概率统计

曹贤通　李海峰　主编

黄河水利出版社

概率统计

曹贤通　李海峰　主编

黄河水利出版社

图书在版编目（CIP）数据

概率统计/曹贤通，李海峰主编. -郑州：黄河水利出版社，1999. 8
ISBN 7-80621-285-X

Ⅰ.概… Ⅱ.①曹… ②李… 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(1999)第 19369 号

责任编辑:杜亚娟	封面设计:谢 萍
责任校对:赵宏伟	责任印制:常红昕

出版发行:黄河水利出版社
地址:河南省郑州市顺河路黄委会综合楼 12 层 邮编:450003
发行部电话:(0371)6302620 传真:6302219
E-mail:yrcp@public2.zz.ha.cn
印 刷:河南地质彩色印刷厂

开 本:850mm×1168mm 1/32	印 张:9
版 别:1999 年 8 月 第 1 版	印 数:1-2 500
印 次:1999 年 8 月 郑州第 1 次印刷	字 数:226 千字

定 价:16.80 元

内 容 提 要

本书为高等工业院校概率论与数理统计教材. 全书共分十章. 前四章是概率论的主要内容, 包括: 随机事件及其概率, 随机变量及其分布, 多维随机变量, 随机变量的数字特征; 后六章是数理统计的主要内容, 包括: 数理统计的基本概念, 参数估计, 假设检验, 回归分析, 正交试验设计, 方差分析. 本书的编写主要是参照高等学校工科数学课程教学指导委员会于 1995 年修订的《概率论与数理统计课程教学基本要求》进行的, 它适用于基本要求中概率多、统计少及概率少、统计多两种类型, 因而高等工业院校各个专业都能选用.

本书叙述详细, 结构紧凑, 行文流畅, 内容精练, 例题较多, 习题按学时及小节划分, 书末附有答案, 便于教学. 本书还可供各类职业高等学校和工程技术人员阅读参考.

前　言

概率统计是概率论与数理统计的总称，它是研究大量发生的随机现象规律性的科学．所谓随机现象，即在一定条件下可能发生也可能不发生的不确定现象．例如，由于许多偶然因素的影响，天气的变化是无常的（有不确定性）；又如，生物种群的出生率和死亡率都是不确定的，因而种群未来的数量也有随机性；再如，要确定仓库中商品的订货策略，由于市场对商品的需求是不确定的，仅遵从某种随机规律，这样就必须建立一个随机性的存贮模型，等等．正是由于随机现象的广泛存在和深刻影响，使得人们早就开始对它进行观察、分析与研究，其最原始的模型来自赌博等；而且早在17世纪中叶已由一批著名数学家、力学家从理论上提出有关概念和着手建立古典概率的基础，并逐步形成了概率论这门学科．随着社会实践的发展，以概率论为基础而应用于各个方面的随机性数学——数理统计便应运而生，并且获得了长足的进步．至今，概率论与数理统计已经成为一门独立的应用性学科，在人类生活（包括政治、军事、经济、生产、生态、工程技术、科学研究等）的各个方面均起到了越来越重要的作用，并且仍在不断充实与发展．因此，概率统计课程的教学愈来愈受到高等工业院校的重视，对教材的要求也愈来愈高．正是在这种情况下，根据高等工业院校教学的需要，结合当前的实际，在积累多年教学经验的基础上，我们编写了这本试用教材，供本科各专业及部分专科选用．

本书内容主要包括传统的概率论基础（前四章）和数理统计

基础（后六章），而且以数理统计为主．各个部分的编写尽量做到既简要明了、通俗易懂，又不失严格的理论性，特别注意介绍每个重要概念的实际背景及其统计意义．在精练地提出有关概念、方法及其应用时，从理论上尽量给出完整的证明，同时提供大量的实际例子；在习题中则注意收集充足的来自于生活的各种问题．希望读者学习本书后，不仅对概率统计的基本理论和计算方法能够理解与掌握，而且对随机数学特有的思维方式有所了解，从而对建立相应的数学模型分析与解决实际问题有所裨益．

本书由郑州纺织工学院和郑州粮食学院的部分教师共同编写，并由曹贤通、李海峰担任主编，吴晓非、江世景、宋长明、王瑞卿、张宏伟、高桂芬、霍振宏、皮上超、陈金环、郭森林、周瑞芳、黄勉之任副主编．完稿后，全书由李海峰、曹贤通二位教授整理、审定．本书的编写得到了郑州纺织工学院有关部门领导和黄河水利出版社编辑同志的大力支持与帮助，谨致谢意．

限于编者水平，不妥之处在所难免，希望广大读者与专家批评指正．

编　者

1999 年 6 月

目　　录

第一章　随机事件及其概率……………………………………（1）

§1　随机试验、随机事件、样本空间 …………………………（1）

§2　频率与概率、古典概型……………………………………（7）

§3　条件概率…………………………………………………（14）

§4　事件的独立性、独立试验概型 …………………………（20）

第二章　随机变量及其分布 ……………………………………（27）

§1　随机变量的概念…………………………………………（27）

§2　离散型随机变量…………………………………………（27）

§3　随机变量的分布函数……………………………………（34）

§4　连续型随机变量…………………………………………（39）

§5　随机变量的函数的分布…………………………………（53）

第三章　多维随机变量 …………………………………………（59）

§1　二维随机变量及其分布函数……………………………（59）

§2　边缘分布…………………………………………………（68）

§3　条件分布…………………………………………………（74）

§4　随机变量的相互独立性…………………………………（80）

§5　两个随机变量函数的分布 …………………… (86)
第四章　随机变量的数字特征 …………………… (97)
§1　随机变量的期望 …………………… (97)
§2　随机变量的方差 …………………… (115)
§3　协方差与相关系数 …………………… (122)
§4　大数定律和中心极限定理 …………………… (128)
第五章　数理统计的基本概念 …………………… (132)
§1　总体、样本及其分布函数 …………………… (132)
§2　统计量及其分布 …………………… (139)
第六章　参数估计 …………………… (150)
§1　点估计的矩法与极大似然法 …………………… (150)
§2　无偏性与有效性 …………………… (158)
§3　区间估计 …………………… (166)
第七章　假设检验 …………………… (178)
§1　假设检验的基本概念 …………………… (178)
§2　单正态总体参数的假设检验 …………………… (180)
§3　双正态总体参数的假设检验 …………………… (185)
§4　拟合优度的 χ^2 检验 …………………… (188)

第八章 回归分析……………………………………………………(193)
§1 一元线性回归 …………………………………………(194)
§2 多元线性回归 …………………………………………(201)
第九章 正交试验设计……………………………………………(208)
§1 正交表及其用法 ………………………………………(208)
§2 混合水平的正交设计 …………………………………(215)
§3 有交互作用的正交设计 ………………………………(219)
第十章 方差分析…………………………………………………(225)
§1 单因素方差分析 ………………………………………(225)
§2 双因素方差分析 ………………………………………(236)
附 录 概率统计常用数表…………………………………………(249)
习题答案……………………………………………………………(264)

第一章　随机事件及其概率

本章首先引入随机事件的概念，给出随机事件概率的定义，进而给出在古典概型意义下随机事件概率的运算规律，最后介绍伯努利试验.

§1　随机试验、随机事件、样本空间

一、随机试验

我们遇到过各种各样的试验.在此，试验是一个意义广泛的术语，它可以是各种各样的科学试验，也可以认为是对某事物的某一特征的观察.举例说明如下：

(1)将一枚硬币连续抛两次，观察正、反面出现的情况.

(2)掷一颗匀称的骰子，观察出现的点数.

(3)记录某传呼台一分钟内接到的呼唤次数.

(4)一口袋中装有红白两种颜色的乒乓球，从袋中任取一只球，观察它的颜色.

(5)在一批灯泡中任取一只，测试其寿命.

(6)记录某十字路口在半小时内所通过的车辆的数目.

上面我们举了 6 个例子，它们有共同的特点.如试验(1)，它有 4 种可能的结果(正、正，正、反，反、正，反、反)，但在我们抛出之前并不能确定会出现哪一种情况；这样的试验可在相同的条件下重复地进行.又如试验(2)，它有 6 种可能的结果，即可能出现的点数为 1,2,…,6 中之一，但在投掷之前不能确定会出现几点；这个试验也可以在相同条件下重复地进行.对其余 4 个试验进行分析可

知，它们也有上述特点.

综上所述，如果试验具有以下几个特点：

E_1：可以在相同的条件下重复地进行；

E_2：每次试验的可能结果不只一个，且能事先知道试验的所有可能结果；

E_3：进行一次试验之前不能确定哪一种结果会出现.

则我们在概率论中称之为**随机试验（简称试验）**.

概率论的主要任务就是通过随机试验来研究随机现象及其规律性.

二、随机事件

在一次试验中，可能出现也可能不出现，而在大量重复试验中却具有某种规律性的结果（或现象），称为此随机试验的**随机事件**，简称**事件**，常用大写字母 $A,B,C,\cdots$ 来表示.

在一随机试验中，它的每一个可能出现的结果都是一个随机事件，它们是这个试验的最简单的随机事件（具有“不能再分性”），我们称这些简单的随机事件为**基本事件**. 由多个基本事件组成的事件称为**复合事件**. 如试验(2)中“出现 1 点”，“出现 2 点”，…，“出现 6 点”均为基本事件；另一类事件，如“出现偶数点”也是一个随机事件，它由“出现 2 点”、“出现 4 点”、“出现 6 点”这三个基本事件所组成（又称它包含三个基本事件），所以是复合事件.

在试验中必然发生的事件称为**必然事件**，用 Ω 来表示；必然不发生的事件称为**不可能事件**，，用 Φ 来表示. 如在试验(2)中“点数不大于 6”是必然事件，“点数大于 6”就是不可能事件. 本来必然事件与不可能事件不具有不确定性，但为了以后讨论方便，将二者也视为随机事件.

下面我们用集合论的观点进行讨论.

三、样本空间

随机试验中所出现的基本事件称为**样本点**,所有基本事件所组成的集合称为该试验的**样本空间**(用 Ω 表示).这样,我们看到,事件就是样本点组成的某个集合,而必然事件 Ω 就是样本点的全体,即样本空间;不可能事件 Φ 则是不含任何样本点的集合,即空集.

例 1.1 写出试验(1)的样本空间.

解 样本空间为

$$\Omega = \{(正,正),(正,反),(反,正),(反,反)\}.$$

下面我们再举一个关于事件的例子.

例 1.2 在试验(1)中,"第一次出现正面的"的事件为$\{$(正,正),(正,反)$\}$,"两次出现同一面"的事件为$\{$(正,正),(反,反)$\}$.

四、事件之间的关系与事件的运算

设试验 E 的样本空间为 $\Omega, A, B, A_k(k=1,2,\cdots)$是试验 E 的事件.

1.事件的包含与相等

若事件 A 发生必然导致事件 B 发生,则称**事件 B 包含事件 A**,记为 $B\supset A$,或 $A\subset B$(见图 1-1).

若 $A\subset B$,且 $B\subset A$,则称**事件 A 与事件 B 相等**,,记为 $A=B$.

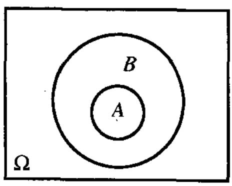

图 1-1

2.事件的和与积

在事件 A 与事件 B 中至少有一个发生,这一事件称为**事件 A 与事件 B 的和**,,记为 $A\cup B$ 或 $A+B$(见图 1-2 阴影部分).

与此类似,若在 n 个事件 $A_1,A_2,\cdots,A_n$ 中至少有一个发生,这一事件称为 $A_1,A_2\cdots,A_n$ 的和,记为 $A_1\cup A_2\cup\cdots\cup A_n$ 或 A_1

$+A_2+\cdots+A_n$,简记为$\bigcup\limits_{k=1}^{n}A_k$或$\sum\limits_{k=1}^{n}A_k$.

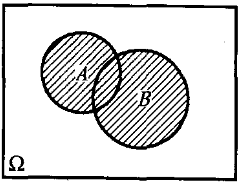

图 1-2

若事件 A 与事件 B 同时发生,这一事件称为**事件 A 与事件 B 的积**,记为 $A\cap B$ 或 AB(见图 1-3 的阴影部分).与此类似,可定义 n 个事件 $A_1,A_2,\cdots,A_n$ 的积为

$$A_1\cap A_2\cap\cdots\cap A_n=\bigcap_{k=1}^{n}A_k$$

或
$$A_1A_2\cdots A_n=\prod_{k=1}^{n}A_k$$

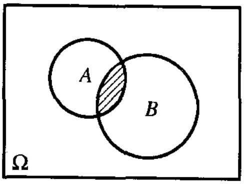

图 1-3

3. 对立事件与事件的差

事件"非 A"(即 A 不发生)称为 **A 的对立事件**,记为 $\overline{A}$.显然有

$$A\overline{A}=\Phi,\quad A+\overline{A}=\Omega$$

若事件 A 发生而事件 B 不发生,这一事件称为**事件 A 与事件 B 的差,记为** $A-B$(见图 1-4).

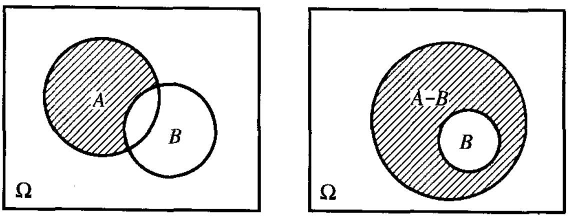

图 1-4

4. 事件的互不相容性

若事件 A 与事件 B 不能同时发生,亦即 $AB=\Phi$,则称**事件 A 与事件 B 是互不相容的**(见图 1-5).

显然 A 与 $\overline{A}$ 是互不相容的.此外,我们也可以定义 n 个事件

的互不相容性. 若 $A_1,A_2,\cdots,A_n$ 两两互不相容,则称 n 个事件 $A_1,A_2,\cdots,A_n$ 是互不相容的.

5. 事件的运算规律

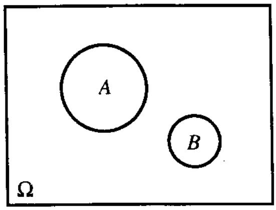

图 1-5

(1) 交换律:$A+B=B+A, AB=BA$.

(2) 结合律:$(A+B)+C=A+(B+C), (AB)C=A(BC)$.

(3) 分配律: $A(B+C)=AB+AC, A+BC=(A+B)(A+C)$.

(4) 重迭律:$A+A=A, AA=A$.

(5) 两次求逆律:$\overline{\overline{A}}=A$.

(6) 互逆律:$A+\overline{A}=\Omega$, $A\overline{A}=\Phi$.

(7) 差化积:$A-B=A\overline{B}$.

(8) 吸收律:若 $A\subset B$,则 $A+B=B, AB=A$;特别地,$A+\Omega=\Omega, A\Omega=A; A+\Phi=A, A\Phi=\Phi$.

(9) 反演律(莫根定理):$\overline{A+B}=\overline{A}\overline{B}, \overline{AB}=\overline{A}+\overline{B}$.

用运算符号把事件联结起来的算式称为事件式,在事件式中,运算的顺序规定如下:第一"逆",第二"积",第三"加"或"差",如有括号,则优先执行括号里的运算.

例 1.3 如图 1-6 所示的电路中,以 A 表示"信号灯亮"这一件事,以 B,C,D 分别表示事件"继电器接点 Ⅰ,Ⅱ,Ⅲ 闭合",则容易知道事件之间有如下关系:

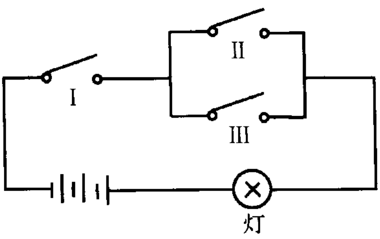

图 1-6

$$BC\subset A, BD\subset A,$$

$BC+BD=A, \overline{B}A=\Phi$(事件 $\overline{B}$

与事件A互不相容).

从以上的讨论中可以看到,概率论中事件之间的关系与运算和集合论中集合之间的关系与运算是一致的.为对比方便,列出下面表格,以便我们把对事件的分析和运算转化为对集合的分析和运算.

记　　号	概　率　论	集　合　论
Ω	样本空间,必然事件	全集
Φ	不可能事件	空集
e	基本事件	元素
A	事件	子集
$\overline{A}$	A的对立事件	A的补集
$A\subset B$	事件A发生导致事件B发生	A是B的子集
$A=B$	事件A与事件B相等	A与B相等
$A\cup B$	事件A与事件B至少有一个发生	A与B的并集
$A\cap B$	事件A与事件B同时发生	A与B的交集
$A-B$	事件A发生而事件B不发生	A与B的差集
$A\cap B=\Phi$	事件A与事件B互不相容	A与B没有相同的元素

习题 1.1

1.对立与互不相容有何异同?试举例说明A,B,C三事件互不相容与$ABC=\Phi$是不是一回事?为什么?

2.设A,B,C是三个事件,试用A,B,C的事件式分别表示下例各事件:(1)A,B,C中至少有一个发生;(2)A,B,C中恰有一个发生;(3)A,B,C中不多于一个发生.

3.在五个数字1,2,3,4,5中任取一数,以$A=\{1,2,3\}$表示取得1,或2,或3,$B=\{2,3,4\}$表示取得2,或3,或4,试问下列各事件式:(1)$\overline{A}B$;(2)$\overline{A}B$;(3)$\overline{\overline{A}\,\overline{B}}$分别表示什么?

4.试问$(A+B)-A=B$是否成立?

§2 频率与概率、古典概型

一、频率

对于随机试验的许多可能的结果,我们最关心的是某些结果出现的可能性有多大.例如,要在某河流上建筑一座防洪大坝,为确定大坝的高度,就要知道河流在造大坝地段每年最大洪水达到某高度的可能性的大小,最大洪水达到某一高度是随机事件.我们希望能将一个随机事件发生的可能性的大小用一个数来表达.为此,我们首先来看下面两个例子:

例 1.4 检查某工厂产品的次品率,其结果如下:

表 2-1

抽出产品件数(n)	5	10	60	150	600	1 200	1 800	2 400
次品数(m)	0	3	7	19	52	109	169	248
次品率($\frac{m}{n}$)	0	0.3	0.117	0.127	0.087	0.091	0.094	0.103

从表 2-1 中的数字可以看出,尽管每次抽到次品数的多少具有偶然性,但随着抽样的大量进行,抽取的产品件数逐渐增多,我们发现次品率在 0.1 附近摆动.

例 1.5 观察某种子的发芽率.从一大批种子中抽取 10 批种子做发芽试验,其结果如下:

表 2-2

种子粒数	2	5	10	70	130	310	700	2 000	3 000
发芽粒数	2	4	9	60	116	282	639	1 806	2 715
发芽率	1	0.8	0.9	0.87	0.892	0.910	0.913	0.903	0.905

从表 2-2 可以看出,发芽率在 0.9 附近摆动.

从以上例子看到，当考虑事件 A 发生的可能性大小时，只要我们在同一条件下做大量的重复试验，事件 A 发生的次数与试验的总次数之比将呈现某种稳定性.在一般情况下，此比值稳定在某一数值附近，而这个数值能反映事件 A 发生的可能性的大小.

一般地，设随机事件 A 在 n 次试验中出现 n_A 次，比值

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

称为事件 A 在这 n 次试验中出现的**频率**.

二、频率的性质

当试验次数 n 固定时，事件 A 发生的频率 $f_n(A)$ 有如下性质：

(1)$0 \leqslant f_n(A) \leqslant 1$.

(2)$f_n(\Omega)=1, f_n(\Phi)=0$.

(3)若 A,B 两事件互不相容，即 $AB=\Phi$，则

$$f_n(A+B) = f_n(A) + f_n(B)$$

证 (1)因为 $0 \leqslant n_A \leqslant n$，所以 $0 \leqslant \frac{n_A}{n} \leqslant 1$，即

$$0 \leqslant f_n(A) \leqslant 1$$

(2)因为 Ω 为必然事件，Φ 为不可能事件，所以 $n_\Omega = n$，$n_\Phi = 0$，故 $f_n(\Omega)=1, f_n(\Phi)=0$.

(3)设在 n 次试验中 A 发生 n_A 次，B 发生 n_B 次，由 $A+B$ 的定义及 A、B 的互不相容性可知：

$$n_{A+B} = n_A + n_B$$

于是

$$f_n(A+B) = \frac{n_A + n_B}{n} = \frac{n_A}{n} + \frac{n_B}{n} = f_n(A) + f_n(B)$$