普通高等教育"十一五"规划教材

概率论与数理统计

王熙照 编著

科学出版社
www.sciencep.com

普通高等教育“十一五”规划教材

概率论与数理统计

王熙照　编著

科 学 出 版 社
北 京

内 容 简 介

本书是高等学校非数学类专业的概率论与数理统计课程的教材. 全书共分 10 章,内容包括:随机事件与概率,随机变量及其分布,随机向量及其分布,数字特征,大数定律和中心极限定理,抽样分布,参数估计,假设检验,方差分析与回归分析初步,Matlab 及其统计工具箱简介. 本书体系结构合理、层次分明,在阐述基本概念和基本思想的基础上,注重对学生基本技能和实际应用能力的训练和培养.

本书适用于普通高校理工类或经管类已具备微积分基础的本科生 36~72 学时教学,也可供相关专业教师和科技工作者参考.

图书在版编目(CIP)数据

概率论与数理统计/王熙照编著. —北京:科学出版社,2009
普通高等教育"十一五"规划教材
ISBN 978-7-03-024800-8

Ⅰ. 概… Ⅱ. 王… Ⅲ. ①概率论-高等学校-教材②数理统计-高等学校-教材 Ⅳ. O21

中国版本图书馆 CIP 数据核字(2009)第 099922 号

责任编辑:赵 靖 / 责任校对:包志虹
责任印制:张克忠 / 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

西源印刷厂 印刷
科学出版社发行 各地新华书店经销

*

2009 年 7 月第 一 版 开本:B5(720×1000)
2009 年 7 月第一次印刷 印张:15 1/4
印数:1—4 500 字数:300 000

定价: 23.00 元

(如有印装质量问题,我社负责调换〈环伟〉)

前　　言

概率论与数理统计是研究自然界、人类社会及技术过程中大量随机现象规律性的一个数学分支.概率论把随机现象抽象为随机变量去研究它一般的规律性,数理统计从收集、整理与分析实际问题中的随机数据出发,对问题作出推断、预测与决策.在很大程度上说,概率论是数理统计的基础,数理统计是概率论的应用,同时又对概率论的研究有很大的推动.

随着现代科学技术的迅速发展,这门学科已得到蓬勃发展,在自然科学、经济、人文、管理、工程技术等众多领域有越来越多的应用.例如,使用概率统计方法可以进行气象预报、水文预报以及地震预报;还可以进行产品的抽样检验,在研究新产品时,为寻求最佳生产方案可以进行试验设计和数据处理;在可靠性工程中,使用概率统计方法可以给出元件或系统的使用可靠性以及平均寿命的估计;在自动控制中,可以通过建立数学模型以便通过计算机控制工业生产,在通信工程中可用以提高抗干扰能力和分辨率等.

若干年来,笔者曾多次担任河北大学数学类和理工类本科生概率论与数理统计课程的教学工作.多年的教学经历使笔者深深地体会到,一本合适的教材对于学生的学习往往会起到事半功倍的效果.但是,一本教材,哪些地方应该详细些,哪些地方应该简略些,又是很难把握的.概率论与数理统计的内容涉及实变函数与测度论,如果从测度论的角度系统地去阐述,非数学类专业的读者将很难理解.本书主要针对非数学类专业的读者,力求在避免出现很专业的数学知识的基础上,给出基本概念的严格定义,给出大多数定理的严格证明.

下面就编写过程中的几点考虑作些说明:

(1) 编写过程中,力求做到“由浅入深,循序渐进,推理严格且简明”,希望读者在掌握基本概念和基本内容的同时,更重要的是掌握解决问题的方法;

(2) 在力求突出重点、加强理论的同时,注意运算技能的培养.书中附有大量的例题和习题,一些有实际应用背景的例题和习题能够增强学生的应用意识,提高解决问题的能力.

(3) 在内容的安排上,严格根据大纲编排内容,分别是随机事件及概率、随机变量及分布、随机向量及分布、数字特征、大数定律和中心极限定理、抽样分布、参数估计、假设检验、方差分析和回归分析.

本书适用于理工类或经管类已具备微积分基础的本科生36～72学时教学.针对不同层次的要求,任课教师可根据学时的限制等因素,酌情选择标注 * 的章节.

第 10 章是 Matlab 及其统计工具箱的介绍，作为进一步学习数值概率统计的基础.书末附有常用的统计分布表及各章习题的答案和提示，供读者参考.

本书在河北大学《概率论与数理统计讲义》的基础上，参考国内外一些同类工作编著而成.已在部分专业试用 3 遍.全书最终稿是与讲授概率统计的教师张群峰、张玉芬、陈爱霞、李杰、赵丽、张辉、董春茹反复讨论修改形成的，在此对他们表示感谢.

由于时间仓促，更受到教学水平和教学经验的限制，书中的不妥和疏漏之处在所难免，恳请广大读者批评指正.

王熙照

2009 年 2 月于河北大学

目　录

第 1 章　随机事件与概率

在自然界、人类社会及科学研究中，大量存在着这样一类现象：在对它们进行观察或试验时，有多种可能结果，且事先不能预知哪一个结果会发生. 例如，掷一枚硬币，可能出现正面或反面；观察某妇产科新生婴儿的性别，可能为男或女；多次测量某圆形零件的直径，可能得到不同的数值；检查一匹布上的疵点数，结果可能是 $\{0,1,2,\cdots\}$ 中的某一个，这类现象称为**随机现象**.

虽然随机现象具有偶然性和不确定性，但在对它们进行大量重复观察或试验时，随机现象会呈现出某种固有的规律性. 例如，在多次重复掷一枚硬币的试验中，出现正面和反面的次数大约各占一半；持续大量地观察某地新生儿的性别，会发现新生儿男女比例基本固定；多次测量某个零件的直径，结果会稳定在某个数值附近等. 这种在大量重复试验或观察中呈现出的固有规律性称为**统计规律性**.

概率论和数理统计是研究随机现象统计规律性的一门数学学科. 为了能够运用数学工具对随机现象进行分析，首先应该选择适当的数学语言描述随机现象，建立恰当的数学模型，为进一步研究随机现象打下基础. 本章主要介绍概率论的基本概念——随机事件及其概率.

1.1　随 机 事 件

1.1.1　随机事件

研究随机现象统计规律性的基本方法是对随机现象进行大量的重复试验或观察，这种观察或试验称为**随机试验**. 例如，掷一枚硬币，观察出现正面还是反面；掷一颗骰子，观察向上的点数；测量某个零件的直径；从一批产品中任取一件，观察其是合格品还是次品等.

随机试验具有以下三个特点：

(1) 在相同的条件下可以重复进行；

(2) 每次试验的可能结果不止一个，并且事先知道试验的所有可能结果；

(3) 每次试验之前不能预知哪一个结果会出现.

通过随机试验来了解随机现象，首先应明确随机试验的所有可能结果. 把随机试验 E 的每个可能结果称为该随机试验的**样本点**，通常用 ω 表示；把所有样本点的集合称为随机试验的**样本空间**，通常用符号 Ω 表示.

例 1　在掷一枚硬币的试验中，有两个可能的结果：出现正面和反面. 令 H 表示出现正面，T 表示出现反面，则样本空间为 $\Omega=\{H,T\}$.

例 2　掷一枚骰子，观察出现的点数，所有可能的结果有 6 个，若用 $i(i=1,2,\cdots,6)$表示出现 i 点，则样本空间为 $\Omega=\{1,2,3,4,5,6\}$.

例 3　检测某液晶显示器上暗点的个数，所有可能的结果有可列无穷多个，则该试验的样本空间为 $\Omega=\{0,1,2,\cdots\}$.

例 4　观察某人向半径为 15cm 的圆形靶子射击的弹着点的位置. 假定没有脱靶的情形，以靶心为原点建立平面直角坐标系，则样本空间为 $\Omega=\{(x,y)\mid x^2+y^2\leqslant 15^2\}$.

例 5　测试某个灯泡的使用寿命，则样本空间为 $\Omega=\{t\mid t\in\mathbf{R},t\geqslant 0\}$.

在研究随机现象时，我们常会关心满足某种条件的样本点的集合. 例如，在检测液晶显示器暗点个数的试验中，常会关心暗点数是否超过某个规定的数目；在两个人掷一颗骰子比点数大小的游戏中，后掷的人关心自己掷出的点数是否超过前面的人掷出的点数等.

随机试验 E 的样本空间 Ω 的子集称为**随机事件**，简称**事件**，通常用 $A,B,C,\cdots$表示. 在一次随机试验中，当且仅当某事件的一个样本点出现时，称该**事件发生**. 例如，

在例 1 的随机试验中，$A=\{H\}$为一个事件，表示出现正面；

在例 2 的随机试验中，$A=\{1,3,5\}$为一个事件，表示掷出了奇数点；

在例 3 中，$A=\{0,1,\cdots,10\}$为一个事件，表示暗点数不超过 10 个；

在例 4 中，$A=\{(x,y)\mid x^2+y^2\leqslant 2^2\}$为一事件，表示弹着点距靶心不超过 2cm.

由单个样本点构成的单点集，称为**基本事件**；在每次试验中都发生的事件称为**必然事件**；在每次试验中都不发生的事件称为**不可能事件**. 在例 1 的试验中有两个基本事件：$\{H\}$和$\{K\}$；在例 2 的试验中有 6 个基本事件：$\{1\},\{2\},\cdots,\{6\}$. 显然 Ω 是必然事件，$\varnothing$是不可能事件.

1.1.2　随机事件的关系和运算

同一个随机试验的事件之间存在着一定的关系.

例如，在例 2 的掷骰子试验中，设事件 $A=\{2,4\}$，$B=\{2,4,6\}$，$C=\{2,4,5\}$，$D=\{1,3,5\}$，则当事件 A 发生时，事件 B 也发生；B 和 C 都发生，意味着 A 发生；B 和 D 不能同时发生等.

一般地，在一个试验中，若事件 A 的发生导致事件 B 发生，则称事件 A **包含于** B 或 B **包含** A，记为 $A\subset B$. 此时，作为样本点的集合，A 是 B 的子集. 若 $A\subset B$ 且 $B\subset A$，则称事件 A 与 B **相等**，记为 $A=B$.

由事件 A 和事件 B 的所有样本点组成的事件，称为 A,B 的**并**或**和**，记为 $A\cup B$ 或 $A+B$. 作为样本点的集合，$A\cup B$ 是 A,B 的并集. $A\cup B$ 发生当且仅当事件 A

和事件B至少有一个发生.

由事件A和事件B的公共样本点组成的事件称为A,B的**交**或**积**,记为$A\cap B$或AB. 作为样本点的集合,$A\cap B$是A,B的交集. $A\cap B$发生当且仅当事件A和事件B同时发生.

两个事件的并和交可以推广到任意有限多个事件及可列多个事件的情形.

由属于A而不属于B的样本点组成的事件,称为A与B的**差**,记为$A-B$. 作为样本点的集合,$A-B$是A与B的差集. $A-B$发生当且仅当A发生而B不发生.

若两个事件不能同时发生,即$AB=\varnothing$,则称事件A和B**互不相容**或**互斥**. 如上例中,B和D互不相容. 若一组事件中的任何两个事件互不相容,则称这组事件互不相容.

若$AB=\varnothing$且$A\cup B=\Omega$,则称事件A和B互为**对立事件**,A的对立事件记为$\overline{A}$. 如上例中,$\overline{B}=D$.

事件的并、交等运算满足以下运算律:

(1) 交换律 $A\cup B=B\cup A, A\cap B=B\cap A$;

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

$(A\cap B)\cap C=A\cap(B\cap C)$;

(3) 分配律 $(A\cup B)\cap C=(A\cup C)\cap(B\cup C)$,

$(A\cap B)\cup C=(A\cap C)\cup(B\cap C)$;

(4) 对偶律 $\overline{A\cup B}=\overline{A}\cap\overline{B}, \overline{A\cap B}=\overline{A}\cup\overline{B}$.

例 6 设A,B,C为某一随机试验的事件,利用它们表示以下事件:

(1) A发生,但B,C都不发生;

(2) 三个事件当中至少有一个发生;

(3) 三个事件中恰有一个发生.

解 (1) $A\overline{B}\overline{C}$. (2) $A\cup B\cup C$. (3) $A\overline{B}\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\overline{B}C$.

例 7 将一枚硬币连掷两次,设H表示出现正面,T表示出现反面,写出该试验的样本空间. 若事件A为"至少出现一次正面",B为"第二次出现反面",试用集合表示事件$A,B,A\cup B,A\cap B,\overline{A}$和$A-B$.

解 样本空间$\Omega=\{HH,HT,TH,TT\}$,$A=\{HT,TH,HH\}$,$B=\{HT,TT\}$,$A\cup B=\{HT,TH,HH,TT\}=\Omega$,$A\cap B=\{HT\}$,$\overline{A}=\{TT\}$,$A-B=\{TH,HH\}$.

1.2 事件的概率

随机事件发生可能性的大小是可以度量的,如从一批产品中随机抽取一件,抽到次品的可能性大小由该批产品的次品率决定;统计资料表明,一些国家新生婴儿

为男孩的可能性略大于 1/2 等. 一般来说,在随机试验中,事件 A 发生的可能性大小称为该事件的**概率**,用 $P(A)$ 表示. 本节首先介绍概率的统计定义、古典概型和几何概型,然后引入概率的公理化定义.

1.2.1 概率的统计定义

设在相同的条件下,进行了 n 次试验. 若事件 A 发生了 n_A 次,则称 $f_n(A)=n_A/n$ 为事件 A 的**频率**. 在大量重复试验中,事件的频率稳定于一个固定的常数,这种性质称为频率的稳定性. 历史上有人做过掷硬币的试验,所得数据如表 1.1 所示.

表 1.1

试验者	掷硬币次数	出现正面次数	频 率
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

由表 1.1 可以看出,随着试验次数的增加,事件"正面朝上"的频率稳定在 0.5 附近.

由事件频率的定义,可知频率具有以下基本性质:

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

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

(3) 若 $A_1,A_2,\cdots,A_k$ 互不相容,则

$$f_n(A_1\cup A_2\cup\cdots\cup A_k)=f_n(A_1)+f_n(A_2)+\cdots+f_n(A_k).$$

在大量重复试验中,事件频率的稳定值作为事件的概率. 这种定义称为概率的统计定义. 按照概率的统计定义,可以认为"正面向上"这个事件的概率为 0.5.

事件频率的稳定性可以由第 5 章的大数定律给予理论上的解释.

1.2.2 古典概型

一个随机试验,若具有以下两个特点,则称为**古典概型**.

(1) 有限性:试验的样本空间为有限集合;

(2) 等可能性:样本空间中每个样本点发生的可能性都相等.

古典概型是概率论历史上最先研究的概率模型.

在古典概型中,设样本空间包含的样本点的个数为 $|\Omega|=n$,事件 A 包含的样本点个数为 $|A|=m$,则事件 A 的概率为

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

例 1 掷一颗质地均匀的骰子,则样本空间包含的样本点个数为 $n=6$. 设 $A=$

$\{2,4,6\}$，则 $P(A)=3/6=0.5$.

例 2 甲、乙两人玩游戏，每人掷 2 颗骰子，比掷得的点数之和，多者为胜. 若甲先掷，第一次和第二次分别掷得 5 和 4，则乙获胜的概率有多大？

解 样本空间 $\Omega=\{(1,1),(1,2),\cdots,(1,6),\cdots,(6,6)\}$，样本点个数为 $n=6\times6=36$. 设 $A=$“乙获胜”，则 $A=\{(4,6),(5,5),(5,6),(6,4),(6,5),(6,6)\}$，它包含了 6 个样本点，故乙获胜的概率为

$$P(A)=\frac{6}{36}=\frac{1}{6}.$$

例 3（抓阄问题） 设在袋中放有 n 个颜色不同的小球，其中，仅有一个为红色. 现从袋中连续随机取球，每次取一个，取后不放回. 求第 $i(1\leqslant i\leqslant n)$ 次取到红球的概率.

解 设 A 表示第 $i(1\leqslant i\leqslant n)$ 次取到红球，将取到的球依次排列. 到第 i 次取完后，共有

$$n\times(n-1)\times\cdots\times(n-i+1)$$

种可能的排列，所求事件包含了其中的

$$(n-1)\times\cdots\times(n-i+1)\times1$$

种，则所求概率为

$$P(A)=\frac{(n-1)(n-2)\cdots(n-i+1)}{n(n-1)(n-2)\cdots(n-i+1)}=\frac{1}{n}.$$

可见，该概率与取到红球的次序无关.

例 4 设 N 件产品中恰有 M 件次品.

(1) 若一次从中任取 n 件，求其中恰有 $k(k\leqslant M)$ 件次品的概率；

(2) 若有放回的连续取 n 次，每次取一件，求取到的 n 件产品中恰有 $k(k\leqslant M)$ 件次品的概率.

解 设 $A=$“取得 k 件次品”.

(1) 无放回抽样：将每次取得的结果看成一个组合，则所有可能的结果有 C_N^n 种，所求事件包含了其中的 $C_M^kC_{N-M}^{n-k}$ 种. 因此所求事件的概率为

$$P(A)=\frac{C_M^kC_{N-M}^{n-k}}{C_N^n}.$$

(2) 有放回抽样：将 N 件产品编上不同的号，有放回的抽样可看成重复排列，共有 N^n 种，其中，恰好有 k 件次品的排列有 $C_N^kM^k(N-M)^{n-k}$ 种. 因此所求事件的概率为

$$P(A)=\frac{C_N^kM^k(N-M)^{n-k}}{N^n}=C_N^k\left(\frac{M}{N}\right)^k\left(1-\frac{M}{N}\right)^{n-k}.$$

例 5　某研究所研制出了一种能快速检测奶粉中是否含有三聚氰胺的方法. 在一次实验中,运用这种方法对 10 份样品进行了检测,都得到了正确的结果. 请问这种检测方法是否可靠?

解　假定这种检测方法不可靠,即对每份样品进行检测都有两个等可能的结果:对和错,分别记为 0 和 1,则对 10 份样品检测所得的结果构成一个以 0 或 1 为分量的 10 维向量. 因此,该实验的样本空间由 2^{10} 个 10 维向量构成. 而 10 次检测都正确对应其中一个向量,即$(1,1,\cdots,1)$,故对 10 份样品检测的结果都正确的概率为 $1/2^{10}=0.000977$.

人们在长期的实践中总结出所谓的"实际推断原理":概率很小的事件在一次实验中是不会发生的. 而现在小概率事件发生了,表明有理由怀疑所作假设的正确性,即可以认为这种方法有一定的可靠性,值得进一步研究.

1.2.3　几何概型

考虑下面的概率问题.

据历史资料记载,在某地 10000m^2 的地域内,有一面积为 200m^2 的古墓. 若在该地域内随机取一点进行勘探,问探到古墓的概率为多大? 因为随机选取一点进行勘探,我们只能认为探到古墓的概率为 $1/50=0.02$. 实际上,这里我们假定该试验具有某种与古典概型相似的等可能性,并利用几何方法给出了一种合理的答案.

一般来说,若一个随机试验的可能结果可用几何上的点表示,并且所有可能结果对应的点充满了直线上的某线段,或平面上的某个区域,或空间的某区域(统一用 Ω 表示). 设事件 A 包含的结果充满了其中的一个子集 S,并且任意一点落在度量相同的子区域内的可能性相同,则可定义 A 的概率为

$$P(A)=\frac{S\text{ 的测度}}{\Omega\text{ 的测度}},$$

这里的测度指的是线段的长度、平面区域的面积、空间区域的体积. 这种处理概率问题的模型称为**几何概型**.

例 6(约会问题)　甲乙二人都随机地在周日上午(8:00～12:00)到同一网吧上网 1h. 求他们相遇的概率.

解　由题设,每人在 8:00～12:00 间的任何时刻等可能到达,从 8:00 开始计时,用 x,y 分别表示他们到达的时刻,则样本空间为

$$\Omega=\{(x,y)\mid 0\leqslant x,y\leqslant 4\}.$$

由题意,样本点(x,y)等可能地落在 Ω 中. 用 A 表示两人相遇,则

$$A=\{(x,y)\mid |x-y|\leqslant 1,(x,y)\in\Omega\},$$

如图 1.1 所示，则由几何概型中概率的算法可知

$$P(A)=\frac{4^2-3^2}{4^2}=\frac{7}{16}.$$

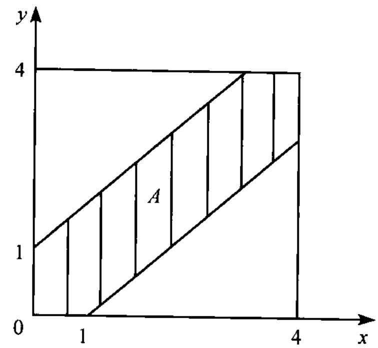

图 1.1

例 7(Buffon 投针问题) 桌面上画有一族平行线，相邻两条平行线间的距离都为 d. 现将一长度为 $l(l<d)$ 的针随机地投向该桌面，求针与其中一条平行线相交的概率.

解 由题设可知，必有一条直线离针的中点最近. 设针的中点离这条直线的距离为 x，针与直线的交角为 α(图 1.2)，则针在桌面的位置可用(x,φ)表示且 $0\leqslant x\leqslant d/2,0\leqslant\alpha\leqslant\pi$. 当且仅当 $x\leqslant(l/2)\sin\alpha$ 时，针与平行线相交(图 1.3).

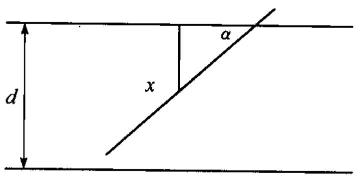

图 1.2

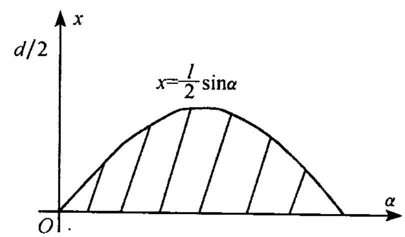

图 1.3

因此，该试验中随机点(x,α)等可能地落在矩形区域

$$\Omega=\left\{(x,\alpha)\mid 0\leqslant x\leqslant\frac{d}{2},0\leqslant\alpha\leqslant\pi\right\}$$

中，针与直线相交就是随机点落在图中阴影区域. 故所求概率为

$$p=\frac{\dfrac{l}{2}\displaystyle\int_0^{\pi}\sin\alpha\mathrm{d}\alpha}{\dfrac{1}{2}d\pi}=\frac{2l}{d\pi}.$$

1.2.4 概率的公理化定义

前面给出了三种不同的概率定义：概率的统计定义、古典概型中的概率定义和几何概型中的概率定义. 这三种定义在各自的情况下反映了概率的本质，即事件发生的可能性的大小. 但这三种定义都有理论上的缺陷. 在概率的统计定义中，概率是事件频率的稳定值，因此概率的确定带有主观性；古典概型的适用范围太小；几何概型中的概率定义可能带来歧义，如贝特朗(Bertrand)奇论(参见复旦大学李贤平《概率论基础》). 因此，如何给出概率的统一、严格的定义成了概率论发展史上的

重大问题.

1933 年，前苏联伟大的数学家柯尔莫哥洛夫(Kolmogrov)概括了前述几种概率定义的共同特征，以测度论为基础，提出了概率的公理化定义. 这个定义为概率论建立了坚实的数学基础，是概率论发展史上的一个里程碑.

为了讨论概率的公理化定义，首先给出事件域的概念.

定义 1.1　设 Ω 是某个随机试验的样本空间，$\mathscr{F}$是其某些子集构成的集合. 如果 $\mathscr{F}$满足以下条件：

(1) $\Omega\in\mathscr{F}$;

(2) 如果 $A\in\mathscr{F}$，则 $\overline{A}\in\mathscr{F}$;

(3) 如果 $A_i\in\mathscr{F}, i=1,2,\cdots$，则 $\bigcup\limits_{i=1}^{\infty}A_i\in\mathscr{F}$，

则称 $\mathscr{F}$为 Ω 的 σ **域或事件域**，称 $\mathscr{F}$中的元素为**事件**，称$(\Omega,\mathscr{F})$为**可测空间**.

定义 1.2(概率的公理化定义)　设$(\Omega,\mathscr{F})$为可测空间，P 是定义在事件域 $\mathscr{F}$上的集合函数. 若 P 满足以下条件：

(1) 非负性　$\forall A\in\mathscr{F}, P(A)\geqslant 0$;

(2) 规范性　$P(\Omega)=1$;

(3) 可列可加性　如果 $A_i\in\mathscr{F}, i=1,2,\cdots$且两两互不相容，则

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i),$$

则称 P 为 $\mathscr{F}$上的**概率测度**，$P(A)$称为 A 的**概率**，$(\Omega,\mathscr{F},P)$称为**概率空间**.

可以验证概率的统计定义、古典定义、几何定义均满足上述三条公理.

通常情况下，可以认为我们感兴趣的样本空间的子集都是事件，而概率是定义在事件集合上的满足非负性、规范性和可列可加性的集合函数.

1.2.5　概率的性质

利用概率的公理化定义，可以推得概率的一些常用的性质.

性质 1.1　不可能事件的概率为 0，即 $P(\varnothing)=0$.

证明　因为

$$\Omega=\Omega\cup\varnothing\cup\varnothing\cup\cdots,$$

所以

$$P(\Omega)=P(\Omega)+P(\varnothing)+P(\varnothing)+\cdots.$$

由概率的非负性及 $P(\Omega)=1$ 可得

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

性质 1.2　概率具有有限可加性，即若 $A_1,A_2,\cdots,A_n$ 互不相容，则

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

证明 因为

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} A_i \cup \varnothing \cup \varnothing \cup \cdots,$$

由概率的可列可加性及性质 1.1 可得

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

性质 1.3 对于任意事件 A 有 $P(\overline{A})=1-P(A)$.

证明 因为

$$A \cup \overline{A} = \Omega, \quad A \cap \overline{A} = \varnothing,$$

所以

$$1 = P(\Omega) = P(A) + P(\overline{A}),$$

移项即得所证.

性质 1.4 设 A,B 为两个事件,若 $A \supset B$,则

$$P(A-B) = P(A) - P(B), \quad P(A) \geqslant P(B).$$

证明 因为当 $A \supset B$ 时,有

$$A = B \cup (A-B), \quad B \cap (A-B) = \varnothing,$$

所以

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

再由概率的非负性得

$$P(A) \geqslant P(B).$$

当 $A \supset B$ 时,有 $P(A) \geqslant P(B)$,这称为概率的**单调性**.

性质 1.5 概率的加法公式

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

证明 因为

$$A \cup B = A \cup (B - A \cap B), \quad A \cap (B - A \cap B) = \varnothing,$$

所以

$$P(A \cup B) = P(A) + P(B - A \cap B).$$

而 $B \supset (A \cap B)$,故由性质 1.4 得

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

因此

$$P(A\cup B)=P(A)+P(B)-P(A\cap B).$$

特别地，当 A, B 互不相容时有

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

一般地，对任意 n 个事件 $A_1, A_2, \cdots, A_n$，可以用数学归纳法证得

$$P(A_1\cup A_2\cdots\cup A_n)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n).$$

例 8　袋中有 5 只红球、4 只白球，从中任取 3 球，求其中至少有一只红球的概率.

解　设 A_i 表示恰有 i 只红球，$i=1,2,3$，则 $A_i(i=1,2,3)$ 互不相容. 因此，所求事件的概率为

$$\begin{aligned}P(A_1\cup A_2\cup A_3)&=P(A_1)+P(A_2)+P(A_3)\\&=\frac{\mathrm{C}_5^1\mathrm{C}_4^2}{\mathrm{C}_9^3}+\frac{\mathrm{C}_5^2\mathrm{C}_4^1}{\mathrm{C}_9^3}+\frac{\mathrm{C}_5^3\mathrm{C}_4^0}{\mathrm{C}_9^3}\\&=\frac{30}{84}+\frac{40}{84}+\frac{10}{84}=\frac{20}{21}.\end{aligned}$$

另解　设 B 表示取到的全是白球，则 B 是 $A_1\cup A_2\cup A_3$ 的对立事件，故有

$$P(A_1\cup A_2\cup A_3)=1-P(B)=1-\frac{\mathrm{C}_4^3}{\mathrm{C}_9^3}=1-\frac{4}{84}=\frac{20}{21}.$$

1.3　条件概率与乘法公式

1.3.1　条件概率

我们知道，随机试验总是在一定条件下进行的. 在前面所讨论的概率问题中，除了这些固定的条件外，没有考虑其他的信息. 但是在有些问题中，常常需要计算在已知某一事件发生的条件下某些事件的概率. 例如，在一项癌症检查中，某人的血样呈阳性，求他确实患有癌症的概率；一个工厂有两个车间生产同一种产品，若在抽样检查中抽到一件次品，求该产品是第一车间生产的概率；等等. 像这种在已知某事件发生的条件下的概率，称为条件概率.

先看一个简单的例子，我们将从中引出条件概率的算式，并用该算式作为条件概率的定义式.