GAILULUN YU SHULI TONGJI CHUBU

概率论与
数理统计初步

上海交通大学应用数学系编

工程数学丛书

上海交通大学出版社

·工程数学丛书·

概率论与数理统计初步

上海交通大学应用数学系编

上海交通大学出版社

内　容　提　要

本书是工程数学丛书之一。内容包括随机事件及其概率、随机变量及其分布、随机变量的数字特征、大数定律和中心极限定理、参数估计、假设检验等。每章配有精选的例题和习题，书末附有习题答案以及常用的概率分布表。可供高等院校理工科各专业作教材用，也可供工程技术人员参考及自学者选用。

概率论与数理统计初步

上海交通大学出版社出版

（淮海中路1984弄19号）

新华书店上海发行所发行

江苏太仓印刷厂印装

开本787×1092毫米　1/32　印张8.75　字数195000

1989年1月第1版　　1989年1月第1次印刷

印数：1—5350

ISBN7-313-00239-4/O·212　　科技书目：188—238

定价：1.75元

序　　言

目前，高等院校的工程数学课程，由于内容多、进度快，因此有些重要的内容不得不一带而过，或者删掉，尤其没有足够多的演题时间，从而使学生对课程内容的理解往往只浮于表面。为了弥补这些不足，编者吸取我校在工程数学教学中积累的有益经验，根据高等学校工科数学课程教学指导委员会（原工科数学教材编委会）的编写要求，针对工科院校的具体特点，编写了一套工程数学教材，并拟配以一册这套教材的习题解答，合称《工程数学丛书》。

丛书中的教材共有 5 册：《线性代数》、《复变函数》、《积分变换》、《概率论与数理统计初步》和《特殊函数与数学物理方程》。这套教材的特点是：内容丰富，说理清楚，重点突出，深浅得当，通俗易懂；对工程数学中的基本概念、基本理论和基本方法的叙述，力求深入浅出、清晰、准确；并配有大量典型例题和类型齐全的习题。本着循序渐进的原则，对全部内容由易到难，由浅入深地作了统筹安排，书中加有“*”号的内容可根据不同情况予以取舍。这对读者逐步地系统掌握工程数学的基本内容，进一步提高分析问题和解决问题的能力均有裨益。

由于上述特点，这套教材具有比较广泛的适用性。除了全日制高等院校以外，函授大学、电视大学、职工业余大学等都可用来作为教材，自学工程数学的广大读者也可选用，从事科研生产的工程师也可参考。

本套丛书主编袁公英，编委有张建元、贺才兴、武霞敏、吴登益、童品苗、陈茵、唐济楫等同志。《概率论与数理统计初步》的概率论部分由吴登益同志执笔编写，数理统计初步部分由童品苗同志执笔编写。在全套教材编写过程中得到校、系领导的关心、帮助和我系广大教师的大力支持，编者在此一并致以衷心的感谢。

由于编者水平有限，加之时间仓促，疏漏与不当之处在所难免，恳请读者和使用本套教材的教师批评指正。

编　者

于上海交通大学

1988 年 10 月 25 日

目　录

引　言

客观世界中发生的现象不外乎两种：一种是确定性现象，一种是随机现象。如在 1 个大气压下水在 100℃ 时必然沸腾，在 0℃ 时必然结冰，就是确定现象。掷一次硬币可能出现正面，也可能出现反面；新生的婴儿可能是男孩，也可能是女孩。这种在一定条件下，具有多种可能结果，但事先又不能确定究竟发生哪一种结果的现象就是随机现象。经典的数学理论如微积分学、微分方程等都是用来研究确定性现象的，而对随机现象却无能为力。随着社会生产与科学的发展，人们对随机现象越来越重视，从而使研究随机现象的概率统计获得了迅速的发展，形成了数学的一个重要分支，它广泛地应用于工业、农业、军事和科学技术中，并且还不断地向其他学科渗透，其势头至今不衰。

经典数学和概率论所研究的对象虽然如此不同，但它们是相辅相成、互相渗透的。如一枚炮弹在空中飞行的曲线——弹道曲线，在认为炮弹是一个质点，并且不考虑阻力影响及其他因素的微小影响下，可以归结为微分方程问题，从而得到一条确定的抛物线。如果空气的阻力规定为与飞行速度成正比，即把空气阻力按一定方式考虑进去，也可以利用微分方程的方法得到一条确定的弹道曲线。然而，在实际发射中，会发现每次的飞行路线都不相同。这个差异是由于炮弹飞行路线受到捉摸不定的空气阻力、炮弹本身的不均匀性、弹身振动等的影响而造成的。经典数学正是把那些看起来是微不足道的次要因素的影响

略去后进行研究的，若将这些次要因素考虑进去就会造成飞行路线的不确定性，那就只能用概率统计的方法加以研究。但必须指出，尽管飞行路线有所偏差，但都是围绕用经典数学方法求出的那条曲线摆动的，如果连这条确定的曲线也无法知道的话，那末考察飞行路线的偏差也就无多大意义了。这就是说某些概率统计的问题必须辅之以经典数学的方法；反之将经典数学研究的某些问题的结果用于实际，则由于大量的被略去的次要因素的作用，会产生与实际情况不同的差异，而这种差异是不确定的，所以又必须用概率统计的方法加以补充解决。因此必须正确认识概率统计方法与经典数学方法间的这种相辅相成的关系。

本书主要是讲概率论，内容包括随机事件及其概率、随机变量及其分布、数字特征、大数定律和中心极限定理。全部讲授约需 30～36 学时。我校某些专业需要一点数理统计的基础知识，而又无法单独设置数理统计这门课程，为此，写了第六章“数理统计初步”，讲授这部分内容约需 8～10 学时。书中带“*”号的内容可以略去不讲。各章配有适量的习题，书末附有习题答案。

第一章　随机事件及其概率

§1　随机事件及其运算

概率论是研究随机现象数量规律的一门数学学科。对随机现象进行研究，就要进行观察、试验。为了叙述方便，我们把对自然现象或社会现象进行的观察或实验，都称为试验。如果一个试验在相同条件下可以重复进行，而每次试验的可能结果不止一个，但在进行一次试验之前却不能断言它出现哪个结果，则称这种试验为随机试验。今后所说的试验都是指随机试验。

在试验中，可能发生也可能不发生的事情，称为**随机事件**，简称**事件**。

例 1　掷一枚硬币，出现正面及出现反面都是随机事件。

例 2　掷一颗骰子（骰子是一个均匀对称的正六面体，各面上分别标有“1”、“2”、“3”、“4”、“5”、“6”点）出现“1”点、“5”点和出现偶数点都是随机事件。

例 3　电话接线员在上午 9 时到 10 时接到的电话呼唤次数，如出现 1 次，2 次…及出现次数在 50 到 100 之间都是随机事件。

例 4　对某一目标发射一发炮弹，弹着点到目标的距离为 0.1 米、0.5 米及 1 米到 3 米之间都是随机事件。

从上面的例子可以看出，在一个试验中，所出现的事件是很多的。例 1 的事件有两个，例 2 的事件有很多个，但却是有限的。例 3 和例 4 的事件却有无穷多个。

在一个试验下，不管事件有多少个，总可以从其中找出这样一组事件，它具有如下性质：

(1) 每进行一次试验，必然发生且只能发生这一组中的一个事件；

(2) 任何事件，都是由这一组中的部分事件组成的。

这样一组事件中的每一个事件称为**基本事件**，用 ω 来表示。基本事件的全体，称为试验的**样本空间**，用 Ω 表示。

在例 1 中，我们取

$$\Omega=\{(\text{出现正面}),(\text{出现反面})\}。$$

在例 2 中，我们取

$$\Omega=\{(\text{出现 1 点}),(\text{出现 2 点}),\cdots,(\text{出现 6 点})\}。$$

在例 3 中，我们取

$$\Omega=\{(\text{出现 0 次}),(\text{出现 1 次}),\cdots\}。$$

在例 4 中，我们取

$$\Omega=\{(\text{弹着点与目标的距离 }\omega)\mid 0\leqslant\omega<+\infty\}。$$

通常，Ω 中的基本事件就是试验中所有可能直接出现的结果。根据这一点，我们可以对试验找出所有的基本事件。

如果我们把一个基本事件视为一个抽象的“点”，那末样本空间就是由这些抽象的“点”组成的空间。根据性质(2)，一个事件就是由 Ω 中的部分点(基本事件)组成的集合。通常用大写字母 $A, B, C, \cdots$ 表示事件，它们是 Ω 的子集。

如果某个 ω 是事件 A 的组成部分，即这个 ω 在事件 A 中出现，记为 $\omega\in A$，读作 ω 属于 A。如果在一次试验中所出现的

ω 有 $\omega\in A$，则称在这次试验中事件 A 发生。

如果 ω 不是事件 A 的组成部分，就记为 $\omega\overline{\in}A$，读作 ω 不属于 A。在一次试验中，所出现的 ω 有 $\omega\overline{\in}A$，则称此次试验 A 没有发生。

很显然，总有 $\omega\in\Omega$，将 Ω 作为事件，则在试验中事件 Ω 总是发生的，故称 Ω 为**必然事件**。它不是随机事件，把它作为事件主要是为了讨论问题的方便。另一个不是随机事件而视为事件的就是不包含任何基本事件的事件，记为 ϕ。由于对一切的 ω 有 $\omega\overline{\in}\phi$，故在试验中，ϕ 总是不发生的，所以称 ϕ 为**不可能事件**。

如果我们有了一些事件，则可以从这些事件得出其他事件来，这就是事件的运算。下面先介绍事件的包含与等价关系，再讨论事件的运算。

如果事件 A 的组成部分也是事件 B 的组成部分，则称事件 **A 包含于事件 B**，或称事件 **B 包含事件 A**，记为 $A\subset B$ 或 $B\supset A$。$A\subset B$ 的直观意义就是如果事件 A 发生必有事件 B 发生。

如果同时有 $A\subset B$，$B\subset A$，则称事件 A 与事件 B **等价**，或称 **A 等于 B**，记为 $A=B$。其直观意义是组成 A、B 的基本事件完全相同，因此可以看作是一样的。

将事件 A 与 B 的组成部分合并（A、B 所共同的基本事件只取一次）而组成的事件称为事件 A 与事件 B 的**并事件**（或**和事件**），记为 $A\cup B$。由于 $A\cup B$ 发生是指属于 $A\cup B$ 的某个基本事件 ω 发生，所以 ω 不属于 A 就属于 B，即表示不是 A 发生就是 B 发生，因而 $A\cup B$ 的直观意义就是 A、B 中至少有一个发生的事件。类似地，我们可以规定可列个事件 $A_1, A_2, \cdots, A_i, \cdots$ 的并，记为

$$A_1\cup A_2\cup\cdots\cup A_i\cup\cdots,\quad 或\quad \bigcup_{i=1}^{\infty}A_i,$$

它表示 $A_1, A_2, \cdots, A_i, \cdots$ 中至少有一个发生的事件。

事件 A、B 的共同组成部分所构成的事件，称为**事件 A 与事件 B 的交**(或**积**)，记为 $A\cap B$。有时也可省去"$\cap$"而简写为 AB。若属于 $A\cap B$ 的某个 ω 发生，那就是 A 与 B 同时发生，所以 $A\cap B$ 的直观意义是 A、B 同时发生的事件，类似地可规定可列个事件 $A_1, A_1, \cdots, A_i, \cdots$ 的交(积)，记为

$$A_1\cap A_2\cap\cdots\cap A_i\cap\cdots, \text{ 或 } \bigcap_{i=1}^{\infty} A_i,$$

它表示 $A_1, A_2, \cdots, A_i, \cdots$ 同时发生的事件。

属于 A 而不属于 B 的部分所构成的事件，称为 A 与 B 的**差**，记为 $A-B$，它表示 A 发生而 B 不发生的事件。

$A\cap B=\phi$，则表示 A 与 B 不可能同时发生，称事件 A 与事件 B **互不相容**。基本事件是互不相容的。

$\Omega-A$ 称为事件 A 的**逆事件**，或称 A 的**对立事件**，记为 $\bar{A}$。它表示 A 不发生的事件。这样可得

$$A\cup\bar{A}=\Omega, \quad A\cap\bar{A}=\phi。$$

前一式表示 A 与 $\bar{A}$ 至少一个发生，后一式表示 A 与 $\bar{A}$ 不能同时发生。

必须指出，直观意义能帮助我们理解，但不能代替等价关系的证明。下面我们来证明

$$\overline{\bigcap_{i=1}^{\infty} A_i}=\bigcup_{i=1}^{\infty}\bar{A}_i。$$

为书写方便，将命题甲能推出命题乙写为甲$\Longrightarrow$乙。要证明上式就是要证明两边的包含关系成立，即

$$\overline{\bigcap_{i=1}^{\infty} A_i}=\bigcup_{i=1}^{\infty}\bar{A}_i \Longleftrightarrow \overline{\bigcap_{i=1}^{\infty} A_i}\subset\bigcup_{i=1}^{\infty}\bar{A}_i \text{ 且 } \bigcup_{i=1}^{\infty}\bar{A}_i\subset\overline{\bigcap_{i=1}^{\infty} A_i}。$$

设 $\omega\in\overline{\bigcap_{i=1}^{\infty} A_i}$，则有

$$\omega \in \overline{\bigcap_{i=1}^{\infty} A_i} \Longrightarrow \omega \overline{\in} \bigcap_{i=1}^{\infty} A_i \Longrightarrow \text{存在 } i_0,$$

使

$$\omega \overline{\in} A_{i_0} \Longrightarrow \omega \in \overline{A}_{i_0} \Longrightarrow \omega \in \bigcup_{i=1}^{\infty} \overline{A}_i,$$

故有

$$\overline{\bigcap_{i=1}^{\infty} A_i} \subset \bigcup_{i=1}^{\infty} \overline{A}_i\text{。}$$

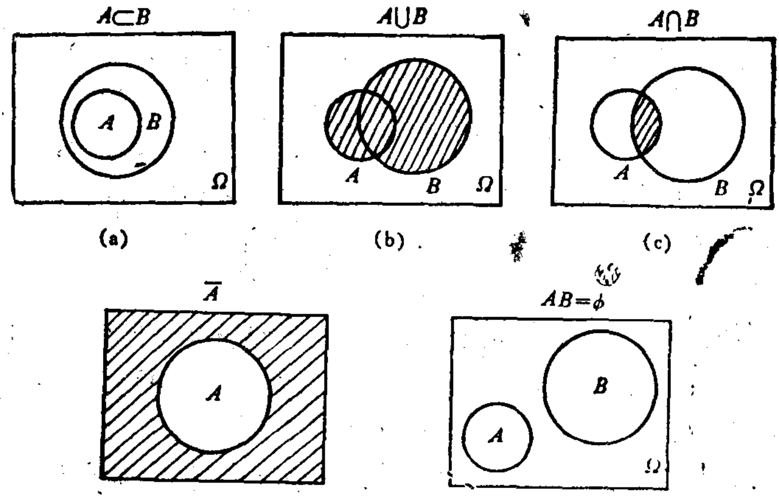
以上证明过程逆推回去也是成立的，于是又有

$$\bigcup_{i=1}^{\infty} \overline{A}_i \subset \overline{\bigcap_{i=1}^{\infty} A_i},$$

从而有 $\overline{\bigcap_{i=1}^{\infty} A_i} = \bigcup_{i=1}^{\infty} \overline{A}_i$。

以上是证明等价关系的一般方法。对于一些比较明显的等价关系，可以由直观意义获得，也可以借助几何直观得到。下面就介绍一下文（Venn）图。

用平面上的一个矩形表示样本空间，矩形内的点表示基本



(d) 图 1.1 (e)

事件，则事件间关系及运算就可用平面上的几何图形表示，如图1.1所示。事件 A、B 用两个小圆表示，阴影部分表示 A 与 B 的各种关系及运算。

由图1.1(b)，可得 $A-B=A-AB$，$(A-B)\cup(B-A)=(A\cup B)-AB$。

例5　一口袋中装有五只乒乓球，其中三只是白色的，两只是红色的。现从袋中取球两次，每次一只，取出后不再放回。写出该试验的样本空间 Ω。若 A 表示取到的两只球是白色的事件，B 表示取到的两只球是红色的事件，试用 A、B 表示下列事件：

(1) 两只球是颜色相同的事件 C，

(2) 两只球是颜色不同的事件 D，

(3) 两只球中至少有一只白球的事件 E。

解　假设每个球是可以区别的，则可以给每个球编号：白$_1$，白$_2$，白$_3$；红$_4$，红$_5$。于是有

$$\begin{aligned}\Omega=\{&(白_1,白_2),(白_1,白_3),(白_1,红_4),(白_1,红_5),\\&(白_2,白_1),(白_2,白_3),(白_2,红_4),(白_2,红_5),\\&(白_3,白_1),(白_3,白_2),(白_3,红_4),(白_3,红_5),\\&(红_4,白_1),(红_4,白_2),(红_4,白_3),(红_4,红_5),\\&(红_5,白_1),(红_5,白_2),(红_5,白_3),(红_5,红_4),\end{aligned}$$

从而，$C=A\cup B$，$D=\bar{C}=\overline{A\cup B}$，$E=D\cup A=\overline{A\cup B}\cup A$，或 $E=\bar{B}$。

若我们认为白球间是无区别的，红球间也是无区别的，则有

$$\Omega=\{(白,白),(白,红),(红,白),(红,红)\}。$$

这样 C、D、E 所表示的事件就更加简单了。

由上可知，同一个试验可以设计不同的样本空间，而设计的好坏全在于能否使得讨论的问题简单。

§2 概率的定义及其计算

一、频　率

既然事件的发生有可能性，自然就有一个可能性大小的问题。人们常常通过实际观察来确定某个事件发生的可能性大小。例如遇到某种天气，人们常会说“今天十之八九要下雨”，这个“十之八九”就是表示“今天下雨”这一事件发生可能性的大小。这也是人们通过大量实践总结出来的，即已经历过n次这种天气，而下雨的天数在这n天中所占比例大约是8/10到9/10。一般说来，一个事件A发生的可能性大小，可用在n次重复试验下，事件A发生的次数n_A与n的比值来反映。我们把

$$F_n(A)=\frac{n_A}{n}$$

称为事件A在n次试验中出现的**频率**。

频率具有如下性质：

1°　对任一事件A，有 $0\leqslant F_n(A)\leqslant 1$；

2°　对必然事件Ω，有 $F_n(\Omega)=1$；

3°　若事件A、B互不相容，则

$$F_n(A\cup B)=F_n(A)+F_n(B)。$$

证　因为$0\leqslant n_A\leqslant n$，所以有$0\leqslant\frac{n_A}{n}\leqslant 1$，由$F_n(A)$的定义就得性质1°。

由$n_\Omega=n$，即可得性质2°。

由于$A\cup B$事件的发生就是A，B两事件中至少一个发生，

又知 A、B 互不相容，故有

$$n_{A\cup B}=n_A+n_B,$$

两端除以 n，就得性质 3°。

性质 3° 还可以推广，若事件 A_1, A_2, …, A_m 两两互不相容，则

$$F_n\left(\bigcup_{i=1}^{m}A_i\right)=\sum_{i=1}^{m}F_n(A_i)。$$

频率虽然在一定程度上反映了事件发生的可能性大小，但它却依赖于人的认识，即会因人而异。因为即使同样做了 n 次试验，n_A 却会不一样，这种差异我们常说成是频率具有随机波动性。但若加深认识（这里就是增加试验次数 n），那末随机波动性将会减小。即随着 n 逐渐增大，$F_n(A)$ 也就逐渐稳定于某个常数 $P(A)$。这个常数 $P(A)$ 就能客观上反映事件 A 发生的可能性的大小。

历史上著名的统计学家蒲丰（Buffon）和皮尔逊（Pearson）曾进行过大量掷硬币的试验，所得结果如下：

试验者	掷硬币次数	出现正面的次数	出现正面的频率
蒲　丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

可见出现正面的频率总在 0.5 附近波动。随着试验次数的增加，它逐渐稳定于 0.5。这个 0.5 就能反映正面出现的可能性的大小。

每个事件都有这样一个常数与之对应。这就是说频率具有