概率统计

张廷杰 编

西南交通大学出版社

概 率 统 计

张廷杰 编

西南交通大学出版社

1995年·成都

（川）新登字018号

内 容 简 介

本书前四章为概率论基础知识，包括随机事件与概率、随机变量的概率分布与数字特征、二元随机变量简介与极限定理；后六章为数理统计的基本知识和常用方法，内容包括随机抽样与样本、参数估计、假设检验、方差分析、回归分析和正交试验设计。每章末均有习题，书后附有习题参考答案。

本书内容简炼，思路清晰，通俗易懂，联系实际，适合作各类成人高校财经类、管理类专业大专教学用书或参考书，亦可供具有微积分和线性代数基础知识的读者作为自学用书，还可供普通高校财经类、管理类专业师生参考。

概 率 统 计

张廷杰 编

*

西南交通大学出版社出版发行

（成都 九里堤）

机械工业出版社京丰印刷厂印刷

*

开本：787×1092 1/32 印张：7.75

字数：179千字 印数：1-1000册

1995年1月第1版 1995年1月第1次印刷

ISBN7-81022-774-2/O·070

定价：8.00元

前　言

概率统计包括概率论和数理统计两部分内容。概率论是研究和揭示随机现象统计规律性的一门数学学科；数理统计扎根于概率论的基础之上，它是研究如何用有效的方法搜集、整理、分析带有随机影响的数据，从而对所考察的问题做出推断和预测，直至为采取某种决策提出依据和建议。本书在简要介绍概率论基本知识的基础之上，着重介绍数理统计的一些基本概念和方法，以期学员在本课程结束之后，既能为一些后续课程打下较好的基础，又能将其直接应用于经济领域和各种定量分析之中去。

本书是作者在北京市丰台区职工大学、北京电视大学丰台工作站、中国统计干部学院北京分院、北京教育学院丰台分院等成人高等院校的讲稿和讲义改编而成。本书力求把握成人学习特点，突出针对性。以应用为目的，以必须够用为度，深入浅出阐述概念，以大量实例和适量习题强化应用，删除一些繁难的论证和计算。本书带“*”号的内容可根据教学需要和学时安排适当取舍删减，前七章基本部分可用40学时讲完，后三章可根据教学需要选讲。

本书通俗易懂，密切联系实际，便于学员自学。适合作各类成人高等学校财经类、管理类专业大专教学用书和教学参考书，也可供具有微积分和线性代数初步知识的读者作为自学用书，还可供普通高等学校有关专业师生作为教学参考书。

全书共十章，第一至第四章为概率论基础知识，内容包括随机事件与概率、随机变量的概率分布与数字特征、二元随机变量简介与极限定理；第五至第九章为数理统计的基本知识和应用范围较广的方法，内容包括随机抽样与样本、参数估计、假设检验、方差分析与回归分析；第十章正交试验设计在工农业生产实际和科学试验中有着广泛的应用。本书在编写和出版过程中，得到陆秀嫒、杨莲子等同志的支持和帮助，作者深表感谢。

本书书稿虽经多年试用，反复修改，但由于作者水平所限，加以脱稿仓促，错误和不足之处在所难免，敬希读者批评指正。

编　者

1994年9月

目　录

第一章　随机事件与概率

概率论是研究和揭示随机现象统计规律性的一门数学学科。数理统计是在概率论的基础上发展起来的，而又以概率论为其理论基础，它是研究如何用有效的方法搜集、整理、分析带有随机影响的数据，从而对所考察的问题做出推断和预测，直至为采取某种决策提出依据和建议，它是适应人类社会实践的需要而产生和发展起来的。

随机事件与概率是概率论中最基本的两个概念，研究概率的性质和计算是概率论中的基本问题。本章主要讨论概率的定义、性质和计算。

§1.1　随 机 事 件

在自然界和人类社会、政治、经济生活中，经常遇到各种各样现象。其中一类现象称为必然现象或确定现象，另一类现象称为随机现象或不确定现象。

必然现象是在一定条件下一定会出现的现象，这类现象的发生与否一般可以事先确定。这种现象也称为确定现象。

例如，在标准大气压下(约为0.1013MPa)，纯水被加热到100℃时会沸腾。这里，“标准大气压”和“100℃”是条件，而“沸腾”是结果。只要这两个条件具备，那么纯水就一定出现“沸腾”这一结果。

再如，有100件产品全是正品，从其中任取1件必然是正品。这里，“100件产品全是正品”是条件，结果是“任

取 1 件是正品”。只要条件具备，结果必然出现。

上面两个例子中的现象都是必然现象。在一定条件下，必然现象只有一个可能结果。因此，必然现象所出现的结果总是肯定的。在相同的条件下，人们对必然现象出现的结果可以事先预言。

在一定条件下，某种现象的结果不止一个，至于哪一个结果出现事先无法确定和预言。具有这种特点的现象称为随机现象，也称为不确定现象。

例如，任意抛出一枚均匀硬币，当硬币落到地面上时有两种可能结果：正面朝上或反面朝上(通常把有币值的一面称为正面)。究竟哪面朝上，只有在硬币落地之后才见分晓，而在硬币落地之前是无法确定的。

又如，在相同条件下，一门大炮多次射击同一目标，炮弹并不都落于同一点。这说明炮弹射击时，炮弹的落地点有多个可能结果。至于炮弹落于何处，在炮弹射击之前是不能确定的。

再如，在10个同类产品中包含 8 个正品、2 个次品，从这些产品中任意抽取 3 个，那么“3 个都是正品”或“至少 1 个是正品”的结果，在抽取之前也是不能确定的。

上述现象都是随机现象。

在自然界和人类社会、政治、经济活动中，任何现象都是在一定条件下发生的。而随机现象的发生除了基本条件外，还要受到一些不能为人们完全控制的次要因素的影响，从而表现出偶然性来。例如大炮射击时，其炮弹落地点除了受炮本身的性能和发射角等基本条件决定之外，还要受到风力、风向、温度、湿度等次要因素的影响，而这些次要因素无时不在改变。因此，即使基本条件不变，射击时炮弹的落

地点事先也无法精确确定。

对随机现象进行一次或少数几次观察或试验，其结果是不能确定的。如果在相同条件下，通过大量观察或试验，每种随机现象都会出现某种规律性。例如，对于一粒或几粒种子，事先不能断定它们能否发芽。但在同一条件下，一批种子的发芽率却是稳定的。

在概率论中，把对随机现象的观察或试验叫做随机试验。在一定条件下，进行随机试验所得到的每一个可能结果称为随机事件，简称为事件，通常用符号A、B、C等表示。在随机试验中，不能再分解的最简单的结果，称为基本事件。由若干个基本事件组合而成的事件，称为复合事件。在一定条件下，每次试验中一定发生的事件称为必然事件，用符号U表示。每次试验中一定不发生的事件称为不可能事件，用符号V表示。

例如，在标有号码1、2、3的三件商品中，任意抽取1件，则“抽中1号”、“抽中2号”、“抽中3号”都是基本事件。“抽中奇数号”也是一个随机事件，但它不是基本事件，它是由“抽中1号”和“抽中3号”这两个基本事件组成的复合事件，只要这两个基本事件中的一个发生，“抽中奇数号”这个事件就发生。“抽中号数不大于3”是必然事件，“抽中号数大于3”是不可能事件。

再如，投掷一颗均匀骰子的试验中，观察其出现的点数，“1点”、“2点”、…、“6点”都是基本事件。“奇数点”也是随机事件，但它不是基本事件，它是由“1点”、“3点”、“5点”这三个基本事件组成的复合事件，只要这三个基本事件中的一个发生，“奇数点”这个事件就发生。“点数小于7”是必然事件，“点数不小于7”

是不可能事件。

应该指出：不论是必然事件、不可能事件，还是随机事件，都是相对于一定试验条件而言的，试验条件发生变化之后，事件的性质也可能发生变化。例如，“点数小于 7”这个事件在投掷一颗骰子的试验中是必然事件，在投掷两颗骰子的试验中就是一个随机事件，而在投掷 8 颗或 8 颗以上的骰子时，它就成为不可能事件了。为了讨论方便，一般将必然事件和不可能事件也当作随机事件，而作为它的两个极端情况。

对于随机试验的每一个基本事件，用包含一个元素 ω 的单元素集合 $\{\omega\}$ 表示；由若干个基本事件复合而成的事件，用包含若干个相应元素的集合 $\{\omega_1, \omega_2, \cdots, \omega_n\}$ 来表示；由所有基本事件对应的全部元素组成的集合，称为样本空间。由于任何一次试验的结果必然出现全部基本事件之一，这样，样本空间作为一个复合事件是必然事件，仍用符号 U 表示。样本空间中的每一个元素称为样本点。因此，可以把随机事件定义为样本点的某个集合。如果把样本空间看作全集，则每一个随机事件就相应于它的一个子集。不可能事件 V 就是空集，必然事件 U 就是全集。因此，可以把事件的关系和运算与集合的关系和运算统一起来。

为了直观，经常用图形表示事件。一般用平面上某一正方形(或矩形)表示必然事件，该区域内的一个子区域表示事件。

样本空间 U 中的事件之间有下述关系和运算：

1. 事件的包含

如果事件 A 发生必然导致事件 B 发生，即事件 A 中的每一个样本点都包含在事件 B 中，则称为事件 B 包含事件 A，或

称为事件 A 包含于事件 B。记作

$$B \supset A \text{ 或 } A \subset B$$

显然，对任一事件 A，有

$$V \subset A \subset U$$

如果事件 A 包含事件 B，事件 B 也包含事件 A，则称事件 A 与事件 B 相等。即事件 A 与事件 B 所含的样本点完全相同。记作

$$A = B$$

2．事件的和(或并)

两个事件 A、B 中至少有一个发生的事件是一个事件“A 或 B”，称为事件 A 与 B 的和(或并)。它是由事件 A 与事件 B 所有样本点构成的集合。记作

$$A + B \text{ 或 } A \cup B$$

3．事件的积(或交)

两个事件 A、B 同时发生的事件是一个事件“A 且 B”，称为事件 A 与 B 的积(或交)。它是由事件 A 与事件 B 的所有公共样本点构成的集合。记作

$$AB \text{ 或 } A \cap B$$

4．事件的差

事件 A 发生而事件 B 不发生，是一个事件，称为事件 A 与事件 B 的差。它是由属于 A 但不属于 B 的那些样本点构成的集合。记作

$$A - B$$

5．互不相容事件(或互斥事件)

如果事件 A 与事件 B 不能同时发生，即 $AB = V$，则称事件 A 与事件 B 为互不相容事件(或互斥事件)。互不相容事件 A 与 B 没有公共的样本点。显然，基本事件之间是互不相

容的。

6．对立事件

如果事件A与事件B中必然有一个发生且仅有一个发生，即事件A与事件B满足条件

$$A+B=U, \quad AB=V$$

则称事件A与事件B为相互对立的事件(或互逆事件)。又称A是B的对立事件(或B是A的对立事件)。记作

$$A=\bar{B}(\text{或 } B=\bar{A})$$

7．完备事件组

如果事件A_1、A_2、…、A_n两两互不相容，且$A_1+A_2+\cdots+A_n=U$，则称A_1、A_2、…、A_n构成一个完备事件组。

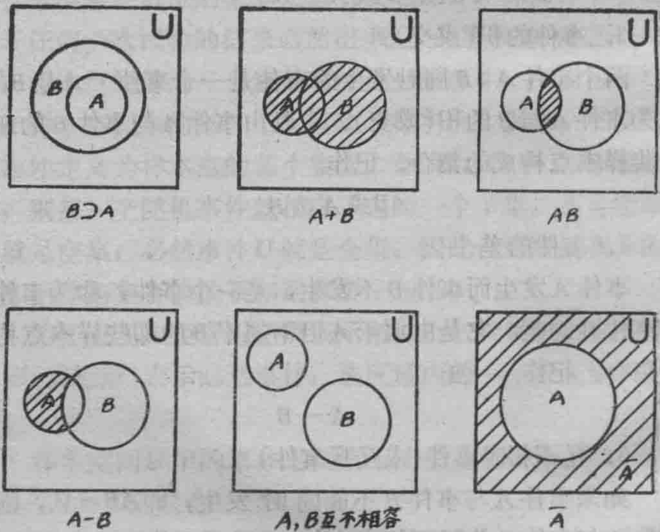

图 1.1

用图形表示事件间的关系和运算见图1.1。

利用集合的运算规则可得到事件间的下述定律：

1．交换律

$AB=BA$； $A+B=B+A$

2．结合律

$(AB)C=A(BC)$； $(A+B)+C=A+(B+C)$

3．分配律

$(A+B)C=AC+BC$； $(AB)+C=(A+C)(B+C)$

4．对偶律

$$\overline{A+B}=\bar{A}\bar{B}；\quad \overline{AB}=\bar{A}+\bar{B}$$

例1.1 掷一颗骰子的试验，观察出现的点数：事件A表示"奇数点"；事件B表示"点数小于5"；事件C表示"点数小于5的偶数点"。用集合的列举法表示下列事件：U，A，B，C，$A+B$，$A-B$，AB，AC，$C-A$，$\bar{A}+B$。

解 $U=\{1, 2, 3, 4, 5, 6\}$

$A=\{1, 3, 5\}$

$B=\{1, 2, 3, 4\}$

$C=\{2, 4\}$

$A+B=\{1, 2, 3, 4, 5\}$

$A-B=\{5\}$

$AB=\{1, 3\}$

$AC=V$

$C-A=\{2, 4\}$

$\bar{A}+B=\{1, 2, 3, 4, 6\}$

例1.2 设事件A_k表示k次取到了合格品($k=1, 2, 3$)，试用事件间的运算表示下列事件：3次都取到了合格

品；3次中至少有1次取到合格品；3次中恰有2次取到合格品；3次中至多有1次取到合格品。

解 3次都取到合格品：$A_1A_2A_3$

3次中至少有1次取到合格品：

$$A_1+A_2+A_3$$

3次中恰有2次取到合格品：

$$A_1A_2\bar{A}_3+A_1\bar{A}_2A_3+\bar{A}_1A_2A_3$$

3次中至多有1次取到合格品：

$$\bar{A}_1\bar{A}_2+\bar{A}_1\bar{A}_3+\bar{A}_2\bar{A}_3$$

§1.2 随机事件的概率

（一）概率的定义

设事件A在n次重复试验中发生了m次，则$\frac{m}{n}$称为事件A发生的频率。记作$f(A)$，即

$$f(A)=\frac{m}{n}$$

其中m称为事件A发生的频数。事件A发生的频率满足不等式

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

如果A是必然事件，有$m=n$，则$f(A)=1$；如果A是不可能事件，有$m=0$，则$f(A)=0$。故必然事件的频率为1，不可能事件的频率为0。

当重复试验的次数n很大时，某事件A发生的频率的数值总会在某个确定的数值附近摆动。并且，当重复试验的次数n越大时，事件A发生的频率就越接近这个数值。历史上有不少人做过多次投掷硬币的试验，设A表示“正面朝上”的事件，表1.1记录了几个人的试验结果。

表1.1

试验者	投掷次数 n	A出现次数 m	A出现频率 $f(A)$
德摩根	2048	1061	0.518
蒲丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
尔皮逊	24000	12012	0.5005

由表1.1可以看出，投掷硬币次数越多时，频率越接近0.5，且逐渐稳定于0.5。这样，0.5这个数反映了事件A出现的可能性的大小。这种特性就是随机事件发生频率的稳定性。

定义1.1 在不变的条件下，重复进行n次试验，事件A发生的频率$f(A)$稳定地在某一常数p附近摆动($0 \leqslant p \leqslant 1$)，且一般说来，$n$越大，摆动幅度越小，则称常数$p$为事件$A$的概率。记作$P(A)$，即

$$P(A)=p$$

这就是概率的统计定义，它描述了在一次试验中事件A发生的可能性的大小。对于每一个随机事件A，总有一个数$P(A)$与之对应。它是客观存在的，而且随着试验次数n不断增加时，频率$f(A)$就逐渐稳定于概率$P(A)$。例如，投掷一枚硬币，事件"正面朝上"的概率为$\frac{1}{2}$。投掷一枚骰子，事件"出现1点"的概率为$\frac{1}{6}$；"出现2点"的概率为$\frac{1}{6}$；"出现奇数点"的概率为$\frac{1}{2}$。

事件的频率与概率是度量事件出现可能性大小的两个统计特征。频率是个试验值，具有随机性；可能取多个不同的值。因此，它只能近似地反映事件出现可能性的大小。概率