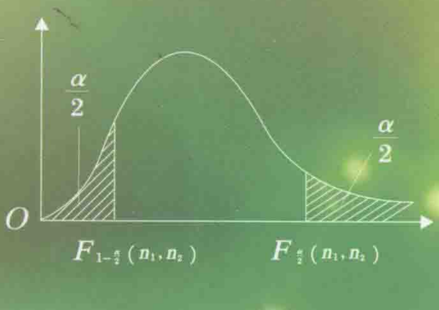
高等学校“十二五”规划教材

概率论与数理统计

主编 杨立夫



西北工業大學出版社

高等学校“十二五”规划教材

概率论与数理统计

主　编　杨立夫
编　者　杨立夫　郭天印　王树勋　田壤　苏晓海

西北工業大學出版社

【内容简介】 本书是作者根据多年来的教学实践经验，结合工科、经管学科各专业对概率论与数理统计课程的基本要求编写而成的。本书主要内容由两部分组成，第一部分为概率论部分，包括随机事件及其概率、随机变量、随机变量的数字特征、大数定律初步和中心极限定理四章内容；第二部分为数理统计部分，包括数理统计的基本概念、参数估计、假设检验、方差分析和回归分析五章内容，最后一章简要介绍了 Matlab 在数理统计中的应用。每章后附有习题，涵盖了研究生入学考试数学一和数学三考试大纲的所有知识点。

本书可作为工科、经管学科各专业的本科生教材，也可供工程技术人员及报考硕士研究生人员参考。

图书在版编目（CIP）数据

概率论与数理统计/杨立夫主编．—1 版．—西安：西北工业大学出版社，2012.8（2014.1 重印）

ISBN 978－7－5612－3423－5

Ⅰ．①概… Ⅱ．①杨… Ⅲ．① 概率论②数理统计 Ⅳ．①021

中国版本图书馆 CIP 数据核字(2012)第 190342 号

出版发行：西北工业大学出版社
通信地址：西安市友谊西路 127 号　　邮编：710072
电　　话：(029)88493844　88491757
网　　址：www.nwpup.com
印 刷 者：陕西向阳印务有限公司
开　　本：727 mm×960 mm　1/16
印　　张：17.625
字　　数：315 千字
版　　次：2012 年 8 月第 1 版　　2014 年 1 月第 2 次印刷
定　　价：36.00 元

前言

随着科学技术的发展，数学课程和其他学科一样，都面临着教学思想的转变及内容更新的问题。引入现代数学思想和方法，实现信息技术与学科课程的整合，变单纯的知识传授为知识、能力、素质的提升是摆在我们面前亟待解决的重大课题。

概率论与数理统计是研究大量随机现象的规律性的数学学科，它已被广泛地应用到自然科学、社会科学、工程技术、国防科技和工农业生产中，并与其他数学学科互相渗透和结合。因此，概率论与数理统计已成为本科阶段各专业学生必修的一门基础课。通过本课程的学习，学生能掌握研究随机现象的基本思想和方法，并具备一定的分析和解决问题的能力。

按照国家教育部高等学校工科数学课程教学指导委员会制定的概率论与数理统计课程基本要求，Ⅱ类(概率少，统计多)所规定内容的广度和深度以及概率论与数理统计课程内容本身的特点，结合我们和其他院校的教学实践经验，在试用多年讲义的基础上，编写了本教材。本教材着眼于介绍概率论与数理统计的基本概念、思想和方法，同时注重其直观背景和实际意义，力争体现下述特点。

(1)在保持该课程体系的基础上，分散难点，以应用广的内容为重点，合理安排，力求内容衔接紧凑，过渡自然。

(2)引入基本概念时，注意揭示其直观背景和实际意义，在阐述基本理论和基本方法时注意阐明概率和统计的意义和思想。

(3)削弱了古典概率的计算，突出概率概念的形成和完善过程。

(4)例题的选配注意结合工程技术实际，并注意能引起学生的兴趣，着力使学生理解基本理论和基本方法是怎样解决实际问题的，以培养学生运用概率统计方法解决实际问题的能力。同时，有些习题本身也是正文的补充和扩展，这些习题有助于学生巩固和进一步掌握有关理论内容，激发学生的学习兴趣。

(5)辟专节简述了概率统计在可靠性理论中的应用，使学生了解如何应用概率统计的理论解决实际问题。

(6)介绍了 Matlab 软件在概率论与数理统计中的应用。相关内容的直观演示和利用 Matlab 求解随机问题的案例分析，使学生学会借助于计算机进行随机模拟和数据处理的方法，同时逐步掌握使用 Matlab 进行数据处理的基本方法和技巧，

激发学习兴趣。

本书内容由两部分组成，第一部分为概率论部分，包括随机事件及其概率、随机变量、随机变量的数字特征、大数定律初步和中心极限定理；第二部分为数理统计部分，包括数理统计的基本概念、参数估计、假设检验、方差分析和回归分析。

全书由杨立夫副教授任主编，郭天印、王树勋、田壤、苏晓海为编写组成员，其中，杨立夫编写了第 2,5,7 章及附表，郭天印编写了第 8,9,10 章，王树勋编写了第 6 章，田壤编写了绪论及第 1 章，苏晓海编写了第 3、4 章并绘制了全部插图，最后由杨立夫、郭天印统一修改定稿。

在本书的编写过程中，参阅了许多兄弟院校的教材及资料，并得到陕西理工学院领导和教务处的大力支持，得到数计学院各位同事的帮助，在此表示衷心的感谢。

由于编者水平所限，书中不足之处，恳请读者、同行专家批评指正。

编　者

2012 年 6 月

目　　录

绪　论

0.1　事物的不确定性

在现实社会和工程领域、科学试验中，存在着大量具有不确定性的现象，这些不确定现象可归纳为两类：随机性和模糊性.

一、随机性

随机性是由于客观事物的因果关系不充分所致，表现为因果律的缺陷，造成结果的不可预知性. 例如，汉中明年 10 月 8 日的天气可能晴，可能阴，也可能有雨；同一射手，几次打靶的成绩，可能不同；等等. 它们的结果无法确定，也就是"随机"的. 这种在一定条件下可能发生种种不同结果的现象被称为**随机现象**.

然而，若对一随机现象进行多次重复观察，人们可以发现其出现的结果呈现出规律性. 例如投掷一枚硬币，它出现的结果可能是正面，也可能是反面，但若多次重复投掷硬币，则出现正面的次数约占一半. 像这种在大量重复试验或观察中所呈现出的固有规律性，叫**统计规律性**.

概率论与数理统计就是研究和揭示随机现象统计规律性的一门数学学科.

二、模糊性

与随机性不同，模糊性是指客观事物的差异在中介过渡中所呈现的"亦此亦彼"性，表现为排中律的缺陷，造成事物的边界不清晰(尽管结果已知). 例如在日常生活中，人们谈论年龄老中青、身材的高矮、体型胖瘦、气候冷热等概念，都具有含义不确切、边界不清晰的模糊性. 这种具有模糊性的现象叫**模糊现象**.

对事物模糊性的描述是建立在模糊集合论基础上的**模糊数学**.

三、确定性现象

与不确定性现象相对立的是确定性现象，即在一定条件下必然发生某一结果的现象. 例如，在标准大气压下，纯水加热到 100℃ 必然沸腾；向上抛一物体必然下落，等等.

研究确定性现象规律性的数学学科是建立在经典集合论基础上的数学分析、几何学、代数学以及微分方程等.

0.2 概率论与数理统计的应用

随机现象的普遍存在决定了概率统计应用的广泛性，它几乎遍及所有科学技术领域和工农业生产等国民经济的各部门.如：

(1)进行气象、水文、地震预报.

(2)产品的抽样验收及质量控制.

(3)研究新产品，为寻求最佳方案进行的试验设计及数据处理.

(4)在可靠性理论中的应用.近年来应用可靠性理论指导设备的设计、制造和维修，产生了巨大的效益.

第1章　随机事件及其概率

1.1　随机事件

一、随机试验

为研究随机现象，就要对客观事物在相同的条件下进行大量的观察. 这里的观察在概率统计中就称为随机试验. 那么，什么是随机试验呢？若一个试验具备以下特点：

(1)允许在相同条件下重复进行；

(2)每次试验的可能结果不止一个，但试验前可以明确该试验的所有可能结果；

(3)进行每次试验之前不能确定哪一结果会出现.

则称这种试验为**随机试验**，简称**试验**，通常用 E 表示.

例如：E_1：抛一枚硬币，观察正反面出现的情况；

E_2：投掷一颗骰子，观察出现的点数；

E_3：记录车站售票处一天内售出的车票数；

E_4：统计某网站每天的点击次数；

E_5：从一批圆钢中任取一条，测量它的抗拉强度 $f\,(\mathrm{kg/mm^2})$.

二、样本空间与随机事件

1. 样本空间

对于一个试验 E，尽管在每次试验之前不能预知试验的结果，但试验 E 的所有可能结果却是已知的，从而称试验 E 的所有可能结果组成的集合为 E 的**样本空间**，记为 S. 样本空间的元素，即试验 E 的每个可能的结果称为**样本点**，记为 e.

例如：E_1 的样本空间为：$S_1=\{$正面(H)，反面$(T)\}$；

E_2 的样本空间为：$S_2=\{1,2,3,4,5,6\}$；

E_3 的样本空间为：$S_3=\{0,1,2,3,\cdots,n\}$，这里的 n 表示售票处一天内准备出售的车票数；

E_4 的样本空间为：$S_4=\{0,1,2,3,\cdots\}$，这里点击次数理论上没有上限；

E_5 的样本空间为：$S_5=\{f \mid f>0\}$.

注意样本点 e 是由试验的目的所确定. 例如试验 E_2，若试验目的改为观察偶数点出现的情况，则样本空间为 $S_2=\{2,4,6\}$. 可见改变了试验的目的，其样本空间也要改变.

2. 随机事件

在每次试验中，有且只有一个结果（样本点）出现. 但在实际中，人们不但对试验的单一结果感兴趣，而且常常对试验的某些结果组成的集合更感兴趣. 例如，试验 E 表示在一批灯泡中任取一只，测试它的寿命. 对该试验 E，若规定这种灯泡的寿命（小时）$t<500$ 为次品，于是我们关心的是这批灯泡的寿命是否有 $t\geqslant 500$. 而满足该条件的结果（样本点）组成样本空间 $S=\{t \mid t\geqslant 0\}$ 的一个子集，即

$$A=\{t \mid t\geqslant 500\}$$

我们称 A 是试验 E 的一个随机事件.

一般地，设试验 E 的样本空间为 S，由 S 中的一些样本点所组成的集合，称为试验 E 的**随机事件**，简称**事件**，随机事件常用大写字母 $A,B,C,\cdots$ 等表示.

值得注意的是，随机事件 A 是样本空间 S 的一个子集，即 $A\subset S$. 在一次试验中，当且仅当 A 中的一个样本点出现时，就说在这一次试验中 A 发生了. 正因为事件 A 是由 S 中的一部分样本点组成的，所以在一次试验中，该事件可能发生，也可能不发生.

对于一个试验 E，在每次试验中必然发生的事件，称为 E 的**必然事件**；在每次试验中都不发生的事件，称为 E 的**不可能事件**. 于是必然事件可以用样本空间 S（是它本身的子集）来表示；不可能事件可用空集 $\varnothing$（是 S 的子集）来表示.

三、事件的关系和运算

在实际问题中，往往不只研究随机试验的一个事件，而要研究很多事件，且这些事件之间又有一定的联系. 由于事件是特殊的集合，因此事件间的关系与运算自然可运用集合论的一些术语、记号来描述. 下面，给出这些关系和运算在概率论中的意义. 假定以下讨论的均为同一试验 E 下的随机事件.

1. 事件的包含与相等

设有两个事件 A 与 B，若事件 A 发生时事件 B 必发生，则称事件 B 包含事件 A，或者说 A 含于 B，记为 $A\subset B$. 此时也称 A 是 B 的**子事件**.

若 $A\subset B$ 且 $B\subset A$，则称 A,B 两事件**相等**，记为 $A=B$.

例如，掷两颗骰子，记

$A=\{$掷出的点数之和大于 10$\}$，　　$B=\{$至少有一颗掷出的点数为 6$\}$

若 A 发生，易见 B 非发生不可，故 $A \subset B$.

又若 $A=\{$两骰子掷出的点数奇偶不同$\}$，$B=\{$两骰子掷出点数之和为奇数$\}$，这两个事件，只是表示说法不同，其实是一样的，故 $A=B$.

2. 事件的和与积

"事件 A 与事件 B 中至少有一个事件发生"这一事件，称为 A,B 的**和事件**，记为 $A\cup B$. 即

$$A\cup B=\{e \mid e\in A \text{ 或 } e\in B\}$$

"事件 A 与事件 B 同时发生"这一事件，称为 A,B 的**积事件**，记为 $A\cap B$ 或 AB. 即

$$AB=\{e \mid e\in A \text{ 且 } e\in B\}$$

类似地，对有限个事件 $A_1,A_2,\cdots,A_n$，"事件 $A_1,A_2,\cdots,A_n$ 至少有一个发生"这一事件称为 $A_1,A_2,\cdots,A_n$ 的**和事件**，记为 $\bigcup\limits_{i=1}^{n}A_i$；"事件 $A_1,A_2,\cdots,A_n$ 同时发生"这一事件称为事件 $A_1,A_2,\cdots,A_n$ 的**积事件**，记为 $\bigcap\limits_{i=1}^{n}A_i$. 也可以类似地定义可列个事件 $A_1,A_2,\cdots,A_n,\cdots$ 的和事件 $\bigcup\limits_{i=1}^{\infty}A_i$ 与积事件 $\bigcap\limits_{i=1}^{\infty}A_i$.

3. 互不相容事件与对立事件

若事件 A 与 B 不能同时发生，即 $AB=\varnothing$，则称 A 与 B 为**互不相容事件**，亦称**互斥事件**.

例如，掷一颗骰子，"出现 2 点"和"出现 5 点"是互斥事件；进行一次试验，"试验成功"与"试验失败"也是互斥事件.

互斥事件的一个特殊情况是对立事件.

设有事件 A，则事件 $B=\{A$ 不发生$\}$ 称为 A 的**对立事件**，记为 $\overline{A}$，即 $B=\overline{A}$.

例如：在掷一颗骰子的试验中，令 $A=\{$掷出偶数点$\}$，$B=\{$掷出奇数点$\}$，即 $A=\{2,4,6\}$，$B=\{1,3,5\}$，故 $B=\overline{A}$；不难有 $A=\overline{B}$. 这说明对立事件是相互的概念. 而且由本例亦可发现：一方面 $AB=\varnothing$，另一方面 $A\cup B=\{1,2,3,4,5,6\}=S$，即 B 包含的试验结果加上 A 包含的试验结果便补全了全部的试验结果，故对立事件亦称为"补事件". 很明显有：$\overline{\overline{A}}=A$.

一般地，若事件 A,B 互为对立事件，则必须满足 $AB=\varnothing$ 且 $A\cup B=S$.

4. 事件的差

"事件 A 发生，事件 B 不发生"这一事件称为 A 与 B 的**差事件**，记为 $A-B$，即

$$A-B=\{e \mid e\in A \text{ 且 } e\notin B\}$$

由差事件的定义可知

$$A-B=A\overline{B}$$

以上各事件的关系及运算的几何表示如图 1.1 所示.

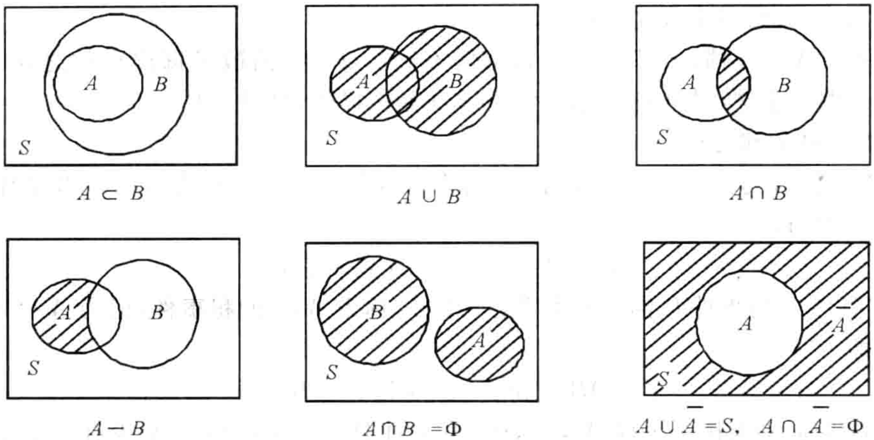


图 1.1

例 1 从一批产品中每次取出一个产品进行检验(每次取出的产品不放回),事件 A_i 表示第 i 次取到合格品 $(i=1,2,3)$.试用事件的运算符号表示下列事件:

(1)三次都取到合格品; (2)三次中至少有一次取到合格品;

(3)三次中恰有两次取到合格品; (4)三次中最多有一次取到合格品.

解 因 $A_i=\{$第 i 次取到合格品$\}$,$i=1,2,3$. 故

(1)三次都取到合格品:$A_1A_2A_3$;

(2)三次中至少有一次取到合格品:$A_1\cup A_2\cup A_3$;

(3)三次中恰有两次取到合格品:$A_1A_2\overline{A}_3\cup A_1\overline{A}_2A_3\cup\overline{A}_1A_2A_3$;

(4)三次中最多有一次取到合格品:$\overline{A}_1\overline{A}_2\cup\overline{A}_1\overline{A}_3\cup\overline{A}_2\overline{A}_3$.

例 2 如图 1.2 所示的电路中有编号为 1,2,3 的 3 个开关,设事件 $A_i=\{$开关 i 闭合$\}(i=1,2,3)$.试用 A_1,A_2,A_3 表示事件 $B=\{$电灯亮$\}$.

解 由图中 3 个开关的联接情况知,要电灯亮必须且只须开关 1,2 中至少有一个闭合且开关 3 闭合,即事件 B 发生相当于事件 A_1,A_2 中至少发生一个且事件 A_3 发生,而事件"A_1,A_2 中至少发生一个"可以表示为

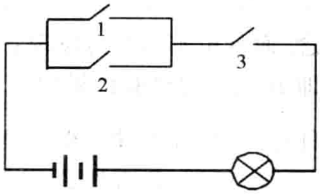


图 1.2

$$A_1\cup A_2 \quad 或\ A_1A_2\cup A_1\overline{A}_2\cup\overline{A}_1A_2 \quad 或\ \overline{\overline{A}_1\overline{A}_2}$$

故事件 B 发生可以表示为以下 3 种不同的形式:

$$B=(A_1\cup A_2)A_3$$

或

$$B=(A_1A_2\cup A_1\overline{A}_2\cup\overline{A}_1A_2)A_3$$

或
$$B=\overline{\overline{A_1}\,\overline{A_2}}A_3$$

5. 事件的运算律

在进行事件的运算时，经常要用到下列运算律：设事件 A,B,C，则有

交换律　$A\cup B=B\cup A$，　$AB=BA$

结合律　$(A\cup B)\cup C=A\cup(B\cup C)$，　$(AB)C=A(BC)$

分配律　$(A\cup B)C=(AC)\cup(BC)$，　$(AB)\cup C=(A\cup C)(B\cup C)$

对偶律(德·摩根律)　$\overline{A\cup B}=\overline{A}\,\overline{B}$，　$\overline{AB}=\overline{A}\cup\overline{B}$

现解释对偶律如下：
$$\overline{A\cup B}=\overline{\{A,B\text{ 中至少有一个发生}\}}=$$
$$\{A,B\text{ 都不发生}\}=\{\overline{A},\overline{B}\text{ 同时发生}\}=\overline{A}\,\overline{B}$$

因为
$$AB=\overline{\overline{A}}\,\overline{\overline{B}}=\overline{\overline{A}\cup\overline{B}}$$

所以
$$\overline{AB}=\overline{\overline{\overline{A}\cup\overline{B}}}=\overline{A}\cup\overline{B}$$

例 3　简化下列各式

(1)　$(A\cup B)-(A-B)$；　(2)　$(A\cup B)(A\cup\overline{B})(\overline{A}\cup B)$.

解　(1)　$(A\cup B)-(A-B)=(A\cup B)(\overline{A-B})=(A\cup B)(\overline{A\overline{B}})=$
$$(A\cup B)(\overline{A}\cup B)=(A\overline{A})\cup B=B.$$

(2)　$(A\cup B)(A\cup\overline{B})(\overline{A}\cup B)=(A\cup(B\overline{B}))(\overline{A}\cup B)=$
$$A(\overline{A}\cup B)=(A\overline{A})\cup(AB)=AB.$$

注意，对事件 A,B，若 $A\subset B$，则有
$$A\cup B=B,\quad AB=A$$

故对任意事件 A，有
$$A\cup S=S,\quad AS=A,\quad A\cup\varnothing=A,\quad A\varnothing=\varnothing$$

还需指出的是，在进行事件运算时，运算的优先顺序是：补，积，和或差；若有括号，则括号内的优先.

1.2　随机事件的概率

随机事件在一次试验中可能发生，也可能不发生. 既然有可能性，就有可能性的大小问题. 概率就是度量事件发生可能性大小的一个数量指标，也就是说，事件 A 发生的可能性大小就是事件 A 的概率. 常用 $P(A)$ 表示.

直观上很容易理解，必然事件发生的可能性是百分之百，故它的概率是 1，即 $P(S)=1$；而不可能事件发生的可能性是 0，故它的概率是 0，即 $P(\varnothing)=0$. 而任一

事件 A 发生的可能性不会小于 0，也不会大于百分之百，所以 A 的概率介于 0 与 1 之间，即

$$0 \leqslant P(A) \leqslant 1$$

那么，怎样获得某事件发生的概率呢？本节就来回答这个问题.

一、古典概型

对一类简单问题，我们可较容易求得事件发生的概率. 例如抛掷一枚硬币，这个试验只有两个可能结果："正面向上"（记为事件 A）或"反面向上"（记为事件 B）. 若这枚硬币质地均匀，又是对称的，则 $P(A)=P(B)=\frac{1}{2}$，因为我们没有理由认为哪种结果出现的可能性更大，也就是说，事件 A 与 B 出现的机会是均等的——这就是**等可能性**.

上例中的随机试验具有以下两个特点：

(1)试验的可能结果只有有限个，即样本空间只包含有限个样本点，即

$$S=\{e_1, e_2, \cdots, e_n\}$$

(2)每个试验结果在一次试验中出现是等可能的.

一般地，称此类随机试验的概率模型为**古典概型**. 它是概率论发展初期的主要研究对象.

定义 1.1 设古典概型试验 E 的所有可能结果为 $\{e_1\}, \{e_2\}, \{e_3\}, \cdots, \{e_n\}$. 若事件 A 恰包含其中的 m 个结果，则事件 A 的概率 $P(A)$ 定义为

$$P(A)=\frac{m}{n} \tag{1.1}$$

由于古典概型的两个特征："有限性"及"等可能性"，不难看出式(1.1)的合理性. 在一次试验中，每个结果出现的机会同为 $\frac{1}{n}$，现在事件 A 包含了 m 个结果，则在一次试验中，当这 m 个结果之一发生时事件 A 发生. 故事件 A 发生的概率应为 $m \cdot \frac{1}{n}=\frac{m}{n}$. 注意到式(1.1)中分子分母的含义，该式又可写成

$$P(A)=\frac{A\text{ 所包含的试验结果的个数}(m)}{\text{试验结果的总数}(n)}$$

由式(1.1)算得事件 A 的概率称为**古典概率**.

可见，对于古典概型，只要清楚试验 E 的所有可能结果总数 n 以及事件 A 所包含的试验结果个数 m，就可以求得事件 A 的概率. 这样就把求概率问题转化为计数问题，故排列组合是计算古典概率的重要工具.

例 1 袋内装有 5 个白球，3 个黑球. 从中任取两个球，计算取出的两个球都

是白球的概率.

解　令 $A=\{$取出的两个球都是白球$\}$. 分析题意知:试验的所有可能结果总数 $n=C_{5+3}^2$,事件 A 所包含的试验结果个数 $m=C_5^2$,由式(1.1)有

$$P(A)=\frac{C_5^2}{C_8^2}\approx 0.357$$

例 2　掷两个均匀骰子,求出现点数之和为 8 的概率.

解　设 X 为第一个骰子掷出的点数,Y 为第二个骰子掷出的点数. 该试验共有 $6\times 6=36$ 个等可能结果,即 $n=36$. 令 $A=\{$两个骰子的点数之和为 8$\}$,即 A 等价于"$X+Y=8$",只有 (2,6),(3,5),(4,4),(5,3),(6,2) 这 5 个结果之一出现时 A 才出现(见图 1.3). 故

$$P(A)=P\{X+Y=8\}=\frac{5}{36}$$

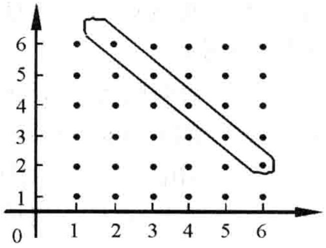


图　1.3

例 3　设有 r 个人($r\leqslant 365$),并且假定每人的生日在一年 365 天中的每一天的可能性是均等的. 问此 r 个人有不同生日的概率是多少?

解　设 $A=\{r$ 个人有不同生日$\}$. 因为 r 个人都以等可能的机会在 365 天中的任一天出生,所以 $n=365^r$. 根据题意,事件 A 包含的试验结果个数为

$$m=365\times 364\times\cdots\times(365-r+1)$$

它恰是 365 个数中任取 r 个数的排列 A_{365}^r ,即

$$m=A_{365}^r=\frac{365!}{(365-r)!}$$

因此所求概率为

$$P(A)=\frac{365!}{(365-r)!\ 365^r}=\frac{365\times 364\times\cdots\times(365-r+1)}{365^r}=$$
$$\left(1-\frac{1}{365}\right)\left(1-\frac{2}{365}\right)\cdots\left(1-\frac{r-1}{365}\right).$$

从古典概率的定义(1.1)中可以得到以下性质:

(1)对于任意事件 A,$0\leqslant P(A)\leqslant 1$;

(2)对于必然事件 S,不可能事件 $\varnothing$,有

$$P(S)=1,\quad P(\varnothing)=0$$

(3)对于互斥事件 A,B,即 $AB=\varnothing$,则有

$$P(A\cup B)=P(A)+P(B)\tag{1.2}$$

例 4　从一批由 95 件正品、3 件次品组成的产品中,接连抽取两件产品(第一