概率论基础

程维虎　赵　旭　来向荣　编著

科　学　出　版　社

0211/95

2018

概率论基础

程维虎　赵　旭　来向荣　编著

RFID

北方工业大学图书馆

C00585850

科 学 出 版 社

北　京

内 容 简 介

本书是作者多年教学工作经验的总结与提炼. 本书介绍概率论的一般理论, 全书共分 5 章, 内容包括: 随机事件及其概率、随机变量和分布函数、随机变量的数字特征、特征函数和极限定理等, 各章后都配有适量的习题, 书后附习题答案与选解. 本书内容符合培养目标的要求, 既重视基本概念的透析、基本理论的阐述、基本方法的介绍, 又特别强调知识发生过程的探索、基本观点的提炼, 联系实际讲清概率模型, 注重基本观点的提炼, 阐述清楚概率论的思想方法, 训练学生正确解决概率问题的能力.本书体系完整, 特色鲜明, 论述严谨, 推到细致, 内容丰富且通俗易懂.

本书可作为高等理工院校与高等师范院校数学类各专业、统计学专业概率论课程的教材, 也可作为其他人员的参考书.

图书在版编目(CIP)数据

概率论基础/程维虎, 赵旭, 来向荣编著. —北京: 科学出版社, 2018.8
ISBN 978-7-03-058433-5

I. ①概… II. ①程… ②赵… ③来… III. ①概率论-高等学校-教材 IV. ①O211

中国版本图书馆 CIP 数据核字(2018) 第 170799 号

责任编辑: 王 静 / 责任校对: 张凤琴
责任印制: 吴兆东 / 封面设计: 陈 敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

北京九州迅驰传媒文化有限公司 印刷
科学出版社发行 各地新华书店经销
*
2018 年 8 月第 一 版 开本: 720 × 1000 1/16
2018 年 8 月第一次印刷 印张: 16 1/4
字数: 333 000

定价: 39.00 元

(如有印装质量问题, 我社负责调换)

前　言

本书是在来向荣、程维虎合编的《简明概率论教程》基础上编写的. 那本教材是根据 1997 年 4 月全国地方重点院校教材编写研讨会的精神, 参照 1977 年 10 月 “高等学校理科数学类教材编写大纲研讨会” 上制定的 “概率论与随机过程编写大纲”, 结合编者多年教学与科研经验编写的, 于 2001 年 3 月出版. 编写该教材的目的是用作高等理工院校或高等师范院校数学类专业概率论课程的教材, 或作为其他人员的参考书.

随着科学技术的飞速发展, 特别是网络化、大数据时代的到来, 概率论课程教学内容需要更新, 特别是要将工程技术、医疗卫生、食品安全、国防科技等交叉学科领域中的应用概率的实例引进教材, 替换那些已经过时的旧例子; 同时, 注重学生概率论思想方法的训养, 为学生日后深造及从事科学研究或其他工作打下坚实的基础. 鉴于此, 也基于作者们近二十年的教学与科研体会, 编写了本书. 本书由程维虎、赵旭和来向荣合作完成. 编写过程中, 编者广泛听取了同行专家们的意见, 尽量吸纳他人之长; 在写作上, 做了如下方面努力:

(1) 尽量保证理论的完整性与严谨性, 对课程中最基本的概念、定理及公式做较全面及严格的叙述, 尽量阐述其实际意义.

(2) 精选能够加深和理解基本概念、定理及公式的例题和习题.

(3) 对一些定理仅给出条件和结论, 证明过程或留给同学作为习题自己思考、完成, 或指出参考之处, 以保证教学重点的完成.

(4) 在保证基本内容完整的同时, 充分反映国内外最新成果.

(5) 尽量引入工程技术、医疗卫生、食品安全、国防科技等交叉学科领域的实例, 加强概率论思想方法的训练.

在本书的编写过程中, 编者得到了同事张旭、闵慧和刘玲三位老师的大力支持和帮助, 他们仔细阅读了本书初稿, 提出了许多宝贵的、建设性修改意见, 为本书润色、添彩颇多; 同校概率统计专家高旅端教授仔细审读了书稿, 斧正了稿中的错谬和瑕疵, 使编者有足够的信心将书稿交付出版.

限于编者水平, 不当乃至谬误之处在所难免, 恳请国内同行及广大读者不吝赐教.

编　者

2018 年 7 月

目　录

此为试读，需要完整PDF请访问：www.ertongbook.com

第 1 章 随机事件及其概率

1.1 基 本 概 念

1.1.1 随机现象、随机试验与事件

在人类社会的生产实践和科学实验中, 人们所观察到的现象大体上分成两类: 一类是事前可以预知结果的, 即在一定的条件满足时, 某一确定的现象必然会发生, 或根据它过去的状态, 完全可以预知其将来的发展状态. 我们称这类现象为确定性现象 (deterministic phenomenon) 或必然现象 (inevitable phenomenon). 例如, 在标准大气压下, 100°C 的纯净水必然沸腾; 带异性电荷的小球必然相互吸引; 冬天过去, 春天就会到来, 等等. 还有一类现象, 它是事前不能预知结果的, 即在相同的条件下重复进行试验时, 每次所得到的结果未必相同, 或者即使知道它过去的状态, 也不能确定其将来的状态. 我们称这类现象为不确定性现象 (uncertain phenomenon), 或偶然现象 (accidental phenomenon), 或随机现象 (random phenomenon). 例如, 掷一颗合格的骰子, 观察所掷的点数, 其结果可能是 $1, 2, \cdots, 6$ 中的一个, 但掷前无法预知所掷的结果; 又如, 某射击运动员用一支步枪在同一地点进行射击训练, 每次射击的成绩 (环数) 可能不同, 且射击前无法预知射击成绩; 再如, 观察某市未来一个月内发生交通事故的次数, 观察前也无法预知观察结果, 等等.

通常, 把对某种现象的一次观察、观测或进行一次科学实验, 统称为一个试验 (one experiment). 如果一个试验在相同的条件下可以重复进行, 且每次试验的结果是事前无法预知的, 则称这种试验为随机试验 (random experiment), 简称试验, 常用字母 E 表示. 本书以后所提到的试验均指随机试验.

进行随机试验, 尽管每次试验的结果不能完全预知, 但全部可能结果所构成的集合通常是已知的. 在上述例子中, 掷一颗骰子, 观察所掷的点数, 全部可能结果所构成的集合为 $\{1, 2, 3, 4, 5, 6\}$; 观察某市未来一个月内发生交通事故次数, 所有可能结果所构成的集合是 $\{0, 1, 2, \cdots\}$. 通常称试验所有可能结果构成的集合为该试验的样本空间 (sample space), 常用字母 Ω 表示. 样本空间的元素 (element) 称为样本点 (sample point), 常用字母 ω 表示.

称样本空间中满足一定条件的子集 (subset) 为随机事件 (random event), 简称事件 (event), 常用字母 A, B, C 或带角标的字母 A_1, A_2, B_1, B_2 等表示. 在 1.3 节中, 我们将对“一定条件”给出严格定义. 这里, 不妨把它理解为“统计规律性

(statistical regularity)”, 即：在大量重复试验中, 该子集 (事件) 的出现具有统计规律性. 这种限制, 是出于数学上的考虑. 举例来讲：如果一个试验的样本空间是区间 $[0,1]$, 将其具有长度的子集看成事件. 那么, 从实变函数论可知, 完全可以构造出子集 $Q \subset [0,1]$, 而 Q 没有长度. 这时, 我们当然不把 Q 看成事件. 实际上, 可以验证：这样的 Q 在试验中的出现没有“统计规律性”.

在试验中, 一个事件的“出现”或“发生”, 是指构成该事件的一个样本点在试验中出现. 在试验中必然出现的事件, 称为必然事件 (certain event)；不可能出现的事件, 称为不可能事件 (impossible event). 它们都无随机性. 把它们看作特殊的随机事件, 是为了讨论上的方便, 如同数学分析中把常量看作特殊的变量一样.

易见, 不可能事件不含任何样本点, 它是一个空集 (empty set, null set), 常用空集 $\varnothing$ 表示; 必然事件包含了样本空间 Ω 中的所有点, 所以, 必然事件也用样本空间的记号 Ω 来表示.

例 1.1.1 掷一颗合格的骰子, 观察所掷的点数.

若用 $\{1\}, \{2\}, \cdots, \{6\}$ 分别表示所掷的点数为 $1, 2, \cdots, 6$, 则样本空间为 $\Omega = \{1, 2, \cdots, 6\}$.

例 1.1.2 抛一枚硬币, 观察正面 (带人头像或国徽的面)、反面 (带币值的面) 出现的情况.

用 H 表示出现正面, T 表示出现反面, 则样本空间 $\Omega = \{H, T\}$.

例 1.1.3 观察某市未来一个月内发生交通事故的次数.

从理论上讲, 每个非负整数 k 都可能是试验的一个结果, 所以, 样本空间 $\Omega = \{0, 1, 2, \cdots\}$.

例 1.1.4 甲、乙相约于某日 19 时至 20 时在某地约会.

若以当日 19 时 0 分为原点, 用 x, y 分别表示甲、乙到达的时刻, 则样本空间 $\Omega = \{(x, y) \mid 0 \leqslant x,\ y \leqslant 60\}$.

例 1.1.5 对某只灯泡做实验, 观察其使用寿命.

用 x 表示灯泡的使用寿命. 从理论上讲, x 可取任意的非负实数, 所以, 样本空间 $\Omega = \{x \mid x \geqslant 0\} = [0, \infty)$.

需要说明的是, 在例 1.1.3 中, 虽然每个城市每个月内发生交通事故的次数是有限的, 不会非常大, 但一般说来, 人们理论上很难定出一个交通事故次数的有限上限. 为了方便, 我们把上限视为 ∞. 这样的处理方法在理论研究中经常被采用. 同样地, 在例 1.1.5 中我们也作了类似的处理.

1.1.2 事件的关系与运算

既然事件是一个集合, 因此, 有关事件之间的关系、运算及运算规则就可按集合间的关系、运算及运算规则来处理. 根据事件“出现”或“发生”的含义, 不难给

出事件的关系与运算的含义.

现设 Ω 是某试验 E 的样本空间, A,B,C 及 $A_k\,(k\geqslant 1)$ 都是事件, 即 Ω 的子集. 则事件之间有如下关系及运算:

(1) 若 $A\subset B$, 则称 B 包含 (contain) A. 它表示: 在试验中, 若 A 出现, 则 B 一定出现.

(2) 若 $A\subset B$, 并且 $B\subset A$, 则称 A 与 B 相等 (identical), 记成 $A=B$.

(3) 令 $A\cup B=\{\omega\,|\,\omega\in A\text{或}\omega\in B\}$, 称其为 A 与 B 的并 (union), 表示 A 与 B 至少一个出现. 类似地, 可定义 n 个事件 $A_1,A_2,\cdots,A_n$ 的并 $\bigcup\limits_{k=1}^{n}A_k$, 表示 $A_1,A_2,\cdots,A_n$ 中至少一个出现; 可列无穷个事件 $A_1,A_2,\cdots$ 的并 $\bigcup\limits_{k=1}^{\infty}A_k$, 表示 $A_1,A_2,\cdots$ 中至少一个出现.

(4) 令 $A\cap B=\{\omega\,|\,\omega\in A\text{且}\omega\in B\}$, 称其为 A 与 B 的交 (intersection), 表示 A 与 B 都出现. 类似地, 可定义 n 个事件 $A_1,A_2,\cdots,A_n$ 的交 $\bigcap\limits_{k=1}^{n}A_k$, 表示 $A_1,A_2,\cdots,A_n$ 都出现; 可列无穷个事件 $A_1,A_2,\cdots$ 的交 $\bigcap\limits_{k=1}^{\infty}A_k$, 表示 $A_1,A_2,\cdots$ 都出现.

有时也简记 $\bigcap\limits_{k=1}^{n}A_k$ 为 $A_1A_2\cdots A_n$.

若 $AB=\varnothing$, 则称 A 与 B 互斥 (mutually exclusive) 或互不相容 (incompatible), 它表示: 在试验中, 若 A 出现, 则 B 不出现; 若 B 出现, 则 A 不出现.

(5) 令 $A-B=\{\omega\,|\,\omega\in A\text{且}\omega\notin B\}$, 称其为 A 与 B 的差 (difference), 表示 A 出现, 但 B 不出现. 特别地, 称 $\Omega-A$ 为 A 的逆事件 (complement event) 或对立事件 (opposite event), 记成 $\overline{A}$.

需要注意的是, 如果 A,B 及 $A_k(k\geqslant 1)$ 是事件, 我们总假设 $\bigcup\limits_{k}A_k$, $\bigcap\limits_{k}A_k$ 和 $A-B$ 也都是事件, 其中 $\bigcup\limits_{k}A_k$ 表示 $\bigcup\limits_{k=1}^{n}A_k$ 或 $\bigcup\limits_{k=1}^{\infty}A_k$, $\bigcap\limits_{k}A_k$ 表示 $\bigcap\limits_{k=1}^{n}A_k$ 或 $\bigcap\limits_{k=1}^{\infty}A_k$.

在进行事件的运算时, 经常用到如下规则:

(1) 交换律 (commutative law): $A\cup B=B\cup A,\ AB=BA$.

(2) 结合律 (associativity law): $(A\cup B)\cup C=A\cup(B\cup C),\ (AB)C=A(BC)$.

(3) 分配律 (distributive law): $A\big(\bigcup\limits_{k}A_k\big)=\bigcup\limits_{k}(AA_k),\ A\cup\big(\bigcap\limits_{k}A_k\big)=\bigcap\limits_{k}(A\cup A_k)$.

(4) 德 · 摩根 (对偶) 律 (De Morgan's law): $\overline{\bigcup\limits_{k}A_k}=\bigcap\limits_{k}\overline{A_k},\ \overline{\bigcap\limits_{k}A_k}=\bigcup\limits_{k}\overline{A_k}$.

此外, 还应注意到 $A-B=A\overline{B},\ A=(AB)\cup(A\overline{B})$ 等.

为直观地理解上述内容, 可用图示法表示事件之间的关系与运算.

假设用平面上的一个矩形表示样本空间 Ω, 矩形内的每个点表示一个样本点, 用两个小圆分别表示事件 A 和 B, 则事件的关系与运算可用 Venn 图 1.1.1 表示,

其中 $A \subset B, A \cup B, AB, A$ 与 B 互斥, $A-B, \overline{A}$ 分别为图中阴影部分.

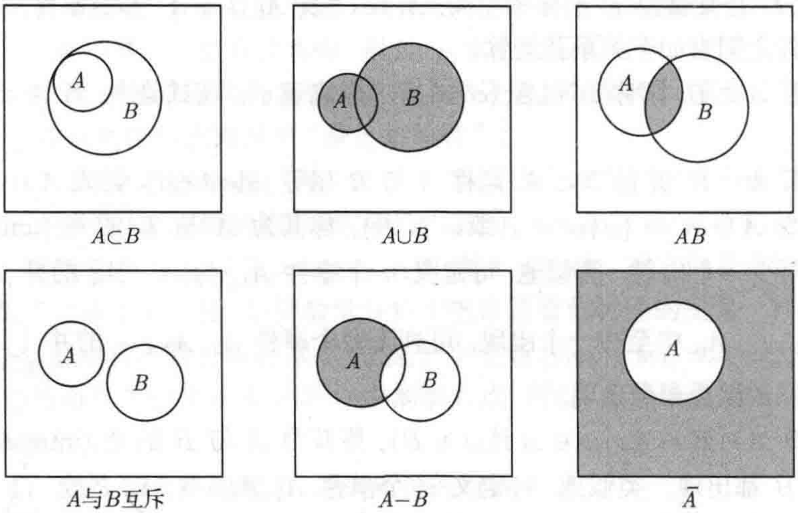

图 1.1.1　事件关系与运算的 Venn 图

1.2　几种概率模型

设 A 是试验 E 的事件. 在一次试验中, A 可能出现, 也可能不出现. 可以用一个数来刻画 A 在一次试验中出现的可能性大小, 称之为 A 的概率 (probability), 记成 $P(A)$. 下面介绍三种最基本的概率模型 (probability model).

1.2.1　古典概型

概率论起源于赌博游戏, 因此, 最先涉及的求概率问题都满足“各可能结果具有等可能性”这一假设. 例如, 在游戏中使用的骰子是匀质的立方体, 以便掷出 1 至 6 各点等可能, 保证游戏的公平. 又如, 一副纸牌, 每一张牌的形状、大小与背面花色均相同, 且在发牌前充分地洗牌, 于是, 发到每张牌也是等可能的.

定义 1.2.1　若随机试验 E 具有如下特点:

(1) 样本空间 Ω 有有限个样本点;

(2) 各样本点在试验中出现是等可能的.

则称这样的试验模型为古典概率模型, 简称古典概型 (classical probability model); 对应的概率为古典概率 (classical probability), 简称概率.

对一个古典概型, 假设试验 E 的样本空间为 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$, 事件 A 含 k 个样本点 $\{\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k}\}$, $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$, 则定义 A 的古典概率为

此为试读, 需要完整PDF请访问: www.ertongbook.com

$$P(A)=\sum_{j=1}^{k}P(\{\omega_{i_j}\})=\sum_{j=1}^{k}\frac{1}{n}=\frac{k}{n}=\frac{A\text{中样本点数}}{\Omega\text{中样本点数}}. \tag{1.2.1}$$

易知, 古典概率有以下性质:

(1) $0\leqslant P(A)\leqslant 1$, A 是试验的任意事件;

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

(3) 若 $m\geqslant 2$, 事件 $A_1,A_2,\cdots,A_m$ 两两互斥, 则有

$$P\Big(\bigcup_{i=1}^{m}A_i\Big)=\sum_{i=1}^{m}P(A_i).$$

从 (3) 易推出, 若 A 为事件, 则 $P(\overline{A})=1-P(A)$. 有时可用此公式简化计算.

例 1.2.1 抛一枚匀质硬币 3 次, 求事件 A_1,A_2 的概率, $A_1=\{$恰有一次出现正面$\}$, $A_2=\{$至少一次出现正面$\}$.

解 我们将每抛 3 次硬币看成一次试验, 并用 (x,y,z) 表示试验的结果 (样本点), 三个坐标值 x,y,z 均可为正面 H 或反面 T. 所以, 样本点总数为 $2\times2\times2=8$, 各样本点的出现是等可能的, $A_1=\{(H,T,T),(T,H,T),(T,T,H)\}$, $\overline{A_2}=\{(T,T,T)\}$. 故

$$P(A_1)=\frac{3}{8},\ \ P(A_2)=1-P(\overline{A_2})=1-\frac{1}{8}=\frac{7}{8}.$$

例 1.2.2 袋中有同型号小球 6 个, 4 个色白, 2 个色红. 从袋中任意取球两次, 每次取一个. 试分别在返回抽取和不返回抽取的情况下, 求事件 A,B,C 的概率. $A=\{$取到两个白球$\}$, $B=\{$取到两个同色球$\}$, $C=\{$取到的两个球中至少有一个是白球$\}$. 所谓返回抽取, 是指在第二次取球前, 已将第一次取到的球放回袋中; 所谓不返回抽取, 是指在第二次取球时, 第一次取到的球已不在袋中.

解 首先考虑返回抽取的情况.

我们把取两次球看成一次试验, 并用 (x,y) 表示试验的结果 (样本点), x 为第一次取到的球, y 为第二次取到的球. 则 x 和 y 均有 6 种取法, 样本点总数为 $6\times6=36$, 各样本点的出现是等可能的, 由 A 含 $4\times4=16$ 个样本点, 得

$$P(A)=\frac{16}{36}=\frac{4}{9};$$

注意到 $\overline{C}=\{$取到两个红球$\}$, 其样本点的两个坐标 x,y 各有两种取法, 所以, $\overline{C}$ 含 $2\times2=4$ 个样本点, 于是,

$$P(\overline{C})=\frac{4}{36}=\frac{1}{9},\ \ P(C)=1-P(\overline{C})=\frac{8}{9};$$

注意到 $B=A\cup\overline{C}$, 且 A 与 $\overline{C}$ 互斥, 所以,

$$P(B)=P(A)+P(\overline{C})=\frac{5}{9}.$$

其次考虑不返回抽取的情况.

仍用 (x,y) 表示样本点, x 为第一次取到的球, y 为第二次取到的球, 则 x 有 6 种取法, y 有 5 种取法, 样本点总数为 $6\times 5=30$, 且各样本点的出现是等可能的. A 的样本点的 x 有 4 种取法, y 有 3 种取法, 所以, A 含 $4\times 3=12$ 个样本点,

$$P(A)=\frac{12}{30}=\frac{2}{5};$$

仿上, $\overline{C}=${取到两个红球}, 其样本点的 x 有两种取法, y 只有一种取法, 所以, $\overline{C}$ 含 $2\times 1=2$ 个样本点,

$$P(C)=1-P(\overline{C})=1-\frac{2}{30}=\frac{14}{15};$$

进一步, 由 $B=A\cup\overline{C}$, 且 A 与 $\overline{C}$ 互斥, 得

$$P(B)=P(A)+P(\overline{C})=\frac{2}{5}+\frac{1}{15}=\frac{7}{15}.$$

例 1.2.3　设 15 件新产品中有 3 件特级品, 将这 15 件新产品平均分配给三个售货点, 求事件 A 和 B 的概率. $A=${每个售货点各分配到 1 件特级品}, $B=${3 件特级品被分配到同一个售货点}.

解　将 15 件产品分到三个点, 使每个点恰有 5 件的分法有 $15!/(5!5!5!)$ 种, 每种分法是等可能的. 把每种分法看成一个样本点, 样本点总数为 $15!/(5!5!5!)$.

把 3 件特级品分到三个点, 使每个点恰有 1 件的分法有 $3!$ 种, 对应于每种这样的分法, 把其余 12 件非特级品分到三个点, 使每个点恰有 4 件的分法有 $12!/(4!4!4!)$ 种, 所以, A 含 $3!12!/(4!4!4!)$ 个样本点,

$$P(A)=\frac{3!12!/(4!4!4!)}{15!/(5!5!5!)}=\frac{25}{91}.$$

把 3 件特级品分到同一个点的分法有 3 种, 对应于每种这样的分法, 把其余 12 件非特级品分到三个点, 使分到特级品的点恰有两件, 另两个点各恰有 5 件的分法有 $12!/(2!5!5!)$ 种, 所以, B 含 $3\times 12!/(2!5!5!)$ 个样本点,

$$P(B)=\frac{3\times 12!/(2!5!5!)}{15!/(5!5!5!)}=\frac{6}{91}.$$

上述计算中用到以下公式: 把 n 个物品分成 k 组, 其中第一组 n_1 个, 第二组 n_2 个, $\cdots$, 第 k 组 n_k 个, $n_1+n_2+\cdots+n_k=n$, 则不同的分法数为

$$\begin{pmatrix} n \\ n_1,n_2,\cdots,n_k \end{pmatrix}=\frac{n!}{n_1!n_2!\cdots n_k!}. \tag{1.2.2}$$

注意到, 当 $k=2$ 时, (1.2.2) 式就是以前学过的组合公式 (combination formula), 此时 $\begin{pmatrix} n \\ n_1, n_2 \end{pmatrix} = \dfrac{n!}{n_1!n_2!}$, 有时简记成 $\mathrm{C}_n^{n_1}$, 即 $\mathrm{C}_n^{n_1} = \dfrac{n!}{n_1!n_2!} = \dfrac{n!}{n_1!(n-n_1)!}$.

思考题 令 $C=\{$恰有两件特级品被分配到同一个售货点$\}$, 验证 $P(C)=\dfrac{60}{91}$. 注意到 $A\cup B\cup C=\Omega$, A,B,C 两两互斥.

例 1.2.4(占位问题) 将 r 个不同的球任意放入编号分别为 $1,2,\cdots,n$ 的 n 个盒中 ($r\leqslant n$), 每个球进入各盒等可能, 每个盒可容纳的球数不受限制. 求事件 A,B,C 的概率, 其中 $A=\{$指定的 r 个盒恰各有 1 个球$\}$, $B=\{$每盒至多有 1 个球$\}$, $C=\{$某指定盒恰有 m 个球$\}$.

解 因第 i 个球可放进 n 个盒中的任一个, 有 n 种放法, $i=1,2,\cdots,r$, 所以, r 个球放入 n 个盒中, 共有 n^r 种放法, 每种放法是等可能的. 把每种放法看成一个样本点, 样本点总数为 n^r. 指定的 r 个盒恰各有 1 个球的放法有 $r!$ 种, 即 A 含 $r!$ 个样本点, 故

$$P(A)=\frac{r!}{n^r};$$

事件 B 与 A 的差异仅在于“各有 1 个球的 r 个盒没有指定”, 而指定 r 个盒有球的方法数为 C_n^r, 由乘法原理可得到, B 含 $\mathrm{C}_n^r r!$ 个样本点, 故

$$P(B)=\frac{\mathrm{C}_n^r r!}{n^r}=\frac{n!}{(n-r)!n^r};$$

为数清事件 C 所含样本点个数, 可先取 m 个球放入指定盒中, 共有 C_r^m 种取法; 然后再把剩下的 $r-m$ 个球任意放入其余 $n-1$ 个盒中, 放法有 $(n-1)^{r-m}$ 种; 再根据乘法原理, 知 C 含 $\mathrm{C}_r^m(n-1)^{r-m}$ 个样本点, 得

$$P(C)=\frac{\mathrm{C}_r^m(n-1)^{r-m}}{n^r}.$$

下面给出占位问题的一个特例.

例 1.2.5(生日问题) 求任意 r 个人生日各不相同的概率.

解 将一年 365 天看成 365 个盒, r 个人去任意占这 365 个盒. “生日各不相同”不就等价于“每盒至多 1 个球”, 即上例中的事件 B 吗? 于是, 所求概率为

$$P(B)=\frac{365!}{(365-r)!365^r}=\frac{365\cdot 364\cdot\cdots\cdot(365-r+1)}{365^r}.$$

而 B 的逆事件 $\overline{B}=\{r$ 个人中至少有两人生日相同$\}$的概率为

$$1-P(B)=1-\frac{365\cdot 364\cdot\cdots\cdot(365-r+1)}{365^r}.$$

表 1.2.1 给出了若干个 $1-P(B)$ 的计算值.

表 1.2.1 r 个人中至少有两人生日相同的概率表

r	10	15	20	25	30	35	40	45	50
p	0.12	0.25	0.41	0.57	0.71	0.81	0.89	0.94	0.97

从表 1.2.1 可以看出: 在 40 人左右的人群里, “十有八九” 会出现 {两人或两人以上生日相同}这一事件.

例 1.2.6 设 N 件产品中有 M 件废品, $N-M$ 件正品, $M<N$, 从 N 件产品中任意抽取 n 件, $n \leqslant N-M$, 试分别在不返回抽取和返回抽取的情况下求事件 A 的概率. $A=\{n$ 件中恰有 k 件废品$\}$, $k=0,1,2,\cdots,l$, $l=\min(M,n)$.

解 首先考虑不返回抽取的情况.

注意到, 对 N 件产品作不返回抽取, 每次取一件, 取 n 次, 得到 N 件产品中 n 件不同的产品, 等同于从 N 件产品中一次取出这 n 件不同的产品. 从 N 件中取出 n 件有 C_N^n 种取法, 每种取法是等可能的. 把每种取法看作一个样本点, 样本点总数为 C_N^n; k 件废品从 M 件废品中取, 有 C_M^k 种取法, 其余 $n-k$ 件正品从 $N-M$ 件正品中取, 有 C_{N-M}^{n-k} 种取法, 所以, A 含 $\mathrm{C}_M^k\mathrm{C}_{N-M}^{n-k}$ 个样本点, 于是

$$P(A)=\frac{\mathrm{C}_M^k\mathrm{C}_{N-M}^{n-k}}{\mathrm{C}_N^n}. \tag{1.2.3}$$

其次考虑返回抽取的情况.

从 N 件中取 n 件, 每次皆可从 N 件中任取一件, 所以有 N^n 种取法, 每种取法是等可能的. 把每种取法看作一个样本点, 样本点总数为 N^n; k 件废品从 M 件废品中取, 有 M^k 种取法, 其余 $n-k$ 件正品从 $N-M$ 件正品中取, 有 $(N-M)^{n-k}$ 种取法, k 件废品出现在 n 次中的方式有 C_n^k 种, 所以, A 含样本点个数为 $\mathrm{C}_n^kM^k(N-M)^{n-k}$,

$$P(A)=\frac{\mathrm{C}_n^kM^k(N-M)^{n-k}}{N^n}=\mathrm{C}_n^k\left(\frac{M}{N}\right)^k\left(1-\frac{M}{N}\right)^{n-k}. \tag{1.2.4}$$

需要指出的是: 古典概率的计算不依赖于样本空间的选取. 如例 1.2.6 中不返回抽取的情形, 若考虑抽取顺序, 从 N 件中抽取 n 件的方式有 A_N^n 种, 而每种方式的出现是等可能的. 把每种方式看成一个样本点, 样本点总数为 A_N^n; k 件废品从 M 件废品中取, 有 A_M^k 种方式, 其余 $n-k$ 件正品从 $N-M$ 件正品中取, 有 A_{N-M}^{n-k} 种方式; k 件废品出现在 n 次中的方式有 C_n^k 种, 所以, A 含 $\mathrm{C}_n^k\mathrm{A}_M^k\mathrm{A}_{N-M}^{n-k}$ 个样本点,

$$P(A)=\frac{\mathrm{C}_n^k\mathrm{A}_M^k\mathrm{A}_{N-M}^{n-k}}{\mathrm{A}_N^n}=\frac{\mathrm{C}_M^k\mathrm{C}_{N-M}^{n-k}}{\mathrm{C}_N^n}.$$

这与 (1.2.3) 式的结果一致.

由两种抽样方式得到的 (1.2.3) 式称为超几何分布 (hypergeometric distribution) 的概率分布; (1.2.4) 式称为二项分布 (binomial distribution) 的概率分布. 二者都是重要的离散型概率分布, 将在第 2 章详细介绍.

思考题 用 (1.2.3), (1.2.4) 式解例 1.2.2.

1.2.2 统计概型

设 A 是试验 E 的事件, 在一定的条件下, 将试验重复 n 次, 如果 A 出现 m 次, 则称 m 为 A 在 n 次试验中出现的频数, $f_n(A)=m/n$ 为 A 在 n 次试验中出现的频率 (frequency). A 的频率表示事件出现的频繁程度, 也可用来表示在一次试验中 A 出现的可能性的大小. 问题是: A 在 n 次试验中的频率首先依赖于试验次数 n; 其次, 当 n 固定时, 若另做 n 次试验, 所得频率也可能同先前所得的频率不同. 但是, 人们发现, 当试验次数 $n\to\infty$ 时, A 的频率具有某种稳定性, 以一个常数为波动中心. 频率的这种性质, 将在第 5 章通过大数定律来阐明. 因此, 对于充分大的 n, 不妨用 A 在 n 次试验中出现的频率 $f_n(A)$ 作为 A 的概率. 这样所得到的概率称为统计概率 (statistical probability), 这种概率模型为统计概率模型 (statistical probability model), 简称统计概型.

统计概率通过试验得到, 不要求样本空间的元素有限, 也不要求样本点的出现具有等可能性. 这是一种一般的方法, 虽不能当作概率的数学定义, 但可用来验证理论和近似计算.

易知, 统计概率具有以下性质:

(1) $0\leqslant f_n(A)\leqslant 1$, A 是试验的任意事件;

(2) $f_n(\Omega)=1$;

(3) 对任意两两互斥事件 $A_1,A_2,\cdots,A_m$, 有 $f_n\left(\bigcup\limits_{i=1}^{m}A_i\right)=\sum\limits_{i=1}^{m}f_n(A_i)$.

1.2.3 几何概型

设 $n=1,2,3$, 试验 E 的样本空间 Ω 是 n 维空间中的一个有有限正度量的集合, 每个样本点在试验中的出现是等可能的. 设事件 A 是 Ω 的有度量的子集. 以 μ 表示度量 (当 $n=1$ 时, μ 为长度; 当 $n=2$ 时, μ 为面积; 当 $n=3$ 时, μ 为体积). 定义 A 的概率 $P(A)=\mu(A)/\mu(\Omega)$, 称这种概率模型为几何概率模型 (geometric probability model), 简称几何概型, 所定义的概率 $P(A)$ 为几何概率 (geometric probability). 当 $n>3$ 时, 也可类似地定义 n 维空间中的几何概型和几何概率.

易知, 几何概率具有以下性质:

(1) $0\leqslant P(A)\leqslant 1$, A 是试验的任意事件;

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

(3) 对有限个或可列无穷个两两互斥事件构成的族 $\{A_i\}$, 有

$$P\Big(\bigcup_i A_i\Big) = \sum_i P(A_i).$$

特别地, 对任意事件 A, 有 $P(\overline{A}) = 1 - P(A)$.

例 1.2.7(会面问题) 甲、乙二人相约于某日晚 19:00 至 20:00 在某地相会, 约定先到者等 20 分钟后离去. 求甲、乙能相会的概率.

解 以 x, y 分别表示甲、乙到达的时刻 (不妨以 19::00 为时间起点, 以分钟为单位记时), 它们可取 [0, 60] 内的任意值. 此时, 样本空间 $\Omega = \{(x,y)\,|\,0 \leqslant x, y \leqslant 60\}$, 其面积 $\mu(\Omega) = 60^2$. 假定各样本点出现等可能. 令 $A =\{$甲、乙能会面$\}$, 则 $A = \{(x,y)\,|\,-20 \leqslant x-y \leqslant 20,\ (x,y) \in \Omega\}$, A 是图 1.2.1 中阴影部分区域, 面积 $\mu(A) = 60^2 - 40^2$. 于是,

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{60^2 - 40^2}{60^2} = \frac{5}{9}.$$

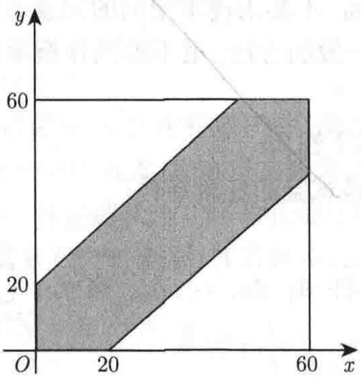

图 1.2.1 会面问题几何概型图

例 1.2.8 见图 1.2.2, 一圆盘以一定速度旋转, 在圆盘的同一平面上作长为 $2h$ 的线段 AB, AB 的中垂线通过圆盘的中心 C, C 到 AB 的距离为 l. 如果从圆盘的周上沿切线方向飞出一粒子, 求粒子能打到 AB 上的概率.

解 设 AD, BE 是切线, 粒子从圆盘的周上的每个点飞出是等可能的, 只有从 $\overset{\frown}{DE}$ 上飞出的粒子能打到 AB 上. 注意到, 当 CD 转到 CE 时, DA 同 EB 重合, CA 同 CB 重合, 所以, $\angle DCE = \angle ACB$, 令其弧度数为 φ, 则 $\tan(\varphi/2) = h/l$, 从而 $\varphi = 2\arctan(h/l)$. 令 $G =\{$粒子打到 AB 上$\}$, 则 $P(G) = \varphi/(2\pi) = \pi^{-1}\arctan(h/l)$.

此为试读, 需要完整PDF请访问: www.ertongbook.com