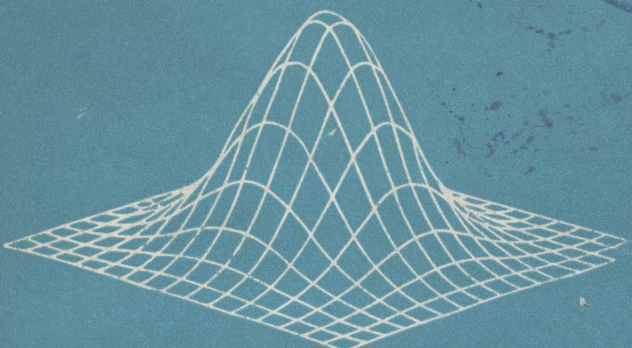
概率论与数理统计

山西省工科院校数学教材编写组



山西高校联合出版社

概率论与数理统计

山西省工科院校数学教材编写组

山西高校联合出版社

概 率 论 与 数 理 统 计

山西省工科院校数学教材编写组

*

山西高校联合出版社出版发行 （太原南内环街31号）

太原千峰科技印刷厂印刷

*

开本：787×1092 1/32 印张：15 字数：324.5千字

1991年8月第1版 1991年8月太原第1次印刷

印数：1—9850册

*

ISBN 7—81032—065—3

O·7 定价：4.50元

出版说明

在山西省教育委员会的指导下，为适应教育改革的需求，山西省五所工科院校联合编写了（1）线性代数（2）线性代数学习指导书；（3）概率论与数理统计；（4）概率论与数理统计学习指导书；（5）高等数学习题课讲义共五本教材与学习指导书，本书是这套书中的一本。

本书是依据国家教委颁发的高等工业学校《概率论与数理统计课程教学基本要求》编写的。全书内容包括概率论与数理统计两部分，共十一章。大部分节后附有练习题，这是要求读者应该掌握的基本内容。每一章后又附有数量和难度适中的习题，它有助于读者对基本内容的进一步理解和深化，书末附有习题答案。讲授全书约需68学时。对于学时较少者，可根据不同专业的情况，略去部分内容。

配合此书而编写的《概率论与数理统计学习指导书》，其内容编排大致与本书相一致，但指导书更突出一般学习方法的指导，与本书相配合，能更有效地帮助读者学好这门课程。

本书由太原重型机械学院张鸿秀、张宝玉二同志任主编，第一、四、六章由太原重型机械学院郭正光同志编写，第二、三章由山西矿业学院左蔚文同志编写，第五、七、八章由张鸿秀同志编写，第九、十、十一章由太原工业大学王学深同志编写。全书的图形由吴夏阳、单志辅、周静卿三同志绘制。

山西大学常学将教授对本书提供了宝贵的意见，为本书

写了序言，在此向他表示深深的谢意。

本书宜作为高等工科院校各专业《概率论与数理统计》课程的通用教材，也可作为各类成人高等教育同类专业的教科书，也是工程技术人员的合适参考书。

由于编者水平有限，书中难免有不少缺点和错误，恳请读者批评指正。

编　者　1991年

序　言

概率论与数理统计是研究随机现象的数量规律的数学分支。它在自然科学、技术科学、社会科学等领域已得到广泛的应用，它作为一门科学已被人们广泛的承认。但是至今也还有一些人只承认因果关系，不承认有概率。认为使用概率统计研究问题只是权宜之计，是在对事物了解不够的情况下才不得已而为之，只有找到因果关系才是科学的成果。包括伟大的物理学家爱因斯坦，在他后半生将近四十年的时间内都在从事统一场论的研究，希图实现他追求的完全的因果性。他用讥讽口吻对量子力学的统计陈述说："这个理论有很大贡献，但是它并不使我们更接近上帝的奥秘，无论如何我相信上帝不是在掷骰子……"。"我正在辛苦地工作，要从广义相对的微分方程，推导出看作奇点的物质粒子的运动方程……"。据说，在他与同时代的量子力学家玻尔论战中，提出一个黑盒子问题，使玻尔难以解决。玻尔在去世时仍在他工作室的黑板上画着这个解决不了的黑盒子，玻尔是长怀无已。但是爱因斯坦的统一场论也未获成功，尽管得到许多重要的成果。可见完全的因果观点，即全知的观点只是一种幻想。把因果与概率结合起来考察事物，将会获得更大的成就。

当然，科学的进步可以减少偶然性。例如，经过设计与技术改造，可以使射击的弹道偏差缩小，但是随机性并未消失，只不过是对改变了试验条件的随机事件进行研究而已。

又如地震预报问题，不管有无板块理论，都没有改变地震的随机性。当然有了板块理论，也许预报会更准确些，那只不过给定条件下，研究发生地震的问题，这是条件概率问题。如果一定要找出地震的因果关系，那只是恩格斯所讥讽的宿命论者，恩格斯说他们认为“这一些事实都是不可改变的因果连锁……，产生太阳系气团时，早就这样构成，以致这些事情只能这样发生，而不能是别样的”。再如温度的概念，是任何人都感觉得到的。然而单个分子是没有温度的，温度是大量分子聚集起来呈现的总体特征。所以温度不是一个一个的研究分子的特性所能获得的概念，即使将每个分子的运动方程都写出来，也得不到总体概念，何况对无数个方程的计算量也就变得无法实现。这种总体特性的研究正是概率论的研究对象。

数理统计方法虽然是从若干个个体（即样本）出发，推断总体的特性，但是它并不追求个体成员共同具有的某种特性，它不同于通常理解的归纳方法——由特殊到一般。例如温度这个集体特性，对单个分子是无意义的，它不是特殊的分子性质归纳出的一般性质。用样本推断总体，做出的结论有一个飞跃过程，不是常用的归纳过程，只能说是另一种归纳形式。

统计推断的方法使人们接受某种结论常常带有一定风险，于是人们常会怀疑统计结论是否保证真实性。什么是真实性？人们都认为真实就是符合客观现实。但是客观现实又是什么呢？难道任何决策都要等到完全弄清什么是现实的真实性才能作出吗？恐怕那将寸步难行。一个有才能的决策者，决不是只走无风险之路的人。统计决策方法有助于他寻找成

功与失败的原因。所以统计学不单作为理论学科，它更是一种工具，一种普遍的方法论。

当今，任何学科都不能排斥概率统计，所以在科学教育中应普遍设置概率统计课程的看法已被人们所接受。事实上，在工科院校已作为基础课进行教学。但要使学生学好这门课程，普遍感到有一定困难。特别是其中随机变量及其分布，在教与学中所遇到的困难，与让学生学习极限理论一样。后者要求学生由学习常量数学，转而接受变量数学，有个学习上的飞跃；而前者则要求学生由学习确定性数学，转而接受研究随机性的数学，必将要求学生在学习上的又一次飞跃。所以对教与学来说一本成功的教材是至关重要的。张鸿秀副教授主编的《概率论与数理统计》一书是总结了编写组成员多年教学成功经验之后写出来的，深入浅出的处理了许多难点。本教材以讲授实用方法为主导，但又不忽视基本理论的培养，坚持以实例引入概念，着重强调统计背景，不使读者落入形式地理解定义定理等，在适当部分提出一些思考问题，以启迪学生的思维。由此种种，使本教材具备了易读、实用、准确的特色。此外，为了适应性更加广泛，较诸现行大纲规定的章节有所增加，诸如方差分析，正交试验设计、回归分析等章节，可满足或兼顾电类、机械类、工业企业管理等专业之需。总之，本教材相对于现有各种流行的版本在内容与形式、文词与结构、均具独到之处，确是一部好的教材。相信人们经过对教材的阅读会感到是一种享受。

常学将

1991.6于太原

目 录

第一章 概率论的基本概念……………………………(1)
§1.1 随机事件的直观意义及其运算………………(1)
§1.2 概率的直观意义及其计算……………………(17)
§1.3 概率的公理化定义及其性质…………………(25)
§1.4 条件概率…………………………………………(31)
§1.5 事件的相互独立性………………………………(41)
§1.6 独立试验概型……………………………………(48)
习题一……………………………………………………(52)
第二章 随机变量及其分布……………………………(55)
§2.1 随机变量的概念…………………………………(55)
§2.2 离散型随机变量…………………………………(59)
§2.3 连续型随机变量…………………………………(70)
§2.4 正态分布及其应用………………………………(85)
§2.5 二维随机变量及其分布…………………………(95)
§2.6 随机变量函数的分布……………………………(119)
习题二……………………………………………………(137)
第三章 随机变量的数字特征…………………………(141)
§3.1 数字期望…………………………………………(142)
§3.2 方差………………………………………………(156)
§3.3 协方差及相关系数………………………………(165)
习题三……………………………………………………(176)

第四章　大数定律与中心极限定理………………(179)
§4.1 大数定律……………………………………(179)
§4.2 中心极限定理………………………………(184)
习题四 …………………………………………(189)
第五章　马尔可夫链………………………………(191)
§5.1 马尔可夫链的基本概念……………………(191)
§5.2 转移概率……………………………………(193)
§5.3 遍历性与平稳分布…………………………(204)
习题五 …………………………………………(212)
第六章　抽样分布…………………………………(217)
§6.1 基本概念……………………………………(217)
§6.2 抽样分布……………………………………(224)
习题六 …………………………………………(236)
第七章　参数估计…………………………………(238)
§7.1 参数点估计的两种常用方法………………(238)
§7.2 估计量的评选标准…………………………(248)
§7.3 区间估计……………………………………(255)
习题七 …………………………………………(271)
第八章　假设检验…………………………………(273)
§8.1 假设检验的基本概念………………………(273)
§8.2 正态总体均值的检验………………………(278)
§8.3 正态总体方差的检验………………………(288)
§8.4 分布拟合检验………………………………(293)
习题八 …………………………………………(304)
第九章　方差分析…………………………………(307)
§9.1 单因素试验的方差分析……………………(309)

§9.2 双因素试验的方差分析…………………………(323)
§9.3 有交互作用的双因素试验的方差分析……(335)
习题九 …………………………………………………(346)
第十章 正交试验设计……………………………(350)
§10.1 正交试验设计的基本方法…………………(350)
§10.2 水平数不相同的正交试验设计……………(365)
§10.3 有交互作用的正交试验设计………………(367)
§10.4 多指标的正交试验设计……………………(373)
习题十 …………………………………………………(374)
第十一章 回归分析…………………………………(378)
§11.1 一元线性回归………………………………(379)
§11.2 一元非线性回归……………………………(405)
§11.3 多元线性回归………………………………(414)
习题十一…………………………………………………(422)
习题答案…………………………………………………(426)
参考书目…………………………………………………(450)
附录 常用数理统计表………………………………(452)
附表1 泊松分布表…………………………………(452)
附表2 标准正态分布表……………………………(454)
附表3 χ^2分布表…………………………………(455)
附表4 t分布表……………………………………(457)
附表5 F分布表……………………………………(458)
附表6 部分常用正交表……………………………(462)
附表7 相关系数临界值r_α表……………………(468)
附表8 正态概率纸…………………………………(469)

第一章 概率论的基本概念

§1.1 随机事件的直观意义及其运算

一 决定性现象与随机现象

人们在生产实践和科学实验中，发现对自然界和社会上所观察到的现象大体可分为两种类型。一类是事前可预料的，即在一定条件下必然发生或必然不发生的现象。我们把这一类现象称之为必然现象或决定性现象。例如，两个三角形三边对应相等，则此两三角形全等；任何两个质点之间都有相互吸引的作用力，引力的大小与两质点质量的乘积成正比，与它们的距离平方成反比；水在直流电的作用下，分解成氢气和氧气。而上述现象的反面，如三边对应相等的两个三角形不全等则必然不会发生。到目前为止，我们已学过的初等数学和高等数学（概率论除外），所研究的对象就是决定性现象中的数量关系。另一类是事前不可预料的，即在相同条件下重复进行观察或试验时，有时出现有时不出现的现象，我们称之为偶然现象或随机现象。例如，抛掷一枚硬币，其结果可能是正面(假设国徽的一面为正面)向上，也可能是反面向上，在一次抛掷之前是无法肯定向上的一面是正面还是反面；一门炮在发射角相同的条件下向某一目标发射多发炮弹，其弹着点也有所不同，类似的例子还可举出许多。

随机现象的特点是：在基本条件不变的情况下，一系列

试验或观察会得出不同的结果，而在一次试验之前是无法预知确切结果的，即呈现出不确定性。然而在进行了多次重复的试验或大量的观察后，人们逐渐发现在一次观察或试验中不能肯定结果的现象内部隐藏着必然的规律性。例如，掷一枚均匀的硬币，当投掷次数增加时，出现正面与反面的机会是差不多的；用炮射击目标时，若射击次数不多，弹着点似乎杂乱无章；当射击次数增多时，就会发现弹着点关于目标的分布是对称的；越靠近中心越密等。正如革命导师恩格斯所说："在表面上是偶然性在起作用的地方，这种偶然性始终受内部隐蔽着的规律支配的，而问题只是在于发现这种规律"①。在大量试验或观察中所呈现出来的规律性称之为统计规律性。概率论的任务就在于揭示和研究随机现象的这种统计规律性。

概率论起源于17世纪，至今已形成具有自己特色的数学分支，它的理论与方法在各领域中得到越来越广泛的应用。因此，学习概率论的知识将有助于我们更好地学习专业课和初步掌握解决某些实际问题的能力。

二 随机试验与事件

通常，我们把对自然现象的观察或进行一次试验，统称为一个试验。如果这个试验在相同的条件下可以重复进行，而且每次试验的结果事前无法预料，我们就称它为一个随机试验。一般用字母E 表示，以下通过几个例子来说明。

例1 试验E_1为掷一枚均匀的硬币，观察正面H（有国徽

①恩格斯：《路德维希·费尔巴哈和德国古典哲学的终结》，人民出版社1972年版，第38页。

的一面），反面T出现的情况。

例2 试验E_2为掷一枚骰子，观察点数出现的情况。

例3 试验E_3为一袋中装有编号1，2，3，4的四个球，从袋中任取一球后，不放回袋中，再从中任取一球观察两次取球的结果。

例4 试验E_4为记录某电话交换台在早8：00～8：10内接到的呼唤次数。

例5 试验E_5为从一批灯泡中，任取一只，测试它的寿命。

上面所举的5个试验的例子，尽管内容各异，但它们有着共同的特点，例如试验E_1，它有两种可能结果，出现H或者出现T，但在抛掷之前不能确定出现H还是T，这个试验可以在相同的条件下重复地进行。又如试验E_2，它有六种可能结果，即出现点数为1，2，…6中之一，但在投掷之前不能确定会出现第几点，这个试验可以在相同的条件下重复地进行。再如试验E_4，因为在某一天早8：00～8：10内接到的呼唤次数可能为0，或1，或2，…，但在试验之前不能确定会有几次呼唤，这个试验也可以在相同的条件下重复进行。总之，这些试验都具有随机试验的三条特性：

1 可以在相同的条件下重复进行；

2 每次试验的可能结果不止一个，并且事前能明确试验的所有可能结果；

3 进行一次试验之前不能确定哪一个结果出现，但每次试验总是出现这些可能结果中的一个。

以下所提到的试验都是指具有上述特征的随机试验。

对于某个随机试验，在一次试验中可能出现也可能不出

现的事情，称为这试验的随机事件，简称为事件，用字母A，B，C，…表之。

例如，在E_1中，“出现H”、“出现T”；在E_2中“出现5点”；在E_3中，“第1次取得球的号码为1”；在E_4中，“电话交换台接到的呼换次数不大于10”；在E_5中，“灯泡的寿命在(1000，1500)小时之间”等，都是相应试验的随机事件。

随机试验E中不可能再分的事件，叫做E的基本事件，它是最简单的事件。

例如，在E_1中，“出现H”、“出现T”是试验的基本事件；在E_3中，若用（1，3）表示第一次取得1号球，第二次取得3号球，其余如（2，4）…可作类似理解，则“两次取球的结果为（1，3）”是试验的基本事件；在E_4中，“交换台接到呼唤次数为6”；在E_5中；“测得灯泡的寿命为1200小时”等，都是相应试验的基本事件。

由若干个基本事件组合而成的事件，称为复合事件。例如，在E_3中，“第一次取得球的号码为1”是一个复合事件，它是由两次取球的结果为（1，2）、（1，3）、（1，4）这三个基本事件所组成，当且仅当这三个事件中有一个发生，则“第一次取得球的号码为1”这一事件就发生。

值的注意的是，一个事件是否为基本事件是相对于试验的目的而言，不是绝对的。例如。在E_5中，若考虑的是灯泡寿命的大小，那么基本事件为“灯泡的寿命为t”$t\geqslant0$，有不可数无穷多个基本事件。若按寿命取值的大小将灯泡分为优质品、合格品或次品，那么就只有三个基本事件，即“取得灯泡为优质品”，“取得灯泡为合格品”，取得灯泡为“次品”。

在每次试验中，一定发生的事件称做必然事件，用符号

Ω表示。在任何一次试验中都不可能发生的事件称做不可能事件，用符号ϕ表示。例如，在E_1中，"出现H或T"是必然事件，"出现H和T"是不可能事件。必然事件和不可能事件本来没有不确定性，也就是说它们不是随机事件，但为了今后讨论问题方便起见，常把这两个决定性现象中的事件看作特殊的随机事件。

三 样本空间

为了便于进一步研究随机试验E，我们将借助于集合论的概念①来描述随机事件。

对于某个随机试验，其中随机事件是各式各样的，并且在不同的试验中的事件又有很大的区别，我们把随机试验E中的所有基本事件所组成的集合称做E的样本空间。记做Ω。Ω中的元素就是试验E的基本事件。基本事件也称样本点，记做ω。

下面写出本节五个例子中试验E_k（$k=1, 2, \cdots, 5$）的样本空间Ω_k。

在E_1中，基本事件有两个，即"出现H"、"出现T"，所以，样本空间Ω_1中含有两个样本点，ω_1—出现H、ω_2—出现T，分别记为H和T。即$\Omega_1=\{H, T\}$。

在E_2中，基本事件有六个，即"出现k点"($k=1, 2, \cdots, 6$)，所以，样本空间Ω_2中含有六个样本点，分别记为ω_k——出现k点，简记为k（$k=1, 2, \cdots, 6$），即$\Omega_2=\{1, 2, 3, 4, 5, 6\}$。

---

① 有关集合论的知识请参阅《概率论与数理统计学习指导书》

在E_3中，由于两次取球的标号不能重复，按排列的知识共有$P_4^2=4\times3=12$种取球的结果，所以，E_3中有12个基本事件。故样本空间Ω_3中含有12个样本点。记ω_k—（i，j）（$i\neq j$，i，$j=1$，2，3，4），即

$$\Omega_3=\{(1,2),(1,3),(1,4),(2,1),(2,3),(2,4),(3,1),(3,2),(3,4),(4,1),(4,2),(4,3)\}$$

在E_4中，由于基本事件是"交换台接到的呼唤次数为K（$k=0,1,2,\cdots$），所以，E_4中有可数无穷多个基本事件、(实际上接到的呼唤次数是有限的，但无法具体限制其数目)，故Ω_4中含有可数无穷多个样本点，记ω_k——交换台接到的呼唤次数为k，简记为k（$k=0$，1，2，…），即

$$\Omega_4=\{0,1,2,3,\cdots\}。$$

在E_5中，由于基本事件为"任取一只灯泡，测试其寿命为t"，而t在区间$[0,+\infty)$内连续取值，所以，E_5中有不可数无穷多个基本事件。故样本空间Ω_5中含有不可数无穷多个样本点。记ω_t——任取一只灯泡，测试其寿命为t，简记为t。则

$$\Omega_5=\{t\mid t\geqslant0\}。$$

同样应当注意的是，样本空间的元素取决于试验的目的。例如在E_5中，如果只考虑取得灯泡的优劣，则$\Omega_5=$｛优质品，合格品，次品｝。

随机试验E的任一事件A，或是基本事件，或是由若干基本事件所组成，故引进样本空间Ω之后，事件A就可看作是样本空间Ω的子集。而事件A发生，当且仅当A所含的基本事件中有一发生，即为子集中的一个样本点发生。例如，在试验