

概率统计 及 随机过程

郭绍建 萧亮壮
张福渊 付丽华 编著

航空工业出版社



112

概率统计及随机过程

郭绍建 萧亮壮 编著
张福渊 付丽华

航空工业出版社

1993

(京)新登字161号

内 容 简 介

本书介绍概率论、数理统计、随机过程三部分内容。概率论(第一至六章)包括:随机事件的概率;随机变量及其分布;随机变量的函数的分布;随机变量的数字特征;大数定律和中心极限定理。数理统计(第七至十二章)包括:统计量及其分布;参数的点估计、区间估计;假设检验;方差分析;回归分析。随机过程(第十三、十四章)包括:随机过程的基本概念;几类重要的随机过程。

本书可以作为理工科大学的教材。也可作为有关专业人员和工程技术人员的参考书。全书内容丰富,由浅入深,并且配备相当数量的问题和习题,便于教学和自学。

概率统计及随机过程

郭绍建 萧亮壮 张福渊 付丽华 编著

航空工业出版社出版发行

(北京市安外小关东里14号)

—邮政编码:100029—

全国各地新华书店经售

北京马起乏印刷厂印刷

1993年7月第1版

1993年7月第1次印刷

开本:787×1092毫米 1/16 印张:17 1/8

印数:1—7000

字数:438.4千字

ISBN 7-80046-535-7/O·011

定价:9.50元

序 言

概率论以及以它为基础的数理统计、随机过程都是研究随机现象的数学学科。学科的产生与发展都出自实际的需求。早在十七世纪文艺复兴时期的欧洲，一些城市已经建立起预防不测事故的商业保险组织。概率论研究的初衷是：怎样利用大量现象的统计规律性来防范个别事件的偶然突发性。

几个世纪以来，经过众多学者潜心研究，概率论、数理统计、随机过程已经成为既互相关联又相对独立的三门严谨的分支学科。它们的应用极为广泛。例如，试验的设计与数据处理；产品的抽样检验；对气象、水文、地震、商品市场的预测预报；最优控制中对随机状态的估计等，无不用概率统计方法。同时，又派生出不少应用学科和边缘学科，如：随机微分方程；随机运筹学；排队论；时间序列分析；可靠性数学；统计物理；生物统计；数学地质，等等。

随着科学技术突飞猛进的发展和生产力大幅度的提高，在各研究领域和工程技术领域，人们愈来愈关注随机模型。因而，研究随机现象的数学理论和方法的应用愈来愈广泛，几乎渗透到了科学技术的各个领域。在我国，七十年代以前的工科大学里，很少开设概率课程，而如今，绝大多数的理工类、财经类、医学类专业都把概率统计作为一门必修课，部分专业还把随机过程作为必修或选修课。足见这门学科的重要地位。不难预见，这些来自实际需求的学科，在应用于实际方面，在学科自身的完善与发展方面，有着广阔的前景。

本书的初稿是笔者在北京航空航天大学使用的自编教材。使用多年，几经修订。笔者吸收了国内外同类书籍的长处，对全书纲目、内容、例题类型，经过集体研究商定，在内容的选取上，既注重基本概念、基本理论和方法，又注重知识性、技巧性、应用性。在内容的编排上，既注重系统性、逻辑性，又注意贯彻由浅入深、循序渐进的教学原则。第一至四章由郭绍建副教授编写，第五、六章由付丽华讲师和萧亮壮副教授编写，第七至十章由萧亮壮副教授编写，第十一至十四章由张福渊副教授编写。经过多遍审阅、修改，才成为现在这本书。

我们谨将本书奉献给读者。希望这本书能够帮助您掌握概率统计及随机过程的基础知识和基本方法。为深入学习研究，为在实践中应用，奠定一定的基础。

在成书过程中，得到董少英副研究员和王日爽教授的支持和帮助。北京理工大学研究生院王式安教授详尽地审阅了全部书稿，提出了极为切实的意见，使我们受益匪浅。特在此致以衷心的感谢！

对于书中缺点和错误之处，恳请同行和读者批评指正。

编 者

1993年5月

目 录

序 言

第一章 随机事件的概率

§1 排列与组合	(1)
§2 随机事件	(4)
§3 事件的概率	(6)
§4 条件概率	(14)
§5 全概率公式、贝叶斯公式	(16)
§6 事件的独立性	(18)
§7 几个例子	(20)
习题一	(23)

第二章 随机变量及其分布

§1 随机变量	(27)
§2 分布函数	(28)
§3 离散型随机变量及其概率分布	(30)
§4 连续型随机变量及其概率密度	(34)
§5 正态分布	(37)
§6 几个例子	(39)
习题二	(42)

第三章 二维随机变量

§1 二维随机变量及其分布	(46)
§2 边沿分布	(50)
§3 条件分布	(53)
§4 相互独立的随机变量	(57)
§5 几个例子	(60)
习题三	(69)

第四章 随机变量的函数的分布

§1 离散型随机变量的函数的分布	(72)
§2 一维连续型随机变量的函数的分布	(74)
§3 二维连续型随机变量的函数的分布	(77)
§4 几个例子	(83)

习题四	(88)
-----	--------

第五章 随机变量的数字特征

§1 数学期望	(92)
§2 方差	(98)
§3 协方差和相关系数	(104)
§4 矩、协方差矩阵	(108)
§5 契比雪夫不等式	(110)
§6 几个例子	(111)
习题五	(114)

第六章 大数定律和中心极限定理

§1 大数定律	(117)
§2 中心极限定理	(118)
习题六	(120)

第七章 统计量及其分布

§1 总体与样本	(121)
§2 样本矩和统计量	(122)
§3 统计量的分布	(123)
习题七	(130)

第八章 参数的点估计

§1 点估计的方法	(132)
§2 点估计量的优良性	(136)
习题八	(139)

第九章 区间估计

§1 置信区间	(141)
§2 正态总体均值和方差的区间估计	(142)
§3 二正态总体均值差和方差比的区间估计	(146)
习题九	(148)

第十章 假设检验

§1 假设检验	(149)
§2 正态总体均值的假设检验	(150)
§3 正态总体方差的假设检验	(157)
§4 总体分布的假设检验	(161)
习题十	(166)

第十一章 方差分析	
§1 单因素方差分析	(168)
§2 双因素方差分析	(176)
§3 有交互作用的双因素方差分析	(181)
习题十一	(185)
第十二章 回归分析	
§1 一元线性回归	(187)
* §2 预测与控制	(196)
§3 可线性化的非线性回归	(200)
习题十二	(202)
第十三章 随机过程的基本概念	
§1 随机过程的定义	(204)
§2 随机过程的概率分布	(206)
§3 随机过程的数字特征	(208)
习题十三	(211)
第十四章 几类重要的随机过程	
§1 正态过程	(213)
§2 独立增量过程	(213)
§3 马氏链	(217)
§4 齐次马氏链的概率分布	(220)
* §5 马尔可夫过程的一般概念	(223)
§6 平稳过程	(225)
§7 平稳过程的相关函数与谱密度	(228)
§8 各态历经性	(232)
习题十四	(235)
习题答案	(239)
附表一 泊松分布表	(252)
附表二 标准正态分布表	(254)
附表三 t 分布表	(256)
附表四 χ^2 分布表	(257)
附表五 F 分布表	(259)
附表六 相关系数临界值 (r_α) 表	(267)
参考文献	(268)

第一章 随机事件的概率

在学习概率论时，需要用到排列与组合的有关知识。在本书开始之际，首先作一些简单介绍。

§1 排列与组合

一、两个基本原理

加法原理：若完成某件事有 n 类不同办法，而第 i 类办法中又有 n_i 种不同方法($i=1, 2, \dots, n$)，则完成这件事共有

$$N = \sum_{i=1}^n n_i \quad (1.1)$$

种不同的方法。

例1 用红、黄、绿色三面旗子来发信号，共能表示多少种不同信号？

解 第一类办法是用一面旗子发信号，包括：红旗、黄旗、绿旗3种信号。第二类办法是用两面旗子发信号，包括：红黄（指先红旗后黄旗）、黄红（指：先黄旗后红旗。下同。）、红绿、绿红、黄绿、绿黄等6种信号。第三类办法是用三面旗子发信号，包括：红黄绿、红绿黄、黄红绿、黄绿红、绿红黄、绿黄红等6种信号。由加法原理，三面不同颜色的旗子表示的不同信号的数目为

$$3 + 6 + 6 = 15 \text{ (种)}$$

乘法原理：若完成某件事需要 n 个不同步骤，而第 i 步有 n_i 种不同方法($i=1, 2, \dots, n$)，则完成此事共有

$$M = \prod_{i=1}^n n_i \quad (1.2)$$

种不同的方法。

例2 从甲村到乙村有两条路径，乙村到丙村有三条路径，丙村到丁村有四条路径，甲村到丙村、丁村都无直接路径，乙村到丁村也无直接路径。问：从甲村到丁村共有多少不同路径？

解 从甲村到丁村必须经过三个步骤：甲到乙、乙到丙、丙到丁。每一步骤的路径分别为2、3、4条。由乘法原理，甲到丁共有

$$2 \times 3 \times 4 = 24$$

条不同路径。

二、排列

定义1 从 n 个不同元素中任意取出 k 个($1 \leq k \leq n$)按任意顺序排成一列，称为排列。

当 $k < n$ 时的排列, 称为选排列; 当 $k = n$ 时的排列, 称为全排列。

从 n 个不同元素中取 k 个的所有不同排列的数目, 称为排列数, 记为 A_n^k 。 n 个元素的全排列数记为 $P_n = A_n^n$ 。

一个排列可以看作在 k 个有次序的位置上各放一个元素构成。第一个位置上的元素有 n 种不同的取法; 在它取定之后, 第二个位置上的元素只有 $n-1$ 种不同的取法; 前两个元素取定后, 第三个位置上的元素只有 $n-2$ 种不同取法; 依此类推, 第 k 个位置上的元素只有 $n-k+1$ 种不同取法。故由乘法原理知, 从 n 个不同元素中取 k 个的排列数

$$A_n^k = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}, k=1, 2, \dots, n \quad (\text{其中规定 } 0! = 1)$$

(1.3)

n 个元素的全排列数

$$P_n = A_n^n = n!$$

例3 在一次文艺晚会上, 预计演出6个歌舞节目和4个相声节目。问编排节目演出次序的方案有多少? 若规定相声节目不能接连演出, 节目次序的方案又有多少?

解 编排10个节目的演出次序是一个全排列问题。共有方案

$$P_{10} = 10! = 3628800 \text{ (种)}$$

若规定相声节目不能接连演出, 节目次序的编排可分两步来实现。首先将6个歌舞节目作全排列, 于是有

$$P_6 = 6! = 720$$

种不同方案, 然后将4个相声节目分别安排在开头、末尾或相邻两个歌舞节目之间, 其安排方案有

$$A_7^4 = 7 \times 6 \times 5 \times 4 = 840 \text{ (种)}$$

由乘法原理, 相声节目不能接连演出时, 10个节目演出次序的方案共有

$$P_6 \times A_7^4 = 720 \times 840 = 604800 \text{ (种)}$$

在上述排列问题中, 参加每一排列的元素是各不相同的, 即不允许重复取同一元素在一个序列中。但在实际中, 重复取同一元素来排列的情况是存在的。如电话号码和各种车辆牌照的号码就是有重复数字的排列。一般地, 从 n 个不同元素中有放回地取 k 个的排列称为可重复的排列。由乘法原理, 可重复的排列数为

$$\underbrace{n \cdot n \cdots n}_{k \text{ 个}} = n^k \quad (1.4)$$

例4 某市的电话号码是六位数字 (不是六位数), 问以字码46为首的电话号码最多有多少?

解 电话号码的前两位数字已经确定, 只要考虑后四位数字。后四位数字的每一位都可能是0, 1, 2, ..., 9这十个数字中的任何一个。从十个数字中取4个, 同一数字可重复取的排列数为 10^4 。所以, 以46为首的六位数字的电话号码最多有一万个。

三、组合

元素的排列, 是要讲究次序的。对于同一组元素, 只要次序有所不同, 就是不同的排

列。然而，有时并不需要考虑元素的排列次序，而只需要考虑以下的组合问题。

定义2 从 n 个不同元素中任意取出 k 个 ($1 \leq k \leq n$)，不管其顺序组成一组，称为从 n 个不同元素中取 k 个的组合。

从 n 个不同元素中取 k 个的所有不同的组合的数目，称为组合数。记为 C_n^k 或 $\binom{n}{k}$ 。

由于从 n 个不同元素中取 k 个进行排列可分两步来实现：首先从 n 个元素中任意取出 k 个进行组合，然后将取出的 k 个元素进行全排列。故由乘法原理有

$$A_n^k = C_n^k \cdot k!$$

从而

$$C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!} \quad (1.5)$$

上式中以 $n-k$ 代替 k 得

$$C_n^{n-k} = \frac{n!}{(n-k)! k!} = C_n^k$$

例5 袋中有6只红球，4只白球。从中一次取4只，问

(1) 共有多少种不同的取法？

(2) 至少取到2只红球的取法有多少？

解 (1) 这是个组合问题。从10只球中一次取4只，所有不同的取法有

$$C_{10}^4 = 210 \text{ (种)}$$

(2) 至少取到2只红球，即可能取到2只、3只或4只红球。由乘法原理知，恰好取到 i 只红球的取法有

$$C_6^i \cdot C_4^{4-i} \text{ (种)} \quad (i=0, 1, 2, 3, 4)$$

故由加法原理知，至少取到2只红球的取法有

$$C_6^2 C_4^2 + C_6^3 C_4^1 + C_6^4 C_4^0 = 185 \text{ (种)}$$

至少取到2只红球的取法还可以从另一角度来考虑：从所有不同的取法中，除去全部取白球和恰好取到1只红球的取法便是至少取到2只红球的取法。即有

$$C_{10}^4 - (C_4^4 + C_6^1 C_4^3) = 210 - 25 = 185 \text{ (种)}$$

例6 将 r 个球随机地逐个放入 n 个盒中 (每盒容纳球的个数不限)，设 $r \leq n$ ，问：

(1) 所有不同的放法有多少？

(2) 恰好两个盒子有球的放法有多少？

(3) 某盒 (指定的一个盒) 不多于两个球的放法有多少？

(4) 恰好一盒有两个球而其余盒不多于一个球的放法有多少？

解 (1) 每个球都可放入 n 个盒的任何一个，即有 n 种不同放法。由乘法原理， r 个球放入 n 个盒的所有不同放法有 n^r 种。

(2) 两个有球的盒子是 n 个盒子中的任意两个，有 C_n^2 种选取方法。对于选定的两个盒子，每盒都有球的放法数等于 r 个球随意放入两个盒的放法数减去全部球放入同一盒中的放法数，故恰好两个盒子有球的放法有 $C_n^2 (2^r - 2)$ 种。

(3) 某盒不多于两个球，即该盒或者没有球，或者只有一个球，或者恰好两个球。而指定的一个盒中恰有 i ($0 \leq i \leq r$)个球的放法可分两个步骤来实现：首先从 r 个球中任取 i 个放入该盒，然后把剩下 $r-i$ 个球随意放入其余的 $n-1$ 个盒中。由乘法原理知，共有

$C_i^r \times (n-1)^{i-1}$ 种不同放法。

故由加法原理得某盒不多于两球的放法有

$$(n-1)^r + C_1^r (n-1)^{r-1} + C_2^r (n-1)^{r-2} \text{ (种)}$$

(4) 首先从 n 个盒任选一个, 然后从 r 个球中任取两个球放入选定的盒中, 再把剩下的 $r-2$ 个球放入其余的 $n-1$ 个盒中使每盒至多有一个球。故由乘法原理知, 所求放法有

$$C_1^n C_2^r A_{n-1}^{r-2} \text{ (种)}$$

§ 2 随机事件

一、随机试验与随机事件

为了叙述方便, 我们把各种各样的科学实验或对某一事物的某种特性的观察统称为试验。如果一个试验在相同的条件下可以重复进行, 而且每次试验的结果事前不可预言, 那么, 称它为随机试验, 简称为试验。用字母 E 或 E_1, E_2, \dots , 表示一个试验。

例如, 投掷一颗匀称的骰子, 观察其出现的点数; 记录某电话交换台在一天内接到的呼叫次数; 在一批灯泡中任取一只, 测试它的寿命等, 都是随机试验。并分别以 E_1, E_2, E_3 表示。

在试验中可能发生, 也可能不发生的事件称为随机事件, 简称事件。如试验 E_1 中, “出现奇数点”和“出现的点数大于3”等都是随机事件; 试验 E_2 中, “接到100次呼叫”和“呼叫次数不超过10”等也是随机事件; 试验 E_3 中, “灯泡寿命超过500小时”和“灯泡寿命在100~200小时之间”等亦是随机事件。以下用字母 A, B, C 或 A_1, A_2, \dots 等表示随机事件。显然, 试验中的每一个可能结果都是一个最简单的随机事件, 称为基本事件。可见, 随机事件是由若干基本事件组成的。随机事件 A 发生当且仅当组成 A 的基本事件有一个发生。

在试验中必然会发生的事件称为必然事件, 记为 S 。不可能发生的事件称为不可能事件, 记为 ϕ 。如 E_1 中, “出现的点数不超过6”是必然事件; “出现的点数大于6”是不可能事件。必然事件和不可能事件实际上并不是随机事件, 但为了讨论方便, 也把它们当作一种特殊的随机事件。

二、样本空间

定义3 试验 E 的全部基本事件组成的集合, 称为试验 E 的样本空间, 记为 S 。

就是说, 试验 E 的基本事件是试验 E 的样本空间中的元素。基本事件又称为样本点。

如前面的试验 E_1, E_2, E_3 的样本空间分别为: $S_1 = \{1, 2, \dots, 6\}$, $S_2 = \{0, 1, 2, \dots\}$, $S_3 = \{t \mid t \geq 0\}$ 。

又如“投掷硬币”这个试验, 它的样本空间 $S = \{\text{正面向上}, \text{反面向上}\}$ 。若以 0, 1 分别表示“正面向上”和“反面向上”这两个基本事件, 则样本空间可简单地表示为 $S = \{0, 1\}$ 。实际中, 只有两种可能结果的试验是很多的, 如检查一件产品是正品或是次品; 射击目标是击中或是不中; 人的身体健康与否等等。这些试验的样本空间都可以用 $S = \{0, 1\}$ 来表示。

引入样本空间的概念之后, 随机事件便是样本空间的子集。特别地, 不可能事件 ϕ 表示

空集，而必然事件 S 表示样本空间。这样，我们就可以引用集合论的有关知识来讨论事件间的关系与运算。

三、随机事件的关系与运算

设 E 的样本空间为 S ， A 、 B 、 C 、 A_i ($i=1, 2, \dots$) 为 E 的事件。

若事件 A 发生必然导致事件 B 发生，则称事件 A 含于事件 B ，或称事件 B 包含事件 A ，记为 $A \subset B$ 或 $B \supset A$ 。若 $A \subset B$ 且 $B \subset A$ ，则称 A 与 B 相等（或称 A 与 B 等价），记为 $A=B$ 。例如，掷骰子的试验中，令 $A = \{\text{出现1点}\}$ ， $B = \{\text{出现奇数点}\}$ ， $C = \{\text{出现非偶数点}\}$ 则有

$$A \subset B, \quad B = C$$

特别地，对任意事件 A 有

$$\emptyset \subset A \subset S$$

若 C 表示“事件 A 与 B 至少有一个发生”这一事件，则称 C 为 A 与 B 之积，记为 $C=A+B$ 或 $C=A \cup B$ 。例如，试验 E_1 中，令 $A = \{1, 3, 5\}$ ， $B = \{1, 2, 3\}$ 则 $A+B = \{1, 2, 3, 5\}$ 。显然，若 $A \subset B$ ，则 $A+B=B$ 。对任意事件 A ， $A+A=A$ ， $\emptyset+A=A$ ， $A+S=S$ 。

若 C 表示“事件 A 与 B 同时发生”这一事件，则称 C 为 A 与 B 之积，记为 $C=AB$ 或 $C=A \cap B$ 。如试验 E_1 中， $A = \{1, 3, 5\}$ ， $B = \{1, 2, 3\}$ ，则 $AB = \{1, 3\}$ 。特别地，若 $A \subset B$ ，则 $AB=A$ 。对任意事件 A ， $AA=A$ ， $AS=A$ ， $\emptyset A = \emptyset$ 。

若事件 A 与 B 不能同时发生，即 $AB = \emptyset$ ，则称事件 A 与 B 是互不相容的或称 A 与 B 是互斥的。如试验 E_1 中， $A = \{1, 2\}$ ， $B = \{3, 4\}$ ，则 $AB = \emptyset$ 。显然，不可能事件 \emptyset 与任何事件 A 互不相容。如果 $A_1, A_2, \dots, A_n, \dots$ 中的任意两个事件互不相容，则称事件 $A_1, A_2, \dots, A_n, \dots$ 互不相容。特别地，若 $AB = \emptyset$ ，且 $A+B=S$ ，则称事件 A 与 B 互逆，又称 A 是 B 的对立事件（或 B 是 A 的对立事件），记为 $A = \bar{B}$ 。（或 $B = \bar{A}$ ）。如试验 E_1 中， $A = \{1, 3, 5\}$ ， $B = \{2, 4, 6\}$ ，则 A 与 B 互逆。

显然， $\bar{\bar{A}} = A$ ， $\bar{\emptyset} = S$ ， $\bar{S} = \emptyset$

若 C 表示“事件 A 发生而 B 不发生”这一事件，则称 C 为 A 与 B 之差，记为 $C=A-B$ 。如试验 E_1 中， $A = \{1, 2, 3\}$ ， $B = \{1, 3, 5\}$ ，则 $C=A-B = \{2\}$ 。特别有， $S-A = \bar{A}$ ， $A-\emptyset = A$ ， $A-B = A-AB = A\bar{B}$ 。

事件的和与积的概念可以推广到有限多个与可列无穷多个事件的情形。即 $A = \sum A_i$ 表示“事件 $A_1, A_2, \dots, A_n, \dots$ 中至少有一个发生”这一事件。 $B = \prod A_i$ 表示“事件 $A_1, A_2, \dots, A_i, \dots$ 同时发生”这一事件。

事件的关系与运算可用几何图形直观地说明。参看图1-1。

由于事件是样本空间的子集，事件之间的运算满足下列规则：

- 1) 交换律 $A+B=B+A$ ， $AB=BA$

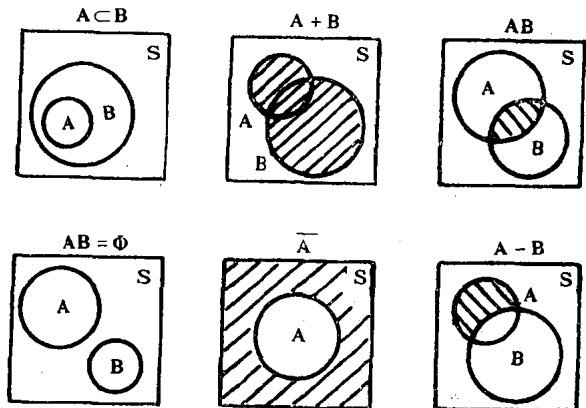


图1-1

- 2) 结合律 $(A+B)+C=A+(B+C)$
 3) 分配律 $(A+B)C=AC+BC$,
 4) 德莫根 (De Morgan) 公式: 对有限个或可列无穷多个 A_i , 恒有

$$\overline{\sum A_i} = \prod \overline{A_i}, \quad \overline{\prod A_i} = \sum \overline{A_i}$$

例1 射击目标三次, 记录射击结果。设 $A_i =$ “第 i 次击中目标”, $i=1, 2, 3$ 。试写出样本空间 S , 并用 A_1, A_2, A_3 表示下列各事件:

- (1) 只第一次击中 (B_1); (2) 只击中一次 (B_2)
 (3) 至少击中一次 (B_3); (4) 击中不多于一次 (B_4)。

解 样本空间 $S = \{A_1A_2A_3, \overline{A_1}A_2A_3, A_1\overline{A_2}A_3, A_1A_2\overline{A_3}, \overline{A_1}\overline{A_2}A_3, \overline{A_1}A_2\overline{A_3}, A_1\overline{A_2}\overline{A_3}, \overline{A_1}\overline{A_2}\overline{A_3}\}$

(1) 显然 $B_1 = A_1\overline{A_2}\overline{A_3}$ 。

(2) “只击中一次”是指: 或者仅第一次击中, 或者仅第二次击中, 或者仅第三击中。所以

$$B_2 = A_1\overline{A_2}\overline{A_3} + \overline{A_1}A_2\overline{A_3} + \overline{A_1}\overline{A_2}A_3$$

(3) “至少击中一次”是指: 可能击中一次, 也可能击中两次, 也可能三次都击中。

于是 $B_3 = A_1\overline{A_2}\overline{A_3} + \overline{A_1}A_2\overline{A_3} + \overline{A_1}\overline{A_2}A_3 + A_1A_2\overline{A_3} + \overline{A_1}A_2A_3 + A_1\overline{A_2}A_3 + A_1A_2A_3$
 或写成 $B_3 = A_1 + A_2 + A_3$

(4) “击中不多于一次”是指: 或者仅一次击中, 或者三次都没中。所以

$$B_4 = A_1\overline{A_2}\overline{A_3} + \overline{A_1}A_2\overline{A_3} + \overline{A_1}\overline{A_2}A_3 + \overline{A_1}\overline{A_2}\overline{A_3}$$

由于 B_4 的逆事件是: “至少击中两次”。所以 B_4 也可表示为

$$B_4 = \overline{A_1A_2} + \overline{A_1A_3} + \overline{A_2A_3}$$

§3 事件的概率

随机事件在一次试验中既可能发生, 也可能不发生。如果在相同的条件下, 把一个试验重复做许多次, 我们一定会发现, 某些事件发生的次数多一些, 而另一些事件发生的次数少一些。例如, 将一颗骰子重复掷100次, 毫无疑问, 事件“出现偶数点”比较事件“出现2点”发生的次数会多得多。那么, 发生次数多的事件在每次试验中发生的可能性大一些, 而发生次数少的事件在每次试验中发生的可能性小一些。问题是: 如何度量事件发生可能性的大小? 对于事件 A , 如果实数 $P(A)$ 满足: (一) 数 $P(A)$ 的大小表示事件 A 发生可能性大小; (二) $P(A)$ 是事件 A 所固有的, 不随人们主观意志而改变的一种度量。那么数 $P(A)$ 称为事件 A 的概率。它是事件 A 发生可能性的度量。

在本节中, 我们首先介绍两类最简单的概率模型, 然后逐步引出概率的一般定义。

一、概率的古典定义

设试验 E 的样本空间 S 只包含有限个基本事件, 即 $S = \{e_1, e_2, \dots, e_n\}$, 并且每个基

本事件发生的可能性相等, 即 $P(e_1) = P(e_2) = \dots = P(e_n)$, 则称这种试验为古典型随机试验, 简称古典概型。下面我们来讨论古典概型中事件 A 的概率 $P(A)$ 。

考虑一个具体的例子: 投掷一颗匀称的骰子, 观察其出现的点数。易知 $S = \{e_1, e_2, \dots, e_6\}$, 其中 e_i 表示出现 i 点, $i = 1, 2, \dots, 6$ 。由于骰子是匀称的, 所以每个基本事件 e_i 发生的可能性相同。这是个古典概型。考虑事件 $A = \{e_1, e_3, e_5\}$ 。因为事件 A 包含的基本事件的个数等于基本事件总数的一半, 并且每个基本事件发生的可能性都相等, 因此认为事件 A 发生的可能性, 即概率 $P(A) = \frac{1}{2}$ 是合理的。它恰好是 A 包含的基本事件的个数除以基本事件的总数所得的结果。

定义 1 设试验 E 的样本空间 $S = \{e_1, e_2, \dots, e_n\}$, 且 $P(e_1) = P(e_2) = \dots = P(e_n)$, E 中事件 A 包含 k 个基本事件, 则称 $P(A) = \frac{k}{n}$ 为事件 A 的概率。即事件 A 的概率等于事件 A 所包含的基本事件的个数与基本事件总数之比值。概率的这种定义称为概率的古典定义。

由概率的古典定义, 容易证明古典概率具有下列性质:

- 1) 对任意事件 A , $0 \leq P(A) \leq 1$,
- 2) $P(S) = 1$
- 3) 若事件 A_1, A_2, \dots, A_m 互不相容, 则

$$P\left(\sum_{i=1}^m A_i\right) = \sum_{i=1}^m P(A_i)$$

证 因为任一事件 A 所包含的基本事件数 k 恒满足 $0 \leq k \leq n$, 故

$$0 \leq P(A) \leq 1,$$

由于必然事件 S 包含了全部 n 个基本事件, 所以

$$P(S) = \frac{n}{n} = 1;$$

设事件 A_i 含有 k_i ($0 \leq k_i \leq n$) 个基本事件, 由定义得

$$P(A_i) = \frac{k_i}{n}, \quad i = 1, 2, \dots, m.$$

由于 A_1, A_2, \dots, A_m 互不相容, 故 $\sum_{i=1}^m A_i$ 含有 $\sum_{i=1}^m k_i$ 个不同的基本事件, 因此

$$P\left(\sum_{i=1}^m A_i\right) = \frac{\sum_{i=1}^m k_i}{n} = \sum_{i=1}^m \frac{k_i}{n} = \sum_{i=1}^m P(A_i).$$

性质 3) 称为概率的有限可加性。

例 1 袋中装有 6 个白球, 4 个红球。从中任取两个, 试求: (1) 取到两个白球的概率; (2) 取到两个相同颜色球的概率。

解 设 $A =$ “取到两个白球”,

$B =$ “取到两个同色球”

设想把 10 个球分别编上不同号码 (如 1~10 号)。从 10 个球中任取两个, 每种不同的取法为一基本事件, 于是基本事件总数为 C_{10}^2 。由于两个白球只能在袋中的 6 个白球中任取, 所

以事件 A 包含的基本事件数为 C_3^2 , 故由定义得

$$P(A) = \frac{C_3^2}{C_{10}^2} = \frac{15}{45} = \frac{1}{3}$$

令 $C =$ “取到两个红球”, 由于 “取到两个同色球” 意味着: 或者 “取到两个白球”, 或者 “取到两个红球”。因此有

$$B = A + C \text{ 且 } AC = \emptyset$$

由于两个红球只能在袋中的4个红球中任取, 因此事件 C 所含基本事件数为 C_4^2 。故由概率的有限可加性及定义得

$$P(B) = P(A + C) = P(A) + P(C) = \frac{1}{3} + \frac{C_4^2}{C_{10}^2} = \frac{7}{15}$$

例2 将5本不同的数学书和3本不同的物理书随意地摆放在书架上。试求: (1) 3本物理书没有两本放在一起的概率, (2) 3本物理书放在一起的概率。

解 8本书的每一种放法为一基本事件, 由于8本书的所有不同放法共有 P_8 种, 故基本事件总数为 $P_8 = 8!$ 。

设 $A =$ “3本物理书没有两本放在一起”, $B =$ “3本物理书放在一起”。

要使3本物理书没有两本放在一起可分两步来实现。首先将5本数学书随意摆放在书架上, 共有 P_5 种不同放法。然后将3本物理书逐一放在相邻的两本数学书之间或放在两端的位置上, 共有 A_3^3 种不同放法。故由乘法原理知, 3本物理书没有两本放在一起的所有不同放法共有 $P_5 \cdot A_3^3$ 种, 即事件 A 含有 $P_5 \cdot A_3^3$ 个基本事件。由概率定义得

$$P(A) = \frac{P_5 \cdot A_3^3}{P_8} = \frac{5}{14}$$

要使3本物理书摆放在一起, 可以把3本物理书当作一本 (如把3本物理书捆在一起), 把它和5本数学书任意摆放在书架上, 共有 P_6 种不同放法。由于3本物理书之间不同的顺序对应不同的放法。所以, 3本物理书放在一起的全部放法有 $P_6 \cdot P_3$ 种。即事件 B 含有基本事件数为 $P_6 \cdot P_3$ 。故由定义得

$$P(B) = \frac{P_6 \cdot P_3}{P_8} = \frac{3}{28}$$

二、概率的几何定义

概率的古典定义是以试验的基本事件总数有限和基本事件等可能发生为基础的。对于试验的基本事件有无穷多个的情形, 概率的古典定义显然不适用了。为了研究基本事件有无穷多个而又具有某种等可能性这样的一类随机试验, 需要用几何方法来引进概率的几何定义。

设 S 是一个可度量的有界区域 (如: 线段, 平面有界区域, 空间有界区域, 等等。), 做随机试验: 向区域 S 内投掷一质点 M , 观察质点 M 的位置。若质点 M 落在 S 内的任意子区域 A 内的可能性大小与 A 的度量 (记作 $L(A)$) 成正比, 而与 A 的位置和形状无关。则称此试验为几何型随机试验, 简称几何概型。

几何型随机试验中, 质点 M 落在 S 内的任意子区域 A 内的可能性大小与 A 的度量成正比, 而与 A 的位置和形状无关, 这就是 “等可能性” 的含义。考虑到等可能性, 并仿照古

典概率的定义，便得到几何概型中，事件 A 的概率的定义方法。

定义5 设几何概型的样本空间为 S ， A 是含于 S 内的任一随机事件，即 $A \subset S$ ，则称

$$P(A) = \frac{L(A)}{L(S)}$$

为事件 A 的概率。其中， $L(A)$ 是事件 A 的度量， $L(S)$ 是样本空间的度量。

就是说，事件 A 的概率等于事件 A 的几何度量与样本空间 S 的几何度量的比值。这样定义的概率称为几何概率。

根据几何概率的定义和几何图形的度量具有可加性，可以证明几何概率具有下列性质：

- 1) 对任意事件 A ， $0 \leq P(A) \leq 1$ ；
- 2) $P(S) = 1$ ；
- 3) 若事件 A_1, A_2, \dots, A_n 互不相容，则

$$P\left(\sum_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$$

- 4) 若事件 $A_1, A_2, \dots, A_n, \dots$ 互不相容，则

$$P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

性质4) 称为概率的可列可加性(完全可加性)。

例3 将一根长为 l 的铁丝任意截成三段，试求此三段铁丝能构成三角形的概率。

解 设铁丝两端的坐标为 0 和 l ，截断点的坐标为 x, y ，则 $0 < x < l, 0 < y < l$ 。(如图1-2) 满足不等式的点 (x, y) 构成的区域 S 是边长为 l 的正方形的内部(如图1-3)，所截三段铁丝能构成三角形的充要条件是任意两段铁丝长度之和大于第三段的长度。即有

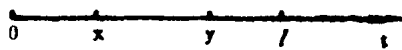


图1-2

$$\begin{cases} x + (y - x) > (l - y) \\ x + (l - y) > (y - x) \\ (y - x) + (l - y) > x \end{cases} \quad (\text{当 } x < y)$$

或

$$\begin{cases} y + (x - y) > (l - x) \\ y + (l - x) > (x - y) \\ (x - y) + (l - x) > y \end{cases} \quad (\text{当 } y < x)$$

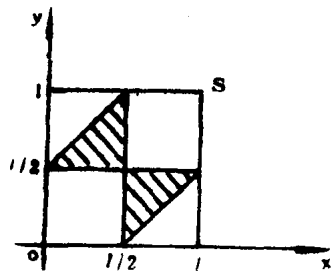


图1-3

即

$$\begin{cases} y > \frac{l}{2} \\ y - x < \frac{l}{2} \\ x < \frac{l}{2} \end{cases} \quad (x < y) \quad \text{或} \quad \begin{cases} x > \frac{l}{2} \\ x - y < \frac{l}{2} \\ y < \frac{l}{2} \end{cases} \quad (y < x)$$

满足两个不等式组之一的 (x, y) 构成的区域 A 是两个腰长为 $\frac{l}{2}$ 的等腰直角三角形的内部(图

1—3中的阴影部分)。于是问题等价于向区域 S 内任意投掷质点,求质点落入区域 A 内的概率。由几何概率的定义得,三段铁丝能构成三角形的概率为

$$p = \frac{L(A)}{L(S)} = \frac{4}{l^2} = 0.25$$

例4 (约会问题) 两人约定于8时至9时在某地会面。先到者等候20分钟,过时就离去。试求两人能见面的概率。

解 设两人到达的时间分别为8时 x 分,8时 y 分。则

$$0 \leq x \leq 60 \quad 0 \leq y \leq 60$$

满足两个不等式的点 (x, y) 构成边长为60的一个正方形区域 S (图1—4),由题意知,两人能见面的充要条件为

$$|x - y| \leq 20$$

满足此不等式的点 (x, y) 构成了 S 内的一个子区域 A (图1—4阴影部分)。于是问题等价于向区域 S 内任意投掷质点,求质点落入区域 A 的概率。故两人能见面的概率为

$$p = \frac{L(A)}{L(S)} = \frac{60^2 - 2\left(\frac{1}{2} \times 40^2\right)}{60^2} = \frac{5}{9}$$

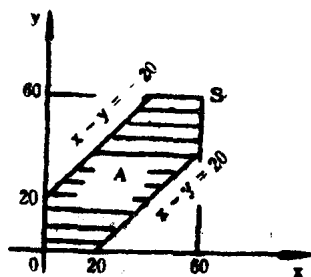


图 1—4

三、概率的统计定义

概率的古典定义和几何定义都要求试验的基本事件等可能发生。但在实际中,许多随机试验并不具有这种性质。如记录某电话交换台在8~9点这段时间内接到的呼叫次数,其全部可能结果为 $S = \{0, 1, 2, \dots\}$ 。显然,这种试验的每一个基本事件的发生不会是等可能的。因此,为了研究这样一类随机试验,就需要引进概率的新的定义方法。为此,先引进随机事件的频率的概念。

定义6 设某试验重复做了 n 次,事件 A 共发生了 n_A 次,则称比值 $\frac{n_A}{n}$ 为 n 次试验中事件 A 发生的频率,记作 $f_n(A)$ 。即

$$f_n(A) = \frac{n_A}{n}$$

频率具有下列性质:

- 1) 对任意事件 A , $0 \leq f_n(A) \leq 1$;
- 2) $f_n(S) = 1$;
- 3) 若事件 A_1, A_2, \dots, A_m 互不相容,则

$$f_n\left(\sum_{i=1}^m A_i\right) = \sum_{i=1}^m f_n(A_i)$$

(读者可从定义出发加以验证。)

事件 A 的频率 $f_n(A)$ 是随着试验次数 n 而变化的不确定的数。但是,当试验次数 n 逐渐增大时,频率 $f_n(A)$ 总是在某确定的常数 p 附近摆动,并且逐渐稳定于该常数 p 。历史上