概率统计基础知识

张尚志　俞宗源　编

概率统计基础知识

张尚志　俞宗源　编

贵州人民出版社

概率统计基础知识

张尚志　俞宗源　编

贵州人民出版社出版

（贵阳市延安中路5号）

贵州新华印刷厂印刷　贵州省新华书店发行

787×1092毫米　32开本　6 印张　121千字

印数 1—7900

1982年5月第1版　1982年5月第1次印刷

书号：15115·140　定价：0.63元

引　言

概率论与数理统计（简称概率统计）是专门研究随机现象的数量规律的科学，至今已渗透到各个学科领域，工业、农业、国防和科研等方面都在广泛应用。

在自然界和人类社会中，广泛地存在着两类不同的现象：确定性现象和随机现象。

所谓确定性现象，就是在一定条件下一定会出现某一种肯定的结果的现象。例如，一个做直线运动的物体，若速度为 v，运动时间为 t，则它经过的路程一定是 vt。又如在常压下，将水加热到100℃就会沸腾。这些都是确定性现象。初等数学、微积分等数学分支都是研究确定性现象的数量规律的。

所谓随机现象，就是在一定条件下，具有多种可能的结果，事先并不能肯定究竟会发生什么结果的现象。因此，在相同条件下，做若干次试验，它们的结果可能是不一样的。例如，即使是一条稳定的生产线上生产出来的维尼纶，其纤维粗细也是有差异的。又如在测量一个物体的长度时，即使我们用同一仪器，重复测量多次，结果也常会有微小的差异，这就是平时我们所说的测量误差造成的。测量误差就是一种随机误差。随机现象，由于人们事先不能断定它将发生什么

样的结果，表面上看来好象是不可捉摸和无规律可循似的。其实不然，实践证明，研究大量的随机现象后，通常总会显露出一定的规律性来。例如，生男还是生女，就个别来说是不确定的，但通过大量的人口统计，男婴与女婴的比例却历年基本不变，都接近二分之一。又如多次测量同一物体，其结果虽略有差异，但当测量次数增到很大时，就会发现测量数据集中在某个值的附近，它们相当对称地分布在这个值的两侧，而且离这个值越远越少。由此可见，虽然个别现象是无规律的，但是大量重复观测的结果却是有规律的，这种规律性叫做统计规律性。因此，随机现象都具有偶然和必然的两重性。概率论与数理统计就是研究随机现象的数量规律的一个数学分支。

由于随机现象广泛存在，所以从各个不同的领域提出了大量的问题，促使概率统计蓬勃发展，也使它在生产和科研的各个领域获得了广泛而富有成效的应用。早先曾流行一种传说，认为概率论是从赌博发展起来的，这当然是一种误解，产生这种误解的原因是因为远在三百多年前，在概率论发展的初期，有些赌徒曾经向数学家提出过一些问题，数学家给他们解决了。例如，甲乙两人相约赌五局，先赢三局者为胜，可得全部赌本，但当甲赢了二局乙赢了一局时，由于某种原因，他们不再赌下去了，问这时赌本该如何分法？后来象分扑克牌、掷骰子等问题，由于它们对阐述古典概型的概念与性质都比较容易为读者所理解，因此不少作者乐于采用这些例子，并不是说这些赌博和游戏是概率论发展的源泉。促使概率论发展成为一门科学，还是因为生产和科学研究的需要。数理统计则是在概率论基础上发展起来的一门应用科学。时

至今日，概率统计已发展成有许多分支在多方面获得应用的学科，在一些工业化先进的国家里，概率统计已十分普及，甚至从事社会科学的人和普通工人也懂得不少概率统计知识。在我国，虽然早在三十年代，已故著名数学家许宝禄教授就已在统计方面作出举世公认的杰出贡献，但在解放前从事概率统计的人寥寥无几，直到一九五六年以后才在大学里开设概率论专门化。此后由于建设事业的发展，这方面的队伍也迅速扩大。尤其是在应用方面，也随之逐渐开展工作。例如产品的抽样检查，试验设计（包括正交设计），全面质量管理，水文、气象、地震、农作物的病虫害等各类预报，工业自动化中数学模型的建立，通讯以及诸如社会资料、医学资料、地质资料的统计分析等方面，都已大量应用概率统计。因此在新编的中学数学课本中也编入了概率统计的若干基本知识。本书就是为初学者提供的一本概率统计的自学读物，在不失科学性和不用高等数学的前提下，对概率统计最基本的内容作一介绍，举了较多的难易程度不同的例子（其中有些在概率论的发展上曾经起过一定的作用）。对于较复杂的例子如果初看时有困难，也可以暂时不看，其他打*号的小节也是如此。

学习概率统计不要指望记住几个公式代代套套就能解决问题，一定要理解深透。此外，概率统计有它独特的思考问题的方法，它对培养我们正确的思想方法也是很有好处的。

本书初稿曾经华东师大茆诗松副教授阅过，承提出宝贵意见，谨致谢意！

限于编者水平，书中定有不少缺点以致错误，恳请同志们批评指正。

责任编辑　何伊德
封面设计　蒋道环
技术设计　荀新馨

书号：15115·140
定价：　0.63　元

目　录

第一章
事件与概率

§1 随机试验与随机事件

一个试验如果事先不能准确地预言它的结果，而且在相同条件下可重复进行，则称此种试验为**随机试验**。

在一定条件下，可能发生也可能不发生的事件称为**随机事件**。

〔例1〕“丢一枚硬币，结果国徽朝上”。这里“丢一枚硬币”是一个随机试验，它的结果可能是国徽朝上，也可能是字面朝上。因此，“丢一枚硬币，结果国徽朝上”是一个随机事件。在条件已经明确，不会引起误解时我们常可把事件写得简单些，如上面的事件就可写成“国徽朝上”。

〔例2〕“某电话总机在上午九时至十时的一小时内接到至少15次呼唤”是一个随机事件，因为在这一小时内这事件可能出现，也可能不出现（当接到的呼唤少于15次时，就称上述事件不出现，或称不发生）。

〔例3〕检查某厂的一批产品时，我们随机地抽取100件，“其中恰有5件次品”，也是一个随机事件，因为我们抽取的100件产品中不一定正好有5件次品。

〔例4〕“贵阳八月份降水量在400～430毫米之间”是

一个随机事件。因为具体观测贵阳某年八月的降水量（相当于一次随机试验）时，可能在400～430毫米之间，也可能不足400毫米或多于430毫米。

类似这样的例子读者自己还可以举出很多。

随机事件是在一定条件下可能发生也可能不发生的事件。它的两个极端就是在一定条件下一定发生的事件——我们称为**必然事件**；以及一定不发生的事件——我们称为**不可能事件**。

例如“某电话总机一小时内接到的呼唤次数大于等于零”，“扔一个骰子得的点子在1～6之间”等是必然事件；而“某电话总机一小时内接到－1次呼唤”，“扔一个骰子得7点”等便是不可能事件。

今后为叙述简便起见，我们将随机事件、必然事件和不可能事件都简称为**事件**，而且在不致引起误解的情况下，也不一定每次都把“一定的条件”明确地写出来。但是必须注意，我们总是在一定条件下讨论随机事件的，否则随机事件的规律性也就无从谈起。

今后，我们用大写的英文字母A、B、C、…或A_1、A_2、…等符号来表示事件，而用Ω表示必然事件，用ϕ表示不可能事件。

§2　基本事件与复合事件

当我们仔细考察事件时，会发现有的事件能分解成一些更基本的事件，有些则不能再分了，我们把不能再分解的（或不必再分解的）事件称为**基本事件**，而把能分解成若干个基本

事件的事件称为**复合事件**。

〔例5〕 我们来考虑三个球a、b、c，将它们随机地放入三个盒中，则总共有27种可能的放法。现将它们列成表1，$\{\ |\ |\ \}$自左至右表示第1、2、3个盒子，如$\{ab|-|c\}$就表示a、b放在第1个盒子内、c放在第3个盒子内，这样表1中的每一种放法都是一个基本事件，因为它们已不能再分解成更基本的事件了。而象事件：

A：“每个盒内放一个球”，

B：“第一个盒内恰放一个球”，

就不是基本事件了，因为A是由22～27这六个基本事件组成的，而B是由10～15以及22～27这12个基本事件组成的。所以，A、B都是复合事件。

表1

1	$\{abc\|-\|-\}$	10	$\{a\|bc\|-\}$	19	$\{-\|a\|bc\}$
2	$\{-\|abc\|-\}$	11	$\{b\|ac\|-\}$	20	$\{-\|b\|ac\}$
3	$\{-\|-\|abc\}$	12	$\{c\|ab\|-\}$	21	$\{-\|c\|ab\}$
4	$\{ab\|c\|-\}$	13	$\{a\|-\|bc\}$	22	$\{a\|b\|c\}$
5	$\{ac\|b\|-\}$	14	$\{b\|-\|ac\}$	23	$\{a\|c\|b\}$
6	$\{bc\|a\|-\}$	15	$\{c\|-\|ab\}$	24	$\{b\|a\|c\}$
7	$\{ab\|-\|c\}$	16	$\{-\|ab\|c\}$	25	$\{b\|c\|a\}$
8	$\{ac\|-\|b\}$	17	$\{-\|ac\|b\}$	26	$\{c\|a\|b\}$
9	$\{bc\|-\|a\}$	18	$\{-\|bc\|a\}$	27	$\{c\|b\|a\}$

又如“抽检100件产品，次品不多于2件”这个事件是由“抽检100件产品，没有次品”，“抽检100件产品，恰有1件次品”，“抽检100件产品，恰有2件次品”这三个基本事件

组成的。所以，它也是一个复合事件。

善于把一个复杂的事件分解为一些较简单的事件是很有用的。下面，我们就来介绍事件之间的关系和运算。

§3 事件的运算

我们已经知道了事件都是由基本事件组成的。我们前述的“事件A出现”，是指组成A的基本事件中某一个出现，并不是指A中每一个基本事件都要同时出现才说A出现。如〔例5〕中若放球的结果是表1中22～27号结果之一时，我们便说事件A（“在每个盒内放一个球”）出现了。

下面我们来定义几种事件间的关系和运算：

1.**包含关系** 果如事件A出现必导致B出现，就说**B包含A**，记作$A\subset B$。如〔例5〕中，当A出现时，是指每个盒中恰放了一个球，那当然第一个盒中恰放了一个球，所以B必出现，即$A\subset B$。

2.**相等关系** 事件A、B，如有$A\subset B$又有$B\subset A$，则称**A与B相等**，记作$A=B$。

3.**事件的和** “两事件A、B中至少有一个出现”也是一个事件，这个事件称为**事件A与B的和**，记为$A+B$或$A\cup B$。例如我们有4张考签，今任取一张，令A是“抽到偶数号考签”，B是“抽到2号考签”，C是“抽到4号考签”，则$A=B+C$。

类似的对任意k个事件A_1，A_2，…，A_k，可定义A_1，A_2，…，A_k的和，那就是“A_1，A_2，…，A_k中至少一个出现”，记为$A_1+A_2+\cdots+A_k$或$A_1\cup A_2\cup\cdots\cup A_k$。也可用求

和号$\sum_{i=1}^{k}A_i$记。

4.**事件的交** "事件A、B同时出现"这也是一个事件，称为$\boldsymbol{A}$**与**$\boldsymbol{B}$**的交**，记作$A\cap B$或AB。例如15张考签，任取一张，令A为事件"抽到偶数号考签"，B为事件"抽到的考签号数是3的倍数"，则$A\cap B$就是事件"抽到的考签号既是偶数又是3的倍数"，即$A\cap B=$"抽到6号或12号考签"。类似的可以定义k个事件A_1，A_2，$\cdots A_k$的交，记为$A_1\cap A_2\cdots\cap A_k$或$A_1A_2\cdots A_k$。

5.**事件的差** "A出现而B不出现"也是一个事件，称为$\boldsymbol{A}$**与**$\boldsymbol{B}$**的差**，记作$A-B$，例如上面的例中$A-B$就是"抽到的考签号是偶数但不是3的倍数"。读者试想，如令C为"抽到的考签号数小于10"，则$A-C$是什么？$A\cap C$呢？

6.**互斥关系** 如果事件A、B不可能同时出现，我们就说事件$\boldsymbol{A}$**与**$\boldsymbol{B}$**是互斥的或不相容的**，用我们已经学过的交来写就是$A\cap B=\phi$（不可能事件）。例如若A是"抽到号数小于5的考签"，B是"抽到号数大于10的考签"，则A与B是互斥的，因为如果抽到的考签号数是小于5的，它就不可能大于10，反之亦然。

7.**事件的逆事件** 对于事件A、B如果有

1°$A\cap B=\phi$、即A、B不可能同时出现；

2°$A\cup B=\Omega$、即A、B一定有一个要出现。

则称A、B为**互逆事件**（亦称**对立事件**），这时事件B就称为事件A的**逆事件**（同样A也称为B的逆事件）。一个事件A的逆事件常用$\bar{A}$表示。当我们把A看作一个集合

时，$\overline{A}$就是A的补集。

注意A、B互逆，则一定互斥，但A、B互斥不一定互逆，如果A、B互斥，则必有$B\subset\overline{A}$。

例如：有1～10号考签，如A=“抽到偶数号考签”，B=“抽到奇数号考签”，C=“抽到1号或3号考签”，则B是A的逆事件，但C只与A互斥而不是A的逆事件。

上面我们定义了事件之间的种种关系和运算，实际上我们把基本事件看成元素，事件就是以基本事件为元素的集合。因此事件的关系和运算与集合的关系和运算是一致的。例如事件A发生，必然导致事件B发生（$A\subset B$），相当于集合论中A是B的部分集合（$A\subset B$）。又如事件A、B同时发生（$A\cap B$），相当于集合论中A与B之交等。所以我们就可以用集合的图形〔这种图形叫韦恩（Venn）图，见图1〕来表示事件之间的关系。

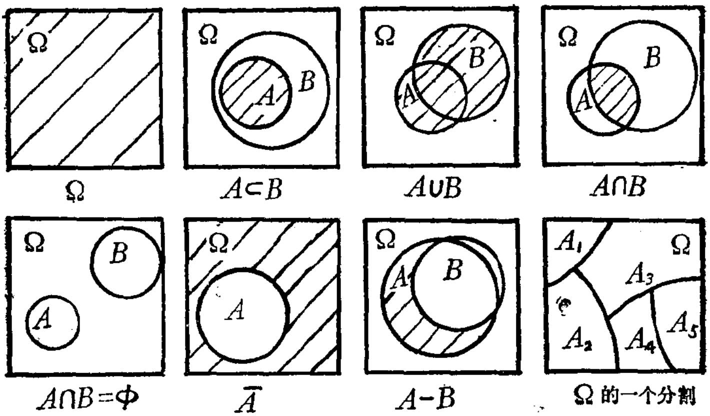

图1　事件关系的韦恩图

在概率论中，通常把基本事件称为**样本点**。事件便是样本点组成的集合，不可能事件ϕ便是空集，必然事件Ω便是全体样本点组成的集合，我们也称Ω为**样本空间**。如果A_1，A_2，…，A_k两两互斥，且$\sum_{i=1}^{k} A_i = \Omega$，则称$A_1$，$A_2$，…，$A_k$是$\Omega$**的一个分割**。

§4 概　　率

随机事件在一次试验中，可能发生也可能不发生，似乎捉摸不定，无规律可循。但是，从大量的试验或观测中，我们就会发现事件的发生还是会呈现出规律性的。这种规律性告诉我们事件出现可能性的大小，刻划事件出现可能性的大小的量便是概率。

一、概率的古典定义

〔例6〕 有50张编好号的考签，一学生从中任抽一张，问结果是“抽到5号考签”的可能性是多大?

这个问题大家都会很快地回答是1/50。那是为什么呢?因为每张考签被抽到的可能性是相同的，此时“抽到1号考签”，“抽到2号考签”，…，“抽到50号考签”便是50个基本事件，我们把“抽到5号考签”记为A。抽一张考签有50种可能结果，而且每种结果都是机会均等的，所以事件A出现的可能性为$1/50=0.02$。

由此可得到**概率的古典定义**

定义 一个随机试验，若