高等院校"十一五"规划教材

概率论与数理统计

张雅文 李晓莉◎主编

中国农业出版社

高等院校“十一五”规划教材

概率论与数理统计

张雅文　李晓莉　主编

中国农业出版社

图书在版编目（CIP）数据

概率论与数理统计 / 张雅文，李晓莉主编．—北京：中国农业出版社，2009.6
高等院校“十一五”规划教材
ISBN 978-7-109-13797-4

Ⅰ．概…　Ⅱ．①张…②李…　Ⅲ．①概率论-高等学校-教材②数理统计-高等学校-教材　Ⅳ．021

中国版本图书馆 CIP 数据核字（2009）第 055016 号

内容简介

本书为高等院校“十一五”规划教材，包括概率论与数理统计两部分内容，前 4 章为概率论部分，主要有随机事件与概率、随机变量及其概率分布、随机变量的数字特征、大数定律与中心极限定理；后 4 章为数理统计部分，主要有统计量及其分布、参数估计、假设检验、方差分析与回归分析。

本书结合作者多年教学经验与实践做了一些有益的处理和尝试，介绍了相关统计方法在 Excel 下的实现，并在各章末增添了相关的阅读材料，使读者更好地了解概率统计的客观背景。本书可供高等院校理工科、经济管理等本科专业作为教材或教学参考书使用。

中国农业出版社出版
（北京市朝阳区农展馆北路 2 号）
（邮政编码 100125）
责任编辑　朱　雷　魏明龙

中国农业出版社印刷厂印刷　　新华书店北京发行所发行
2009 年 6 月第 1 版　2009 年 6 月北京第 1 次印刷

开本：820mm×1080mm　1/16　　印张：15
字数：370 千字
定价：25.00 元

前言

概率论与数理统计是研究随机现象客观规律性的数学学科，是高等院校理工类及管理类专业的一门必修基础课。它的基本概念、理论和方法在自然科学和社会科学中有着广泛的应用。通过本课程的学习，使学生初步掌握处理随机现象的基本思想和方法，培养学生运用概率统计方法分析和解决实际问题的能力。正是基于这样的教学目的，我们在编写本教材时既保持了理论的严谨性，也兼顾了通俗性和直观性。在选材和叙述上尽量从实际出发，由浅入深、由具体到抽象、由一般到特殊，力求语言精练，通俗易懂。通过介绍相关统计方法在 Excel 下的实现，使学生能够利用计算机完成相应的统计计算，从而增强学生解决实际问题的能力，对学生创新能力的培养和综合素质的提高也起到一定的作用。本书的另一个特色是在各章末增添了相关的阅读材料，使学生能更好地了解概率统计的客观背景，并从中体会到概率论与数理统计这门学科的无穷魅力。

本书内容符合教育部“概率论与数理统计”教学指导委员会课程建设的指导意见和“十一五”规划教材的要求，可供高等院校理工科、经济管理等本科专业作为教材或教学参考书使用。

本书由张雅文和李晓莉主编，参加编写人员及编写章节如下：李晓莉编写第 1、8 章及各章末的阅读材料，王开永编写第 2、3 章，王林芳编写第 4 章，李秋芳编写第 5 章及统计方法在 Excel 下的实现，张雅文编写第 6 章，程毛林编写第 7 章。

全书最后由张雅文、李晓莉统稿，书中附表及插图由李秋芳完成。

限于作者的水平，本书难免有不当之处，敬请同行及专家批评指正。

编　者

2009 年 4 月

目录

第一章

随机事件与概率

概率论是研究随机现象统计规律的数学分支，具有十分丰富的内容．本章主要介绍概率论中的基本概念、概率的基本性质以及概率计算中常用的几个公式．重点是掌握概率的概念、随机事件概率的计算、概率基本公式及应用，难点是如何利用概率的性质和基本公式建立实际问题的概率模型，并解决实际问题．

§1.1 样本空间与随机事件

1.1.1 随机现象与随机试验

在自然界和人类社会活动中，存在着两类不同的现象，一类为在一定的条件下必然会发生或者必然不会发生的现象，如：在地球上，太阳每天从东边升起，从西边落下；在标准大气压下把水加热至 100 ℃必然沸腾等等，这类现象称之为必然现象．

另一类现象表现为在一定的条件下，可能会出现这样的结果也可能会出现那样的结果，观察或试验的结果不确定．如：抛掷一枚质地均匀的硬币，可能是正面朝上，也可能是反面朝上；从一批产品中任取一件，可能是合格品，也可能是不合格品；一射手向一目标射击，可能击中也可能击不中等等，这类现象称之为**随机现象**．随机现象在个别的试验或观察中结果是不确定的，但在大量重复的试验中，会呈现某种规律性，比如我们说“甲射手比乙射手的射击水平高”，即意味着“在过去大量的射击训练比赛中，甲射手的命中率比乙射手的命中率要高”．概率论与数理统计就是研究随机现象统计规律性的数学学科．

研究随机现象的统计规律性，必然要对客观现象进行大量的观察或试验，我们把在相同的条件下可以重复观察或试验的随机现象称为**随机试验**（random experiment），简称**试验**，通常用字母 E 表示．

随机试验满足如下条件：

（1）试验可以在相同条件下重复进行；

（2）每次试验的可能结果不止一个；

（3）试验前哪个结果出现不确定，但试验的所有可能结果是明确的．

例 1.1.1 随机试验的例子：

E_1：掷一枚硬币，考察哪个面向上；

E_2：掷一枚骰子，观察出现的点数；

E_3：观察一分钟内某十字路口通过的汽车数量；

E_4：考虑某种型号的电子元件的使用寿命．

1.1.2 样本空间与随机事件

对于一次随机试验，我们把所出现的每一个基本可能结果称之为**样本点**（sample point），记为 ω，而所有的样本点构成的集合称之为**样本空间**（sample space），记为 Ω.

例 1.1.2 表示下面随机试验的样本空间：

（1）掷一枚硬币，考察哪个面向上，样本点有两个，$\omega_1=$“正面”，$\omega_2=$“反面”，样本空间是 $\Omega_1=\{$正面，反面$\}$；

（2）掷一枚骰子，观察出现的点数，有 6 个样本点，则样本空间为 $\Omega_2=\{1, 2, 3, 4, 5, 6\}$；

（3）观察一分钟内某十字路口通过的汽车数量，汽车数量可以是 0，1，2，…，所以样本空间表示为 $\Omega_3=\{0, 1, 2, \cdots, n, \cdots\}$；

（4）考虑某种型号的电子元件的使用寿命，如果以 t 表示元件使用寿命，则所有的可能结果为实数空间 $\mathbf{R}$ 的一个子集，即样本空间为 $\Omega_4=\{t \mid t\geqslant 0, t\in\mathbf{R}\}$．

注意：（1）样本空间所包含的样本点是与随机试验的目的有关的，如：掷一枚骰子，若考虑出现的点数，则样本空间为 $\Omega=\{1, 2, 3, 4, 5, 6\}$，若考虑出现点数的奇偶性，则样本空间是 $\Omega=\{$奇数，偶数$\}$．试验条件都相同，但由于考虑目的不同，样本空间可能不同；

（2）样本空间中的样本点可以是数，也可以不是数；

（3）样本空间中的样本点个数可以是有限多个，如例 1.1.2 中的 Ω_1，Ω_2，也可以是无穷多个，无穷又可以是无穷可列，如例 1.1.2 的 Ω_3，或无穷不可列，如例 1.1.2 中的 Ω_4.

对于一个随机试验，我们经常要考虑某个试验结果是否会发生，比如，考虑掷一枚骰子出现的点数是否大于 3，考虑某种型号的电子元件的使用寿命是否不大于 100 小时等等．我们称随机试验的可能结果为**随机事件**（random event），简称**事件**，一般用大写的英文字母 A，B，C，…记之．

任何随机事件都是样本空间的子集，如在例 1.1.2 中，“掷一枚骰子点数大于 3”可以用集合 $A=\{4, 5, 6\}$ 表示，A 为样本空间 $\Omega_2=\{1, 2, 3, 4, 5, 6\}$ 的子集；“电子元件的使用寿命不大于 100 小时”可以表示为集合 $B=\{t, 0\leqslant t\leqslant 100\}$，$B$ 为 $\Omega_4=\{t \mid t\geqslant 0, t\in\mathbf{R}\}$ 的子集．

随机事件又可以分为基本事件和复合事件，由样本空间中的单个样本点 ω 构成的单点集 $\{\omega\}$ 称为**基本事件**，而由若干个基本事件组合而成的事件称为**复合事件**．如“掷一枚骰子，考虑出现的点数”，则 $\{1\}$，$\{2\}$，…，$\{6\}$ 都是基本事件，而“点数大于 3”可以用 $\{4, 5, 6\}$ 表示，是一个复合事件，它是由三个基本事件 $\{4\}$，$\{5\}$，$\{6\}$ 复合而成的．

在每次试验中必然发生的事件称为**必然事件**，记为 Ω. 这是因为在任何一次试验中，总会出现 Ω 的某个基本事件，所以任何一次试验 Ω 必然会发生，因此用 Ω 表示必然事件是合理的．同理我们把在每次试验中都不会发生的事件称为**不可能事件**，记为 $\varnothing$. 严格地讲，必然事件与不可能事件已经不具备随机性，但为研究问题的方便，我们把它们作为随机事件的两个极端情形．

1.1.3 事件间的关系与运算

在同一样本空间中，往往要考虑许多事件，有些事件比较简单，有些则比较复杂．概率论的一个重要研究课题就是希望通过对比较简单事件的分析，去分析了解复杂事件．由于事件可以表示为

样本空间的子集，所以事件之间的关系和运算完全可以归结为集合间的关系和运算.

1. 事件的包含

若事件 A 发生必导致事件 B 发生，则称事件 B **包含**事件 A，或称事件 A **包含于**事件 B，记作 $B\supset A$ 或 $A\subset B$. 显然，对任意事件 A，$\varnothing\subset A\subset\Omega$.

2. 事件的相等

若事件 B 包含事件 A，且事件 A 包含事件 B，即 $B\supset A$ 且 $A\supset B$，则称事件 A 与事件 B **相等**，记作 $A=B$.

3. 事件的并

设 A 与 B 是任意两个事件，称"两个事件 A 与 B 至少有一个发生"为事件 A 与事件 B 的并，记作 $A\cup B$.

事件的并可以推广到有限个或可列无穷多个事件的情形：

"n 个事件 A_1，A_2，…，A_n 至少有一个发生"称为 A_1，A_2，…，A_n 这 n 个事件的**并**，记作 $A_1\cup A_2\cup\cdots\cup A_n$ 或 $\bigcup\limits_{i=1}^{n}A_i$；

"可列无穷多个事件 A_1，A_2，…，A_n，…至少有一个发生"称为可列无穷多个事件的并，记作 $\bigcup\limits_{i=1}^{\infty}A_i$.

4. 事件的交

设 A 与 B 是任意两个事件，称"两个事件 A 与 B 同时发生"为事件 A 与 B 的**交**，记作 $A\cap B$ 或 AB.

同样，事件的交也可以推广到有限个或可列无穷多个事件的情形：

"n 个事件 A_1，A_2，…，A_n 同时发生"称为这 n 个事件的交，记作 $A_1A_2\cdots A_n$ 或 $\bigcap\limits_{i=1}^{n}A_i$；

"可列无穷多个事件 A_1，A_2，…，A_n，…同时发生"称为这可列无穷多个事件的交，记作 $\bigcap\limits_{i=1}^{\infty}A_i$.

5. 事件的差

设 A 与 B 是任意两个事件，称"事件 A 发生而事件 B 不发生"为事件 A 与 B 的**差事件**，记作 $A-B$.

6. 事件的互不相容

若事件 A 与 B 不能同时发生，即 $AB=\varnothing$，则称事件 A 与 B 是互不相容的（或互斥的）.

推广到 n 个事件的情形，若 n 个事件 A_1，A_2，…，A_n 中任意两个事件互不相容，即 $A_iA_j=\varnothing(1\leqslant i<j\leqslant n)$，则称这 n 个事件**互不相容**.

7. 事件的对立

若两个事件满足 $AB=\varnothing$ 且 $A\cup B=\Omega$，则称事件 A 与 B 是对立的，并称事件 B 是事件 A 的**对立事件**（或**逆事件**）；同样，事件 A 也是事件 B 的**对立事件**，记作 $B=\overline{A}$ 或 $A=\overline{B}$.

于是有 $\overline{\overline{A}}=A$，$A\overline{A}=\varnothing$，$A\cup\overline{A}=\Omega$，而 A 与 B 的差事件 $A-B$ 也可以表示为 $A\overline{B}$.

若用平面上某个矩形区域表示样本空间 Ω，矩形区域内的点表示样本点，则上述事件的关系及运算可以用韦恩图直观地表示出来，如图 1.1.1 所示.

事件之间的运算律与集合之间的运算律也完全类似：

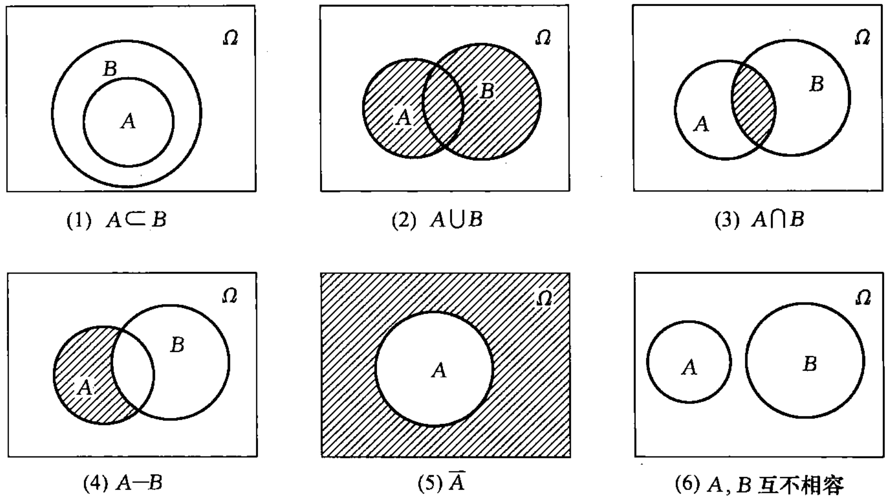

图 1.1.1

8. 事件的运算律

交换律 $A\cup B=B\cup A$；$A\cap B=B\cap A$.

结合律 $A\cup(B\cup C)=(A\cup B)\cup C$；$A\cap(B\cap C)=(A\cap B)\cap C$.

分配律 $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$；$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$.

对偶律（De Morgan 律） $\overline{A\cup B}=\overline{A}\cap\overline{B}$；$\overline{A\cap B}=\overline{A}\cup\overline{B}$.

在熟练掌握事件间的关系与运算，以及事件的运算律的基础上，对具体问题进行具体分析，可用基本事件表达复合事件，简单事件表达复杂事件 .

例 1.1.3 （1）“A 发生，而 B 与 C 都不发生”可表为 $A\overline{B}\overline{C}$ 或 $A(\overline{B\cup C})$；

（2）“A，B，C 中恰有一个发生”可表为 $A\overline{B}\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\overline{B}C$；

（3）“A，B，C 中恰有两个发生”可表为 $AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$；

（4）“A，B，C 中不多于一个发生”可表为 $\overline{A}\overline{B}\overline{C}\cup A\overline{B}\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\overline{B}C$ 或 $\overline{AB\cup BC\cup CA}$.

例 1.1.4 三只考签由三个考生轮流有放回抽取一次，每次取一只，设 A_i（$i=1$，2，3）表示“第 i 只考签被抽到”，试用 A_i 表示“至少有一只考签没有被抽到”这一事件 .

解 因为 A_i（$i=1$，2，3）表示“第 i 只考签被抽到”，所以“至少有一只考签没被抽到”可表为 $\overline{A}_1\cup\overline{A}_2\cup\overline{A}_3$ 或 $\overline{A_1A_2A_3}$.

§1.2 随机事件的概率及计算

对于一个随机试验，我们不仅要知道它可能会出现哪些结果，更重要的是研究各种结果发生的可能性的大小，从而揭示随机现象内在的规律性 . 对随机事件发生可能性大小的数量描述即为我们常说的**概率**（probability）. 本节主要介绍概率论发展的历史上，人们针对不同情况，从不同角度对事件的概率给出的几种定义，并给出相应的概率计算公式和性质 .

1.2.1 概率的统计定义

定义 1.2.1 在相同的条件下重复进行 n 次试验，若事件 A 发生了 M_n 次，则称比值 $\dfrac{M_n}{n}$ 为事

件 A 在 n 次试验中出现的**频率**（frequency），记为

$$f_n(A)=\frac{M_n}{n}. \tag{1.2.1}$$

频率的性质：

(1) **非负性：对任意** A，**有** $f_n(A)\geqslant 0$；

(2) **规范性：** $f_n(\Omega)=1$；

(3) **可加性：若** A，B **互不相容，则** $f_n(A\cup B)=f_n(A)+f_n(B)$.

在大量的重复试验中，频率常常稳定于某个常数，称为**频率的稳定性**.

通过大量的实践，我们很容易看到，若随机事件 A 出现的可能性越大，一般来讲，其频率 $f_n(A)$ 也越大. 由于事件 A 发生的可能性大小与其频率大小有如此密切的关系，加之频率又有稳定性，故可通过频率来定义概率.

定义 1.2.2（概率的统计定义） 在相同的条件下，进行独立重复的 n 次试验，当试验次数 n 很大时，如果某事件 A 发生的频率 $f_n(A)$ 稳定地在 $[0, 1]$ 上的某一数值 p 附近摆动，而且一般来说随着试验次数的增多，这种摆动的幅度会越来越小，则称数值 p 为事件 A 发生的**概率**，记为 $P(A)=p$.

概率的统计定义一方面肯定了任一随机事件的概率是存在的；另一方面又给出了一个近似计算概率的方法，在实际应用的许多场合中，甚至常常简单地把频率当作概率使用. 但定义要求试验次数足够大才能得到频率的稳定性，由于经济成本的限制，尤其是一些破坏性试验，不可能进行大量重复的试验，这些都限制了它的应用. 此外，由于数学意义上的不够严格，该定义在理论研究上的使用也很有限. 定义所刻画的事件 A 发生的概率与其频率大小的密切关系在第 4 章极限理论中我们将给出严格论述.

1.2.2　古典概型

定义 1.2.3 若一个随机试验满足：

(1) 样本空间中只有有限个样本点（有限性）；

(2) 每个样本点的发生是等可能的（等可能性），

则称该随机试验为**古典型随机试验**.

由于这一概型最早被人们所研究，故又称之为**古典概型**. 古典概型在概率论的研究中占有相当重要的地位，对它的讨论有助于直观地理解概率论的许多基本概念和性质，是概率论发展初期的主要研究对象. 下面给出古典概型的概率定义：

定义 1.2.4（概率的古典定义） 设古典型随机试验的样本空间为 $\Omega=\{\omega_1, \omega_2, \cdots, \omega_n\}$，若事件 A 中含有 $k(k\leqslant n)$ 个样本点，则称 $\frac{k}{n}$ 为 A 发生的**概率**，记为

$$P(A)=\frac{k}{n}=\frac{A\text{ 中含有的样本点数}}{\Omega\text{ 中总的样本点数}}. \tag{1.2.2}$$

古典概率的性质：

(1) **非负性：对任意** A，$P(A)\geqslant 0$；

(2) **规范性：** $P(\Omega)=1$；

(3) **可加性：若** A **和** B **互不相容，则** $P(A\cup B)=P(A)+P(B)$.

计算 Ω 和 A 的样本点数时通常要利用排列组合的知识，注意避免重复和遗漏.

例 1.2.1 从标号为 1，2，…，10 的 10 个同样大小的球中任取一个，求下列事件的概率：A=“抽中 2 号”，B=“抽中奇数号”，C=“抽中的号数不小于 7”.

解 令 i 表示“任取一球为 i 号球”，$i=1, 2, \cdots, 10$，则 $\Omega=\{1, 2, 3, \cdots, 10\}$，而事件 A 中包含 1 个样本点，事件 B 中包含 5 个样本点，事件 C 中包含 4 个样本点，所以有 $P(A)=\frac{1}{10}$，$P(B)=\frac{5}{10}$，$P(C)=\frac{4}{10}$.

例 1.2.2（抽签问题） 设一个袋中有 m 只红球，n 只黑球，它们除颜色外没有区别．现有 $(m+n)$ 个人依次从袋中随机地取出一球，求第 k 个人取到红球的概率.

解 设 A_k=“第 k 个人取到红球”，$k=1, 2, \cdots, m+n$.

首先计算样本空间 Ω 的样本点总数：当 $(m+n)$ 个人抽取 $(m+n)$ 个球时，相当于对 $(m+n)$ 个球进行全排列，故所有的排列总数为 $(m+n)!$；

再计算事件 A_k 所包含的样本点：“第 k 个人取到红球”这个事件可以分两步：首先取一个红球分给第 k 个人，一共有 m 种分法，剩下的 $(m+n-1)$ 个球再进行全排列，即事件 A_k 共有 $m\cdot(m+n-1)!$ 个样本点，

所以
$$P(A_k)=\frac{m\cdot(m+n-1)!}{(m+n)!}=\frac{m}{m+n},\ k=1, 2, \cdots, m+n.$$

从结果来看，事件 A_k 发生的概率与 k 无关，这说明不管第几个抽，抽到红球的可能性都相同. 这也是抽签方法广泛应用于各种场合的原因.

【思考】 请读者考虑一下，本题是否还有另外的解法.

例 1.2.3（抽样问题） 在工业生产过程中，经常采用下面两种抽样方式进行产品检验，一种称为有放回抽样，即每次抽取一件，检验完后仍将产品放回，再进行下一次抽取；另一种称为不放回抽样，即每次抽一件，检验完后不再将产品放回，再抽取下一件.

设有 100 件产品，其中有 95 件正品，5 件次品，分别按照上述两种抽样方式抽取 10 件产品，求其中恰有 2 件次品的概率.

解 (1) 有放回抽样：

由于每次抽取的产品仍然放回，所以每次都面临的是 100 件产品，那么 10 次抽取共有 100^{10} 种取法，即样本点总数为 100^{10}，

设 A_1=“从 100 件产品中有放回依次抽取 10 件产品，其中恰有 2 件次品”，即“10 次抽取中有 8 次取得了正品，2 次取得了次品”，而 8 次取得的正品都是在 95 件正品中取得的，有 95^8 种取法，2 件次品是在 5 件次品中取到的，故有 5^2 种取法；又因为 2 件次品可以是 10 次抽取中的任何两次，所以有 C_{10}^2 种情况．因此，事件 A_1 包含的样本点总数为 $C_{10}^2\times95^8\times5^2$.

依古典概率的定义得
$$P(A_1)=\frac{C_{10}^2\times95^8\times5^2}{100^{10}}=C_{10}^2\left(\frac{95}{100}\right)^8\left(\frac{5}{100}\right)^2\approx0.0746.$$

(2) 不放回抽样：

由于每次抽取的产品不再放回，所以第一次抽取时有 100 件产品，第二次抽取时有 99 件产品，…，依此类推，那么 10 次抽取相当于从 100 个元素中取 10 个元素的不允许重复排列，共有

P_{100}^{10}种取法，即样本点总数为 P_{100}^{10} .

设 A_2= “从 100 件产品中不放回依次抽取 10 件产品，其中恰有 2 件次品”，即 “10 次抽取中有 8 次取得了正品，2 次取得了次品”，而 8 次不放回取得的正品应有 P_{95}^{8}种取法，2 次取到的次品有 P_5^2 种取法 . 又因为 2 件次品可以是 10 次抽取中的任何两次，所以有 C_{10}^2种情况 . 因此，事件 A_2 包含的样本点总数为 $C_{10}^2P_{95}^8P_5^2$.

依古典概率的定义得

$$P(A_2)=\frac{C_{10}^2P_{95}^5P_5^2}{P_{100}^{10}}\approx 0.070\,2.$$

值得注意的是，若是从 100 件产品中，一次取出 10 件，其中恰有 2 件次品的概率与 $P(A_2)$ 相等，即设 A_3= “从 100 件产品中任取 10 件产品，其中恰有 2 件次品”，由于 10 件产品是一次抽取的，因此可以不考虑抽取顺序 . 故

$$P(A_3)=\frac{C_{95}^5C_5^2}{C_{100}^{10}}\approx 0.070\,2.$$

其实利用排列组合的性质，不难验证

$$P(A_2)=\frac{C_{10}^2P_{95}^5P_5^2}{P_{100}^{10}}=\frac{C_{10}^2\times C_{95}^5\times 5!\times C_5^2\times 2!}{C_{100}^{10}\times 10!}=\frac{C_{95}^5C_5^2}{C_{100}^{10}}=P(A_3).$$

例 1.2.4（盒子模型） 设有 n 个球，每个球被等可能地放到 N 个不同的盒子中的任一个，假设每个盒子所能容纳的球无限 . 试求：

(1) 指定的 $n(n\leqslant N)$ 个盒子中各有一球的概率 p_1；

(2) 恰好有 $n(n\leqslant N)$ 个盒子各有一球的概率 p_2.

解 因为每个球都以相同的可能性放到 N 个盒子中的任一个，所以 n 个球放的方式共有 N^n 种.

(1) 因为放球的盒子已经被指定，所以只要考虑把 n 个球放到 n 个盒子中的放法，其可能种数为 $n!$ ，故所求概率为

$$p_1=\frac{n!}{N^n}.$$

(2) 该问题与问题 (1) 的差别是放有球的 n 个盒子要在 N 个盒子中任意选取，所以可以分为两步走：第一步，首先在 N 个盒子中任取 n 个盒子，共有 C_N^n 种取法；第二步，把 n 个球放到 n 个已选中的盒子中，其可能种数为 $n!$. 由乘法原则，共有 $C_N^n n!$ 种放法，因此所求概率为

$$p_2=\frac{C_N^n n!}{N^n}=\frac{N!}{N^n(N-n)!}.$$

盒子模型是一类重要的概率模型，可以应用到许多实际问题 . 下面的生日问题就是历史上著名的一例 .

例 1.2.5（生日问题） 考虑由 n 个人组成的班集体，问至少有两人生日相同的概率是多少（一年以 365 天计）?

解 记 A= “至少有两人生日相同”，首先考虑其逆事件 $\overline{A}$= “n 个人生日全不相同”，把人看作 “球”，一年的 365 天看作 365 个 “盒子”，那么问题就归结为盒子模型，则 $P(\overline{A})=\dfrac{C_N^n n!}{N^n}$ ，因此

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

下面给出对于不同的 n，$P(A)$ 的计算结果（表 1.2.1）：

表 1.2.1 例 1.2.5 部分计算结果

n	10	20	30	40	50	60
$p(A)$	0.116 0	0.405 8	0.696 3	0.882 0	0.965 1	0.992 2

由表中结果可以看出，当一个集体的人数达到 60 人时，至少有两人生日相同的概率超过 99%，这有些出乎人们意料．

【思考】 考虑由 n 个人组成的班集体，问至少有一人与班长的生日相同的概率是多少（一年以 365 天计)？

概率的古典定义只能解决满足古典概型要求的概率计算，若样本空间的样本点是等可能的，但是有无穷多个且充满了一个几何体时，可以利用几何方法解决概率的计算．

1.2.3 几何概型

先看一个简单的例子：

例 1.2.6 如果在一个 50 000 平方千米的海域里有表面积达 40 平方千米的大陆架贮藏着石油，假如在海域里随意选取一点钻探，问钻到石油的概率是多少？

解 在该题中由于选点的随机性，可以认为该海域中各点被选中的可能性是一样的，因而所求概率自然可以认为是贮油海域的面积与整个海域面积之比，即 $p=\frac{40}{50\ 000}$.

在这类问题中，试验的一个可能结果是某几何空间中的一个点，这个区域可以是一维、二维、三维的，甚至可以是 n 维的，而试验的每个结果都是等可能的，且所有可能结果充满一个几何区域，记为 Ω. 我们所关心的事件 A 对应这个区域的一个子区域 D，那么事件 A 发生可能性的大小描述为：落在区域 D 的可能性与区域 D 的测度（长度、面积、体积等）成正比而与其位置及形状无关．

定义 1.2.5（概率的几何定义） 若以 A 记“在区域 Ω 中随机地取一点，而该点落在区域 D 中”这一事件，则其概率定义为

$$P(A)=\frac{S_A}{S_\Omega}. \tag{1.2.3}$$

其中，S_Ω 为样本空间 Ω 的几何度量，S_A 为事件 A 所表示的区域 D 的几何度量．

几何概型的性质：

(1) **非负性：** $\forall A$，$P(A)\geqslant 0$；

(2) **规范性：** $P(\Omega)=1$ ；

(3) **可列可加性：若** A_1，A_2，**…两两互不相容，则** $P\left(\sum_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}P(A_n)$．

例 1.2.7 假如某公交车每 10 min 一班，某同学随机到达车站等车，问：

(1) 该同学等车时间不超过 5 min 的概率是多少？

(2) 该同学等车时间介于 4 min 到 5 min 之间的概率是多少？

解　该问题属于几何概型，样本空间为$\Omega=\{x \mid 0\leqslant x\leqslant 10,\ x\in \mathbf{R}\}$.

(1) 令$A=$ “等车时间不超过 5 min”，则$P(A)=\dfrac{5}{10}=0.5$;

(2) 令$B=$ “等车时间介于 4 min 与 5 min 之间”，则$P(B)=\dfrac{1}{10}=0.1$.

【思考】　该同学等车时间恰好等于 5 min 的概率是多少？由此你可以得到什么结论？

例 1.2.8（会面问题）　两人相约 7 点到 8 点在某地会面，先到者等候另一人 20 min，这时就可离去，试求这两人能会面的概率？

解　以x，y分别表示两人到达时刻（7 点设为零时刻），则样本空间Ω可以表示为

$$\Omega: 0\leqslant x\leqslant 60,\ 0\leqslant y\leqslant 60.$$

令$A=$ “两人能够会面”则会面的充要条件为

$$A: |x-y|\leqslant 20.$$

这是一几何概率问题，样本空间Ω可以表示为二维空间的一个边长为 60 的正方形，如图 1.2.1 所示，事件A为图中阴影部分，所求概率$P(A)=\dfrac{60^2-40^2}{60^2}=\dfrac{5}{9}$.

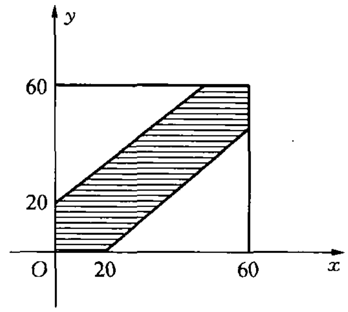

图 1.2.1

例 1.2.9（蒲丰投针问题）　平面上画有间隔为$d(d>0)$的等距平行线，向平面内任意投掷一枚长为$l(l<d)$的针，求针与任一平行线相交的概率.

解　以x表示针的中点与最近一条平行线的距离，又以φ表示针与此直线间的夹角，如图 1.2.2 所示，则样本空间Ω为：$0\leqslant x\leqslant \dfrac{d}{2}$，$0\leqslant \varphi\leqslant \pi$，表示为$x-\varphi$平面上的一个矩形，如图 1.2.3 所示，其面积$S_\Omega=\dfrac{\pi d}{2}$.

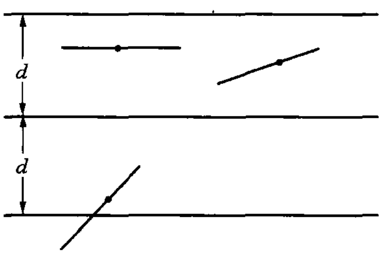

图 1.2.2

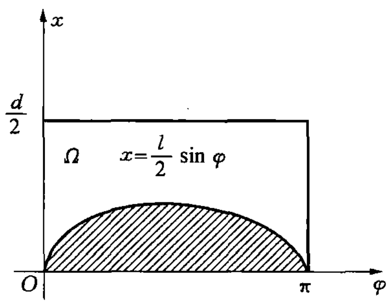

图 1.2.3

令$A=$ “针与平行线相交”，事件A发生当且仅当$0\leqslant x\leqslant \dfrac{l}{2}\sin\varphi$，$0\leqslant \varphi\leqslant \pi$，如图 1.2.3 所示的阴影部分，其面积是$S_A=\displaystyle\int_0^\pi \frac{l}{2}\sin\varphi \mathrm{d}\varphi$，故