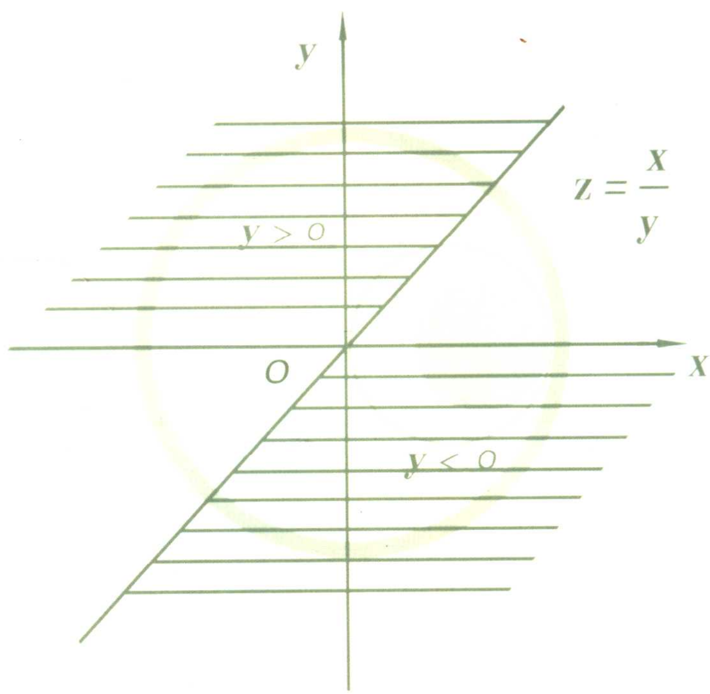
高等教材

全国高等农林院校教材

概率论与数理统计

吴瑞武　郑大川　主编



中国林业出版社

概率论与数理统计

全国高等农林院校教材

概率论与数理统计

吴瑞武　郑大川　主编

中 国 林 业 出 版 社

内 容 简 介

本教材主要内容分为两部分：概率论部分和数理统计部分。概率论内容包括概率论的基本概念，一维和多维随机变量及其分布，数字特征等。数理统计包括样本及抽样分布，参数估计，假设检验，方差分析和回归分析等。结合编者多年的教学经验，精心配置了适当的例题和丰富的习题，每章后的习题都附有参考答案，以便读者自学。

图书在版编目（CIP）数据

概率论与数理统计/吴瑞武，郑大川主编. —北京：中国林业出版社，2007. 1
全国高等农林院校教材
ISBN 978-7-5038-4475-1

Ⅰ. 概… Ⅱ. ①吴… ②郑… 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（2006）第 151794 号

中国林业出版社·教材建设与出版管理中心

策划编辑：牛玉莲　　责任编辑：杜建玲
电话：66188720　66170109　　传真：66170109

出版发行　中国林业出版社（100009　北京西城区德内大街刘海胡同 7 号）
E-mail：cfphz@ public. bta. net. cn　电话：66184477
网　址：http：//www. cfph. com. cn
经　销　新华书店北京发行所
印　刷　北京百善印刷厂
版　次　2007 年 1 月第 1 版
印　次　2007 年 1 月第 1 次
开　本　850mm × 1168mm　1/16
印　张　11. 5
字　数　242 千字
定　价　17. 00 元

全国高等农林院校“十一五”规划教材

《概率论与数理统计》编写人员

主　　编　吴瑞武　郑大川

副 主 编　刘　雯　杨建红

编写人员　(按姓氏笔画排序)

左珊珊　刘　雯　杨建红　吴瑞武

郑大川　夏梓祥　黄勇林　梁　兵

前 言

概率论与数理统计是数学分支中一门应用性很强的学科，它在科学研究及工农业生产、金融、医药卫生、经济管理等方面都有着广泛的应用，其独特的作用也越来越受到人们重视。概率论与数理统计也是高等院校各类专业必须开设的一门重要课程。

本教材在内容编写上力求理论体系完整，结构简明，条理清晰。突出基本概念、基本原理，强调对学生应用能力的培养。

本教材主要内容分为两部分：概率论部分和数理统计部分。概率论内容包括概率论的基本概念，一维和多维随机变量及其分布，数字特征等。数理统计包括样本及抽样分布，参数估计，假设检验，方差分析和回归分析等。结合编者多年的教学经验，精心配置了适当的例题和丰富的习题，每章后的习题都附有参考答案，以便读者自学。

本书由云南农业大学吴瑞武、郑大川担任主编，刘雯、杨建红担任副主编，参加编写的有云南农业大学左珊珊、夏梓祥，云南机电职业技术学院黄勇林，成都电子机械高等专科学校梁兵。

限于编者水平，书中难免有不妥和错误之处，恳请专家、读者指正。

编　者

2006. 10

目 录

前 言

第1章　概率论的基本概念

1.1　随机试验与随机事件

1.1.1　必然现象与随机现象

概率论与数理统计是研究和揭示随机现象统计规律的一门数学学科。在自然界和人类社会实践活动中，人们观察到的现象是多种多样的，但是大体可以分为两类。一类现象在一定的条件下是必然发生的，称这类现象为必然现象。例如，在标准大气压下，水加热到100℃必然沸腾；磁极相同的磁铁相互排斥；在没有外力作用下，作匀速直线运动的物体的运动状态保持不变，等等。但在自然界和社会现象中还广泛存在着与必然现象有着本质区别的另一类现象。例如用同一仪器测量同一物体的质量，所得结果彼此总是略有差异；在某生产线上用同种工艺生产出来的灯泡的寿命也有差异；抛一枚骰子，观察出现的点数，可能出现1，2，…，6点，但在每次抛掷之前无法肯定抛掷的结果是什么。总之以上所举的这些现象的一个共同的特点是：在基本条件不变的情况下，一系列的试验或观察会得到不同的结果。在个别的试验或观察中，它可能会出现这样的结果，也可能会出现那样的结果，而且在试验或观察之前不能预先知道确切的结果。但在人们长期实践积累和研究中发现，虽然对这类现象孤立地、少量地观察时似乎是偶然的、毫无规律可言的，但在大量重复试验或观察下，它的结果却呈现出某种固有的规律性。例如，多次重复抛一枚硬币得到正面朝上和反面朝上的结果大致各占一半；多次测量一个物体的质量，测量值的平均值总在某固定常数附近波动，等等。这种在大量重复试验或观察中所呈现出来的规律性称为统计规律性。而这类在个别试验或观察中呈现不确定性，而在大量重复试验或观察中又具有统计规律性的现象称为随机现象。

正如恩格斯所言："在表面上是偶然性在起作用的地方，这种偶然性始终是受内部的隐蔽着的规律支配的，而问题只是在于发现这些规律。"概率论与数理统计正是探索看似偶然的随机现象中的必然规律的有力工具。

1.1.2　随机试验

对随机现象的研究必然要联系到对客观事物进行"观察"或"试验"。在这里"试验"广义地包含各类科学实验以及对自然现象、社会现象的观察。下面举一些试验的例子。

【例1.1】 E_1：抛一枚质地均匀的硬币，观察其落地后哪面朝上；

E_2：抛一枚均匀的骰子，观察其出现的点数；

E_3：用 100 粒种子做发芽试验，观察其发芽的粒数；

E_4：记录某人在靶场对 10 环靶进行一次射击的环数；

E_5：从一批电池中任意抽取一只，测试其寿命。

上面所举例子有着随机现象共同的特点。例如试验 E_1 有两种可能结果，正面朝上或反面朝上，但在抛掷之前无法预知，且这个试验可以在相同的条件下重复进行。又如试验 E_5 知道电池的寿命(单位：h) $t \geqslant 0$，但在测试之前不能确定其寿命到底有多长，这一试验也可在相同条件下重复进行。总而言之，这些试验概括起来具有以下 3 个特点：

①可以在相同的条件下重复进行；

②每次试验出现的可能结果不止一个，进行一次试验之前无法确定出现的结果；

③可以事先预知试验的所有可能结果。

在概率论中，将具有上述 3 个特点的试验称为随机试验。概率论与数理统计研究的试验指的是随机试验。随机试验是产生随机现象的过程，人们是通过研究随机试验来研究随机现象的。

1.1.3　样本空间、样本点

定义 1　某一随机试验的可能结果的全体称为该随机试验 E 的样本空间，记为 Ω。而样本空间的元素，即 E 的每个可能结果，称为样本点。

在【例 1.1】中试验 E_k($k=1,2,\cdots,5$)对应的样本空间 Ω_k 为：

$\Omega_1=\{$正面朝上，反面朝上$\}$

$\Omega_2=\{1,2,3,4,5,6\}$

$\Omega_3=\{$有 i 粒发芽$(i=1,2,\cdots,100)\}$

$\Omega_4=\{$射中 i 环$(i=1,2,\cdots,10)\}$

$\Omega_5=\{t \mid t \geqslant 0\}$

从【例 1.1】中可以看到，样本空间的元素可以是有限多个。也可以是无穷多个。样本空间的元素取决于试验的目的，不同的随机试验其样本空间是不同的；即使是同一随机试验，试验的目的不同，样本空间也不一样。如对于 E_2，抛一枚均匀的骰子，观察其出现的点数，则其样本空间为 $\Omega_2=\{1,2,3,4,5,6\}$；同样的试验如果换为观察骰子出现点数的奇偶性，样本空间相应变为 $\Omega=\{$奇数，偶数$\}$。但是，无论怎样构造样本空间，作为样本空间的元素——样本点，必须具备两个基本属性：①互斥性，即任意两个样本点不会在同一试验中出现；②完备性，即每次试验一定至少出现一个样本点。

1.1.4　随机事件

有了样本空间的概念，就可以定义随机事件了。还是先从考察一个例子开始。

【例 1.2】　甲、乙二人掷一枚骰子比较点数的大小。假设甲先掷得 4 点，轮到乙掷，则乙掷后可能输，可能平，也可能赢。如果把这 3 个可能结果分别记作 $A,B,$

C,则

乙输—— $A=\{1,2,3\}$

平局—— $B=\{4\}$

乙赢—— $C=\{5,6\}$

而该随机试验的样本空间为 $\Omega=\{1,2,3,4,5,6\}$，显然 A,B,C 均为 Ω 的子集。

一般地,称随机试验 E 的样本空间 Ω 的子集为 E 的随机事件,简称事件。在每次试验中,当且仅当这一子集中的一个样本点出现时,就称该事件发生。

特别地,由一个样本点组成的单点集,称为基本事件。例如,试验 E_1有两个基本事件{正面朝上}和{反面朝上};试验 E_2有 6 个基本事件 $\{1\}$,$\{2\}$,…, $\{6\}$ 。

样本空间也作为一个事件,是 Ω 自身的子集,而且在每次试验中必然出现 Ω 中的某个样本点,也就是在每次试验中它总是发生的,所以常称 Ω 为必然事件。类似地,空集$\varnothing$也作为一个事件,空集不包含任何样本点,也可看做 Ω 的子集,而且在每次试验中都不发生,称为不可能事件。

1.1.5 事件的关系与运算

在实际生活中,往往需要同时考察几个在同样条件下的事件以及它们之间的关系,从而更深刻地认识事件的本质。下面讨论事件间的关系及事件的运算。由于事件是一个集合,因此事件间的关系及运算自然可以按照集合之间的关系和运算来处理。

设试验 E 的样本空间为 Ω ,而 $A,B,C,A_i(i=1,2,\cdots,n)$ 为 E 的事件,它们之间有下列关系和运算:

①若 $A\subset B$,则称事件 B 包含事件 A ,这时事件 A 发生必然导致事件 B 发生;若 $A\subset B$ 且 $B\subset A$,即 $A=B$,则称事件 A 与事件 B 相等。

②事件 $A\cup B$ 称为事件 A 与事件 B 的和事件。当且仅当 A,B 中至少有一个发生时,事件 $A\cup B$ 发生。

类似地,当 n 个事件 $A_1,A_2,\cdots,A_n$中至少有一个发生时,称和事件 $A_1\cup A_2\cup\cdots\cup A_n=\bigcup_{i=1}^{n}A_i$ 发生。

③事件 $A\cap B$ 称为事件 A 与事件 B 的积事件(或交)。当且仅当 A,B 同时发生时,事件 $A\cap B$ 发生, $A\cap B$ 也可记为 AB 。

类似地,当 n 个事件 $A_1,A_2,\cdots,A_n$中同时发生时,称积事件 $A_1\cap A_2\cap\cdots\cap A_n=\bigcap_{i=1}^{n}A_i$ 发生。

特别地,当 $A\cap B=\varnothing$时,称事件 A 与事件 B 为互不相容事件。即事件 A 与事件 B 不能同时发生。类似地,若 $A_i\cap A_j=\varnothing(i\neq j;i,j=1,2,\cdots,n)$,则称 n 个事件 $A_1,A_2,\cdots,A_n$两两互不相容。基本事件是两两互不相容的。

④事件 $A-B$ 称为事件 A 与事件 B 的差事件,当且仅当事件 A 发生而事件 B 不发生。

⑤若 $A\cup B=\Omega$ 且 $A\cap B=\varnothing$,称事件 A 与事件 B 互为逆事件(对立事件)。即对于每次试验而言,事件 A 与事件 B 不能同时发生,但必有一个发生且仅有一个发生。A 的对立事件记为 $\bar{A}$,$\bar{A}=\Omega-A$ 。

用图 1-1 ~ 图 1-6 可以直观地表示以上事件之间的关系与运算。

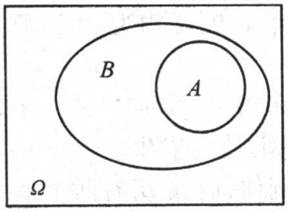


图 1-1 $A\subset B$

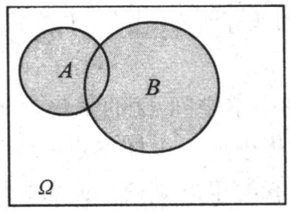


图 1-2 $A\cup B$

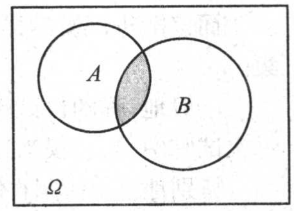


图 1-3 $A\cap B$

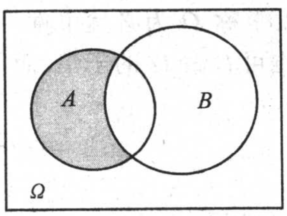


图 1-4 $A-B$

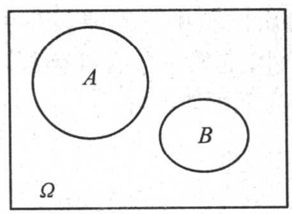


图 1-5 $A\cap B=\varnothing$

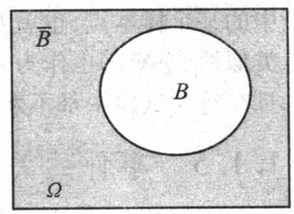


图 1-6 $\bar{B}$

在进行事件的运算时,经常用到以下规律。设 A,B,C 为事件,则有

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(De Morgan)律:$\overline{A\cup B}=\bar{A}\cap\bar{B}$;$\overline{A\cap B}=\bar{A}\cup\bar{B}$ 。

【例 1.3】 掷一枚骰子,记出现点数小于 5 的事件为 A ,记出现奇数点的事件为 B ,求 $A\cup B$,$A\cap B$,$A-B$,$\overline{A\cup B}$。

解 据题意 $A=\{1,2,3,4\}$,$B=\{1,3,5\}$,则

$$A\cup B=\{1,2,3,4,5\}$$
$$A\cap B=\{1,3\}$$
$$A-B=\{2,4\}$$
$$\overline{A\cup B}=\bar{A}\cap\bar{B}=\{5,6\}\cap\{2,4,6\}=\{6\}$$

【例 1.4】 设有事件 A,B,C ,写出下列事件:

① A,B,C 至少有一个不发生;

② A,B,C 都不发生;

③ A,B,C 恰有一个发生;

④ A,B,C 恰有两个发生。

解

① A,B,C 至少有一个不发生:$\bar{A}\cup\bar{B}\cup\bar{C}$ (或 $\overline{ABC}$);

② A,B,C 都不发生:$\bar{A}\bar{B}\bar{C}$;

③ A,B,C 恰有一个发生：$A\bar{B}\bar{C}\cup\bar{A}B\bar{C}\cup\bar{A}\bar{B}C$；

④ A,B,C 恰有两个发生：$\bar{A}BC\cup A\bar{B}C\cup AB\bar{C}$ 。

1.2 频率与概率

1.2.1 概率的直观意义

由于随机事件是样本空间的一个子集，而且样本空间可以包含不止一个事件，在观察一个随机试验的各种事件时，一般会发现，并非所有事件出现的可能性都相等，有些事件出现的可能性大些，有些事件出现的可能性则小些。

在生产实践中，研究随机现象时不仅要知道可能会出现哪些事件，更重要的是要了解这些事件出现可能性的大小。例如，根据工厂各部门的机器设备发生故障的可能性的大小，就可以合理安排生产计划、配备设备管理和维修人员。又如，为了设计能让打字更便捷的计算机键盘，人们需要知道每一个英文字母的使用率，等等。因此需要有一个刻画事件发生可能性大小的数量指标，这个指标称为随机事件的概率，常用 p 表示。

概率产生的背景是事件发生的频率。为了给出概率的统计定义，首先讨论一下随机事件的频率。

1.2.2 随机事件的频率及概率的统计定义

定义 1 在相同的条件下，重复 n 次试验 E，在这 n 次试验中，事件 A 发生的次数 m 称为事件 A 发生的频数，比值 $\frac{m}{n}$ 称为事件 A 发生的频率，并记为 $f_n(A)$，即 $f_n(A)=\frac{m}{n}$。

频率具有以下性质：

① $0\leqslant f_n(A)\leqslant 1$；

② $f_n(\Omega)=1$；

③ 若事件 $A_1,A_2,\cdots,A_k$ 是两两互不相容事件，则

$$f_n(A_1\cup A_2\cup\cdots\cup A_k)=f_n(A_1)+f_n(A_2)+\cdots+f_n(A_k)$$

可见，$f_n(A)$ 是与 A,n 有关的一个数，它的值具有随机性。但对于大量重复试验来说，随机试验的结果具有明显的统计规律性。因此，一般而言，当 n 很大时，$f_n(A)$ 是趋于稳定的。

【例 1.5】 抛一枚质地均匀的硬币观察其出现正面的次数，当抛硬币的次数不多时，正面出现的频率是不稳定的，但是随着抛掷次数的增多，频率越来越明显地呈现出稳定性。历史上曾有人做过大量多次抛硬币的试验，表 1-1 是试验的部分结果。

表 1-1

试验者	抛硬币的次数 n	出现正面的次数 m	$f_n(A)$
德·摩根(De morgan)	2048	1061	0.5181
蒲丰(Buffon)	4040	2048	0.5069
皮尔逊(K. Pearson)	12000	6019	0.5016
皮尔逊(K. Pearson)	24000	12012	0.5005

可见,当试验次数充分多时,正面出现的频率在0.5附近摆动,并随着 n 的增大逐渐稳定于0.5。

【例1.6】 考察某种子的发芽率。从一大批种子中抽取10批种子做发芽试验,其结果如表1-2所示。

表 1-2

种子总粒数	2	5	10	70	130	310	700	1500	2000	3000
发芽种子数	2	4	9	60	116	282	639	1339	1806	2715
发芽种子频率	1	0.8	0.9	0.857	0.892	0.910	0.913	0.893	0.903	0.905

从表1-2可以看出,发芽率在0.9附近摆动。

从上述两例的统计结果可以看出:当 n 较小时,频率 $f_n(A)$ 在0与1之间随机波动,其幅度较大;而当 n 逐渐增大时,频率 $f_n(A)$ 逐渐稳定于某个常数;对于每一个事件 A 都有这样一个客观存在的常数与之相对应。频率具有稳定性这一事实,不断地为人类的实践所证实,它揭示了隐藏在随机现象中的统计规律性。因而用频率稳定值来表示事件发生的可能性的大小是合适的。称这个"频率的稳定值"为事件发生的概率。频率稳定于概率的事实说明了随机现象中的偶然性与必然性的辩证的统一。下面给出概率的统计定义。

定义2 在相同的条件下,重复进行 n 次同一试验,事件 A 发生的频率 $f_n(A)$ 在某个常数值 p 附近摆动。而且一般来说这种摆动的幅度会随着试验次数的增多而越来越小,则称频率的稳定值 p 为事件 A 发生的概率,简称为事件 A 的概率,记作

$$P(A)=p$$

由频率出发所定义的事件 A 的概率常称为统计概率,所以定义2也就称为概率的统计定义。

虽然事件的频率与概率都是事件出现可能性大小的度量,但是频率是试验值,依赖于试验次数,而且即使试验次数相同,频率也可能取不同的值,频率只具有相对的稳定性。概率是先于试验而客观存在的理论值,它是一个确定的常数,其大小取决于事件本身固有的规律性。概率的统计定义是在总结统计资料的基础上给出的,反映了概率的统计性质,比较直观。但是在进行理论研究和实际应用中,不可能对每一个事件都做大量的重复试验以从中获得频率的稳定值,也不知道试验的

次数 n 该取多大才能获得频率的稳定值。同时如果要使满足的 n 足够大,需要做大量的重复试验,即使这样也不能保证每次试验的条件都完全一致。这是概率的统计定义的不足之处。

1.2.3 概率的公理化定义及其性质

对事件的概率给出一般定义的最好方法是用一组公理来确定其真实含义。下面给出的概率定义是柯尔莫哥洛夫于 1933 年提出的。这个定义明确了概率的基本概念,使其成为一个严谨的数学分支,对概率论的迅速发展起到了积极的推动作用。

定义 3 设 E 是随机试验,Ω 是它的样本空间。若 E 的每一事件 A 都对应一个实数 $P(A)$,如果 $P(A)$ 满足下面 3 条公理:

①非负性: $P(A) \leqslant 0$;

②规范性: $P(\Omega)=1$;

③可列可加性:若事件 $A_1,A_2,\cdots,A_k,\cdots$ 是两两互不相容事件,则

$$P(A_1 \cup A_2 \cup \cdots \cup A_k \cup \cdots)=P(A_1)+P(A_2)+\cdots+P(A_k)+\cdots$$

$P(A)$ 称为事件 A 的概率。

由概率的定义,可以推得概率的一些重要性质。

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

证 在公理③中,取 $A_n=\varnothing(i=1,2,\cdots,n)$,则有

$$P(\varnothing)=P(\varnothing)+P(\varnothing)+\cdots+P(\varnothing)+\cdots$$

根据公理①, $P(\varnothing)$ 是非负实数,所以必有 $P(\varnothing)=0$。

性质 2 (有限可加性)若事件 $A_1,A_2,\cdots,A_n$是两两互不相容事件,则

$$P(A_1 \cup A_2 \cup \cdots \cup A_n)=P(A_1)+P(A_2)+\cdots+P(A_n)$$

证 在公理③中,取 $A_i=\varnothing(i=n+1,n+2,\cdots)$,则有

$$\begin{aligned}P(A_1 \cup A_2 \cup \cdots \cup A_n) &= P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \varnothing \cup \varnothing \cup \cdots) \\ &= P(A_1)+P(A_2)+\cdots+P(A_n) + P(\varnothing) \\ &\quad + P(\varnothing) + \cdots \\ &= P(A_1)+P(A_2)+\cdots+P(A_n)\end{aligned}$$

公理③与性质 2 揭示了概率的最重要的性质——可加性,它是研究概率的基础。公理③是可列可加性,性质 2 是有限可加性。

性质 3 若 A 与 $\bar{A}$ 互为对立事件,则

$$P(A)=1-P(\bar{A})$$

证 因为 $A \cup \bar{A}=\Omega$,且 $A\bar{A}=\varnothing$,基于性质 2 有

$$1=P(\Omega)=P(A+\bar{A})=P(A)+P(\bar{A})$$

所以 $P(A)=1-P(\bar{A})$。

性质 3 可以用文字表述为“对立事件的概率之和为 1”,这个性质在概率计算中经常用到。

性质 4 设 A,B 是两个事件,若 $A \subset B$,则