浙江省高等教育重点建设教材

概率论与数理统计

PROBABILITY AND MATHEMATICAL STATISTICS

许冰／主编

ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

概率论与数理统计

浙江省高等教育重点建设教材

概率论与数理统计

PROBABILITY AND MATHEMATICAL STATISTICS

许 冰／主 编

ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

图书在版编目(CIP)数据

概率论与数理统计／许冰主编．—杭州：浙江大学出版社，2007.8

ISBN 978-7-308-05431-7

Ⅰ.概… Ⅱ.许… 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 130168 号

概率论与数理统计

许　冰　主编

卓文新　何静慧　副主编

责任编辑	周卫群
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	浙江大学出版社电脑排版中心
印　　刷	富阳市育才印刷有限公司
开　　本	787mm×960mm　1/16
印　　张	12.25
字　　数	220 千
版 印 次	2007 年 8 月第 1 版　2007 年 8 月第 1 次印刷
印　　数	0001—4000
书　　号	ISBN 978-7-308-05431-7
定　　价	20.00 元

浙江大学出版社发行部邮购电话　（0571）88072522

总　序

妣健敏

随着高等教育大众化，成人高等教育面临的形势和任务也发生了很大的变化，成人高等教育日益呈现出新发展趋势：教育对象大众化；教育内容职业化；教育体系一体化；教育方式多样化；办学主体社会化。这些发展趋势，对成人高等教育的教学内容、教学方法、管理体制提出一系列改革与创新的要求，而教材改革则是其中最为迫切的改革要求。

经过细致的调查和分析，我们认为成人高等教育的教材，既要保证成人高等教育的规格，又要体现社会对成人高校的要求，才能适应已经变化和飞速发展的成人教育实际：首先，教材内容上，必须以实用性、应用性为主，以区别于普通高校教材侧重理论的特点。成人高等教育的培养目标已逐步转为广大在职者的学历补偿教育，职前职后的培训教育，因此，成人高等教育教材的理论深度，内容的难易程度要适中，同时偏向应用性、实用性。其次，在教材内容安排上，要分单元、分层次，充分考虑成人的学习特点，使教材内容前后连贯，以便从教材内容上充分体现对学习过程的层次控制。第三，在学习效果的检测上，以学生自己测试为主，将传统一次性考试变为对单元，学习时段的多层次考试。因此，材料要根据不同课时和单元配备习题成测试，以突

妣健敏　浙江大学副校长，教授、博导。

出学生自己检测学习效果的特点。

按照上述思想，我们聘请一线专家和经验丰富的学科带头人精心编写了这套成人高等教育教材。相信这套精心策划、认真编写出版的系列教材会得到广大院校的认可。

2006年12月

目 录

第1章 概率计算

未来因多变而不确定，才使世界充满多彩而神奇！

概率论，作为对不确定性的一种度量，将展示其科学而有效的方法和工具，用以揭开未来不确定性的面纱，实现人类探索未知的追求.

§1.1 随机事件的集合描写

随机现象：

下周末到园丁公园游玩，天气晴好还是阴雨？

商大公司的股票，明天的价格涨还是跌？

这些自然、人、自然与人，相互交融的不确定现象，虽然现在还未知发生哪种结果，但是，随着时间的推移，其出现的结果是确定可知的，如：天气好或坏，股价涨或跌，恋情结束年限.

这种事先未知，事后可知的现象，称为随机现象；

事先未知，事后出现的确定结果，称为随机事件（是一个过程！）；

事后所有可能出现的结果，组成一个样本空间.

为明确随机事件与样本空间的定义，便于数学的定量分析，现将它们作抽象的

集合描写：

用 $\Omega = \{\tilde{\omega}_t : t \in T\}$ 表示样本空间，则随机事件 A，可表示为 $A \subset \Omega$.

如：下周末天气 $\Omega=\{$晴好，阴雨$\}$，$A=\{$晴好$\}$

商大股票的价格 $\Omega=\{$涨，跌$\}$，$A=\{$跌$\}$

恋情结束 $\Omega=\{$3 年之内，3 至 6 年，6 年以上$\}$，$A=\{$3 年之内$\}$

这里，具体的随机事件与抽象集合相互对应表达，其结果是：不仅在于随机事件升华为抽象描写，而且使抽象描写的集合得到具体含意.

随机事件的具体含意赋予抽象集合生动的现实意义，而抽象集合描写使随机事件的数学度量成为可能.

由此，有必要描述，事件与集合相互对应表达的一些基本知识.

1. 事件间的运算 —— 并与交

并：事件 A 与 B 至少发生一个 $\Leftrightarrow A\cup B=\{\omega\in A$ 或 $\omega\in B\}$

交：事件 A 与 B 同时发生 $\Leftrightarrow A\cap B=\{\omega\in A$ 而且 $\omega\in B\}\triangleq AB$

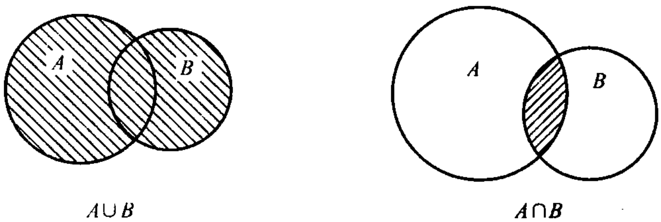


图1.1

2. 事件间的关系

(a) 包含与相等

包含：事件 B 发生必然导致 A 发生 $\Leftrightarrow B\subset A\Leftrightarrow\{\omega\in B\Rightarrow\omega\in A\}$

相等：$A=B\Leftrightarrow A\subset B$ 与 $A\supset B$ 同时成立.

图1.2

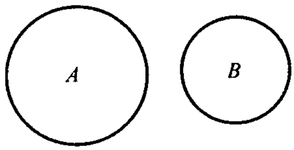
不可能事件和必然事件分别用：$\varnothing$ 和 Ω 表示.

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

(b) 互不相容与对立

A 与 B 是互不相容事件 $\Leftrightarrow A\cap B=\varnothing$.

A 的对立 $\triangleq \overline{A} = \{\omega \in \Omega$ 而且 $\omega \notin A\}$,
这时 $A \cap \overline{A} = \varnothing$,而且 $A \cup \overline{A} = \Omega$.



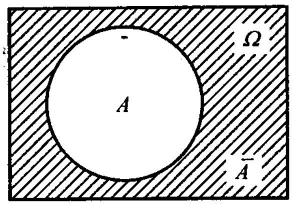


图1.3

直觉地,两事件对立,一定互不相容.反之未必,A 与 B 互不相容,不必要求 $A \cup B = \Omega$.

3. 若干公式

$A \cup B = B \cup A$ $\qquad$ $A \cap B = B \cap A$

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

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

例 1.1 跳伞表演员朝地面上跳伞,降落地有三个同心圆,半径分别为 1、2、3 米,用 A_i 记降落在半径为 i 的圆内,$i = 1,2,3$. 试说明下列事件含义:

(1) $\overline{A_3}$, (2) $A_1 \cup A_2 \cup A_3$, (3) $A_1A_2A_3$, (4) $\overline{A_1}A_2$

解:(1) 降落在半径为 3 米的圆外.

(2) 降落在半径为 3 米的圆内.

(3) 降落在半径为 1 米的圆内.

(4) 降落在半径为 1 米和 2 米的两个圆周围成的圆环内.

例 1.2 试用集合及运算符号,表示下列随机事件:

(1)A 发生,而 B 不发生 $\qquad$ (2)A、B、C 至少发生一个

(3)A、B、C 最多发生一个 $\qquad$ (4)A、B、C 中恰好发生两个

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

(4) $AB\overline{C} \cup \overline{A}BC \cup A\overline{B}C$

例 1.3 在电炉上安装了 4 个温控器,其显示温度的误差是随机的.在使用过程中,只要有两个温控器显示的温度不低于临界温度 t_0,电炉就断电.记 $T_{(1)} \leqslant T_{(2)} \leqslant T_{(3)} \leqslant T_{(4)}$ 为 4 个温控器显示的温度按递增顺序排列的温度值,试用 t_0 和 $T_{(i)}, 1 \leqslant i \leqslant 4$ 表示出"电炉断电"事件 E.

解：$E = \{T_{(3)} \geqslant t_0\}$

§1.2 随机事件的概率度量

直觉问题：

随机事件能否用数学来度量？

——两个统计实例

例 1.4 在众多英文小说中，表面上，小说内容，作家个性，使 26 个英文字符出现是随机的，但是，人们通过大量的实证研究，发现 26 个英文字符出现的频率是稳定的.

字母	空格	E	T	O	A	N	I	R	S
频率	0.2	0.105	0.072	0.0654	0.063	0.059	0.055	0.054	0.052
字母	H	D	L	C	F	U	M	P	Y
频率	0.047	0.035	0.029	0.023	0.0225	0.0225	0.021	0.0175	0.012
字母	W	G	B	V	K	X	J	Q	Z
频率	0.012	0.011	0.0105	0.008	0.003	0.002	0.001	0.001	0.001

数据引自 L. Brillouin, Science and Information Theory, New York, 1956.

例 1.5 尽管每个小孩在出生前不知是男还是女，但是，人类的延续一直能均衡男女出生比率. 这与历史上许多科学家重复抛掷一个均匀硬币的试验结果一致.

试验者	抛掷次数 n	正面数 m	频率 m/n
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

数据引自，格涅坚科，概率论教程，高等教育出版社.

随机现象产生的事件，其特点是

事先 —— 未知将产生何种结果 —— 随机性，

事后 —— 其结果是完全知道的 —— 确定性.

事先的随机性，不再是微积分中，确定性函数 $y = f(x)$ 所能描述的范围.

事后的确定性，使事先描述存在可能性，至少可以猜测哪种结果将产生的可

能性较大.

概率度量:

实例 1.4 和 1.5 表明随机事件可能做出度量.

—— 如何决定度量值?

通常称随机事件的度量值为概率值,其方法主要有:

(1) 统计定义

一个随机实验重复 n 次,其中 A 事件共发生了 f_A 次,则 A 事件发生的概率定义为

$$P(A) = f_A/n$$

叫做事件 A 发生的频率或称经验概率.

经验概率计算方式需要客观实验数据. 因此,也称为客观概率. 例如欲调查民众对于杭州城市东扩规划的看法. 通过对市民电话的民意调查. 令事件 A 表示赞成城市东扩规划. 在 100 位随访者中,有 65 位赞成,因此 $P(A) = 0.65$. 若样本增加到 1000 位时,发现有 714 位赞成. 此时 $P(A) = 0.714$.

以频率作为随机事件的概率度量,其缺点为,随机实验往往无法在环境不变的条件下长期进行,不同的调查对象,概率值会存在差异.

上面例子 1.4 和 1.5 中的频率值可以作为概率值的统计定义.

(2) 主观定义

有些事件发生的可能性,常由人们对此事件的经验或心里的感觉而判断. 如特殊疾病,很难根据其样本计算频率. 只有通过当事人或专业人士的经验认定,预测其发生的可能性. 又如 A 先生第一次参加百米径赛,依其以往的经验和认知,自我评估约有八成的夺冠,因此可设

$$P(A\text{ 先生胜}) = 0.8$$

即为一种认定的主观概率.

但是,经验或感觉不仅会随空间和时间的变化而改变,而且哪怕是在同一时空内也会因人而异. 因此经验或主观认定出来的概率,会颇多争议.

(3) 古典定义

若以 $\#\Omega$ 表示在样本空间 Ω 内,发生所有事件的计数(如个数、长度、面积和体积等),$\#(A)$ 表示属于 A 事件的计数. 则 A 事件发生的概率定义为

$$P(A) = \#(A)/\#(\Omega) \tag{1.1}$$

叫做随机事件发生的古典概率或称理论概率.

计算古典概率的方法,通常利用数学的工具,如排列组合知识,进行逻辑分析和抽象的数学推理.

作为说明，例如：

一个小孩出生是男孩的事件记作 A，A 的概率，历史上有许多统计学家做过大量的实证工作，得到结论是

$$P(A) = 22/43$$

这是一个典型的统计定义概率值.

现在，用概率古典定义，由合理的假设：

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

容易得到理论推导：

$$P(A) = 1/2$$

这可以算是一个典型的古典(理论)定义概率值.

例 1.6 投掷一颗骰子两次，所有可能的结果为

$\Omega = \{(1,1),(1,2),\cdots,(5,6),(6,6)\}$，此时 $\#(\Omega) = 36$.

若 A 表示点数和为 4 的事件集合，即 $A = \{(1,3),(3,1),(2,2)\}$.

则 A 事件发生的概率为

$$P(A) = \#(A)/\#(\Omega) = 3/36 = 1/12$$

例 1.7 生日问题 统计专业两个班共 64 个学生，问其中至少两个同学生日相同的概率有多大？

假设每人生日在一年 365 天中的每一天有相同的可能性：$\#\Omega = 365^{64}$，64 人都不同生日：$\#A = C_{365}^{64}64!$，那么 64 人生日都不同的概率为：

$$\frac{\#A}{\#\Omega} = \frac{C_{365}^{64}64!}{365^{64}}$$

由此，至少有两人生日相同的概率：

$$1 - \frac{C_{365}^{64}64!}{365^{64}} = 0.997$$

找同日出生的人不是一件很难的事！

人数	30	40	50
至少有两人生日相同的概率	0.71	0.89	0.97

例 1.8 配对问题

A a　　B b　　Cc　　D d　　E e　　F f

从上面成对因子中任取 4 个，求下面事件的概率：

1) 没有配对，2) 恰有一个配对

3) 恰有两个配对

$$p_1 = \frac{C_6^4C_2^1C_2^1C_2^1C_2^1}{C_{12}^4} = \frac{16}{32} \qquad p_2 = \frac{C_6^1C_2^2C_5^2C_2^1C_2^1}{C_{12}^4} = \frac{16}{33} = 0.48 \qquad p_3 = \frac{C_6^2C_2^2C_2^2}{C_{12}^4} = \frac{1}{3}$$

- 至少有一配对?
- 最多有一配对?

1777 年法国科学家蒲丰提出了下列著名问题.

例 1.9 平面上画着一些平行线,它们之间的距离都等于 a,向此平面任投一长度为 $l(l<a)$ 的针,试求此针与任一平行线相交的概率.

解 以 x 表示针的中点到最近的一条平行线的距离. φ 表示针与平行线的交角. 则有

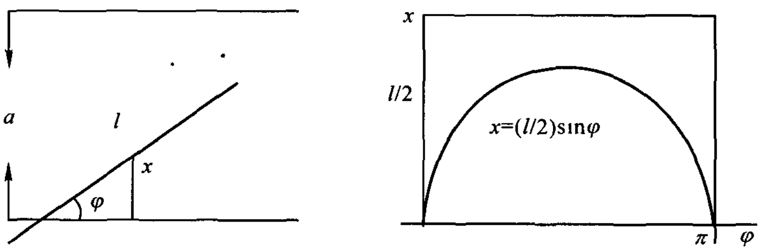


图1.4

$\Omega=\{(x,\varphi):0\leqslant x\leqslant a/2,0\leqslant\varphi\leqslant\pi\}$

$A=\{(x,\varphi):0\leqslant\varphi\leqslant\pi,0\leqslant x\leqslant(l/2)\sin\varphi\}$

于是

$$P(A)=\frac{\# A}{\# \Omega}=\frac{\frac{l}{2}\int_0^\pi \sin\varphi d\varphi}{\frac{1}{2}a\pi}=\frac{2l}{\pi a}$$

由于最后的答案与 π 有关,因此不少人想利用它来计算 π 的数值,其方法是投针 N 次,计算针与线相交的次数 n,再以频率值 n/N 作为概率 P 之值代人上式,求得

$$\pi=\frac{2lN}{an}$$

下表是一些试验的相关资料.

实验者	年份	针长	投掷次数	相交次数	π 实验值
Smith	1855	0.6	3204	1218	3.1554
Fox	1884	0.75	1030	489	3.1595
Lazzerini	1901	0.83	3408	1808	3.1415929

数据引自: Gridgeman. N. T. Geometric probability and the number π. Scripta Mathematica, 25 (1960), 183—195.

值得注意的是这里采用的方法:建立概率模型,与感兴趣的量——如π——相关,然后设计适当的随机实验,并通过这个实验的结果来确定这个量.

现在,利用计算机,按照上述思想建立起的 Monte-Carlo 方法,已有广泛的应用.

例1.10 公平分奖 甲、乙两人均为九段围棋棋手,先胜三局者赢得10万元奖金.现已赛三局,甲2胜1负而因故终止比赛,问如何分10万元奖金,才算公平?

解

(1) 平均各分5万 甲欠公平

(2) 考虑已赛出的成绩 甲6.7万 乙3.3万

(3) 已赛出和未赛出的成绩一起考虑

记甲 胜:1 负:0

已赛出,事实上,只是3种可能:110,101,011 中的1种

将赛出,事实上,只有3种可能:1,01,00 中的1种

注意关键词:可能

已赛出甲2胜1负的3种(可能)情形之一:110,101,011

未来将赛出的两局的4种可能结果:10,11,01,00,这四种结果的出现机会均等,其中前三种结果出现,甲将获胜,因此甲胜3/4,乙胜1/4.

据此,公平分法:甲7.5万元,乙2.5万元

概率的古典定义,从公式(1.1)可见,至少要求:

1) 样本空间整体的可计算性:$\#\Omega=\#\{\bar{\omega}_t:t\in T\}<\infty$

——分母计量时的存在性.

2) 样本空间内各基本事件发生有相等的可能性:$P\{\bar{\omega}_t\}=P\{\bar{\omega}_s\},t,s\in T$

——分子与分母各自计量时的可加性.

为使概率计算的分析推演过程,合乎人类的共同认知与思考水准,必须给定一些概率公理,以便合理进行理论与实务的演算.因为在认知一致的架构上,概率的表示与应用才有意义.

1933年,前苏联柯尔莫哥洛夫院士提出了概率公理化结构,使概率论成为一门严谨的科学理论.

概率公理

在样本空间Ω中,不是所有的$A\subset\Omega$,都有唯一确定的概率值$P(A)$,因此事先需要做一个限制,建立概率空间(Ω,F,P),使得任意属于F的事件A,有唯一确定的概率值$P(A)$,这里

$F=\{A:A\subset\Omega\}$ 限制为满足下述条件的σ域: