概率统计入门

梁之舜　邓永录

广东科技出版社

概率统计入门

梁之舜　邓永录

广 东 科 技 出 版 社

概率统计入门

梁之舜　邓永录

*

广东科技出版社出版

广东省新华书店发行

广东韶关新华印刷厂印刷

787×1092毫米　32开本　6·75印张　120,000字

1983年11月第1版　1983年11月第1次印刷

印数1－13,000册

统一书号7182·35　定价0.72元

前　　言

为了适应广大数学爱好者、科技工作者、中学教师迫切希望学点概率统计知识的要求，我们编写了这本小册子，作为概率统计的入门读物．

我们编写本书时，力求使它通俗易懂，具有一定的实用性和趣味性，不仅注意阐明概念和结果的直观背景，而且给出涉及多方面应用的例子，另一方面，为了使读者正确地理解、掌握概率统计的基本概念和方法，并且有利于进一步的学习，我们还注意在读者可能接受的水平上适当保持本学科理论的严谨性和系统性，同时初步熟悉一些与概率统计有关的较新内容(如蒙特卡洛方法、可靠性的数学理论、排队论和随机点过程理论中的一些最简单的概念和方法)。

本书读者最好具备初等微积分的知识．但是，具有中学程度的读者，只须直接承认和利用书中用到微积分推证的结果，暂时搁置其推证过程，也可以学习并掌握本书的基本内容．书中有*号的段落，读者在第一次阅读时可略去，这样不会影响对本书基本内容的掌握．

本书初稿曾用作广东省数学学会概率统计讲座的讲义．这次出版时，杨维权同志仔细校阅原稿，提出了一些修改意见，并增写了附录二，何远江同志仔细地校阅了清样。谨此致谢。

目　录

引　言

人们在观察和研究客观世界各种现象和事物间的因果关系时，发现有两类不同的现象和规律，即必然性现象与偶然性现象(或称随机现象)，以及与之相联系的必然性规律与偶然性规律(或称统计规律性)．所谓必然性现象就是在一定的条件下，必然会发生某一种结果的现象．例如，在一个大气压下水在100℃时必然沸腾，在0℃时必然会结冰；从高处让一物体自由下落时，它肯定会落到地面，而且物体在空中经过的路程可由公式$S=\frac{1}{2}gt^2$准确地算出(在没有空气阻力的假定下)．而偶然性现象则是在一定的条件下，具有多种可能的结果，但事先不能断定究竟发生那一种结果的现象．例如，我们随意地抛掷一个均匀的硬币，结果可能出现正面或反面；新生的婴儿可能是男孩，也可能是女孩；灰色的兔与棕色的兔交配，产生的后代可能是灰色、黑色、肉桂色和棕色；对某零件用同一计量工具测量其长度或重量时，由于种种偶然因素的影响(如测量仪器受气候的影响、测量者生理上或心理上的变化等)不可避免地会产生随机的误差，因此，重复进行测量时得到的结果一般是不相同的，有些偏大，有些偏小；在密封的容器里盛着一定体积的气体，其中气体分子由于其它分子的冲击，而引起运动速度和方向的随机变化等等．所有这些现象都是我们事先不能断定其结果的．

必然性现象服从必然性规律的支配，这是不难被人们所理解的。这种规律在数量上常常可以通过数学公式或方程式来表示（例如前面提到的自由落体公式$S=\frac{1}{2}gt^2$）。对于随机现象，由于人们事先不能断定它将发生什么样的结果，从表面上看好象是不可捉摸、纯粹是偶然性起支配作用而谈不上有什么规律性的。其实不然，革命导师恩格斯说："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。"（《路德维希·费尔巴哈与德国古典哲学的终结》）。实践证明，如果对同类的随机现象进行大量重复观察，人们可以从中发现一种几乎确定的规律性，也就是所谓偶然性规律或统计规律性。这是随机现象所特有的一种规律性。就以上面举出的例子来说，多次随意抛掷一枚均匀的硬币，出现"正面"和"反面"的次数约略相等，在一段时间内一个地区出生的男婴和女婴的比例近似是1:1；灰色的兔与棕色的兔交配产生的后代中，灰色、黑色、肉桂色、棕色的比例大概是9:3:3:1。再拿测量零件的长度或重量为例，在消除了错误与系统误差之后，由于还存在着随机的测量误差，因此，通过几次观测而得到的几个观测值不可能完全一样。当观测次数n不大时，观测值的分布杂乱无章，看不出有什么规律性，但当观测次数增加时，就开始呈现一定的规律性，n越大，规律性越显著。一般说来，观测值服从正态分布，即它们在平面坐标图上的分布呈对称的钟形，最高点（对应于最多的观测值）在该零件的长度或重量的真值附近。我们又知道，气体对容器壁的压力等于单位时间内撞击在容器壁单位面积上的气体分子的总影响。从分子物理学的观点来看，气体是由许许多

多的分子组成的，这些分子在不断地运动着，而且在运动过程中彼此影响着．因此，就单个分子而言，其运动的速度、方向是随机的，没有什么规律性．但从宏观来看，对于一定温度和体积的气体，它对容器壁的压力可以认为是稳定的（即等于一常数值）．这是因为气体分子数足够大，使得各单个分子的运动所具有的随机性在集体作用下相互抵消了．这一点在后面讨论大数定律时还要进一步阐明．

综上所述，我们知道个别随机现象虽然是无规律的，但大量性质相同的随机现象，却是有统计规律性的．概率论和数理统计就是研究随机现象的数量规律的科学，它从数量方面研究大量同类现象所特有的一种规律——统计规律性，其目的在于根据实际的观测资料和某些已知规律性去寻找和研究与之有关的一些未知规律性，并据此作出某些人们所希望得到的科学判断和预测．

概率统计这一数学分支，是从十七世纪中叶开始形成的，它的产生和发展紧密地和整个社会的生产与科学的发展相联系着．当时西方处于资本主义社会初期，关闭的封建社会经济逐渐被航海商业经济所突破，在当时科学技术还不发达的条件下，航海的风险是相当大的．因此在航海事业及伴随着它而发展起来的社会保险事业中，人们迫切希望估计出现各种不幸事故和自然灾害的可能性．同时，由于社会发展而出现的大量征兵、征税中提出来的人口统计和其它统计问题也需要利用一些概率统计方法．此外，天文学、物理学以及其它科学的发展，还要求对测量误差的规律及估计方法进行研究．这些都是促使概率统计诞生并为其进一步发展开辟道路的主要源泉。有些人仅仅看到在概率统计产生的初期，个别数学家在讨论某些概率问题时涉及到一些有关赌博的背景，

就认为概率统计起源于赌博，这是不十分确切的。

近年来随着科学技术的迅速发展，人们对客观世界的认识不断深化，需要考虑各式各样的随机因素的问题愈来愈多，因而概率统计的重要性日益为更多的人所认识。例如，过去人们常常是利用一般的微分方程去描述炮弹飞行的轨道，但是现在想要精确地确定和控制导弹、人造地球卫星和宇宙飞船的轨道时，仅仅利用一般的微分方程是远远不够的，这时必须用到随机微分方程和其他概率统计方法。迄今，概率统计除了在农业、工业、交通运输、测量学、地质学、地理学、天文学、气象学、物理学、化学、电子技术、通信技术、自动化科学、生物学、医学、经济学、军事科学，以及各尖端技术中获得越来越广泛的应用之外，它还日益渗透到科学技术的各个领域中，产生了许多新的分支和边缘学科。因此，把概率统计的基本概念和方法在我国逐步普及，让广大科学技术工作者和人民群众了解它、掌握它，这对实现四个现代化和提高人民群众的科学文化水平都是很有必要的。

最后要指出，概率统计作为整个数学学科的一部分，它也是从数量概念出发，对我们所关心的现象的一个数量侧面加以研究并找出其中的统计规律性，因此，它只是各门具体科学的辅助工具。

概率统计不能代替其它科学去研究各种自然现象和社会现象的本质。只有具体分析被研究现象的特点和在反映这些现象本质的科学规律的基础上，才能正确地运用概率统计这个数学工具。同时，又因为概率统计是研究大量同类现象的统计规律性的，我们不可能用它来对个别现象作出任何绝对肯定或否定的预测，因此，也不能用它来代替其它数学分支

(如微分方程、代数学、几何学等)去研究个别现象所服从的必然性数量规律。总之，我们对概率统计的重要性应有充分的认识，但又不能片面地强调和夸大，这才是我们学习和研究这门学科所应持的正确态度。

第一章　随机事件和概率

§1-1　随机事件的概念及其运算

概率论中经常要用到的一个基本概念是“随机事件”，它与前面提到的“随机现象”是紧密相联的。所谓随机现象，就是在一定的条件下有多种可能结果，但人们事先不能断定出现那一种结果的现象。这种随机现象的结果以及这些结果的集合（较严格地说，是这些结果中满足某些条件的集合）就称做随机事件（以后常常简称做“事件”）。例如，前面提到的随机地抛掷一枚均匀的硬币，其可能结果“出现正面”和“出现反面”是随机事件，在新生婴儿的例子中，“婴儿是男孩”和“婴儿是女孩”是随机事件；某射手用步枪对距离为100米的靶子进行射击，可以看作是随机现象，射击的结果“中靶”和“脱靶”是随机事件；某电话总机在早上8点到9点接到的呼唤次数是一随机现象，在该时间区间内“没有呼唤”、“有五次呼唤”、“有不多于两次呼唤”……都是随机事件；对某工厂生产的产品抽出10件作质量检查是随机现象，检查的结果其中“没有次品”、“有两件次品”、“有多于两件次品”、……、等都是随机事件。

随机事件具有这样的特点，即它们在一定的条件下所作的大量重复试验中，就其各别来说，出现与否是具有偶然性的，但就其总体来说，则是蕴含近似的必然性即统计规律性

的．在概率统计中研究随机事件，就是撇开它的具体特性而在数量上考察其统计规律性．

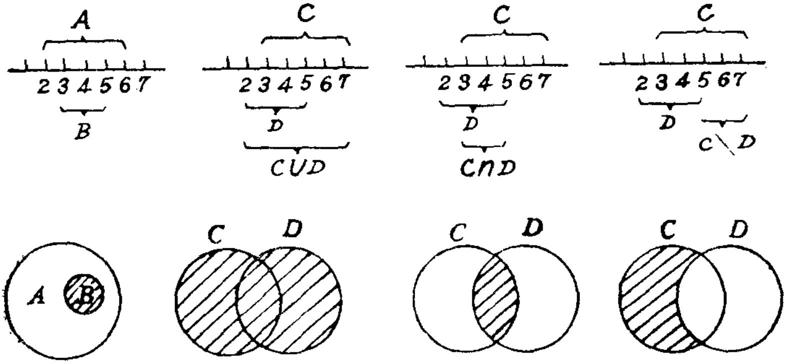
作为随机事件的两种极端情形，我们还要引入必然事件和不可能事件这两个概念．前者是在一定的条件下进行大量重复试验时每次都必然发生的事件，我们用Ω来表示它．后者则恰恰相反，即在一定的条件下进行大量重复试验时每次都不会发生的事件，我们用ϕ来表示它．例如，在一个大气压下“水加热到100℃时沸腾”就是必然事件，而在同样条件下“水加热到60℃时沸腾”则是不可能事件．易见这两个概念是相互联系着的，因为在一定的条件下，必然事件的反面是不可能事件，反之亦然．

事件可以是数量性质的，即试验结果可以直接由测量计数而得．如电话总机接到的呼唤次数；抽查某工厂产品的次品数等等．事件也可以是属性性质的，如抛掷硬币出现正面或反面；新生婴儿是男孩或女孩；射击命中与否等等．当然，这两类事件可用某种关系使其相互转化．例如，抛掷硬币出现正面记为1，反面记为0；射击结果按一定规则计算环数时，就可将它们转化为数量性质的事件．反之，如在抽查某工厂的10件产品时，若规定次品数不多于两件为合格，否则为不合格的话，我们就得到了属性性质的事件．

事件又可以分为简单事件和复合事件，前者指对于所研究的具体问题来说是不能再分解的事件，后者则指由简单事件复合而成的．例如，在抛掷硬币的例子中“出现正面”和“出现反面”都是简单事件．在抽查某工厂的10件产品的例子中，“没有次品”和“有1件次品”是简单事件，“有多于两件次品”则是复合事件，它由“有3件次品”、“有4件次品”、……、“全部10件都是次品”等8个简单事件复合而成．不过

要提醒大家注意，事件区分为简单和复合，是对所研究的具体问题来说的，具体问题不同，其区分的标准和结果也不同。譬如说，对于同一个抽查产品的例子，若我们关心的是10件产品中的次品数，简单事件和复合事件的区分方法，就是刚才所说那样。但若我们规定了次品数不多于两件为合格，否则为不合格，而检查的目的是确定产品是否合格。那么，这时就只有两种可能结果——“合格”或“不合格”，相应的简单事件也就只有两个。

如上所述，每一事件都有它的反面，事件和事件又可以复合，所以我们不能孤立地研究个别的事件，而应当从事件与事件之间的关系来进行研究，因此，需要考虑事件之间的关系及运算。我们仍用具体例子来说明。在抽查某工厂的10件产品中的次品数这个例子里，若以A表示事件“有2～6件次品”，B表示事件“有3～5件次品”，则易见事件A包含事件B，即当事件B发生时，事件A一定发生，这时记为$A\supseteq B$（图1-1(a)）。如果关系$A\supseteq B$与$B\supseteq A$同时成立，则称事件A和事件B等价；记为$A=B$。



(a)$A\supset B$　(b)$C\cup D$　(c)$C\cap D$　(d)$C\setminus D$

图1-1

如果以C表示事件“有 3 ～ 7 件次品”，D表示事件“有 2 ～ 5 件次品”，则事件C与事件D之和是事件 “有 2 ～ 7 件次品”，即“C与D至少有一个发生”，我们记之为$C \cup D$（图 1－1 (b)）。

事件C与事件D之积是事件“有 3 ～ 5 件次品”，即 “C与D同时发生”，我们记之为$C \cap D$或CD(图 1－1 (c))。

事件C与事件D之差是事件“有 6 ～ 7 件次品”，即“C发生而D不发生”，我们记之为$C \setminus D$(图 1－1 (d))。

如果事件A和事件B不能同时发生，即$A \cap B = \phi$，则称A与B互不相容(或互斥)。例如，事件“次品数为奇数” 与事件“次品数为偶数”是互不相容的，因为10件产品中的次品数可取 0 ， 1 ， 2 ，…，10等11个可能值，但一经检查后次品数即为确定，它若是奇数就不能是偶数，反过来亦如此。当A与B互不相容时，它们的和$A \cup B$有时特别记为$A + B$。

如果事件A与事件B互不相容，而且它们中间必有一个发生，即$A + B = \Omega$，则称B是A的对立事件（当然，A也是B的对立事件)，并以$B = \bar{A}$表示之。例如，抛掷硬币的试验中“出现正面”和“出现反面”是互为对立事件；抽查10件产品中“没有次品”和“至少有一件次品”也是互为对立事件。但“有一件次品”和“有两件次品”这两事件虽然互不相容，却非互为对立，因为它们不满足二者必有一个发生的条件。

类似地，我们可以定义 n 个事件 A_1，A_2，…，A_n 的和 $A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup\limits_{i=1}^{n} A_n$；积$A_1 \cap A_2 \cap \cdots \cap A_n = \bigcap\limits_{i=1}^{n} A_i$。由于实际问题和理论研究的需要，我们还可以把事件的运算推广到无穷多个事件A_1，A_2，…，A_n，…的情形，它们的和与积分别记为$\bigcup\limits_{i=1}^{\infty} A_i$与$\bigcap\limits_{i=1}^{\infty} A_i$。例如，为试制某种新产品作一连

串的试验，以A表示事件“新产品试验成功”，A_i表示事件“第i次试验获得成功”，则有$A=\bigcup\limits_{i=1}^{\infty}A_i$。我们还不难推知$A$的对立事件“新产品试验失败”$\overline{A}=\bigcap\limits_{i=1}^{\infty}\overline{A}_i$，这里$\overline{A}_i$是$A_i$的对立事件，即事件“第$i$次试验失败”。

如果一组事件A_1，A_2，…，A_n满足条件

$$\bigcup_{i=1}^{n}A_i=\Omega,$$

则称$A_1, A_2, \cdots, A_n$构成一个事件完备系．若还满足条件

$A_i\cap A_j=\phi$，　$i\neq j$，i，$j=1$，2，…，n，则称之为互不相容的事件完备系．这个概念以后常常要用到．在抽查10件产品的次品数这个例子中，事件“没有次品”、“有一件次品”、“有两件次品”、……、“有10件次品”等十一个简单事件就组成一个互不相容的事件完备系．要注意这样的事件系不是唯一的．譬如，在同一例子中“没有次品”和“至少有一件次品”这两事件也组成一个互不相容的事件完备系．一般说来，两个互为对立的事件都组成一个互不相容的事件完备系．

§1-2　概率的定义和算法

我们研究随机事件，目的在于揭示它们内在的统计规律性，研究它们发生的可能性有多大，只有这样做才能适应人类认识世界和改造世界的需要．例如，知道了电话总机“每小时接到几次呼唤”的可能性，对于线路、设备、人员等的配置是有实际意义的．同样，如果知道某工厂产品的“次品率少于某一百分比p”的可能性有多大的话，对于该工厂产

品质量的估计和指导生产也有很大作用。因此，需要确定一个能够定量地刻划随机事件A发生的可能性的指标$P(A)$，我们把这个数量指标$P(A)$称做随机事件A发生的概率（又称几率、或然率）。既然这个指标是统计规律性的一种数量表现形式，那末，它就应当是事件本身所固有的、不随人们主观意志而改变的一种客观属性。而且，人们可以利用事件的特性，或者通过大量重复试验以及其它方法来认识它。

下面介绍几种常见的确定概率的方法。

（一）古典法

在某些特殊情形中，我们可以利用物理模型本身所具有的对称性，运用演绎的方法去计算事件的概率，这就是确定概率的古典方法（有人称之为概率的古典定义或古典模型）。这方法的具体描述是：对于某种随机现象，记每次试验中它可能出现的结果为E_i，如果

（1）试验结果的数目是有限的，其全体为$\{E_1, E_2, \cdots, E_n\}$，每次试验必出现其中的一个；

（2）$E_1, E_2, \cdots, E_n$互不相容，并且具有等可能性。

我们把每一个试验结果E_i（$i=1, 2, \cdots, n$）称做基本事件。

例如，在随机地抛掷一个均匀硬币的例子中，由硬币的均匀性我们自然可以假定这试验具有对称性，即“出现正面”与“出现反面”的可能性相等，于是这两事件发生的概率都等于1/2。又如在0，1，2，…，9这十个数字中随机地取出一个，于是每次试验可得这十个数字中的一个，即共有十个可能的试验结果。因为这时十个数字所处的地位是一样的，所以我们也可以假定每一个数字出现的可能性相等，即每一数字出现的概率均等于1/10。