● 大学公共数学系列教材

概率论与数理统计

第二版

武汉大学数学与统计学院
齐民友 主编
刘禄勤 王文祥 龚小庆 编

高等教育出版社
HIGHER EDUCATION PRESS

概率论与数理统计

武汉大学数学与统计学院

齐民友 主编

刘禄勤 王文祥 龚小庆 编

高等教育出版社·北京

HIGHER EDUCATION PRESS BEIJING

内容简介

本书在大学公共数学系列教材《概率论与数理统计》(2002年第一版)的基础上修订而成。主要改动有：重写了第一章，改写了第五章、第七章和第八章，其余各章的部分内容作了改动和增删，习题作了较多调整和补充，更正了第一版中的错误。全书有如下特色：

1. 坚持数学理论的完整性和严谨性，并尽量阐述其实际意义；

2. 重点放在对基本概念的准确理解、对常用方法的熟练掌握上；

3. 坚持理论与实际相结合的原则，详细叙述了概率统计中主要概念和方法产生的背景和思路，注重培养学生对随机现象的理解和概率统计直觉；

4. 在保持传统体系和经典内容的同时，注意渗透和吸收现代概率统计新的思想、概念和方法；

5. 本书内容紧扣全国硕士研究生入学统一考试“数学一”和“数学三”的考试大纲，不仅有针对性地在例题和习题中收录了考研的各种题型，而且在书末给出了一个附录，提供了几套考研客观试题。

本书可作为高等学校理科(非数学、非统计学专业)、工科、经济、管理等各专业概率论与数理统计课程的教材。

图书在版编目(CIP)数据

概率论与数理统计/齐民友主编；刘禄勤，王文祥，龚小庆编.
—2版.—北京：高等教育出版社，2011.8
(大学公共数学系列教材)
ISBN 978-7-04-032516-4
Ⅰ.①概… Ⅱ.①齐…②刘…③王…④龚… Ⅲ.①概率论-高等学校-教材②数理统计-高等学校-教材 Ⅳ.①O21
中国版本图书馆CIP数据核字(2011)第129023号

策划编辑 兰莹莹　责任编辑 兰莹莹　封面设计 于文燕　版式设计 杜微言
插图绘制 尹文军　责任校对 杨雪莲　责任印制 张泽业

出版发行	高等教育出版社	网　址	http://www.hep.edu.cn
社　址	北京市西城区德外大街4号		http://www.hep.com.cn
邮政编码	100120	网上订购	http://www.landraco.com
印　刷	北京机工印刷厂		http://www.landraco.com.cn
开　本	787mm × 960mm　1/16		
印　张	20.5	版　次	2002年8月第1版
字　数	380千字		2011年8月第2版
购书热线	010-58581118	印　次	2011年8月第1次印刷
咨询电话	400-810-0598	定　价	28.00元

物 料 号　32516-00

第二版前言

本书在大学公共数学系列教材《概率论与数理统计》(2002 年第一版) 的基础上修订而成, 可作为高等学校理科 (非数学、非统计学专业)、工科、经济、管理等各专业概率论与数理统计课程的教材, 适用于全国硕士研究生入学统一考试数学一和数学三所涵盖的各专业。

本次再版修订的主要依据是本书第一版使用过程中师生们反馈的意见和建议。主要改动有: 重写了第一章, 改写了第五章、第七章和第八章, 第二章至第四章和第六章部分内容作了调整和增删, 各章习题作了调整和补充, 更正了第一版中的错误。为方便读者使用, 书末常见分布表的内容也作了扩充。尽管改动较大, 但仍保持了第一版的特色。

本书第一章至第五章由刘禄勤执笔, 第六章和第七章由王文祥执笔, 第八章由龚小庆 (现工作单位为浙江工商大学) 执笔; 各章习题由龚小庆编写; 附录 1 客观题由王文祥编写。全书由刘禄勤修改定稿。

感谢使用本书第一版的师生们, 他们为本次修订提供了很多好的意见和建议, 感谢武汉大学数学与统计学院和高等教育出版社对本书再版的关心和支持, 感谢樊启斌副院长和兰莹莹编辑给予的帮助和指导。在编写过程中, 我们参考或引用了部分优秀概率统计教材 (见本书参考文献) 中的一些例题和习题, 谨此致谢!

由于编者水平有限, 书中错误之处在所难免, 恳请读者指正。

编者

2011 年 2 月于武汉大学

第一版前言

本书是我们在武汉大学多年来的教学实践的基础上,并参照全国工学、经济学硕士研究生入学统一考试数学(一)和数学(三)对概率论和数理统计部分的基本要求编写的,可作为高等学校理工科(非数学专业)、金融管理等各专业的概率论与数理统计课程的教材,也可作为报考硕士研究生人员和实际工作者的参考书。

概率论与数理统计是研究随机现象的统计规律的一门学科,它有别于数学其他分支的重要的一点在于,初学者往往对一些重要的概率统计概念的实质感到疑惑不解.考虑到这个原因以及概率统计应用性很强的特点,我们在取材与写作上,在如下几个方面做了努力:

1) 用较多的篇幅详细地叙述了概率统计中一些主要概念和方法产生的背景和思路,从直观分析入手逐步过渡到严格的数学表述。

2) 坚持数学理论的完整性和严谨性,对基本的概念、定理和公式作严格、准确、规范的叙述,并尽量阐述其实际意义。

3) 重点放在对基本概念的准确理解、对常用方法的熟练掌握上、对那些证明起来较为困难的结论(需要较高深的数学理论),尽可能给出一个虽然不太严格,但有助于学生领会证明思路的形式上的“证明”,或是尽可能在一些特殊条件下加以论证。这样既使学生在一定程度上“知其所以然”,又保证了教学重点的完成。

4) 坚持理论与实际相结合的原则,注重培养学生对随机现象的理解和概率统计直觉。为此,我们不仅从实例出发引入基本概念,还精选了大量能够加深理解基本概念、定理和公式的例题和习题,目的在于使学生对实际事物中的随机性生产敏感、培养学生的概率统计直觉能力。

5) 本书内容紧扣全国硕士研究生入学统一考试数学(一)和数学(三)的考试大纲。采用考试大纲规范的术语和符号,不仅有针对性地在例题和习题中收录了考研的各种题型,而且在书末给出了一个附录,提供了一些历年工学、经济学硕士研究生入学考试的客观题。对考研中的重点和难点内容,我们尽可能地进行了细致的处理。

6) 本书提供了较多的反例,帮助学生正确理解基本概念、准确掌握基本

结论。

7) 在保持传统体系和经典内容的同时，注意渗透和吸收现代概率统计新的思想、概念和方法，有些结果是第一次出现在此类教材中。

根据我们的教学经验，讲完本书除少数带 * 号以外的全部内容，大约需要 54 学时；如果只讲授前面五章 (概率论部分)，则只需 36 学时。为了便于学生自学，我们配备了较多的例题，教师可根据需要选择其中的一部分在课堂讲授。

本书由刘禄勤组织编写，大纲和体系由集体讨论而定。第一章由龚小庆执笔，第二、三、四、五章由刘禄勤执笔，第六、七、八章的初稿由王文祥执笔并由龚小庆修改；各章的习题及解答由龚小庆编写；附录的客观题及解答由王文祥编写。全书由刘禄勤修改定稿。

本丛书主编齐民友教授和概率统计系高付清教授仔细审阅了全书，提出了许多宝贵意见。在编写的过程中，我们参考了较多的有关文献，对于引用了其中例题或习题的书籍，我们均列入书末的参考文献中。本书的编写自始至终得到高等教育出版社和武汉大学数学与统计学院的大力支持，尤其是徐可编辑和陆君安教授给予了不少帮助。对此，我们一并表示衷心的感谢！

由于编者的水平有限，书中不当乃至错误之处在所难免，恳请读者不吝赐教。

编者

2002 年 7 月于武汉大学

目　录

第一章　随机事件与概率

在自然界和人类社会中存在着两类现象: 一类是确定性现象, 指的是在一定条件下必然会发生的现象; 另一类是随机现象, 其特点是在一定条件下可能出现这样的结果, 也可能出现那样的结果, 且在试验和观察之前, 不能预知确切的结果. 确定性现象的例子: 在一个标准大气压下, 纯净水加热到 100°C 必然沸腾; 在地球上向上抛出的重物必然下落. 微积分、线性代数等学科是研究确定性现象的数学工具. 随机性现象的例子: 抛掷一枚质地均匀的硬币, 硬币落地后有可能反面朝上, 也有可能正面朝上; 掷一颗骰子, 掷出的点数可能是 $1, 2, \cdots, 6$ 中的任何一个; 在没有政策变化等重大事件发生的条件下, 观察下一天上海股市的上证指数, 它可能会涨, 也可能会跌; 某人购买了一张彩票, 可能中奖, 也可能不中奖; 某射手向远处的靶子发射一发子弹, 可能击中目标, 也有可能脱靶. 随机现象在自然界和人类社会中广泛存在, 对它的研究有着重要的理论意义和应用价值.

随机现象虽然给人的感觉是“纯属偶然”、难以捉摸, 似乎没有规律可言, 但事实上人们发现很多随机现象依然存在着固有的规律性, 这种规律性体现在对同一随机现象的大量重复观察之中. 比如, 大量重复地抛掷一枚质地均匀的硬币, 会发现正面朝上的次数所占的比例会越来越接近于 50%. 下表列出了历史上一些科学家在抛掷硬币试验中得到的相关数据.

表 1.1　历史上一些著名的抛掷均匀硬币的试验

实验者	抛掷次数	正面朝上次数	正面朝上频率
德摩根	2 048	1 061	0.518 1
蒲丰	4 040	2 048	0.506 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5
维尼	30 000	14 994	0.499 8

从表中我们可以看到, 随着试验次数的增加, 正面朝上的次数所占的比例逐渐稳定于 50%. 这种在大量试验中呈现的稳定性, 我们将其称之为统计规律性. 概率论与数理统计正是揭示和研究随机现象统计规律的数学分支.

§1.1 随机事件

1.1.1 随机试验与事件

研究随机现象, 首先要对研究对象进行观察. 概率论中将对随机现象的观察或为观察随机现象而进行的实验称为**随机试验 (random experiment)**. 随机试验应具备以下三个特征:

1) 可在相同的条件下重复进行;

2) 试验的所有可能的结果不止一个, 并且是事先知道的;

3) 试验之前无法知道哪一个结果会出现.

随机试验常用字母 $\mathcal{E}$ 表示. 称随机试验的每一个可能的结果为**样本点 (sample point)**或**基本事件**, 常用字母 ω 表示; 称随机试验所有可能的结果组成的集合为**样本空间 (sample space)**, 一般用字母 Ω 表示. 下面是一些随机试验的例子:

$\mathcal{E}_1$: 抛一枚硬币, 观察正面、反面出现的情况. 样本空间 $\Omega_1 = \{H,T\}$, 其中 H 表示正面,T 表示反面.

$\mathcal{E}_2$: 将一枚硬币抛掷三次, 观察正面 H、反面 T 出现的情况. 样本空间为

$$\Omega_2 = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

$\mathcal{E}_3$: 掷一颗骰子, 观察出现的点数. 样本空间 $\Omega_3 = \{1,2,3,4,5,6\}$.

$\mathcal{E}_4$: 将一颗骰子抛掷两次, 观察出现点数的情况. 样本空间为

$$\Omega_4 = \{(i,j) : 1 \leqslant i,j \leqslant 6\}$$

$\mathcal{E}_5$: 观察某一天某城市共发生车祸的次数. 样本空间 $\Omega_5 = \{0,1,2,3,\cdots\}$.

$\mathcal{E}_6$: 在电视机厂仓库里任意抽取一台电视机, 测试它的寿命. 样本空间 $\Omega_6 = [0,\infty)$.

在随机试验中, 人们不仅关心某个样本点是否出现, 更关心满足某一特定条件的样本点是否出现. 例如, 在随机试验 $\mathcal{E}_5$ 中, 人们往往关心是否出现了“该天车祸数超过 20 起”这一事件, 或者关心是否出现了事件“该天车祸数不超过 3 起”. 满足某一特定条件的样本点组成了样本空间 Ω 的子集. 今后, 我们把样本空间 Ω 的某些子集 (满足某一特定条件的样本点或是具有某种特征的样本点组成的子集) 称为**随机事件**, 简称为**事件 (event)**. 以 $\mathcal{E}_5$ 为例, 事件“该天车祸数超过 20 起” $= \{21,22,23,\cdots\}$, 而事件“该天车祸数不超过 3 起” $= \{0,1,2,3\}$. 在 $\mathcal{E}_6$ 中, 如果电视机的寿命超过 10 000 h 被认为是合格品, 则事件“所抽取的电视机是合格品”可用 (10 000, $+\infty$) 这一子集来表示. 又如在

$\mathcal{E}_4$ 中, 事件“两次掷出的点数之和等于 4”可用子集 $\{(1,3),(2,2),(3,1)\}$ 来表示. 事件作为样本空间 Ω 的子集, 一般用大写的英文字母 $A,B,C,\cdots$ 来表示.

设 A 为一个事件, 如果在一次具体的试验中出现的样本点 ω 属于 $A(\omega \in A)$, 则称**事件 A**(在该次试验中)**发生**. 如在上述 $\mathcal{E}_4$ 中, 若骰子抛掷两次出现样本点 $(3,1)$, 则事件 $\{(1,3),(2,2),(3,1)\}$ 即“两次掷出的点数之和等于 4”在该次试验中发生; 若出现样本点 $(6,2)$, 则事件 $\{(1,3),(2,2),(3,1)\}$ 在该次试验中没有发生. 又如在上述随机试验 $\mathcal{E}_5$ 中, 若某天该城市共发生车祸 30 起, 即出现样本点 30, 则事件“该天车祸数超过 20 起”即 $\{21,22,23,\cdots\}$ 发生, 而事件“该天车祸数不超过 3 起”即 $\{0,1,2,3\}$ 没有发生.

任何随机试验都有两个特殊的事件, 一个是样本空间 Ω, 它在任何一次试验中均必然发生, 称为**必然事件**; 另一个是空集 $\varnothing$, 它在任何一次试验中均不发生, 称为**不可能事件**. 其实这两个事件都没有随机性, 只是为了方便, 我们把它们作为随机事件的两个极端包括到随机事件中来.

1.1.2 事件的关系与运算

既然事件是样本空间的子集合, 因此事件之间也有与集合之间一样的关系及运算,关键是要用概率论的语言来解释这些关系和运算. 以下设 $A,B,A_n\ (n \geqslant 1)$ 均是样本空间 Ω 中的事件.

1) **事件的包含与相等**. $A \subset B$ 或 $B \supset A$ 表示事件 A 被事件 B 包含或事件 B 包含了事件 A, 其概率论涵义是: 若事件 A 发生 (即 $\omega \in A$), 则事件 B 发生 (即 $\omega \in B$). 若 $A \subset B$ 且 $A \supset B$, 则称事件 A,B 相等, 记为 $A=B$.

2) **事件的并**. $A \cup B$ 表示“事件 A 与事件 B 中至少有一个发生”这一新事件. 这是因为:$A \cup B$ 发生 $\Leftrightarrow \omega \in A \cup B \Leftrightarrow \omega \in A$或 $\omega \in B \Leftrightarrow A,B$ 中至少有一个发生. 称 $A \cup B$ 为事件 A 和事件 B 的**和事件**. 类似地,$\bigcup\limits_{n=1}^{\infty} A_n$ 表示“事件 $A_1,\cdots,A_n\ (n \geqslant 1)$ 中至少有一个发生”这一新事件.

3) **事件的交**. $A \cap B$ 表示“事件 A 和事件 B 同时发生”这一新事件. 这是因为:$A \cap B$ 发生 $\Leftrightarrow \omega \in A \cap B \Leftrightarrow \omega \in A$且 $\omega \in B \Leftrightarrow A,B$ 同时发生. 称 $A \cap B$ 为事件 A 和事件 B 的**积事件**. 类似地,$\bigcap\limits_{n=1}^{\infty} A_n$ 表示“事件 $A_1,\cdots,A_n\ (n \geqslant 1)$ 同时发生”这一新事件. 通常我们简记 $A \cap B$ 为 AB.

4) **事件的差**. $A-B$ 表示“事件 A 发生但事件 B 不发生”这一新事件. 这是因为,$A-B$ 发生 $\Leftrightarrow \omega \in A-B \Leftrightarrow \omega \in A$但 $\omega \notin B \Leftrightarrow A$ 发生但 B 不发生. 称 $A-B$ 为事件 A 和事件 B 的**差事件**.

5) **逆事件或对立事件**. 记 $\overline{A}=\Omega-A$. $\overline{A}$ 表示“A 不发生”这一新事件. 这

是因为,$\overline{A}$ 发生 $\Leftrightarrow \omega \in \overline{A} \Leftrightarrow \omega \notin A \Leftrightarrow A$ 不发生. 称 $\overline{A}$ 为 A 的**逆事件**或 A 的**对立事件** (**complementary events**).

类似地,$A \cap B = \varnothing$ 或 $AB = \varnothing$ 表示“A 和 B 同时发生”是不可能事件,即表示 A 和 B 不可能同时发生;$A \cup B = \Omega$ 表示“A 和 B 中至少有一个发生”是必然事件, 即表示 A 和 B 中必然有一个发生. 若 $AB = \varnothing$, 则称事件 A 与事件 B **不相容**. 若多个事件 $A_1, \cdots, A_n, \cdots$ 两两不相容 (即 $i \neq j$ 时 $A_iA_j = \varnothing$),则称它们**互不相容**(**mutual exclusion**).

我们已经看到事件的运算实际上就是集合的运算, 因此, 事件的运算满足和集合的运算一样的运算规律:

1) 交换律:$A \cup B = B \cup A,\ \ AB = BA$;

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

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

4) 德摩根律:$\overline{A \cup B} = \overline{A} \cap \overline{B},\ \overline{AB} = \overline{A} \cup \overline{B},\ \overline{\bigcup\limits_{n=1}^{\infty} A_n} = \bigcap\limits_{n=1}^{\infty} \overline{A}_n,\ \overline{\bigcap\limits_{n=1}^{\infty} A_n} = \bigcup\limits_{n=1}^{\infty} \overline{A}_n$;

5) 不相容分解:$A \cup B = A \cup \overline{A}B,\ \ A = AB \cup A\overline{B}$.

其中的公式 4), 5) 是值得牢记的.

正确地用字母表示事件的关系与运算是相当重要的.

例 1.1.1. *设 A, B, C 是随机事件, 则*

“A 与 B 发生,C 不发生”可以表示成 $AB\overline{C}$;

“A, B, C 至少有两个发生”可以表示成 $AB \cup AC \cup BC$;

“A, B, C 中恰好发生两个”可以表示成 $AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC$;

“A, B, C 中至多一个发生”可以表示成 $\overline{A}\,\overline{B}\,\overline{C} \cup A\overline{B}\,\overline{C} \cup \overline{A}B\overline{C} \cup \overline{A}\,\overline{B}C$.

1.1.3　事件域

我们已经把事件理解为 Ω 的某些子集, 但是否能把 Ω 的一切子集都看作事件呢? 这倒不一定, 因为若考虑的“事件”太多, 会使问题变得很复杂而不便处理; 若考虑的“事件”太少, 则不能满足我们的需要. 所以, 把哪些子集视为事件, 要根据试验的目的来决定, 但这并不等于说一点规则都不要. 由前面的学习我们知道, 在一个试验中, 如果把 A, B 视为事件, 则自然要求把 $A \cup B$, AB, $A - B$ 等也视为事件, 也就是要求事件的并、交、差也为事件. 因此, 若把所有的事件放在一起, 用符号 $\mathscr{F}$ 表示, 则 $\mathscr{F}$ 应对集合的并、交、差运

算封闭.

现在, 我们可以抽象出“事件”的数学定义了.

定义 1.1.1. 设 Ω 为样本空间,$\mathscr{F}$ 为 Ω 的某些子集组成的集合, 如果它满足

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

2) $A \in \mathscr{F} \Rightarrow \overline{A} \in \mathscr{F}$;

3) $A_n \in \mathscr{F}, n \geqslant 1 \Rightarrow \bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{F}$.

则称 $\mathscr{F}$ 为 (Ω 上的)**事件域**或 **σ 域 (σ-field)**, 称 $\mathscr{F}$ 中的集合为**事件 (event)**.

可以证明, 事件域 $\mathscr{F}$ 中的有限或可列无穷多个集合经过有限或可列无穷次交、并、差的运算后所得之集合仍然在事件域 $\mathscr{F}$ 中, 也就是说事件的有限或可列无穷多次交、并、差还是事件.

最小的事件域是 $\{\varnothing, \Omega\}$, 最大的事件域是 Ω 的所有子集全体. 对 $A \subset \Omega$, 包含 A 的最小事件域为 $\{\varnothing, A, \overline{A}, \Omega\}$.

在理论学习与研究中, 事件域是事先给定的; 在应用上, 事件域则要根据试验的目的来确定. 若 Ω 是有限集合或可数无穷集合, 我们总取它的所有子集全体为事件域, 即**有限样本空间和可列无穷样本空间的每个子集均为事件**.

§1.2 频率与概率

1.2.1 频率的稳定性与概率的定义

随机事件在一次具体的随机试验中可能发生, 也可能不发生, 带有偶然性, 但随机事件发生的可能性的大小却是客观存在的. 例如, 在上节随机试验 $\mathcal{E}_1$ 中, 若硬币是均匀的, 从直观上容易看到出现正面和出现反面的可能性应该相等; 在上节随机试验 $\mathcal{E}_3$ 中, 若骰子是质量均匀的正六面体, 从直观上也容易看到出现 $1, 2, 3, 4, 5, 6$ 中任何一个的可能性都应该相同; 飞机坠机的可能性远比汽车发生车祸的可能性小, 这是众所周知的事实. 正如一根木棒有长度, 一块土地有面积一样, 一个随机事件发生的可能性的大小是它自身固有的属性. 我们用概率来表示这个“可能性的大小”, 也就是说概率是随机事件发生的可能性大小的数量化指标. 我们用 $P(A)$ 表示事件 A 的概率.

怎样确定一个事件的概率呢? 为了确定木棒的长度, 我们用带刻度的尺子去测量它, 用测量值作为木棒真实长度的近似值. 下面我们将看到, 为了确定一个事件的概率, 可通过试验的方法, 用频率这把“尺子”去“测量”它, 用频率作为事件真实概率的近似值.

定义 1.2.1. 设 Ω 为随机试验 $\mathcal{E}$ 的样本空间, 事件 $A \subset \Omega$. 在相同条件下将试验 $\mathcal{E}$ 重复做 n 次, 以 $n(A)$ 表示事件 A 在这 n 次试验中发生的次数, 称 $\frac{n(A)}{n}$ 为 A 在这 n 次试验中发生的频率 (**frequency**), 记为 $f_n(A)$, 即

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

由表 1.1 可知, 在抛掷均匀硬币的试验中, 随着试验次数的增加, 正面朝上的频率稳定在 $\frac{1}{2}$ 附近波动, 且波动幅度越来越小. 大量的实践表明, 只要在相同的条件下重复做同一个随机试验, 随着试验次数 n 的增加, 任何事件 A 发生的频率 $f_n(A)$ 都会稳定地在某一数值 p 左右波动, 且当 n 越来越大时, 这种波动的幅度一般会越来越小. 频率的波动中心 p 即是该事件的概率 $P(A)$. 频率的这种特性称为**频率的稳定性**, 它说明当试验次数趋于无穷大时, 一个事件的频率收敛到该事件的概率 (我们将在第五章中严格证明这一点). 频率的稳定性有着非常重要的意义, 一方面它说明了概率的客观存在性; 另一方面, 它启发我们: 既然频率收敛到概率, 概率就应该具有频率所具有的基本性质. 因频率具有如下性质:

1) 非负性:$f_n(A) \geqslant 0$;

2) 规范性:$f_n(\Omega) = 1$;

3) 可列可加性: 对任何可列无穷多个互不相容的事件 $A_1, \cdots, A_k, \cdots$,

$$f_n\left(\bigcup_{k=1}^{\infty} A_k\right) = \frac{n\left(\bigcup\limits_{k=1}^{\infty} A_k\right)}{n} = \frac{\sum\limits_{k=1}^{\infty} n(A_k)}{n} = \sum_{k=1}^{\infty} f_n(A_k)$$

所以, 概率也应具有上述三条性质.

现在, 我们可以提炼出"概率"的公理化定义了.

定义 1.2.2. 设 Ω 为样本空间,$\mathscr{F}$ 为 Ω 上的事件域 ,P 为定义在 $\mathscr{F}$ 上的实值函数 . 如果 P 满足

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 : 对 $\mathscr{F}$ 中任何可列无穷多个互不相容的事件 $A_1, \cdots, A_k, \cdots$,

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k) \tag{1.2.2}$$

则称 P 为 $\mathscr{F}$ 上的**概率测度** , 简称为**概率 (probability)**; 称 $(\Omega, \mathscr{F}, P)$ 为**概率空间 (probability space)**; 称 $P(A)$ 为事件 A 的概率.

下面我们给出概率空间的几个简单例子.

例 1.2.1. 考虑上节随机试验 $\mathcal{E}_3$, 样本空间为 $\Omega=\{1,2,3,4,5,6\}$, 事件域 $\mathscr{F}$ 取为 Ω 的所有子集全体. 若骰子是均匀的, 令 $P(A)=\dfrac{A\text{中元素的个数}}{6}, A\in\mathscr{F}$. 则易见 $(\Omega,\mathscr{F},P)$ 为概率空间. 一般地, 若样本空间 $\Omega=\{\omega_1,\cdots,\omega_N\}$ 由有限多个样本点构成, 取事件域 $\mathscr{F}$ 为 Ω 的所有子集全体, 令 $P(A)=\dfrac{A\text{中元素的个数}}{N}, A\in\mathscr{F}$, 则 $(\Omega,\mathscr{F},P)$ 为概率空间. 在这个概率空间中, 每个基本事件等可能发生. 这样的概率空间我们将在下一节专门研究.

例 1.2.2. 考虑上节随机试验 $\mathcal{E}_5$, 样本空间为 $\Omega=\{0,1,2,\cdots,k,\cdots\}$, 事件域 $\mathscr{F}$ 取为 Ω 的所有子集全体. 设 $\{a_k,k\geqslant 0\}$ 为非负实数序列, 满足 $\sum\limits_{k=0}^{\infty}a_k=1$. 令

$$P(A)=\sum_{k\in A}a_k,\ A\in\mathscr{F}$$

上式中求和是对所有下标属于 A 的那些 a_k 求和, 比如当 $A=\{2,3,6,9\}$ 时,$P(A)=a_2+a_3+a_6+a_9$, 则 $(\Omega,\mathscr{F},P)$ 为概率空间. 在这个概率空间中, 基本事件发生的概率为 $P(\{k\})=a_k,\ k\geqslant 0$.

例 1.2.3. 设样本空间 $\Omega=[0,1]$, 事件域 $\mathscr{F}$ 由 $[0,1]$ 中所有可求长 (即具有长度) 的子集构成. 令 $P(A)$ 等于 A 的长度 ,$A\in\mathscr{F}$. 则易见 $(\Omega,\mathscr{F},P)$ 为概率空间. 类似地, 若样本空间 $\Omega=[0,1]\times[0,1]$ 为平面中单位正方形, 事件域 $\mathscr{F}$ 由 $[0,1]\times[0,1]$ 中所有具有面积的子集构成. 令 $P(A)$ 等于 A 的面积,$A\in\mathscr{F}$. 则易见 $(\Omega,\mathscr{F},P)$ 也为概率空间. 这样的概率空间我们也将在下一节专门研究.

1.2.2 概率的基本性质

下面我们介绍概率的一些基本性质, 它们都是从概率定义中的三条公理推导出来的, 具有基本的重要性.

定理 1.2.1. 设 $(\Omega,\mathscr{F},P)$ 为概率空间. 概率 P 具有如下性质 :

1) 不可能事件的概率为 $0: P(\varnothing)=0$;

2) 有限可加性 : 对 $\mathscr{F}$ 中任何有限多个互不相容的事件 $A_1,\cdots,A_n$,

$$P\left(\bigcup_{k=1}^{n}A_k\right)=\sum_{k=1}^{n}P(A_k) \tag{1.2.3}$$

3) 可减性：若 $A,B\in\mathscr{F}$, $A\subset B$, 则 $P(B-A)=P(B)-P(A)$;
4) 单调性：若 $A,B\in\mathscr{F}$, $A\subset B$, 则 $P(A)\leqslant P(B)$;
5) 逆事件概率公式：若 $A\in\mathscr{F}$, 则 $P(\overline{A})=1-P(A)$.

证. 1) 令 $A_1=\Omega, A_n=\varnothing\,(n\geqslant 2)$, 由概率的定义, $1=P(\Omega)=P\left(\bigcup\limits_{n=1}^{\infty}A_n\right)=\sum\limits_{n=1}^{\infty}P(A_n)=1+\sum\limits_{n=2}^{\infty}P(\varnothing)$, 故必有 $P(\varnothing)=0$.

2) 令 $A_{n+i}=\varnothing\,(i\geqslant 1)$, 由概率的可列可加性和已证的 $P(\varnothing)=0$, 有

$$P\left(\bigcup_{k=1}^{n}A_k\right)=P\left(\bigcup_{k=1}^{\infty}A_k\right)=\sum_{k=1}^{\infty}P(A_k)=\sum_{k=1}^{n}P(A_k)$$

3) 注意 $B=A\cup(B-A)$, $A\cap(B-A)=\varnothing$, 由已证的有限可加性 (2), $P(B)=P(A)+P(B-A)$, 移项即得所需结论.

4) 由已证的 (3), $P(B)=P(A)+P(B-A)\geqslant P(A)$.

5) 注意 $A\cap\overline{A}=\varnothing$, $A\cup\overline{A}=\Omega$, 由已证的有限可加性知, $P(A)+P(\overline{A})=P(A\cup\overline{A})=1$. □

我们已经知道, 必然事件发生的概率为 1, 不可能事件发生的概率为 0, 但逆命题并不成立. 事实上, 由例 1.2.3可知, 事件 $\{0.5\}$ 的概率 (即长度) 为 0, 但它不是不可能事件 (即不是空集 $\varnothing$); 事件 $(0,1]$ 发生的概率 (即长度) 为 1, 但它不是必然事件 (即不是全空间 $\Omega=[0,1]$).

定理 1.2.2. (加法公式) 设 $(\Omega,\ \mathscr{F},\ P)$ 为概率空间.

1) 若 $A,B\in\mathscr{F}$, 则

$$P(A\cup B)=P(A)+P(B)-P(AB) \tag{1.2.4}$$

2) 若 $A_1,\cdots,A_n\in\mathscr{F}$, 则

$$P\left(\bigcup_{k=1}^{n}A_k\right)=\sum_{k=1}^{n}P(A_k)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n) \tag{1.2.5}$$

证. 1) 注意 $A\cup B=A\cup(B-A)=A\cup(B-AB)$, 且 $A\cap(B-AB)=\varnothing$, 由概率的有限可加性和可减性知:

$$\begin{aligned}P(A\cup B)=&P(A\cup(B-AB))\\=&P(A)+P(B-AB)\\=&P(A)+P(B)-P(AB)\end{aligned}$$

2) 用数学归纳法即可证明. □

例 1.2.4. 设 $p=P(A)\,,q=P(B)\,,r=P(A\cup B)$. 求 $P\big(A\overline{B}\big)$ 及 $P\big(\overline{A}\,\overline{B}\big)$.

解. 由 $r=P(A\cup B)=P(A)+P(B)-P(AB)=p+q-P(AB)$ 知 $P(AB)=p+q-r$. 所以,

$$P\big(A\overline{B}\big)=P(A\cap(\Omega-B))=P(A-AB)=P(A)-P(AB)=r-q$$

$$P\big(\overline{A}\,\overline{B}\big)=1-P\Big(\overline{\overline{A}\,\overline{B}}\Big)=1-P(A\cup B)=1-r$$

1.2.3 概率的连续性和次可加性 *

在这一小节中, 我们将学习概率更深入的一些性质.

如果事件序列 $\{A_n:n\geqslant 1\}$ 满足 $A_1\subset A_2\subset A_3\subset\cdots$, 则称该事件序列是**单调增**的; 如果事件序列 $\{A_n:n\geqslant 1\}$ 满足 $A_1\supset A_2\supset A_3\supset\cdots$, 则称该事件序列是**单调降**的. 通常称 $\bigcup\limits_{n=1}^{\infty}A_n$ 为单调增事件序列 $\{A_n:n\geqslant 1\}$ 的极限, 称 $\bigcap\limits_{n=1}^{\infty}A_n$ 为单调降事件序列 $\{A_n:n\geqslant 1\}$ 的极限. 下一定理说明, 不论 $\{A_n:n\geqslant 1\}$ 是单调增还是单调降,A_n 的概率均收敛到它的极限的概率, 故称概率具有 (上、下)**连续性**.

定理 1.2.3.* 设 $(\Omega,\mathscr{F},P)$ 为概率空间. 概率 P 具有如下性质:

1) 下连续性: 若 $\{A_n:n\geqslant 1\}$ 是单调增的事件序列, 则

$$\lim_{n\to\infty}P(A_n)=P\left(\bigcup_{n=1}^{\infty}A_n\right)\tag{1.2.6}$$

2) 上连续性: 若 $\{A_n:n\geqslant 1\}$ 是单调降的事件序列, 则

$$\lim_{n\to\infty}P(A_n)=P\left(\bigcap_{n=1}^{\infty}A_n\right)\tag{1.2.7}$$