高等学校试用教材

概率论与数理统计

成都地质学院《概率论与数理统计》编写小组

地质出版社

高等学校试用教材

概率论与数理统计

成都地质学院
概率论与数理统计编写小组

地质出版社

内容简介

本书是为高等院校地质类各专业编写的试用教材。全书分为四篇，共十九章。第一篇与第二篇共八章，是课程的必修部份，分别介绍了概率论与数理统计的基础知识，并在每章后附有习题。第三篇与第四篇共十一章，是课程的选修部份，分别介绍多元分析的一些方法，以及关于随机过程和信息论的一些最基本的概念与知识。

本书可供地质技术人员参考。

概率论与数理统计

成都地质学院

概率论与数理统计编写小组

责任编辑：潘麟生　金绍棠

*

地质部教育司教材室编辑

地质出版社出版

地质印刷厂印刷

新华书店北京发行所发行·各地新华书店经售

*

开本850×1168 1/32·印张：15·字数：390,000

1981年4月北京第一版·1981年4月北京第一次印刷

印数1-8790册·定价2.70元

统一书号：15038教95

前　　言

投掷一枚硬币到桌面上，究竟硬币的哪面向上是不肯定的；用天平衡量同一物体几次，各次称得的重量也不见得准确一致，彼此略有偏差，并且指定的两次其偏差有多大也不能确切预知；容器里的气体，其分子在运动着，相互碰闯，对于个别的分子其运动轨道也难以事前确知。诸如此类的现象都是**随机现象**。其所以不肯定、不能事前确知结果，乃由于在一定的基本条件下，总有另外某些次要（随机的）因素的影响所致。

怎样来研究随机现象呢？

假如把所有次要因素与基本因素一视同仁、包罗无遗地去分析它们对现象的影响，则势必浩繁不堪、无济于事。因此，对于随机现象的研究必须别开途径。

虽然随机现象的个别结果是随机的、不确定的，但是大量的同类随机现象却通常总是呈现出一种完全确定的规律性。例如将一枚硬币投掷多次，则这硬币各面向上的次数就近乎相等；用天平称同一物体多次，所得重量的平均值总是游移于一个定值的附近；容器里的气体，如果分子足够多，则气体就以一定的压力作用于器壁。

恩格斯说：**"凡表面上看去是偶然性在起作用的地方，其实这种偶然性本身始终是服从于内部的隐藏着的规律的"**。（《费尔巴哈与德国古典哲学的终结》，人民出版社，第38页。）

单独一次实验结果的不肯定（随机）性，在重复多次实验的结果中却呈现出数量的规律性，**概率论**就是研究随机现象的数量规律性的一门学科，运用它来对随机现象进行分析，可以阐明这种规律性。

数理统计扎根于概率论，致力于为研究工作者设计实验方

案、提供分析法则，而且针对特定的实验问题构造出相适应的数学模型来进行分析。迄今还有许多重要问题的解决有待统计理论的进一步发展与扩充。乍看起来，统计理论好象就是概率论的自然引伸；但不知何故，在漫长的岁月中，统计理论几乎完全为概率论学者所忽略。因而统计理论与统计方法的发展至今还方兴未艾，而且由于科学研究的发展，所需的相应统计工具也愈益复杂而专门化起来。地质科学工作者为了对地质科学的专题进行研究、对特定的实验进行分析，往往要改创出切贴它的统计理论与统计方法，这就要求不仅对一般的数理统计知识有所了解，并且应阐熟精通、运用自如，必要时还要能从事改革创新，才能顺利地完成预定的任务。

概率论与数理统计是应实践的需要而发展起来的。十七世纪初叶，著名的物理学家伽利略就企图把物理测量的误差看成随机的以估计其概率。十七世纪中叶，巴斯加（Pascal,1623-1662）与费马（Fermat,1601-1665）等研究随机博弈，逐步建立了概率、数学期望等等重要概念及其基本性质与运算法则。后来，概率论在各种学科中得到了广泛的应用，原因是它蕴含着不少对实际问题的数据处理很有用的法则和依据。从简单的随机博弈开始研究，甚至现在还往往以它为例来形象地说明概率论的基本定律与法则，理由是大量的随机现象因素烦琐复杂，而人们对客观事物的认识只能是由浅入深、以简驭繁。

后来很多人对概率论的发展作出了杰出的贡献。整个十八世纪与十九世纪初叶，概率论简直发展到了使人着迷的程度，有些人就把它滥用于不当用之处，因而在十九世纪二、三十年代，概率论的巨大魅力曾一度为失望与怀疑所代替。幸亏许多学者成功地进行了卓越的理论探讨，才奠定了现代概率论的基础，而且大大地发展了它。随机过程论就是它的一个新的分支，并且诞生了有别于数理统计而具有独特方法的学科——信息论。

特别在十九世纪以来，好多新的学科就是在数理统计的基础上发展起来的。例如，里尔（Lyell）在1830至1833年间，出版

了他的著作《地质学原理》，共三卷，其中运用数理统计的论据确定了第三纪岩石地层的次序。他同贝类学家德夏依斯（Deshayes）把各个地层中化石的种加以辨认与记录，并确定其中迄今在海洋中还活着的种的百分率，据以制定更新世、上新世、中新世、始新世等地层的名称。但当他的命名为学术界公认以后，他所用的数理统计方法却几乎湮没无闻了。生物科学中也有类似的例子。生物学家达尔文（1809-1882）在乘船去南美考察的旅程中，可能由于阅读里尔著作第二卷的启发而肇始了他后来所提出的进化论。孟德尔在1866年所发表的关于豌豆杂交的研究，也是一个数理统计问题。

因此，在十九世纪人们迫切地去为数理统计谋求牢固的基础。本来是数学物理学家的皮尔逊（Pearson，1857-1936），费了将近五十年的精力去从事数理统计的认真研究，把数学用于进化论，并且创办了生物测量杂志（Biometrika）及一所数理统计学校，他的学生哥塞特（Gosset，1876-1937）是在一家酿酒厂工作的科学家，鉴于将大样本理论用于小样本实验工作总是格格不入，于是他另作研究，从弄乱了的卡片中进行抽样、计算，他将其成果发表在1908年的生物测量杂志上，署名为“Student”。后来，Student分布或 t 分布已成为数理统计、实验工作中的一种基本工具。大多数目前常用的数理统计方法，则是在二十世纪中发展起来的。

为了适应我国社会主义现代化的需要，在高等地质院校的数学教学计划中必须加强概率论与数理统计的内容，因此迫切需要有一本相应的有关基础内容的教科书，本书就是为此目的而编写的。全书共分四篇。第一篇是概率论基础，叙述了随机事件、随机变量、分布函数、数字特征、极限定理等基本概念与基础理论；第二篇是数理统计基础，叙述了参数估计与假设检验的基本理论，并在此基础上介绍了有广泛应用的方差分析与回归分析的基本内容；第三篇是多元统计分析，介绍了在地质工作中常用的几种多元分析方法；第四篇是随机过程与信息论简介，扼要介绍

随机过程的一些基本概念与基本理论，以及有关信息论的几个最基本的概念。

本书由叶乃膺教授任主编，康继鼎任付主编。第一、二、三、四、十三诸章及附录由李南赣执笔；第五、六、七、八、九、十、十四、十五诸章由师万瑞执笔；第十一、十二两章由黄南强执笔；第十六、十七、十八、十九诸章由康继鼎执笔。

本书由合肥工业大学潘麟生、河北地质学院金绍棠担任主审。参加审稿的还有中南矿冶学院温欣深，贵州工学院顾悦，云南大学何湘藩以及南京大学洪再吉等同志。审稿同志都认真审阅了原稿并提出了很多改进意见。本书在编写过程中得到我数学教研室许多同志的支持与帮助，尤其是杨万治与彭光复二同志承担了大量的誊写工作，我院绘图室的同志代为绘图，付出了辛勤的劳动。统此深志谢忱。

限于编者水平，编写时间仓促，不妥之处在所难免，望广大读者不吝批评和指正。

成都地质学院

概率论与数理统计编写小组

1980.11.21

目　　录

第一篇　概率论基础

第二篇　数理统计基础

第三篇　多元统计分析的一些方法

附　　录

第一篇　概率论基础

第一章　随机事件及其概率

§1　随 机 事 件

一、随机事件的概念

自然界中，有许多现象，我们完全可以预言它们在一定条件下是否出现，例如："在标准大气压下，水加热到100℃时必定沸腾"，就是必然出现的现象；"同性电互相吸引"就是必然不出现的现象。前一种在一定条件下必然出现的现象叫做必然事件；后一种在一定条件下必然不出现的现象叫做不可能事件。然而，自然界中还有许多现象，它们在一定的条件下可能出现，也可能不出现，这类现象称为随机事件。例如："掷一硬币出现正面"、"在某未知矿区钻到100米深时见矿"等都是随机事件（简称为事件）。

概率论是数学的一个分支，它研究的对象是随机事件的数量规律性。

我们常常是通过随机试验来观察随机事件的。那么，什么是随机试验呢？所谓随机试验是指一个试验，如果不能准确地预言它的结果，而且在相同条件下可以重复进行，就称这一试验为随机试验。例如事件"出现正面"就由随机试验"抛掷硬币"来观察；事件"见矿"就由随机试验"钻100米深"来观察。随机试验的每一个可能结果，都称为一个基本事件，随机事件系由若

干个基本事件组成。例如，我们的随机试验是掷一枚硬币，这时共有两个基本事件："正面"与"背面"；又如，我们的随机试验是掷两次硬币，这时共有四个基本事件：A（正$_1$，背$_2$）、B（正$_1$，正$_2$）、C（背$_1$，正$_2$）、D（背$_1$，背$_2$）。如果我们在上述的投掷中，考虑"一次是正面，另一次是反面"的事件，这就是由A与C组成的复合事件。

以下，"必然事件"用Ω表示，"不可能事件"用ϕ表示，"随机事件"则用大写拉丁字母A,B,C，…表示。

二、事件的包含与相等

如果事件A出现必导致事件B出现，则称事件B包含事件A记作$A\subset B$或$B\supset A$。

例如A表示"某物体平均温度不大于18℃"，B表示"某物体平均温度小于20℃"。这时事件A发生必导致事件B发生，即$A\subset B$。

若事件A包含事件B，并且事件B包含事件A，则称事件A与事件B相等，记作$A=B$。

例如在检查某圆柱形产品中，要求它的长度和直径符合规格才算合格，这时若A表示"产品合格"，B表示"直径和长度都符合规格"，则$A=B$。

三、事件的和、差、积

对于两个事件A与B而言，"二事件A与B中至少有一个出现"也是一个事件，称此事件为A与B的和（或并），记作$A+B$（或$A\supset B$）。

例如A表示"在一分钟内，某电话总机接到呼唤的次数在5至100之间"，B表示"在一分钟内，某电话总机接到呼唤的次数在16至180之间"。于是"在一分钟内，某电话总机接到呼唤的次数在5至180之间"就是事件$A+B$。

对于两个事件A与B而言，"A发生而B不发生"的事件称为A与B之差，记作$A-B$。

例如A表示"呼唤次数不小于6"，B表示"呼唤次数不小

于7”，那末$A-B$就表示“呼唤次数为6”

对于任意两个事件A与B，“A与B同时发生”的事件称为A与B之积（或交），记作$A\cdot B$（或$A\cap B$）。

例如A表示“某金属含量为5%～15%”，B表示“某金属含量为10%～20%”，那么$A\cdot B$表示“某金属含量为10%～15%”。

事件和与事件积的概念不难类似地推广到有限个事件的情形。设有事件A_1，A_2，……A_n，我们以

$$A_1+A_2+\cdots+A_n=\sum_{i=1}^{n}A_i$$

表示“A_1，A_2，…A_n诸事件至少发生其一”的事件；而以

$$A_1\cdot A_2\cdots A_n=\prod_{i=1}^{n}A_i$$

表示“A_1，A_2，…，A_n诸事件同时发生”的事件。

四、互不相容事件与对立事件

如果二事件A、B满足关系

$$A\cdot B=\phi,$$

也即是说，如果A与B不可能同时出现，则称A与B是互不相容的。例如事件“某金属的含量为5%～15%”与事件“某金属的含量为20%～25%”就是互不相容的。

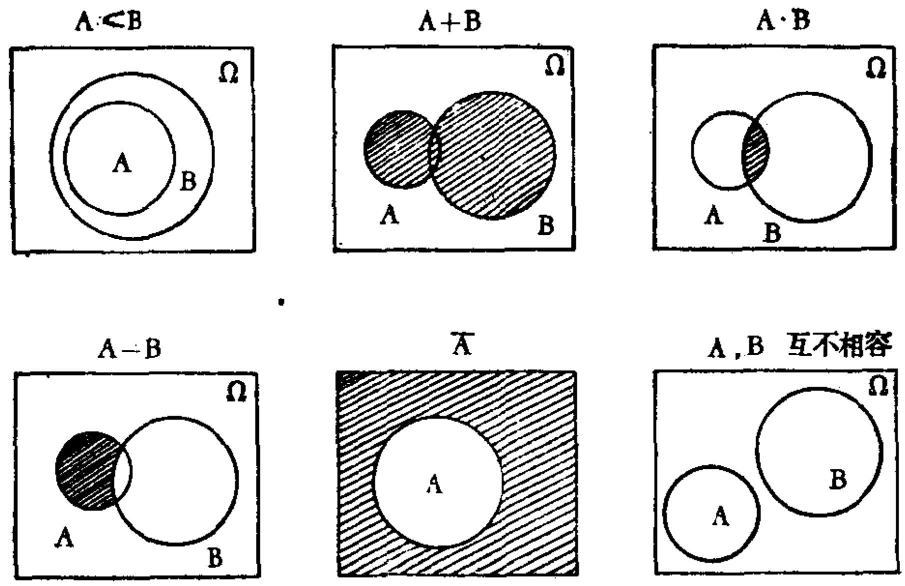
二事件A与B如果满足关系

$$A+B=\Omega,\quad A\cdot B=\phi,$$

也即是说，A与B中必出现其一，但A与B不能同时出现，则称A与B互补，或者说A与B互为对立事件，记作$A=\overline{B}$或$B=\overline{A}$。

例如事件“某金属的含量大于10%”与事件“金属含量不大于10%”就是对立事件。

事件之间的关系，可以用下列的图示来表达。



(其中A+B,A·B,A−B,$\bar{A}$分别为图中阴影部分)

图 1-1

§2 随机事件概率的定义

随机事件是客观存在的，我们不仅要会判断某个事件是否是随机事件，而且更重要的是进一步研究它发生的可能性的大小。例如，修筑大水坝时，必须根据历年水文资料来估计今后若干年内发生特大洪水的可能程度，这对于设计水坝有着重大的现实意义；又如对于矿床中的金属含量、岩层厚度等的可能分布有所了解，那将会有利于选择合适的普查和勘探工作方法，这是实际工作所必须的。

实践证明，事件发生的可能性大小是事件本身所固有的一种客观属性。我们常常可以用一个数字来描述随机事件发生的可能性大小，这个数字就是随机事件的概率。

一、概率的统计定义

例1·1 检查大批的产品，当被检查的产品长度介于13.60cm到13.90cm内，则产品为合格品，否则为次品。我们分别抽取5件、10件、60件、150件、600件、900件、1200件、1800件来检查，其结果

如下表所示：

总　　数	5	10	60	150	600	900	1200	1800
合格产品数	5	7	53	131	548	820	1091	1631
合格品的频率	1	0.7	0.883	0.873	0.913	0.911	0.909	0.906

$$合格品的频率=\frac{合格品数}{总数}$$

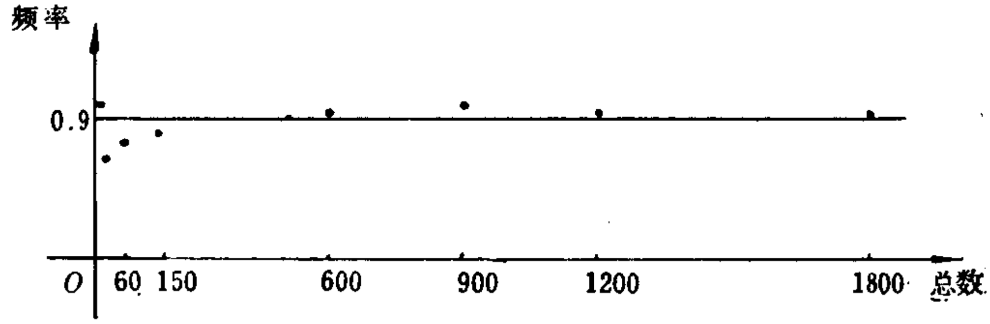


图 1-2

虽然，在抽出的产品中，次品数目是随机的，然而可以看出：随着抽查数目的增加，合格品的频率虽有摆动，但总是在常数0.9左右，而只有在极少数情况下，合格品的频率才与0.9有较大的差异。这种频率在某个值左右摆动而势于稳定的性质，完全是由其本身的客观规律性所决定。由此看来，我们不妨把0.9作为抽得合格品的概率。

例1·2　过去，曾有人在相同条件下，把同一枚质地均匀的硬币重复掷许多次，他所得到的实验结果如下：

掷硬币次数 n	正面出现次数 v	出现正面的频率 $\frac{v}{n}$
4040	2048	0.5069
24000	12012	0.5005

我们可以看出，当掷硬币的次数增加时，出现正面的频率$\frac{v}{n}$