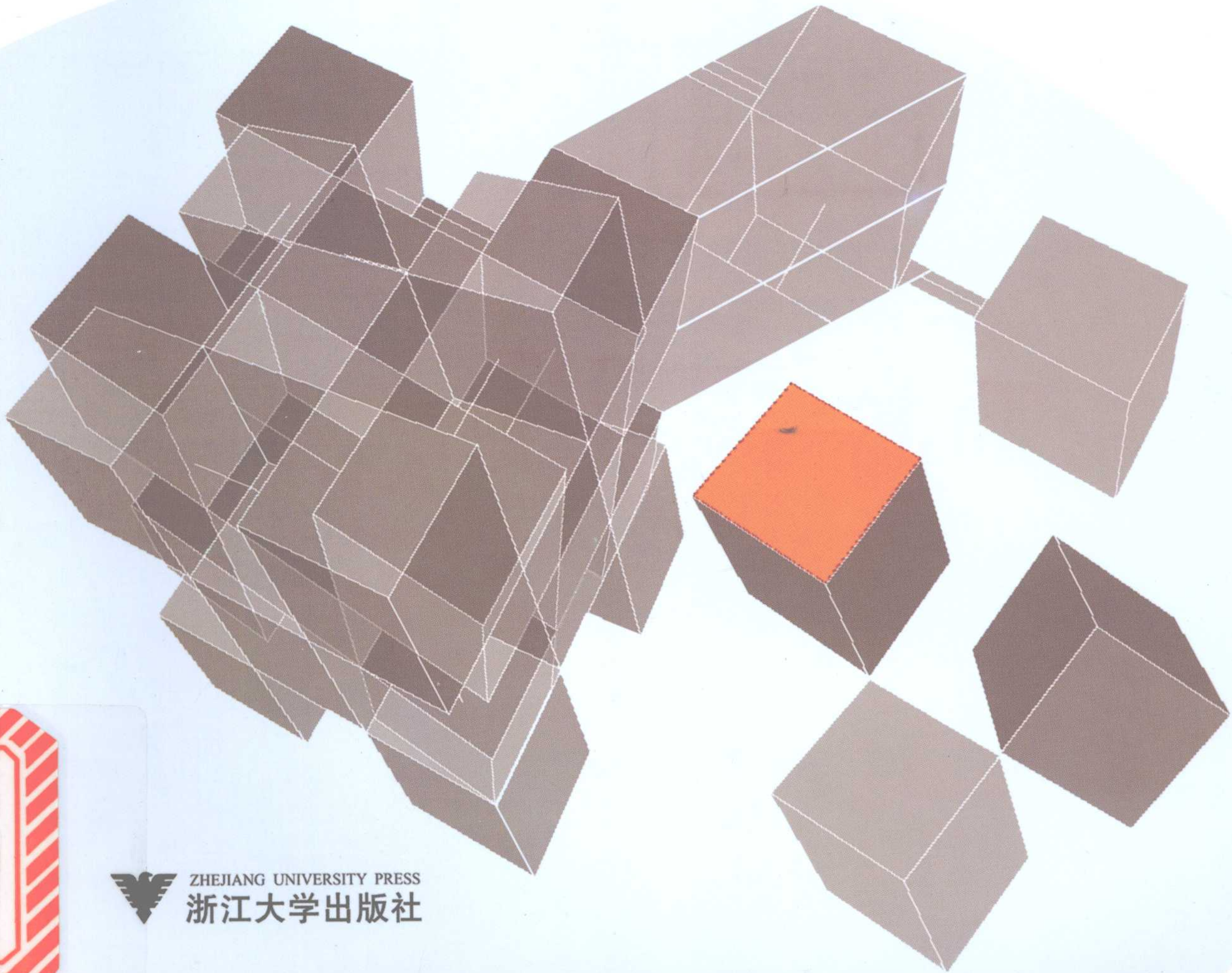
# 概率论与数理统计

龚小庆 王炳兴 主编



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

# 概率论与数理统计

龚小庆　王炳兴　主编



浙江大學出版社

**图书在版编目(CIP)数据**

概率论与数理统计/龚小庆，王炳兴主编．—杭州：浙江大学出版社，2007.8
ISBN 978-7-308-05539-0

Ⅰ．概… Ⅱ．①龚…②王… 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(2007)第 139211 号

# 概率论与数理统计

**主　　编** 龚小庆　王炳兴

**责任编辑** 徐素君
**封面设计** 刘依群
**出版发行** 浙江大学出版社
(杭州天目山路 148 号　邮政编码 310028)
(E-mail：zupress@mail.hz.zj.cn)
(网址：http://www.zjupress.com)
**排　　版** 杭州大漠照排印刷有限公司
**印　　刷** 杭州杭新印务有限公司
**开　　本** 787mm×960mm　1/16
**印　　张** 14.25
**字　　数** 305 千
**版 印 次** 2007 年 9 月第 1 版　2007 年 9 月第 1 次印刷
**书　　号** ISBN 978-7-308-05539-0
**定　　价** 19.80 元



浙江大学出版社发行部邮购电话(0571)88072522

# 前　言

概率论与数理统计作为现代数学的一个分支，是专门研究随机现象的统计规律性的一门学科，具有其特殊性。概率统计一方面具有应用性很强的特点，另一方面在数学理论上又显得比较抽象并且涉及的数学工具也较多。初学者往往对一些重要的概率统计的概念感到疑惑不解，尤其是在学习数理统计时常常会有"入宝山而空归"的感觉。考虑到这些因素，我们在取材和写作上，在以下几个方面作了努力：

(1) 用较多的篇幅详细地叙述了概率统计中的一些主要概念和方法产生的背景和思路，从直观入手逐步过渡到数学表述。

(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、准确和规范的叙述，并尽量阐述其实际意义。

(3) 由于公共数学本身的特点，本教材的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.

(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的直觉。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产生敏感，培养学生的概率统计直觉能力。

(5) 注重思想方法的介绍。概率统计不仅是一门数学理论，而且还具有世界观的性质。具备正确的概率统计的思想方法是大学生应该具备的一种基本修养和素质。因此本书特别注重阐释统计的思想、问题的背景和统计方法产生的历史(如极大似然法和显著性检验的思想)，以使学生对统计的思想方法有一个系统的了解。

(6) 本书内容紧扣全国硕士研究生入学数学(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，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.

(7) 为了帮助学生正确理解基本概念，准确掌握基本结论，本书提供了较多的反例。

(8) 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，有些思想和观点是第一次出现在此类教材中。

本教材是在我们多年教学实践的基础上并参照教育部关于全国非数学专业(理工、经济和管理等)硕士研究生考试数学(一)和数学(三)对概率论和数理统计部分的基本要求编写的，可作为高等学校非数学专业的理工科和金融经济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科研工作者的参考书。

根据我们的教学经验，授完本书大约需要 51 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。

本书由浙江工商大学统计与数学学院组织编写，大纲和体系由集体讨论而定。第一、六、七、八章由龚小庆教授执笔，第二、三、四、五章由王炳兴教授执笔，最后由龚小庆教授统一修改定稿。

在编写的过程中，我们参考了较多文献，为此我们均在书末的参考文献中列出。本书的编写自始至终得到浙江大学出版社和浙江工商大学数学与统计学院的大力支持，尤其是苏为华教授、朱灵教授、许冰教授和王海敏副教授给予了许多的帮助，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

2007 年 5 月于杭州

# 目 录

第四章　随机变量的数字特征



目　录

第五章　大数定律与中心极限定理



第六章　抽样分布



第七章　参数估计



第八章　假设检验

## 附　录



## 参考文献

# 第一章 随机事件与概率

自然界和人类社会中发生的现象大体上可分为如下两种类型.

一类是确定性现象. 这类现象的特点是,一旦某些确定的条件给定,某一特定的结果将必定会发生,或者已知它过去的状态,它将来的发展状态也被完全确定. 譬如,在一个标准大气压下,水在 100℃时一定沸腾;日全食和月全食的发生时刻可以准确预测;物体失去支撑就会坠落,等等. 可以说正是这一类现象的存在,导致了人们相信自然界中一定存在着秩序和规律的信念,这种信念又进一步导致了近代科学的发展.

另一类现象则是不确定性现象或随机现象. 这类现象的特点是,即使在相同的条件下重复进行,每次试验所得的结果也会不相同,或者已知它过去的状态,它将来的发展状态仍然无法确定. 譬如,抛掷一枚质地均匀的硬币,硬币落地后既有可能正面朝上也有可能反面朝上;两支足球队之间每场比赛的结果也不会完全相同,而且即使根据以前两队相互之间的成绩,也仍然无法准确地预测比赛的结果;明天的股市可能会涨也可能会跌,等等. 另外,有些事情即使已经发生了,但是在你知道结果之前,它们仍然具有不确定性. 譬如,硬币落地后虽然结果已经确定,但是在你观察之前还是无法确定硬币是正面朝上还是反面朝上;在一定的技术条件下,工厂生产某产品的边际成本是给定的,但是在竞争对手看来仍然具有多种可能性,等等. 因此,不确定性现象中的"不确定性"包含有两方面的含义,其一是客观结果的不确定性,其二是主观猜测的不确定性,后者往往融入了观察者个人的信念.

基于对随机现象与事件发生可能性大小的认识来做出决策,正是人类智能的表现. 一个理性的人在处理一个具体问题的时候,往往会根据各种结果发生的可能性的大小以及利害得失来作决策. 譬如,在交通日益发达的今天,人们会想到出门在外车祸发生的可能性与日俱增. 但是不出门又显然是不行的,于是人们便想出各种方法制定各种规则以便尽可能地减小车祸发生的可能性. 事实上,在一个法制健全的现代社会中,发生车祸的可能性是很小的. 为了让"万一"出了车祸以后所产生的损失减低到最小程度,就又有了保险业的发展. 保险业的保费并不是根据一件单独的事件来决定的,而是根据类似事件发生的可能性的大小来决定的,其理论基础就是本书所要讲授的概率论.

在人们的日常谚语中,也包含了大量如何应对随机现象的思想,如"晴带伞,饱带饭"就是提醒人们要对可能发生的不利事件进行必要的防范.

因此,在研究随机现象时必须采用一种与以往研究确定性现象不同的思维方式. 我

们不能期望将复杂的随机现象简单化为确定性现象来研究，而必须承认随机因素是不可避免的. 在决策时，我们必须考虑各种可能的结果. 在互相矛盾的各种可能结果中往往并不存在一个十全十美的决策，我们所能做的只是做出一个尽可能合理的决定. 由于随机因素的存在，再合理的决策也难免会有失误，但是我们要使失误的可能性尽可能的小，同时又能以较大的可能性获得预期的利益.

随机现象虽然给人的感觉是难以捉摸、不好把握，似乎丧失了确定性模型中所固有的客观性. 但是人们发现很多随机现象依然存在着固有的规律性，这种规律性往往是在大量的试验中呈现出来的. 譬如，重复抛掷一枚质地均匀的硬币，你会发现正面和反面朝上的次数所占的比例会越来越接近于 50%. 下面是历史上一些科学家所做的著名的抛硬币试验以及相关的数据：



**表 1.1　历史上一些著名的抛硬币试验**

| 实　验　者 | $n$ | $n_H$ | $f_n(H)$ |
|---|---|---|---|
| 德·摩根 | 2048 | 1061 | 0.5181 |
| 蒲　丰 | 4040 | 2048 | 0.5069 |
| K. 皮尔逊 | 12000 | 6019 | 0.5016 |
| K. 皮尔逊 | 24000 | 12012 | 0.5005 |
| 维　尼 | 30000 | 14994 | 0.4998 |

表中 $n$ 表示抛掷硬币的次数，$n_H$ 表示出现正面的次数，$f_n(H)=n_H/n$ 表示出现正面的次数所占的比例. 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将逐渐稳定于 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. 正是这种规律性的存在，使得我们利用数学工具来研究随机现象成为可能. 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数学分支.

## §1.1　基 本 概 念

### 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或试验. 这里所说的观察或试验，意义比较广泛，可以是各类科学试验，也可以是对某些事物的某些特征的观察. 下面是一些试验的的例子：

$E_1$：抛一枚硬币，观察正面 $H$、反面 $T$ 出现的情况；

$E_2$：将一枚硬币抛掷两次，观察正面 $H$、反面 $T$ 出现的情况；

$E_3$：将一枚硬币抛掷三次，观察正面 $H$、反面 $T$ 出现的情况；

$E_4$：掷一颗骰子，观察出现的点数；

$E_5$：在电视机厂的仓库里，随机地抽取一台电视机，测试它的寿命；

$E_6$：记录某一天城市发生车祸的次数.

以上所述的试验具有如下的共同特点：

(1) 可在相同的条件下重复进行；

(2) 每次试验的可能结果不止一个，但是能事先明确试验的所有可能的结果；

(3) 试验之前不能确定哪一个结果会出现.

一般地，我们称满足以上三个特点的试验为随机试验，简称为试验.

由随机试验的第二条性质知，那些事先不知道所有可能结果的随机现象不属于概率论的研究范围.



**定义 1.1.1** 随机试验 $E$ 的所有可能结果组成的集合称为 $E$ 的样本空间，记为 $S$. $S$ 的元素，即 $E$ 的一个可能结果，称为样本点或基本事件.

上面提到的各试验 $E_k$ 的样本空间 $S_k$ 分别为

$S_1 = \{H, T\}$

$S_2 = \{HH, HT, TH, TT\}$

$S_3 = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

$S_4 = \{1, 2, 3, 4, 5, 6\}$

$S_5 = \{t: t \geqslant 0\}$

$S_6 = \{0, 1, 2, 3, \cdots\}$

再设 $E_7$ 表示试验：将一枚硬币抛掷两次，观察正面出现的次数，则此时样本空间为

$$S_7 = \{0, 1, 2\}$$

同样是将硬币抛掷两次，但是 $S_2$ 和 $S_7$ 截然不同. 这说明，试验的目的决定试验所对应的样本空间.

今后，我们将把样本空间的某个子集(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”. 以 $E_5$ 为例，如果电视机的寿命超过 10000 个小时被认为是合格品，则“所抽取的电视机是合格品”这一事件可以用 $S_5$ 的子集 $A = \{t: t > 10000\}$ 来表示. 又如在 $E_2$ 中，如果令 $A =$“至少出现一次正面”这一事件，则 $A$ 可写成

$$A = \{HT, TH, HH\}$$

再譬如，在 $E_6$ 中，“发生了奇数次车祸”这一事件等同于 $S_6$ 的子集$\{1, 3, 5, 7, \cdots\}$.

设 $A$ 为一个事件，如果在试验中出现的样本点 $\omega \in A$，则称事件 $A$(在该次试验中)发生. 例如在 $E_5$ 中，若抽出的电视机的寿命为 10500 小时，则事件“所抽取的电视机是合格品”$=\{t: t > 10000\}$在该次试验中发生；若抽取的电视机的寿命为 9500 小时，则(在该次试验中)事件“所抽取的电视机是合格品”$=\{t: t > 10000\}$没有发生. 在 $E_6$ 中，若发生了 9 次车祸，则事件$\{1, 3, 5, 7, \cdots\}=$“发生了奇数次车祸”发生；若发生的车祸数

为 2，则$\{1,3,5,7,\cdots\}$=“发生了奇数次车祸”这个事件没有发生.

一般地，我们用英文字母表中前面的大写字母(可以带下标)表示事件，如用 $A$，$B$，$C$，或 $A_1$，$A_2$，$B_2$，$D_{11}$等表示事件.

样本空间 $S$ 有两个特殊的子集：一个是 $S$ 本身，由于它包含了所有可能的结果，所以在每次试验中它总是发生的，我们将其称为必然事件；另一个子集是空集$\varnothing$，它不包含任何元素，因此在每次试验中都不发生，我们将其称为不可能事件.

### 1.1.2 事件间的关系与运算

由于事件是样本空间的一个子集，因此本节所涉及到的事件之间的关系与运算就是集合间的关系与运算. 虽然在中学时就已经学过一些集合方面的知识，但是读者仍然要特别关注本节的内容. 在这里重点需要做的就是能正确地将集合论中的符号翻译成概率论的语言.

首先，如上节所述，我们有如下的对应关系：

$$\text{事件 } A \text{ 发生} \Leftrightarrow \text{试验 } E \text{ 的结果 } \omega \in A$$

其中符号“$\Leftrightarrow$”表示“当且仅当”. 从这个关系出发，就可以进一步地讨论事件间的关系与运算.

1. 包含关系

符号“$A \subset B$”在集合论中意味着：“若 $\omega \in A$，则必有 $\omega \in B$”，因此依据上边的对应关系，其在概率论中的含义为：“若事件 $A$ 发生，则事件 $B$ 必发生”. 这时我们就称事件 $B$ 包含事件 $A$.

因此，符号“$A \subset B$”在概率论中就意味着“事件 $B$ 包含事件 $A$”.

如果 $A \subset B$ 且 $B \subset A$，则我们称事件 $A$ 与事件 $B$ 相等，并记为 $A = B$.

2. 和运算

符号“$A \cup B$”在集合论中意味着：“$\omega \in A \cup B \Leftrightarrow \omega \in A$ 或 $\omega \in B$”，因此其在概率论中的含义是：“$A \cup B$ 发生 $\Leftrightarrow A$ 发生或 $B$ 发生 $\Leftrightarrow A$ 与 $B$ 至少有一个发生”.

我们将 $A \cup B$ 称为 $A$，$B$ 的和事件，它表示“$A$ 与 $B$ 至少有一个发生”这一新事件. 类似地，我们将“$\bigcup_{k=1}^{n} A_k$”称为 $n$ 个事件 $A_1,A_2,\cdots,A_n$ 的和事件，它表示“$A_1,A_2,\cdots,A_n$ 至少有一个发生”这一事件；将“$\bigcup_{k=1}^{\infty} A_k$”称为可列个事件 $A_1,A_2,\cdots,A_n,\cdots$ 的和事件，它表示“$A_1,A_2,\cdots,A_n,\cdots$ 至少有一个发生”这一事件.

3. 积运算

符号“$A \cap B$”或“$AB$”在集合中的含义是：“$\omega \in A \cap B \Leftrightarrow \omega \in A$ 且 $\omega \in B$”，因此其在概率论中意味着：“$A \cap B$ 发生 $\Leftrightarrow A$ 发生且 $B$ 发生 $\Leftrightarrow A$，$B$ 同时发生”.

我们将 $A \cap B$ 或 $AB$ 称为 $A$，$B$ 的积事件，它表示“事件 $A$ 与 $B$ 同时发生”这一事件.

类似地,称 $\bigcap_{k=1}^{n} A_k$ 为 $n$ 个事件 $A_1,A_2,\cdots,A_n$ 的积事件,它表示"$A_1,A_2,\cdots,A_n$ 同时发生"这一事件;称 $\bigcap_{k=1}^{\infty} A_k$ 为可列个事件 $A_1,A_2,\cdots,A_n,\cdots$ 的积事件,它"表示 $A_1,A_2,\cdots,A_n,\cdots$ 同时发生"这一事件.

4. 差运算

符号"$A-B$"在集合论中是指:"$\omega \in A-B \Leftrightarrow \omega \in A$ 且 $\omega \notin B$",因此在概率论中就意味着:"$A-B$ 发生 $\Leftrightarrow A$ 发生但 $B$ 不发生".

称 $A-B$ 为 $A$ 与 $B$ 的差事件,它表示"事件 $A$ 发生而事件 $B$ 不发生"这一新事件.

5. 互不相容或互斥

符号"$A \cap B = \varnothing$"或"$AB = \varnothing$"在集合论表示"$A$ 与 $B$ 不相交即没有公共元素",因而在概率论中就表示"$A$ 与 $B$ 同时发生是不可能事件".

一般地,如果 $AB=\varnothing$,我们就称事件 $A$ 与 $B$ 互不相容或互斥,它表示事件 $A$ 与 $B$ 不可能同时发生.

6. 对立事件或逆事件

符号"$B=\overline{A}$"在集合论中表示 $B$ 是 $A$ 的补集,即 $B=S-A$,它又可以表示为:

$$A \cup B = S \text{ 且 } AB = \varnothing$$

因此,其在概率论中的含义为:事件 $A$ 与 $\overline{A}$ 有且只有一个发生.

称 $B=\overline{A}$ 为 $A$ 的对立事件或逆事件,称 $A$ 与 $\overline{A}$ 的关系为互相对立或互逆.

从上面我们可以看出事件的运算实际上就是集合的运算,因此,和集合的运算一样,事件的运算同样应该满足下面的几条规律:

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

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

**德·摩根律** $\overline{A\cup B}=\overline{A}\cap\overline{B},\ \overline{A\cap B}=\overline{A}\cup\overline{B}$

$\overline{\bigcup_{n\geqslant 1} A_n}=\bigcap_{n\geqslant 1}\overline{A}_n,\ \overline{\bigcap_{n\geqslant 1} A_n}=\bigcup_{n\geqslant 1}\overline{A}_n$

正确地用符号表示事件的关系与运算是相当重要的,在很多时候往往成为解题的关键.

**例 1.1.1** 设 $A,B,C$ 是随机事件,则

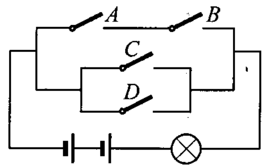
$AB\overline{C}$ 表示"$A$ 与 $B$ 发生,$C$ 不发生";

$AB\cup AC\cup BC$ 或 $ABC\cup\overline{A}BC\cup A\overline{B}C\cup AB\overline{C}$ 表示"$A,B,C$ 至少有两个发生";

$AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$ 表示"$A,B,C$ 恰好发生两个";

$\overline{A}\overline{B}\overline{C}\cup A\overline{B}\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\overline{B}C$ 表示"$A,B,C$ 中有不多于一个事件发生".

**例 1.1.2** 如图所示的电路,以 $A$, $B$, $C$, $D$ 分别表示开关$A$, $B$, $C$, $D$ 闭合,$E$ 表示"信号灯亮",则有 $E=AB\cup C\cup D$.




该例表明,在实际问题中,事件之间相互关系的确定有时不必直接借助于集合,而只须从概率论本身的含义出发即可.

## §1.2 频率与概率



研究随机现象不仅要知道可能出现哪些事件,还要知道各种事件出现的可能性的大小.我们把衡量事件发生可能性大小的数量指标称作事件的概率.事件 $A$ 的概率用 $P(A)$来表示.

现在我们要提出的问题是,对于一个给定的随机事件,它发生可能性大小的数量指标——概率,该如何确定呢?

为了从数学上对概率这个概念给出严格的定义,先讨论一个与此相关的概念——频率.

**定义 1.2.1** 在相同条件下,进行了 $n$ 次试验,在这 $n$ 次试验中事件 $A$ 发生的次数 $n_A$ 称为事件 $A$ 的频数,比值 $n_A/n$ 称为事件 $A$ 发生的频率,并记为 $f_n(A)$,即

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

从实际经验以及本章一开始就给出的表 1.1 中可知,频率具有如下一些特点:

(1) 频率的大小能体现事件发生可能性的大小,频率大则发生的可能性也应该大;反之,频率小则发生的可能性也小.事实上,在很多实际问题中我们就是用频率来衡量事件发生可能性大小的.譬如体育比赛中,体育评论员常用"投篮的命中率","射门的命中率"等来表示运动员在某一时段的水平.

(2) 频率有一定的随机波动性.譬如在表 1.1 中我们看到,当抛投硬币的次数不同时得到的频率常常会不一样,事实上,有时甚至投同样多次硬币由于抛投的时间地点不同可能也会得到不同的频率,这样就使频率缺乏科学度量单位所具有的客观性.

(3) 但是,另一方面,当试验的次数逐渐增多时,频率又具有稳定性.从表 1.1 中我们可以看出,随着 $n$ 的增大,$f_n(H)$ 稳定在 50% 左右.

如何来理解频率的波动性与稳定性呢?

给定一根木棒,谁都不会怀疑它有自身的"客观"长度,问题是它的长度是多少?在实际过程中,我们可以用尺或仪器来测量,但是不论尺或仪器有多精确,反复测量得到的数值总多多少少有一些差异,这类似于前面所说的频率的波动性.但是如果我们对大量重复测量的结果取平均值,这个平均值却总是稳定在真实的"长度"值的附近,这又有点类似于频率的稳定性.这个类比将不但有助于我们去理解概率与频率之间的内在联

系，而且还可以更进一步地揭示出客观世界广泛的统一性：概率与长度、面积等度量一样，都具有“测度”(measure) 的性质.

从上面的类比中，我们可以认为，频率实际上是概率的一个“测量”. 在这个测量过程中频率所呈现出的稳定性反映了概率的客观性，因此，如果我们能找出频率那些稳定的性质，那么从这些性质中便可看出概率的本质属性来.

经过归纳总结，可以得到频率的如下三条最基本的性质，即

(1) 非负性 任意事件 $A$ 的频率非负：$f_n(A) \geqslant 0$

(2) 规范性 必然事件 $S$ 的频率为 1：$f_n(S) = 1$

(3) 有限可加性 若 $A_1, A_2, \cdots, A_m$ 是一组两两不相容的事件，则有

$$f_n(\bigcup_{i=1}^{m} A_i) = \sum_{i=1}^{m} f_n(A_i)$$

上面的三条性质可直接由频率的定义来证明，请读者完成该证明.

因为频率的本质是概率，因此频率所满足的这三条性质同时也必须是概率所具有的性质. 但是由于理论上还要考虑到可列无穷多个事件的关系和运算，因此对上面的三条性质作适当的推广后，给出如下的定义.

**定义 1.2.2** 设 $E$ 为随机试验，$S$ 为相应的样本空间，$\mathbf{F}$ 为所有事件组成的集合，对于 $\mathbf{F}$ 中的每一个事件 $A$，分别赋予一个实数，记为 $P(A)$，如果集合函数 $P(\cdot)$ 满足下列条件：

(1) 非负性　对于每一个事件 $A$，$P(A) \geqslant 0$

(2) 规范性　$P(S) = 1$

(3) 可列可加性　若 $A_1, A_2, \cdots, A_n, \cdots$ 是一组两两不相容的事件，则有

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

则我们称 $P$ 为 $\mathbf{F}$ 上的概率；对任意的 $A \in \mathbf{F}$，称 $P(A)$ 为事件 $A$ 的概率.

由定义可知，概率 $P$ 是一个非负的、规范的和可列可加的实值集函数.

在历史上，对概率的理解一直存在着各种不同看法，譬如有从频率的角度来理解的(如本书中所主要采用的观点)，也有从主观信念的角度来理解的(如贝叶斯学派的主观概率)，等等，但是不论从哪个角度来理解概率这个概念，大家都承认上面三条是概率必须满足的最基本的性质. 这三条性质就像公理一样已被数学家们所普遍接受. 因此，上面的定义又被称为概率的公理化定义.

由定义，可推出概率的如下一些重要性质.

**性质 1** 不可能事件的概率为零，即 $P(\varnothing) = 0$

**证明** 令 $A_n = \varnothing (n = 1, 2, \cdots)$，则 $\bigcup_{i=1}^{\infty} A_i = \varnothing$，且 $A_i A_j = \varnothing, i \neq j$，由概率的可列可加性，得

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

而实数 $P(\varnothing)\geqslant 0$,故由上式知 $P(\varnothing)=0$. □

**性质 2** (有限可加性)若 $A_1,A_2,\cdots,A_n(n\geqslant 2)$ 是一组两两不相容的事件,则有

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

**证明** 利用可列可加性及性质 1,令 $A_i=\varnothing(i=n+1,n+2,\cdots)$ 即可. □

**性质 3** 设 $A\subset B$,则有

$$P(B-A)=P(B)-P(A)$$
$$P(B)\geqslant P(A)$$

**证明** 注意到 $B=A\cup(B-A)$,用性质 2 和概率的非负性即可证明. □

注意,一般情况下我们有

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

**性质 4** $P(A)\leqslant 1$

**证明** 由 $A\subset S$,再由性质 3 即得. □

**性质 5** $P(\overline{A})=1-P(A)$

**证明** 由于 $\overline{A}\cup A=S,A\overline{A}=\varnothing$,再由概率的规范性和有限可加性,得

$$1=P(S)=P(A\cup\overline{A})=P(A)+P(\overline{A})$$

移项后即证. □

**性质 6** (加法公式)对于任意两个事件 $A,B$ 有

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

**证明** 因为 $A\cup B=A\cup(B-AB),A(B-AB)=\varnothing$.
再由性质 2,3,有

$$P(A\cup B)=P(A)+P(B-AB)=P(A)+P(B)-P(AB)$$ □

该性质可推广到多个事件的和:

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

上式可由归纳法证得.

**性质 7** (概率的次可加性)对于任意的 $n$ 个事件 $A_1,A_2,\cdots,A_n$,有

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

证明请读者自己完成.

**例 1.2.1** 设 $P(A)=\frac{1}{3}$, $P(B)=\frac{1}{2}$. (1) 若 $AB=\varnothing$,求 $P(B\overline{A})$;(2) 若 $A\subset B$,求 $P(B\overline{A})$;(3) 若 $P(AB)=\frac{1}{8}$,求 $P(B\overline{A})$.

**解** (1) 因为 $AB=\varnothing$, $P(B\overline{A})=P(B)-P(AB)=P(B)=\frac{1}{2}$

(2) 因为 $A\subset B$,所以 $AB=A$,故

$$P(B\overline{A}) == P(B)-P(AB)=P(B)-P(A)=\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$$

(3) $P(B\overline{A})=P(B)-P(AB)=\frac{1}{2}-\frac{1}{8}=\frac{3}{8}$ ◇

## §1.3 等可能概型

概率论所讨论的问题中,有一类问题最简单直观,这类问题所涉及到的试验具有下面两个特征:

(1) 试验的样本空间的元素只有有限个.

(2) 试验中每个基本事件发生的可能性相同.

一般地,我们把具有上述两个特征的试验称为等可能概型或古典概型. 例如,抛一枚质地均匀的硬币,或者出现正面或者出现反面,只有两种结果,且每种结果出现的可能性相同. 又如抛一颗骰子,观察出现的点数,则共有 6 种可能结果,且每一种结果出现的可能性相同.

设试验的样本空间 $S=\{\omega_1,\omega_2,\cdots,\omega_n\}$,则由假设有

$$P(\omega_1)=P(\omega_2)=\cdots=P(\omega_n) \tag{1.3.2}$$

又基本事件是两两不相容的,于是

$$1=P(S)=P(\bigcup_{i=1}^{n}\omega_i)=\sum_{i=1}^{n}P(\omega_i) \tag{1.3.3}$$

由(1.3.2),(1.3.3),立即可得

$$P(\omega_i)=\frac{1}{n}, i=1,2,\cdots,n$$

于是,若 $A=\{\omega_{i_1},\omega_{i_2},\cdots,\omega_{i_k}\}$,其中 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$,则有

$$P(A)=\sum_{j=1}^{k}P(\omega_{i_j})=\frac{k}{n}=\frac{A\text{ 所包含的基本事件数}}{S\text{ 中基本事件总数}} \tag{1.3.4}$$

因此,要计算任何一个事件的概率,关键是要计算:(1) 样本空间所含的基本事件数 $n$;