经典数学系列

最好玩的概率事

齐浩然　编著

金盾出版社

经典数学系列

最好玩的概率事

齐浩然　编著

金盾出版社

内 容 提 要

本书通过讲述事件的方式，让你对概率有更加透彻的理解。它将带你进入到奇妙的可能性世界，带你一起去讨论关于概率论起源的故事。其实一些看似不合理的条件概率蕴藏着很多的知识，从本书之中你将会学到很多很好玩的有关概率的知识。

图书在版编目（CIP）数据

最好玩的概率事 / 齐浩然编著．—北京：金盾出版社，2015．5
（经典数学系列）
ISBN 978-7-5186-0015-1

Ⅰ．①最…　Ⅱ．①齐…　Ⅲ．①概率—青少年读物　Ⅳ．①O211．1-49

中国版本图书馆 CIP 数据核字（2015）第 021827 号

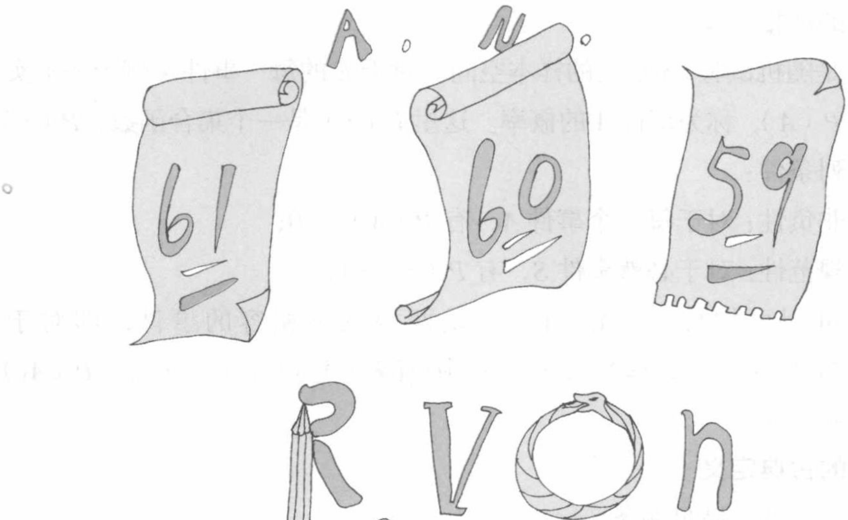
金盾出版社出版、总发行

北京市太平路 5 号（地铁万寿路站往南）
邮政编码：100036　电话：68214039　83219215
传真：68276683　网址：www.jdcbs.cn
北京市业和印务有限公司印刷、装订
各地新华书店经销
开本：700×1000　1/16　印张：11.75　字数：208千字
2015 年 5 月第 1 版第 1 次印刷
印数：1～10 000 册　定价：29.50 元

目录 contents

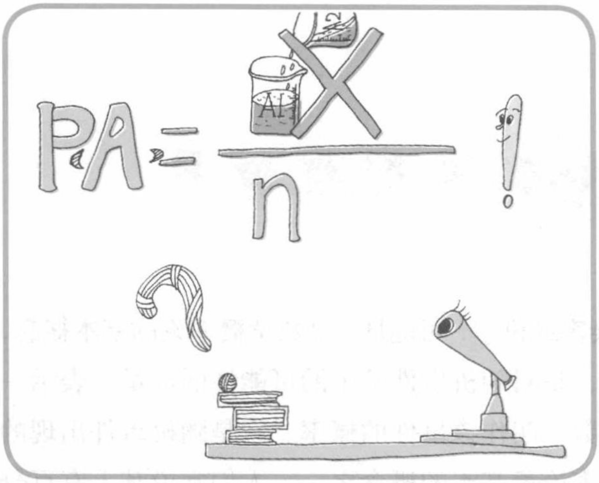
进入奇妙的可能性世界

概率，又称或然率、机会率或机率、可能性，是数学概率论的基本概念，是一个在 0 到 1 之间的实数，是对随机事件发生的可能性的度量。表示一个事件发生的可能性大小的数，叫作该事件的概率。它是随机事件出现的可能性的量度，同时也是概率论最基本的概念之一。人们常说某人有百分之多少的把握能通过这次考试，某件事发生的可能性是多少，这都是概率的实例。但如果说一件事情发生的概率是 $1/n$，不是指 n 次事件里必有一次发生该事件，而是指此事件发生的频率接近于 $1/n$ 这个数值。



定义

概率的频率定义



随着人们遇到问题的复杂程度的增加，等可能性逐渐暴露出它的弱点，特别是对于同一事件，可以从不同的等可能性角度算出不同的概率，从而产生了种种悖论。另一方面，随着经验的积累，人们逐渐认识到，在做大量重复试验时，随着试验次数的增加，一个事件出现的频率，总在一个固定数的附近摆动，显示一定的稳定性。R.von 米泽斯把这个固定数定义为该事件的概率，这就是概率的频率定义。从理论上讲，概率的频率定义是不够严谨的。A.H. 柯尔莫哥洛夫于 1933 年给出了概率的公理化定义。

概率的严格定义

设 E 是随机试验，S 是它的样本空间。对于 E 的每一事件 A 赋予一个实数，记为 $P(A)$，称为事件 A 的概率。这里 $P(\cdot)$ 是一个集合函数，$P(\cdot)$ 要满足下列条件：

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

（2）规范性：对于必然事件 S，有 $P(S)=1$；

（3）可列可加性：设 A_1，A_2……是两两互不相容的事件，即对于 $i \neq j$，$A_i \cap A_j=\phi$，（i，$j=1$，2……），则有 $P(A_1 \cup A_2 \cup \cdots\cdots)=P(A_1)+P(A_2)+\cdots\cdots$

概率的古典定义

如果一个试验满足两条：

（1）试验只有有限个基本结果；

（2）试验的每个基本结果出现的可能性是一样的。

这样的试验，称为古典试验。

对于古典试验中的事件 A，它的概率定义为：

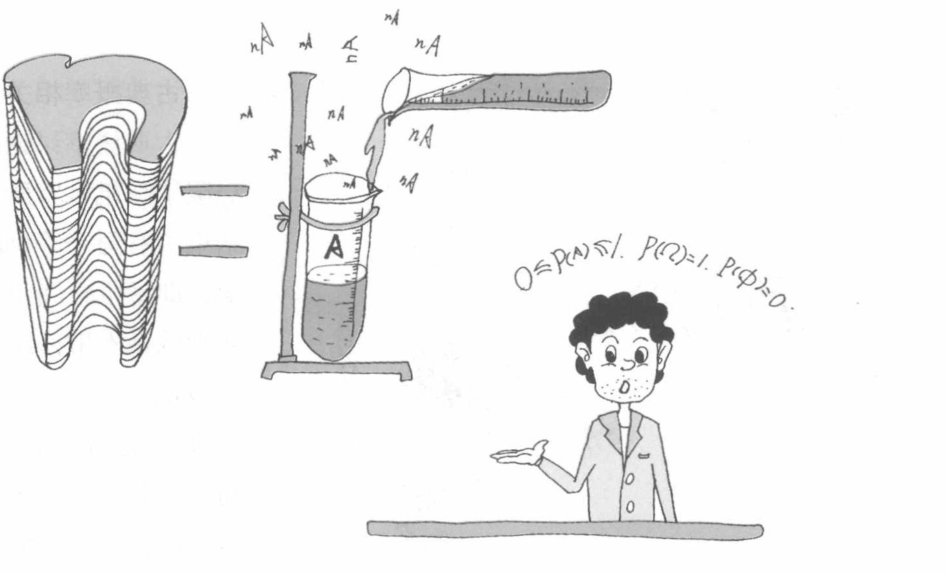
$P(A)=m/n$，n 表示该试验中所有可能出现的基本结果的总概率数目。m 表示事件 A 包含的试验基本结果数。这种定义概率的方法称为概率的古典定义。

概率的统计定义

在一定条件下，重复做 n 次试验，nA 为 n 次试验中事件 A 发生的次数，如果随着 n 逐渐增大，频率 nA/n 逐渐稳定在某一数值 p 附近，则数值 p 称为事件 A 在该条件下发生的概率，记做 $P(A)=p$。这个定义成为概率的统计定义。

在历史上，第一个对“当试验次数 n 逐渐增大，频率 nA 稳定在其概率 p 上”这一论断给以严格的意义和数学证明的是早期概率论史上最重要的学者雅各布·伯努利（公元 1654 年 ~ 1705 年）。

从概率的统计定义可以看到，数值 p 就是在该条件下刻画事件 A 发生可能性大小的一个数量指标。



由于频率 nA/n 总是介于 0 和 1 之间，从概率的统计定义可知，对任意事件 A，皆有 $0 \leqslant P(A) \leqslant 1$，$P(\Omega)=1$，$P(\Phi)=0$。

Ω、Φ 分别表示必然事件（在一定条件下必然发生的事件）和不可能事件（在一定条件下必然不发生的事件）。

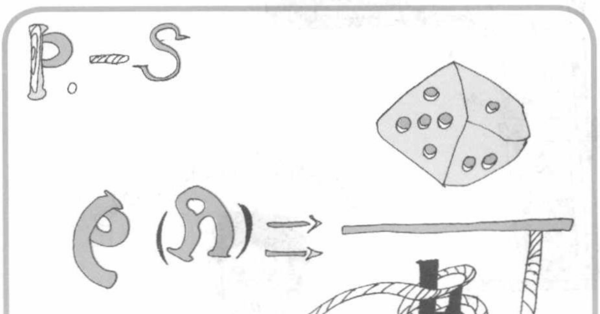
历史

第一个系统地推算概率的人是 16 世纪的卡尔达诺（Cardano）。此文记载在他的著作《*Liberde Ludo Aleae*》中。书中关于概率的内容是由 Gould 从拉丁文翻译出来的。

Cardano 的数学著作中有很多给赌徒的建议。这些建议都写成了短文。例如:《谁，在什么时候，应该赌博？》《为什么亚里士多德谴责赌博？》《那些教别人赌博的人是否也擅长赌博呢？》等。

然而，首次提出系统研究概率的是在帕斯卡和费马来往的一系列信件中。这些通信最初是由帕斯卡提出的，他想找费马请教几个关于由 ChevvalierdeMere 提出的问题。ChevvalierdeMere 是一知名作家，路易十四宫廷的显要，也是一名狂热的赌徒。问题主要是两个：掷骰子问题和比赛奖金分配问题。

两大类别



古典概率相关

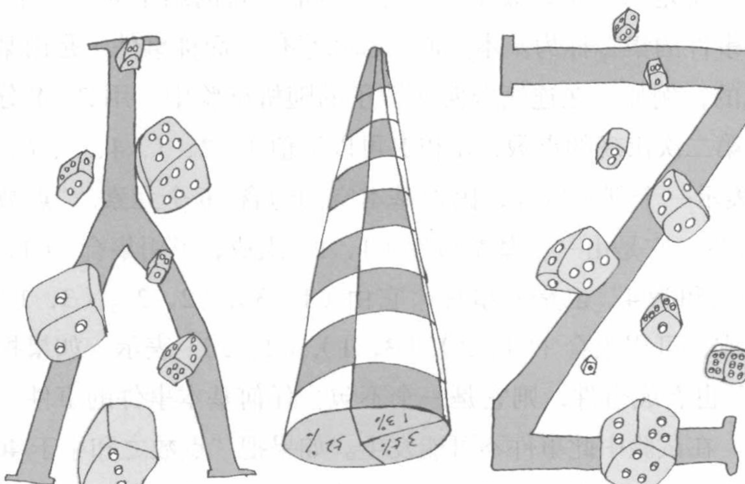
古典概率讨论的对象局限于随机试验所有可能结果为有限个等可能的情形，即基本空间由有限个元素或基本事件组成，其个数记为 n，每个基本事件发生的可能性是相同的。若事件 A 包含 m 个基本事件，则定义事件 A 发

生的概率为 $P(A)=m/n$，也就是事件 A 发生的概率等于事件 A 所包含的基本事件个数除以基本空间的基本事件的总个数，这是 P.–S. 拉普拉斯的古典概率定义，或称之为概率的古典定义。历史上古典概率是由研究诸如掷骰子一类赌博游戏中的问题引起的。计算古典概率，可以用穷举法列出所有基本事件，再数清一个事件所含的基本事件个数相除，即借助组合计算可以简化计算过程。

几何概率相关

若随机试验中的基本事件有无穷多个，且每个基本事件发生是等可能的，这时就不能使用古典概率，于是产生了几何概率。几何概率的基本思想是把事件与几何区域对应，利用几何区域的度量来计算事件发生的概率，布丰投针问题是应用几何概率的一个典型例子。

在概率论发展的早期，人们就注意到古典概率仅考虑试验结果只有有限个的情况是不够的，还必须考虑试验结果是无限个的情况。为此可把无限个试验结果用欧式空间的某一区域 S 表示，其试验结果具有所谓“均匀分布”的性质，关于“均匀分布”的精确定义类似于古典概率中“等可能”这一概念。假设区域 S 以及其中任何可能出现的小区域 A 都是可以度量的，其度量的大小分别用 $\mu(S)$ 和 $\mu(A)$ 表示。如一维空间的长度，二维空



间的面积，三维空间的体积等。并且假定这种度量具有如长度一样的各种性质，如度量的非负性、可加性等。

几何概率的严格定义

设某一事件 A（也是 S 中的某一区域），S 包含 A，它的量度大小为 μ（A），若以 P（A）表示事件 A 发生的概率，考虑到“均匀分布”性，事件 A 发生的概率取为：$P(A)=\mu(A)/\mu(S)$，这样计算的概率称为几何概率。

若 Φ 是不可能事件，即 Φ 为 Ω 中的空的区域，其量度大小为0，故其概率 $P(\Phi)=0$。

独立试验序列

假如一串试验具备下列三条：

（1）每一次试验只有两个结果，一个记为“成功”，一个记为“失败”，$P\{成功\}=p$，$P\{失败\}=1-p=q$；

（2）成功的概率 p 在每次试验中保持不变；

（3）试验与试验之间是相互独立的。

则这一串试验称为独立试验序列，也称为 bernoulli 概型。

必然事件与不可能事件

在一个特定的随机试验中，称每一可能出现的结果为一个基本事件，全体基本事件的集合称为基本空间。随机事件（简称事件）是由某些基本事件组成的，例如，在连续掷两次骰子的随机试验中，用 Z，Y 分别表示第一次和第二次出现的点数，Z 和 Y 可以取值 1、2、3、4、5、6，每一点（Z，Y）表示一个基本事件，因而基本空间包含 36 个元素。“点数之和为2”是一事件，它是由一个基本事件（1，1）组成，可用集合 {(1，1)} 表示“点数之和为 4”也是一事件，它由（1，3），（2，2），（3，1）3 个基本事件组成，可用集合 {(1，3)，(3，1)，(2，2)} 表示。如果把“点数之和为 1”也看成事件，则它是一个不包含任何基本事件的事件，称为不可能事件。在试验中此事件不可能发生。如果把“点数之和小于 40”看成一事件，它包含所有基本事件，在试验中此事件一定发生，所以称为必然

事件。若 A 是一事件，则“事件 A 不发生”也是一个事件，称为事件 A 的对立事件。实际生活中需要对各种各样的事件及其相互关系、基本空间中元素所组成的各种子集及其相互关系等进行研究。

举个例子：小明要在 4 个抽屉中放入 5 个球，其中有一个抽屉会有 2 个球，这就是必然事件。

再举个例子：小明要在 5 个抽屉中放入 3 个球，如果说其中每个抽屉都有球，那么，这就是不可能事件。

随机事件，基本事件，等可能事件，互斥事件，对立事件

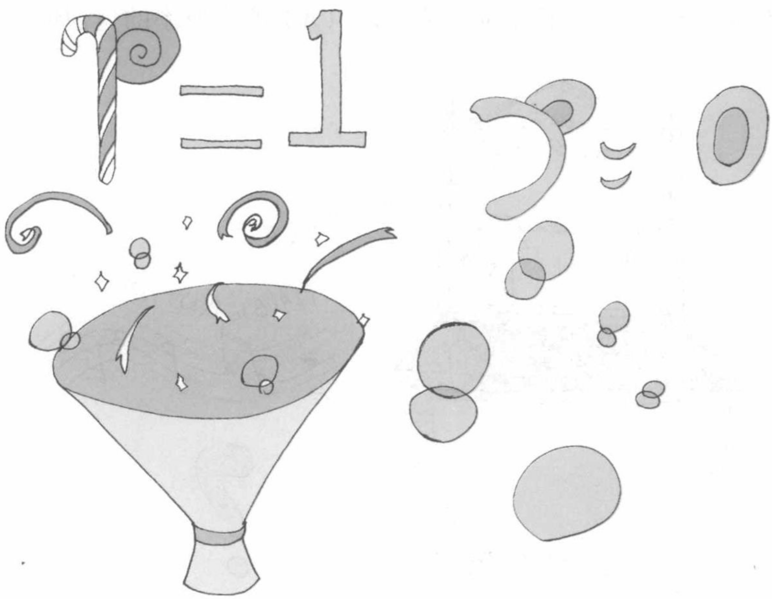
在一定的条件下可能发生也可能不发生的事件，叫作随机事件。

一次实验连同其中可能出现的每一个结果称为一个基本事件。

通常一次实验中的某一事件由基本事件组成。如果一次实验中可能出现的结果有 n 个，即此实验由 n 个基本事件组成，而且所有结果出现的可能性都相等，那么这种事件就叫作等可能事件。

不可能同时发生的两个事件叫作互斥事件。

必有一个发生的互斥事件叫作对立事件。



即 P（必然事件）=1

P（可能事件）=（0–1）（可以用分数）

P（不可能事件）=0

性质

性质 1：$P(\Phi)=0$

性质 2：（有限可加性）。当 n 个事件 $A_1\cdots A_n$ 两两互不相容时：$P(A_1\cup\cdots\cup A_n)=P(A_1)+\cdots+P(A_n)$

性质 3：对于任意一个事件 A：$P(A)=1-P$（非 A）

性质 4：当事件 A，B 满足 A 包含于 B 时：$P(B-A)=P(B)-P(A)$，$P(A)\leqslant P(B)$

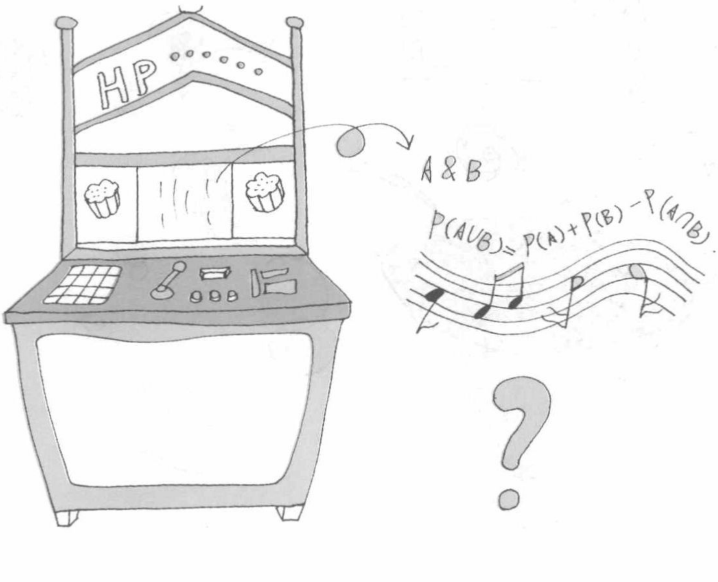
性质 5：对于任意一个事件 A，$P(A)\leqslant 1$

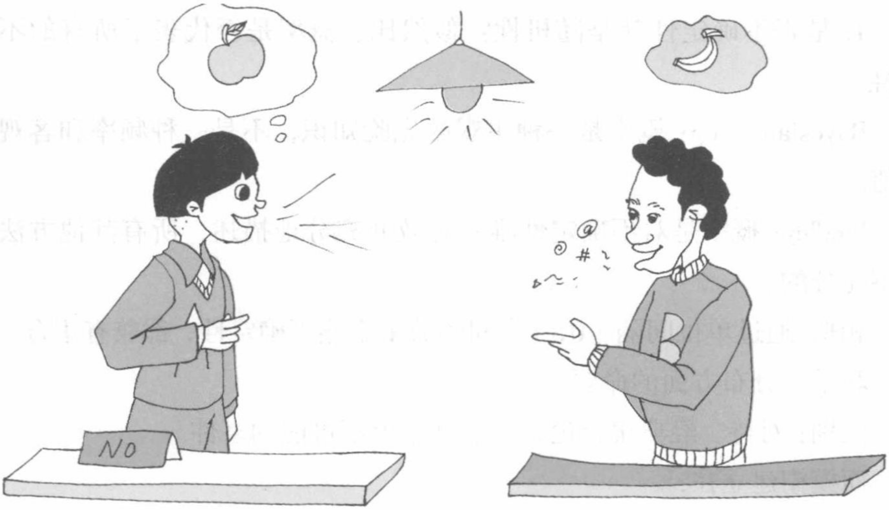
性质 6：对任意两个事件 A 和 B，$P(B-A)=P(B)-P(AB)$

性质 7：（加法公式）。对任意两个事件 A 和 B，$P(A\cup B)=P(A)+P(B)-P(A\cap B)$

$P(A)$ 是客观的，而 $Fn(A)$ 是依赖经验的。

统计中有时也用 n 很大的时候的 $Fn(A)$ 值当概率的近似值。





三个基本属性

1. 非负性：任何事件 A，$P(A) \geqslant 0$

2. 完备性：$P(\Omega)=1$

3. 加法法则：如事件 A 与 B 不相容，即如果 $AB=\phi$，则 $P(A+B)=P(A)+P(B)$

加法法则

如事件 A 与 B 不相容，$A+B$ 发生的时候，A 与 B 两者之中必定而且只能发生其中之一。独立重复地做 n 次实验，如记事件 A 发生的频数为 μA、频率为 $Fn(A)$，记事件 B 发生的频数为 μB、频率为 $Fn(B)$，事件 $A+B$ 发生的频数为 $\mu A+B$、频率为 $Fn(A+B)$，已知：$\mu A+B=\mu A+\mu B$，$\therefore Fn(A+B)=Fn(A)+Fn(B)$，它们的稳定值也应有：$P(A+B)=P(A)+P(B)$［加法法则］如事件 A 与 B 不相容，即如果 $AB=\phi$，则 $P(A+B)=P(A)+P(B)$ 即：两个互斥事件的和的概率等于它们的概率之和。请想一下：如 A 与 B 不是不相容，即相容的时候呢？进一步的研究得：$P(A+B)=P(A)+P(B)-P(AB)$ 这被人称为："多退少补"！

模糊和概率

1. 是否不确定性就是随机性？似然比、概率是否代表了所有的不确定性？

Bayesian camp：概率是一种主观的先验知识，不是一种频率和客观测量值。

Lindley：概率是对不确定性唯一有效并充分地描述，所有其他方法都是不充分的。

相似:通过单位间隔［0，1］间的数来表述不确定性，都兼有集合、相关、联系、分布方面的命题。

区别：对待。经典集合论，代表概率上不可能的事件。

而模糊建立在：

（1）是否总是成立的？

考虑能否逻辑上或部分地违背“无矛盾定理”（Aristotle 的三个‘思考定理’之一，同时包括排中定理、同一性定理这些都是非黑即白的经典定理。）模糊（矛盾）的产生，就是西方逻辑的结束。

（2）是否可以推导条件概率算子？