21世纪高等学校规划教材

概率论与数理统计

张雁芳 王刘禾 主编
刘洋洋 刘浪 副主编

中国工信出版集团
人民邮电出版社
POSTS & TELECOM PRESS

21世纪高等学

概率论与数理统计

张雁芳 王刘禾 主编
刘洋洋 刘浪 副主编

人民邮电出版社
北京

图书在版编目（CIP）数据

概率论与数理统计 / 张雁芳，王刈禾主编．-- 北京：人民邮电出版社，2015.9
21世纪高等学校规划教材
ISBN 978-7-115-39940-3

Ⅰ．①概… Ⅱ．①张… ②王… Ⅲ．①概率论－高等学校－教材②数理统计－高等学校－教材 Ⅳ．①021

中国版本图书馆CIP数据核字(2015)第182689号

内 容 提 要

本书以培养应用技术型人才为目标，以教育部高等学校教学指导委员会2014年数学课程教学要求为标准，借助大量实例系统讲解概率统计的基本概念、基本理论和基本计算方法，以增强学习的直观性和应用性，从而提高学生的学习兴趣和实际动手能力．本书前四章是概率论基本内容，在中学概率感性认识基础上进行系统、深入讲解，同时为数理统计准备必要的理论基础；后四章是在概率论基础上侧重分析如何用统计方法分析、解决带有随机性的实际问题．两部分内容配合紧密，每章还有大量习题作为正文的有机组成部分．本书可以作为应用技术型院校各类专业学生、自考生的概率统计教材，也可作为相关专业技术分析人员的自学或参考书．

◆ 主　　编　张雁芳　王刈禾
　副 主 编　刘洋洋　刘　浪
　责任编辑　张孟玮
　执行编辑　李　召
　责任印制　沈　蓉　彭志环

◆ 人民邮电出版社出版发行　　北京市丰台区成寿寺路11号
　邮编　100164　　电子邮件　315@ptpress.com.cn
　网址　http://www.ptpress.com.cn
　北京艺辉印刷有限公司印刷

◆ 开本：787×1092　1/16
　印张：11.25　　　　2015年9月第1版
　字数：262千字　　　2015年9月北京第1次印刷

定价：29.80元

读者服务热线：(010)81055256　印装质量热线：(010)81055316
反盗版热线：(010)81055315

前言

概率论与数理统计作为大学各专业都开设的一门基础学科，是数学类课程中应用性较强的一门课程，在自然科学、社会科学、工程技术、农业生产、医疗卫生等领域都有着广泛的应用. 这门学科的思维方式与传统的确定性数学的思维方式不太一样，它是建立在随机现象的统计规律上的，是寻找偶然中的必然，概率统计观念已成为现代观念的重要组成部分. 我国新课改从小学就开始了概率统计观念的教育，对随机现象的辨别和简单的分析，已成为绝大部分大一新生可以完成的事情，但是他们的认知基本还停留在感性认识阶段，对概率统计的基本理论、基本方法掌握不系统、不深入，没办法灵活应用. 而目前，高等教育领域有一大批普通本科高校将向应用技术型高校转型，对知识的应用提出了更高的要求，教育部高等学校数学课程教学指导委员会（以下简称数学教指委）也根据转型要求于 2014 年对《大学数学课程教学基本要求》进行了修订. 修订后的大纲更突出应用、淡化理论. 为做好基础教育和高等教育衔接，适应转型发展，本书作者严格按照数学教指委的要求，在自己多年的教学基础上，依据“理论够用、加强应用”的原则，为当前应用技术型人才培养提供参考教材.

本书依据 2014 版《大学数学课程教学基本要求》组织内容，着眼于系统介绍概率论与数理统计的基本思想、基本概念、基本理论和基本应用，让对概率统计有一定感性认识的学生形成系统观念；在内容设计上按照由简单到复杂、由特殊到一般的探索思路，逐步展开，借助大量实例揭示理论和概念的本质，同时在实例中进行应用示范，以开拓学生的思路与视野，培养学生转化知识进行应用的能力，在应用中进一步激发学生学习兴趣和热情. 在材料组织安排与例题的选配中，既考虑到为讲授者留有个人发挥的空间，又便于他们根据教学要求与学生情况对教材内容进行取舍.

本书可以作为应用技术型院校各类专业学生、自考生的概率统计教材，也可作为相关专业技术分析人员的自学或参考书. 讲授本书全部内容约 54 学时，其中概率论部分 26 学时，数理统计部分 28 学时.

本书主编为张雁芳、王刘禾，副主编为刘洋洋、刘浪.

本书编写过程中，得到了湖北文理学院及湖北文理学院理工学院有关领导和教师的支持与帮助，在此谨致谢意.

限于编者的水平和精力，书中难免存在错误和不妥之处，敬请广大读者批评指正.

编　者

2015 年 6 月

目录

引 言 概率统计漫谈

当今社会，是一个数据时代，人类与世界交流，数据是唯一的语言，小到工厂管理，大到科学研究，无不需要数据给我们指明前进的方向. 数据能够帮助我们认识世界，做出决策和预测. 概率统计作为数据分析的通用语言，成为数据时代预测未来的根基，也成为现代素养的一部分.

这样的结果应该是赌徒梅累当年没有预料到的. 梅累当年向数学家帕斯卡求教，仅仅是为了解决赌局中赌资的分配问题，谁知道可怕的数学家们前仆后继把这个问题研究成了一种强大的工具，工程、经济、金融等各个领域随处可见它的身影. 只要存在随机现象的地方就离不了概率统计.

历史上，早期的概率和统计几乎无太多关联，沿着各自的轨迹发展. 赌徒之问引发了概率的研究，人口、兵力、文化水平等社会问题导致了统计的产生. 回答赌徒之问的帕斯卡、费马、惠更斯被誉为概率论的创始人，开启用数学方法描述社会现象先河的政治经济学之父威廉·配第被马克思称为统计学创始人. 帕斯卡、费马完整地解决了“分赌注问题”，并建立了概率论的一个基本概念——数学期望. 拉普拉斯利用高等数学知识将古典概率向近代概率推进，他明确了概率的古典定义，引入了更有力的数学分析工具，并证明了“棣莫弗—拉普拉斯定理”，这是最早的中心极限定理. “如果我们能把一切事件永恒地观察下去，则我们终将发现：世间的一切事物都受到因果律的支配，而我们也注定会在种种极其纷纭杂乱的事象中认识到某种必然.”写下这段话的伯努利发现了“大数定律”的极限定理，从而推动了概率统计的融合，被誉为概率论的奠基人. 在帕斯卡、费马、惠更斯、拉普拉斯、泊松、高斯、柯尔莫哥洛夫、麦克斯韦、玻尔兹曼、吉布斯等一代又一代数学家继续努力下，概率统计由解决一个个问题逐渐发展为一门学科，成为联系宏观与微观的桥梁、确定性与不确定性的中介. 现在，概率统计已经成为动力学、系统论、协同学等众多学科的重要组成部分，成为心理学等社会科学研究中的重要方法.

概率统计作为研究随机现象规律的学科，它为人们认识客观世界提供了重要的思维模式和解决问题的方法. 实际上，概率和统计是这种工具的不同侧面，虽然它们之间有交融的部分. 概率是概率论的简称，一般研究随机事件发生的可能性大小、统计独立性和更深层次上的规律性. 统计，这里指的是数理统计，一般研究如何有效地收集、整理和分析受随机因素影响的数据，并对所考虑的问题做出推断或预测，为采取某种决策和行动提供依据或建议. 统计推断以概率知识为基础，通过局部或部分推断整体. 概率为统计学的发展提供了理论基础.

从方法论上讲，统计是推理，概率是归纳. 通俗来讲，概率论研究的是一个白箱子，你知道这个箱子的构造（里面有几个红球、几个白球，也就是所谓的分布函数），然后计算下一个摸出来的球是红球的概率. 而统计学面对的是一个黑箱子，你只看得到每次摸出来的是红球还是白球，然后需要猜测这个黑箱子的内部结构，例如红球和白球的比例是多

少（参数估计）？能不能认为红球 40%，白球 60%（假设检验）？概率论中的许多定理与结论，如大数定律、中心极限定理等保证了统计推断的合理性. 做统计推断一般都需要对那个黑箱子做各种各样的假设，这些假设都是概率模型，统计推断实际上就是在估计这些模型的参数. 概率统计不仅是科学研究中具有重要意义的理论，也是一种具有普遍意义的思想方法. 大家在学习的时候，要真正理解各个公式背后的现实生活意义，能够用概率统计的方法解决一些实际问题.

第 1 章 概率论基础

在日常生活中，随机现象非常普遍，比如每期福利彩票的中奖号码. 概率论与数理统计就是研究随机现象规律性的学科. 从本章开始，将引入概率论的基本概念、基本性质，并逐步开展概率统计理论和方法的研究.

1.1 随机事件

1.1.1 随机试验、样本空间、事件

1. 随机试验

为了找到某种随机现象的规律，就需要对这种随机现象进行研究. 研究的方法就是进行试验，一般是进行大量的试验，根据试验结果的统计规律，构造概率模型，作出客观、量化的判断.

这里的试验不同于物理、化学中的实验，这里的试验称为**随机试验**，主要目的是研究随机现象. 具体来讲，所谓随机试验（random experiment，一般记为 E）是指满足下述条件的试验：

(1) 试验在相同的条件下可以重复进行；

(2) 每次试验可出现不同的结果，但最终出现哪种结果，试验之前不能确定；

(3) 事先知道试验可能出现的全部结果.

【例 1.1】 随机试验的例子.

E_1：掷一枚匀称的骰子并观察顶面出现的点数.

E_2：抛一枚匀称的硬币 4 次，观察正面出现的次数.

E_3：抛一枚匀称的硬币 4 次，观察正反面出现的序列.

E_4：在一条生产线上制造零件，一天（24h）生产的次品零件的数目.

E_5：10 个零件含有 3 个次品，一个接一个（不放回）取出零件直至最后一个次品零件取出时，从这批零件中取出的零件总数.

E_6：制造零件直至 10 个正品生产出来时制造的零件总数.

E_7：生产一个灯泡，将它插入插座，记录直至灯灭所经过的时间（以小时计），即记录它的寿命.

E_8：向坐标平面区域 D：$x^2+y^2\leqslant 100$ 内随机投掷一点（假设点必落在 D 上），观察落点 M 的坐标.

E_9：测量某个零件的尺寸，观察测得的尺寸与规定尺寸的偏差 x(mm).

2. 样本空间

定义 1.1 对我们所研究的每一个试验 E，我们定义 E 的一切可能结果的集合为它的**样本空间**（sample space），一般记为 Ω. 每一可能结果称为**样本点**（sample point），常记作 ω.

【例 1.2】给出例 1.1 中随机试验的样本空间，把试验 E_i 的样本空间记为 Ω_i.

$\Omega_1=\{1, 2, 3, 4, 5, 6\}$.

$\Omega_2=\{0, 1, 2, 3, 4\}$.

$\Omega_3=\{$由 $a_1a_2a_3a_4$ 组成的一切序列$\}$，这里每个 $a_1=H$ 或 T 由第 i 次投掷时出现正面还是反面而定.

$\Omega_4=\{0, 1, 2, \cdots, N\}$，这里 N 是在 24h 内能够生产零件的最大数目.

$\Omega_5=\{3, 4, 5, 6, 7, 8, 9, 10\}$.

$\Omega_6=\{10, 11, 12, \cdots\}$.

$\Omega_7=\{t \mid t\geqslant 0\}$.

$\Omega_8=\{(x, y) \mid x^2+y^2\leqslant 100\}$.

$\Omega_7=\{x \mid a<x<b\}$.

样本点和样本空间是概率论中的两个基本概念. 随着对所讨论问题的兴趣不同，同一随机试验可以有不同的样本空间，如 E_2、E_3. 所以讨论问题前必须先确定相应的样本空间.

请注意，样本空间的样本点即试验的结果不一定是数. 样本空间的样本点的个数有 3 种情况：个数有限，如 $\Omega_1\sim\Omega_5$；个数无穷多个但可数可列，如 Ω_6；个数无穷且不可数，如 $\Omega_7\sim\Omega_9$.

3. 事件

样本空间的某些样本点组成的集合称为**随机事件**（random event），简称**事件**，一般用大写字母 A、B、C 等表示.

【例 1.3】按要求列出例 1.1 中随机试验对应的各个事件. 用 A_i 来表示相应于试验 E_i 的一个事件.

A_1：出现偶数点，即 $A_1=\{2, 4, 6\}$.

A_2：出现两次正面，即 $A_2=\{2\}$.

A_3：正面多于反面，即 $A_3=\{HHHH, HHHT, HHTH, HTHH, THHH\}$.

A_4：全部是正品，即 $A_4=\{0\}$.

A_5：取出的零件总数不多于 5 个，即 $A_5=\{3, 4, 5\}$.

A_6：制造的零件总数不少于 12 个，即 $A_6=\{12, 13, 14, \cdots\}$.

A_7：灯亮持续 12h 以上，即 $A_7=\{t \mid t>12\}$.

A_8：落在第一象限，即 $A_8=\{(x, y) \mid x^2+y^2\leqslant 100$ 且 $x>0, y>0\}$.

A_9：偏差小于 0.1 (mm)，即 $A_9=\{x \mid a<-0.1<x<0.1<b\}$.

关于事件的定义，有以下几点说明.

(1) 如果在一次试验中出现了某个样本点 ω，而 $\omega\in A$，则称事件 A 发生了.

(2) 由样本空间 Ω 中的单个元素组成的子集称为基本事件.

(3) Ω 本身称为必然事件，空集 $\varnothing$ 称为不可能事件.

(4) 当样本空间 Ω 是有限或可数可列时，每一个子集都可以看作是一个事件. 当 Ω 是无穷不可数时，不是每一个可能的子集都能作为一个事件，但一般不是事件的子集在实际中基本不会出现，所以在这里我们不必关心. 在此我们可以假定，我们所考虑的子集都是事件.

1.1.2 事件间关系与运算

假设以下讨论都是在同一样本空间 Ω 中进行的.

1. 事件的关系

(1) 包含关系.

如果 B 中的样本点都是 A 中的样本点，则称 B 包含于 A，或称事件 A 包含 B，也称 B 为 A 的子事件，记作 $A \supset B$（或 $B \subset A$）. 用概率论语言描述：事件 B 的发生必然导致事件 A 发生.

$B \subset A$ 一个等价的说法为，如果事件 A 不发生，则事件 B 必然不发生.

为方便起见，规定对于任一事件 A，有 $\varnothing \subset A \subset \Omega$.

(2) 相等关系.

若 $A \supset B$ 同时 $B \supset A$，则称事件 A 与 B 相等，记作 $A = B$.

(3) 互不相容关系.

如果 A 与 B 没有相同的样本点，即 $AB = \varnothing$，则称 A 与 B 互不相容（或称为互斥）. 用概率论语言描述为：事件 A 与 B 不能同时发生，即若事件 A 发生，则事件 B 必然不发生，反之亦然. 但是，有可能事件 A 与 B 都不发生.

基本事件是两两互不相容的.

(4) 互逆关系.

如果 A 与 B 互不相容且它们中必有一事件发生，即 $AB = \varnothing$ 且 $A \cup B = \Omega$，则称事件 A 与 B 是对立的（或互逆的），称事件 A 是事件 B 的对立事件（或逆事件）；同样，事件 B 也是事件 A 的对立事件（或逆事件），记为 $B = A$ 或 $\overline{A} = \overline{B}$.

对立事件必为互不相容事件，而互不相容事件未必是对立事件.

2. 事件间的运算

现在我们利用集合理论将已知的集合（即事件）A、B 通过各种运算以得到新的集合（事件）. 下面就来叙述它们.

(1) 事件 A 与 B 的和（并），记作 $A + B$（也可记作 $A \cup B$），称为和事件.

其含义为“由事件 A 与 B 中所有样本点（相同的只计入一次）组成的新事件”. 用概率论语言描述为事件 A 与 B 至少有一个发生.

(2) 事件 A 与 B 的积（交），记作 AB（也可记作 $A \cap B$），称为积事件.

其含义为“由事件 A 与 B 中公共样本点组成的新事件”. 用概率论语言描述为事件 A 与 B 同时发生.

(3) 事件 A 与 B 的差，记作作 $A - B$，称为差事件.

其含义为“由在事件 A 中而不在 B 中的样本点组成的新事件”. 用概率论语言描述为事件 A 发生，但 B 不发生. 显然 $A - B = A\overline{B} = A - AB$.

(4) 事件 A 的补运算，记作 $B = \overline{A}$，称为 A 的对立事件，或者 A 的**逆事件**.

其含义为“由在样本空间 Ω 中但不在事件 A 中的样本点组成的新事件 B”，即 $AB=\varnothing$ 且 $A+B=\Omega$. 用概率论语言描述为事件 A 不发生.

3. 事件运算的性质

事件的关系与运算满足集合论中有关集合运算的一切性质.

（1）**交换律**：$A+B=B+A$，$AB=BA$.

（2）**结合律**：$(A+B)+C=A+(B+C)$，$(AB)C=A(BC)$.

（3）**等幂律**：$A+A=A$，$AA=A$.

（4）**吸收律**：若 $A\subset B$，则 $AB=A$，$A+B=B$.

（5）**分配律**：$(A+B)C=AC+BC$，$(AB)+C=(A+C)(B+C)$.

（6）**德·摩根（De Morgan）律**：$\overline{A+B}=\overline{A}\,\overline{B}$，$\overline{AB}=\overline{A}+\overline{B}$.

对于多个事件，甚至对于无限可列个事件，德·摩根律也成立. 用概率论语言描述为任意多个事件至少有一个发生的反面是这些事件都不发生；任意多个事件都发生的反面是这些事件中至少有一个不发生.

从集合角度讲，任意一个事件都是样本空间的一个子集，事件之间的关系我们也可以借助 Venn 图来理解. 事件既可以用语言描述，也可以用集合表示，读者要学会把集合论的写法与事件运算的有关意义互相翻译，要学会利用事件的运算把复杂事件分解成简单事件.

【例 1.4】 若 A、B、C 是某个随机试验的三个事件，则

{事件 A 发生而 B 与 C 都不发生} 表示为 $A\overline{B}\,\overline{C}$ 或 $A-B-C$ 或 $A-(B+C)$；

{事件 A 与 B 都发生而 C 不发生} 表示为 $AB\overline{C}$ 或 $AB-C$ 或 $AB-ABC$；

{A、B、C 三个事件都发生} 表示为 ABC；

{A、B、C 三个事件恰好发生一个} 表示为 $A\overline{B}\,\overline{C}+\overline{A}B\overline{C}+\overline{A}\,\overline{B}C$；

{A、B、C 三个事件恰好发生两个} 表示为 $AB\overline{C}+A\overline{B}C+\overline{A}BC$；

{A、B、C 三个事件至少发生一个} 表示为 $A+B+C$ 或 $A\overline{B}\,\overline{C}+\overline{A}B\overline{C}+\overline{A}\,\overline{B}C+AB\overline{C}+A\overline{B}C+\overline{A}BC+ABC$.

【例 1.5】 一电路系统由元件 A 与 B 并联所得的线路再与元件 C 串联而成（如图 1.1 所示）. 若以 A、B、C 表示相应元件能正常工作的事件，那么事件 $W=$ {系统能正常工作} $=$ {元件 A 与 B 至少一个能正常工作并且 C 能正常工作} $=(A+B)C$ 或者 $AC+BC$.

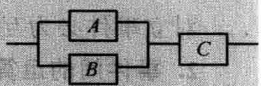

图 1.1 某电路系统

在此题中可以利用电路的相关知识来验证德·摩根律：

$$\overline{W}=\overline{(A+B)C}=\overline{C}+(\overline{A+B})=\overline{C}+(\overline{A}\,\overline{B}).$$

本书所讨论的试验的一个基本特征是在进行试验之前我们不知道哪个结果会发生. 也就是说，如果 A 是相应于某一试验的一个事件，那么我们不能断言 A 将发生还是不发生. 因此很重要的一个问题是尝试把事件 A 与一个数联系，这个数在某种意义上能量度出事件 A 发生的可能性大小，这个目的引导我们得到概率理论.

1.2 古典概型与概率

事件发生（出现）可能性大小是用概率来描述的. 概率的概念是逐步形成和完善起来

的. 最初人们讨论的是古典概型试验中事件发生的概率，即古典概率.

1.2.1 古典概型、随机抽球问题

1. 古典概型

所谓**古典概型**是指样本空间中的样本点的个数是有限的且每个样本点（组成的基本事件）发生的可能性是相同的，简称为有限性与等可能性. 对于古典概型试验，概率定义如下.

定义 1.2 设试验 E 是古典概型的，其样本空间 $\Omega=\{\omega_1, \omega_2, \cdots, \omega_n\}$. 若事件 A 由 k 个样本点组成，则定义 A（发生）的概率为 $\frac{k}{n}$，记为 $P(A)$，即

$$P(A)=\frac{A \text{ 中样本点数}}{\Omega \text{ 中样本点数}}=\frac{k}{n},$$

并称这样定义的概率为**古典概率**，称概率这样的定义为古典定义.

显然，古典概率的计算要借助加法原理、乘法原理以及排列与组合的相关知识得出 k 和 n，进而求得相应概率.

根据定义可得古典概率满足以下 3 个性质.

(1) 对任意事件 A，有 $0\leqslant P(A)\leqslant 1$.

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

(3) 设 $A_1, A_2, \cdots, A_m$ 为两两互斥的 m 个事件，则 $P(\bigcup\limits_{i=1}^{m} A_i)=\sum\limits_{i=1}^{m} P(A_i)$.

这三条分别称为概率的有界性、规范性与有限可加性.

下面讨论几种常见的古典概型计算及其应用.

2. 随机抽球问题

随机抽球问题是概率论中第一种典型的古典概型. 我们先讨论一个相对比较简单的抽球问题.

【例 1.6】 一口袋装有 6 只球，其中 4 只黑球、2 只红球. 从袋中取球两次，每次随机地取一只. 考虑两种取球方式：(a) 第一次取一只球，观察其颜色后放回袋中，搅匀后再取一球. 这种取球方式叫作**放回抽样**. (b) 第一次取一球不放回袋中，第二次从剩余的球中再取一球. 这种取球方式叫作**不放回抽样**. 试分别就上面两种情况求：

(1) 取到的两只球都是黑球的概率.

(2) 取到的两只球颜色相同的概率.

(3) 取到的两只球中至少有一只是黑球的概率.

解 设 $A=$｛取到的两只球都是黑球｝，$B=$｛取到的两只球都是红球｝，$C=$｛取到的两只球中至少有一只是黑球｝. 易知｛取到两只颜色相同的球｝$=A+B$.

在袋中分两次取两只球，每一种取法为一个基本事件，显然此时样本空间中仅包含有限个元素，且由对称性知每个基本事件发生的可能性相同，因而可利用古典定义来计算事件的概率.

(a) 放回抽样的情况.

第一次从袋中取球有 6 只球可供抽取，第二次也有 6 只球可供抽取. 由乘法原理知，共有 6×6 种取法. 即样本空间中元素总数为 6×6. 对于事件 A 而言，由于第一次有 4 只黑

球可供抽取，第二次也有 4 只黑球可供抽取，由乘法原理共有 4×4 种取法，即 A 中包含 4×4 个元素. 同理，B 中包含 2×2 个元素. 于是

$$P(A)=\frac{4\times4}{6\times6}=\frac{4}{9};\ P(B)=\frac{2\times2}{6\times6}=\frac{1}{9}.$$

由于 $AB=\varnothing$，得

$$P(A+B)=P(A)+P(B)=\frac{4}{9}+\frac{1}{9}=\frac{5}{9},$$

$$P(C)=P(\bar{B})=1-P(B)=\frac{8}{9}.$$

(b) 不放回抽样的情况，由读者自己完成.

【例 1.7】 袋中有 a 只黑球，b 只红球，k 个人依次在袋中取一只球，(1) 作放回抽样；(2) 作不放回抽样，求第 $i(i=1,2,\cdots,k)$ 人取到黑球（记为事件 B）的概率 $(k\leqslant a+b)$.

解 (1) 放回抽样的情况，显然有

$$P(B)=\frac{a}{a+b}.$$

(2) 不放回抽样的情况. 各人取一只球，每种取法是一个基本事件，共有 $(a+b)(a+b-1)\cdots(a+b-k+1)=\mathrm{A}_{a+b}^{k}$ 个基本事件，且由于对称性知每个基本事件发生的可能性相同. 当事件 B 发生时，第 i 人取的应是黑球，它可以是 a 只黑球中的任一只，有 a 种取法. 其余被取的 $k-1$ 只球可以是其余 $a+b-1$ 只球中的任意 $k-1$ 只，共有 $(a+b-1)(a+b-2)\cdots[a+b-1-(k-1)+1]=\mathrm{A}_{a+b-1}^{k-1}$ 种取法，于是 B 中包含 $a\cdot\mathrm{A}_{a+b-1}^{k-1}$ 个基本事件，故由定义 1.2 得

$$P(B)=a\cdot\mathrm{A}_{a+b-1}^{k-1}/\mathrm{A}_{a+b}^{k}=\frac{a}{a+b}.$$

值得注意的是 $P(B)$ 与 i 无关，即 k 个人取球，尽管取球的先后次序不同，各人取到黑球的概率是一样的，大家机会相同（例如在购买福利彩票时，各人得奖的机会是一样的）. 另外，还值得注意的是放回抽样的情况与不放回抽样的情况下 $P(B)$ 是一样的.

【例 1.8】 设袋中有 N 个红球，M 个黑球，现有放回地从袋中摸球，求：

(1) 在 n 次摸球中恰好摸到 $k(k\leqslant n)$ 个黑球的概率.

(2) 第 k 次才摸到黑球的概率.

(3) 如果摸球是不放回的，求在 n 次摸球中恰好摸到 $k(k=0,1,2,\cdots,\min(M,n))$ 个黑球的概率.

解 (1) 由于袋中有 $N+M$ 个球，且摸球是有放回的，故每次摸球都有 $N+M$ 种可能（这里设想球是编了号的，即是可辨的），设 A 表示事件“有放回地摸球，在 n 次摸球中恰好摸到 $k(k\leqslant n)$ 个黑球”，则

$$P(A)=\frac{\mathrm{C}_n^k M^k N^{n-k}}{(N+M)^n}=\mathrm{C}_n^k p^k(1-p)^{n-k},\ k=0,1,2,\cdots,n,\ 其中\ p=\frac{M}{N+M}.$$

由于 $\mathrm{C}_n^k p^k(1-p)^{n-k}$ 是二项展开式 $(p+1-p)^n=\sum\limits_{k=0}^{\infty}\mathrm{C}_n^k p^k(1-p)^{n-k}$ 的一般项，所以称 $\mathrm{C}_n^k p^k(1-p)^{n-k}$ 为二项分布的概率公式.

(2) 设 B 表示事件“有放回摸球中第 k 次才摸到黑球”，则

$$P(B)=\frac{MN^{k-1}}{(N+M)^k}=p(1-p)^{k-1},\ k=1,\ 2,\ \cdots,\ 其中\ p=\frac{M}{N+M}.$$

由于 $p(1-p)^{k-1}$ 是几何级数 $\sum_{k=1}^{\infty}p(1-p)^{k-1}$ 的一般项，所以称 $p(1-p)^{k-1}$ 为几何分布的概率公式.

(3) 设 C 表示事件“无放回地摸球，在 n 次摸球中恰好摸到 k 个黑球”，则

$$P(C)=\frac{C_M^k C_N^{n-k}}{C_{N+M}^n},\ k=0,\ 1,\ 2,\ \cdots,\ \min(n,\ M).$$

$\frac{C_M^k C_N^{n-k}}{C_{N+M}^n}$式即所谓**超几何分布的概率公式.**

二项分布模型和超几何分布模型最主要的区别在于是有放回抽样还是不放回抽样. 有放回抽样时，每次抽取的总体没有改变，因而每次抽到某物的概率都是相同的，可以看成是独立重复试验，此种抽样是二项分布模型；而不放回抽样，取出一个则总体中就少一个，因此每次取到某物的概率是不同的，此种抽样为超几何分布模型.

在实际中，产品的检验、疾病的抽查、农作物的选种等问题均可化为随机抽球问题. 如某人有 6 把钥匙，其中 3 把大门钥匙，但是他忘记了哪 3 把是大门钥匙，只好不放回随机试开，求他第 $k(1\leqslant k\leqslant 4)$ 次才打开大门的概率与在 3 次（试开）内打开大门的概率. 这个问题中如果把“钥匙”换成“球”，“大门钥匙”换成“黑球”，则问题就变成摸球问题，其中事件“第 $k(1\leqslant k\leqslant 4)$ 次才打开大门”变成“第 $k(1\leqslant k\leqslant 4)$ 次才摸到黑球”，事件“在 3 次（试开）内打开大门”变成事件“在前 3 次内摸到黑球”. 我们选择抽球模型的目的在于使问题的数学意义更加突出，而不必过多地交待实际背景.

1.2.2 随机分球问题

“分球入盒”问题是概率论中另一种重要的古典概型. 实际中的投信、分配、住宿、不定方程的求解问题都与“分球入盒”问题具有相通之处.

【例 1.9】 将 n 只不同编号的球放入 $N(N\geqslant n)$ 个盒中，每球以相同的概率被放入盒中，每盒容纳球数不限，试求：

(1) 恰有 n 个盒中各有一球的概率.

(2) n 个球都在一个盒中的概率.

(3) 至少有 2 个球在同一个盒中的概率.

解 因为每个球有 N 种放法，n 个球有 N^n 种放法，即样本点总数为 N^n.

(1) 设 A 表示事件“恰有 n 个盒中各有一球”. N 个盒中有 n 个盒各有 1 个球，是哪 n 个盒子？可能是前 n 个，也可能是后 n 个，共有 C_N^n 种不同情形. 对于某种指定的情形（如前 n 个盒中各有 1 个球），第一个球有 n 种放法，第 2 个球有 $n-1$ 种放法，依此类推，第 n 个球有 1 种放法，再由排列组合的乘法原理知，A 中有 $C_N^n n!$ 个样本点，所以

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

(2) 设 B 表示事件“n 个球都在一个盒中”，哪一个盒子？是 N 个之中的一个. 这有 C_N^1 种可能，即 B 中有 N 个样本点，所以

$$P(B)=\frac{N}{N^n}=\frac{1}{N^{n-1}}.$$

(3) 设 C 表示事件"至少有 2 个球在同一个盒中"，显然，$C=\overline{A}$，所以

$$P(C)=1-\frac{C_N^n n!}{N^n}.$$

实际应用：假设每人的生日在一年 365 天中的任一天是等可能的，即都等于 1/365，那么随机选取 $n(n\leqslant 365)$ 个人，他们的生日各不相同的概率为

$$\frac{C_{365}^n n!}{365^n}=\frac{365\cdot 364\cdots(365-n+1)}{365^n},$$

那么，n 个人中至少有两个人生日相同的概率为

$$p=1-\frac{365\cdot 364\cdots(365-n+1)}{365^n}.$$

经计算可得表 1.1 的结果.

表 1.1　n 个人中至少有两个人生日相同的概率

n	20	23	30	40	50	64	100
p	0.411	0.507	0.706	0.891	0.970	0.997	0.9999997

从表 1.1 可看出，在仅有 64 人的班级里，"至少有两人生日相同"这一事件的概率与 1 相差无几，因此，如做调查的话，几乎总是会出现的. 读者不妨试一试.

【例 1.10】 将 15 名新生随机地平均分配到 3 个班级中去，这 15 名新生中有 3 名是优秀生. 问：

(1) 每一个班级各分配到一名优秀生的概率是多少？

(2) 3 名优秀生分配在同一个班级的概率是多少？

解　15 名新生平均分配到 3 个班级中的分法总数为 $C_{15}^5C_{10}^5C_5^5=\frac{15!}{5!\times 5!\times 5!}$，每一种分配法为一基本事件，且由对称性易知每个基本事件发生的可能性相同.

(1) 将 3 名优秀生分配到 3 个班级使每个班级都有一名优秀生的分法共 3! 种. 对于这每一种分法，其余 12 名新生平均分配到 3 个班级中的分法共有 $\frac{12!}{4!\times 4!\times 4!}$ 种. 因此，每一班级各分配到一名优秀生的分法共有 $\frac{3!\times 12!}{4!\times 4!\times 4!}$ 种. 于是，所求概率为

$$p_1=\frac{3!\times 12!}{4!\times 4!\times 4!}\Big/\frac{15!}{5!\times 5!\times 5!}=\frac{25}{91}.$$

(2) 将 3 名优秀生分配在同一班级的分法共有 3 种. 对于这每一种分法，其余 12 名新生的分法（一个班级 2 名，另两个班级各 5 名）有 $12!/(2!\times 5!\times 5!)$ 种. 因此，3 名优秀生分配在同一班级的分法共有 $(3\times 12!)/(2!\times 5!\times 5!)$ 种，于是，所求概率为

$$p_2=\frac{3\times 12!}{2!\times 5!\times 5!}\Big/\frac{15!}{5!\times 5!\times 5!}=\frac{6}{91}.$$

"分球入盒"是应用非常广泛的一种古典概型，座位问题、下电梯问题等都可转化为分球入盒模型. 这里的球都是可辨的，还有球不可辨的情况下分球入盒问题，有兴趣的读