

成人高等教育自学辅导丛书

概率论与 数理统计 自学指导

胡林 主编

谭鼎 翟连林

熊廷煌 张友极

编著

水利电力出版社

成人高等教育自学辅导丛书

概率论与数理统计自学指导

胡 林 主编

谭 翔 程连林 熊廷煌 张友极



水利电力出版社

2R29/21

内 容 提 要

本书共分十章。前六章介绍了概率的基础部分，内容包括：随机事件与样本空间的概念，随机事件间的关系及其运算、重要的常见分布、随机变量数学特征的分布函数以及随机向量的数字特征和分布函数。后四章主要介绍了数理统计部分，内容包括：参数的点估计与区间估计，假设检验、方差分析、回归分析。

本书内容通俗易懂，每节配有习题，每章后配有自我检查题、及小结，书末有习题提示与解答便于读者自学。

本书除供各类成人高等学校学生学习外，还可供普通高等学校的学生和教师参考。

成人高等教育自学辅导丛书 概率论与数理统计自学指导

胡·林·主·编

谭嘉、福连林、顾廷煊、张友振

水利电力出版社出版

(北京三里河路8号)

新华书店北京发行所发行 各地新华书店经售

北京联华印刷厂印刷

787×1092毫米 32开本 12.875印张 283千字

1986年6月第一版 1986年6月北京第一次印刷

印数00001—20800册 定价 2.70元

书号 7149·6040

前 言

当前，在经济体制改革和新技术革命挑战的形势下，智力开发的重要性更加突出了。我们迫切需要有一支高水平的职工队伍，以加速实现技术现代化，管理现代化，提高经济效益，这就要求在普遍提高职工的政治、文化、技术、业务素质质的同时，尽快从现有职工中培养造就大批的专业技术干部和管理干部，形成一支在数量上能基本满足要求，质量上能掌握现代科学技术和经营管理知识，专业配套的职工队伍。可以说：大力加强职工教育，培养各类人才，是摆在我们面前的一项十分重要而又急迫的任务。

这套“成人高等教育自学辅导丛书”就是根据当前加强职工教育的形势和需要而专门组织编写的。

“丛书”以“面向实际，面向生产，为提高职工队伍素质，提高经济效益服务”作为编写指导思想；内容紧密结合成人高等教育理工类（或财经类）部分课程的教学大纲和电视大学及一些函授大学、职工大学、业余大学的教材，在布局、选材、体例和编写形式上尽量适应成人自学的特点。所以，非常适用于理工、财经类电视大学、职工大学、业余大学学员作为学习辅导书，或函授大学作为函授教材；对于广大自学读者，则是帮助他们通过自学高等考试的一种自学读本。

为了切合读者的实际需要，提高学习效果，丛书中的每一册都包括基本概念、重点和难点解释、典型例题分析

结提示以及思考与练习等几部分内容，并配有适量的作业测验题（附答案）和电大试题选解。

这套丛书共包括十一门课程，十三册：

高等数学自学指导（上、下册）

线性规划自学指导

线性代数自学指导

概率论与数理统计自学指导

常微分方程自学指导

逻辑代数与BASIC语言自学指导

复变函数自学指导

微积分自学指导（财经类）

普通物理自学指导（上、下册）

普通化学自学指导

物理化学自学指导

本册《概率论与数理统计自学指导》，全书共分十章。前六章介绍了概率的基础部分，首先从古典概型入门，建立概率的概念，介绍事件的关系和运算以及相应的概率定理；接着以常见分布为先导引入分布的一般概念；最后介绍随机变量的数字特征和随机向量的基本知识。后四章介绍了基本的数理统计方法和原理。内容包括：参数的点估计与区间估计，假设检验、方差分析和回归分析。

本书内容通俗易懂，每节配有问题，每章后配有自我检查题，及小结，书末有习题提示与解答便于读者自学。

本书的编写得到了王湘浩教授的关怀与指导，在此向他表示深深的谢意，马少华同志为本书整理及习题解答作了大量的工作，在此表示感谢。

参加这套书编写工作的都是有经验的高等学校教师或成

人教育工作者，其中有些同志还讲授过电视大学的有关课程或担任过电大辅导课主讲教师。“丛书”融汇了他们多年的教学经验和心得体会，更鲜明地具有电视教学及自学、辅导、函授多用的特色。

在编写过程中，我们得到各课程的有关教授和专家的关怀和指导，有些同志直接参予审阅、整理等工作，在此一并表示深切的谢意。

组织编写这类面向成人读者，自学，辅导，电教、函授多用的大专读本还是第一次，欢迎读者对“丛书”的内容、布局、结构、形式等提出宝贵意见，以帮助我们改进工作，提高“丛书”质量。

胡 林

1985年7月

序

这是一本难得的好教材，它对没有机会进大学或者在电大、函大、职大、夜大学习而缺乏辅导教师的数学爱好者及理工科学员特别适合。本书的编者一直从事数学基础课教学，具有丰富的教学经验，这本书充分反映了编者们这些宝贵的经验。

一般教材需要经过教师在课堂上的再创造才能很好地发挥作用，而本书则体现出教师讲课时的这种创造，使读者感到好象是在课堂里亲聆教师的讲授。

经验告诉我们，学生感到困难、深入不下去的主要原因是缺乏自学能力，不善于给自己提出问题，本书在这一方面也下了很大功夫。编者在展开教材内容的同时，引导读者去提出问题、探讨问题、解决问题，提供了一条怎样深入学习的途径。

本书也是很好的教学参考书，它不仅在内容上涉及面广，而且更重要的是提供了一种可取的教学方法，这对于教师，特别是刚从事教学工作的青年教师具有很好的借鉴作用。

由于本书切合学员的学习实际，虽然增加了一定篇幅，但并不显得繁琐。相反，在日益蓬勃发展的成人教育事业中，本书必将发挥积极的作用，为广大读者所欢迎。特代为作序。

王湘浩

1985，4，长春

目 录

序

第一章 古典概型	1
1.1 随机事件与样本空间	1
1.2 概率的古典定义	5
1.3 计算方法与基本解法	9
1.4 技巧之一——模型变通	18
1.5 几何概率与统计概率	26
第二章 事件代数与概率定理	32
2.1 事件的关系与运算	32
2.2 概率加法定理	45
2.3 条件概率与概率乘法定理	50
2.4 事件独立性	56
2.5 全概公式与逆概公式	64
第三章 二项分布与泊松分布	78
3.1 贝努里概型与二项分布	78
3.2 二项分布的泊松逼近	88
3.3 泊松流与泊松分布	95
3.4 几何分布与多项分布	98
第四章 随机变量及其分布	104
4.1 离散型随机变量的概率分布	104
4.2 分布函数	112
4.3 连续型随机变量的概率密度	117
4.4 正态分布	127

4.5 随机变量函数的分布	135
第五章 随机变量的数字特征	146
5.1 数学期望	148
5.2 方差	159
5.3 切比雪夫不等式与矩	167
第六章 随机向量	174
6.1 二维随机向量及其分布	174
6.2 条件分布与随机变量的独立性	185
6.3 随机向量的分布函数与随机向量函数的分布	195
6.4 随机向量的数字特征	205
6.5 二维正态分布	217
6.6 大数定律与中心极限定理	221
第七章 参数估计	233
7.1 数理统计概述	233
7.2 参数的点估计	240
7.3 参数的区间估计	253
第八章 假设检验	265
8.1 假设检验的基本思想	265
8.2 u -检验	273
8.3 t -检验	277
8.4 χ^2 -检验和 F -检验	283
第九章 方差分析	289
9.1 单因素的方差分析	289
9.2 双因素的方差分析	300
9.3 平方和的简化算法	304
9.4 方差分析公式的证明	310
第十章 回归分析	317
10.1 回归分析的研究对象	317

10.2	回归口诀	318
10.3	回归口诀溯源	328
10.4	相关系数及其显著性检验	330
10.5	回归线的精度	336
10.6	一元非线性回归和二元线性回归	340
10.7	回归口诀的适用范围	350

附录

提示与解答

第一章 古典概型

1.1 随机事件与样本空间

一、随机事件

公司招标，十家争投，这会有多大？有奖储蓄，恰中头奖，这机会又有多大？出差他乡，偶遇故知，这机会又有多大呢？——现在，就让我们带着数学工具去考察这一机会世界吧！

用数学工具处理问题的第一步，是将对象概念化。我们把象“十家争投”，“恰中头奖”，“偶遇故知”这样一些“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；而把那些“在一定条件下必然发生的事件”和“在一定条件下必然不发生的事件”分别称为必然事件和不可能事件。比如，“在标准大气压下，水加热到 100°C 时，沸腾”就是一个必然事件；而“在标准大气压下，水加热到 100°C 时，不沸腾”就是一个不可能事件。此例也说明，不可能事件不过是否定形式的必然事件。

我们常用大写的拉丁字母 A ， B ， C 等表示随机事件，用 U 和 V 分别表示必然事件与不可能事件。为方便起见，我们还把必然事件和不可能事件当作两种极端情形的随机事件，并将随机事件简称为事件。实际上，偶然现象与必然现象并非绝然无缘。许多我们现在认为是必然现象的，起先却曾认为是偶然现象。比如，设想一个内地人到了西藏高原，发现

在那里烧水，不到 100°C 就沸腾。如果他不懂得这与气压有关，便会觉得“水烧到 100°C 时沸腾”是一个随机事件。可见，随着人们的认识从“必然王国”向“自由王国”的转变，原来视为偶然现象的，可能转化为必然现象。

当然，由于客观事物的错综复杂，以及人们的认识在一定时期的局限性，随机现象的存在仍然具有普遍性和长期性。我们应当学会对它特有的研究方法。概率论就是研究随机现象数量规律性的一个数学分支。

二、样本空间

研究随机现象，首先必须进行大量的观察或实验，以便从中发现统计规律性。为方便起见，我们把观察或实验统称为试验。

一个试验 E ，如果事先不能准确地预言它的结果，并且可以在相同条件下重复进行，就称 E 为随机试验，简称试验。

随机试验的每个（不能再分的）可能结果，称为该随机试验的样本点（或基本事件）。全体样本点的集合称为样本空间（或基本事件空间）。样本空间常用大写希腊字母 Ω 表示；而用其小写 ω 表示样本点。

例1 随机试验 E_1 ：掷一枚硬币，把出现正面（国徽一面）和出现反面作为可能结果，并分别记作 $\omega_{\text{正}}$ 和 $\omega_{\text{反}}$ ；

E_1 的样本空间： $\Omega_1 = \{\omega_{\text{正}}, \omega_{\text{反}}\}$

例2 随机试验 E_2 ：先后掷两次硬币，把依次所得的结果作为一个样本点。记

ω_{00} = “第一次出现正面，第二次也出现正面”

ω_{01} = “第一次出现正面，第二次出现反面”

ω_{10} = “第一次出现反面，第二次出现正面”

ω_{11} = “第一次出现反面，第二次也出现反面”

E_2 的样本空间： $\Omega_2 = \{\omega_{00}, \omega_{01}, \omega_{10}, \omega_{11}\}$

例3 随机试验 E_3 ：先后掷两次硬币，把两次（不计次序）所得的结果作为一个样本点，记

ω_0 = “两次都出现正面”

ω_1 = “有一次出现正面”

ω_2 = “两次都出现反面”

E_3 的样本空间： $\Omega_3 = \{\omega_0, \omega_1, \omega_2\}$

上面的例子启发我们两点值得注意之处：

1. 要知道样本空间，就得先给定随机试验；而所谓“给定”随机试验，就是要指明“把什么作为一个样本点”。

2. 指明样本点的关键是从问题的需要出发，明确“不能再分”的含意。例3与例2的区别就在于：由于不计次序， ω_1 就是不能再分的（不计次序的先后掷两次相当于同时掷两枚一样的硬币）。

例4 随机试验 E_4 ：观察某电话交换台一天中接到的呼唤次数，把每一种可能的次数作为一个样本点。则样本空间为 $\Omega_4 = \{0, 1, 2, \dots\}$ 。由于很难指定一个数作为呼唤次数的上界，所以我们不妨把 Ω_4 看作是一个无穷样本空间（即有无穷多个样本点），并且是可列的（即这些样本点可以按照某种次序排列起来）。

例5 随机试验 E_5 ：观察某地区的气温。我们可以取样本空间为 $\Omega_5 = (-\infty, \infty)$ 或 $\Omega_5 = [a, b]$ 。这个样本空间包含无穷多个样本点，它们充满了一个区间，而不是可列的。

从上面的例子，我们看到样本空间有两大类型：一类是可列的（有穷或无穷），一类是不可列的。前者称为离散样

本空间，后者称为非离散样本空间。

三、再看随机事件

现在，我们可以用样本空间的概念来定义事件。

首先，每个样本点都是一个事件，即基本事件。比如例 2 中的 ω_{00} , ω_{01} 等。

其次，比如在例 2 中，“第一次出现正面”也常常是我们感兴趣的一个事件。如果将它记作 A ，则 $A = \{\omega_{00}, \omega_{01}\}$ 。换句话说， A 是 Ω_2 的一个子集。再如 Ω_2 的下列两个子集

$$B = \{\omega_{01}, \omega_{10}\} \text{ 和 } C = \{\omega_{00}, \omega_{01}, \omega_{10}\}$$

分别表示事件“恰有一次出现正面”和“至少有一次出现正面”而 $\Omega = \{\omega_{00}, \omega_{01}, \omega_{10}, \omega_{11}\}$ 和空集 Φ 则分别表示必然事件 U 和不可能事件 V 。一般地，我们有

定义 样本点的某个集合（即样本空间的某个子集）称为事件。称某事件发生（或出现），当且仅当它们所包含的某一个样本点出现。

例如，当事件 ω_{00} （“两次都出现正面”）发生时，我们可以说事件 C （“至少有一次出现正面”）发生了；反过来，说事件 C 发生了，就意味着 $\omega_{00}, \omega_{01}, \omega_{10}$ 中的某一个发生了。

习 题 1.1

1. 什么叫随机试验、样本空间和随机事件？对样本点要求“不可再分”是什么意思？

2. 某袋中装有 4 只白球和 2 只黑球，考虑依次从中摸出两个球所可能出现的事件。为明确计，对球进行编号：4 只白球分别编为 1, 2, 3, 4 号；2 只黑球分别编为 5, 6 号。用数对 (i, j) 表示“第一次摸得 i 号球，第二次摸得 j 号球”这一个样本点。试写出全部样本点，并写出下列事件

所包含的样本点;

$A =$ “第一次摸出黑球”;

$B =$ “两次都摸出黑球”;

$C =$ “至少有一次摸出黑球”。

1.2 概率的古典定义

一、有限样本空间

我们把样本点有限的样本空间称为有限样本空间。把这种较为简单的样本空间作为讨论的出发点是较为适宜的。

有限样本空间的样本点具有两个明显的性质:

(一) 有限性 (个数有限);

(二) 互不相容性。即每次试验中, 至多只有一个样本点出现, 而不可能有多于一个的样本点同时出现。

二、古典概型 (等可能有限样本空间)。

让我们从更简单的情形出发来展开讨论。

定义 如果我们考察的随机现象, 其样本空间是有限的, 并且各样本点的发生或出现是等可能的 (即机会均等), 则称这类随机现象的数学模型是古典概型。

例如上节的例 1, 可以认为是古典型的。因为硬币两面的形状和质量大致是对称的和相等的。所以可以认为“出现正面”和“出现反面”这两个样本点是等可能的。由此又推知例 2 也是古典型的: 可以认为 ω_{00} , ω_{01} , ω_{10} , ω_{11} 出现的机会都是 $1/4$ 。但是例 3 不是古典型的。因为 ω_0 , ω_1 , ω_2 出现的机会分别是 $1/4$, $2/4$ 和 $1/4$ 。对此, 历史上曾出现过争论, 大数学家达朗贝尔 (D'Alembert) 就曾误认 ω_0 , ω_1 和 ω_2 是等可能的。

判别一个概率模型是不是古典型，关键是看“等可能性”条件满不满足；而对此，又通常是根据实际问题中的某种对称性进行理论分析，而不是通过实验来判定的。

三、概率的初等描述

我们研究随机现象，搞清样本空间，只是第一步。我们感兴趣的主要问题是：所要考虑的随机事件在一次试验中出现的“可能性”有多大？

对于一个随机事件 A ，我们可以用一个数 p 来表示它在一次试验中出现的可能性大小。我们就把这个数称为事件 A 的概率，并记作 $P(A)$ ：

$$P(A) = p$$

那么，对于一个给定的事件 A ，到底用一个怎样的数作为它的概率呢？即怎样从数量上来规定 $P(A)$ 呢？要一般地回答这一问题还不容易。但在某些特殊情形（比如古典概型）的随机试验中，事件的概率容易合理地规定，并能满足实际需要。

四、概率的古典定义

古典概型是在概率论发展初期就被重点研究，尔后又证明在概率论中占有基本重要地位的一类概率模型。1812年，法国数学家拉普拉斯（Laplace）给出了古典概型的概率定义。现在被称为概率的古典定义：

设 Ω 是等可能有限样本空间，其样本点总数为 n 。而事件 A 所包含的样本点数为 m ，则规定

$$P(A) = \frac{m}{n} = \frac{A \text{ 所含的样本点数}}{\text{样本点总数}}$$

比如，对于上节的例 2，我们有

$$P(\omega_{00}) = P(\omega_{01}) = P(\omega_{10}) = P(\omega_{11}) = 1/4$$

$$P(A) = P(\text{“第一次出现正面”}) = 2/4 = 1/2$$

$$P(B) = P(\text{“恰有一次出现正面”}) = 2/4 = 1/2$$

$$P(C) = P(\text{“至少有一次出现正面”}) = 3/4$$

特别地，有

$$P(U) = P(\text{“}\omega_{00}, \omega_{01}, \omega_{10}, \omega_{11}\text{中至少有一个出现”}) = 4/4 = 1$$

$$P(V) = P(\text{“}\omega_{00}, \omega_{01}, \omega_{10}, \omega_{11}\text{中一个都不出现”}) = 0/4 = 0$$

从概率的古典定义，显然可以一般地推出概率的下列性质：

(一) 对于任意事件 A ，有

$$0 \leq P(A) \leq 1$$

(二) $P(U) = 1$ ， $P(V) = 0$

事实上，设样本点总数为 n ， A 所含的样本点数为 m ，则 $0 \leq m \leq n$ ，所以 $0 \leq m/n = P(A) \leq 1$ ；又必然事件包含了所有样本点，不可能事件不包含任何样本点，所以 $P(U) = \frac{n}{n} = 1$ ， $P(V) = \frac{0}{n} = 0$ 。

我们看一个稍复杂的例子：设有一批产品共 100 件，其中有 5 件次品。现从中任取 50 件，问：无次品的概率是多少？

首先，从 100 件产品中任取 50 件，我们知道共有 C_{100}^{50} 种不同的取法，我们可以把每一种取法作为一个样本点，则样本点总数 C_{100}^{50} 是有限的；而且我们没有理由认为其中哪些取法的可能性比别的取法的可能性大，换句话说，我们应当认为 C_{100}^{50} 个样本点出现的可能性是相等的。因此，这个问题