普通高等教育公共基础课程用书

应用数理统计

YINGYONG SHULI TONGJI

刘剑平　朱坤平　陆元鸿
编著

華東理工大學出版社
EAST CHINA UNIVERSITY OF SCIENCE AND TECHNOLOGY PRESS

应用数理统计

刘剑平　朱坤平　陆元鸿　**编著**

華東理工大學出版社
EAST CHINA UNIVERSITY OF SCIENCE AND TECHNOLOGY PRESS

图书在版编目（CIP）数据

应用数理统计/刘剑平，朱坤平，陆元鸿编著. —上海：华东理工大学出版社，2012.5
ISBN 978-7-5628-3264-5

Ⅰ. ①应… Ⅱ. ①刘… ②朱… ③陆… Ⅲ. ①数理统计 Ⅳ. ①O212

中国版本图书馆 CIP 数据核字（2012）第 068597 号

应用数理统计

编　　著/ 刘剑平　朱坤平　陆元鸿
责任编辑/ 郭　艳
责任校对/ 张　波
封面设计/ 肖　车
出版发行/ 华东理工大学出版社有限公司
　　社　址：上海市梅陇路 130 号，200237
　　电　话：(021) 64250306（营销部）　64252174（编辑部）
　　传　真：(021) 64252707
　　网　址：press. ecust. edu. cn
印　　刷/ 常熟华顺印刷有限公司
开　　本/ 787mm×1092mm　1/16
印　　张/ 16.25
字　　数/ 398 千字
版　　次/ 2012 年 5 月第 1 版
印　　次/ 2012 年 5 月第 1 次
书　　号/ ISBN 978-7-5628-3264-5/O·244
定　　价/ 39.00 元

本 书 编 委 会

主编　刘剑平　朱坤平　陆元鸿

编委　刘剑平　朱坤平　陆元鸿　钱夕元
曹宵临　俞绍文　林爱红　姬　超
闫中凤　孙　叶　李　平　樊国号
雷倩倩　叶炎钧

前言

数理统计是一门研究如何有效地收集、整理和分析受随机因素影响的数据，对所考查的问题作出推断，进而为制定决策和采取行动提供科学依据的学科．现在，数理统计方法已经在工业、农业、国防、科研、经济、管理、社会、医学、生物、考古、地质、气象等领域得到了广泛的应用．因为数理统计方法密切联系实际，所以学习和掌握数理统计方法对于理工科的本科生和研究生都是十分有益和必要的．

本书是作者根据多年的教学经验，参考国家教育部制定的“工学硕士研究生应用数理统计课程教学基本要求”，为工科研究生编写的一本数理统计教材．教材讲稿在我校本科生及研究生中反复试用，广受好评．

考虑到数理统计是一门实用性很强的学科，而工科学生包括研究生学习本课程的目的也在于用数理统计的方法来解决实际问题，因此，本书以讲清基本概念、原理和方法为主，避免烦琐的理论推导．全书力求简明扼要，清晰易懂，以便于读者通过自学就可以掌握本书的基本内容．

本书内容的选材和编排密切结合了研究生教学的特点．目前，国内关于研究生教材应如何编写还有很大的争议．一个普遍的共识是，与国外教材相比，国内的研究生教材大多更强调教材内容的逻辑性、严谨性，而缺乏探索性和启发性．为了启发读者对教材基本内容有更深入的理解和探索，在本书每章后都有一个延伸阅读和思考题，并指明了参考文献．全书共七章，内容包括概率论基础、抽样与抽样分布、参数估计、假设检验、回归分析、方差分析和正交试验设计、多元统计应用等，可根据需要供32—48学时的教学使用．有概率论基础知识的读者可以直接从第二章开始学习，学时少的可删除第七章多元统计应用，以及4.4节正态分布的概率纸检验、5.5节逐步回归分析等内容．

本书由刘剑平、朱坤平、陆元鸿老师主编，在编写的过程中，得到了华东理工大学教材建设委员会及教务处的大力支持，得到了院、系领导的关心和指导，在此表示衷心的感谢．同时，我们还要感谢教学组的全体成员，他们在本书的编写过程中提供了宝贵的建议．

由于编者水平所限，书内难免有疏漏差错之处，欢迎读者批评指正．另外，本书配有讲课的PPT，需要的读者请来函索取．作者电子邮箱 liujianping60@163.com 或 kpzhu@ecust.edu.cn.

目 录

4 假设检验

5 回归分析

6 方差分析和正交试验设计

7 多元统计应用

1

概率论基础

1.1 随机事件与概率

自然界和人类社会中存在着许多现象,其中有一些现象,只要满足一定的条件,就必然会发生.例如,在标准大气压下,纯水加热到100℃必然沸腾;向空中抛一枚硬币,硬币必然会下落.这些现象有一个共同特点,即事前人们完全可以预言会发生什么结果.我们称这类现象为**确定性现象**或**必然现象**.

但是在自然界和人类社会中,还存在着与必然现象有着本质差异的另一类现象.例如,抛一枚硬币,硬币可能正面向上也可能反面向上;一门大炮对目标进行远距离射击,可能击中也可能击不中.这些现象的一个共同特点是,在同样的条件下进行同样的观测或实验,有可能发生多种结果,事前人们不能预言将出现哪种结果.这类现象被称为**随机现象**或**偶然现象**.

表面上看来,随机现象的发生,完全是随机的、偶然的,没有什么规律可循,但事实上并非如此.对一次或少数几次观测或实验而言,随机现象的结果,确实是无法预料的,是不确定的.但是,如果我们在相同的条件下进行多次重复的实验或大量的观测,就会发现,随机现象的结果的出现,具有一定的规律性,例如,各个国家、各个时期的人口统计资料显示,新生婴儿中男婴和女婴的比例大约总是1∶1.又如,向水平、光滑的地面多次重复抛一枚均匀的硬币,发现正面向上与反面向上出现的次数近似相同.我们称这种规律性为随机现象的**统计规律性**.概率统计理论,就是研究随机现象统计规律性的一门学科.

1.1.1 随机试验与事件

为了方便起见,我们把对某种自然现象进行的一次观测或所做的一次实验,统称为一个试验.如果一个试验具有下列三个特性,就称这种试验为**随机试验**(Random Experiment).

(1)试验可以在相同条件下重复进行;

(2)每次试验,可以出现多种结果,总共有可能出现哪几种结果,是可以事先明确知道的;

(3)每一次试验,实际只出现一种结果,至于出现哪一种结果,在这一次试验结束之前,是无法预知的.

我们常用字母 E 来表示随机试验.一般把随机试验简称为**试验**.

进行试验的目的在于研究试验结果出现的规律. 对于一个试验 E,它的每一种可能出现的最简单的结果,称为**基本事件**或**样本点**,习惯上用 ω 表示.

对于一个试验 E,它所有的基本事件组成的集合称为**基本事件空间**或**样本空间**(Sample Space),记为 Ω.

例 1 设试验 E 为:掷一颗骰子并观察向上那面的点数,则有样本点

$$\omega_i=\{\text{出现 } i \text{ 点}\}(i=1,2,\cdots,6),$$

于是样本空间 $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}$.

例 2 设试验 E 为:在相同条件下接连不断地向同一个目标射击,直到第一次击中为止,观察直到击中为止所需要的射击次数. 则有样本点

$$\omega_i=\{\text{到击中为止需要射击 } i \text{ 次}\}(i=1,2,3,\cdots),$$

于是样本空间 $\Omega=\{\omega_1,\omega_2,\cdots\}$.

例 3 设试验 E 为:观察某地每日的最低气温 x 和最高气温 y,已知当地每日气温最低不低于-20℃,最高不高于 50℃. 以(x,y)表示一次观察结果,显然有$-20\leqslant x\leqslant y\leqslant 50$,则样本空间可表示为 $\Omega=\{(x,y)\mid -20\leqslant x\leqslant y\leqslant 50\}$,$\Omega$ 中的每一元素都是样本点.

以上各例说明随机试验的样本空间有以下三种情况:

(1)有有限个可能的结果;

(2)有可列无穷个(即无穷多个,但可依某种次序排成一列)可能的结果;

(3)有不可列无穷个(例如像一条线段上的点数那么多)可能的结果.

在一次随机试验中,通常关心的是带有某些特征的现象是否发生. 比如在例 1 中,

$$A=\{\text{出现 4 点}\},B=\{\text{出现点数为偶数}\},C=\{\text{出现点数不超过 4}\}.$$

A 是一个基本事件,而 B 和 C 则由多个基本事件组成(事实上,$B=\{\omega_2,\omega_4,\omega_6\}$,$C=\{\omega_1,\omega_2,\omega_3,\omega_4\}$),相对于基本事件,就称它们为**复合事件**. 无论基本事件还是复合事件,它们在试验中发生与否,都带有随机性,所以都称作**随机事件**(Random Event),简称**事件**. 习惯上,常用大写字母 A,B,C 等表示. 在试验中,如果事件 A 中包含的某一个基本事件 ω 发生,则称 A 发生,记为 $\omega\in A$.

样本空间 Ω 包含了全体基本事件,而随机事件是由具有某些特征的基本事件所组成. 由此可见,**任一随机事件都是样本空间的 Ω 的一个子集**. 样本空间有两个平凡的子集,即 Ω 本身和空集$\varnothing$. 如在例 1 中,“出现点数小于 10”就表示事件 $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}$,它在每次试验中一定发生. 而“出现点数小于 0”,就不包含任何样本点,即$\varnothing$中在每次试验中都不可能发生.

在每次随机试验中一定会发生的事件,称为**必然事件**. 相反地,如果某事件一定不会发生,则称为**不可能事件**.

必然事件与不可能事件没有“不确定性”,因而严格地说,它们已经不属于“随机”事件了. 但是,为了讨论方便起见,我们还是把它们包括在随机事件中,作为特殊的随机事件来处理.

1.1.2 事件的关系和运算

一个样本空间Ω中,可以有很多随机事件.人们通常需要研究这些事件间的关系和运算,以便通过较简单事件的统计规律去探求复杂事件的统计规律.

在讨论事件的关系和运算时,我们总是假定它们是同一个随机试验的事件,即它们是同一个样本空间Ω的子集.因为只有在这样的假定下,讨论它们之间的关系和运算才有意义.

(1)事件B包含事件A 如果事件A的发生必然导致事件B的发生,则称事件B包含事件A,或称事件A包含在事件B中,记作$B\supset A$或$A\subset B$.

(2)事件A与B相等 如果事件A包含事件B,而且事件B又包含事件A,则称A与B相等,记为$A=B$.

(3)事件A与B的和 "事件A与B中至少有一个发生"也是一个事件,称此事件为事件A与B的和(也称为事件的并),记作$A\cup B$或$A+B$.

"n个事件$A_1,A_2,\cdots,A_n$中至少有一个发生"这一事件称为事件$A_1,A_2,\cdots,A_n$的和,记作$A_1\cup A_2\cup\cdots\cup A_n$(或$A_1+A_2+\cdots+A_n$),简记为$\bigcup\limits_{i=1}^{n}A_i$(或$\sum\limits_{i=1}^{n}A_i$).

(4)事件A与B的积 "事件A与B同时发生"也是一个事件,称此事件为A与B的积(也称为事件的交),记作AB或$A\cap B$.

"n个事件$A_1,A_2,\cdots,A_n$同时发生"这一事件称为事件$A_1,A_2,\cdots,A_n$的积,记作$A_1A_2\cdots A_n$(或$A_1\cap A_2\cap\cdots\cap A_n$),简记为$\prod\limits_{i=1}^{n}A_i$(或$\bigcap\limits_{i=1}^{n}A_i$).

(5)事件A与B互不相容 如果事件A与B不可能同时发生,则称事件A与B互不相容或互斥.此时必有$AB=\varnothing$.

如果n个事件$A_1,A_2,\cdots,A_n$中任意两个事件都互不相容,即

$$A_iA_j=\varnothing \quad (1\leqslant i<j\leqslant n),$$

则称这n个事件是**互不相容**的或**互斥**(Exclusive)的.

显然任意一个随机试验E的所有基本事件都是互斥事件.并且,不可能事件$\varnothing$与任何事件都互斥.

(6)事件A与B相互对立 如果事件A与B满足

$$A\cup B=\Omega, AB=\varnothing,$$

则称事件B是事件A的逆事件或对立事件.容易看出,当B是A的逆事件或对立事件时,A也是B的逆事件或对立事件,所以,这时也称A与B是相互**对立**(或**互逆**)的事件,记为$B=\overline{A}$或$A=\overline{B}$,事件A的对立事件$\overline{A}$表示A不发生.

显然,$\overline{\Omega}=\varnothing,\overline{\varnothing}=\Omega,\overline{\overline{A}}=A$,并且$A\subset B$的充要条件是$\overline{A}\supset\overline{B}$.

(7)事件A与B的差 "事件A发生而事件B不发生"这一事件称为A与B的差,记为$A-B$.

例4 设一个工人生产了4个零件.A_i表示他生产的第i个零件是正品($i=1,2,3,4$),试用A_i表示下列事件:

(1)没有一个是次品;(2)至少有一个是次品;

(3)只有一个是次品;(4)至少有三个不是次品.

解　(1)$A_1A_2A_3A_4$;(2)$\overline{A_1A_2A_3A_4}$;

(3)$\overline{A_1}A_2A_3A_4 \cup A_1\overline{A_2}A_3A_4 \cup A_1A_2\overline{A_3}A_4 \cup A_1A_2A_3\overline{A_4}$;

(4)$\overline{A_1}A_2A_3A_4 \cup A_1\overline{A_2}A_3A_4 \cup A_1A_2\overline{A_3}A_4 \cup A_1A_2A_3\overline{A_4} \cup A_1A_2A_3A_4$.

关于事件的运算有如下的规律:

(1)交换律　$A\cup B=B\cup A, AB=BA$;

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

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

(4)德摩根[①]定律(对偶原理)

$$\overline{A_1\cup A_2\cup\cdots\cup A_n}=\overline{A_1}\,\overline{A_2}\cdots\overline{A_n}$$
$$\overline{A_1A_2\cdots A_n}=\overline{A_1}\cup\overline{A_2}\cup\cdots\cup\overline{A_n}.$$

我们知道,任一随机事件都是样本空间的一个子集,所以事件之间的关系与运算和集合之间的关系与运算是完全类似的.因而,可以借助于集合的知识来证明事件的运算规律.下面仅从事件运算含义的角度来解释对偶律的第一个公式.

$A_1\cup A_2\cup\cdots\cup A_n$ 表示事件 $A_1,A_2,\cdots,A_n$ 至少有一个发生,$\overline{A_1\cup A_2\cup\cdots\cup A_n}$表示它的否定,即不是 $A_1,A_2,\cdots,A_n$ 至少有一个发生,这等价于 $A_1,A_2,\cdots,A_n$ 都没发生,即 $\overline{A_1}\cap\overline{A_2}\cap\cdots\cap\overline{A_n}$.

1.1.3　概率及性质

对于随机现象,只考虑它的所有可能结果是没有什么意义的.我们所关心的是各种可能结果在一次试验中出现的可能性究竟有多大,从而就可以在数量上研究随机现象.

定义 1　**对于随机事件 A,若在 n 次试验中发生了 μ_n 次,则称比值 $\frac{\mu_n}{n}$ 为随机事件 A 在 n 次试验中发生的频率,记为 $f_n(A)$,即**

$$f_n(A)=\frac{\mu_n}{n},$$

其中 μ_n 称为频数.易知频率 $f_n(A)$ 具有以下性质:

(1)非负性　$0\leqslant f_n(A)\leqslant 1$;

(2)规范性　$f_n(\Omega)=1$;

(3)有限可加性　若事件 A,B 互不相容(即 $AB=\varnothing$),则

$$f_n(A\cup B)=f_n(A)+f_n(B).$$

例如,试验 E 为抛一枚质地均匀的硬币.若抛 20 次,硬币出现 11 次"正面向上"(记为事件 A),即 $n=20,\mu_n=11$.此时,事件 A 在 20 次试验中出现的频率为

$$f_{20}(A)=\frac{11}{20}=0.55;$$

① 德摩根(De Morgan,1806—1871 年),英国数学家、逻辑学家.

再重复进行 40 次试验,事件 A 出现的频数为 20,则

$$f_{40}(A)=\frac{20}{40}=0.5,$$

此即事件 A 在 40 次试验中发生的频率.

当然,我们还可以重复上千次、上万次的试验,分别记录事件 A 发生的频数,计算出其频率. 人们发现,尽管重复试验的次数不同,事件 A 发生的频数也各有差异,但其频率却稳定在一个固定的数值(0.5)左右,而且随着试验次数的增多,这种稳定性愈加明显. 为了验证这种频率的稳定性,历史上有不少统计学家曾做过"抛硬币"的试验,试验结果如下表.

实验者	抛硬币次数	出现正面次数	频率
蒲丰①	4 040	2 048	0.506 9
皮尔逊②	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

频率的稳定值的大小,反映了事件 A 发生的可能性的大小. 因此,可以给出下列定义:

定义 2 在大量重复进行同一试验时,随着试验次数 n 的无限增大,事件 A 发生的频率 $f_n(A)=\frac{\mu_n}{n}$ 会稳定在某一常数值附近. 这个常数值是随机事件 A 发生的可能性大小的度量,称为事件 A 的概率(Probability),记作 $P(A)$.

由于上面给出的概率定义是通过对频率的大量统计观测得到的,通常称为概率的统计定义. 这个定义虽然比较直观,但实际上我们不可能用它来计算事件的概率. 因为按照定义,要真正得到频率的稳定值,必须进行无穷多次试验,这显然是做不到的. 因此,我们还要另外寻找一些不用凭借试验就可以计算出事件发生的概率的方法.

概率论的基本研究课题之一就是寻求随机事件的概率. 我们先讨论一类最早被研究、也是最常见的随机试验. 这类随机试验具有下述特征:

(1)全部可能结果只有有限个;

(2)这些结果的发生是等可能的.

这种数学模型通常被称为**古典模型**(Classical Model of Probability).

由古典概型随机试验的特征可以看出,如果一个试验 E,在它的样本空间 Ω 中共有 n 个基本事件:$\omega_1,\omega_2,\cdots,\omega_n$,则每一基本事件在一次试验中发生的可能性都是$\frac{1}{n}$. 对任一随机事件 A 来说,如果 A 包含了其中的 k 个基本事件,则 A 发生的可能性应该是$\frac{1}{n}$的 k 倍,即$\frac{k}{n}$,所以事件 A 的概率

$$P(A)=\frac{k}{n}=\frac{A\text{ 包含的基本事件数}}{\Omega\text{ 中的基本事件总数}}=\frac{A\text{ 包含的样本点数}}{\Omega\text{ 中的样本点总数}}.$$

---

① 蒲丰(George-Louis Leclerc de Buffon,1707—1788 年),法国数学家.

② 卡尔·皮尔逊(Karl Pearson,1857—1936 年),英国统计学家,数理统计的奠基人.

这个定义称为**概率的古典定义**.

例 5 某种福利彩票的中奖号码由 3 位数字组成,每一位数字都可以是0～9中的任何一个数字,求中奖号码的 3 位数字全不相同的概率.

解 设事件 $A=\{$中奖号码的 3 位数字全不相同$\}$.

由于每一位数有 10 种选择,因此 3 位数共有 10^3 种选择,即基本事件总数为 10^3 个. 要 3 位数各不相同,相当于要从 10 个数字中任选 3 个做无重复的排列,共有 P_{10}^3 种选择,即 A 包含的基本事件数为 P_{10}^3 个. 因此,

$$P(A)=\frac{P_{10}^3}{10^3}=\frac{10\times9\times8}{1\ 000}=\frac{18}{25}.$$

例 6 掷一颗均匀的骰子,求出现偶数点的概率.

解 设 $\omega_i=\{$出现 i 点$\}(i=1,2,\cdots,6)$,则样本空间

$$\Omega=\{\omega_1,\omega_2,\cdots,\omega_6\}.$$

故基本事件总数为 6,令

$$A=\{\text{出现偶数点}\},$$

显然 $A=\{\omega_2,\omega_4,\omega_6\}$. 所以 A 中含有 3 个基本事件,从而

$$P(A)=\frac{3}{6}=\frac{1}{2}.$$

例 7(分房问题) 设有 n 个人,每个人都等可能地被分配到 N 个房间中的任意一间去住($n\leqslant N$),求下列事件的概率:

(1)指定的 n 个房间各有一个人住;

(2)恰好有 n 个房间,其中各有一个人住;

(3)某个指定的房间中有 $k(\leqslant n)$个人住.

解 因为每个人有 N 个房间可供选择,所以 n 个人住的方式共有 N^n 种.

(1)指定的 n 个房间各有一个人住,其可能的选择方式为 n 个人的全排列 $n!$,于是

$$P_1=\frac{n!}{N^n};$$

(2)n 个房间可以在 N 个房间中任意选取,共有 C_N^n 种方式,对选定的 n 个房间,由前述可知共有 $n!$ 种分配方式,所以恰有 n 个房间其中各有一个人住的概率为

$$P_2=\frac{C_N^n n!}{N^n}=\frac{N!}{N^n(N-n)!};$$

(3)k 个人可以在 n 个人中任意选择,共有 C_n^k 种方式,另外 $n-k$ 个人可以在其他房间中任意入住,共有$(N-1)^{n-k}$种方式,所以某指定的房间有 k 个人住的概率为

$$P_3=\frac{C_n^k(N-1)^{n-k}}{N^n}.$$

在统计力学中常把相空间分成数目很大的几个小区域或相格，因此，每个粒子总要落入一个相格中，这样整个系统的状态就可由粒子在相格中的分布确定. 如果把粒子看作是不可分辨的，那么上述例题的模型对应于玻色-爱因斯坦①(Bose-Einstein)统计；如果粒子是不可分辨物，并且每一个相格中最多只能放一个粒子，这就得到费米-狄拉克②(Fermi-Dirac)统计.

在古典概型中考虑的是试验结果为有限个的情形，这在实际应用中具有很大的局限性，还需要进一步考虑试验结果(即样本点数)有无穷多个的情形.

例如，试验 E 是向平面区域 Ω 内任意投点，点必须落在区域 Ω 内，并且落在 Ω 内任一点处都是等可能的. 这里“等可能”的含义是：设在区域 Ω 中有任意一个小区域 A，如果它的面积是 S_A，则点落入 A 中的可能性大小仅与 S_A 成正比，而与 A 的位置和形状无关. 于是，“点落入区域 A”这一事件的概率为

$$P=\frac{S_A}{S_\Omega}.$$

一般地，如果试验的样本空间含有无限多个样本点，但可以理解为一个可度量的几何图形，其度量值可以理解为一几何量(如长度、面积、体积等)，并且试验中任一随机事件 A 发生的概率与表示 A 的子区域的几何度量 μ_A 成正比，则事件 A 发生的概率为

$$P(A)=\frac{\mu_A}{\mu_\Omega}=\frac{\text{区域}A\text{的几何度量}}{\text{区域}\Omega\text{的几何度量}},$$

这个定义称为概率的几何定义. 这种类型的概率问题称为**几何概型**(Geomeffic Model of Probablity).

例 8(会面问题)　甲乙两人相约 7 点到 8 点在某地会面，先到者等候另一人 20 分钟，如果超过 20 分钟对方仍未到达就离去不再等候，试求这两人能会面的概率.

解　设甲于 7 点 x 分到达会面地点，乙于 7 点 y 分到达会面地点. 已知 $0\leqslant x\leqslant 60$，$0\leqslant y\leqslant 60$，所有可能的结果，即样本空间 Ω 可表示为

$$\{(x,y)\mid 0\leqslant x\leqslant 60, 0\leqslant y\leqslant 60\}.$$

它在平面直角坐标系中对应于一个边长为 60 的正方形，面积为 $S_\Omega=60^2$.

设 A 是甲乙两人能会面的事件，两人能会面的充分必要条件为 $|x-y|\leqslant 20$，所以它可表示为 $A=\{(x,y)\,\big|\,|x-y|\leqslant 20, 0\leqslant x\leqslant 60, 0\leqslant y\leqslant 60\}$. 与 A 对应的区域即图 1-1 中用阴影标出的部分，它的面积等于 $S_A=60^2-40^2$.

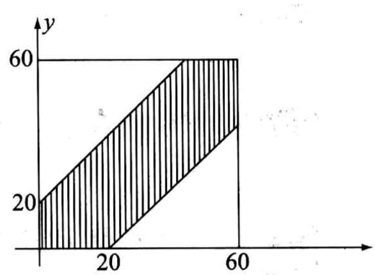

图 1-1

① 爱因斯坦(Albert Einstein，1879—1955 年)，生于德国，1933 年加入美国国籍. 爱因斯坦是 20 世纪最伟大的科学家之一，由于其对光电定律和理论物理方面的贡献被授予 1921 年诺贝尔物理学奖. 但其最重要的贡献是著名的爱因斯坦相对论.

② 保罗狄拉克(Paul Adrie Maurice Dirac，1902—1984 年)，英国理论物理学家量子力学奠基人之一，因狄拉克方程获 1933 年诺贝尔奖.

根据几何概率的定义,所求概率为

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

前面我们针对不同的问题,分别讨论了概率的频率定义,古典定义和几何定义及其计算方法. 可以看到,它们有一些共同的属性:非负性、规范性、有限可加性. 这些共同的属性为我们建立概率的公理化定义提供了理论基础.

下面我们引入柯尔莫哥洛夫①的概率公理化定义

定义 3　设试验的样本空间为 Ω,对试验的任一随机事件 A,定义实值函数 $P(A)$,如果它满足如下三个公理:

公理 1(非负性)　$P(A)\geqslant 0$;

公理 2(规范性)　$P(\Omega)=1$;

公理 3(可列可加性)　对于可列无穷多个互不相容的随机事件 $A_1,A_2,\cdots,A_n,\cdots$,有

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

则称 $P(A)$ 为事件 A 发生的概率.

从概率的三个公理出发,可以证明它的一些重要性质.

性质 1　不可能事件的概率是 0,即

$$P(\varnothing)=0.$$

性质 2　(有限可加性)若随机事件 $A_1,A_2,\cdots,A_n$ 互不相容,则

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

性质 3　对任一事件 A,有

$$P(A)=1-P(\overline{A}).$$

性质 4　对任意两个事件 A,B,有

$$P(A-B)=P(A)-P(AB).$$

性质 5　对任意的两个事件 A,B,若 $A\supset B$,则

$$P(A-B)=P(A)-P(B).$$

推论　若 $A\supset B$,则 $P(A)\geqslant P(B)$.

性质 6　对任意的两个事件 A,B,有

$$P(A\cup B)=P(A)\cup P(B)-P(AB).$$

利用数学归纳法,可以将性质 6 推广到任意有限个事件的情形.

① 柯尔莫哥洛夫(Колмогоров, A. H, 1903—1987 年),20 世纪最有影响的前苏联数学家,在函数论、信息论、动力系统和有限自动机等研究领域都有突出贡献,在 20 世纪世界数学家排名的问卷调查中位居第一. 1929 年发表的论文《测度的一般理论和概率论》建立了概率论的公理化体系.

推论　对任意 n 个事件 $A_1, A_2, \cdots, A_n$，有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leqslant i < j \leqslant n} P(A_i A_j) + \cdots + (-1)^{n-1} P(\prod_{i=1}^{n} A_i).$$

这个公式称为概率的一般加法公式(Additive Law of Probability)，经常使用$n=3$时的公式

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC).$$

利用这些基本性质，可以方便地计算某些事件的概率.

例 9　在所有的两位数 10～99 中任取一个数，求：(1)这个数能被 2 但不能被 3 整除的概率；(2)这个数能被 2 或 3 整除的概率.

解　设 $A=\{$所取数能被 2 整除$\}$，$B=\{$所取数能被 3 整除$\}$，则事件 $A-B$ 表示所取的数能被 2 但不能被 3 整除；又 AB 表示既能被 2 又能被 3 整除，即能被 6 整除的数. 因为所有的 90 个两位数中，能被 2 整除的有 45 个，能被 3 整除的有 30 个，能被 6 整除的有 15 个，所以我们有

$$P(A) = \frac{45}{90}, P(B) = \frac{30}{90}, P(AB) = \frac{15}{90}.$$

于是有

$$(1) P(A-B) = P(A) - P(AB) = \frac{45}{90} - \frac{15}{90} = \frac{1}{3}.$$

$$(2) P(A \cup B) = P(A) + P(B) - P(AB) = \frac{45}{90} + \frac{30}{90} - \frac{15}{90} = \frac{2}{3}.$$

1.1.4　条件概率及独立性

在实际生活中，人们经常需要求解在特定前提条件下某个事件发生的概率问题. 例如，盒中有大小相同的 10 个球，其中 4 个红球，6 个白球，从盒中依次取出 2 个球，取后不放回. 设 A 为第二次取出的球为白球，根据抽签原理，$P(A)=\frac{6}{10}$. 如果已知第一次取出的球为红球，记为 B，在此前提下再求第二次取出的球为白球的概率，显然这个概率与无任何前提的事件 A 发生的概率是不同的，它与事件 A 和 B 都有关系，它是在 B 已经发生的条件下事件 A 发生的概率，记为 $P(A|B)$，它表示的是从 3 个红球和 6 个白球的盒中任取一球为白球的概率，即$P(A|B)=\frac{6}{9}$. 我们再来考察 $P(A|B)$与事件 A 与事件 B 的关系. 事实上，

$$P(AB) = \frac{4 \times 6}{10 \times 9}, \frac{P(AB)}{P(B)} = \frac{4 \times 6}{10 \times 9} / \frac{4}{10} = \frac{6}{9} = P(A|B).$$

定义 4　**设随机事件 B 的概率 $P(B)>0$，则在事件 B 发生的条件下，事件 A 发生的条件概率(Conditional Probability)$P(A|B)$为**

$$P(A|B) = \frac{P(AB)}{P(B)}.$$