Probability For Applications

应用概率论

曹培慎◎编著

陕西师範大學出版总社

陕西师范大学教材建设基金资助出版

Probability for applications

应用概率论

曹培慎◎编著

陕西师范大学出版总社

图书代号 JC17N0869

图书在版编目(CIP)数据

应用概率论/曹培慎编著. —西安:陕西师范大学出版总社有限公司,2017.10
ISBN 978-7-5613-9393-2

Ⅰ.①应… Ⅱ.①曹… Ⅲ.①概率论—应用—高等学校—教材 Ⅳ.①O211.9

中国版本图书馆CIP数据核字(2017)第140330号

应用概率论

YINGYONG GAILÜLUN

曹培慎 编著

责任编辑/ 张俊胜 王舒涵
责任校对/ 聂会珍
封面设计/ 鼎新设计
出版发行/ 陕西师范大学出版总社
(西安市长安南路199号 邮编 710062)
网 址/ http://www.snupg.com
经 销/ 新华书店
印 刷/ 兴平市博闻印务有限公司
开 本/ 787mm×960mm 1/16
印 张/ 20.75
字 数/ 303千
版 次/ 2017年10月第1版
印 次/ 2017年10月第1次印刷
书 号/ ISBN 978-7-5613-9393-2
定 价/ 42.00元

前　言

概率论研究的是随机现象中的规律性。

目前国内的同类教材大都是以《概率论与数理统计》命名，将概率论和数理统计结合在一起讲授。作为教学计划中一学期的课程，教学内容显得较多。同时，经管类院系还开设《统计学》课程，与数理统计的内容又有很多的重复。鉴于此，本教材以《应用概率论》命名，按 54 课时设计，在内容上只讲授概率论部分，同时增加了在经济、金融领域经常用到的几种随机变量的分布类型，如帕累托分布、Logistic 分布、几何正态分布等，还增添了矩母函数、特征函数这些重要的研究随机变量的工具，对于条件期望和条件方差也进行了稍微细致的讨论，对于在实际中会遇到的断尾分布和截尾分布也尝试性地进行了讨论。

在写作上，本教材没有一味追求理论的严密性，而是尽量以通俗的语言阐述概念和定理的含义，以便能和人们的生活经验相衔接。在教材中增加了更多的函数图像和逻辑关系示意图，以增强对于问题的理解。全书的主线是随机事件发生的概率的计算。随机事件的概率是$(0,1)$之间的一个数，由于人们不容易理解大小是 e^{-1} 或者是$\sqrt{7}/10$的概率，所以对于概率，本教材最后都约化成小数的形式。

在应用中，现在的数据处理软件大多都是英文的，为了能够顺利地使用这些软件，本教材添加了专业名词的英文。基本概念或者说专业名称是构成一门学科的基本要素，本教材在每章末整理了该章的重要名词。

在例题的讲解上，本教材对于较难理解的例题，在求解前面加了解析，对于结论有价值的例题，在求解后面加了评注。事实上对于每一道例题都应该分为解析、求解、评注三部分。对于一个问题，利用不同的方法进行求解和相互验证，能够加深对于问题的理解。

概率论是与生活联系非常紧密的数学分支。随着电子技术的进步，教学内容和教学方式也应该随之改变。例如一些计算问题已经变得很容易，像二项分布的泊松近似、正态近似的重要性已经降低。而如何应用理论解决实际问题的重要性应该提高。本教材设计了 22 个 Excel 专栏，利用 Excel 进行计算、画图、随机模拟。图形能够直观地显示特性，进行随机模拟增加感性认识也是理解概率论的重要途径，在教学中可以贯穿 Excel 的使用。本书使用的是 Excel 2007。在 Excel 2016 中，函数名称有所变化，还增加了更多的函数。

根据教学实际情况，对于全书中带有星号的部分，可以作为选学内容。

在编写过程中，参考与借鉴了同类的教材。本教材的出版也得到了陕西师范大学教材建设基金的资助，在此一并致以衷心的感谢。

限于编者的水平，本书一定存在不少错误与不足，希望读者批评指正。我的电子信箱是 caodd@ snnu. edu. cn.

作　者

目 录

第一章　随机事件及其概率

世界是由万事万物组成的，而认识世界是人类的天性。人们对不确定现象中所包含规律性的认识和研究成果总结为概率论。从确定世界到不确定世界需要思维和研究方式的转换。学习概率论是一段充满乐趣和挑战的旅程。

1.1　随机现象和随机事件

一、随机现象

通过对世界上的事物/现象的长期观察，人们认识到这些事物/现象可以分为两类：一类是在一定的条件下，结果是确定的或者说必然的；另一类是在一定的条件下，结果是不确定的或者说随机的。前者如每天早晨太阳从东方升起；给定正方形的边长为2，则其面积必然为4；一个口袋中装有10个完全相同的白球，从中任取一个球出来，它必然为白球。人们称这类现象为**确定现象**（determination phenomenon）。后者如在同样的条件下，一次又一次地抛硬币，结果始终是有两种可能，有时是正面向上有时是反面向上，但事前不能确定是正面向上还是反面向上，也就是结果是不确定的或者说随机的，这样的结果可以称为**随机结果**。世界上存在着大量的类似现象，如掷骰子、投飞镖、买彩票和股票等等，这些现象都是在一定的条件/原因下结果是随机的，人们事前不能确定结果会如何，称这类现象为**随机现象**（random phenomenon）。

在生活中，人们经常要面向未来进行决策，需要知道明天的情况，如明天是否会下雨？某种股票明天的价格是多少？这些事情在事前都不能确定其结果。人们把它们都看作是随机现象。从这个意义上可以说，人们其实是生活在一个随机世界中。一般地，人们偏爱的是确定现象，因为其结果在事前是已知的，人们能够做出合适的决策。但在另外一些事情如金融投资问题上，人们偏爱的却

是随机现象,在不确定的结果中才蕴含着获利的机会。

确定现象和随机现象之间的界限是模糊的。有时确定现象和随机现象之间可以相互转化。理论上的确定现象,在实际中经常表现为随机现象。如零件的尺寸,在设计/理论上是一个确定的值,但在生产过程中会受到很多难以控制的因素的影响,生产出来的零件在尺寸上总有一些误差,使得零件尺寸表现为随机现象。另一方面,随着人们对随机现象研究的深入,对原因影响结果的机制认识得越来越清楚,该现象将变成确定现象。目前,之所以说随机现象大量存在、人们生活在一个随机世界中,表明人们对世界的认识还远远不够深入。

二、随机现象中的规律性

人们需要认识、研究随机现象。对于像抛硬币这样的随机现象,人们经过大量重复的试验、观察,发现虽然事前不能确定是哪种结果,但正面向上和反面向上出现的概率/可能性是相同的。

在这样的随机现象中,随机结果出现的可能性具有一定的稳定性或者说规律性。这表明随机现象也具有固有的数量规律性,人们研究随机现象的着眼点,不再是什么情况下会出现什么结果这样的对应关系,而是转向了随机结果出现的规律。可以说,**概率论**(probability theory)的研究对象就是随机现象中结果出现的规律性,这种规律性就是随机结果出现的可能性,可能性是用概率这个概念来表达的。

三、随机试验

为了对随机现象的结果所具有的数量规律性进行研究,人们需要对随机现象进行反复的观察。历史上,研究随机现象中所具有的规律性的最著名的试验是抛硬币的试验。

表 1.1　历史上抛硬币试验的记录

试验者	抛掷次数(n)	正面次数(r_n)	正面频率$\left(\frac{r_n}{n}\right)$
De Morgan	2048	1061	0.5181
Buffon	4040	2048	0.5069
Pearson K	12000	6019	0.5016
Pearson K	24000	12012	0.5005

从表 1.1 可以看出,不管是谁抛一枚均匀的硬币,当试验次数逐渐增多时,

频率呈现出稳定于 0.5 的倾向。频率的这种稳定性就是其中的规律性。

把对随机现象的结果进行的反复观察称为**随机试验**(randomized trial),简称为试验。试验不同于**实验**(experiment)。根据《现代汉语词典》的解释:实验是为了检验某种科学理论或假设而进行某种操作或从事某种活动。而试验是为了观察某事的结果或某物的性能而从事某种活动。除了上面这个例子,再如观察某射手对目标的多次射击结果、观察某电话交换台在一天内收到的呼叫次数等等,都是试验。这样的试验具有下列几方面的特点:

1. **可观察性** 试验结果是可以观察的,所有可能结果是明确的;

2. **可重复性** 试验原则上可以在相同条件下反复进行;

3. **随机性** 每次试验事前无法预知将要出现那个结果。

通过随机试验可以认识很多随机现象中的规律性,但并不是所有的随机现象都可以通过试验的方法来研究:有时无法做试验,有时试验并不具有这些特点。通过试验的方法来研究随机现象中的规律性还是很受局限的。

四、样本空间

对于能够通过试验来研究的随机现象来说,其试验将要出现的结果事前是不能确定的,但所有可能的结果事前却是明确的。把试验的每一个可能结果称为一个样本点,用 ω 表示。由于所有样本点也是明确的,把它们的全体/集合称为**样本空间**(sample space),用 Ω 表示。样本空间其实是对试验结果的描述。

例 1.1 在抛硬币的试验中,有两个样本点:正面向上、反面向上。样本空间为{正面向上、反面向上}。为了方便,如果记 ω_1 = 正面向上,ω_2 = 反面向上,则样本空间为 $\Omega=\{\omega_1,\omega_2\}$.

例 1.2 在掷骰子的试验中,有 6 个样本点:1 点、2 点、…、6 点,样本空间为{1 点、2 点、…、6 点}或{1,2, …,6}。如果记 $\omega_1=1$ 点,$\omega_2=2$ 点,…,$\omega_6=6$ 点,则样本空间为 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_6\}$.

例 1.3 观察某电话交换台在某天收到的呼叫次数,其样本点可能有可数无穷多个:0 次、1 次、2 次、…,样本空间为{0 次,1 次,2 次,…}或 $\Omega=\{0,1,2,\cdots\cdots\}$.

例 1.4 观察一个新灯泡的使用寿命,其样本点有不可数无穷多个:t 小时,$0\leqslant t\leqslant+\infty$,样本空间为{$t$ 小时|$0\leqslant t\leqslant+\infty$}或$\{t|0\leqslant t\leqslant+\infty\}$.

注:可数的与不可数的

对于一个集合来说,其元素个数有有限和无限两种情况。对于个数无限的

情况,如果其个数能和自然数集建立起一一对应关系,则称为是可数的/可列的,如整数集、有理数集;反之称为不可数的/不可列的,如实数集。

五、随机事件

在试验中,人们除了关注各种不同结果出现的概率外,常常还关注试验结果所具有的特征。如掷骰子的试验中,人们还关注出现的点数是否为偶数。"点数为偶数"是结果的一种特征。在对随机现象的研究中,把具有某种特征的结果称为一个**事件**(event)。事件是对结果具有某种特征的描述。对于一个随机试验,可以探讨结果的多种特征,因而事件有很多。为了方便,通常用大写字母 A、B、C…或 A_1,A_2,A_3…来表示事件。

例 1.5 在掷骰子的试验中,

"点数是 6"是一事件,可记为 A,

"点数小于 5"是一事件,可记为 B,

"点数是小于 5 的偶数"是一事件,可记为 C,

"点数小于 7"也是一个事件,可记为 D,

"点数小于 0"同样是一个事件,可记为 E.

事件 A 对应于唯一可能的结果,即样本点 6 点,这样的事件是基本事件,在该试验中,共有 6 个基本事件。事件 B 对应的结果可能为 1 点、2 点、3 点和 4 点,这样的事件是复合事件。事件 C 对应的可能结果为 2 点和 4 点,也是复合事件。事件 D 对应所有可能的结果,而事件 E 没有对应任何一个结果。

在一次试验中,如果出现的结果具有事件所描述的特征,则称该事件在这次试验中发生;如果出现的结果不具有事件所描述的特征,则称该事件在这次试验中没发生。在掷骰子的试验中,如果掷了一次,出现的结果是样本点 6 点,则在这次试验中,事件 A 发生,而事件 B、事件 C 没发生;如果出现的结果是样本点 3 点,则在这次试验中,事件 A 没发生,事件 B 发生,而事件 C 没发生。

像 A,B,C 这样的事件,在一次试验中有可能发生也有可能没发生,称为**随机事件**(random event)。随机事件包括基本事件和复合事件。在一次试验中,不管结果是那个样本点,事件 D 都发生,这样的事件是**必然事件**;而事件 E 一定不会发生,这样的事件是**不可能事件**。必然事件和不可能事件是确定事件。概率论关注的是随机事件,简称为事件。

六、事件的集合表示

样本空间是试验的所有可能结果的全体,实际上是所有样本点的集合,每一个样本点是该集合的元素。随机事件包含了样本空间中的一些元素,自然就可以用样本空间的子集来表达。如例 1.5 中,各事件可相应地记作

$A=\{6\},B=\{1,2,3,4,5\},C=\{2,4\},D=\{1,2,3,4,5,6\}$

而事件 E 是空集,所以 $E=\varnothing$.

事件可以用语言描述,也可以用集合表示。将事件用集合表示之后,给人们认识事件带来了极大的便利。在概率论中,常用一个长方形表示样本空间 Ω,用其中的一个圆表示事件 A,这样的图形称为维恩图。

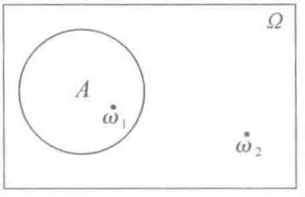


图 1.1 维恩图

七、事件之间的关系和运算

在一个随机试验中,为了通过基本事件来研究复合事件,需要研究事件之间的关系和运算。有了事件的集合表示,事件之间的关系和运算就类似于集合之间的关系和运算。

1. 事件的包含

如果属于事件 A 的每一个样本点一定也属于事件 B,则事件 A 发生必然导致事件 B 发生,称事件 A 是事件 B 的子事件,或者称事件 A 包含于事件 B,或者称事件 B 包含事件 A,其关系如图 1.2 所示,记作

$A\subset B$ 或 $B\supset A$.

对于任意事件 A,有 $\varnothing\subset A\subset\Omega$.

2. 事件的相等

如果事件 A 与事件 B 包含的样本点完全相同,则事件 A 包含事件 B 同时事件 B 也包含事件 A,这意味着事件 A 与事件 B 总是同时发生或同时不发生,则

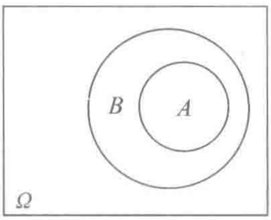
称事件 A 与事件 B 相等,记作 $A=B$.

3. 事件的并(union)

由事件 A 和事件 B 的样本点构成的事件称为它们的并(或和),其运算如图 1.3 所示,记作事件 $A\cup B$(或 $A+B$)。这意味着在一次随机试验中,事件 A 或事件 B 发生。

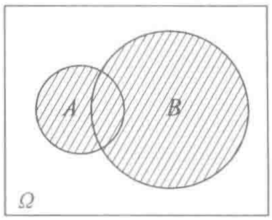
4. 事件的交(intersection)

由既属于事件 A 又属于事件 B 的样本点构成的事件称为它们的交(或积),其运算如图 1.4 所示,记作事件 $A\cap B$(或 AB)。这意味着在一次随机试验中,事件 A 和事件 B 同时发生。



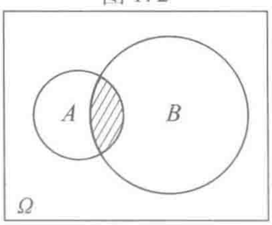
$A\subset B$

图 1.2



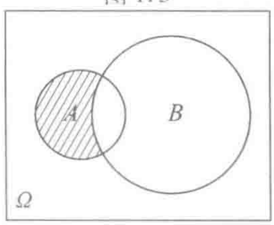
$A\cup B$

图 1.3



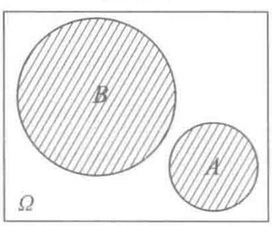
$A\cap B$

图 1.4



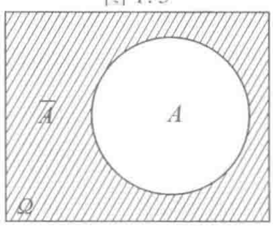
$A-B$

图 1.5



$A\cap B=\varnothing$

图 1.6



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

图 1.7

5. 事件的差

由属于事件 A 但不属于事件 B 的样本点构成的事件称为它们的差，其运算如图 1.5 所示，记作事件 $A-B$。这意味着在一次随机试验中，事件 A 发生但事件 B 不发生。

6. 互不相容事件

如果事件 A 与事件 B 不可能同时发生，也就是 AB 是不可能事件，则称事件 A 与事件 B 是**互不相容**的(incompatible)或互斥的，其关系如图 1.6 所示。这意味着 $AB=\varnothing$。否则，称事件 A 与事件 B 是**相容**的(compatible)。

例如，在掷骰子的试验中，事件“点数小于3”和“点数大于4”是互不相容事件。

7. 对立事件

由不属于事件 A 的样本点构成的事件称为事件 A 的**对立事件**(opposite)，其关系如图 1.7 所示，记作 $\overline{A}$。根据定义，有 $\overline{A}=\Omega-A$，$\overline{(\overline{A})}=A$。这意味着在一次试验中，事件 A 与事件 $\overline{A}$ 只能发生其中之一。

对立事件的另一种定义是，如果事件 A 与事件 B 满足：$A\cap B=\varnothing$，且 $A\cup B=\Omega$，则称 A 与 B 互为对立事件。

需要特别注意的是，互不相容事件(或互斥事件)不同于对立事件。对立事件一定是互不相容事件，但互不相容事件不一定是对立事件。

8. 可数个事件的并与交

事件的并与交运算可以推广到有限个或可列个事件，例如有事件 $A_1, A_2\cdots$，则称 $\bigcup\limits_{i=1}^{n}$ 为它们的有限并，称 $\bigcup\limits_{i=1}^{\infty}A_i$ 为它们的可列并，称 $\bigcap\limits_{i=1}^{n}$ 为它们的有限交，称 $\bigcap\limits_{i=1}^{\infty}A_i$ 为它们的可列交。

例 1.6　设 A,B,C 是某个随机现象中的三个事件，则

(1)事件“A 与 B 发生，C 不发生”可表示为：$AB\overline{C}=AB-C$；

(2)事件“A,B,C 中至少有一个发生”可表示为：$A\cup B\cup C$，正是它们的有限并或者说和事件；

(3)事件“A,B,C 中至少有两个发生”可表示为：$AB\cup AC\cup BC$；

(4)事件“A,B,C 中恰好有两个发生”可表示为:$AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$;

(5)事件“A,B,C 同时发生”可表示为:ABC,正好是它们的有限交或者说积事件;

(6)事件“A,B,C 都不发生”可表示为:$\overline{A}\,\overline{B}\,\overline{C}$;

(7)事件“A,B,C 不全发生”可表示为:$\overline{A}\cup\overline{B}\cup\overline{C}$.

这些结果也可以用维恩图来直观显示。

9. 完备事件组

设 $A_1,A_2,\cdots,A_n,\cdots$ 是有限或可数个事件,如果满足:

$A_iA_j=\varnothing,i\neq j,i,j=1,2,\cdots$,且 $\cup A_j=\Omega$,

则称 $A_1,A_2,\cdots,A_n,\cdots$ 是一个**完备事件组**。其实,$A_1,A_2,\cdots,A_n,\cdots$ 是对样本空间 Ω 的一个分割或者划分。显然,A 和 $\overline{A}$ 就构成一个完备事件组,也是对样本空间最简单的分割。

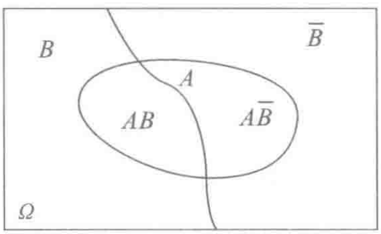


图 1.8　两个事件对样本空间的分割

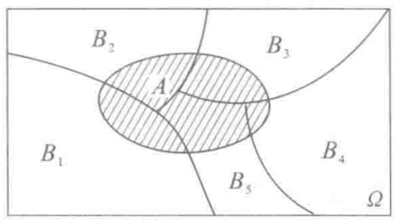


图 1.9　五个事件对样本空间的分割

八、随机事件的运算律

把样本空间的子集和事件等同起来后,事件的运算就类似于集合的运算。其中的运算律,可以归纳为:

1. 交换律

$A\cup B=B\cup A;A\cap B=B\cap A$

2. 结合律

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

3. 分配律

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

4. 对偶律

$\overline{A\cup B}=\overline{A}\cap\overline{B};\overline{A\cap B}=\overline{A}\cup\overline{B}$

这些运算律在对复合事件的分析中频繁地用到。

九、事件域

经常要讨论一些有联系的事件,这些事件组成的集合称为**集合类**,或者称为**事件族**,记为 $\mathscr{F}$. 如果这个事件族的构成满足下列条件,将是一个有意义的结构:

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

2. 若 $A\in\mathscr{F}$,则 $\overline{A}\in\mathscr{F}$;

3. 若 $A_1,A_2,\cdots\in\mathscr{F}$,则 $\bigcup\limits_{i=1}^{\infty}A_i\in\mathscr{F}$。

这时把事件族 $\mathscr{F}$ 称为一个事件域,或者称为 σ - 域或 σ 代数。

例 1.7 不难验证,下面几个事件族都是 σ - 域:

(1) $\mathscr{F}_1=\{\varnothing,\Omega\}$。这是 Ω 上最小的 σ - 域;

(2) $\mathscr{F}_2=\{\varnothing,A,\overline{A},\Omega\}$。这是由 A 生成的 σ - 域;

(3) $\mathscr{F}_3=\{A|A\subset\Omega\}$。这是由 Ω 的所有子集构成的,是 Ω 上最大的 σ - 域。

随着对样本空间分割的不同,形成的事件域就不同。在以后的讨论中,一般总默认 $\mathscr{F}$ 已经给定。

习题 1.1

1. 写出下列随机试验的样本空间 Ω:

(1) 记录一个班一次数学考试的平均分数(设以百分制计分)。

(2) 生产产品直到有 10 件正品为止,记录生产产品的总件数。

(3) 对某工厂出厂的产品进行检查,合格的记上"正品",不合格的记上"次品",如连续查出了 2 件次品就停止检查,或检查了 4 件产品就停止检查,记录检查的结果。

(4) 在单位圆内任意取一点,记录它的坐标。

2. 从某班学生中任选一名学生,设 $A=\{$选出的人是男生$\}$,$B=\{$选出的人是数学爱好者$\}$,$C=\{$选出的人是班干部$\}$,试问下列运算结果分别表示什么事件?