余两个数字任意, 因此 A_2 的样本点数为 $C_9^1 \cdot 10^2 \cdot C_5^1$, 所以

$$P(A_2) = \frac{C_9^1 \cdot 10^2 \cdot C_5^1}{10^4} = 0.45.$$

(3) 四个数字中 0 恰好出现两次, 必须有某两个数字为 0, 可能的情况为 C_4^2 ; 另两个不能为 0, 有 9^2 种情况, 故 A_3 的样本点数为 $C_4^2 \cdot 9^2$, 所以

$$P(A_3) = \frac{C_4^2 \cdot 9^2}{10^4} = 0.0486.$$

例 1-10 某批产品共 N 件, 其中有 M 件次品, 无放回地从中任取 n 件产品, 恰好有 k 件次品的概率是多少?

解 将从 N 件产品中任抽 n 件产品的所有可能结果取作样本空间, 总的抽法有 C_N^n 种. 以 A 记抽取的 n 件产品中恰有 k 件次品的事件, A 的样本点数的计算: 先要从 M 件次品中抽取 k 件, 可能的抽法有 C_M^k 种, 又要从 $N-M$ 件正品中抽取 $n-k$ 件, 同理有 C_{N-M}^{n-k} 种取法, 从而随机地抽取 n 件, 恰好有 k 件次品的取法共有 $C_M^k \cdot C_{N-M}^{n-k}$ 种, 因此所求概率为

$$P(A) = \frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^n}, \quad k=0, 1, \dots, \min(M, n).$$

这一例子可以说明古典概型在产品抽样检验中的应用. 在实际产品检验中, 可以采用普查和抽查两种方式. 但许多工厂产量很高, 每天生产的产品数以万计, 对这些产品进行全面的普查通常是不可能的也是不经济的. 另外, 有的产品检验带有破坏性, 如灯泡寿命的检验、棉纱断裂强度的检验等, 因此经常采用的检验方法是抽样检验, 根据抽出来的若干件产品的情况去推断整批产品的质量情况.

例 1-11 设有 n 个互不相同的球, 每个球都以相同的概率 $\frac{1}{N}$ 落到 $N (N \geq n)$ 个格子的每一格中(假设每个格子无限深), 求:

(1) 某指定的 n 个格子中各有一球的概率;

(2) 恰有 n 个格子中各有一球的概率.

解 (1) 以 A 表示“某指定的 n 个格子中各有一球”的事件, 由于每个球可以落入 N 个格子的任一格中, 所以 n 个球在 N 个格子中的分配法共有 N^n 种, 而 A 的样本点数显然为 $n!$, 所以

$$P(A) = \frac{n!}{N^n}.$$

(2) 以 B 表示“恰有 n 个格子中各有一球”的事件, 由于 n 个格子可以任选, 共有 C_N^n 种选法, 对于每一种选定的 n 个格子, 球的分配法有 $n!$ 种, 所以

$$P(B) = \frac{C_N^n \cdot n!}{N^n} = \frac{P_N^n}{N^n}.$$

这一例子是古典概型中的典型问题, 例如概率论历史上的“生日问题”就可以借用这一模型解决. 所谓生日问题, 即要求参加集会的 n 个人生日各不相同的概率. 这里不妨假设一年有 365 天, 每人在 365 天内每天出生的概率相等, 且 $n \leq 365$ (否则生日各不相同的概率显然为 0). 若把一年的 365 天看作 $N=365$ 个格子, n 个人看作 n 个球, 则 n 个人生日各不相同的概率为

$$\frac{P_{365}^n}{365^n}.$$